Chapter 1
Experience in System Design for Human-Robot Teaming in Urban Search & Rescue*

G.J.M. Kruijff, M. Janíček, S. Keshavdas, B. Larochelle, H. Zender (DFKI GmbH)
N. Smets, T. Mioch, M. Neerincx, J. van Diggelen (TNO)
F. Colas, M. Liu, and F. Pomerleau (ETH Zürich)
T. Svoboda, T. Petriček, (CTU Prague)
F. Pirri, M. Gianni, P. Papadakis, A. Sinha (ROMA)
P. Balmer, N. Tomatis (BLUE)
R. Worst, T. Linder, H. Surmann, V. Tretyakov (Fraunhofer)
S. Corrao (VVFF)
S. Pratzler-Wanczura, M. Sulk (FDDO)

Abstract The paper describes experience with applying a user-centric design methodology in developing systems for human-robot teaming in Urban Search & Rescue. A human-robot team consists of several robots (rovers/UGVs, microcopter/UAVs), several humans at an off-site command post (mission commander, UGV operators) and one on-site human (UAV operator). This system has been developed in close cooperation with several rescue organizations, and has been deployed in a real-life tunnel accident use case. The human-robot team jointly explores an accident site, communicating using a multi-modal team interface, and spoken dialogue. The paper describes the development of this complex socio-technical system per se, as well as recent experience in evaluating the performance of this system.

1.1 Introduction

Urban Search & Rescue is a domain where robots have the potential to make a difference [26]. They can go where people cannot. To help assess a situation, determine an approach to deal with it, even before humans have gone in.

To make this possible, we do need more autonomy in the robot [3], in perceiving the environment, in navigating it. However, disaster areas are harsh places. We inevitably experience what Woods et al [35] termed "***(Robin) Murphy's Law****: any deployment of robotic systems will fall short of the target level of autonomy, creating or exacerbating a shortfall in mechanisms for coordination with human problem holders*." Adaptive autonomy is one way of trying to address this problem [30, 24], making explicit the inherent interdependence between humans and robots [13].

* The paper reports research funded by the EU FP7 ICT program, Cognitive Systems & Robotics unit, under contract 247870, "NIFTi" (http://www.nifti.eu)

Still, this is all for naught if the humans in the team do not *accept* a robot's autonomous capabilities and intelligence. Recent experience in Fukushima (S. Tadokoro, p.c.) and in our own end user studies underline this. A robot's abilities, behaviour, and possible achievements need to be transparent to a human operator: Whether the robot is doing something, what it is doing and why, whether it thinks it has achieved a goal (or not). If an operator is unclear about what to expect from the robot, he or she is unlikely to delegate control to the robot. Instead, no matter what the robot is able to do autonomously, the operator will revert to tele-operation.

And that's not quite what anybody wants. We see this as an issue of (lacking) transparency in experience, behavior and intentions [6]. Robot behavior needs to be transparent, to allow for a proper management of user expectations. A gap between these expectations, and what actually happens, can seriously affect the interaction [23, 17]. A lack of transparency reduces acceptability, which might explain why human-robot interaction (HRI) is a bottleneck in USAR [25].

The problem gets exacerbated in the context of USAR. Humans and robots perform under stress, in complex environments. Situations, interactions, plans change. And with that, expectations change. What we are looking at is not characterizing a gap between expectations before and after a human has interacted with a robot, as is typically done in studies on HRI [23, 17]. Instead, we need to address expectation management online. As situations change, affecting the dynamics of the team, the robot needs to adapt its behavior, and the way it presents that behavior to continue to provide adequate and effective transparency; cf. e.g. [27].

In this paper, we try to further understand the problem. We do not offer a solution; but we discuss a way in which we believe we can come to *understand* the problem better, and design systems that can eventually address the problem in real-life. We present a user-centric design methodology (§1.2) which draws in end users (first responders from several organizations across Europe) and their experience into the entire R&D process. Following this methodology, we discuss how we design our systems (human-centric, §1.3), and how we experiment with them and evaluate them (with end users, under real-life circumstances §1.4).

1.2 User-centric design methodology

We adopt a user-centric design methodology, in several respects. Firstly, we include users in all the phases of the development cycle: Requirements analysis, component- and system development, and experiments & evaluations. Users are from various rescue services (Fire Department of Dortmund/Germany, Vigili del Fuoco/Italy). Together, we formulate requirements for hard- and software functionality, and develop physically realistic use cases in which we can experiment with and evaluate our approaches. Figure 1.1 illustrates one such use case, namely a tunnel accident.

Involving users throughout the yearly development cycle does more than just telling us what they need (requirements), and whether our systems do the job (evaluations). Their involvement provides us with a deeper insight into their needs, their

(a) (b) (c)

Fig. 1.1 NIFTi tunnel accident use case: (a) Setting (b) UAV and UGV in action; (c) control center.

procedures, and what happens out in the field. This is another aspect of the human user-centric design approach we follow. We build systems which can assist humans, doing so in ways that mimic human understanding, and operational procedure. The hypothesis being that this makes robot behavior more transparent to the user.

In the system design, the human perspective is pervasive throughout the representations the robot builds, and the way it determines its behavior. The conceptual understanding of the environment provides a human-like view on the environment, and the inference of spatially grounded affordances results in robot behavior that mimics standard procedure. When it comes to human-robot interaction and planning, humans are explicitly modeled as actors, and action and interaction are planned in ways that conform to human operational practice.

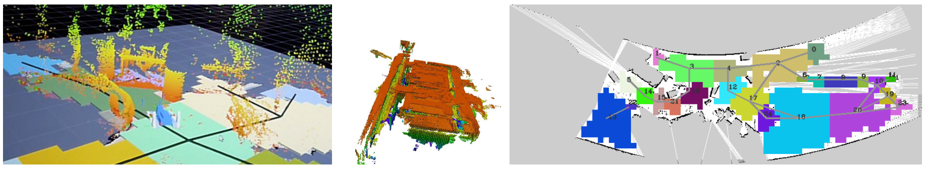
1.3 Socio-technical system design

We approach design from a socio-technical perspective. It concerns the entire system of robots, humans, and how they work together. We focus on four questions:

1. How to model situation awareness which (a) bridges the gap between a robot's quantitative, and a human's qualitative sense of space, (b) facilitates use by a geographically distributed team, and (c) provides the basis for individual or joint action (4)? See §1.3.1.
2. How to model the impact of situations in task- and team-work which influence user performance, given that (a) humans typically perform under stress in USAR missions, and (b) stress alters interaction patterns (3)? See §1.3.2.
3. How to model user-adaptive human-robot communication, to adjust how, what, and when a robot communicates given an awareness of the current operative situation (1) and its effects on human performance (2)? See §1.3.2.
4. How to model flexible temporal planning and execution, to guide how a robot plans and executes its own actions under different conditions (1)? See §1.3.3.

1.3.1 Intelligence in situation awareness

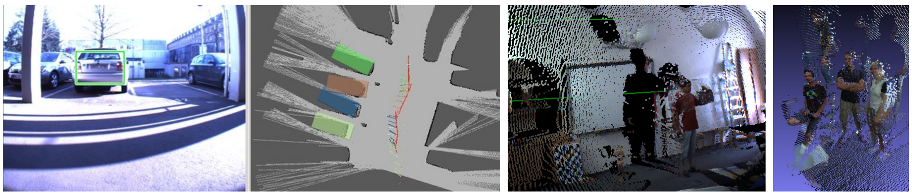
A robot builds up a situation awareness which bridges the gap between its own quantitative forms of perception, and a human qualitative understanding of space. The robot builds up a qualitative structure of dynamic space, and can make inferences about possible actions situated in that space. Mapping therefore builds up several layers of abstraction. First we try to build an accurate metric representation of the environment based on the 3D rolling laser sensor mounted on our robot. Based on this metric representation, we then segment the navigable space into coherent areas linked in a navigation graph. Going 3D requires both to have an efficient 3D representation of the environment and to be able to estimate the 6 degrees-of-freedom pose of our robot. The representation of the map is made using fast and flexible octrees [36]. Figure 1.2 shows an example of a 3D point cloud representation in the tunnel accident use case, and an octrees-based 3D map. To avoid part of the distortions, the 3D point clouds are registered into the map only when the robot is static. Preliminary results show that in most cases the distortion when the robot is moving is not too large, but localization may jump from local optima and induce point cloud deformation due to the pose interpolation. The 6 DOF pose estimate is based on a robust 2D map when the robot lies in a mostly 2D part of the environment. We rely on fast and efficient 3D pose estimation to handle 3D environments [32].



(a) (b) (c)

Fig. 1.2 3D mapping: (a) 3D point cloud in tunnel, (b) 3D map (indoor environment), (c) Topological segmentation of tunnel, with navigation graph (in grey)

For the topological segmentation, we take as input the map of the environment. Previously we performed topological extraction based on spectral clustering and mutual information [22]. To better handle changes in the map, both due to exploration and due to actual changes, we use incremental topological segmentation. Figure 1.2(c) illustrates the result of this new method in the tunnel environment.



(a) (b) (c)

Fig. 1.3 Car detection using visual features and 2D mapping (a), 3D laser point clouds (b) and fusion with visual data from omnicam (c)

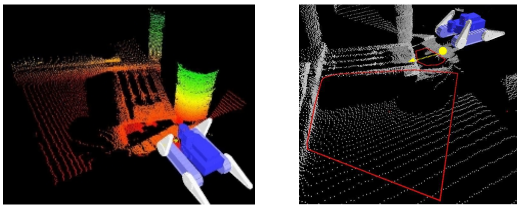
We use 3D point clouds and robot positioning to improve vision. Image-based detection of rear parts of cars in a tunnel accident works relatively well [37], see Figure 1.3. Estimating the 3D positions of cars proved more difficult, especially the orientation. To deal with 3D instability we associate 2D features with the 3D metric maps. Figure 1.3(b,c) shows an example of assigning image colors to the 3D point clouds. The 2D object detector creates a probabilistic map over the image, and attributes this to points in a 3D point cloud. The 3D information provides an absolute scale, which allows for discarding many false positives. The result of connecting visual perceptions with the 2D- and 3D map representations is that we now obtain grounded observations of objects in the scene.

We use these object observations to perform Functional Mapping, a form of spatial inference [14]. Given an object, and an action to be performed, functional mapping infers areas around the object, where the action can be performed relative to the object. This combines logical inference over associated ontologies for objects and their internal structure, and for actions; and geometric inference. In the tunnel accident, functional mapping infers that being in a particular position relative to a car window facilitates looking into that car. The projection of the areas into 3D space is based on real-time map data and the observed 3D pose of the object. Functional mapping thus combines top-down inferencing, from a priori knowledge of expected objects, and bottom-up inferencing from real-time observations.

Inferring functional areas serves several purposes. First of all, when fire-fighters explore a disaster site, they themselves move between functional areas to make their observations [15]. We observed the same behaviour when fire-fighters tele-operated robots to explore an accident, [14]. Making the robot follow similar behaviour makes that behaviour more transparent to an operator working with the robot. Secondly, we use the inference of functional areas to determine optimal vantage points for the robot to perform an observation. Finally, these functional areas serve in maintaining common situation awareness between the robot's metrical environment knowledge and the pilot's qualitative understanding of the environment. Thus, when a pilot instructs the robot to "go to the car", it goes into a functional area, rather than naively trying to go to (the center of) the car.

Finally, we use map information to perform terrain analysis for traversability, particularly negative obstacle and gap detection. Our approach has two main stages; (i) Application of image morphological and contour detection algorithms and (ii) application of Principal Component Analysis in the orientation domain of the gap contours [28] and extraction of the optimal traversability path. Reasoning with respect to the traversability of the detected gaps is done considering the dimensions and morphological adaptation capabilities of the robot. A representative example of gap detection and analysis is given in Fig. 1.4.

Adapting the robot's morphology concerns adjusting its articulated components to reduce instabilities that could tip it over [29]. To optimally adapt its morphology with respect to the terrain we consider maximizing the surface contact of the tracks with the ground. (This aims to maximize the traction efficiency of the robot which in parallel results in minimized pressure on the tracks.) Using a set of various ter-



(a) (b)

Fig. 1.4 Gap detection and analysis; (a) Top view of the 3D point cloud and (b) detected gaps together with traversability direction.

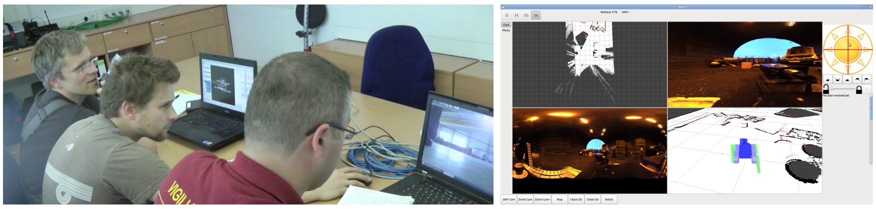
rains classes we first learn the optimal configurations of the robot offline, using a simulation environment (Gazebo) [10], to employ them later on in the real scenario.

1.3.2 Intelligence in interaction

HRI is regarded one of the major bottlenecks in rescue robotics [25, 26]. Teleoperating a robot is highly demanding. More autonomy can be a way out of this. But as we already argued, more autonomy requires more transparency, to facilitate common ground and coordination. And that requires communication.

Unfortunately, most models of HRI have so far been relatively limited in their use of spoken dialogue, one of the most natural means for humans to interact. Also, these models typically do not ground communication in the social structure, to explain why actors (need to) interact, and what information is to be exchanged. We are working on an approach that takes the social structure and the collaborative ("intentional") context explicitly into account [19, 18, 20]. The approach is based in previous collaborative views on dialogue processing [1, 11, 2]. Our approach improves on these by dealing explicitly with uncertain, incomplete information, as is typical for spoken dialogue, and particularly situated dialogue.

We have integrated (limited) spoken dialogue into our multi-modal GUI for human-robot interaction. A user can use dialogue to instruct the robot to move to particular waypoints or landmarks (possibly selected in the GUI), or drive in specific directions [21], similar to [8]. Based on insights in human-human interaction in human-robot teams for USAR (NJEx 2011, §1.4), and the recent experience in the end user evaluations at SFO (§1.4), we see there is particularly a need for the robot to *produce* contextually appropriate feedback to maintain transparency. (The range of utterances which a robot needs to understand is relatively limited in this domain.) Using our recent experimental data, we are investigating the relation between when what is to be communicated by the robot to someone (communication patterns) – and task context, and the user's estimated stress and workload. This should provide an insight in not just *what* to say, but also *how* to say it best such that it is easy to understand by the user under the given circumstances.



(a) (b)

Fig. 1.5 Team-based, multi-modal GUI with multiple perspectives (a) and multiple info-views (b)

1.3.3 Intelligence in team cooperation

Human-robot teams are typically geographically dispersed. For team cooperation this requires the entire system to integrate different views on the environment, (e.g. UAV, UGV, in-field operators), and to facilitate different perspectives and needs [33]. Below we briefly describe the planning approach we use for a robot to share and coordinate control with other team members, to support coordinated execution.

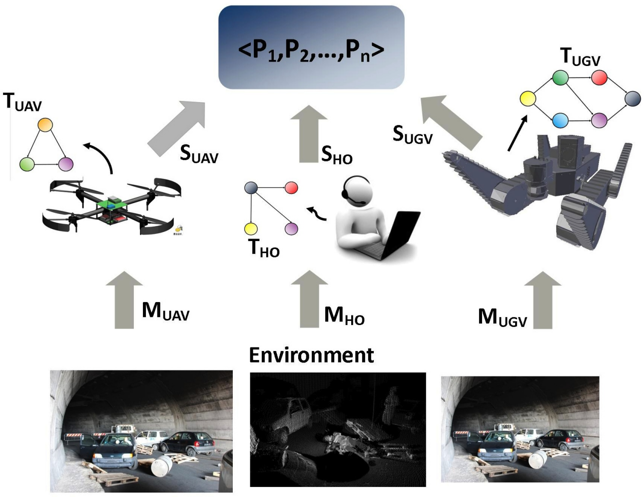


Fig. 1.6 Properties $P_1, ..., P_n$ are defined on inner states S of each team unit, to give a uniform representation of the multi-agent system. M denotes a perceptual model of a unit, T temporal model of unit activities.

The dynamics of the UGV and UAV can be modeled separately by defining two different temporal declarative models in the Temporal Flexible Situation Calculus (TFSC) [7]. The UAV can act in strict cooperation with the UGV, so the TFSC model needs to know the states of both system components, via a common model.

The hybrid framework combines temporal constraint reasoning and reasoning about actions. The flexible behaviours of the UAV and UGV are specified in a compact representation by temporal constraint networks T_{UAV} and T_{UGV}, with the possibility to include a network T_{HO} corresponding to an in-field human operator-rescuer. These causal and temporal relations, and their constraints, are learned by continuous interaction with humans, via demonstration and by collected observations of successful processes in controled contexts [31, 15]. The networks are mapped into a structure managing time, resources and actions, (model-based control). The model accounts for timelines with time flexibly assigned to each component, to satisfy priorities for both resources and tasks, and which rely on online acquisition of sensor data [9]. The whole set is managed by an execution monitor which continuously checks the environment models $\{M_{HO}, M_{UAV}, M_{UGV}\}$ and the inner states $\{S_{HO}, S_{UAV}, S_{UGV}\}$. The execution loop ensures that the network is kept uptodate, and consistent. The inner states S_{UAV} and S_{UGV} represent the internal loop which checks on all of the machine components, namely both of the UAV and UGV. The human-robot team shares the information about the environment and the mission, combining together their models of the current percepts. To integrate the different abilities of the UAV, the UGV, and the users, a set of properties $P_1, ..., P_n$ are defined on top of the inner states of the team units bridging the different dynamic models (Figure 1.6).

1.4 Field experiments & evaluations

During 2011, we performed several experiments with end users, operating our systems under realistic circumstances. Almost needless to say, we observed problems along the way which are familiar to anyone operating in the field, see e.g. [5]: Network problems, hardware failures, issues in keeping complex distributed systems synchronized, logging all the data for future analysis. Below we focus primarily on the lessons in human-robot collaboration for USAR we learnt.

1.4.1 January 2011: Pilot experiments at FDDO

The first pilot experiments with users took place at the training center of the Fire Department of Dortmund (FDDO), in January 2011. Users (all professional first responders) operated a UGV in a "tunnel accident"-like environment (several crashed cars, a motor bike, debris, set up in a large garage). We ran experiments with two different types of fully tele-operated UGVs: The TNO "Generaal" robot [12], and an ActivMedia P3-AT. The UGVs differed primarily in the sensory data they made available to the user and in the way interaction with the robot was supported. The "Generaal" has a specially designed telepresence control, consisting of a headtracking head-mounted display, whereas the P3-AT was operated via the NIFTi touchscreen [21]. The P3-AT was the focus of the pilot experiment, as it used the NIFTi interface and sensory analyses. As was to be expected, many technical problems arose: Signal loss, insufficient battery power, insufficient bandwidth for video-based feedback, damages due to obstacles, and cold [5]. Out of the three users using the P3-AT who tried the 15-minute accident exploration task, two had to be cut short after 10 and 13 minutes respectively.

This experiment provided insights similar to what is found in e.g. [4]. We could observe that users spent about half their time navigating, and about one third of the remaining time trying to find pathways [21]. By observing the paths taken by the fire fighters in the scenario, we found out that they were similar to those followed by actual fire fighters in similar scenarios [15]. A surprising observation was that users were satisfied with the robot's video feed. The highly-compressed poorly-lit images were typically shown in low 400x800 resolution. The users claimed though that the quality was sufficient, even when update frequency was too low ($< 8Hz$) for safe tele-operation. The experiments did reveal that tele-operation increased the cognitive load of the user [21]. This was one reason to develop more autonomous robot behavior (§1.3). We were worried that with so much time spent on navigation tasks rather than on observation, user situation awareness would be poor. However, sketch maps that they drew on white boards during the scenario showed that they located most objects within one meter of their actual location, and that the relative positioning of the objects with one another was often entirely correct.

It was therefore difficult to explain why tele-operation was so difficult if the users were so well aware of their surroundings. We have since improved the visualization in the GUI, to bring closer the robot's model environment to a user's mental map. We tried to do this for example by adding a to-scale 3D model of the robot in the map generated by the robot and allowing for more and better views of the generated maps. We also provided means of controlling the robot other than tele-operation, e.g., by spoken commands.

1.4.2 July 2011: NIFTi Joint exercises (NJEx) at FDDO

Given that our first experiments mostly focused on single-robot single-operator missions, we wanted to study human-robot teaming in more detail. We organized a more complex joint exercise event involving project partners, and end users (FDDO, VVFF). During this event, teams of several humans, a NIFTi-specific outdoor UGV, and a UAV explored several complex environments. This included a multi-story residential building "on fire." Missions of 45 minutes were performed by a team consisting of a mission commander, a UGV operator, and a UGV/UAV mission specialist (in a remote control post), and in-field a UAV operator and two safety directors for the UGV and the UAV (in-field, line-of-sight). Team members included both first responders and scientists.

For NJEx 2011 we particularly focused on human-human interaction within the human-robot team, as the robots were fully tele-operated. Interesting observations here were that communication particularly concerned the communication of situation awareness ("we see a victim under the shelves in the room at the end of corridor, right"), and goals ("we are going to look under the staircase, at the end of the corridor, left"). Nearly all of the information exchanged was explicitly situated (or located) in the environment. The mission commander mostly communicated situation awareness, to maintain common ground within the team, whereas the UGV operator would indicate the next actions that the UGV would perform. Planning exploration tasks was typically done within the control post and the coordination with the in-field team was done through the mission commander. The UAV operator's task was to fly to a particular point, with an explicitly communicated purpose – typically, what kinds of observations the control post would like to make. Video feeds from the UAV were inspected by the UAV mission specialist in the command post, with the mission commander providing feedback to the UAV Operator. The two safety directors had the best awareness of the situations around the robots, as they were in line-of-sight. The protocol was such that they were not allowed to describe the environment, except if to avoid damage to the robot.

Analysis so far has yielded that the mission commander and the UGV operator generate the most radio traffic, with one or the other taking on a leading role. In the most effective teams, this was always the mission commander. Furthermore, variations in stress levels could be detected, particularly for team members with high radio traffic, i.e., the mission commander, UGV operator, and the UGV safety commander. We observed that in low stress situations, loosely defined roles and communication protocols can have a slight negative impact on the team performance. However, in more complex situations (e.g., time pressure, high cognitive load, stress), the lack of protocol can break the team cohesion altogether. Face-to-face and radio communication get overloaded, team members get orders from multiple people, situation awareness becomes more local and poorly shared.

It is thus imperative that human team members follow strict rules from the onset of the mission until the end. Adding an autonomous robot to this team thus means it needs to "fit in." It must be socially accepted. The next experiments studied the introduction of such a robot in a complex USAR mission.

1.4.3 December 2011: End-user evaluations at SFO

This field trial is the third and most recent one. The scenario was again a tunnel car accident and was located inside a real tunnel. The area spanned 25 meters into the tunnel by a width of 10 meters, filled with debris, pallets, barrels, wrecked vehicles, and smoke. The users had 40 minutes to assess the situation with one UGV and one UAV. The human team members were the mission commander and the UGV pilot (in the command post), and the UAV pilot (in-field).

The team members in the control post had access to a variety of information sources, in a multi-screen multi-modal user interface set-up. The views included robot-specific interfaces, for example the UGV Operator Control Unit (OCU), and qualitative views for team-level situation awareness (TREX, [34]). Communication between the command post and the in-field UAV operator was via hand-held radio, through the mission commander. The UGV operator communicated with the in-field UGV using the OCU, which allowed spoken dialogue. The UGV was capable of autonomous navigation, and could also use spoken dialogue to provide observations and basic feedback on actions (action-possibility, action-onset).

Coming into this field trial, we had high expectations. We had defined clear roles for the human participants, we had a robust outdoors robot with enough power, space, and bandwidth for the scenario, we had much improved visualization, we had task and path planners, and we had several different levels of autonomy to reduce the amount of tele-operation, and we ensured that the robot could support a basic dialogue with the operator. From our experience in January 2011, we believed that reducing the need to tele-operate the robot to only the times when autonomous navigation does not work well would free up some time for the user to observe the environment or perform other tasks. However, reality was quite different.

Figure 1.7 shows the typical path followed by a user during the 40-minute scenario. The first section in orange shows that the user was using multi-modal dialogue to direct the robot to waypoints. The green section shows where the robot was run with less autonomy, using spoken commands for small movements (e.g. go left, go forward) or even just manual control. This graph illustrates that autonomous behavior of the robot was used less than expected. Although it is too early to draw definite conclusions, explanations for the low usage seem to revolve around the transparency of the autonomy, and the trust towards the robot's capacities. It was hard for the operator to build up enough trust, as too often, the robot's actions were not successful or interrupted because they could not be achieved – without the robot producing adequate (multi-modal) feedback.

All users started with operating the robot in a high degree of autonomy, using multi-modal dialogue. However, they all took back control as soon as the path became more difficult to navigate, and many more objects to observe became visible (after about 5 meters into the tunnel). This change in sharing control could be grounded in technical reasons (low speed or failures of the autonomous navigation) or on social reasons (lack of trust and understanding about autonomous navigation, due to lack of transparency). While autonomy per se remains crucial to achieving success in robot-assisted USAR, we becomes clear here is that we need to find clear and understandable ways to present this autonomy to the users. The robot's state, behavior, and capabilities must be transparent to the human operator. It must be clear what the robot knows, what it is currently doing and why, and what its planned actions are. And, perhaps even more importantly from the viewpoint of expectation management, it is crucial that the robot communicates why it can *not* perform certain actions, or succeed performing them, rather than just failing. If a user is unclear about what to expect from the robot, that user is unlikely to delegate control to the robot. Instead, the operator keeps control, no matter how capable the robot is. In summary, transparency is needed for understanding and trust, and trust is needed for autonomy. Our future field experiments will focus on understanding how to make the system more transparent, especially in situations of high stress and cognitive load.

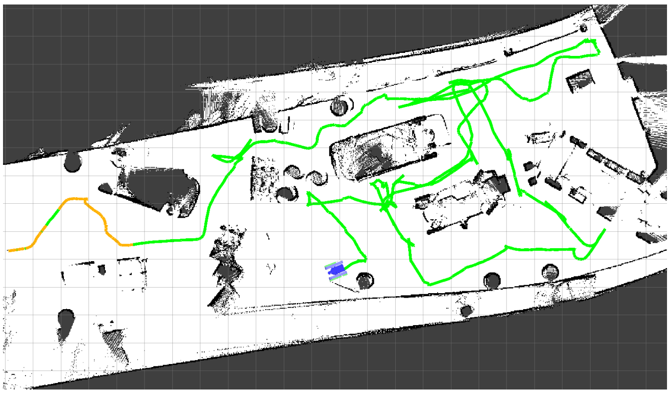


Fig. 1.7 Path taken in semi-autonomous mode vs. tele-operation

1.5 Conclusions

Developing, experimenting, and evaluation USAR robots together with professional users who have much at stake in this domain has turned out to be extremely revealing. In some sense, reality bites. What we believed to be the main issue at stake (autonomy) might well be overshadowed by the problems we are facing in making robot intelligence *acceptable*. Human-robot interaction as "the bottleneck" points into the direction we need to look. We face a socio-technical issue: the entire complex of a robot that can truly behave as a team member in a human-robot team (cf. also [16]). And before we can even talk of common ground, of collaboration, one of the most fundamental lessons we have learnt recently is that this all stands and falls with that robot's autonomous behavior being transparent.

Now that we have are slowly beginning to achieve an acceptable level of robot autonomy, it is time to focus our efforts on making this autonomy accepted. In the last field trial, we used a Wizard-of-Oz (WoZ) rather than an automatic speech recognizer (simply to avoid unnecessary complications in the experiment). We plan on pushing this type of setup further, to see how we can control the perception and usage of autonomous behavior. We would like to identify how technical limitations and failures affect the user's perception, compared to how transparency affects the perception of limitations, failures, as well as (situated) capabilities.

References

1. J. F. Allen and C. R. Perrault. Analyzing intention in utterances. *Artificial Intelligence*, 15(3), 1980.
2. J.F. Allen and D.J. Litman. Discourse processing and common sense plans. In P.R. Cohen, J. Morgan, and M. Pollack, editors, *Intentions and Communication*. The MIT Press, 1990.
3. A. Birk and S. Carpin. Rescue robotics - a crucial milestone on the road to autonomous systems. *Advanced Robotics*, 20(5):595–605, 2006.
4. J.L. Burke, R.R. Murphy, M. Coovert, and D. Riddle. Moonlight in Miami: An ethnographic study of human-robot interaction in USAR. *Human Computer Interaction*, 19((1–2)):85–116, 2004.
5. J. Carlson and R.R. Murphy. How UGVs physically fail in the field. *IEEE Transactions on Robotics*, 21(3):423–437, 2005.
6. H. Clark. *Using Language*. Cambridge University Press, 1996.
7. Alberto Finzi and Fiora Pirri. Representing flexible temporal behaviors in the situation calculus. In *Proceedings of the 19th international joint conference on Artificial intelligence*, IJCAI'05, pages 436–441, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.
8. T. Fong, C. Thorpe, and C. Baur. Multi-robot remote driving with collaborative control. *IEEE Transactions on Industrial Electronics*, 54(4):699–704, 2003.
9. M. Gianni, P. Papadakis, F. Pirri, M. Liu, F. Pomerleau, F. Colas, K. Zimmerman, T. Svoboda, T. Petricek, G. J. M. Kruijff, H. Zender, and Khambhaita. A unified framework for planning and execution-monitoring of mobile robots. In *Proceedings of the AAAI-11 Workshop on Automated Action Planning for Autonomous mobile Robots*, 2011.
10. M. Gianni, P. Papadakis, F. Pirri, and M. Pizzoli. Awareness in mixed initiative planning. In *AAAI Fall Symposium Series*, 2011.
11. B. J. Grosz and C. L. Sidner. Attention, intention and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.
12. C. Jansen and J. B. F. van Erp. Telepresence control of unmanned systems. In M. Barnes and F. Jentsch, editors, *Human-Robot Interactions in Future Military Operations*, pages 251–270. Ashgate Publishing Limited, 2010.
13. M. Johnson, J.M. Bradshaw, P.J. Feltovich, R.R. Hoffman, C. Jonker, B. van Riemsdijk, and M. Sierhuis. Beyond cooperative robotics: The central role of interdependence in coactive design. *IEEE Intelligent Systems*, pages 81–88, May/June 2011.
14. S. Keshavdas, H. Zender, G.J.M Kruijff, M. Liu, F. Colas, K. Zimmermann, T. Petříček, and T. Svoboda. Functional mapping: Spatial inferencing to aid human-robot rescue efforts in unstructured disaster environments. In *Proceedings of the 2012 AAAI Spring Symposium on Designing Intelligent Systems*, 2012.
15. H. Khambhaita, G.J.M. Kruijff, M. Mancas, M. Gianni, P. Papadakis, F. Pirri, and M. Pizzoli. Help me to help you: How to learn intentions, actions and plans. In *Proc. AAAI Spring Symposium Help Me Help You: Bridging the Gaps in Human-Agent Collaboration*, March 2011.
16. G. Klein, D.D. Woods, J.M. Bradshaw, R. Hoffman, and P. Feltovich. Ten challenges for making automation a team player in joint human-agent activity. *IEEE Intelligent Systems*, 19(6):91–95, November-December 2004.
17. T. Komatsu and S. Yamada. Adaptation gap hypothesis: How differences between users expected and perceived agent functions affect their subjective impression. *Journal of Systemics, Cybernetics and Informatics*, 9(1):67 – 74, 2011.
18. G.J.M. Kruijff and M. Janíček. Using doctrines for human-robot collaboration to guide ethical behavior. In *Proceedings of the AAAI 2011 Fall Symposium Robot-human team-work in dynamic adverse environments*. AAAI, 2011.
19. G.J.M. Kruijff, M. Janíček, and P. Lison. Continual processing of situated dialogue in human-robot collaborative activities. In *Proceedings of the 19th International Symposium on Robot and Human Interactive Communication (RO-MAN 2010)*. 2010.

20. G.J.M. Kruijff, M. Janíček, and H. Zender. Situated communication for joint activity in human-robot teams. *IEEE Intelligent Systems*, 2012. forthcoming.
21. B. Larochelle, G.J.M. Kruijff, N. Smets, T. Mioch, and P. Groenewegen. Establishing human situation awareness using a multi-modal operator control unit in an urban search & rescue human-robot team. In *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2011.
22. M Liu, F Colas, and R Siegwart. Regional topological segmentation based on mutual information graphs. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.
23. M. Lohse. The role of expectations and situations in human-robot interaction. In K. Dautenhahn and J. Saunders, editors, *New Frontiers in Human-Robot Interaction*, pages 35–56. John Benjamins, 2011.
24. C. Miller and R. Parasuraman. Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors*, 49:57–75, 2007.
25. R.R. Murphy. Human-robot interaction in rescue robotics. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 34(2):138–153, 2004.
26. R.R. Murphy, S. Tadokoro, D. Nardi, A. Jacoff, P. Fiorini, H. Choset, and A.M. Erkmen. Search and rescue robotics. In B. Siciliano and O. Khatib, editors, *Springer Handbook of Robotics*, pages Part F, 1151–1173. Springer Verlag, 2008.
27. T. Oron-Gilad and P.A. Hancock. Multi-modal information display under stress. In P.A. Hancock and J.L. Szalma, editors, *Performance Under Stress*, Human Factors in Defence, pages 251–270. Ashgate, 2008.
28. Panagiotis Papadakis, Ioannis Pratikakis, Stavros Perantonis, and Theoharis Theoharis. Efficient 3d shape matching and retrieval using a concrete radialized spherical projection representation. *Pattern Recognition*, 40(9):2437 – 2452, 2007.
29. E. G. Papadopoulos and D. A. Rey. A new measure of tipover stability margin for mobile manipulators. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3111–3116, 1996.
30. 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*, 1(2):43–68, 2007.
31. Fiora Pirri. The well-designed logical robot: Learning and experience from observations to the situation calculus. *Artificial Intelligence*, 175(1):378 – 415, 2011. John McCarthy's Legacy.
32. F Pomerleau, S Magnenat, F Colas, M Liu, and R Siegwart. Tracking a depth camera: Parameter exploration for fast icp. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011.
33. P.M. Salmon, N.A. Stanton, G.H. Walker, and D.P. Jenkins. *Distributed Situation Awareness: Theory, Measurement, and Application to Teamwork*. Human Factors in Defence. Ashgate, 2009.
34. J. van Diggelen, K. van Drimmelen, A. Heuvelink, P.J.M. Kerbusch, M.A. Neerincx, S.M.A van Trijp, E.M. Ubink, and B. van der Vecht. Mutual empowerment in mobile soldier support. *Special Issue of the International Journal of Battlefield Technology on Human Factors and Battlefield Technologies*. submitted.
35. D.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*, 34(2):210–218, 2004.
36. K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard. OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems. In *Proc. of the ICRA 2010 Workshop on Best Practice in 3D Perception and Modeling for Mobile Manipulation*, Anchorage, AK, USA, May 2010. Software available at `http://octomap.sf.net/`.
37. K. Zimmermann, D. Hurych, and T. Svoboda. Improving cascade of classifiers by sliding window alignment in between. In *The 5th International Conference on Automation, Robotics and Applications*. IEEE, 2011.