Proceedings Second International Symposium on New Frontiers in Human-Robot Interaction, K. Dautenhahn, J. Saunders (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Towards HRI on the Move with Mixed Initiative

Cristian Bogdan (*), Dominik Ertl (o), Michael Göller (+), Anders Green (*) and **Kerstin Severinson Eklundh (*)** [1]

Abstract. When a human and a robot are traveling in each other's vicinity, they typically have reduced recognition capabilities while they interact or communicate. In this paper we posit that the design of a human-robot interface has to consider such reduced capabilities to allow for robust human-robot interaction on the move. As we will illustrate, this brings an important set of design challenges, even more so when the user interface is multimodal and takes mixed initiative into account. In beginning to address such challenges, we are charting a design space for UIs that are dedicated to HRI on the move. The design space is based on chronological communication phases. Furthermore, we point out several modality specifics related to movement, for modalities such as GUI, speech input, speech output and gesture recognition. Finally, we discuss several UI constraints for HRI on the move: walking cognitive workload, time constraints and reduced perception capabilities for the human or the robot. Our example is a semi-autonomous robot trolley in a simulated shopping environment.

1 Introduction

A lot of efforts in human-robot interaction (HRI) today assume a more or less static human-robot positioning, whereby a robot is standing still and expects commands from a human who is also stationary while giving these commands. After these commands are issued, movement of the full robot or human body, or just body parts may occur. In this paper we are interested not just in intermittently moving robots or users (like the one described in [5, 6]), but settings where both robot and user are traveling simultaneously and continuously, and have to interact while on the move. In the course of a European funded project (`http://www.commrob.eu`) working on a semi-autonomous robot we found the need to re-consider HRI in regard to the full body movement of both the human and the robot. For example, using a touchscreen or speech input is much harder for a user in a crowded environment when the robot is allowed to move (displace itself) during interaction. It is easier to interact with a static robot in a such an environment, if the robot moves from time to time but may not be interacted with when "on the move". This is an even more challenging task when both the human and the robot can propose own initiative for interaction (mixed initiative) in a socially acceptable manner, while on the move.

So, to begin addressing some of these "HRI on the move" challenges in the context of mixed initiative we present chronological communication phases charting a design space. We will take our examples from a robot usage situation which we have worked on in the CommRob project, which corresponds to a well-known everyday scenario: traveling with a semi-autonomous robot through a crowded supermarket while shopping. We have worked with a robot capable of supporting such usage, the interactive Behavior Operated Trolley (InBOT, [2]) that we use as development platform and demonstrator in the course of our project. This robot allows for multimodal communication (GUI, speech input and output, and gesture recognition) and safe navigation in a crowded human every-day environments. Figure 1 shows an improved version of the robot trolley.

Figure 1. Robot Trolley with multimodal UI.

During shopping, both the user and the robot can take initiative. We have taken this particular kind of mixed initiative as our *perspicuous instance* of "HRI on the move". Multiple modalities can be used artfully to allow the robot to indicate its initiative in ways that do not disturb the user, yet manage to get the message across. Within a lab-simulated supermarket environment, multimodal, mixed initiative interaction is used by our robotic trolley to help the customer to find the desired products without extensive search and to relieve the customer from the burden to push the shopping trolley from his own force all the time, especially if the trolley is heavily loaded or the customer is an elderly person. The user can interact with the robot by way of a touchscreen, speech input (with a headset), and speech output. The robot has several sensors to localize the user.

We performed a number of user studies and experiments where

[1] (*) Royal Institute of Technology School of Computer Science and Communication, Stockholm, Sweden, {cristi,green,kse}@csc.kth.se, (o) Institute of Computer Technology, Vienna University of Techology, Vienna, Austria, ertl@ict.tuwien.ac.at (+) FZI Research Center for Information Technology, Karlsruhe, Germany, goeller@fzi.de

a total of 30 users have done simulated shopping with our robot in a laboratory environment . The problems they encountered and the issues we found while trying to improve the interaction led us to delimit several categories of "HRI on the move challenges" which we will present below.

The remainder of the paper is as follows. First, we present mixed initiative in the context of a moving robot. Than we introduce the communication phases charting our HRI design space. Third, we present user input and output interfaces for HRI on the move. At last, we discuss constraints like walking cognitive load, time constraints and reduced perception capabilities for HRI on the move.

2 Mixed Initiative on the Move

Mixed initiative allows either the human user or the robot to start an interaction. This allows for more natural interaction between a human and a robot [3]. We present here two examples of mixed initiative from our shopping trolley context.

Suggestion of product: While on route to some product, the robot suggests an item that the user is probably interested in. This can be an item from a shopping list which is being passed by at the very moment. Such a shopping list is administrated via the touchscreen of the robot. For suggesting such a product the robot can choose between the different modalities available to it. In this case slowing down and uttering a verbal comment (e.g. 'we are passing' followed by product name) can be most familiar to the user. This could be assisted by turning the robot in the direction of the item which would catch more attention, thus, reinforcing speech and GUI output.

Suggestion of route: The robot suggests an alternative route which probably poses advantages for the user, e.g., because it is shorter than the originally planned route or the current route became crowded or blocked. Another interesting application can be leading a visually impaired person, who is holding on to the robot's force sensitive handle bar, along a planned route and this way acting as an intelligent white cane. Here the most suitable suggestion might be slowing down a little and slightly turning in the new intended direction while uttering a speech output. This provides the user with an extra modality to react to the robot's action: the resistance force on the handle bar.

3 Communication Phases Charting a Design Space

To illustrate the challenges of mixed initiative while on the move, we propose a design space conceptualization to support a robot initiative around the chronological phases of the human-robot communication: the robot initiative, the human perception of initiative, the human reaction, the robot detection of that reaction, and the robot response to the human reaction. The design space can be understood as a form of help and guidance for a robot interaction designer who wishes to design socially acceptable human-robot interaction for robots who will accompany their users while on the move. We will now show how the design space for supporting robot initiative can be supported by considering the chronological communication phases and by examining each phase in turn. A similar charting can be done for the symmetrical case, when the initiative comes from the human. Throughout, we will consider our theme, "HRI on the move" and the strains put by robot movement, time and other dynamic constraints on the HRI design for our and similar settings. The chronological phases of the human robot communication are depicted in Figure 2.

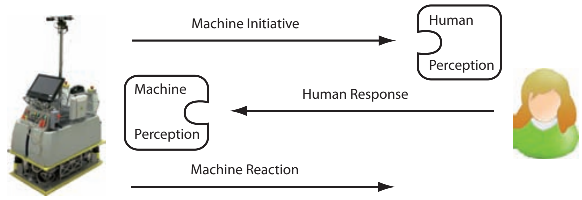

Figure 2. Communication Phases for HRI on the Move.

Machine initiative is the first phase of the process. The design considerations for this phase are depending on the application, i.e. what the initiative is taken about. We have exemplified above two such applications and the possible design decisions in regard to modalities used (speech, GUI) and how to use them. In a lot of cases of interest in our shopping case, this first communication is done on the move.

Human perception of machine initiative follows, after a certain time, the initial initiative. The time needed for the user to notice the initiative, as well as other aspects of the human perception are important factors to consider in a human-robot interaction set-up where human attention may be directed to other components of the environment. If the human is walking along with the moving robot, their attention is affected by the need to look at where to walk. Further robot actions may come too early for the user to comprehend these actions in their logical relation to the initiative.

Human response to the machine initiative can be of different natures. It can range from *tacit perception* whereby the user notices the initiative but does not react in any way that can be understood by the machine, to *acknowledgement* whereby the user reacts to the initiative without attempting to explicitly respond by interacting with the machine, to *interactional response* in which the user interacts with the machine to accept, strengthen, weaken, or reject the machine initiative. For example, for a product suggestion of our robot, the user can acknowledge by slowing down their walking speed, and/or looking at the product, without explicitly interacting with the trolley, yet both such acknowledging cues can help the trolley in shaping its further actions. An interactional response would occur when the user is stopping the trolley via a GUI touchscreen or stopping the trolley by scanning the offered product with a bar code scanner. For route suggestion the user can amplify or reduce the robot turn using the robot's haptic handle. Let us re-emphasize that all such responses are made on the move.

In case of user acknowledgment without further interaction, a *default action* may be defined during design. In some cases acknowledgment without explicit interaction may be enough to consider the robot initiative as accepted, while in other cases it may only account as partial acceptance, thereby weakening the initiative or canceling it altogether. The difference between acknowledgment and explicit interactional response has played a role for us also in evaluating the movement designs with mock-up robots [1] and currently with the InBOT platform. We found that it is much easier to detect whether the initiative was agreed upon by the user (since in that case the user explicitly responds to it) than it is to detect whether it is noticed or acknowledged, especially if the measuring has to be done on the move.

Design plays a role also in *facilitating interactional response*. Not only the launching of the initiative can be designed but also the levers offered to the user for responding to it can be made more available. In the case of a product offering, the user could not only respond by speech or touchscreen GUI interaction, but also by scanning the offered product with a portable scanner, which may be more handy than reaching for the touchscreen. To reject an offer, the user could also employ movement by pushing a trolley which slows down while offering a product or by turning back to the initial direction a trolley that turns to suggest a route.

Such prompting for user interaction is also useful when the user has the initiative, e.g. when the user asks for the trolley to move to a destination, the trolley can start slowly and accelerate progressively, to allow for the user to stop the action or alter the action. While slowly accelerating the robot sensing capabilities are already affected by the robot moving.

Machine detection of human response is fundamental in providing a socially competent robot behavior. Previous work motivates the need for and gives an initial characterisation of a model for communicative contact and perception feedback [7]. While explicit interactional response is straightforward to detect, passive acknowledgement may play an important role in the initiative interaction, yet it may be harder to detect, especially while moving. For example in the case of product offering via slow-down, the user slowing down themselves can mean that the offer is acceptable (though not surely accepted), while the user not slowing down is almost certain to mean offer rejection or lack of notice about the offer. Similarly in our evaluation work (both with a mock-up robot and with the InBOT) we have seen users looking at the offered product but not following this up with any explicit interaction. This inspired us the idea that user gaze detection (while the robot is moving) can be a useful indicator to automatically detect acknowledgment, which can further help in continuing the slow-down robot action for an eventual explicit interaction response. Another detection of acknowledgment (and acceptance) is the posture of the user reaching out to the shelf to grab the product, which can be detected by image analysis. None of these acknowledgment cues represent certainties, yet probability-based techniques similar to those used in speech or gesture recognition can be employed to combine these cues in order to provide a more natural coordination between user and robot.

The product-offering example thus poses at least three important sensing challenges: sensing user moving speed, sensing the user posture in relation to the shelf, and sensing the user gaze; all these are important challenges since detections need to take place while the robot itself moves. Furthermore this example is among the most illustrative to suggest that the interaction design of machine initiative does not end with the design of the initiative itself, but also includes detection of user response or uptake for the machine to be able to continue the interaction 'properly'.

Machine reaction to human response involves the machine follow-up of the initiative based on the user response and its machine detection. Several cases need to be considered here: the user may be acknowledging, accepting, strengthening, weakening or rejecting the machine initiative. All these possibilities may lead to design decisions for a certain application. Furthermore, machine detection of acknowledgment can be followed by detection of acceptance or strengthening (or indeed weakening or rejection) so the machine can react to acknowledgment in a way that is consistent with the following reaction in case of subsequent explicit user interaction.

For example, in case of product offering acceptance, we might want to design the trolley to stop, or to do even more by placing itself in a position which facilitates easier product moving from the shelf to the trolley. Performing such movement may introduce time constraints which may or may not be acceptable in the given communication situation. In case of sensing user acknowledgment without any detection of acceptance or rejection, we could decide to make the robot prolong its slowed-down course, to give more time to the user to decide on the machine initiative.

4 User Input and Output Interfaces on the Move

We have already introduced some challenges for implementing mixed initiative while on the move, in chronological approach, which also describes how an ideal implementation would look like. In what follows, we will take a modality-oriented approach to comment on such "on the move" difficulties as they result from our partial (and less than ideal due to issues we will present) implementation of these mixed-initiative design considerations. We will now illustrate them, as well as the difficulties we expect with the design space aspects that we have not yet implemented.

4.1 Using a GUI with a Touchscreen

Our robot has a touchscreen for GUI input and output. Such a GUI can be used while both the human and the robot are standing. However, it can be used as well when the human and the robot are both moving which poses several problems. We encountered a large number of instances where, although walking close to the robot, the users repeatedly missed targets on the touchscreen. We soon realized that the touchscreen transparent surface has a certain thickness and if it is not looked at from a proper angle (due to walking a bit behind), there is a refraction which makes the users touch the screen lower than they should, hence missing the target. Furthermore, when trying to correct this action, users move their fingers on the screen but it is the first touch that matters for the touchscreen device driver, so the only way to correct the action is to take the finger off the screen and touch again, at the right spot, which no user managed to learn.

This rather simple example shows one interesting issue: our touchscreen is not designed for use on the move. While on a static usage scenario the required viewing angle (with low refraction) is easy to "find" through a correct sitting position, or, lacking that, through trial and error, hitting the targets right on the move is much less successful. While this may indicate that a new touchscreen generation could address such issues for use while moving/traveling, our point with this example is to illustrate that "use on the move" poses specific requirements to input and output devices.

4.2 Using Speech Input and Output

It may be argued (and in HRI it has been argued) that the touchscreen is not the most appropriate device to use on the move, and that stopping the trolley could be performed via, e.g., speech. We soon realized that speech recognition while on the move is a quite different matter from static speech input, as it is made more difficult by the sounds of the robot engines and wheels, by the sound of user steps, by the words being uttered differently by the walking users, environmental sounds, etc. Speech output needs to be adjusted as well, with higher volume due to the sounds specific to robot moving and user walking.

4.3 Using Gesture Recognition

Recognition of gestures performed by a human user is still a matter of research, even more when the user and robot are moving [4]. Such a system recognizes symbolic gestures like "Stop" or "Come To Me" as well as deictic gestures like "Go There", accompanied by showing the direction. This modality relies on the vision system capabilities that are mounted on the moving robot that has to detect and recognize gestures of a (moving) human. It is not only a matter of the underlying computing vision algorithms, as the constraints have to be considered in the interaction design of a robots' multimodal UI as well. Typically, the user has to be in front of the cameras and may not walk around just waving her hands. When the human and the robot are moving (at least one of them), the system may not recognize the gestures correctly. We recognized the required light settings for the vision system to distinguish and recognize colored objects properly as another issue for gesture recognition systems. Optimal light can be achieved comparably easy in a laboratory environment. However, in a real-world example this can lead to a malfunctioning gesture recognition when the human and the robot enter, e.g., a shadowy area.

5 Constraints for HRI on the Move

Besides the modality oriented set of difficulties for mixed-initiative HRI on the move, we identified more constraints for designing such interaction, which we discuss below.

5.1 Walking Cognitive Workload

Slowing down when the robot "shows" something to the user is a functionality of our robot, yet only later we came to reflect on why this is appropriate. Our current understanding is that walking poses a certain cognitive workload which reduces the attention that the user can invest for looking in other directions than the walking direction, looking at specific details like the touchscreen, etc. Reducing the movement speed can thus be necessary if important aspects of the environment must be considered, or if other important facts must be considered in the human-robot communication subject. Again, like for other aspects of the interaction on the move, we wish to consider this cognitive load aspect in a more general way, as a generic condition for HRI on the move.

5.2 Time Constraints

In order for the human-robot communication to take place in a socially acceptable manner, some operations need to be performed within a certain time window. The robot slow-down exemplified a situation where the robot tries to give more time for the user to consider an offer and react.

We can exemplify "latest" time constraints. Some interactions have to be performed in a specific short time frame ("let's go to the right", "oh, stop a second", "no - left, not right"). These are cases where the complete interaction cycle has to take place in few seconds. This means: detection of the need of interaction, taking the initiative and start a communicative act, perception of the communicative act by the other partner, interpretation of the act, performing an action to respond to the communication, detection of the action by the first partner, interpretation by the first partner.

"Earliest" time constraints exist too. Sometime there has to pass a time span which is needed before a reaction to a communicative act is possible or before taking the initiative to start a communicative act is possible. This depends on the available modalities. For example: the available modality is a GUI on a touchscreen: *Robot:* "here is a special offer, if you want me to stop, press the screen!", but the user is 3m away from the robot therefore it takes some seconds for the user to reach the screen.

If both types of constraints (earliest and latest) appear in the same situation, the robot designer must provide heuristics for determining which one is more important, and how does that affect the communication to the user.

5.3 Reduced Machine Perception Capabilities

We have already introduced the reduced human perception capabilities on the human side due to walking. On the technical side, many sensors behave differently (generally more poorly) while on the move. Cameras give fuzzy images, they may give shaky pictures when on a moving robot due to a shaking platform. As mentioned already, noise from the robots drive system disturbs speech recognition. And TimeOfFlight (3D) cameras have great problems with moving objects. Such camera system creates distance data with help of the time-of-flight (TOF) principle. Summed up, electromechanical systems like robots have reduced perception capabilities while on the move.

6 Related work

The Autonomous City Explorer project (`http://www.ace-robot.de`) aims to create a robot that will navigate in an unstructured urban environment and find its way through interaction with humans. Interaction while moving is not as much in focus as four our project yet we believe that the ACE project can constitute a further instance of the "HRI on the move" problematic.

In other instances that illustrate challenges posed by "communication on the move" Leet et al. [10] consider speech-based email during vehicle driving and its effects on attention to the road, and Keskinpala et al. [9] look at interacting with robots via portable devices such as PDAs.

7 Conclusion

In this work we present the challenges we identified when designing mixed-initiative multimodal human-robot interaction while on the move. A major ingredient of these challenges is movement: robot and user movement, their relative distance, relative speed, their slowing down, speeding up, etc. We presented chronology (interaction flow)-oriented, modality-oriented and constraint-oriented charting and illustration of such challenges. Based on such reflections, we believe that considering interaction on the move is an important theme to be addressed by HRI in the future by e.g. working on more intelligent/more robust communication/interaction models, which anticipate communication failures and constrain violations on the move.

ACKNOWLEDGEMENTS

This research has been carried out in the CommRob project (`http://www.commrob.eu`), partially funded by the EU (contract number IST-045441 under the 6th framework programme).

REFERENCES

[1] Bogdan, C.; Green, A.; Hüttenrauch, H.; Räsänen, M.; Severinson Eklundh, K.: Cooperative Design of a Robotic Shopping Trolley, In: Proceedings of COST-298, 2009

[2] Göller, M; Kerscher, T; Zöllner, J. M.; Dillmann, R.: Setup and Control Architecture for an Interactive Shopping Cart. In: Proceedings of the 14th International Conference on Advanced Robotics (ICAR), 2009

[3] Adams, J. A.; Rani, P.: Mixed-initiative interaction and robotic systems. In: Proceedings of AAAI Workshop on Supervisory Control of Learning and Adaptative Systems, 2004

[4] Burger B.; Ferrane, I; Lerasle, F.: Multimodal Interaction Abilities for a Robot Companion. In: Proceedings of International Conference on Computer Vision Systems, ICVS 2008

[5] Kosuge, K.; Hirata, Y.: Human-Robot Interaction. In: IEEE International Conference on Robotics and Biomimetics, ROBIO 2004, pp.8-11, 22-26 Aug. 2004

[6] Goodrich, M. A.; Schultz, A. C.: Human-robot interaction: a survey, In: Journal of Found. Trends Hum.-Comput. Interact., vol. 1, num. 3, 2007, p. 203–275, Now Publishers Inc.

[7] Green, A.: The need for contact and perception feedback to support natural interactivity in human-robot communication. In Proceedings of IEEE 16th International Symposium on Robot and Human Interactive Communication (RO-MAN 2007), pages 552557, Jeju, Korea, August 26-29 2007.

[8] Kaindl, H.; Falb, J.; and Bogdan, C.: Multimodal communication involving movements of a robot. In CHI '08 Extended Abstracts on Human Factors in Computing Systems (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 3213-3218.

[9] Keskinpala, H.K., Adams, J.A., Kawamura, K.: PDA-based human-robotic interface, IEEE International Conference on Systems Man and Cybernetics, 2003

[10] Lee, J.D.. Caven B., Haake S., Brown T.: Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. Human Factors, 43 (4), 2001