Mixing Robotic Realities

Mauro Dragone
School of Computer Science and Informatics
University College Dublin
Belfield, Dublin 4, Ireland
+353 1 716 2855
mauro.dragone@ucd.ie

Thomas Holz
School of Computer Science and Informatics
University College Dublin
Belfield, Dublin 4, Ireland
+353 1 716 2491
thomas.holz@ucd.ie

Gregory M.P. O'Hare
School of Computer Science and Informatics
University College Dublin
Belfield, Dublin 4, Ireland
+353 1 716 2472
gregory.ohare@ucd.ie

ABSTRACT
This paper contests that Mixed Reality (MR) offers a potential solution in achieving transferability between Human Computer Interaction (HCI) and Human Robot Interaction (HRI). Virtual characters (possibly of a robotic genre) can offer highly expressive interfaces that are as convincing as a human, are comparably cheap and can be easily adapted and personalized. We introduce the notion of a *mixed reality agent*, i.e. an agent consisting of a physical robotic body and a virtual avatar displayed upon it. We realized an augmented reality interface with a Head-Mounted Display (HMD) in order to interact with such systems and conducted a pilot study to demonstrate the usefulness of mixed reality agents in human-robot collaborative tasks.

Categories and Subject Descriptors
I.2.9 [**Artificial Intelligence**]: Robotics – *autonomous vehicles.*

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence – *intelligent agents, multiagent systems.*

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – a*rtificial, augmented, and virtual realities.*

General Terms
Design, Experimentation, Human Factors.

Keywords
Human-Robot Interaction, Autonomous Agents, Intelligent User Interface, Augmented Reality, Mixed Reality Agents.

1. INTRODUCTION
Robots have escaped from the shackles of academic research laboratories and now permeate daily life in the form of entertainment robots, household appliances, or assistive technologies in vast numbers [12]. The pervasive situated robot demands that robots behave in a socially competent manner [11], exhibiting social and emotional intelligence, in order for human and robot to understand and anticipate the other's intention [2][9].

IUI'06, January 29–February 1, 2006, Sydney, Australia.
ACM 1-59593-287-9/06/0001.

While implicit communication of a machine's intention - even in a very abstract form - can appear as convincing as those of a human [1], robotic interfaces still remain restricted due to physical constraints and thus fail to convey subtler meanings of intention or emotion. First steps towards building a realistic human-like companion with rich visual expressiveness have been taken [7], but still suffer from limitations and high cost. In stark contrast, however, virtual characters (possibly of a robotic genre) can offer highly expressive interfaces, are comparably cheap and can be easily adapted and personalized. Although HCI and HRI share common issues [13], they are by no means identical and should be regarded as distinct fields of research [2].

This paper contests that Mixed Reality (MR) offers a potential solution to this discrepancy between HCI and HRI. We introduce the notion of a *mixed reality agent*, i.e. an agent consisting of a physical robotic body and a virtual avatar displayed on top of it.

2. SYSTEM ARCHITECTURE

2.1 Development Framework
In delivering the mixed reality agents vision we commission the *Social Situated Agent Architecture* (SoSAA) [5], an agent-based middleware for the development of distributed autonomous systems, and associated tools [4][6]. Cognitive capabilities in SoSAA (i.e. deliberation and planning) are based upon *Agent Factory* (AF) [3], an Agent Prototyping Environment inspired by the *Belief-Desire-Intention* (BDI) [10] agent theory. BDI systems use the mental attitudes of Beliefs, Desire and Intentions in order to represent, respectively, the information, motivational, and deliberative states of the agents. The role of these attributes is to provide the agent with a usable description of both present and future states of the agent's environment.

Robotic control in SoSAA is achieved through the integration of this BDI intentional layer with a reactive, behaviour-based robotic system. SoSAA comprises a design methodology – based upon a taxonomy of goals – by which the robot functionalities are organized using their area of competence and distributed across a number of micro-agents. Higher-level goals are described in abstract, *platform-agnostic* terms. The micro-agents assigned to these goals supervise the activation of those serving lower levels goals. At the lowest levels of abstraction, micro-agents manage specific robot perceptual and actuating capabilities (i.e. activating or adjusting behaviour generation modules).

For practical applications, we consider scenarios in which one or more humans, each wearing a HMD, are working upon a collaborative task with one or a team of robots. Employing our augmented reality interface in a robotic development environment equips the researchers with a live representation of the robots' activities while they are working on a designed task. The augmented reality equipped experimenter can query the robots mental state, debug the live system, and experience a rich visual feedback while observing the real robotic system at work. Animations and other visual clues in the virtual avatar are used to attract the attention of the experimenter on particular situations.

2.2 Mixed Reality Agents

Figure 1 shows a typical view of a mixed reality agent scenario. A user, wearing a computer and a see-through HMD with associated digital camera and microphone, can see the live scene with the real robots, augmented by superimposing synthetic imagery showing the associated avatars and the other components of the augmented reality interface. The overlay between real and virtual images is realized by tracking the position and orientation of the HMD of each observer in the coordinate frame of each observed robot. This information is then used to align the image of the virtual character with the associated real robot.

Figure 1. HMD Interface

As a simple and cost-effective way to implement this tracking we use ARToolkit [8], a software library for the recognition and pose estimation of square markers within a camera image. We arranged five different markers upon the visible faces of a cube (as seen in Figure 1, between the robot and the snowman avatar) to make the robot traceable from all angles.

Since proper overlay of virtual images onto the user's field of vision requires exact knowledge about the position and gaze of the user, we decided to avail of this knowledge and use gaze direction to influence the robot's behaviour, to spatially reference virtual and real objects, and to identify one individual out of a team. The user's gaze is observed in the frame of reference of the cube, and by projecting this vector onto the assumed flat ground plane beneath the robot the intersection point can be calculated and used as, for example, a way point for the robot.

The functionalities of a mixed reality agent emerge from the collaboration of a network of distributed agents: agents in control of the robotic platforms (*robotic agents*), agents managing the user interfaces (*user interface agents*) and agents in control of the virtual avatars (*avatar agents*). In the current implementation, for each user both the user interface agent and the avatar agent reside on the viewer's wearable computer and are connected to the other participating robots and users via the underlying messaging system. More details of this messaging system and the tracking can be found in [6]. Each agent category is now briefly described.

2.2.1 Robot Agent

The robot agent manages the robot conduct. It receives observation updates and user commands and updates other participants regarding its internal state (i.e. sending telemetry and sensorial information when requested or reporting its BDI mental state to the observer). At any given time multiple goals can be active within the BDI system. When the robot agent commits to a new plan, this can contain multiple sub-goals (i.e. plan preconditions) thus leading to subsequent activation of secondary agents, each serving a particular sub-plan. In order to arbitrate between multiple plans, a priority is assigned to each goal.

2.2.2 User Interface Agent

The user interface agent controls the display of text and other 2D graphic overlays in the HMD. Through them, the user can be informed of details of the task and the state of the other participants (both robots and humans). This agent also processes user utterances through the IBM ViaVoice™ speech recognizer and the Java Speech API (http://java.sun.com/products/java-media/speech/). The vocal input triggers a predefined set of commands in order to configure the local interface (i.e. show the mental state of a particular robot) or issue requests to the robots. To do this, the recognized text is matched and unified with a set of templates and the result is transformed into a correspondent ACL directive, e.g. *request(?robot, follow(right))*.

2.2.3 Avatar Agent

The avatar agent is responsible for the rendering of the virtual persona. It monitors the activities of the real robot and manages the visual appearance of its virtual counterpart. Its purpose is to achieve a degree of cohesion within the mixed reality agent by exhibiting behavioural consistency to the observer. This is achieved by activating and managing the animation cycle of a set of expressions, gestures and other animations in harmony with the reception of robot messages or other meaningful events in the robotic agent (i.e. behaviour or plan selection) and in the user interface. In doing so, the avatar agent considers a model of both the avatar (i.e. its identity, its body form and the available animations) and the user (i.e. his identity and preferences) in order to dynamically bind classes of events to the specific visual clues.

2.3 HMD Visualization

In the example in Figure 1, the observer is monitoring two robots. For this particular user, a snowman avatar is used in conjunction with the first robot and a bunny avatar (in the picture moving away from the user) is associated with the second. Debugging information concerning the goal, the behaviour and the state of the robots are displayed at the bottom of the screen.

As the left robot is closer to the user gaze (evidenced with the cross as the centre of the image), the interface automatically tunes to the data broadcast by this robot. Alternatively, the user can explicitly state the robot he is interested in through the menu at the top-left of the screen. Finally, two vertical white bars mark the limits for the user's centre of attention

The robot in Figure 1 has just acknowledged the presence of the observer and triggered a *FaceObject* behaviour using the position of the observer as a reference point. Under the control of the behaviour, the robot rotates on the spot to face the observer. At this point the virtual avatar greets the observer by activating a *Greet* behaviour which in this case results in waving its hat to inform the user that it is now ready to listen to the user's request.

Avatars can be equipped with a wide range of facial expressions and animations (see [5], [6]), which can be associated with the actuation of certain behaviors when commitments are invoked. For example, nodding or saluting to acknowledge a user's command, or shrugging its shoulders if the input has not been understood. Being virtual artifacts, these animations are not merely limited to "natural" human-like forms, but can also include other virtual objects. A flash bulb could for example be displayed when the robot agent is committing to a new plan.

3. DISCUSSION AND FUTURE WORK

In order to realize practical demonstrations of mixed reality agents, we conducted a pilot study in human-robot collaborative tasks, namely the retrieval of a soccer ball by a robot, aided by the user through vocal instructions. We compared task completion time and the number of instructions imparted on the robot for three different groups where the user had: (1) no visual representation, (2) visual representation of mental attitudes in the form of 2D text only, and (3) 2D text and virtual persona.

The quantitative results obtained in the experiment, however, were largely inconclusive due to the limited scale and the set-up of the study. Another reason for this might be due to the inexpressiveness of the employed avatar. A larger and more rigorous user study is necessary in order to confirm and verify our findings. This will include a formal introduction to the system and the behavioural capabilities of the robot, an assessment of the spatial and perceptional skills of the users and a questionnaire to investigate the usability and other aspects of the system.

4. CONCLUSION

Within this paper we have explored the use of mixed reality robotics as an instrument for more effective Human Robotic Interaction (HRI).

We have illustrated the possibilities offered by robotic visual personalisation that matches both the robot's behaviour and the specific robot-user relation. The robotic form can thus be customised as the virtual component can assume multiple forms and project different identities to different users. In addition anthropomorphic behaviours can be ascribed to mixed reality robotic entities to facilitate compelling and engaging interaction. One such example is the lifting of the hat to greet the user or the shrugging of the shoulders if uncertain about an instruction. These character overlays can thus convey a wealth of information that could not easily be achieved by their robotic cousins.

The work described herein is pioneering and has many potential applications. If intelligent robots are to be achieved then roboticists should not be influenced by artificial, self-imposed interaction constraints but should exploit both physical and digital capabilities of the new intelligent machines. The work described herein goes some way to achieving this goal.

5. ACKNOWLEDGMENTS

This research is based upon works supported by the Science Foundation Ireland under Grant No. 03/IN.3/1361.

6. REFERENCES

[1] Bartneck, C. Affective Expressions of Machines. *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI '01)*, 2001, 189-190.

[2] Breazeal, C. Social interactions in HRI: The robot view. *IEEE Trans. on Systems, Man, and Cybernetics, 34* (2004), 81-186.

[3] Collier, R.W. Agent Factory: *A Framework for the Engineering of Agent-Oriented Applications*. Ph.D. Thesis, University College Dublin, Ireland, 2001.

[4] Dragone, M., Duffy, B.R., O'Hare, G.M.P. Social Interaction between Robots, Avatars and Humans. *Proc. of the 14th IEEE Int. Workshop on Robot and Human Interactive Communication (RO-MAN '05)* Nashville, USA, Aug. 2005.

[5] Dragone, M., Holz, T., Duffy, B. R., O'Hare, G.M.P. Social Situated Agents in Virtual, Real and Mixed Reality Environments. *Proc. of the Int. Conf. on Intelligent Virtual Agents (IVA '05)* (Kos, Greece, September 2005).

[6] Dragone, M., Holz, T., Duffy, B. R., O'Hare, G.M.P. Ubiquitous Realities through Situated Social Agents. *Proc. of the 5th Intl. Working Conference of Computer Animation and Social Agents (CASA '05)* (Hong Kong, October 2005).

[7] Hanson, D., Olney, A., Prilliman, S., Mathews, E., Zielke, M., Hammons, D., Fernandez, R., Stephanou, H.E. Upending the Uncanny Valley. *Proc. of the 20th National Conf. on Artificial Intelligence and the 17th Innovative Applications of Artificial Intelligence Conf.*, 2005, 1728-29.

[8] Kato, H., Billinghurst, M., Poupyrev, I. *ARToolKit*. Technical report, Hiroshima City University, 2000.

[9] Breazeal, C., Kidd, C.D., Lockerd, A., Hoffman, G., Berlin, M. Effects of nonverbal communication in human-robot interaction. *Proc. of IROS '05* (Edmonton, Canada, 2005).

[10] Rao, A.S., Georgeff M.P. Modelling Rational Agents within a BDI Architecture. *Proc. of the 2nd Int. Conference on Principles of Knowledge Representation & Reasoning* (1991), 473-484.

[11] Reeves, B. and Nass, C. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, 1998.

[12] United Nations Economic Commission for Europe. *UNECE issues its 2004 World Robotics survey*. UNECE Press Release (Geneva, October 20, 2004).

[13] Yamato, J., Shinozawa, K., Naya, F., Kogure, K. 2001. Evaluation of Communication with Robot and Agent: Are robots better social actors than agents? *Proc. of the 8th IFIP TC.13 Int. Conf. on Human-Computer Interaction (INTERACT '01)* (Tokyo, Japan, July 9-13), 690-691.