WHAT IS MIXED REALITY, ANYWAY? CONSIDERING THE BOUNDARIES OF MIXED REALITY IN THE CONTEXT OF ROBOTS

JAMES E YOUNG[1,2], EHUD SHARLIN[1], TAKEO IGARASHI[2,3]

[1]*The University of Calgary, Canada,* [2]*The University of Tokyo, Japan* [3]*JST ERATO, Japan,*

Abstract. Mixed reality, as an approach in human-computer interaction, is often implicitly tied to particular implementation techniques (e.g., see-through device) and modalities (e.g., visual, graphical displays). In this paper we attempt to clarify the definition of mixed reality as a more abstract concept of combining the real and virtual worlds – that is, mixed reality is not a given technology but a concept that considers how the virtual and real worlds can be combined. Further, we use this discussion to posit robots as mixed-reality devices, and present a set of implications and questions for what this implies for mixed-reality interaction with robots.

Keywords. Human-robot interaction, mixed reality, human-computer interaction

1   Introduction

Mixed reality is a popular technique in human-computer interaction for combining virtual and real-world elements, and has recently been a common technique for human-robot interaction. Despite this popular usage, however, we argue that the meaning of "mixed reality" itself is still vague. We see this as a challenge, as there is a great deal to be gained from mixed reality, and a clear definition is crucial to enable researchers to focus on what mixed reality offers for interaction design.

In this paper, we attempt to clarify the meaning of mixed reality interaction, and follow by relating our discussion explicitly to human-robot interaction. In short, we propose that mixed reality is a concept that focuses on how the virtual and real worlds can be combined, and is not

tied to any particular technology. Based on our definition we posit that robots themselves are inherently mixed-reality devices, and demonstrate how this perspective can be useful for considering how robots, when viewed by a person, integrate their real-world manifestation with their virtual existence. Further, we outline how viewing robots as mixed reality interfaces poses considerations that are unique to robots and the people that interact with them, and raises questions for future research in both mixed reality and human-robot interaction .

2   Considering Boundaries in Mixed Reality

> *Mixed Reality* – "**Mixed reality** refers to the merging of real and virtual worlds to produce new environments and visualisations where physical and digital objects co-exist and interact in real time."[1]

The above definition nicely wraps the very essence of what mixed reality is into a simple statement – mixed reality merges physical and digital worlds. In contrast to this idea-based perspective, today mixed reality is often seen as a technical implementation method or collection of technologies. In this section, we attempt to pull the idea of mixed reality away from particular technologies and back to its abstract and quite powerful general essence, and highlight how this exposes some very fundamental, and surprisingly difficult, questions about what exactly mixed reality is. In particular, we show how robots, and their inherent properties, explicitly highlight some of these questions.

We start our discussion by presenting research we conducted (Young and Sharlin, 2006) following a simple research

---

[1] http://en.wikipedia.org/wiki/Mixed_reality, retrieved 11/11/09.

question: given mixed reality as an approach to interaction, and, robots, we asked ourselves: "if we completely ignore implementation details and technology challenges, then what types of interactions does mixed reality, as a concept, enable us to do with robots?" In doing this, we forced ourselves to focus on what mixed reality offers in terms of interaction possibilities, rather than what we can do with a given implementation technology, e.g., a see-through display device, or the ARToolkit[2] tracking library. We formalized this exploration into a general idea for mapping such an interaction space, and presented exemplary techniques (Young and Sharlin, 2006) – we present the core of this work below, where the techniques serve as interaction examples to be used throughout this paper.

2.1 THE MIXED REALITY INTEGRATED ENVIRONMENT (MRIE)

Provided that technical and practical boundaries are addressed, the entire three-dimensional, multi-modal real world can be leveraged by mixed reality for integrating virtual information. One could imagine a parallel digital, virtual world superimposed on the real world, where digital content, information, graphics, sounds, and so forth, can be integrated at any place and at any time, in any fashion. We called such an environment the "mixed-reality integrated environment", or the MRIE (pronounced "merry") (Young and Sharlin, 2006), and present it as a conceptual tool for exploring how robots and people can interact using mixed reality. Specifically, we used the MRIE as a technology-independent concept to develop a taxonomy that maps mixed-reality interaction possibilities (Young and Sharlin, 2006), and used this taxonomy to devise specific interaction techniques. For our current discussion, we quickly

---

[2] http://www.hitl.washington.edu/artoolkit/

revisit two of the interaction techniques we proposed in our MRIE work: bubblegrams and thought crumbs (Young and Sharlin, 2006).

Bubblegrams – based on comic-style thought and speech bubbles, bubblegrams are overlayed onto a physical interaction scene, floating next to the robot that generated it. *Bubblegrams* can be used by the robot to show information to a person, and can perhaps be interactive, allowing a person to interact with elements within the bubble (Figure 1).

Thought Crumbs – inspired by breadcrumbs from the Brothers Grimm's *Hansel and Gretel* [3] , *thought crumbs* are bits of digital information that are attached to a physical, real-world location (Figure 2). A robot can use these to represent thoughts or observations, or a person could also leave these for a robot to use. These can also perhaps be interactive, offering dynamic digital information, or enabling a person or robot to modify the though crumb.
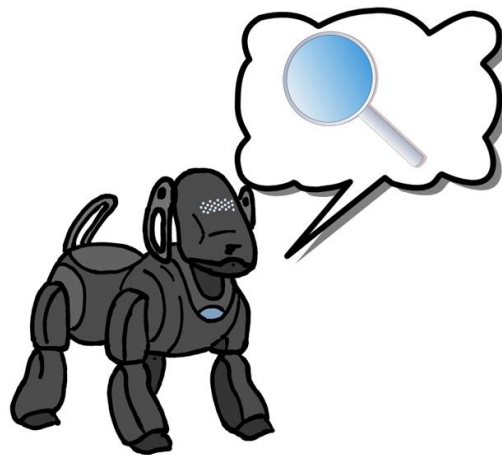

Figure 1. bubblegrams

[3] http://en.wikipedia.org/wiki/Hansel_and_Gretel

Figure 2. thought crumbs, in this case a robot leaves behind a note that a person can see, modify, or interact with later

2.2 BASIC IMPLEMENTATION

Our original bubblegrams implementation (Figure 3) uses either a head-mounted or a tablet see-through display, where the head mounted display setting was used for viewing only, and interaction was only possible through the tablet setting. Using a vision algorithm, the location of the robot is identified in the scene and the bubble is drawn on the display beside the robot. A person can interact with the bubble using a pen on the tablet PC (Young et al., 2005).
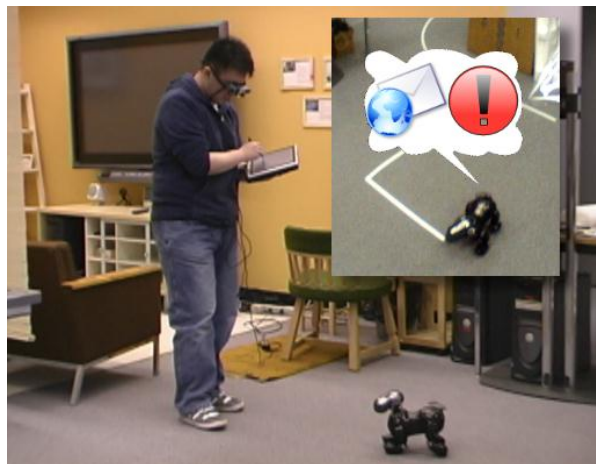
Figure 3. bubblegrams see-through device implementation

Few would argue that this is a mixed-reality system, as it fits a very common mixed-reality implementation mould – see-through display with computer graphics superimposed over real-world objects. However, consider the case where an interface designer does not want to use a bulky hand-held display and opts to replace the graphical bubbles with, perhaps, a display attached to the robot. This display would show the exact same information as in the prior interface but would not require the person to carry any actual equipment – is this still mixed reality?

Perhaps the designer later decides to replace the display with a series of pop-out cardboard pieces, with a clever set of retractable cut-outs and props – possibly mounted on springs to add animation effects. While we concede that there are important differences with this approach, such as a greatly-reduced level of flexibility, this display still represents digital, virtual information and superimposes it in the real world in much the same way (conceptually) as the previous method – is this still mixed reality?

The thought crumbs implementation (Figure 4) uses RFID tags for messages, where the physical tag itself denotes the location of the message, and the message information is stored within the tag. The tags also have human-readable outward appearances, and are supplemented with infrared
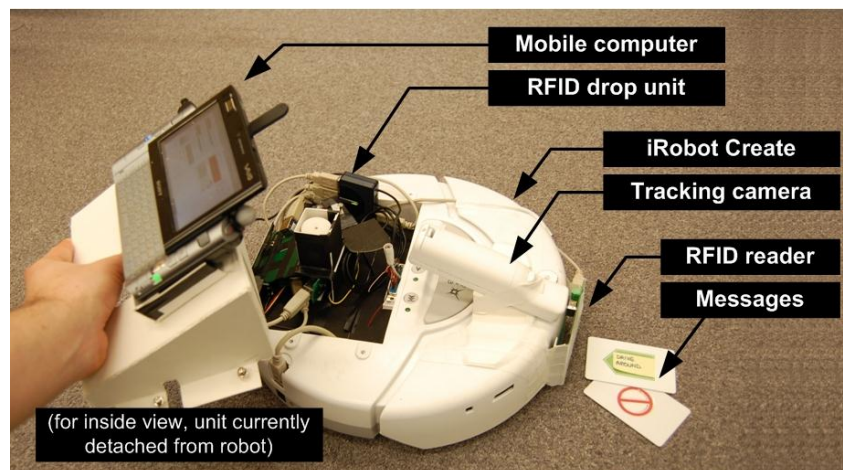
Figure 4. RFID Thought Crumbs implementation

lights so the robot can locate the tags from a distance (Marquardt et al., 2009). In a similar effort, *Magic Cards* (Zhao et al., 2009), paper tags are used by both the person and the robot. A robot can leave representations of digital states or information at meaningful real-world locations as paper printouts, and can read cards left by people, enabling a person to interact with the robot's virtual state through working with physical cards.

Our original thought crumbs discussion (Section 2.1) introduced it as a mixed-reality interaction technique, and in both the implementations shown here virtual information (pending robot commands, system state, robot feedback, etc) is integrated into the physical world through their manifestations. Overall the core concept of the interaction is the same as the original idea, but are these implementations, without any superimposed visual graphics, mixed reality?

The above discussion highlights how easy it is to draw lines on what kinds of interaction or interfaces count as mixed reality, based solely on the implementation technology. We fear that this can serve as a limiting factor when exploring mixed-reality techniques for interaction with robots, and argue that mixed reality should not be limited to or

limited by any particular technology, implementation technique, or even modality (graphics, audio, etc). We see the concept of mixed reality itself as a very powerful approach to interaction, one that can serve as motivation for a plethora of interaction techniques and possibilities far beyond what is possible by the current technical state-of-the-art.

3   Defining Mixed Reality

Should mixed reality be viewed as an interaction device or mechanism, similar to a WiiMote or a tabletop? Or as an implementation tool such as C# or ARToolkit[4] that enables the superimposing of computer graphics via a display device onto the real world? Is mixed reality limited to particular modalities, such as graphics, or, can it include other modalities such as sound or haptic interaction? Or, is mixed reality a more-general approach, such as ubiquitous computing, that spans particular implementations, tools, or modalities?

The common-use definition of mixed reality is difficult to pinpoint, but we believe that it is summed up by our earlier quote (Section 2). Note that this definition itself reveals a muddled stance. On the one hand it clearly describes the general idea of merging of real and virtual worlds. On the other hand, it explicitly focuses the definition toward the modality, "visualizations". This limits and shapes the perspective offered by the definition, where we argue that mixed reality transcends the modalities.

3.1   MILGRAM AND KISHINO

In 1994 Milgram and Kishino presented what is considered to be a seminal discussion of mixed reality (Milgram and Kishino, 1994). This paper's self-proclaimed primary contribution is a taxonomy of graphical,

---

[4] http://www.hitl.washington.edu/artoolkit/

visual displays, and as such the tone of the paper surrounds mixing graphical and real-world environments.

On closer inspection, however, the theoretical discussion of the paper, including the well-known "virtuality continuum", leaves the visual focus behind and is careful to abstract to the more general case. They say mixed reality is combining "real" objects, those that have an "actual objective existence", with "virtual" objects, those objects which exist "in effect, but not formally or actually." Further, the authors directly state that their focus on visual displays is largely related to the current state of technology, and outline that, as technology allows, mixed reality will include, for example, "auditory displays" and "haptic displays" (Milgram and Kishino, 1994). Below we attempt to relate this broader view of Milgram and Kishino's model to current state of the art in tangible, physical and robotic interaction.

3.2   MIXED REALITY AND TANGIBLE-USER INTERFACES

Much of the work in tangible computing revolves around the observation that we, as computer users, are simultaneously living in two realms: the physical one and the virtual one. Tangible user interfaces, then, are devices that are designed to "augment the real physical world by coupling this digital information to everyday physical objects and environments" (Ishii and Ullmer, 1997).

A modality-independent, general definition of mixed reality can be applied to a number of interaction approaches, with physical/tangible interaction being a straightforward extension. Strong parallels can be found between the motivation and meaning behind tangible-user interfaces and the general mixed reality approach of combining the virtual and the physical. Particularly if we discard the technology used or the communication modality (graphics, haptics, aural, etc) it becomes clear that both approaches are similarly attempting to find ways to combine the virtual and the real.

With this we do not mean to lessen tangibles or to imply any debasement as a research area, but to rather bring tangibles, and common understanding of mixed reality, under the same general theoretical foundation. We hope that this unification can help provide focus to the real challenges (and real contributions) that are being faced by these fields. Particularly, we are interested on focusing on interaction, more specifically human-robot interaction, and not any particular implementation tools or technologies.

3.3   REVISITING THE MEANING OF MIXED REALITY INTERACTION

We see mixed reality as the concept of meshing the virtual and physical worlds into one interaction space. If we accept this definition, then there is an immediate problem of scope. For example, would not a mouse, as it couples physical input to virtual cursor state, or even a monitor, which gives a real view (via the photons it emits) of a virtual space, be a mixed reality device? This wide scope raises the question of how this broad definition can be useful or even desirable.

Mixed reality, as a concept, helps to push thinking toward the combination between the virtual and the real. It is useful as a sensitizing concept, or as tool to explicitly focus on the point of meshing. While the mouse is an amazingly successful interface in general, mixed reality highlights the mouse's limitations to mesh the virtual and the real – the link is unidirectional (no inherent physical feedback from the virtual world) and limited to the mouse's two-dimensional relative movements.

As another example, the *Magic Cards* interface described above (Zhao et al., 2009) uses physical print-out cards as a representation of a robot command or feedback message. Mixed reality points out that the paper (and printer) is the medium and sole contact point for bridging the virtual and the physical, and pushes us to consider how real information (e.g., location, real-world tasks) and virtual information (e.g., robot commands, robot feedback) can be linked through this interface. The same analysis

applies for the thought crumb implementation presented earlier, (Marquardt et al., 2009), where RFID tags couple digital information with a particular real-world location (denoted by the location of the tag itself).

While this wide scope may sometimes make it difficult to draw lines on what mixed reality constitutes, thinking of interaction as mixed reality is useful as a tool that explicitly pushes us to consider the mapping between virtual objects, views, or states and the real-world and physical manifestations.

3.4   WHAT MIXED REALITY PROVIDES

The idea of mixed reality as we present it provides only a simple, overarching perspective on interaction and is itself a very limited tool for examining, describing, and exploring interaction. That is, our approach does not supplant existing frameworks, categorizations, or interface design practices. Rather, mixed reality is a point of view from which existing tools can be applied.

For example, we do not consider how to approach interaction or interface design or evaluation, in either the real or virtual worlds. Existing design philosophies, heuristics, and so forth, still apply; mixed reality points toward the meshing point between the virtual and the real.

Further, we do not discuss how such a meshing point could be considered, targeted, mapped, and so forth, as this is already an active area of work in HCI. For example, mixed-reality work like Milgram and Kishino's virtuality continuum (Milgram and Kishino, 1994), tangible computing work such as Sharlin et al.'s consideration of input-/output-space coupling (Sharlin et al., 2004), or even by concepts such as Dourish's "embodied interaction," where the meaning of interaction (and how interaction itself develops meaning) is considered within the tangible and social real-world context (Dourish, 2001). Our approach on mixed reality shows how work such as this can be brought together under a common conceptual foundation.

To summarize, we view mixed reality not as a given technology or technique but as an interaction concept that considers how the virtual and real worlds can be combined into a unified interaction space. Therefore, rather than trying to decide if an interface incorporates mixed reality or not, we recommend that mixed reality itself be used as a tool to help directly consider the convergence points where the virtual and real meet.

4   Robots and Mixed Reality

So far, most of the mixed-reality discussion in this paper could be applied without any particular concern for robots. In this section, we outline how robots bring unique considerations to the table for mixed reality.

4.1   AGENCY

Robots are unique entities in that they have clearly-defined physical, real-world manifestations, can perform physical actions, and can act with some level of autonomy – this sets them apart from other technologies such as the PC (Norman, 2004). These real-world actions can easily be construed as life-like, and people have a tendency to treat robots similar to living entities, for example by anthropomorphizing, and give robots names, genders, and ascribe personalities (Forlizzi and DiSalvo, 2006, Sung et al., 2007). As part of this, people have been found to readily attribute intentionality and agency to robots and their actions. While people attribute agency to, e.g., video game characters and movies (Reeves and Nass, 1996), robots' real-world abilities and presence give them a very distinct, physically-embedded sense of agency that sets robots apart from other technologies.

In some ways, then, interacting with a robot has similarities with interacting with an animal or a person (Young et al., 2008a). The robot itself is seen as an independent, capable entity, and there is a sense of ownership and responsibility that ties the interactions with the robot, and the results of the interactions, back to the robot "entity" itself.

12

4.2   MIXED-REALITY ENTITIES

Robots are mixed reality entities, simultaneously virtual and real. They are virtual in that they are, essentially, a computer with virtual states, abilities, calculations, and a wide range of data in any number of formats. They are real entities in their physical manifestation, where they can interact with the world through this manifestation, both manipulating the world (output) and sensing it (input). As such, we argue that robots are, by their very nature, mixed-reality entities, as a large part of what makes them a robot is how they span the virtual and real worlds – the robot *itself* is a direct coupling of the virtual and the real.

Robots, as mixed reality interfaces, have a very explicit coupling between their virtual and real components. Due to agency, the various (virtual and real) components of the robot are directly attributed to (perhaps owned by) the individual, underlying conceptual agent (robot). The agent itself is directly tied to both the physical and virtual manifestations. This series of connections, supported by agency, means that interacting with robots is fundamentally different from interacting with interfaces that do not have agency; we attribute our interactions with the virtual and physical components directly to the underlying agent.

5   Discussion

We have argued for a wide view on mixed reality, and that robots themselves are inherently mixed-reality devices. What exactly this implies for human-robot interaction with mixed reality is not yet clear, and this is an important area for future consideration. In this section, we outline a few particular questions and challenges raised by this framing that we feel are important to consider.

Ownership and Boundaries – the consideration that robots have a strong sense of agency, coupled with their explicit, physical manifestation, raises questions of ownership and boundaries. For one, robots can (through technical means) claim ownership and enforce interaction constraints on mixed reality elements (*Young and Sharlin, 2006*). However, does this idea of robot / non-robot / human ownership of mixed-

reality entities and items make sense to people in practice? If so, how can such ownership be mitigated and organized? Does this relate to concepts of virtual ownership we are familiar with, such as file permissions, private blogs, or even online finances? Similarly, are their implied boundaries in both the physical world and virtual worlds surrounding the robot as they may surround a living entity, such that, even without explicit ownership, people are careful about interacting in the robots *personal* space? Finally, is there a conceptual difference between the robot's mixed-reality *thoughts* (observations, etc), and ones drawn from the larger virtual world, such as the internet?

Agency – robots are not the only mixed-reality entities to have agency, with a simple example being animated, graphical mixed-reality characters. In this paper we argue that robotic agency is unique for various reasons, but this stance needs to be investigated further: is robot agency different enough from animated mixed-reality characters to merit special consideration? We are currently exploring this through comparing an animated system (*Young et al., 2008b*) to a very similar robotic system (*Young et al., 2009*). Further, if this is the case, what does this difference mean for the design of and interaction with mixed-reality interfaces? Following, the above personal-space concerns explicitly apply to the physical body (and perhaps any virtual manifestation) of the robot – do people have reservations about meddling with the robot itself as they may have for animals or people?

Interaction – if robots are simultaneously virtual and real entities, then what does this mean for mapping interaction with the robot? For example, is there a difference between on-robot-body techniques, such as embedded displays, direct haptic interaction (e.g., handshake), or robot sounds, and off-body techniques, such as projected displays, or thought crumbs left behind? How can people interact with these different types of interfaces?

6 Conclusion

In this chapter, we made the argument for moving the ideas of mixed reality away from the constraints of any particular implementation method

or technique, or interaction modality – mixed reality is simply the mixing of the virtual and the real. Robots, then, fall under this wide perspective as inherently mixed reality devices that simultaneously exist in both realms. This perspective enables us to focus directly on the points of meshing between the virtual and the real, and the interface challenges and decisions related to making this meshing happen the way we want it to.

There are still many questions and challenges to be answered surrounding this outlook. Viewing robots as mixed reality devices does not change what we can do with robots, but it does provide us with a perspective that highlights how a robot exists both in the virtual and real realms, and, we hope, encourages us to consider what this means for interaction.

References

Dourish, P. (2001). *Where the Action Is: The Foundation of Embodied Interaction*. The MIT Press, Cambridge, MA.

Forlizzi, J. and DiSalvo, C. (2006). Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, 2006. HRI '06, Salt Lake City, USA, March 2–4, 2006*, pages 258–256, New York, NY, USA. ACM, ACM Press.

Ishii, H. and Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In *ACM Conference on Human Factors in Computing Sysems, 2007. CHI '97, Atlanta, GA, March 22–27, 1997*, pages 234–241, New York, NY, USA. ACM, ACM Press.

Marquardt, N., Young, J. E., Sharlin, E., and Greenberg, S. (2009). Situated messages for asynchronous human-robot interaction. In *adjunct proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (Late-Breaking Abstracts), 2009. HRI '09, San Diego, US, March 11–13, 2009*, New York, NY, USA. ACM, ACM Press.

Milgram, P. and Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems*, E77-D(12):1321–1329.

Norman, D. (2004). *Emotional design: why we love (or hate) everyday things*. Basic Books, New York, USA.

Reeves, B. and Nass, C. (1996). *The Media Equation: How people treat computers, television, and new media like real people and places*. CSLI Publications, Center for the Study of Language and Information Leland Standford Junior University, Cambridge, UK, first paperback edition.

Sharlin, E., Watson, B., Kitamura, Y., Kishino, F., and Itoh, Y. (2004). On tangible user interfaces, humans and spatiality. *Personal and Ubiquitous Computing*, 8(5):338–346.

Sung, J.-Y., Guo, L., Grinter, R. E., and Christensen, H. I. (2007). "my roomba is rambo": Intimate home appliances. In *UbiComp 2007: Uniquitous Computing*, volume 4717/2007 of *Lecture Notes in Computer Science*, pages 145–162. Springer Berlin / Heidelberg, London.

Young, J. E., Hawkins, R., Sharlin, E., and Igarashi, T. (2008a). Toward acceptable domestic robots: Lessons learned from social psychology. *Int. J. Social Robotics*, 1(1).

Young, J. E., Igarashi, T., and Sharlin, E. (2008b). Puppet master: Designing reactive character behavior by demonstration. In *ACM SIGGRAPH / EG SCA '08.*, pages 183–191, Germany. EG Press, EG Association.

Young, J. E., Sakamoto, D., Igarashi, T., and Sharlin, E. (2009). Puppet master: A technique for defining the actions of interactive agents by demonstration. In *Proc. HAI Symposium 2009, Dec. 2009, Tokyo (in Japanese), Presentation*.

Young, J. E. and Sharlin, E. (2006). Sharing Spaces with Robots: an Integrated Environment for Human-Robot Interaction. In *Proceedings of the 1st International Symposium on Intelligent Environments, 2006. ISIE '06, Cambridge, UK, April 5–7, 2006*, pages 103–110, Cambridge, UK. Microsoft Research Ltd., Microsoft Research Ltd. Press.

Young, J. E., Sharlin, E., and Boyd, J. E. (2005). Implementing bubblegrams: The use of haar-like features for human-robot interaction. In *Proceedings of the second IEEE Conference on automation Science and Engineering, 2006. CASE '06, Shanghai, China, October 8–10, 2006*, pages 298–303, Los Alamitos, CA, USA. IEEE Computer Society, IEEE Computer Society Press.

Zhao, S., Nakamura, K., Ishii, K., and Igarashi, T. (2009). Magic cards: A paper tag interface for implicit robot control. In *Proceedings of the ACM Conference on Human Factors in Computing Systems, 2009. CHI 2009, Boston, US, April 4–9, 2009*, New York, NY, USA. ACM, ACM Press.

Index

Biography

James (Jim) E. Young is currently a Ph.D. Candidate at the University of Calgary, under the supervision of Dr. Ehud Sharlin (University of Calgary) and Dr. Takeo Igarashi (Tokyo University). James is a researcher

with both the University of Calgary's uTouch research group (part of the Interactions Lab, Calgary, Canada) and JST Erato Laboratory (Tokyo, Japan).

His research focuses on human-robot interaction, tangible interfaces, human-computer interaction, and computer vision. James has led several unique research projects exploring the use of mixed-reality techniques in human-robots interaction, and designing robotic behaviors based on programming-by-demonstration.

James completed his B.Sc. at Vancouver Island University in 2005. His senior project, titled "Space-Model Machine" and directed by Dr. David Wessels, was on modeling logic problems in multi-agent systems. James is currently pursuing his Ph.D. in Computer Science at the University of Calgary and is expected to graduate in early 2010.

Ehud Sharlin has been faculty at the University of Calgary's Computer Science Department since October 2004, following his position as faculty with the Human Interface Engineering Laboratory at Osaka University between 2003 and 2004. He is currently running the uTouch research group, and is a member of the University of Calgary Interactions Lab. His research interests are directed at interaction with physical objects and entities: human-robot interaction, tangible user interfaces, mixed reality and computer game interfaces.

Ehud completed his Ph.D. in Computing Science in July 2003 at the University of Alberta in Edmonton, Canada, under the supervision of Dr. Ben Watson and Dr. Jonathan Schaeffer, his M.Sc. in in 1997 (Magna Cum Laude), and his B.Sc. in 1990, both in Electrical and Computer Engineering from Ben-Gurion University, Israel, under the supervision of Dr. Jonathan Molcho..

Between 1991 and 1998 Ehud worked as a senior researcher and research director with several Israeli R&D labs, his main research themes during this period were image processing, computer vision and tracking algorithms, electro-optical design, numerical simulation, and user studies.

Takeo Igarashi is an associate professor at the computer science department, at the University of Tokyo. He is also directing the Igarashi Design Interface Project, JST/ERATO. He received his Ph.D from the department of information engineering, the University of Tokyo in 2000. His research interest is in user interface in general and his current focus is on interaction techniques for 3D graphics and robots. He received the ACM SIGGRAPH 2006 significant new researcher award and the Katayanagi Prize in Computer Science.