Moving a Media Space into the Real World through Group- Robot Interaction

Jim Young, Gregor McEwan, Saul Greenberg, and Ehud Sharlin
Department of Computer Science, University of Calgary
2500 University Drive NW., Calgary, Alberta, CANADA T2N 1N4
T+1 403 220 6087
{jyoung, mcewan, saul, ehud}@cpsc.ucalgary.ca

ABSTRACT

New generation media spaces let group members see each other and share information. However, they are separate from the real world; participants cannot see beyond the video, and they cannot engage with people not attending to the computer. To solve this problem, we use a robot as a physical surrogate for a media space group, which allows this distance-separated group to extend their interactions into the real world. Through video, all media space group members see a first-person view of what the robot sees. All have opportunity to control it: where it walks, where it looks, and even the sound it makes. The robot becomes a physical tele-embodiment of the group, representing it for people who may not physically be part of the group but are collocated with the robot.

Categories and Subject Descriptors

K.4.3 [Computers and Society]: Organizational Impacts - *Computer-supported collaborative work*.

Keywords

Media spaces, casual interaction, group-robot interaction.

1. INTRODUCTION

Unplanned casual interaction is critical to how small groups work together [5][13]. While these happen naturally in co-located settings, casual interactions are difficult when people are separated by even small distances. In response, groupware developers have designed a myriad of informal awareness and casual interaction tools; each tool typically provides mechanisms for displaying awareness information that leads to casual interactions between distributed members. Several systems work by creating a virtual space that is shared by all group members, where all can see and interact within it, e.g., Instant Messengers [9], chat rooms / MUDS [3], and video-based media spaces [1]. The problem is that these systems are separate from the real world; participants cannot see beyond the computer, or engage in people outside of it. That is, the virtual world is separate from the social practices of the people in each individual environment.

Our approach to mitigating this problem is through Human-Robot Interaction (HRI), where the mobile robot acts as a substitute for the distant person. The research question we are interested in is: can robots serve as a way for a distance-separated group to extend their media space interactions into the real world? That is, can we create a tele-embodiment that is *many* to one, where all see, hear and optionally control what the robot does as it moves and interacts with people in the real world? To answer this question, we added group-robot interaction capabilities to the Community Bar media space [7] via a new media item called *AIBO Surrogate.* After summarizing related work in HRI and the Community Bar system, we describe how *AIBO Surrogate* works.

2. HUMAN ROBOT INTERACTION

HRI addresses human issues in how people control and interact with robots. Existing HRI themes explore how groups control a robot, how it is perceived as a social entity, and how a robot can be part of a two-way video/audio tele-embodiment.

Group robot control. In some cases, a group needs to work together to control a robot as it performs a task, e.g., when a pilot and a sensor operator control an unmanned aerial vehicle [14][4]. Highly coordinated shared interaction is usually required because of the complexity of robot control (due to its many degrees of freedom) and because the various task parameters are beyond the abilities of a single operator. Research issues include how the group coordinates their control through HRI awareness [4]. Other scenarios of shared human-robot interaction involve a group of uncoordinated humans and a single robot, e.g., when a group orders food from a single robotic waiter [14]. Here, the research is on how the robot manages conflicts arising from their requests.

Robot tele-embodiment. The Personal Roving Presence (PRoP) is an excellent example of how an untethered tele-robot provides the sensation of tele-embodiment in a remote space [10]. PRoP is a mobile robot that includes two-way video and audio. The controlling person can control the distant robot's movement, and see and hear via video/audio what the robot sees. A screen attached to the robot displays a video of the controller so that the remote person can see who the robot represents. Physical camera direction and an attached hand pointer gives gaze and gesture awareness. GestureMan [6] is somewhat similar, although in this case the device is built to support remote instruction that requires precise pointing and frequent mutual monitoring. PRoP, GestureMan and other tele-embodiment systems are usually one to one, i.e., one person is controlling/viewing the system, where that robot serves as that person's physical avatar in the distant space.

Our work was designed with the above HRI themes in mind, where we leverage a robot as a controllable social entity. The difference is that we emphasize group-human awareness through a robot: the robot becomes a surrogate of the media space group within the real world, where it represents selected aspects of the group's shared being, presence, awareness and tasks.

3. COMMUNITY BAR

Community Bar (CB) is groupware media space intended to support causal interaction within a small distributed group; its

Cite as:
Young, J., McEwan, G., Greenberg, S. and Sharlin, E. Moving a Media Space into the Real World through Group-Robot Interaction. Report 2006-827-20, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, T2N 1N4. March 17th.
Includes Video (Duration 1:50) plus 4 page paper

design rationale, interface and implementation is described fully in [7][8]. Figure 1 illustrates what it looks like.

In essence, CB presents itself as a peripheral sidebar display [2], divided into *Places*. For example, Figure 1 shows 2 places called '*mike test*' and '*CSCW class*'. Each place represents a sub-group and displays their communication, their tools, and their shared information. These are visualized through a number of *media items* presented at three levels of granularity. The media item's *tile* view is always visible in the sidebar, and represents things like people's presence (as live video, photos or names), public conversations (as public chat dialogues or sticky notes), or publicly shared information (e.g., web pages of common interest and photos). For example, in Figure 1 tiles in the '*mike test*' Place show two people present, a running text conversation, as well as a photo and web page comprising public information posted by group members. Individuals can choose to explore and interact with that information by mousing over a tile, which displays its *tooltip grande* next to it. For example, Figure 1 shows the tooltip grande for the *AIBO Surrogate* control (discussed in Section 4). Finally, a person can click on the tooltip grande's title bar to raise the *full view window,* which displays even more information and interaction capabilities. Figure 2 illustrates the full view for the *AIBO Surrogate* control, also discussed in Section 4.

For each media item, its tile view generally shows awareness information; its tooltip grande shows more detailed information and allows partial interaction; while the full view shows all the information, communication, and interaction possibilities. Fundamental to the philosophy of the Community Bar is that all media items within a place are publicly visible to all people in that place, i.e., it serves as a virtual communal shared setting.

4. THE AIBO SURROGATE

We created a new Community Bar media item, called the *AIBO Surrogate*, which bridges the CB media space into the real world through group-robot interaction. The robot is a non-threatening Sony AIBO robot dog that is typically located in the physical spaces occupied by one or more CB users, e.g., a shared laboratory for researchers, an office corridor for office workers, or an extended family's home. Multiple spaces can be occupied by multiple robots. The main idea is that the robot acts as a controllable physical surrogate for the group, where the robot can wander through the space, interact with the people that inhabit it, and see/transmit visual information within it. We describe how this works by walking through the system's interface.

4.1 Tile View

The *AIBO Surrogate*'s *tile view* is illustrated in the middle of the CB sidebar in Figure 1. Similar to other tile views, it provides the CB group with peripheral awareness information. This particular tile shows a real time low-resolution streaming video of what the robot sees at this moment. That is, the group acquires a first-person (actually a first-dog) point of view of robot activity in the distant space. Through this video stream, the group is aware of when the robot is moving and when its point of view changes (e.g., as it moves its head). If a group member is currently controlling the robot, the *in use indicator* is checked.

Using Figure 1 as an example, the group sees through the tile view that the robot is currently in use, i.e., that it is being controlled by another group member. The chat also mirrors this interaction, where one person tells the other "I'm using the *AIBO Surrogate* to look for Rob", and the other responds with a suggestion about where to find him and an offer to take over control. Through the first-person video, the group also sees the robot's view as it moves through the space to Rob's desk, and when it looks up at Rob as he is seated at that desk. Figure 3 shows what is actually happening in the real world: the AIBO dog is behind Rob looking up at him.

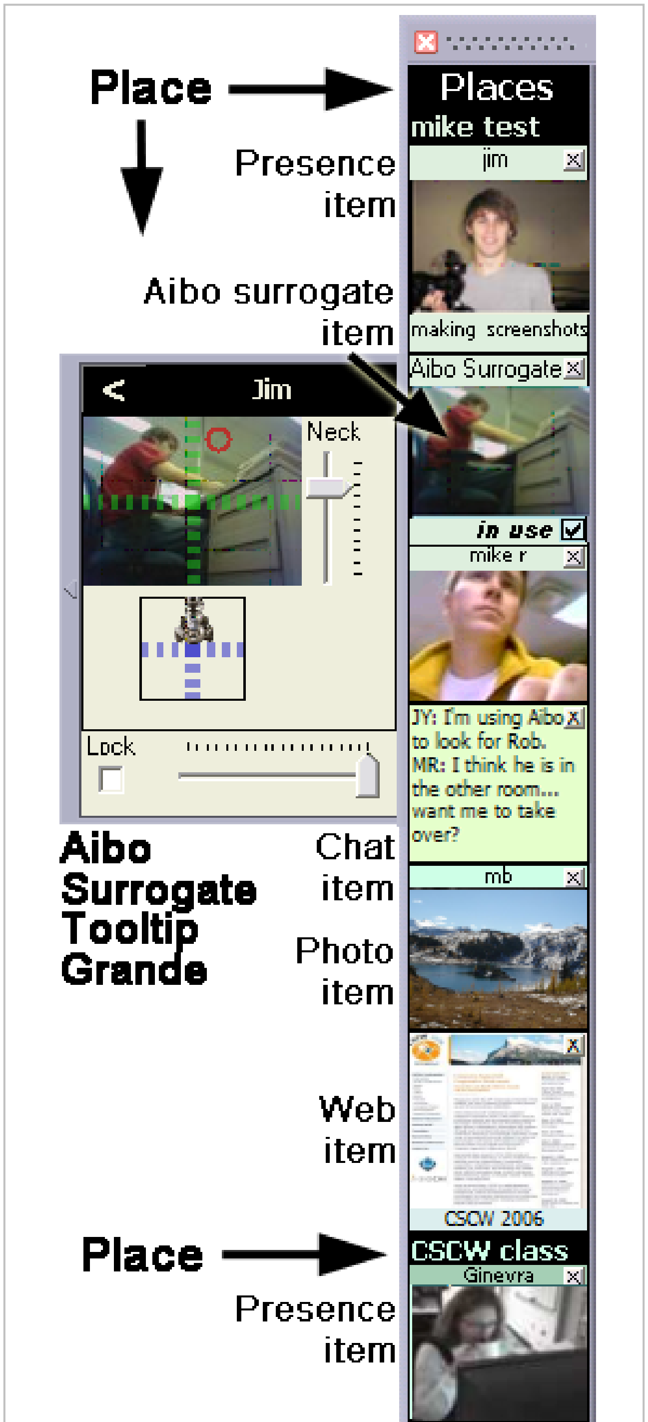

Figure 1. Community Bar: 2 places, a variety of media item tiles, and the AIBO Surrogate tile and tooltip grande.

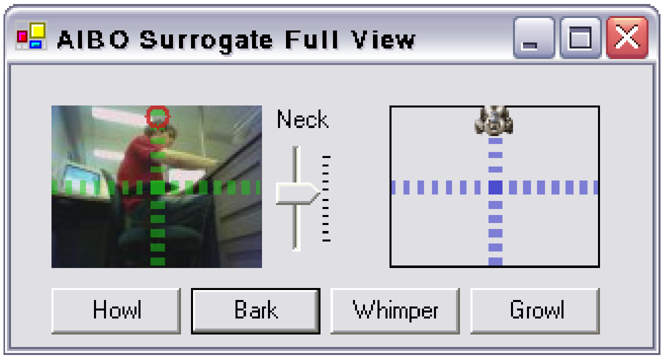

Figure 2. Full Window view of the AIBO Surrogate Item

Figure 3. How AIBO Surrogate appears in the real world

4.2 Tooltip Grande

The *AIBO Surrogate*'s *tooltip grande* (Figure 1, left side) also displays the robot's video stream, but adds three controls for manipulating the robot: a *neck tilt control*, an *interactive look control*, and a *walk control*.

The video now offers a somewhat higher resolution video stream than in the tile view. Thus participants can quickly mouse over and raise this tooltip grande to see more detail.

The *interactive look control*, overlaid atop this live video stream as green crosshairs, allows the users to point the robot's head in a given direction by clicking a spot within this video stream. The green crosshairs represent the axis of control, where the X axis represents the head's absolute pan value and the Y axis represents the head's absolute vertical tilt value. Clicking in the center of the image makes the AIBO look straight ahead, while clicking in the top right corner would make the AIBO look to the extreme top-right limits of its view. The red dot appears to indicate the AIBO's current head posture; this is a gaze indicator that informs the viewer about the limits of the AIBO's head movement. For example, because the red dot is near the top middle of the image, the viewer knows that the robot cannot tilt its head up much further, but that there is still room to pan the robot's head sideways. Also, the center of the cross-hairs serve as a direction indicator that supplies the robot's human controller with feedback information crucial to making the AIBO walk in a correct direction. Of course, the video itself provides valuable feedback to the controller, letting him/her adjust motion and gaze direction on the fly.

The *neck tilt control,* located on the right of the tooltip grande, is another way for the group member to control the robot's up and down neck tilt by adjusting the slider position. This slider position corresponds directly to the position of the neck, where the top and bottom positions reflect the neck being at the highest and lowest positions respectively.

The *walk control*, shown at the bottom of the tooltip grande as blue crosshairs and a robot dog icon, lets the person control the robot's movement. Clicking left or right of center controls how quickly the robot turns left or right. Clicking above or below the center controls how quickly the robot walks forward or backward. Clicking the center tells the robot to stop moving. Thus the controller can make the AIBO walk and turn in any direction by clicking and holding on a desired position. Control feedback of the current walk parameters is provided by the robot icon snapping to the mouse position and following the mouse as it moves. The robot maintains the target speeds until the person selects another location, or until the click is released: in the latter case, the icon returns to centre and the robot stops moving.

If any of these controls are being used, the *in use indicator* on the tile view is checked, but only as the action is occurring. If an action only takes a moment to do, then the *in use* indicator is only briefly checked. This approach lets group members interleave command actions rather than follow a strict floor control strategy.

Using Figure 1 as an example, one person in the group had noticed in the tile view that the robot is looking at Rob seated at his desk. To see who is nearby Rob, that person can raise the tooltip grande, and quickly adjusts the neck to pan left or right.

4.3 Full View

The *full view window* (Figure 2) is similar to the tooltip grande, but gives a larger walk control (for fine-grain manipulation) and adds several additional robotic controls through a series of *notification* buttons.

Controllers can click the various notification buttons to direct the robot to emit a sound: a howl, bark, whimper, and growl. These are used by the controller to attract attention and to communicate intent. People nearby the robot gain awareness information not only by seeing the AIBO move and change its gaze, but by hearing it 'speak' to them. These notifications can be used by remote users as a means of communicating a little extra information. For example, the howl is loud and indicates urgency, the bark is a simple and neutral way to get attention, the whimper indicates a plea or request, while the growl indicates anger or annoyance.

4.4 What People in the Physical World See

Through tele-presence, the robot acts as a surrogate for the group in the physical space. Ideally, people in the real world would also see the robot as a surrogate – a social extension - to the group. If one person sees the robot moving around, that should be as if the group as an entity was moving through the space. If one sees the robot looking up at them, they should realize that the group can see him or her. If the robot exhibits social mannerisms that indicate an attempt to make contact (searching behaviors, head postures and gestures, eye gaze, sounds), then that person may realize that someone is using body language to say something. A person may realize from the robot's behavior that someone is trying to find and contact them, and may respond by going onto the Community Bar to find out who it is and to initiate contact. Ideally, the person should be able to speak directly to the robot and thus to the group, but this capability is not yet implemented.

5. IMPLEMENTATION

The AIBO ERS 7 is an off the shelf programmable robot produced by Sony [11]. The *AIBO Surrogate* owner item communicates with and controls the robot by the Tekkotsu framework software [12]. This framework provides a network interface to various AIBO elements, including head pointing, walking, video stream, and noise notifications.

The *AIBO Surrogate* media item was developed using the media item plug-in capabilities of the Community Bar and its underlying distributed model-view-controller networking and data sharing capabilities [8]. Like many Community Bar media items, *AIBO Surrogate* has both an *owner* and an *audience* variant: the owner is the person who posted the media item, while the audience includes all others in the CB group. In our design, this separation is required for technical reasons rather than for giving different powers to the owner *vs.* the audience. Because of underlying network issues involved with communicating with the AIBO robot, only the owner item can connect to and communicate with the robot: all audience items get and receive robot commands and information indirectly, as these are relayed to them by the owner item. There is no difference from the user's perspective, as all have the same user interface and control capabilities.

6. PRELIMINARY EVALUATION

The *AIBO Surrogate* is a fully functional proof of concept. However, it is still a 'toy' system, as the robot moves too slowly for people to want to move it large distances. Still, we have used it informally in the university laboratory on multiple occasions to get a sense of what it would be like to use in practice.

From a UI perspective, the controlling media space group members managed to use all functionality of the interface with no instruction. They expressed interest in the system, saying that it is not only easy to control, but fun to use. In addition, people were able to self-manage simultaneous control issues, i.e., on the rare occasion when people tried to send simultaneous commands to the robot (e.g., one says go left, the other to go right), they quickly noticed and corrected this situation. Good wish list request ideas included:

- a higher resolution video stream to make navigation easier,
- scene construction, as the AIBO looks around, a larger field of view image is reconstructed from the video frames taken.

From a media space perspective, responses by people who were co-located with the robot were mixed. Some noticed when the AIBO robot dog was headed their way, and gave it attention or waved happily. That is, they treated the robot as if it were a social surrogate of the group. Others disliked the robot for privacy reasons. In one instance, a woman mentioning that she was glad she was not wearing a skirt (the robot dog's ground-level point of view means it is always looking upwards at people). Another person used garbage cans to block the AIBO from entering his region of the lab. This conflict is typical of media space situations, where the benefits of having group awareness are tempered by privacy concerns. The actual adoption / non-adoption of such a system as well as how it is used in practice will likely depend on the cultural practices that a group develops over time.

7. CONCLUSION

The AIBO robotic dog becomes a mechanical extension to the media space group, an awareness tool shared by all members. The people co-located with the AIBO acquire a physical awareness of the robot's (and thus the group's) actions within their space. These people are aware of where the robot is going and can see what it is doing. In addition, the media space group can explicitly contact others in the physical world by controlling the *AIBO Surrogate*'s sounds (howl, bark, whimper, growl). We stress that this is group-robot interaction: as one member controls the robot, other people in the CB group are aware of what the controller is doing, and can even partake in the interaction. Overall, we view the *AIBO Surrogate* as an effective physical two-way awareness tool shared between the people currently on the media space and those people co-located with the robot.

Acknowledgements. Research partially funded by Nectar NSERC Research Networks Grant. Thanks to Min Xin for assistance.

8. REFERENCES

[1] Bly, S.A., Harrison, S.R., and Irwin S. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment, *Comm. ACM*, 3(1), (1993), 28-47.

[2] Cadiz, JJ, Venolia, G.D., Jancke, G., and Gupta, A. Designing and Deploying an Information Awareness Interface. *Proc ACM CSCW*, (2002), 314-323

[3] Curtis, P., Nichols, D. A. MUDs Grow Up: Social Virtual Reality in the Real World*. Proc 39th IEEE COMPCON* (1994), 193-200.

[4] Drury, J. L., Scholtz, J., and Yanco, H. A. (2003). Awareness in Human-Robot Interactions. *Proc IEEE Conf Systems, Man and Cybernetics* (2003).

[5] Kraut, R., Egidio, C., Galegher, J. Patterns of Contact and Communication in Scientific Research Collaboration. In *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work.* LEA (1990) 149-181.

[6] Kuzuoka, H., Ishimoda, G., Nishimura, Y., Suzuki, R., Kondo, K. Can the GestureCam be a Surrogate?. *Proc ECSCW* (1995) 181-196.

[7] McEwan, G., and Greenberg, S. Supporting Social Worlds with the Community Bar. *Proc ACM Group* (2005), 21-30.

[8] McEwan, G., Greenberg, S., Rounding, M. and Boyle, M. Groupware Plug-ins: A Case Study of Extending Collaboration Functionality through Media Items. Report 2006-822-15, Comp. Science, U. Calgary, Canada (2006)

[9] Nardi, B.A., Whittaker, S., and Bradner, E. Interaction and Outeraction: Instant Messaging in Action, *Proc ACM CSCW* (2000), 79-89.

[10] Paulos, E. and Canny, J. PRoP: Personal Roving Presence. *Proc ACM CHI* (1998).

[11] Sony. AIBO ERS-7M3. http://sony.net/Products/ aibo/, March 06

[12] Tira-Thompson, E. *Tekkotsu: A Rapid Development Framework for Robotics*, Master's thesis, Robotics Institute, Carnegie Mellon University, May, 2004.

[13] Whittaker, S., Frolich, D., and Daly-Jones, O. Informal workplace communication: What is it like and how might we support it? *Proc ACM CSCW*, (1994).131-138

[14] Yanco, H. and Drury, J. Classifying Human-Robot Interaction: An Updated Taxonomy. *Proc IEEE Conf Systems, Man and Cybernetics*, (2004)