Avatars at a Meeting

Safurah Binti Abdul Jalil
Department of Computer Science
University of Auckland
sabd058@aucklanduni.ac.nz

Jingwen Huang
Department of Electrical and Computer Engineering
University of Auckland
jhua082@aucklanduni.ac.nz

Marin Markovich
Department of Electrical and Computer Engineering
University of Auckland
mmar149@aucklanduni.ac.nz

Brabbyn Osburn
Department of Electrical and Computer Engineering
University of Auckland
bosb005@aucklanduni.ac.nz

Mike Barley
Department of Computer Science
University of Auckland
mbar098@cs.auckland.ac.nz

Robert Amor
Department of Computer Science
University of Auckland
trebor@cs.auckland.ac.nz

ABSTRACT
The development of remote avatars has recently generated increased research and commercial interest. Current approaches utilize simple remote-user-guided screens to represent the remote participant. Though humanoid robotic systems are significantly more expensive this work investigates the added benefit from utilizing such a robot. Two recent projects examined the potential of humanoid robotic systems to operate as a remote avatar within a meeting context and their impact on meeting dynamics and interactions. These projects identified the utility of human-like gestures as a significant benefit of humanoid robots within such a setting as well as a range of disruptive impacts due to the operational mode of humanoid robots.

Categories and Subject Descriptors
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors.

Keywords
Telepresence, Remote avatar, Gesture, Meeting.

1. INTRODUCTION
Attendance at meetings is a normal part of almost every job, though getting to a meeting is not always easy to achieve. Coordinating several people to the same place at the same time typically requires individual attendees travelling from one place to another, often incurring significant time costs. If meeting attendees could participate in a meeting remotely without significant loss of ability to be effective within the meeting context then there could be significant time savings for many attendees, as well as beneficial environmental impacts such as less travel and hence a reduction in traffic congestion and pollution.

There are currently many products and services available to users wishing to remotely be involved with a meeting. The most traditional method is simple conference calls. This involves the conference operator or host who sets up the call and then the other uses connect to the host [1]. Conference calls are good for a small number of people as it can be difficult to identify the current speaker [2]. Also when more than one person is talking then no one can be easily understood, this is a significant drawback of this as a conferencing method [2]. A more modern approach is video conferencing which addresses some of these issues. There are many video conferencing services available. One of the most popular is Skype which uses the internet to send and receive both video and audio. Skype allows for video conferencing with up to ten different participants [3]. Video calling addresses the issue of identifying the current speaker by providing visual cues which also reduces the likelihood of participants talking over each other. Studies have shown that the addition of video results in more fluent conversation over just straight audio conferencing [4]. More subtle information can also be conveyed through facial expressions which improves the level of communication [5].

Telepresence is a term used to describe a set of technologies that allow a person to feel as if they are present without physically being present [6]. The first telepresence system was TeleSuite, initially developed in 1993 for a branch of resorts in the U.S.A. [7]. The TeleSuite was augmented with mobile robotics to support demonstrations outside the TeleSuite room; this introduced the new field of telepresence robots.

Telepresence robots have the advantage of being mobile and maneuverable, while providing all the essential features of a traditional telepresence system. A standard telepresence robot is usually equipped with a camera, multiple microphones and loudspeakers. These devices allow the robot to capture the video and audio of its surroundings and send it to the remote user. The robot can also be teleoperated, allowing the user to explore a large environment and communicate with people at various locations.

In this project we examine the use of humanoid robots for a very restricted form of telepresence (i.e., a formal meeting). In this context we argue that their ability to perform human-like gestures will allow for greater utility within the meeting than existing commercial telepresence robots. As our available humanoid robot

CHINZ '12, July 02 - 03 2012, Dunedin, New Zealand

(a Nao) has particularly slow mobility we have had to investigate a context where the robot can be mostly located in one position.

2. BACKGROUND

In computing, the term avatar can be loosely defined as an interactive character used to represent a person. The usage of the term avatar as "audiovisual bodies that people use to communicate with each other" [9] was made popular by Stephenson in the cyberpunk novel "Snow Crash" [10]. From this perspective, avatars would be computer generated virtual characters which users can control, such as those in video games. But with Hollywood movies such as Avatar (www.avatarmovie.com), Gamer (gamerthemovie.com) and Surrogates (chooseyoursurrogate.com), the notion of remotely controlled avatars has now been extended to a physical representation rather than restricted to just virtual characters. This heightens the public's acceptance of telepresence applications.

2.1 Gesture

When working within a meeting setting the communicative intent of gestures is of significant importance, and should be offered by the remote avatar. Hassink and Schopman highlight the distinction between a gesture and other body movements being that a gesture is a body movement that is used to deliver a message [11].

This is not unlike in the field of animation where as part of the twelve principles of animation introduced in "Disney Animation: The Illusion of Life" [12], an animated character's action is divided into three parts: anticipation (preparation for an action); the action itself; and follow through (the termination of an action). We take this into consideration when keying in the poses for our remote avatar.

Affective gestures were also examined for our remote avatar. These can be defined as "relating to an external expression of emotion associated with an idea or action" [13], and we used this term to indicate non-verbal gestures in our gesture set that are more emotionally inclined such as frustration and delight, as opposed to a more neutral gestures such as raising an arm to alert others of a wish to speak.

2.2 Commercial Telepresence Robots

A small number of telepresence systems are available commercially and were investigated in this project.

Anybots is a Californian based communications company founded in 2001 who provide business class telepresence robots [14]. Anybots now sell the QB telepresence robot as illustrated in Figure 1. The QB can glide around a room quietly and smoothly, while giving its remote user complete visibility and access to the surroundings. The adjustable head allows the QB various planes of view. Users controlling the robot can use high-definition zoom to get a closer look of people and objects. QB is also equipped with a LCD display that shows the face of the remote user.

VGo communications is an American based company that specializes in telepresence robots and released their line of telepresence robots (called VGo) early in 2011. VGo has an integrated camera, microphones, and video display, all on a light-weight, motorized, remote-controlled platform. VGo is 4 feet tall to work as well with people who are sitting or standing [15].

3. NAO ROBOT

In contrast to the QB and VGo telepresence systems we are utilizing a humanoid robot within a meeting context. The humanoid robot available to the project team for this work is the Nao robot. Nao is an autonomous, programmable, humanoid robot developed by the French company Aldebaran Robotics (Figure 2) [16].

Figure 1. Anybot's QB Telepresence robot [14]

Figure 2. Nao humanoid robot [16]

The robot stands at 58cm tall, weighs 5kg, has a battery life of approximately 1.5hr, and can 'walk' at 0.3km/hr. We considered that its light weight would make it portable and easy to set up within a meeting context, and the slow ambulatory pace would not be a major disadvantage in a normal meeting setting.

4. GESTURE DEVELOPMENT

In considering the different communication purposes of our humanoid robot (and informed by Hassink and Schopman [11]), we have divided our consideration of gestures required into three main categories: non-verbal gestures; speech supporting gestures; and physical contact gestures. Within each of these categories we examined typical meeting processes and determined a set of gestures that appeared as if they would provide for coverage of the needs of the remote meeting participant (see Table 1).

To test the utility of such gestures within a meeting we implemented a subset of these gestures on the Nao robot to match a simple meeting scenario that could be tested with class participants. Since the Nao does not have controllable facial features, we rely solely on its body gestures. Because of this deficit, we chose to exaggerate these gestures to convey particular emotions in the meeting. Even that comes with further constraints. For example, we found that when doing the trial runs with the Nao that: the robotic arms and legs are fairly rigid and can't easily be moved to the pose we intended; the Nao only has three fingers and they can only open or close all at once; the Nao's torso cannot be rotated; only the head can rotate to indicate facing direction; etc. Table 2 shows the meeting gesture set that was finally implemented for use in a test scenario.

Table 1. Categorised gestures useful in meetings

Categories	Intentions
Non-verbal Gestures	**Affective** □ Approval / agreement □ Disapproval / disagreement □ Positive (delighted, happy, etc.) □ Negative (impatient, frustrated, angry, disappointed, etc.) **Alert** □ Request to speak □ Distress signal **Greet** □ Hello □ Goodbye
Speech Supporting Gestures	**Indicating sizes** □ E.g. Big, large, small, little, tall. **Pronouns** □ I / me / myself □ You / yourself **Directions** □ I / me / myself □ You / yourself
Physical Contact Gestures	**Human-initiated touch** □ Pat **Robot-initiated touch** □ Tap **Cooperative touch** □ Hand-shake □ High-five

Table 2. The meeting gesture set for the Nao

Hello everyone	Surprised
Request to speak	Disappointed
I	Disappointed & Sad
You	Darn
Big	Relief
Small	I don't know
Agree	Wipe forehead once
Disagree	Wipe forehead twice
Thinking	Wait for signals

5. EVALUATING NAO IN A MEETING

In order to demonstrate how Nao works in a meeting environment and get our human users to better evaluate the interactions with a telepresence robot in a meeting, we staged a live meeting scenario with Nao as one of the participants. To achieve this we scripted a meeting about a project inception that contained sufficient interactions for Nao to perform the featured gestures in our meeting gesture set. The remote user controlled the Nao utilizing a set of pre-recorded gestures made available through a menu in the Choregraphe control software that comes with the Nao. We video recorded the meeting for reviewing purposes (see: www.youtube.com/watch?v=HdcjkFlCS1I).

Following the scripted meeting the participants were provided with a questionnaire to detail their perceptions of the remote avatar in the meeting. From this limited usability evaluation a number of issues were identified by the participants and the developers of the meeting gesture set. In general we found that participants could relate to the concept of using the Nao to represent a remote participant, though the interactions were often forced and the Nao was provided with greater time to 'perform' than a normal meeting participant would have been. The following issues were identified.

5.1 Timing of Gestures

Participants sometimes found the gestures that the Nao did were not necessarily well timed. For example, if a gesture takes a bit too long to be invoked by the remote user, or if the gesture takes some time to complete. If another participant is trying to speak at the same time as a gesture starts on the Nao then it introduces a degree of distraction to the speaker. Also, the speaker may have to wait for the Nao to complete a significant portion of a gesture to understand its meaning, impacting on their flow of conversation in the meeting.

5.2 Use of Simple Gestures

Some of our participants found that it is easier to understand the simpler gestures. They find a gesture that includes too many fine movements to be confusing. For example, all of our users could understand the greeting gestures – Nao raises and waves its right hand. But they could not really figure out one of our speech supporting gestures – 'Small', with the Nao slowly moving its hands closer together.

We theorise that this is because the waving hand gesture is commonly recognised and used by people in daily life. So it becomes easy for our meeting participants to adopt the message. However, people might use different gestures when they want to indicate that something is 'Small', some might use one hand and their fingers some might use both hands. But they are normally using quite a small scale of movement. Because of some of the posing limitations of the Nao, we had to made movements exaggerated, which might have made the movement not as natural as human ones.

5.3 Size of the Nao

Some users identified that the size of the Nao might be too small to be used as a meeting avatar. The size of the Nao is slightly bigger than a children's robot toy. It is easy to give people the impression that the Nao is just the type of robot that should be doing entertaining things. However, as a meeting is mostly formal a toy looking robot might not be easy for other participants to imagine that is the same as the remote user.

The small size of Nao can limit its ability to interact directly with human meeting attendees. A study conducted by Walters et al. concluded that shorter humanoid robots (such as the Nao) were viewed as being less conscientious than taller robots and were thus perceived as being childlike [17]. Participants in this study reported that due to its size and construction the Nao had a toy like appearance.

5.4 Sound Distraction at the Meeting

The Nao robot is moved by the internal motors located at different joints of the Nao's body. Whenever the Nao moves the motor's moving sound is obvious, especially in a quiet meeting environment. Our meeting attendees reported that such motor sounds brought some degree of distraction to the meeting. It consistently drew the attention of attendees even when the Nao was just moving its head to make observations instead of showing a gesture.

5.5 Live Streaming Video Quality

From the remote participant's point of view, it is important for them to know how the meeting is going. As an attendee who is sitting in a meeting, we would normally look at the person who is speaking to observe body language and facial expressions. We made the Nao perform the same gaze gesture by using the Nao's sound recognition feature. It 'hears' the source of the sound and tries to face the correct direction so that the remote attendee could then see the person speaking. However, we found that if the next speaker was in the opposite direction to where the Nao was looking then the Nao would need to rotate its head to the right direction very quickly in order to look at that person. During such a movement the video image quality is blurry and it was easy to cause some motion sickness effect to the remote participant who is concentrating on the streaming video.

We also identified that several of the Nao's gestures utilized head movement and this impacted the video view of the remote participant for the period that the gesture was being undertaken. For example, nodding the Nao's head would render the view of the meeting incomprehensible until the nodding stopped.

5.6 Physical Issues with the Nao

The Nao robot has a 55 Watt-Hours battery that gives the robot approximately 1.5 hours of usage. However, this battery life can be significantly shorter (~45 minutes) if the Nao is performing lots of movements; in the case of the remote avatar this equates to their gestures. While 45 minutes may be long enough for most meetings the main limiting factor for meeting length with a Nao robot is its battery life.

Also, due to the closed design of the Nao's body it is prone to overheat with use. When the Nao is in use, all of the joints must be stiff to keep it standing upright or holding another position. For the joints to be stiff a constant current must flow through the brushed DC motors which cause them to heat up. After approximately 25 minutes the Nao will shut itself down to prevent damage from high temperatures. If the joints are being used (for example to perform gestures) it will overheat slightly more quickly. This has a big effect on the usability of Nao as a remote avatar as once it overheats it takes at least 5 minutes to cool down enough to run again.

6. CONCLUSIONS

This project examined the use of a humanoid robot as a remote avatar within a meeting context. Using the development environment available with the Nao robot system we were able to successfully prototype a set of gestures alongside voice and video communication with the remote participant. The telepresence system was tested with a simple meeting scenario using student participants, and was followed up with evaluations and an examination of the recorded video of the scenario.

The evaluation of this system showed that meeting participants would accept a humanoid robot in a meeting scenario as a remote avatar. Though as the chosen robot (Nao) had no ability to change its facial appearance it was difficult to make the association between the robot and the remote participant.

An analysis of our chosen gestures showed that we still need to improve some of our gestures to be more immediately recognizable and so that their duration can be shortened. Many of the gestures added little to the normal flow of the meeting and further research into necessary gestures would likely determine that a very small set would be sufficient for a meeting context.

The physical characteristics of the available humanoid robot impacted the efficacy of gestures and also had an impact on the normal running of a meeting (e.g., noisy movements). Other humanoid robots might not impact as significantly in these areas.

7. REFERENCES

[1] Investopedia. 2011. *Conference Call basics*, investopedia.com, [Accessed August. 20, 2011]

[2] Hyder, M., Haun, M. and Hoene, C. 2010. Placing the Participants of a Spatial Audio Conference Call. In *Proc of the 2010 7th IEEE Consumer Communications and Networking Conference (CCNC),* 9-12 January, 1-7.

[3] Skype. 2011**.** *Group video calling - host real-time video conferences on Skype***.** Available: http://www.skype.com/intl/en-us/business/group-video?intcmp=gvcCS-bus [Accessed August. 20, 2011]

[4] Daly-Jones, O., Monk, A. and Watts, L. 1998. Some advantages of video conferencing over high-quality audio conferencing: fluency and awareness of attentional focus. *Intl. Journal of Human-Computer Studies*, 49, 1, 21-58.

[5] Bruce, V. 1996. The role of the face in communication: Implications for videophone design. *Interacting with Computers*, 8, 166-176.

[6] Minsky, M. 1980. Telepresence. *OMNI Magazine*.

[7] Lichtman, H.S. 2006. Telepresence, Effective Visual Collaboration and the Future of Global Business at the Speed of Light. *HPL, Human Productivity Lab Magazine*.

[8] VGo Communications. 2011. Available methods to achieve "being there". Retrieved from: http://www.vgocom.com/solutions on (August 27, 2011).

[9] Nijohlt, A., Zwiers, J. and Peciva, J. 2009. Mixed reality participants in smart meeting rooms and smart home environments. *Personal and Ubiquitous Computing,* 13, 1, 85-94. http://dx.doi.org/10.1007/s00779-007-0168-x

[10] Stephenson, N. 1992. *Snow Crash*. Bantam Books, N.Y., USA.

[11] Hassink, N. and Schopman, M.G. 2006. *Gesture recognition in a meeting environment.* Master's thesis. University of Twente, Enschede, The Netherlands. http://essay.utwente.nl/56211/

[12] Thomas, F. and Johnston, O. 1984. *Disney Animation: the illusion of life*. Abbeville Press, N.Y., USA.

[13] Microsoft Encarta. 2011. Affective definition. http://uk.encarta.msn.com/encnet/features/dictionary/DictionaryResults.aspx?lextype=3&search=affective

[14] Anybots, Inc. 2011. Anybots, about us. Retrieved from: http://www.anybots.com/about on (September 5, 2011).

[15] VGo Communications. 2011. VGo, Overview. Retrieved from: http://www.vgo.com/productsandservices on (September 5, 2011).

[16] Aldebaran Robotics. 2011. Aldebaran Robotics, the creators of Nao. Retrieved from: http://www.aldebaran-robotics.com/en on (June 20, 2011).

[17] Walters, M., Koay, K.L., Syrdal, D., Dautenhahn, K. and te Boekhorst, R. 2009. Preferences and Perceptions of Robot Appearance and Embodiment in Human-Robot Interaction Trials. In *symposium at the AISB09 convention*, 8, 136-143.