

Mixing Virtual and Physical Participation: The Future of Conference Attendance?

Shervin Shirmohammadi¹, Shun-Yun Hu², Wei Tsang Ooi³, Gregor Schiele⁴, and Arno Wacker⁵

¹University of Ottawa, Canada, Email: shervin@site.uottawa.ca

²Academia Sinica, Taiwan, Email: syhu@iis.sinica.edu.tw

³National University of Singapore, Singapore, Email: ooiwt@comp.nus.edu.sg

⁴Digital Enterprise Research Institute (DERI), Galway, Ireland, Email: gregor.schiele@deri.org

⁵University of Kassel, Germany, Email: arno.wacker@uni-kassel.de

Abstract—The rapid advancement and ubiquity of social virtual environments is bringing geographically distant users to interact as if they are in the same physical location, leading to the emergence of new application areas that adopt virtual environment technologies. In this article, we present our experiences of enhancing The 3rd International Workshop on Massively Multiuser Virtual Environments (MMVE 2010) by virtual environment technology to allow participants to attend remotely. Based on a survey conducted with both the virtual and physical participant, we find that virtual participation is a valuable addition for conference hosting, but informal social interactions may still be missing and require further support.

I. INTRODUCTION

Since the beginning of virtual reality, the possibility to construct networked computer-based worlds has been an important vision of computer science [1]. Today, this vision is getting closer to reality. Users can interact with each other in virtual worlds such as World of Warcraft or Lineage II, similar to the way they would in the real world. Recent advances in Broadband Internet and computing technology have given rise to a new class of Networked Virtual Environments (NVEs) [1], where multiple concurrent users can participate in the simulated world by connecting over the Internet. Users can interact with each other in real-time, giving rise to new social spaces like Second Life, There.com, or Sony Home. While we have seen many applications of NVEs in entertainment, their true potential is yet to be explored. One such potential is to allow participants of a conference or workshop to meet virtually and to save costs both for the organizers (local accommodations) and the participants (travel costs). An NVE's sense of immersion could be effective in achieving this goal, allowing multiple participants to see each other, form discussion groups, and socialize. Such NVE-based virtual conferences have been attempted before. One of the earliest is the MASSIVE system that supports online meetings, where participants report heightened enjoyment (e.g., the medium and experience is fun) [2]. The tool, however, is experimental and not used for real venues. Later systems have reported collaborative applications in NVEs, called Collaborative Virtual Environments (CVEs) [3], [4]. CVEs, however, focus more on the co-creation, co-manipulation, or collaborative achievement aspects of online interactions, as opposed to high fidelity information exchange in conferencing. Perhaps the largest

and most complete work in this area is that of [5], where the authors report their experience with a 500-person 3-day conference held purely virtually in Seconds Life, consisting of keynotes, poster sessions, and social sessions, but no paper sessions. While some advantages were observed in using NVE technology, the consensus was that the system had sufficient shortcomings to conclude that a purely virtual conference is not yet an appropriate substitute for the real thing, summarized by this attendant's statement: "It was the first virtual meeting to cross the threshold to being useful! ... But I'd still rather travel 30 hours." [5].

But what about enhancing a workshop or conference with NVE? That is, not to do the workshop purely virtually, but to enhance it with NVE technology so that some participants and presenters can attend remotely. In theory, this would still lead to cost savings for both the organizers and some of the attendants. Having organized three years of the International Workshop on Massively Multiuser Virtual Environments (MMVE), we felt that the technology is ripe enough for serious real-world applications, beyond socializing and games. Therefore, as proponents of virtual environments, we decided to use an actual NVE to enhance our workshop. Enhancing a conference has a long history, from the early text-based chat rooms, to today's video and voice discussions, to large-scale webcasting. The two main design dimensions are the levels of interactivity and scalability (Table I), which also represent the nature of the interactions: presentational versus conversational. Conversations typically require back and forth information exchange and a high degree of interactions, while presentations consist mostly of unidirectional information flow, suitable for a large audience.

While these conventional enhancements increase a conference's quality of experience, they are still far from how people would interact in a real space: immersion, sense of presence, and more natural interactions similar to the real world are still missing. In response, recent work has looked at more advanced techniques to enhance such meetings. For example, to empower virtual participants who are typically marginalized, "back channels" (essentially a live bulletin board) through which people can pose questions have been used to enhance Q&A sessions to determine the audience's most popular questions [6], empowering virtual participants to have a stronger

TABLE I
DESIGN DIMENSIONS OF CONVENTIONAL ENHANCEMENTS TO CONFERENCES/WORKSHOPS

	Low Interaction (presentations)	High Interaction (discussions)
Small-scale (2-10 people)	Stored video streaming (e.g., TED, YouTube tech talk)	Video/VoIP meeting (e.g., Skype meeting)
Large-scale (50+ people)	Live video broadcasting (e.g., webcasting)	Text chat room (e.g., IRC channels)

say in what question should be asked. Also, to engage user attention more effectively, Shami et al. have successfully used avatars (human-like virtual representations of users) in online meetings to foster interaction and team bonding [7]. In some NVEs, it is possible to mix physical and virtual participants, for example, Tang et al. utilize video overlays to project the arms and hands of remote participants in a shared work space, in order to overcome presence disparity [4]. More recently, Bessiere et al. report on using Second Life to support business users in certain collaborative tasks [8], although they find that prior training is needed, and motivation was low as people saw virtual worlds as time-consuming games.

Despite the interesting research results of the above work, we did not find any reported literature on using NVEs to enhance an academic venue such as a workshop to aid remote participation, and so we decided to experiment with this concept in our workshop. Our main goal was to enhance a physical workshop with virtual participation to widen the potential audience. Specifically, we wanted to enable presenters unable to attend physically to present their papers to both physical and virtual attendees, and allow discussions to be held directly between virtual and physical participants. Unlike studies with purely-virtual conferences, our results are more positive and show that a gradual move towards a mixed venue system might be a better approach.

Our paper is structured as follows. In the next section, we describe our setup at the workshop site and how we mixed the physical and virtual environments. After that we present the results of a small user study, which we conducted with the participants of our workshop and discuss our findings and lessons learned. Finally, we wrap up the paper with a short conclusion and outlook on future work.

II. WORKSHOP SETUP AND DEPLOYMENT

After considering various options, we decided to use the Unreal-engine based web.alive tool from Avaya [9], instead of the more well-known tools such as Second Life, as Second Life is not primarily designed for conference presentations, while web.alive is designed with business meetings in mind and includes full-room speech, moderation, and high resolution slide display. Previous studies have shown that high end computing needs and setup effort of tools such as Second Life are a barrier to their usage in conferences, and that a more light-weight setup helps the adoption of NVEs in such settings [10]. This is confirmed by [5], where close to half of the participants agreed that Second Life did not work well. Web.alive is quite light in memory and CPU footprint, and works as a plugin in browsers, making it much easier to install and accessible for our target users.

As our attempt was an initial experiment on the feasibility of virtual participation, we decided that the remote participants must receive some a priori training. The people invited included organizers who were unable to attend in person (including most of this article's authors), presenting authors (we had John Miller from Microsoft Research, who was unable to board the plane at the last minute), or interested colleagues of the presenting authors. Another reason to limit the virtual participants was due to the system's limit (the free-trial version supported 8 maximum concurrent users). In all, we had 9 virtual participants (not all concurrently) coming from Canada, USA, Germany, and Singapore. Contrary to studies that business users are not easily motivated to use NVE for work purposes [8], in our scenario the motivation was stronger, as the alternative would have been complete absence. However, some participants did not show up for orientation, leading to problems described later.



Fig. 1. Double-screen configuration at the workshop showing the virtual environment (left screen) and the virtual speaker's slides (right screen)

One important question was how to actually setup equipment in Taipei, given the capabilities of the remote system. Our intended requirements were: (1) to have the physical audience see the virtual participants, so there is a sense of an extended space; (2) to allow virtual participants see and hear presentations from physical presenters; and (3) to allow physical participants present as normally as possible. With the help of local organizers, we setup a separate projector screen alongside the main slides screen to show the virtual environment (Figure 1). Local attendees could thus have a view of who was participating remotely. The web.alive system allowed a user to share his or her desktop for others to see, and also upload slides in PDF format, which can be viewed

within the virtual environment. We first tried the desktop sharing feature but found it too slow and unresponsive for an engaging presentation, even with our 10 Mbps download and 1.56 Mbps upload network at the workshop. We thus decided that presentations should be displayed using the built-in PDF slide features. Note that for the virtual attendees, the physical audience was represented by a single avatar of the NVE operator (explained next); i.e., the virtual participants were not seeing a live video feed of the physical participants, although they could hear the physical audience.

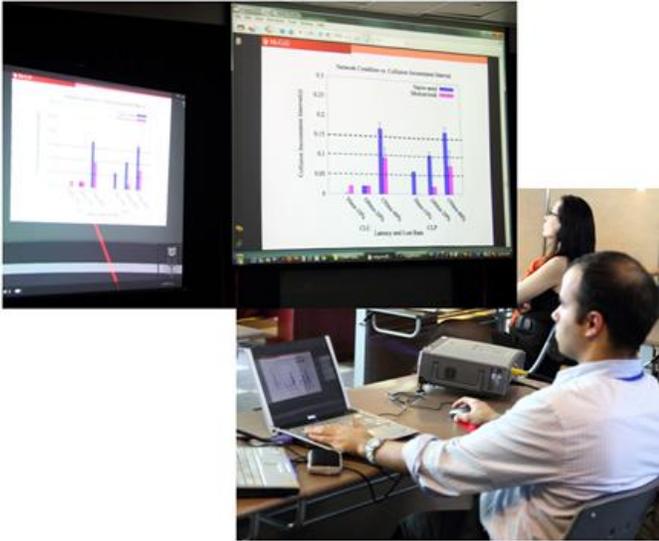


Fig. 2. The NVE operator, advancing slides and using the virtual laser pointer (left) in synchrony with the physical speaker.

To ensure that the presentations were synchronized, we allocated an NVE operator to help both physical and virtual presenters (Figure 2). The NVE operator would control the physical speaker’s avatar in the NVE, upload slides to the system before each presentation, and closely synchronize slide changes with the physical presenter. web.alive allows any avatar to flash a virtual laser pointer at the screen, so our NVE operator was also able to highlight specific parts on each slide to the virtual participants, if the physical presenter did so (as seen in Figure 2 left). Each physical presenter thus was able to present normally using his or her own laptop. In contrast, during a virtual presentation, the NVE operator would turn the slides on-site for the physical audience.

To ensure that remote participants can navigate and use web.alive, we invited the remote attendees to two test sessions before the actual workshop for briefing on basics such as: navigation, conversation, asking questions (by raising virtual hands), zooming into slides, and speaking in front of the virtual microphone. This pre-workshop orientation appeared quite useful, since two of the virtual attendees who missed it did not know how to efficiently navigate, speak, or see slides more clearly.

During our early test trials, voice was found to be important for the immersion between people, which is consistent with

TABLE II
CHARACTERISTICS OF USERS IN THE SURVEY

Participant Type	Virtual	Physical	Overall
Avg. age (year old)	29	33	31
Avg. number of conference attendance / year	1.8	2.3	2.0
Avg. hours using Social Virtual Environments / week	5	1.9	3.3
Avg. number watching online technical talks / month (e.g., iTunes U, YouTube)	3.6	3.4	3.5

earlier findings [3], [5]. So voice communication between the physical and virtual participants was supported by hooking the conference room’s microphone and speaker system to the NVE operator’s laptop. This way, whenever the physical speaker talks, the voice is sent to web.alive’s server and can be heard by all virtual participants, while any words spoken in the NVE can also be heard by all physical attendees through the conference room’s speaker system. The extra equipment cost about USD \$340 for the workshop day, which was affordable for the organizers.

III. EXPERIENCE AND EVALUATION

The day after the workshop and while memories were fresh, we sent an online survey to all physical and virtual participants. We had 19 people (9 remote participants and 10 physical attendees) respond to our survey questions, no later than 3 days after receiving the survey. Their characteristics are shown in Table II. The specific questions asked and the participant responses are shown throughout this section.

It should be noted that the participants’ interpretation of “Social Virtual Environments” was using Facebook (73.7%), Twitter (12.1%), and/or LinkedIn (10.5%). This shows that none actually used an NVE on a regular basis. The experiment was conducted during the workshop, which was held from 11:00 to 18:00 local time in Taipei, while the virtual participants attended from Singapore (10:30 to 18:00 local time), USA (19:00 to 23:30 local time), Canada (22:00 to 23:00 local time), and Germany (9:30 to 11:00 local time). The results of the survey are shown and analyzed next.

A. Virtual Participants Feedback

Figure 3 shows a typical screen shot of what a virtual participant would see. Here, three participants are viewing the presentation given by John Miller, himself participating virtually. The avatar in white is the NVE operator inside the virtual environment.

The resolution and clarity of the graphics (important for presentations) were quite good due to the zooming capability. As can be seen from Figure 4, a virtual participant can zoom into a slide at great magnifications to see the small details more clearly.

We collected feedback from the virtual participants on how effective it was to attend the workshop. We asked the virtual participants to rate their ability to see, hear, and speak through web.alive from 1 to 5, with 1 being the worst and 5 being the



Fig. 3. Screen shot of the virtual conference room with John Miller from Microsoft presenting his talk.

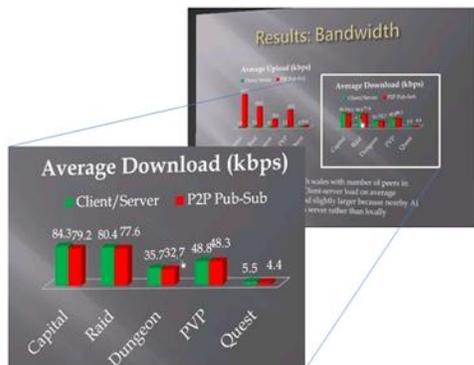


Fig. 4. Visual granularity and zooming abilities

best. We also gave participants the opportunity to comment freely on their experience. The summary is shown in Figure 5. We also asked the virtual participants to comment freely on the merits of mixed attendance. The average rating for the overall experience in virtually attending the workshop was 3.6. The virtual participants gave comments that the experience is “encouraging” and “exceeds expectations.” Next, we have highlighted in bold the main findings, starting with positive feedback followed by negative ones.

Excellent readability The virtual participants indicated that they were able to read the presentations in the NVE clearly. Participants commented that the ability to read slides clearly is an advantage of virtual attendance, since zooming is not possible in the real world. One participant even indicated the ability to zoom makes virtual attendance better than physical.

Good comprehension experience Interestingly, no one singled out the ability to listen to the speaker or the audience as a disadvantage compared to physically attending a workshop or watching video presentations remotely, even though there were occasional lost audio due to network congestion. The virtual participants also gave a positive rating for the ability to understand the talks and to follow the ensuing discussions. This shows that the NVE worked fairly well as an alternative to physical attendance, in contrast to reports with other systems where paper presentations (keynotes, specifically) were

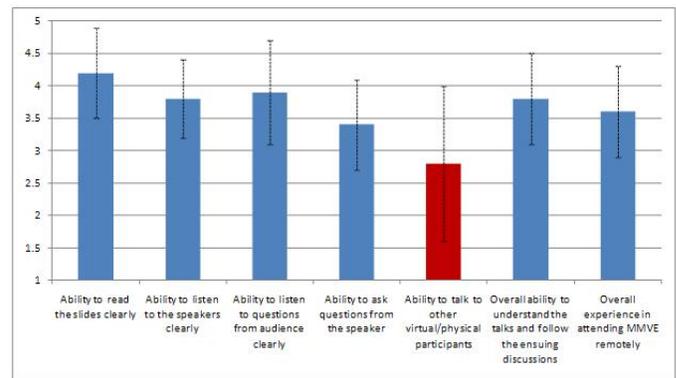


Fig. 5. Survey from nine virtual participants, with 95% confidence interval. Red bar indicates the only component scoring less than acceptable.

problematic and people preferred live video feed [5].

Freedom to multitask Some virtual participants noted that the ability to multitask in a comfortable environment with accessible local resources is a big plus. This is also consistent with the results in [5]. Interestingly, another participant indicated this as a disadvantage, as he was subjected to local disruptions while in his office, and was unable to concentrate on the talks. **More engaged than videos** When asked to compare with viewing offline video presentations later as opposed to virtually attending, many virtual participants pointed out that the ability to ask questions and interact with the speaker is a major advantage. On the other hand, some participants felt that the inability to pause or replay, and the need to participate live during specific times is a drawback, though both can be said about physical attendance as well, and are therefore not a disadvantage of virtual attendance.

Lack of video/animation in presentations One of the drawbacks was the inability to see any video demos by presenters. One participant commented this as the biggest drawback. Although web.alive does support desktop projection, it was not used due to bandwidth constraints. Participants’ feedback indicates that this feature would significantly improve the quality of experience in the NVE.

Avatar shortcomings Another reported drawback was less-than-convincing impressions of the avatars. Most body language and facial expressions of the speaker are not transmitted in the NVE. One virtual participant described his experience with the speaker’s avatar as “creepy”, “feels like a mannequin is speaking.” Another negative comment relates to the avatar of the NVE operator. Since that avatar represents the entire physical world and is linked to a single microphone recording the voices of all physical participants, at least one virtual participant found it confusing that the NVE operator’s avatar speaks with multiple voices, sometimes of different genders. The virtual participants also found the interface for interaction (e.g., press a key to raise hand) as unintuitive.

Low awareness Some virtual participants complained that they don’t know what’s happening in the physical room unless the NVE operator explains it to them. This was important because sometimes a presentation was interrupted due to

technical problems (for example, at some point sunlight was shining onto the whiteboard and curtains had to be pulled down), but the virtual participants did not know the reason for the interruption and wondered why the speaker has stopped.

Post-presentation isolation Some virtual participants were disappointed by the silence in the virtual environment after presentations finished (when the speaker left the microphone) and felt left out of the discussions and socializing that went on in the physical workshop. This is very important because the ability to network socially with others, catch up with colleagues, or hold spontaneous discussions after presentations is one of the main reasons that workshops are held. As we will see later, the physical participants felt the same way about the virtual ones. This sense of disconnectedness between the two environments in the post-presentation mode is an important finding which we will elaborate on later.

Difficulties in asking questions The virtual participants gave lower average rating for their ability to ask questions and talk to other participants, with an average of 3.4 and 2.8 respectively, the latter marked with red in Figure 5 as the only score with less than acceptable average. This is consistent with previous observations that remote participants are often marginalized compared with physical participants [4], [6]. One virtual participant mentioned that for Q&A, the need to un-mute his physical microphone (muted to avoid local echoes) and move towards the virtual microphone was awkward. However, this is not unusual in physical conferences, as attendees who want to ask questions after a keynote must also go to microphones.

Lack of significant social interactions Only a pair of virtual participants socialized with each other before the sessions. This may appear to contradict previous reports where using avatars in meetings has lead to social interactions [7]. However, the result in those reports could be attributed to the fact that the participants already knew each other from real-life, as evident from the personal nature of the jokes exchanged, whereas in our case most people did not know each other. In fact, [5] reports that almost half of participants (47.6%) do not get into a conversation with strangers, and that 60% do not meet new people. It also reports that people mostly do not hang out between sessions. Considering that [5] provided rooms and time slots in their program specifically for socializing, as well as the 3-day duration (compared to our 1 day) which increases the chance of people eventually talking to each other, our results are more or less consistent with previous reports.

B. Physical Participants Feedback

Based on our observations, the overall reactions of the physical participants were quite positive. It was apparent that people saw something new and interesting when the second screen showed the remote participants, and many immediately started to take pictures (as seen in Figure 1 center). The physical participants' evaluations are summarized in Figure 6. It can be seen that the ability to hear and speak to the virtual attendees was quite high, likely due to the good speaker system allowing the virtual attendee's voice to be heard as

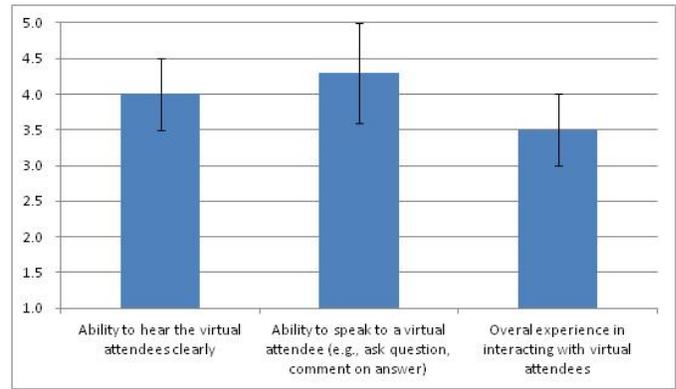


Fig. 6. Survey from 10 physical participants, with 95% confidence interval.

if the person was in the same room. However, in terms of overall experience, it was only a little better than acceptable. This is surprising, considering that the physical participants had no problem seeing, hearing, or talking to the virtual participants. Additional comments revealed that this lower rating was attributed to limited social interactions with the virtual attendees beyond the presentation; i.e., while interaction during the presentation was fine for both participants, the post-presentation “social” interactions were little to non-existent between the two environments. Some also commented that while remote participation certainly would benefit attendees unable to participate physically, they do not feel that it should completely replace a physical gathering (similar to the findings in [5]), since many important social interactions are simply unavailable in a virtual setting.

C. Verdict

Based on the above, we can issue a summarized verdict as one positive and one negative high level finding.

The main positive conclusion is that all participants agreed that presentations and the ensuing Q&A sessions, whether given by a virtual or physical speaker, were quite practical barring minor difficulties. This means that for presenters or listeners unable to attend physically, virtual attendance is a serious and viable alternative. We believe that this will have important impacts on workshop organization in two aspects. First, having virtual presenters will enable workshop organizers to invite busy speakers or reduce “no shows”, which increases the workshop’s quality. Second, for presenters or listeners, attending virtually will save travel time and money, which may increase both the motivation and the number of participants.

However, the findings also suggest that NVEs cannot yet satisfactorily connect the physical and virtual spaces in workshops when it comes to social aspects and presence. Both physical and virtual attendees indicated a clear disconnect between the two environments after a presentation was finished. Also, the virtual participants could not see the physical workshop room and did not know what is currently going on in the room, unless the NVE operator would explain it to them.

This reduced their sense of presence significantly. In the next section, we discuss some possible solutions to these problems.

IV. CONCLUSION AND FUTURE WORK

Despite the limited user size, our study has yielded valuable insights to mixed participation in a workshop. Given NVEs' nature to break away from barriers of time and space, applying them to traditional areas of human interactions (in our case, academic venue) yields new opportunities and issues. Mixed participation allows users to attend a venue that they would otherwise not be able to, e.g. due to financial or time constraints. While it cannot (yet) fully replace physical conferences, as a supplement it has many promises and seems to work better than a purely-virtual approach. Clearly, it also opens up new issues for conference organization, e.g., financial issues like how to cover the additional equipment costs, or legal issues like how to comply to privacy laws or general misuse of the system for illegal activities. These non-technical issues must be subject to serious thoughts before virtual participation can become a standard component of conferences.

For organizers interested to try a similar setup as ours, we highly recommend running a live tutorial with the actual system well in advance of the event, and make it mandatory for all virtual participants to attend. This will greatly reduce the number of small but important problems some participants might face during the actual event (consistent with the experiences in [8]). Also, we recommend that the number of virtual participants be initially limited (especially the virtual presenters), for easier management in the NVE.

For future work, remedying the experienced post-presentation disconnectedness between physical and virtual attendees is a big challenge that requires further research. In fact, "to make interactions with remote people and environments nearly the same as interactions with local people and environments" is one of three grand challenges identified by the ACM SIGMM community for the next decades [11], which involves computer vision, human-computer-interaction, multimedia communications, tele-presence hardware and software technology, psychology, and social studies. We confirm that this is indeed a grand challenge that will take many years of research to overcome.

In future instantiations of our workshop we are planning to test some approaches to increase the level of interactions between users and to immerse them more into the mixed venue. For instance, we think about projecting onto a virtual whiteboard a live video feed of the physical room showing the physical attendees and the room itself. This bird eye's view of the venue would give the virtual attendees a better understanding of what is happening in the room. Also, physical participants who carry a laptop or handheld computing device could log in to the virtual room and be represented by their own avatar, as opposed to the single avatar that represents the entire physical audience. In case an indoor localization system is available, their avatars could even move around naturally. It

would be quite interesting to study the socialization and interaction of such "dual citizens" who are concurrently present in both the physical and virtual worlds.

Another related possibility is to equip NVEs with the abilities to capture, transmit, and render gestures and facial expressions. These features could close the gap between virtual and physical participants significantly, since more natural interactions would be possible.

It would also be interesting to scale up our experiment to a larger number of virtual participants (in the order of dozens or hundreds), or to a conference-scale gathering. We believe the user behavior and experience in such a scenario would be somewhat different.

Finally, another under-investigated aspect is the open question of whether modeling a virtual environment after real-world metaphors is the right approach. For example, [8] reports a meeting held inside a virtual apple tree was more successful than a virtual meeting room. In our case, the presentation aspect was strong so a virtual meeting room made sense. For more interactive sessions such as poster presentations, however, perhaps the "virtual" aspects could enhance and change the interactions in unexpected ways.

V. ACKNOWLEDGEMENTS

The authors would like to thank the organizers of NetGames 2010 for logistics support and setup of the remote participation experiment, as well as Hesam Rahimi for operating the NVE. We also acknowledge the kind support of Nuket Nowlan from 3D Virtual Crafting and Brent Hadden from Avaya, and we thank all the participants of NetGames and MMVE workshops for their participation and feedback.

REFERENCES

- [1] S. Singhal and M. Zyda, *Networked Virtual Environments*. ACM Press, 1999.
- [2] C. Greenhalgh and S. Benford, "Massive: A collaborative virtual environment for teleconferencing," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 2, no. 3, 1995.
- [3] G. Goebels, V. Lalioti, and M. Gobel, "Design and evaluation of team work in distributed collaborative virtual environments," in *Proc. VRST 2003*, 2003.
- [4] A. Tang, C. Neustaedter, and S. Greenberg, "Videarms: Embodiments for mixed presence groupware," in *Proc. HCI 2006*, 2006.
- [5] T. Erickson, N. S. Shami, W. A. Kellogg, and D. W. Levine, "Synchronous interaction among hundreds: An evaluation of a conference in an avatar-based virtual environment," in *Proc. SIGCHI conference on Human Factors in Computing Systems (CHI)*, 2011.
- [6] D. Harry, J. Green, and J. Donath, "backchan.nl: Integrating backchannels in physical space," in *Proc. SIGCHI conference on Human Factors in Computing Systems (CHI 2009)*, 2009.
- [7] N. S. Shami, L.-T. Cheng, S. Rohall, A. Sempere, and J. Patterson, "Enhancing distributed corporate meetings with lightweight avatars," in *Proc. SIGCHI conference on Human Factors in Computing Systems (CHI 2010)*, 2010.
- [8] K. Bessiere, J. B. Ellis, and W. A. Kellogg, "Acquiring a professional 'second life': problems and prospects for the use of virtual worlds in business," in *Proc. CHI EA'09*, 2009.
- [9] "Avaya's web.alive," <http://avayalive.com/>, 2010.
- [10] S. Rohall, L.-T. Cheng, and J. B. Patterson, "Lessons for deploying virtual worlds within the enterprise," in *Proc. Workshop on Virtual Worlds, Collaboration, and Workplace Productivity, CSCW 2008*, 2008.
- [11] L. Rowe and R. Jain, "Future directions in multimedia research," *ACM Trans. on Multimedia Computing, Communications, and Applications (TOMCCAP)*, vol. 1, no. 1, pp. 3–13, 2005.