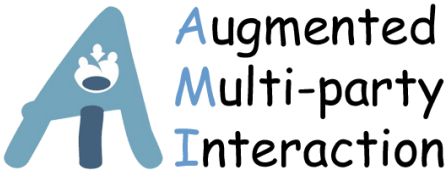

FP6-506811

AMI

Augmented Multiparty Interaction

Integrated Project
Information Society Technologies

D2.3 Infrastructure for remote meetings

Due date: 31/12/2005 **Submission date:** 13/01/2006
Project start date: 1/1/2004 **Duration:** 36 months
Revision: 1

IDIAP

Project co-funded by the European Commission in the 6th Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

D2.3 Infrastructure for remote meetings

Editor : Mike Lincoln (UEDIN)

Abstract: This report gives a description of the hardware and software infrastructure installed in the AMI instrumented meeting rooms to facilitate remote meetings.

1 Overview

To date, the AMI project has concentrated on the analysis of pre-recorded meetings conducted by co-located participants. This analysis has lead to the development of a number of off-line algorithms which extract information from the meeting recordings (such as speech transcriptions, participant focus of attention, participant location etc.) which can be used to allow the rapid browsing and querying of the database of meetings. While work on such algorithms continues, increased effort within AMI is being directed toward the development of realtime systems which can provide additional information to meeting participants. In particular, a 'remote meeting assistant' which will provide additional information to remote participants (for instance alerting them when a particular topic is being discussed within a meeting) is being developed. In order to allow remote participation within meetings, additional infrastructure has been added to the IDIAP and UEDIN instrumented meeting rooms. This document describes the systems used to conduct remote meetings within the instrumented meeting rooms (IMR) and details the additional infrastructure now installed. It also provides an example of how a real-time version of a system used in the off-line analysis of the meeting (a microphone array beamformer) has been incorporated into a remote meeting scenario.

2 Spiderphone audio conferencing

The Spiderphone audio conferencing system (www.spiderphone.com) is used by AMI partners to hold telephone conferences. Users connect to calls using a traditional telephone, however, the system provides a web based interface which provides a number of advanced features compared to a traditional telephone conference system. The web interface allows users to see who is connected to the call and highlights the participant who is currently talking. Users may give PowerPoint presentations to those in the call and share files and applications with other participants. Importantly, the system can also be set to record the conference allowing meetings to be reviewed at a later date. To aid the use of Spiderphone within the IMR, a Polycom Soundstation II audio conference phone has been installed in the UEDIN meeting room.

3 Visual Nexus video collaboration

3.1 Overview

Video conferenceing capabilities have been added to the meeting rooms using Visual Nexus (http://www.visualnexus.com/en). Visual Nexus is a suite of software products running on standard PC hardware which provide a secure online meeting environment in which users may communicate and collaborate. As well as providing video and audio communication between users, the system also allows the sharing of data and applications such that groups may work concurrently on the same documents. The Visual Nexus system has the advantage over other video conferencing systems of being fully standards compliant, making the development of additional components relatively straightforward. Visual Nexus joined the AMI project as a partner in January 2006 and are planing a number of developments to the system based on the requirements of the project, including a system to allow the recording of a conference, and improvements to the application sharing system.

3.2 Software

The Visual Nexus system consists of 3 components :

- *The Meeting server:* The meeting server controls the connections between video conference clients, allowing users to connect to pre-arranged meetings in virtual 'meeting rooms' or to request ad-hoc

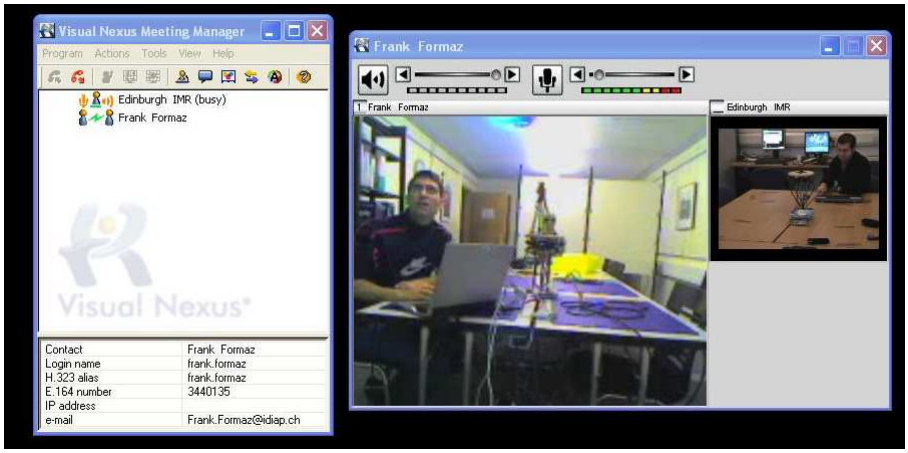

Figure 1: Visual nexus endpoint software showing the meeting room views

meetings with online individuals. Voice data from each participant is mixed into a single channel to provide realistic meeting audio and upto 8 video channels can be provided to each client or 'endpoint'. The server supports connections from any H.323 compliant endpoint for audio and video and uses the T.120 standard for data and application sharing. At present the AMI project is using a demonstration meeting server run by Visual Nexus.

- *The Meeting Endpoint:* The meeting endpoint is a windows application which allows users to join multiparty meetings. The endpoint is H.323 compliant for audio and video transfer, and allows the simultaneous viewing of upto 8 participants. Audio and video codecs are based on the eConf system from France Telecom and provide high quality audio and video even over low bandwidth connections eg. 192 kb/s over domestic ADSL. Data and application sharing is provided via the T.120 standard. Using the application sharing facility, meeting participants may request control of an application running on one of the participants machines and edit and input data within that application. The endpoint also provides a shared whiteboard, and an instant messenger system allowing text messages to be sent to one, or a group of participants. Figure 1 shows a screenshot of the endpoint in use in the meeting rooms.

- *The secure transport layer* This module is required at each site to provide a secure tunnel for the H.323 data through the host site's firewall.

3.3 Hardware

Additional hardware has been added to the meeting rooms to facilitate video conferences. The Edinburgh room has an additional HP-dc7100 3GhZ P4 PC to run the meeting endpoint and secure transport layer. The PC is equipped with wireless mouse and keyboard so that participants seated at the meeting room table some distance from the unit may still operate it. Audio input is provided from one of the omni-directional microphones previously installed in the room as an element in a microphone array, and audio output is provided by a pair of Genelec powered speakers. Video Input is provided by a Sony DCR-HC14E digital video camera mounted at the end of the room beside the screen for the data projector, and the conference is displayed using the data projector already installed in the room. The use of a digital video camera provides significantly better resolution than a traditional web cam. The equipment is shown in figure 2.

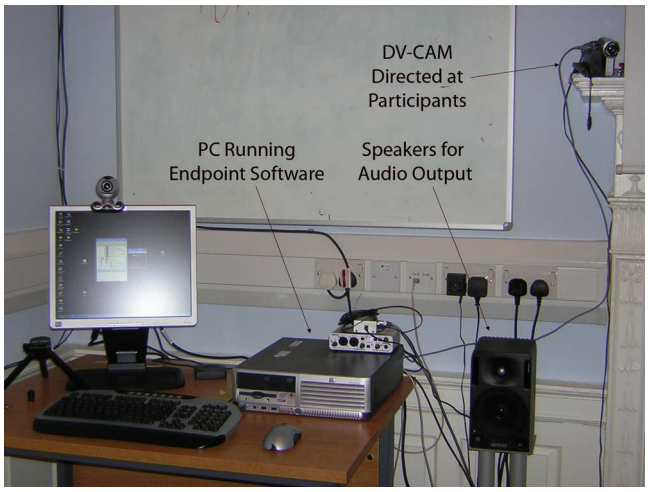

Figure 2: Additional equipment installed in UEDIN instrumented meeting room to facilitate video conferencing.

In the IDIAP room, a laptop running the meeting point and secure transport layer is used. The already installed data projector is used as external display for the conference and an additional Logitech webcam is placed in front of the data projector screen for the video input. The audio output is provided by a pair of powered speakers connected to the laptop. A more permanent solution including a PC with a better audio and video inputs will be implemented in a near future.

4 Use of real-time beamforming for remote meetings

To demonstrate the integration of the original instrumented meeting room equipment with the additional infrastructure, and to show the application of previous AMI research to remote meetings, a real-time beamformer based on that used in the analysis of recorded meetings is under development. The beamformer runs on the audio capture equipment used for the recording of the AMI meetings database detailed in deliverable D2.2. The system is being implemented as a VST plugin on the audio capture PC - VST is an industry standard for developing audio plug-ins for use in VST compatible host applications. Such plug-ins are typically used for real-time effects processing in recording studios and live performances, however, the architecture is also suitable for the development of any realtime audio DSP algorithm, such as the beamformer. The system is based on the steered response power beamformer described in [1]. Beamformed output is generated for each of locations in the room where participants are likely to be positioned (4 seats, the white board and the presentation space). The location with the highest energy is then used as audio input to the video conferenceing system. This should help reduce one of the problems of the current audio setup - that of feedback from the loudspeakers to the omnidirectional microphone. At present feedback is reduced by means of an echo canceling system supplied with the Visual Nexus software. However, due to the proximity of the speakers to the microphone, the system frequently over attenuates participants making them inaudible. The beamformed output is attenuated in directions other than the 'look' direction (including the direction of the loudspeakers) and as such feedback will be reduced and the echo cancelation system should attenuate remote participants less.

References

[1] Lincoln, M., McCowan, I., Vepa, J., Maganti, H.: The multi-channel wall street journal audio visual corpus (mc-wsj-av): Specification and initial experiments. In: Proc. ASRU05. (2005)