Web-system for Mobile Meeting Participant Support

Al.L. Ronzhin, M.V.Prischepa, V.Yu. Budkov
SPIIRAS
39, 14th line, St. Petersburg, Russia
{ronzhinal, prischepa, budkov }@iias.spb.su

Abstract

Web- system for the mobile meeting is a viable alternative to traditional face-to-face meetings. E-meetings are popular in businesses because of their cost savings. To provide quick and effective engagement to the meeting activity, the remote user should be able to perceive whole events in the meeting room and have the same possibilities like participants inside. The technological framework of the developed intelligent meeting room includes many different multimedia devices. The developed web-based application for remote user interaction with equipment of the intelligent meeting room and organization of E-meetings were tested with Nokia mobile phones.

Index Terms: E-meeting, smart space, remote collaboration, meeting support technology, multimodal interfaces.

I. INTRODUCTION

Web-system for teleconference and distant learning (E-meeting and E-lecture) become more popular in business, research, education and other organizations. Such systems allow us to save cost and provide self-paced education and convenient content access and retrieval. However the main part of the secretary work on documentation and connection of remote participants is performed by a human-operator. Another constraint of E-meeting systems is the capacities of the communication network and the capacities of the audio and video user devices that influence on the user interface features and sufficiency of remote participant possibilities.

Key issue of remote communication is high uncertainty, caused by the spatial and temporal distance between co-participants [1]. The physical and psychological barriers that exist in hybrid meetings make difficult for remote partners to attend a selected conversational flow, which the participants share the same room and to initiate a new topic of discussion. Thus the main task of the research is to make remote meetings more engaging by giving remote participants more freedom of control in discussions and decision making processes.

Research projects AMI [2], CHIL [3], AMIGO [4], CALO [5] were targeted to study various aspects of arrangement of meetings or teleconferencing in smart environments and development of meeting support technology, multi-modal meeting browsers, as well as automatic audio-, and video-based summarization systems. Meeting support includes (semi-) automatic retrieval of information needed for support remote participation in hybrid meetings, in which one or more participants are remote and others stay in a shared room [6]. Development of technologies for automatic selection of the best camera view, switching on the projector or whiteboard output and selection of microphone of the current speaker is capable to improve audio-visual support for a remote mobile participant.

The novelty of the proposed web-based system for meeting consists in designing a user interface adaptive to mobile device features of remote participant. Section 2 introduces a model of media stream fusion used for designing actual content of the web-based application. Database specification for meeting translation system are considered in Section 3. The experiment results are discussed in Section 4.

II. MODEL OF COMPILATION OF WEB-PAGE GRAPHICAL INTERFACE FOR MEETING TRANSLATION SYSTEM

The developed intelligent meeting room is equipped by projector, touchscreen TV for smart desk application, several cameras for video monitoring of audience and presentation area, personal web-cameras for analysis of behavior of participants sitting at the conference table, as well as three microphone array for sound processing. A combination of the desktop web-cameras and microphone arrays provides both spatial localization of sound sources and record of participants' speech with high quality. Multimedia content compiled from audio and video signals captured by the referred devices are used in the web-based application for organization of hybrid E-meetings.

Graphical interface of the web-page, on which remote participants could observe a meeting organized inside the intelligent meeting room, contains several basic forms $F = \{F_1, F_2, ...F_{N_F}\}$, where N_F is a number of the forms depending on current meeting state and features of browser used in a client device. Content of the forms could be changed during meeting, but it always includes a graphical component from a set $G = \{G_1, G_2, ...G_{N_G}\}$, where N_G is a number of used components (in the current version $N_G = 10$):

G_1 is a component representing current image on the projector;

G_2 is a component representing current image on the smart desk;

G_3 is a component representing current image captured by the video camera directed to main speaker;

G_4 is a component representing current image captured by the video camera directed to audience;

G_5 is a component representing a assemble of current images captured by personal web-cameras directed on participants sitting at the conference table in the meeting room;

G_6 is a component representing current image captured by a web-camera assigned with a participant, which currently gives a speech comment;

G_7 is a component representing an indicator of speech duration;

G_8 is a component representing a clock with time labels of the current meeting;

G_9 is a component representing a logo of the current meeting;

G_{10} is a component representing main data about the current meeting.

The component G_5, which consists of actual images of participants sitting at the conference table, is compiled by an analysis of states of personal web-cameras and presence of participants and faces on frame. Let's identify an set of images from the web-cameras as $W = \{W_1, W_2, ...W_{N_W}\}$, where N_W is a number of the web-cameras mounted on the conference table (in the developed meeting room $N_W = 10$). Then the component G_5 consists of the images captured by turned-on web-cameras, in which a participant is detected: $G_5 = \bigcup_{i=1}^{N_W} (W_i | c_{2i} \wedge c_{3i} = 1)$. Taking into account limited sizes of the forms used for representing the component G_5, the number of displayed participants is reduced by an analysis of his speech activity c_{4i} and/or presence of his face in the frame c_{5i}. Particularly, the form $F_1^{meeting}(MD)$ for mobile device contains up to three participant images, so both parameters are used for selection of more active participants:

$$F_1^{meeting}(MD) = \bigcup_{i=1}^{N_W} (W_i | c_{2i} \wedge c_{3i} \wedge c_{4i} \wedge c_{5i} = 1) \quad .$$

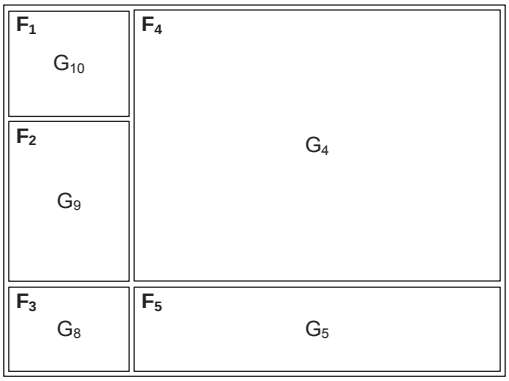
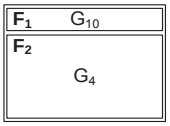
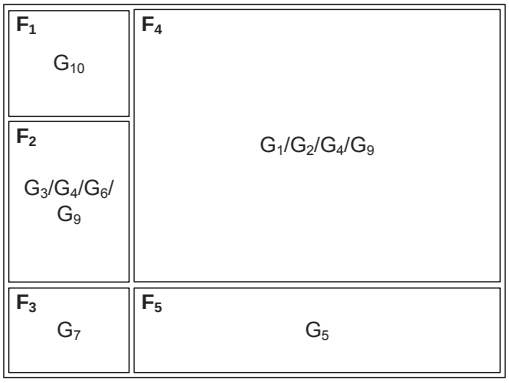
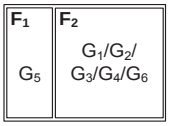
The enumerated components connected with corresponding source, which transmits graphical data (the projector – a presentation slide; the smart desk – window with handwriting sketches; the video and web-cameras – frames with an image; the software services – time indicators, logo

and other data about meeting). Receiving new data on a source leads to updating content of corresponding form in the web-page.

Two main states in meeting process were selected and corresponding notation of forms was used for the current version of web-page software: (1) registration (preparations before meetings), forms F^{reg}; (2) presentations (main part of meeting), $F^{meeting}$. Further the number of the meeting states will be increased taking into account peculiarities of participant behavior and necessity of use of specific technical equipment during the discussion, voting and other formal stages.

Another important factor influencing on the web-page content is the display resolution and correspondingly maximal size of browser window used for remote view of the meeting. So two classes of device, which have especially different sizes of screen, and corresponding notation of the forms were selected: (1) personal computer, forms $F(PC)$, (2) mobile device, forms $F(MD)$. Table 1 shows basic variants of the web-page layout depending on the current state of meeting and type of the client device.

Table 1. The layout variants of the web-page for E-meeting translation.

Meeting state	Screen of client device	
	Personal computer ($F(PC)$)	Mobile device ($F(MD)$)
Registration (F^{reg})		
Presentations ($F^{meeting}$)		

Symbol "/" designates that several variants of graphical content are possible in the form. For instance, the current image G_1 on the projector or the current image G_2 on the smart desk will be represented in fourth form $F_4^{meeting}(PC)$ during presentations on the web-page browsed by a personal computer. Values of parameters of hardware and software are determined by a query of its states via TCP/IP protocol or by means of Object Linking and Embedding Automation. Behavior of participants sitting at the conference table, as well as the main speaker located in the presentation area is analyzed by developed technologies of sound source localization, video tracking of moving objects, seeking and tracking of human face.

Let's consider the process of graphical content compilation in the forms. Each graphical form F_j on the web-page is described by the following tuple $F_j = \langle l_j, u_j, w_j, h_j, g_j \rangle$, where l_j is upper

left corner position of the form at the abscissas axis, u_j is upper left corner position of the form at the ordinates axis, w_j is the form width, h_j is the form height, g_j is a graphical content of the form, which is actual and was chosen from the set G. Sizes of the forms could be changed depending on the current features of browser used in the client device. In the forms $F_2^{meeting}(PC)$, $F_4^{meeting}(PC)$, $F_2^{meeting}(MD)$ the number of graphical component is changed depending on current situation. In other forms the graphical component numbers are kept during the whole meeting. Selection of the current graphical component $g \in G$ for the referred forms is realized by a logical-temporal model of compiling the graphical interface web-page[7].

III. DATABASE SPECIFICATION FOR MEETING TRANSLATION SYSTEM

Data from all sources are coming to the meetings control module and writing in the database table. Table 2 described format of database table.

Table 2. Description of database table fields.

Field	Type	Function
ID	Number	ID of table writing
LayoutType	Number	Current state of event
LayoutChanged	True or False	Shows changing of event's state
FormsInformationChanged	Set of numbers is separated by comma. Serial number of numerals in data is form serial number, number in this set shows data last change.	Shows changing of data in forms
FormsInformationTypes	Set of numbers is separated by semicolumn.	Describe type of data, which are represented in form
FormsInformationData	Set of data is separated by semicolumn. If data presents like array, it is separated by comma.	Store the data, which are represented in form

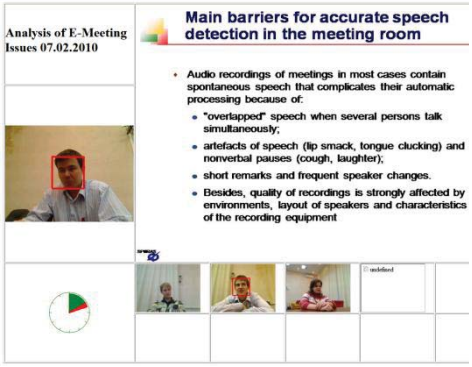
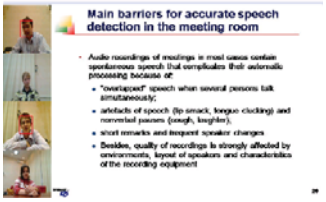
Data for web-pages for mobile device and personal computer are stored in individual tables, because it has different in forms for displaying, which are implemented in these devices. Receiving of information on a web-page is realized like this method: AJAX call a PHP-script, which receives data from database. Received data retries in web-page script like a set of arrays. At first ID of saved data on web-page and ID of new data are compared. If they are matched, processing stops. After that, changing of event's state is defined by analyzing of the LayoutChanged field. In case, if this field is changed the structure of form location reconstruct, if it needs. Compares the number of recent changes to the data in a form with numbers obtained from the field FormsInformationChanged, if they do not match it, depending on the type of data specified in the FormsInformationTypes, and the number of forms, for which they are intended, it handles the data received from the field FormsInformationData, and display them in shape. Data may be presented like: text information (for example, topic of event), array with parameters for a module, images file name (image of current slide). If data sanded like image file name, so in case of data type and form, image changes. For example first name and last name are added for displaying participant's image. Currently audio stream transfer is realized for devices, which support Flash and ActionScript language. All audio steams received from web-cameras based on

conference-table are transmitted to the server of media data transmitting. Channel number with active participant is transmitted to the flash module, which is integrated on web-page, after that flash module connecting to the server and choose required channel.

IV. EXPERIMENTS

Logical-temporal model of compilation of web-page graphical interface was tested on a personal computer and several models of smartphones donated by the Nokia company. Table 3 presents some examples of web-page content during a meeting. The sound source localization system and multichannel speech activity detection system were used for selection of source of audio stream transmitted to remote participant. Speech of a presenter is recorded by the microphone of the web-camera closest to the presentation area. The built-in microphone of the web-camera assigned with the presently active participant sitting at the conference table is used for recording his/her speech.

Table 2. Content examples of the web-based application for meeting.

Meeting state	Screen of client device	
	Personal computer	Mobile device
Registration		
Presentations		

V. CONCLUSION

Model of compilation of web-page graphical interface for meeting translation system was developed. Model describe variant of interface layout for mobile devices and personal computers. The proposed model of compilation of web-page graphical interface allows remote participants to perceive whole events in the meeting room. Model has been tested in the course of lectures and meetings in the smart room of SPIIRAS.

ACKNOWLEDGMENT

This work is supported by The Russian Program “Research and Research-Human Resources for Innovating Russia in 2009-2013” (contract П2360), RFBR (projects 08-08-00128-a, 08-07-90002-CT_a) and Nokia.

REFERENCES

[1] Yankelovich, N., Kaplan, J., Simpson, N., Provino, J.: Porta-person: telepresence for the connected meeting room. In: Proceedings of CHI 2007, pp. 2789–2794 (2007)
[2] McCowan, I., Gatica-Perez, D., Bengio, S., Moore, D., Bourlard, H.: Towards computer understanding of human interactions. In: Aarts E et al. (ed) Ambient intelligence, Lecture Notes in Computer Science, Springer, Heidelberg, pp 235–251 (2003)
[3] Waibel, A., Steusloff, H., Stiefelhagen, R.: CHIL—Computers in the human interaction loop. In: Proc. of 5th IAMIS Workshop (2004)
[4] Danielson, T., Panoke-Babatz, U. et al.: The AMIGO project: Advanced Group Communication Model for Computer-based Communication Environment. In Proc. of CSCW'86 (1986)
[5] Kaiser, E.: Can Modeling Redundancy in Multimodal, Multi-Party Tasks Support Dynamic Learning? CHI '05 Workshop: CHI Virtuality 2005 (2005)
[6] Op den Akker, R., Hofs, D., Hondorp, H., Akker, H., Zwiers J., Nijholt, A.: Supporting Engagement and Floor Control in Hybrid Meetings. Springer, LNAI 5641, pp. 276-290 (2009)
[7] Al.L. Ronzhin, V.Yu. Budkov, A. Karpov. Multichannel System of Audio-Visual Support of Remote Mobile Participant at E-Meeting // Springer-Verlag Berlin Heidelberg, S. Balandin et al. (Eds.): NEW2AN/ruSMART 2010, LNCS 6294, 2010, pp. 62–71.