Reviewing Multimedia Meeting Records: Current Approaches

Simon Tucker Steve Whittaker
University of Sheffield
Regent Court, 211 Portobello St.
Sheffield, S1 4DP
+44 (0) 114 222 2664
{s.tucker,s.whittaker}@sheffield.ac.uk

ABSTRACT

There has been increasing interest in recording and annotating meetings for post-hoc review. Typically, a rich meeting record is formed from multichannel audio and video recordings, along with whiteboard markings, personal notes and presented slides. These records are then supplemented with automatically derived data such as a speech transcript, emotions, and gestures. This media rich meeting representation is then presented to user via a meeting browser. This paper examines the current state of meeting browsers and organises browsers according to the types of data that they typically focus on. Following this current meeting browsers are critiqued and areas for future research are discussed.

1. INTRODUCTION

A number of large scale projects (e.g. [1;16;26]) are investigating the collection and processing of multimodal meeting data. As collection tools and post-processing techniques become more sophisticated the range of data collected and extracted during the course of a meeting becomes both greater and more complex. For example, in the AMI project [1] a typical meeting recording is intended to consist of: 7 video recordings, 22 audio recordings, whiteboard capture, personal notes capture, presented slides capture, along with various derived data including automatically generated transcripts, named entities, dialogue acts, meeting actions, emotion recognition, topic segmentation, individual actions, focus of attention, person location and observations of interest.

Consequently, it is important that the browser efficiently exploits and effectively presents these annotations and data. This entails addressing two issues. Firstly, whilst it is has been shown that data such as the speech transcript and audio recordings will be of use to users (e.g. [17]), it is less clear in which domains such data as emotion recognition or focus of attention will be useful. Secondly, once it has been determined *which* data should be presented there is the further question of *how* this data should be presented.

It is clear, therefore, that the field of meeting browsers will change dramatically as these issues are addressed. As a means of organising the current state of meeting browsers this paper describes a grouping of meeting browsers according to their focus [30] and outlines how current meeting browsers fit into this grouping. Building on this, problems associated with current meeting browsers are described, as well as future directions for research.

2. ORGANISING MEETING BROWSERS

There are several ways of grouping meeting browsers; however when looking at the history of meeting browsers a good means of categorising the current set of browser is by the main *focus* of such browsers. The focus of a browser is defined to be either the main device for navigation within the interface or the primary mode of presenting the meeting data to the user.

Given this definition and the range of data described above four classes of browsers present themselves. Firstly, there are browsers whose focus is mostly audio, including both audio presentation and navigation via audio. Secondly, to complement this there are browsers whose focus is video; again this group includes browsers which use video for both navigation and display. The third set of browsers are focused on meeting artefacts such as whiteboard markings, presented slides and personal notes. The final set of browsers are focused on data derived from the raw meeting recordings. Analysis of meetings has traditionally focused on the content and structure of discussions so this group of browsers is generally concerned with browsing discourse and discourse properties.

These groupings are shown in Table 1. Note that a high-level grouping has been placed on the browsers. The primary focus of perceptual browsers is on audio and visual data and contains the audio and visual browsers. The primary focus of semantic browsers is on textual data, whether it originates from recordings made during the meeting or from post-processing. Given this first-generation grouping each group can now be described in more detail, in the order in which they were presented above.

Table 1: Overview of Taxonomy including typical indexing elements for each class of browser.

PERCEPTUAL	SEMANTIC
Audio	*Artefacts*
• Speaker Turns • User Markings	• Presented Slides • Personal Notes
Video	*Discourse*
• Keyframes • Participant Behaviour	• Transcript • Mode of Discourse

2.1 Audio Browsers

This section discusses browsers whose main focus is *audio*. We separate these browsers into two main subcategories. The first subcategory consists of audio browsers with detailed visual indices; the second category is audio browsers with limited, or no visual feedback.

Both Kimber *et al.* [20] and Hindus and Schmandt [15] describe a meeting browser whose primary means of navigation is via a visual index generated from speaker segmentation. The view presented to the listener is of participant involvement in the meeting - users are able to navigate to each speaker segment and can also navigate between neighbouring speaker segments.

Degen *et al.* [10] describe an indexed audio browser designed for visually reviewing recordings made with a personal tape recorder. The tape recorders allow users to mark salient points whilst recording, the marked recordings then being digitised for review on a computer. The computer interface affords several methods of browsing the recordings. Firstly, users can arbitrarily jump to any part of the recording, and can also navigate using the markings they made during the recording phase. The visual representation of the recording is of amplitude against time, displayed as a vector or colour plot. Users can also zoom in and out of this display and also have the ability to speed up playback (see the discussion surrounding SpeechSkimmer below).

A key element to these browsers is that the visual representations allow users to immediately see the structure of a meeting. This view, however, is dependent on the browsing environment allowing visual representations to be made. There are situations and devices which do not allow for this visual feedback, so that 'pure' audio browsing requires a substantially different interface.

SpeechSkimmer [2] is a system for interactive 'skimming' of recorded speech with no visual representation. Skimming is defined as system controlled playback of samples of original audio. A four level skimming system is implemented, each level compressing the speech further, whilst attempting to retain salient content. The first level is unprocessed playback, the second shortens pauses, whilst the third level plays back only speech which follows significant pauses. The final level uses an emphasis detector to select salient segments of the speech to present to the listener. On top of these skimming levels is a mechanism which allows the playback speed to be altered whilst maintaining the pitch of the speaker. In this way the playback speed can be increased without a significant loss in perception. It should also be noted that the interface allows users to skim backwards in a recording - in this mode short segments of speech are played forwards but in reverse order.

Because of their nature, audio browsers are largely implemented in hardware devices and so can be argued to be distinct from meeting browsers making use of multimodal data. It has been seen, however, that these browsers have overcome some of the limitations of just audio and are able to provide means of browsing audio using computed indices and speed-up techniques. As a complement to this, the following section describes browsers whose primary focus is video.

2.2 Video Browsers

The following class of browsers focus on *video*. Note that whilst each of these browsers have audio and video components, the main component for presentation or navigation in each case is video.

Foote *et al.* [12] describe a simple video browser with two primary modes of navigation. Firstly the user has the ability to jump arbitrarily to any section of the meeting, or to jump between index points which are precalculated from properties of the audio and video. The same indexing, when converted to a continuous 'confidence' measure can also be used to control the playback speed. For example, the playback speed could be related to gesture recognition so that portions of the meeting with significant gestures are played at different speeds, and index marks are made according to these significant gestures.

Girgensohn *et al.* [14] describe video interfaces centred on the use of *keyframes*. Keyframes are static images which have been automatically selected from continuous video according to some heuristic. In the video browsing system, keyframes are chosen according to an importance score, depending on the rarity and duration of each shot. Frames are then sized according to their importance (so that keyframes of higher importance are larger) and are placed linearly on the page. The resulting interface is then similar to a comic book or Japanese Manga drawings. This method can be used to produce a single summary of a full meeting and the user can playback salient portions of the meeting by selecting keyframes, or by choosing a point on a horizontal time line. A similar keyframe-based system was also developed at CMU [6].

A more complex video focused meeting browser is described by Lee *et al.* [23]. A novelty for this system is that it does not require a dedicated meeting room; instead, capture is performed by a single device, encompassing a camera which captures a panoramic video of the meeting and four microphones to record audio. A real-time interface allows meeting participants to examine audio and video during the meeting, as well as making notes during the course of the meeting. The meeting is then archived and processed in preparation for browsing.

The browsing interface has a large number of navigational options. Central to the interface is the video screen, showing both the panorama and a close-up of the currently speaking participant. Users can navigate via a number of indexes, including representations of speaker transitions and visual and audio activity. There is also the opportunity to review an automatically produced transcript of the meeting, and to navigate the meeting via this transcript. A final option for navigating the meeting is a set of automatically generated keyframes. The interface also allows the user to review any notes made during the meeting and to examine any artefacts produced from the meeting.

We note that this class of browser is relatively small, mainly because video is largely supplemented with other browsing devices and is rarely solely used as a means of navigation. Furthermore, often meeting data does not contain the salient visual events that are useful for video browsing. The browsers described above, however, have shown that there is potential for making use of video as a means of browsing meeting data, although its value is not yet determined [6].

2.3 Artefact Browsers

The final browser classification based on data collected during the meeting is that of the *artefact* browser. We use the term artefact to describe any physical item recorded during a meeting which isn't audio or video. Browsers in this class fall into two subclasses: those which focus on presented slides and those which focus on notes taken by meeting participants. An important difference between this class of system and video or audio browsers is that artefacts are usually *searchable*, making it possible to both browse and search data. We will discuss each subclass in turn.

Cutler *et al.* [9] describe a meetings browser in which the central component is captured images from a whiteboard. The interface also contains a participant and whiteboard index, allowing users to jump to particular segments of the meeting, or to review segments which relate to specific elements of the whiteboard annotations. Furthermore, two video components are included - a panorama of all the participants and a close up view of the current speaker. In addition to these components the browser also allows the user to speed up playback, and also to skip the contributions of selected participants.

The browsers described above have focused their attention on presenting *community artefacts* - those which can be altered or viewed by all meeting participants. The following set of browsers in this class examine more private artefacts; specifically, they make use of notes made by participants as a means of indexing and browsing meetings.

Whittaker *et al.* [35] outline the *Filochat* system which combines an audio recorder with a tablet for taking notes as a means of constructing a meeting record. The tablet acted as a virtual notebook and allowed users to store several pages of notes and organise them into sections. Users can then use the notes they have taken to jump to the relevant portion of the conversation. The interface also affords the ability to manually navigate the audio by jumping forwards and backwards. The system was tested and found to be useful both in the field and in lab experiments.

The use of notes to assist recall of meetings was also investigated by Moran *et al.* [27]. The data used for the study was collected from meetings chaired by a single person, in which audio and notes taken both on a shared whiteboard and by the meeting chair were recorded and timestamped. The meeting records were then used by the meeting chair to make technical decisions on the basis of what was said at the meeting. A detailed study of how the chair used the meeting record over a large period of time was made and identified not only how the meeting record was used, but also how the chair's use of the meeting record changed over time. This analysis identified, for example, that the chair would often annotate his notes with the word "ha", meaning that something interesting had occurred and that it would be useful to revisit this section of the meeting during the review process.

The browsers described above examine browsing of artefacts, specifically slides and notes taken by participants. User notes are a powerful means of indexing meetings, since they become both a user-defined index and a means of clarifying any confusing notes taken. Equally, however, a slide index allows users to clarify any confusion originating in presentations. We now discuss the final class of meeting browsers.

2.4 Discourse Browsers

The final class of meeting browsers are focused on *derived elements* of meetings, specifically components such as ASR transcripts or participant contributions. This class of browsers is loosely segregated into those which focus on the transcript, those which focus on participant contributions and those whose focus is a combination of derived and raw data. Because they present ASR transcripts, these systems, like artefact browsers, not only allow browsing but also offer the ability to search.

Bett *et al.* [3] describe a meeting browser in which the transcript is given as much prominence as a video component. The interface also contains a participant index, which indexes single or groups of speakers. In addition to these components the browser also allows the user to construct audio, video or text summaries, using text processing, for complete meetings or salient segments of the meetings. The summary is based on the transcript data and the audio and video streams are segmented accordingly to fit with the reduced transcript. The browser also supports search of a large meetings archive and indexing of discourse features and detected emotions.

The *Ferret* browser [31] also features the transcript alongside video and participant indexes (see Figure 1). A key feature of the browser is that additional temporal annotations can be added or removed at will. For example, it is possible to add automatically derived agenda and interest indices whilst browsing a meeting. The interface is contained in a web browser, and so the transcript can be searched much like a web page, using the browser facilities. As with other browsers, users can navigate through the meeting by clicking on the transcript or by using the derived indices. The index view is customisable and can be viewed at a variety of different zoom levels.

Also included in this class is a browser described by Lalanne *et al.* [22]. Here, the transcript is supplemented with audio and video controls, as well as a view of any documents currently being discussed. Furthermore, the meeting is indexed according to participants and properties of the documents and discourse occurring throughout the meeting. A key element to this interface is that every component is time synchronised, so that any changes or transitions in one component is automatically reflected in all the other components of the interface.

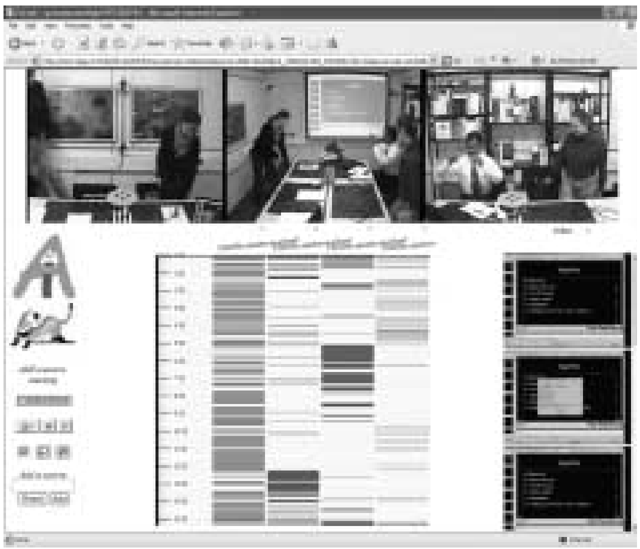

Figure 1. Example of Ferret interface [31].
See http://mmm.idiap.ch

Lisowska [25] (see also [21] for a good review of IM2 meeting browsers) describes an interface to a meeting archive which makes use of a library metaphor for presentation. Each meeting is represented as a book on a shelf, a meeting itself is represented as pages within one of these books. The interaction between the user and the system is done through natural language prompts and, furthermore, the system allows the user to search the full meeting archive in order to locate meetings of interest.

Cremers *et al*. [8] describe an index-centric browser, but one which greatly increases the number of indices available to the user. Further to the indices is a search mechanism which allows users to also search the indices to locate points of interest. The indices include emotions, gestures, items, people, locations, topics, meeting locations. It should be noted that the search is across a number of meetings. The interface affords a video display combined with audio playback for the chosen meeting segment.

Since they make use of both raw and derived data, browsers in this category tend to have a more complex interface than those discussed in the previous classes. At the interface level, however, the index-centric approach taken by these browsers prevents the complexity of the annotations from leading to a complex interface.

3. SUMMARY

Returning to the grouping shown in Table 1, it can be seen that there is increasing complexity both vertically downwards and horizontally to the right. Thus the discourse browsers, because of the larger range and type of data that are focused upon, are generally more complex than browsers in other classes

Despite the range of data and complexity of the browsers examined here it can be argued that the mode of access and intended use is largely similar across all the browsers. The expected use of the browsers is index centric random access: users make use of a number of indices to both navigate the meeting recording and to identify points of interest. There are exceptions to this mode of access: for example, several browsers make use of a search facility.

Given the overview and analysis of meeting browsers presented above, the following sections highlight problems and potential areas for future research in browser development.

3.1 Browser Evaluation

In our previous analysis of meeting browsers we highlighted the lack of evaluation of meeting browsers - generally browsers are developed without any formal evaluation. Any evaluation that does take place focuses on specific aspects of the browser interface rather than examining the overall browser experience. In the interim, substantial effort has been made in designing and testing a browser-independent evaluation (BET) [32]. Briefly the BET comprises two stages. In the first stage a small number of judges watch meeting recordings using a standard media player and make a number of observations of interest. These observations of interest relate to outcomes of the meeting which the meeting participants themselves found interesting. Each observation of interest is stated as a true-false pair of statements. In the second phase a large number of participants are successively given one side of an observation of interest. The participants must then use a designated browser to determine whether they have been given a true or false statement.

The browser is then scored according to the ability of the participant to correctly classify the observation of interest in the minimum amount of time. Since the meeting recordings and observations of interest can be made into an evaluation corpus, browser designers are able to evaluate their browsers independently of each other and to produce comparable scores. Once the materials are available and the BET procedure standardised it can be seen that any new browser can be easily compared to the field. The BET has limits however; for example, it cannot be used to evaluate the utility of a browser in determining the gist of a meeting, nor does it address subjective user preferences

3.2 Text Processing

Whilst there is an increased use of search capabilities in meeting browsers, the index-centric focus tends to overlook the use of other text-processing techniques. To further compound this problem producing a summary of a meeting recording requires a different approach to that used to produce textual summaries which have been developed for written documents. However, it can be seen that some form of summarisation would not only be useful for assessing single meetings but would also be desirable for analysing multiple meeting recordings.

It can also be seen that further text processing would be advantageous for meeting browsers. Several studies [36;17;24] have shown that users of meeting browsers are largely interested in determining personal action items or key decisions, both things which text processing techniques can be used to identify and both things which are largely lacking in the current generation of meeting browsers.

3.3 Motivations

As briefly mentioned in the introduction, there is also a need to determine the informational needs of users of meeting browsers. There are questions about what information different classes of meeting participants (e.g. attendees, non-attendees, managers etc.) wish to extract from meeting recordings. Furthermore there are questions as to the ways in which these disparate users want to access this data.

There has been some research which address these issues [17;36] and some meeting browsers have been constructed using principles derived from these studies [25;8] These studies, however, do not examine the informational needs with regards to the more complex annotations (e.g. emotion, gesture etc.) that are being extracted from meeting recordings. Future browsers should build on the analysis work being undertaken and also investigate how best to make use of the more complex annotations.

3.4 Multi-Meeting Analysis

It is clear that meetings are not isolated events. Ideas and decisions can occur over a series of meetings and thus future meeting browsers should attempt to move away from the current perspective of just examining single meetings. The current approach to leveraging information across several meetings is to focus on using search procedures to identify a relevant meeting from a set of recordings and then using some form of viewing mechanism to review the meeting.

It can be argued that a more integrated view of multiple meetings will be advantageous for users of meeting browsers. This is a difficult area to examine since, whilst already having experience in reviewing records for single meetings (via minutes or personal notes), meeting participants currently have no means of easily reviewing a series of meetings. However it can be argued that, because of the flow of information across several meetings, a multi-meeting analysis tool will be beneficial to meeting participants.

4. CONCLUSION

This paper has revisited the state of the art of meeting browsers. A preliminary grouping of meeting browsers was outline, in which meeting browsers are categorised according to their focus. Four groups of browsers were described - those that focus on audio, video, artefacts and derived or discourse features. Following a description of each of these four groups, four areas for future research were described. It is clear that, as more complex annotations and derived data forms become more mature, so the complexity and maturity of the meeting browsers making use of this data will increase.

5. REFERENCES

[1] AMI Project, http://www.amiproject.org/

[2] B. Arons (1997), "SpeechSkimmer: A System for Interactively Skimming Recorded Speech", *ACM Transactions on Computer-Human Interaction*, pp 3-38.

[3] M. Bett, R. Gross, H. Yu, X. Zhu, Y. Pan, J. Yang, and A. Waibel (2000), "Multimodal Meeting Tracker", *Proceedings of Proceedings of RIAO*, Paris, France, April.

[4] J.A. Brotherton, J.R. Bhalodia, and G. Abowd (1998), "Automated Capture, Integration and Visualization of Multiple Media Streams", *Proceedings Of The IEEE International Conference on Multimedia Computing And Systems*, pp 54-63.

[5] P. Chiu, J. Boreczky, A. Girgensohn, and D. Kimber (2001), "LiteMinutes: An Internet-Based System For Multimedia Meeting Minutes", *Proceedings of 10th WWW Conference*, pp 140-149, Hong Kong, 1-5 May.

[6] M.G. Christel, M.A. Smith., C.R. Taylor and D.B. Winkler (1998), "Evolving Video Skims Into Useful Multimedia Abstractions", *Proceedings of CHI '98*, 18-23 April 1998.

[7] S. Colbath, F. Kubala, D. Liu and A. Srivastava (2000), "Spoken Documents: Creating Searchable Archives From Continuous Audio", *Proceedings of the 33rd Hawaii International Conference on System Sciences*, Hawaii.

[8] A. Cremers and B. Hilhorst (2005), “What was discussed by whom, how, when and where?” Personalized browsing of annotated multimedia meeting recordings, In: *Proceedings of HCII 2005*, Las Vegas, Nevada, USA.

[9] R. Cutler, Y. Rui, A. Gupta, J.J. Cadiz, I. Tashev, L. He, A. Colburn, Z. Zhang, Z. Liu, and S. Silverberg (2002), "Distributed Meetings: A Meeting Capture And Broadcasting System", *Proceedings of 10th ACM International Conference on Multimedia*, pp 503-512, Juan-les-Pins, France, 1-6 December.

[10] L. Degen, R. Mander, and G. Salomon (1992), "Working With Audio: Integrating Personal Tape Recorders And Desktop Computers", *Proceedings of CHI '92*, pp 413-418, Monterey, CA, USA, 3-7th May.

[11] M. Flynn and P. Wellner, (2003), “In Search of a Good BET: A proposal for a Browser Evaluation Test”, *IDIAP-COM 03-11*.

[12] J. Foote, G. Boreczky, and L. Wilcox (1998), "An Intelligent Media Browser Using Automatic Multimodal Analysis", *Proceedings of ACM Multimedia*, pp 375-380, Bristol, UK, 14th September.

[13] W. Geyer, H. Richter, L. Fuchs, T. Frauenhofer, S. Daijavad, and S. Poltrock (2001), "A Team Collaboration Space Supporting Capture And Access Of Virtual Meetings", *Proceedings of 2001 International ACM SIGGROUP Conference On Supporting Group Work*, pp 188-196, Boulder, Colorado, 30-3 September-October.

[14] A. Girgensohn, J. Borczky, and L. Wilcox (2001), "Keyframe-based User Interfaces For Digital Video", *IEEE Computer*, pp 61-67.

[15] D. Hindus and C. Schmandt (1992), "Ubiquitous Audio: Capturing Spontaneous Collaboration", *Proceedings of 1992 ACM Conference on Computer-Supported Cooperative Work*, pp 210-217, Toronto, Ontario, Canada, 1-4 November.

[16] IM2 Project, http://www.im2.ch/

[17] A. Jaimes, K. Omura, T. Nagamine, and K. Hirata (2004), Memory cues for meeting video retrieval. In: *Proceedings of the CARPE '04*, pp. 74-85, 15 October 2004, New York, USA.

[18] R. Kazman and J. Kominek (1997), "Supporting the Retrieval Process In Multimedia Information Systems", *Proceedings of the 30th Annual Hawaii International Conference On System Sciences*, pp 229-238, Hawaii.

[19] R. Kazman, R. Al-Halimi, W. Hunt, and M. Mantei (1996), "Four Paradigms for Indexing Video Conferences", *IEEE Multimedia*, pp 63-73.

[20] D.G. Kimber, L.D. Wilcox, F.R. Chen, and T. Moran (1995), "Speaker Segmentation For Browsing Recorded Audio", *Proceedings of CHI '95*, May 7-11.

[21] D. Lalanne, A. Lisowska, E. Bruno, M. Flynn, M. Georgescul, M. Guillemot, B. Janvier, S. Marchand-Maillet, M. Melichar, N. Moenne-Loccoz, A. Popescu-Belis, M. Rajman, M. Rigamonti, D. von Rotz, P. Wellner (2005), The IM2 Multimodal Meeting Browser Family, IM2 Technical Report.

[22] D. Lalanne, S. Sire, R. Ingold, A. Behera, D. Mekhaldi, and D. Rotz (2003), "A Research Agenda For Assessing The Utility Of Document Annotations In Multimedia Databases Of Meeting Recordings", *Proceedings of 3rd International Workshop on Multimedia Data And Document Engineering*, Berlin, Germany, September 8th.

[23] D. Lee, B. Erol, J. Graham, J.J. Hull, and N. Murata (2002), "Portable Meeting Recorder", *Proceedings of ACM Multimedia*, pp 493-502.

[24] A. Lisowska (2003), Multimodal interface design for the multimodal meeting domain: Preliminary indications from a query analysis study, *Technical Report IM2*.MDM Report 11, IM2.

[25] A. Lisowska, M. Rajman, and T. Bui (2004), Archivus: A system for accessing the content of recorded multimodal meetings. In: Bengio, S., Bourlard, H. (Eds.) *Lecture Notes In Computer Science*, 3361, pp 291-304.

[26] M4 Project, http://www.m4project.org/

[27] T.P. Moran, L. Palen, S. Harrison, P. Chiu, D. Kimber, S. Minneman, W. Melle, and P. Zellweger (1997), ""I'll get that off the audio": A Case study of salvaging multimedia meeting records", *Proceedings of CHI '97*, Atlanta, Georgia, 22-27 March.

[28] S. Mukhopadhyay, and B. Smith (1999), "Passive Capture And Structuring Of Lectures", *Proceedings of Seventh ACM International Conference on Multimedia*, pp 477-487, Orlando, Florida.

[29] D.K. Roy and C. Schmandt (1996), "NewsComm: A Hand-Held Interface for Interactive Access To Structured Audio", *Proceedings of CHI '96*.

[30] S. Tucker and S. Whittaker (2005), Accessing Multimodal Meeting Data: Systems, Problems and Possibilities, In: Bengio, S., Bourlard, H. (Eds.) *Lecture Notes In Computer Science*, 3361, 1-11.

[31] P. Wellner, M. Flynn, and M. Guillemot, (2004), Browsing Recorded Meetings With Ferret, In: Bengio, S., Bourlard, H. (Eds.) *Lecture Notes In Computer Science*, 3361, 12-21.

[32] P. Wellner, M. Flynn, S. Tucker and S. Whittaker (2005), A Meeting Browser Evaluation Test, Presented at the *Conference on Human Factors in Computing Systems (CHI)*, Portand, Oregon, USA, 2nd-7th April.

[33] S. Whittaker, J. Hirschberg, B. Amento, L. Stark, M. Bacchiani, P. Isenhour, L. Stead, G. Zamchick, and A. Rosenberg (2002), "SCANMail: A Voicemail Interface That Makes Speech Browsable, Readable and Searchable", *Proceedings of CHI 2002*, Minneapolis, Minnesota, USA, 20-25th April.

[34] S. Whittaker, J. Hirschberg, J. Choi, D.Hindle, F. Pereira, and A. Singhal (1999) “SCAN: designing and evaluating user interfaces to support retrieval from speech archives”, *Proceedings of SIGIR99 Conference on Research and Development in Information Retrieval*, pp 26-33.

[35] S. Whittaker, P. Hyland, and M. Wiley (1994), "Filochat: Handwritten Notes Provide Access To Recorded Conversations", *Proceedings of Chi '94*, Boston, Massachusetts, USA, 24-28th April.

[36] S. Whittaker, R. Laban and S. Tucker (2005), Analysing Meeting Records: An Ethnographic Study and Technological Implications, Presented at *MLMI 2005*, Edinburgh, UK.

[37] B. Wrede and E. Shriberg (2003), “Spotting “Hot Spots” in Meetings: Human Judgements and Prosodic Cues”, *Proceedings of the European Conference on Speech Communication and Technology*, pp 2805-2808, Geneva, Switzerland.