FaericWorld: Browsing Multimedia Events Through Static Documents and Links

Maurizio Rigamonti, Denis Lalanne, and Rolf Ingold

DIVA Group, Department of Informatics of University of Fribourg
CH-1700 Fribourg, Switzerland
`first name.last name@unifr.ch`

Abstract. This paper describes a novel browsing paradigm, taking benefit of the various types of links (e.g. thematic, temporal, references, etc.) that can be automatically built between multimedia documents. This browsing paradigm can help eliciting multimedia archives' hidden structures or expanding search results to related media. The paper intend to present a novel model for browsing any kind of multimedia archives and further focuses on an archive of meetings recordings, in order to illustrate the advantage of our method to perform cross-meetings and in general cross-documents browsing. First of all, the structure of meeting datasets is presented, describing in particular the media implied, the annotations used for cross-document linking and the major mining techniques integrated in this work. Then, the paper presents at a glance the visual browser we developed that combines searching and browsing by links. Further, the performances of the actual system are discussed, i.e. the automatic indexing and linking processes for the two different meeting corpora, as well as the access and browsing performances. Finally, the paper presents the major unsolved issues and our perspectives for future works.

Keywords: Multimedia browsing, multimedia indexing, multimodal alignments, information visualization, information retrieval, multimedia meetings archives.

1 Introduction

With the new trend in recording events such as meetings and conferences, a huge amount of multimedia data, connected in various ways, is becoming available within large digital libraries. Although those new types of data are rich in content, they often miss high-level abstractions to allow robust indexing and retrieval. For this reason, multimedia data is often hard to retrieve using a standard google-like interface. In the last few years, different works have tackled this issue, aiming at creating new techniques for indexing meeting recordings and for browsing in annotated corpora [18].

Our main idea is based on two observations: 1) static documents such as newspapers and scientific publications are easily indexed using their textual content and 2) multimedia data can be linked to static documents through document alignment methods [14,18]. Thus, it becomes possible to access multimedia content by first searching textual documents and then browsing into the archive through links. The challenge is then to combine

both searching and browsing process in interactive visualizations that does not separate the two processes. Furthermore, still using this link paradigm, this article explains how our system can help replaying a meeting and observe over time its relationships with other documents or meetings, highlighting recurrent topics and themes. In other words, our novel browsing paradigm tackles the issue recently presented by Tucker and Whittaker: *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* [26].

Firstly, this paper presents the context of our work. Then, the section 3 describes the concepts on which is based our engine for indexing corpora of meetings and for representing relationships between documents through links. In section 4, we describe our browser and, in particular, the three steps for accessing multimedia information: searching, browsing and playing. The section 5 is dedicated to our system performances and gives detailed information on the corpora characteristics and links. Finally, the conclusion wraps up the paper and presents perspective works.

2   Indexing and Browsing in Multimedia Archive

Although several works and researches have been proposed in order to index and browse in multimedia archive and existing search engine are rather powerful, retrieving multimedia information is still difficult and related user interfaces are still in early phase of development [21]. Our assumption is that it is necessary to tackle two main challenges for improving search engines: 1) indexing of multimedia documents must take into account the implicit and explicit relationships between different media and 2) browsing techniques need to evolve thanks to these relationships. The following subsections present related works for respectively indexing and further browsing in multimedia archives.

2.1   Indexing Multimedia Documents

Nowadays, search engine are performing well for indexing and retrieving textual content, but their capability for managing multimedia data is still primitive. This lack derives from the difficulty of annotating and indexing media that are poor in semantic content, i.e. images, videos and audio files. Currently, we distinguish four methodologies for creating useful annotations for indexing multimedia documents: manual, automatic, both manual and automatic and, finally, alignment-based.

Browsers like YouTube [28] use manual annotations for indexing videos. Producing such high-level annotations presents two big drawbacks: 1) users personal perception of the document does not guarantee that the labels are entirely satisfying and 2) professional labeling is too expensive.

Various works tried to overcome this inconvenience and tried to integrate automatic methods for indexing images, audio and video streams, but habitually they do not take into account multiple media. For instance, Swain recapitulates different techniques for indexing media in [23]. Moreover, indexing of multimedia data is often based on low-level features that lack of semantic information. Smith et al. proposed to use model vectors for indexing multimedia documents [22], where each vector is correlated with a semantic concept detector. This method produces interesting results, but it is restraint to a limited set of concepts.

Other works combine analysis techniques with manual annotations, in order to re-
duce and simplify interventions of users. M4Note is a multimodal system for video
recording that automatically computes low-level features and that allows users to eas-
ily add manual annotations thanks to electronic ink and speech recognition [11]. In
the same way, Campanella et al. described a system that visualizes features automati-
cally extracted from videos, in order to support users in defining clusters of shots [5].

A recent trend consists in aligning different media, with the purpose of indexing
multimedia data. Alignment techniques enable the transfer of semantic information
from richer to poorer medium. MUMIS [10] indexes videos of football matches align-
ing different textual and multilingual sources, such as news, commentaries, etc. The
results of different documents analysis' techniques are combined in respect of specific
rules, in order to improve the indexing of all these sources. Behera extracts from vid-
eos the slideshows presented in a meeting and aligns them with the original docu-
ments [2]. His method augments both media: documents are augmented with temporal
dimension and videos gain semantic indexes. Likewise, Mekhaldi and Lalanne pro-
posed an alignment technique based on similarity, which links the textual documents
presented in a meeting with the transcript of spoken dialogs [18].

In our work, we integrate the last technique in order to align corpora of meetings
and to create the new annotations necessary for indexing multimedia documents poor
in semantic information.

2.2   Browsing Multimedia Archives

Today, in Google-like search engines a lot of work is dedicated to the improvement of
indexing techniques. In fact, these engines do not completely explore the browsing
mechanism, because users generally retrieve the interesting information already after
the submission of the first correct query. At opposite, browsing means exploring an
archive without having an exact knowledge of the documents it contains. Conse-
quently, relatively recent projects tried to create relationships between documents, in
order to structure information and to overcome searching lacks. Citeseer [3] and
Scholar [20] interconnect scientific publications with different methods: users are able
to read one paper and then to access similar articles. Moreover, the linking mecha-
nism allows proposing the most linked paper in a specific domain as entry point.
Similarly, LinkedIn [16] creates a network of interconnected persons, where links are
social relations. Kartoo [9] calculates thematic clusters with the documents resulting
from a query, whereas Alice in Wonderland [1] uses links for representing the rela-
tionships among words in a book. However, none of these systems have explored the
use of relationships between multimedia documents, which is the most important
aspect of our system.

Browsing mechanism has also been explored in recent multimedia browsers, where
users explore one or more meetings [15, 25]. The JFerret-based family browsers [12,
27] have been developed in order to navigate between the synchronized documents of
one meeting, using in particular transcript and audio streams. Archivus [17] is another
browser that proposes a searching mechanism under constraint for exploring a
collection of meetings. Finally, FriDoc [15] allows to navigate in one synchronized
meeting using static document as main artifact. However, all these systems considers
browsing at the intra-meeting level and thus do not consider browsing of the whole

meeting corpus, which is critical when meetings are linked in time, thematically or by people, places, etc. The system we present in this article proposes to tackle this issue.

3 Cross-Linking Multimedia Data

FaericWorld is the system presented in this paper, which deals with corpora of multimedia data. In particular, cross-linked multimedia data is called a *world* and consists of meetings recordings, which are composed of heterogeneous categories of documents. Each document in the archive is a triple of raw data, annotations and links. Raw data are media sources (videos and audio files, PDF documents, etc.), multimedia documents (websites and slideshows) and persons (personal information, photos, etc.). The raw data is enriched with annotations, either manually added by users or automatically calculated thanks to analysis techniques. The use of annotations for indexing documents allows to access data in a structured manner and, consequently, to create more precise indexes [4]. For instance, two words are thematically more significant when belonging to one article's title in a PDF file rather than to an entire document. Moreover, annotations allow augmenting media that are poor in textual content, such as videos and audio files. Finally, links elicit the similarity in terms of content, properties or time of one document with the rest of the documents in the world.

The particularity of this implicitly linked dataset and the novelty of the browsing task prefixed in our work have implied the creation of a new relational engine for indexing and retrieving information. Above all, browsing by links involves that FaericWorld pre-calculates and stores a huge amount of information that describes relationships and similarities between documents.

The following subsections present firstly the different annotations used in our system, and secondly, how to create links between the different media.

3.1 Extracting Annotation from Media

This subsection presents the different annotations that are extracted from the media composing a multimedia meeting corpus and their properties.

~ *Static documents* are analyzed thanks to XED tool [19], in order to automatically extract their physical structures (the layout with textual blocs and line). Moreover, some documents have been manually annotated with logical structures, which describe the hierarchy and the logical functions of physical structures. Both annotations allow to access fragments [4] of the static document, which become the main vector of thematic information.

~ *Audio* recordings are transcribed either manually or with automatic tools. The transcript is composed of utterances, which are characterized by start and end timestamps, the full name or the id of the speaker, and the textual content. Actually, annotations on audio are rich of temporal and thematic information.

~ *Videos* are poorly annotated with only information about the related audio track and a label containing speaker name or id.

~ **Persons** are a special kind of objects, entirely defined from its annotations, which are personal and thematic information such as first and last name, email and telephone number.

~ **Meetings** are annotated with a descriptor containing the location, the day and the hour of the recording. This information automatically provided is then manually completed with meeting name, content, type (i.e. belonging or not to a scenario) and participants. Last but not least, the descriptors contain a list of all media presented or recorded during the meeting.

~ **Multimedia documents** are enriched with all the annotations that could be extracted from individual media.

3.2   Creating Links Between Documents

Annotating documents is a preprocessing step that not only interests indexing, but also the creation of links: in fact, each type of annotation is useful for creating specific categories of links. Table 1 summarizes the links considered by FaericWorld and the type of document that produces these relationships.

Table 1. Each category of document is generating different types of link

		cross-document links			
	type	thematic		strict (un-weighted)	
		(weighted)	temporal	Reference	hyperlink
Document		Static document, audio, person, event, multimedia	Audio, video, meeting, multimedia	Static document, multimedia	Static document, multimedia

Thematic links are calculated using alignment techniques [18]: each document of the corpus is compared with other ones in the world, accessing the structured content through annotations. For instance, each physical bloc of a static document is aligned with each speech transcript's utterance computed from an audio recording of the meeting dialogs. This technique allows calculating a similarity score between the content of both documents. Higher is the number of similar and discriminating words in the text, higher will be the score.

Temporal links are generated for time-based documents. For instance, a meeting contains several documents discussed, viewed or created at the same time by participants [13]: the whole documents coexist at the same moment within a particular meeting and, consequently, possess temporal relationships. In this manner, non time-based documents such as static documents and persons acquire a temporal dimension thanks to previous thematic links. Moreover, temporal links can be combined with thematic links to achieve a finer granularity: for instance, when a paragraph in a static document is thematically linked with an utterance, the first one is also synchronized with the timestamp of the second.

References are implicit links to other documents (bibliographies, citations, etc.), whereas *hyperlinks* are explicit links, extracted from textual document and web pages by syntactic analysis.

FaericWorld creates all the links presented in this subsection automatically, at the time the documents are imported in the world. At the importation stage, a new document is aligned with the whole world and the resulting links are computed and stored. Several links of the same type can exist between two documents, because the alignment process takes into account documents' parts.

4   A Combined Searching, Browsing and Playing

The navigation in FaericWorld corpora takes full benefit of the links computed between its different documents. In fact, the task of retrieving information does not require users to have any exact knowledge neither about searched documents nor about indexes allowing for accessing the data. Instead, links allows grouping documents by themes and after the selection of a specific topic allow consulting all the related documents. For instance, Bruno wants to replay what Florian said about the evaluation part of his article during a precedent meeting. So, he first retrieves the article using the keywords of its title. He then opens the document and clicks on the evaluation part, in order to access to the related parts of meetings, further selects the ones in which Florian is speaking and finally plays the corresponding audio/video streams to hear his opinion.
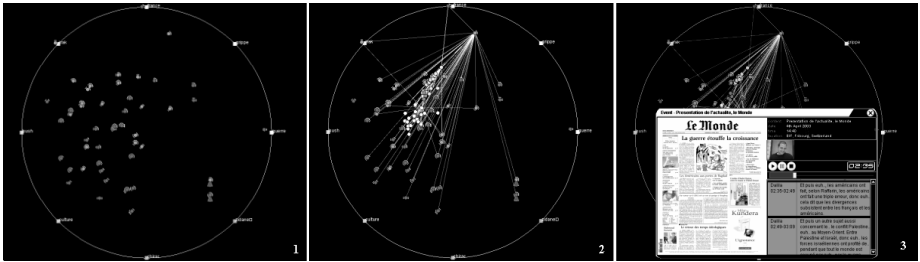

Fig. 1. (1) Searching documents of interest, (2) browsing results through connected documents and (3) playing results are the three main functionalities offered by our system that we propose to combine in a single interactive visualization

This section describes the interactive visualizations proposed by FaericWorld for representing the multimedia corpora, the connection and similarities between meetings and documents, and their relationships through time while playing. Figure 1 illustrates the three main functionalities of FaericWorld, i.e. searching (1), browsing (2) and playing (3), which are discussed in the following sub-sections.

4.1   Searching Through the Archive

The entry point in the archive uses a radial visualization [8], a classical view already used in different works for visualizing large archives of documents (for instance, in

[6]). A query composed of several keywords can be submitted to the system and the set of resulting documents is displayed in the RadViz as illustrated in the center of figure 2. A default query is proposed in order to display from the beginning a preliminary access to the meetings world and to represent an overview of its thematic structure. The default query is composed of the most recurrent terms that belong to disjoint sets of documents.
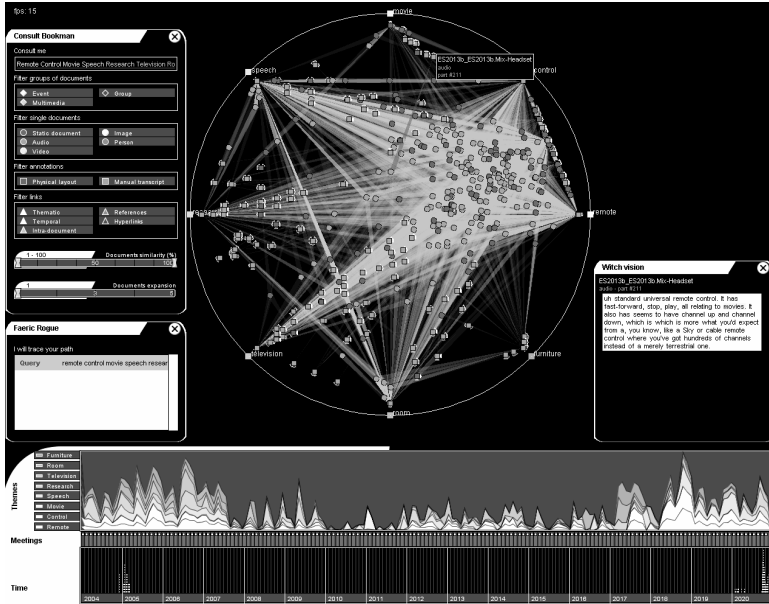

Fig. 2. The screenshot presents an overview of FaericWorld browser, with its two main visualizations and the dialog boxes. The first dialog on the left allows submitting queries, filtering document types, etc. The right dialog is specialized for document previews.

Documents position in the RadViz is defined using their *tf.idf* value (Term Frequency-Inverse Document Frequency, i.e. the ratio between the term frequency in the document over its frequency in the whole archive) for each term of the query. Documents in the center of the RadViz will thus contain all the terms of the query. If the term frequency of a word in a document is very high comparatively to its frequency in the whole corpus, the document will be strongly attracted by the related anchor.

Likewise, the parts of documents (e.g. utterances of speech transcripts or physical blocks of PDF documents) matching the query are displayed in the RadViz and linked to their father document (fig. 3). A mouse over a document of interest for the user, or over a part of document, allows to preview its content and to visualize additional information such as name, number of links, recurrent terms, etc.

A drawback of the radial visualization is the representation of queries with an even number of terms. For instance, sometimes a document is drawn in the visualization's center even if it contains only part of the searched terms, because these terms have the same weight and they are situated exactly at opposite poles on the RadViz. In order to resolve this ambiguity, we used *shapeVis* [24], similar to star coordinates, for repre

Fig. 3. Document parts are visualized as squares in RadViz and linked to the document they belong to, which is represented as a circle. Their colors depend on the document type.

senting the tf.idf value of the document. Figure 4 shows the double interest of this visualization: it represents 1) the terms of the query contained in the document as well as 2) their frequency relatively to the whole corpus of documents (the pie is very accentuated when the frequency of the term in the world is very low).
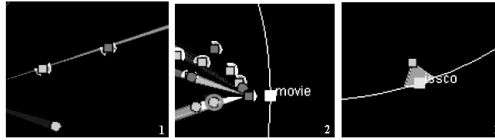

Fig. 4. The whole images are extracted from the same screenshot. (1, 2) The figures show fragments of document with small pies because the terms frequency in the corpus is high. (3) At opposite, the fragment of this image contains the totality of the terms in the archive and, consequently, the pie is accentuated.

Overlapping of documents provokes the second problem in this radial visualization. In some case and especially for atomic queries, several documents could occupy the same location. Currently, FaericWorld does not provide a valid solution to overcome this lack.
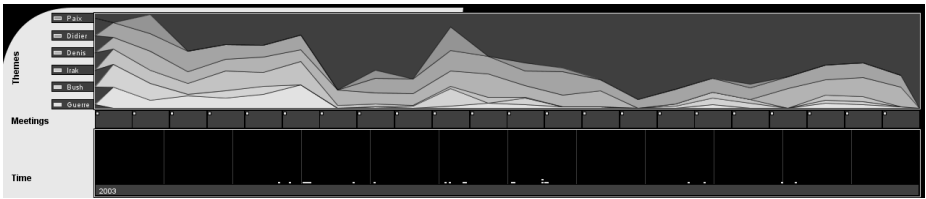

Fig. 5. The second visualization shows themes evolution during time for the whole corpus

The second main visualization in FaericWorld is based on *ThemeRiver* [7] and shows the evolution of themes in the whole archive throughout time (Fig. 5). The bottom part is a representation of the calendar, which indicates when meetings have been recorded. In the middle, the meetings are chronologically sorted one after the other and choosing a meeting highlights its documents in the RadViz (brush and link mechanism). Finally, the top plot represents the evolution of themes defined in the query. Each term is represented in the ordinates axe, whereas meetings occupy the abscissas one. The width of each river depends on the occurrence of each term in the documents linked to the meeting (e.g. speech transcript, attached PDF documents, etc.), weighted by their *tf.idf* values. The meeting having the largest rivers is potentially the most interesting one relatively to the submitted query (for instance, the

second meeting in fig. 5). In a recent implementation, we also offered users the possibility to zoom within the ThemeRiver at a meeting level, displaying the evolution of themes within a single meeting, minute after minute.

4.2   Browsing the Results Using Thematic and Temporal Links

The major contribution of this paper is the use of links to discover new documents. The entry point presented in the previous subsection contains a set of results depending on the query. Sometimes, users have already found the documents of interest, but often the query mechanism is not enough to satisfy their needs. Moreover, when media such as audio or video are poorly indexed, searching is not useful. For these reasons, we propose to use links between documents as a new way to browse in a multimedia archive.
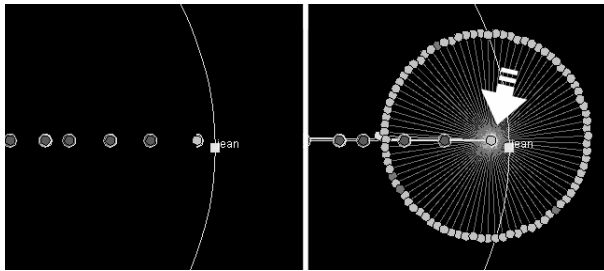

Fig. 6. The user finds an interesting document (1) and she discovers all related documents (2)

When users are interested in a document returned by the query, they can expand the search's results with the linked documents. The radial visualization is thus enriched with these similar documents. Their distribution depends on the relationships
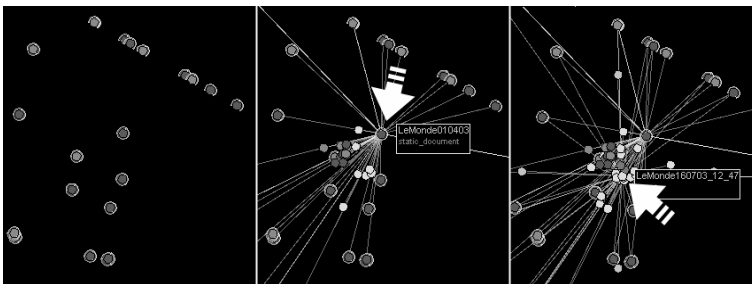

Fig. 7. Some of the new documents are similar or connected to already visualized ones

they have with already visualized documents. On Figure 6, the new documents only share relationships with the selected document and thus, are simply organized around it (the white arrow indicates the clicked document of interest).

The figure 7 illustrates a different case: when a new document is linked with several documents already in the RadViz, its location is defined at the center of mass of all these other documents' locations.

Moreover, users are able to decide that one document displayed in the RadViz is not interesting: hiding this document implies that the most similar documents can be eliminated too from the visualization. Similarity is pre-calculated for each category of links (see section 3.2) and the desired threshold can be changed, thanks to sliders that filter accordingly the search space. In the near future, we will integrate a functionality allowing users to select two or more documents: FaericWorld will then display the documents that optimize the distance between the selected ones.
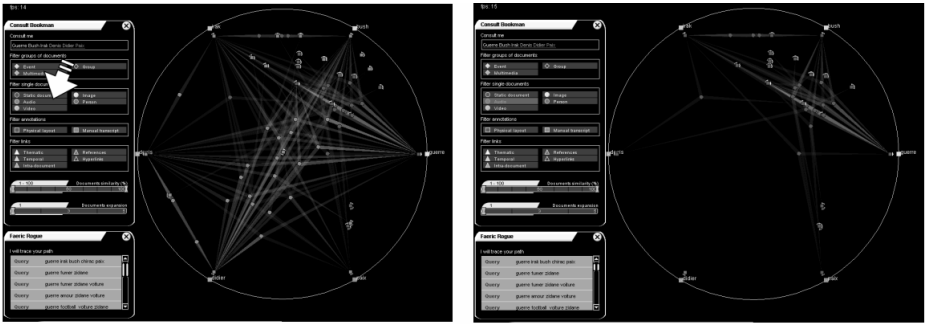

Fig. 8. The image shows all the documents corresponding to the submitted query (1). After filtering, all audio files are hidden (2).

Finally, filtering capabilities have been included in order to facilitate browsing in the world and to configure the RadViz when a large amount of document is displayed (fig. 8). In particular, filters allow selecting the categories of documents or links to visualize. Similarly, two scrollbars control the similarity thresholds: when they are activated, the links with a too low value are hidden and vice versa. Furthermore, when a document discovered while browsing not any more possessed visible links, it is removed from the set of displayed documents.

4.3   Viewing and Playing: Browsing in Time

Finding and consulting a document is only a part of the browsing task. In fact, users are frequently stimulated to extend the browsing experience while consulting a document. Consequently, another contribution of our system consists in using links in order to propose users new related documents while viewing the content of a document. Meetings have a particular property: all the media they contain are strictly synchronized to the meeting time. At opposite of non time-based documents, which are always connected to the same set of related documents, the themes constantly evolve during a meeting throughout time and, consequently, the set of linked documents changes.

Figure 9 is a mockup that illustrates a meeting player. All the media of the meetings are synchronized and when a user for instance clicks a paragraph of a PDF document discussed during the meeting, the transcript of the dialogs moves to the

instant when this document was in the verbal focus. This is done through the document/speech alignment that directly benefits from the thematic links. At contemporary, the audio file and the videos are synchronized thanks to temporal links. The meeting view is a screenshot of JFriDoc, a document-centric meeting browser we developed to navigate through meetings.
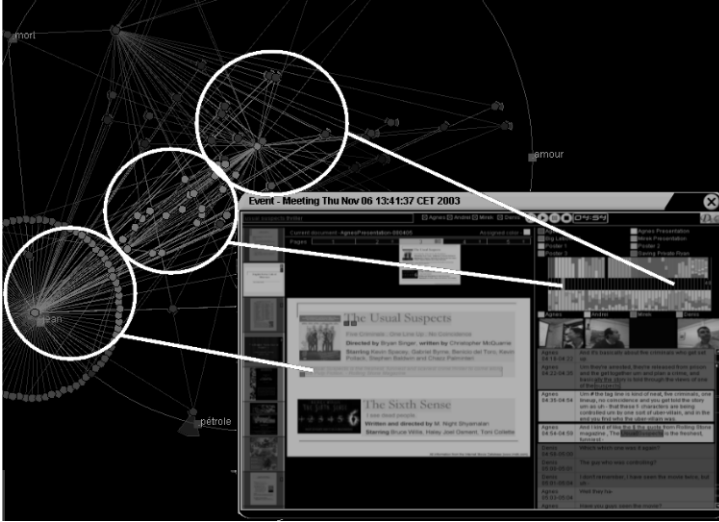

Fig. 9. In the integrated meeting browser, media can be synchronously played, thanks to the temporal alignments. Further, since topics evolve during a meeting, links and relations to the documents in the whole meetings archive dynamically change over time.

Finally, figure 9 illustrates cross-meetings and cross-documents links outside the current meeting. The thematic space of a meeting evolves during time, and thus also the thematic links with non-focused documents. For instance, during the meeting different static documents are projected or discussed and thus different links are activated when the corresponding theme changes.

5   Performances

The performance of FaericWorld has been tested. Currently, two corpora have been integrated in the system: IM2.DI corpus and AMI corpus. The first one consists in 22 meetings recorded in French. The data taken into account by FaericWorld are newspapers that have been manually annotated with logical structures, videos of participants with manual labels declaring speaker identity, audio files enriched with manual transcript, and, finally, meeting descriptors. This corpus has been used for developing the whole system.

FaericWorld system has further been validated with the second corpus, which contains 171 meetings in English. The most part of them belongs to groups of 4 meeting, sharing an identical scenario. The corpus is rich of manual annotations but only part

of them has been selected. In particular, FaericWorld uses utterances of audio transcripts (performed manually), tags on videos (id codes for participants) and meetings' descriptors. Each meeting contains PowerPoint presentations, technical reports and summaries, which have been automatically analyzed and annotated with physical structures.

Both corpora of meeting have been automatically imported, creating simultaneously specific annotations for some category of media such as static documents. At import stage, multiple copies of the same document were eliminated. The retained documents were fragmented into their parts thanks to structural annotations (for instance, the physical structure for static document and the utterances composing speech transcript of meeting dialogs). Links were created thanks to thematic alignment of parts, lexical analysis and media synchronization. FaericWorld automatically indexed and structured IM2.DI and AMI corpora. Table 2 summarizes the major characteristics of the loaded corpora.

Table 2. FaericWorld automatically indexed and structured IM2.DI and AMI corpora. The table summarizes corpora and worlds characteristics.

	Corpora	
	IM2.DI	AMI
Meetings	22	171
Imported documents	245	3.644
Unique documents	176	1.697
Fragments (indexed parts of document)	4.278	113.905
Links between documents (similarity threshold: 10%)	38.747	133.139.945

The PC used for importing the corpora and for accessing indexed data has a Pentium 4 CPU at 2.40 GHz, with 512 MB of RAM. Importing time for IM2 and AMI corpora correspond respectively to 50 minutes and 26 hours. Calculation time augments with the amount of imported documents: the alignment is the bottleneck that, however, could be resolved thanks to parallel calculations. With AMI corpus, submitting a query of 8 words and preparing the visualization takes in mean less than half a second. Retrieving and visualizing the documents linked to a document of interest requires in general less than 1 second. At the time of writing, no formal user evaluation of the browser has been performed, but a discount usability evaluation has shown that the RadViz visualization is sometimes hard to interpret. At opposite, the ThemeRiver has been well accepted by users.

6 Conclusion

This paper presents a novel approach, based on cross-media links for browsing through an archive of multimedia documents. Automatically computed links open new thematically and temporally related spaces. Instead of refining search queries step by step, users can select an entry point in the archive and then follow links to jump from one thematic space to another one. Moreover, they are able to discover in-

teresting multimedia documents, such as video and audio, generally hard to index and retrieve because of the difficulty to extract high semantic abstractions. With our approach, multimedia documents become more visible, thanks to their relationships with other documents.

Moreover, browsing is an incremental experience - a discovery process - that requires a shallow consultation of documents, playing parts of meetings, etc. in order to crystallize knowledge. In our system, browsing through links updates continuously the visualization of the archive, in order to expand user's navigation.

Future works will consist in extending user visual and interactive experience. Currently, FaericWorld uses a RadViz in order to visualize the entire world, using an automatic query as entry point. In fact, it is difficult to distinguish thematic groups of documents, because of theirs' high fragmentation. To overcome this problem, we will develop a new entry view that aims at highlighting clusters of documents using their similarity. Another work in progress is related to the diversity of media and annotations involved in our worlds, which implies the development of specific views for visualizing and interacting with each category of document. Up to now, only meetings and static documents profit of interactive views. Moreover, we shall introduce in the system edition capabilities in order to modify, create and destroy annotations and links. Edition capabilities are as well necessary for creating new documents, which can be shared among users as access views to the archive. Finally, a user evaluation shall be performed, in order to measure the usability of links for visually representing multimedia archives and for guiding the user during a browsing task.

References

1. Alice in Wonderland, TextArc, http://www.textarc.org
2. Behera, A.: A Visual Signature-based Identification Method of Low-resolution Document Images and its Exploitation to Automate Indexing of Multimodal Recordings. University of Fribourg, Switzerland, thesis Nr. 1529
3. Bollacker, K.D., Lawrence, S., Lee Giles, C.: CiteSeer: an autonomous web agent for automatic retrieval and identification of interesting publications. In: Proc. Of 2nd International Conference on Autonomous Agents, pp. 116–123. ACM Press, New York (1998)
4. Callan, J.P.: Passage-level evidence in document retrieval. In: Proc. of the 17th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 302–310. Springer, Heidelberg (1994)
5. Campanella, M., Leonardi, R., Migliorati, P.: An intuitive graphic environment for navigation and classification of multimedia documents. In: ICME 2005. Proc. Of Multimedia and Expo, pp. 743–746. IEEE Press, Los Alamitos (2005)
6. Carey, M., Heesch, D.C., Rüger, S.M.: Info Navigator: A Visualization Tool for Document Searching and Browsing. In: Proc. DMS'03, pp. 23–38 (2003)
7. Havre, S., Hetzler, E., Whitney, P., Nowell, L.: ThemeRiver: visualizing thematic changes in large document collections. In: IEEE Transactions on Visualization and Computer Graphics, pp. 9–20. IEEE Press, Los Alamitos (2002)
8. Hoffman, P., Grinstein, G., Marx, K., Grosse, I., Stanley, E.: DNA visual and analytic data mining. In: Proc. Of Visualisation'97, pp. 437–441. IEEE Press, Los Alamitos (1997)

9. Kartoo, http://www.kartoo.com
10. Kuper, J., Saggion, H., Cunningham, H., Declerck, T., de Jong, F., Reidsma, D., Wilks, Y., Wittenburgh, P.: Intelligent Multimedia Indexing And Retrieval through Multi-source Information Extraction and Merging. In: Proc of IJCAI, pp. 409–414 (2003)
11. Goularte, R., Camacho-Guerrero, J.A., Inácio Jr, V.R., Cattelan, R.G., Pimentel, M.d.G.C.: M4Note: a Multimodal Tool for Multimedia Annotations. In: Proc. of WebMedia and LA-Web, pp. 142–149. IEEE Press, Los Alamitos (2004)
12. Integrated JFerret Browser and Overlapped Speech Browser, In: Demonstration Session Guide, MLMI'06, http://groups.inf.ed.ac.uk/mlmi06/MLMI-2006-DemoSessionFinal.pdf
13. Lalanne, D., Sire, S., Ingold, R., Behera, A., Mekhaldi, D., Von Rotz, D.: A research agenda for assessing the utility of document annotations in multimedia databases of meeting recordings. In: Proc. of 3rd International Workshop on Multimedia Data and Document Engineering, in conjunction with VLDB-2003, pp. 47–55 (2003)
14. Lalanne, D., Ingold, R., Von Rotz, D., Behera, A., Mekhaldi, D., Popescu-Belis, A.: Using Static Documents as Structured and Thematic Interfaces to Multimedia Meeting Archives. In: Renals, S., Bengio, S. (eds.) MLMI 2005. LNCS, vol. 3869, Springer, Heidelberg (2006)
15. Lalanne, D., Lisowska, A., Bruno, E., Flynn, M., Georgescul, M., Guillemot, M., Janvier, B., Marchand-Maillet, S., Melichar, M., Moenne-Loccoz, N., Popescu-Belis, A., Rajman, M., Rigamonti, M., von Rotz, D., Wellner, P.: The IM2 Multimodal Meeting Browser Family. IM2 technical report (2005)
16. LinkedIn, https://www.linkedin.com
17. Lisowska, A., Rajman, M., Bui, T.H.: ARCHIVUS: A System for Accessing the Content of Recorded Multimodal Meetings. In: Proc. of the Joint AMI/PASCAL/IM2/M4 Workshop on Multimodal Interaction and Related Machine Learning Algorithms, pp. 291–304 (2004)
18. Mekhaldi, D.: A Study on multimodal document alignment: bridging the gap between textual documents and spoken language. University of Fribourg, Switzerland, thesis Nr. 1521
19. Rigamonti, M., Bloechle, J.L., Hadjar, K., Lalanne, D., Ingold, R.: Towards a canonical and structured representation of PDF documents through reverse engineering. In: Proc. of ICDAR'05, pp. 1050–1054 (2005)
20. Scholar, http://scholar.google.com
21. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn., p. 652. Addison-Wesley, Reading (2004)
22. Smith, J.R., Naphade, M., Natsev, A.(P.): Multimedia Semantic Indexing Using Model Vectors. In: ICME 2003. Proc. Of Multimedia and Expo, vol. II, pp. 445–448. IEEE Press, Los Alamitos (2003)
23. Swain, M.J.: Searching for Multimedia on the World Wide Web. In: Proc. of Multimedia Computing and Systems, vol. I, pp. 32–37. IEEE Press, Los Alamitos (1999)
24. Theisel, H., Kreuseler, M.: An enhanced spring model for information visualization. In: Proc. of Eurographics 98, vol. 17(3), pp. 335–344. Blackwell Publishing, Malden (1998)
25. Tucker, S., Whittaker, S.: Accessing Multimodal Meeting Data: Systems, Problems and Possibilities. In: Bengio, S., Bourlard, H. (eds.) MLMI 2004. LNCS, vol. 3361, pp. 1–11. Springer, Heidelberg (2005)
26. Tucker, S., Whittaker, S.: Reviewing Multimedia Meeting Records: Current Approaches. In: Multimodal multiparty meeting processing workshop, ICMI 2005, International Conference on Multimodal Interfaces (2005)
27. Wellner, P., Flynn, M., Guillemot, M.: Browsing recorded meetings with Ferret. In: Bengio, S., Bourlard, H. (eds.) MLMI 2004. LNCS, vol. 3361, pp. 12–21. Springer, Heidelberg (2005)
28. YouTube, http://www.youtube.com