Intelligent Search in a Collection of Video Lectures

Angela Fogarolli

University of Trento,
Dept. of Information and Communication Tech.,
Via Sommarive 10, 38050 Trento, Italy
`afogarol@dit.unitn.it`

1   Abstract

In recent years, the use of streamed digital video as a teaching and learning resource has become an increasingly attractive option for many educators as an innovation which expands the range of learning resources available to students by moving away from static text-and-graphic resources towards a video-rich learning environment. Streamed video is already widely used in some universities and it is mostly being used for transmitting unenhanced recordings of live lectures.

What we are proposing is a way of enriching this video streaming scenario in the eLearning context. We want to extract information from the video and the correlated materials and make them searchable. Thus, the aim of this thesis is to create a semantically searchable collection of video lectures.

In the literature, surprisingly little information can be found about speech and document retrieval in combination with lecture recording. There are interesting examples of e-lecture creation and delivery e.g. [5], audio retrieval of lecture recording [3] that explore automatic processing of speech, or systems such as the eLecture portal [1] which indexes the audio and also the text of the lecture slides. But to the best of our knowledge there is no system which combines and synchronizes the different modalities in a searchable collection.

What we propose is enabling the search and navigation through the different media types presented in a frontal lecture with the addition of the video recording. In video indexing domain Snoek and Worring [4] have proposed to define multimodality as "the capacity of an author of the video document to express a predefined semantic idea, by combining a layout with a specific content, using at least two information channels". The channels or modalities of a video document described in [4] are the visual, auditory and textual modality.

We believe that using more than one modality – as explained in [2] – could increase productivity also in the context of e-learning, where it is really frequent to scan for information. Our main focus is to enable search on two modalities; in particular we will index the auditory modality of video lecture content based on transcription obtained with automatic speech recognition tools and on textual modality using text indexing on the related materials. Furthermore, we do not just want to present an enhanced version of the current state of the art in

e-lecture retrieval but we also envision to add semantic capabilities to the search functionalities in order to provide a superior learning experience to the student (personalized search, personalized learning path, relevant contextualization, automatic video content profiling...).

The application we are proposing is a way to provide a tool for students to enable more flexibility in e-Lectures consumption. The student could seek inside a collection of learning lectures and related materials (desktop activities recording, PowerPoint presentations, interactive whiteboard tracks...). The search could also be personalized to meet the student demands. For each hit the system would display the lectures video-recording and the temporally synchronized learning materials. Another benefit we want to archive using Semantic Web techniques is to present a profile information of the content of the video lecture, this could lead to an improvement of the state of the art in the video-indexing field allowing automatic profile annotation of the content of the video.

The research work specifically addressed by this thesis will investigate the following challenges:

- Finding an innovative way for mastering the gap between information extraction and knowledge representation in our context. For each video and related learning resource an RDF representation would be extracted. The created graph would be navigated during the search task to find the requested information and suggest related topics using ontology linkage. We will use ontologies for high level lecture description and query understanding.
- Automatic content description of the presented learning material. A textual description of the content of a video result, lecture or course would be presented at the user. This could be realized presenting a profile information of the knowledge extracted from the video and the related material.
- Evaluation of the tool value for improving the student performance and for shortening the learning time we will conduct before and after the semantic enhancement.

References

[1] Christoph Hermann, Wolfgang Hürst, and Martina Welte. The electure portal: An advanced archive for lecture recordings. In *Informatics Education Europe Conference*, 2006.

[2] D. Jones, Shen A., and D W., Reynolds. Two experiments comparing reading with listening for human processing of conversational telephone. In *Interspeech 2005 Eurospeech Conference*, 2005.

[3] A. Park, T. Hazen, and J. Glass. Automatic processing of audio lectures for information retrieval: Vocabulary selection and language modeling. In *ICASSP*, March 2005.

[4] C.G.M. Snoek and M. Worring. Multimodal video indexing: A review of the state-of-the-art. In *Multimedia Tools and Applications*, number 25, pages 5–35, 2005.

[5] Z. Zhu, C. McKittrick, and W. Li. Virtualized classroom automated production, media integration and user-customized presentation. In *Multimedia Data and Document Engineering*, July 2004. Semantic Web.