A Novel Lecture Browser Using Key Phrases and Stream Graphs

Korbinian Riedhammer, Martin Gropp, Elmar Nöth

Lehrstuhl für Mustererkennung, Universität Erlangen-Nürnberg
Martensstraße 3, 91058 Erlangen, GERMANY
`korbinian.riedhammer@informatik.uni-erlangen.de`

Abstract—We present a novel lecture browser that utilizes ranked key phrases displayed on a stream graph to overcome the shortcomings of traditional extractive (query-based) summaries. The system extracts key phrases from the ASR transcripts, performs an unsupervised ranking, and displays an initial number of phrases on the stream graph. This graph gives an intuition of *when which* key phrase is spoken, and *how dominant* it is throughout the lecture. The user can select the phrases to be displayed and furthermore adjust the ranking of the all phrases. All user interactions are logged to a server to improve the ranking algorithms and provide user specific rankings.

I. INTRODUCTION

Most work on speech summarization focuses on extractive summarization where salient utterances are concatenated to a summary. Given broadcast news or other read speech, this can result in fairly well readable and informative summaries. However, meetings, lectures and similar forms of spoken documents contain rather spontaneous speech with all the errors, filled pauses and other artifacts. As a result, concatenated utterances typically form a rather incoherent and hardly readable conglomerate.

For this demonstration, we have a user in mind who is willing to spend some time on the document given that he finds the regions of interest quickly. For example, consider a student who missed a lecture but has access to the full recording. The "optimization" problem is now to watch as least of the recording as possible while getting all the important information. Therefor, the student needs to know *what* topics were covered and *when* they came up during the lecture.

In previous work, we could show that key phrases are a great base to construct summaries [1] and that a human refinement of those further increases the performance [2]. Also, we could show for lecture data that the automatically extracted key phrases are very similar to human annotated key phrases [3].

This demonstration implements the proposal in [3] and allows the user to

- Select key phrases for display using drag and drop.
- Add, remove and rerank key phrases using drag and drop to emphasize personal preferences.
- Use the stream graph as an intelligent timeline to navigate within the video.

All user interactions are logged to collect a database to improve both general and personalized ranking algorithms.

Figure 1. Screen shot of the browsing interface; *top-right:* selected phrases (colored as in the stream graph) – *bottom-right:* available phrases.

II. SYSTEM DESCRIPTION

A. Key Phrase Extraction

The ASR output is piped through a sentence detector, tokenizer, POS tagger, word form normalizer, stop words filter and topic segmenter. Then, noun phrases are extracted as candidates and ranked based on their length, frequency and cooccurrences.

B. Visualization

The stream graphs are based on [4] and show for the selected key phrases where they occur and how dominant they are at that time. The interface was implemented using the Google Web Toolkit (GWT).

REFERENCES

[1] K. Riedhammer, B. Favre, and D. Hakkani-Tür, "Long Story Short – Global Unsupervised Models for Keyphrase Based Meeting Summarization," *Speech Communication*, vol. 52, no. 10, pp. 801–815, 2010.
[2] ——, "A Keyphrase Based Approach to Interactive Meeting Summarization," in *Proc. IEEE Workshop on Spoken Language Technologies (SLT)*, 2008, pp. 153–156.
[3] M. Gropp, E. Nöth, and K. Riedhammer, "A Novel Lecture Browsing System Using Ranked Key Phrases and StreamGraphs," in *Proc. Int'l Conference on Text, Speech and Dialogue (TSD)*, 2011, pp. 17–24.
[4] L. Byron and M. Wattenberg, "Stacked Graphs – Geometry & Aesthetics," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 14, no. 6, pp. 1245–1252, 2008.