Browsing within Lecture Videos Based on the Chain Index of Speech Transcription

Ravi Raipuria
August 18, 2012

1 Philosophy

In today’s world, the recorded video lecture is used more frequently by institutions for distance education purposes. However, detailed browsing in lecture videos is not supported due to lack of indexing. So, the main objective of this paper is to provide an efficient indexing approach for recorded video lectures.

2 Summary

Indexing of a video lecture requires two main steps:

• Split the video into coherent segments
• Segments have to specify the topic with semantic description

2.1 Status of Video Lecture Indexing

There are four kinds of indexing techniques for video lectures:

• Indexing can be done manually after the lecture by professionals
• Users of the video can do the indexing themselves by tagging
• Slides can be used as a index for related video lectures
• Speech can also be used as a resource for automatic indexing

2.1.1 Manual Indexing

In this indexing, both segmentation and annotations are done manually. Description Logics (DL) are used to formalize the semantic annotations in manual indexing. DL allows the knowledge of an application domain to be represented in a structured way. In DL conceptual knowledge of an application is represented in terms of concepts (Set of individuals) and rules (Relation between individuals). Four tasks have to be done for manual indexing.

• Create the taxonomy manually
• Creating dictionaries for mapping NL words to concepts and roles.
• Segmentation of video into parts called learning object (LO)
• Annotating the video segments with concepts and rules
2.1.2 Tagging

Tagging is a process of annotating the document with an unstructured list of keywords. Collaboration tagging can be used for generation and enrichment of video meta-data to support content based retrieval. This technique will take lot of time to annotating a document and it depend on the many learners who have the ability to annotate.

2.1.3 Slides

Slides represent the most of the information in video segments. Slides can also be use for indexing of related video lectures. For synchronizing between slides and videos, a log file is maintained during the presentation which store the timestamps(start and finish) of each slide.

2.1.4 Transcripts of Speech Recognition Engines

Speech is used as a resource of index. Each transition of vocabulary (eg: pause) could mark a segment boundary. So, the video can be divided into parts. The main task is to convert the speech into text and divide into sequence of topically coherent segments.

2.2 Chain Indexing

In chaining, We create a List L containing the all distinct terms identifies by SRE. Now a chain is constructed consisting of all repeating ranging from the first to the last appearance of the term in the lecture. Chain is divide into subparts when there is long break in terms(say 2 to 10 min). The process for chaining is work as follows [1]:

- Take term T1 from List L.
- Build Clusters
- Store the chain data in database as a chain index.
- Take the next term from L and repeat whole process.

References