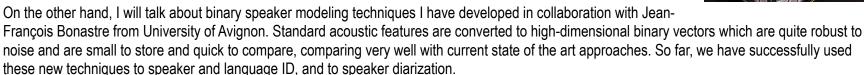Friday December 9th, 2011 • 11:00-12:00 • Interschool Lab, 7th floor Schapiro CEPSR

Identification of similar videos and speakers

Xavier Anguera

Telefónica Multimedia Research, Barcelona

This talk covers two topics I have been working on in the last two years. On the one hand, I will talk about video-copy detection using multimodal features. The objective of a video-copy detection system is to find video segments that have been copied from an originally (known) video. These segments might have been heavily tampered with, but are considered copies if a user is still able to consider both videos as essentially the same. I will explain our multimodal (audio + video) approach, describing the novel audio and video features we applied to the task.

On the other hand, I will talk about binary speaker modeling techniques I have developed in collaboration with Jean-François Bonastre from University of Avignon. Standard acoustic features are converted to high-dimensional binary vectors which are quite robust to noise and are small to store and quick to compare, comparing very well with current state of the art approaches. So far, we have successfully used these new techniques to speaker and language ID, and to speaker diarization.

Xavier Anguera Miro: Ing. [MS] 2001 UPC University (Barcelona, Spain), [MS] 2001 European Masters in Language and Speech, Dr. [PhD] 2006 UPC University, working on speaker diarization for multimicrophone recordings. From 2001 to 2003 he helped develop multilingual TTS systems for Panasonic Speech Technology Lab in Santa Barbara, CA. From 2004 to 2006 he was a research visitor at the International Computer Science Institute (ICSI) in Berkeley, CA. Since 2007 he is with Telefonica Research in Barcelona, Spain working on multimedia information indexing and retrieval, with emphasis on multimodal analysis. His areas of interest cover audio, video and music indexing and retrieval, speaker recognition and diarization and pattern matching algorithms.

Host: John Wright <johnwright@ee.columbia.edu> / Dan Ellis <dpwe@ee.columbia.edu>