



IMSC
Integrated
Media Systems
Center

INTEGRATED MEDIA SYSTEMS CENTER
A National Science Foundation
Engineering Research Center at the
UNIVERSITY OF SOUTHERN CALIFORNIA

PRINCIPAL INVESTIGATOR

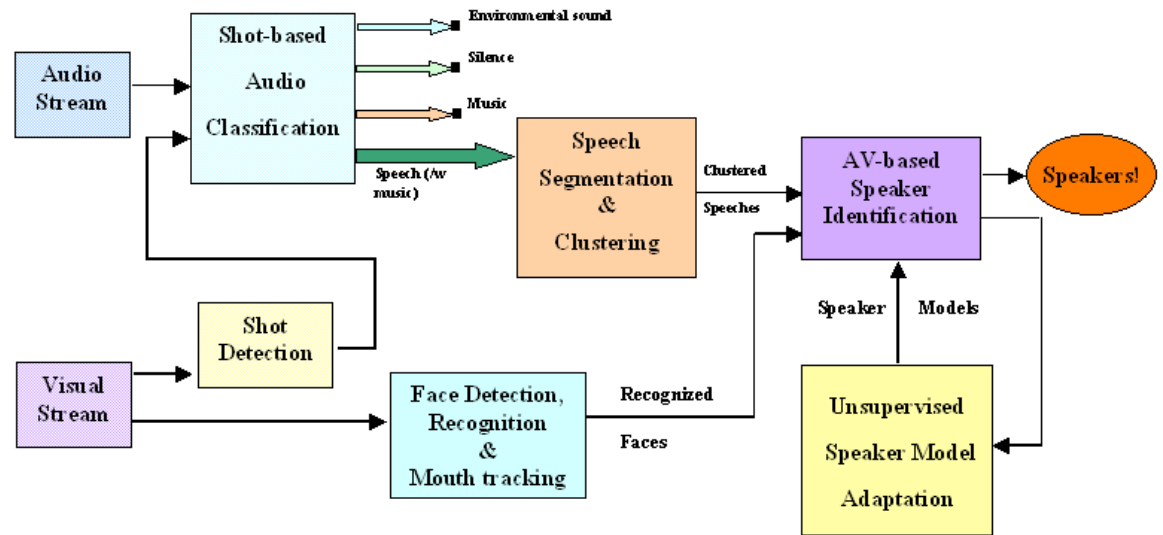
C.-C. Jay Kuo
cckuo@sipi.usc.edu

OTHER USC RESEARCHERS

Shrikanth Narayanan
shri@sipi.usc.edu

USC STAFF

Unsupervised Speaker Identification for Movie Content Analysis



Block diagram of the system:

USC STUDENTS, DEGREES

Ying Li (Ph. D)

BRIEF DESCRIPTION OF DEMONSTRATION

The system can detect speech shots in a movie and automatically identify the speaker. Both audio and visual cues are employed and subsequently combined in a probabilistic framework to identify speakers. Particularly, audio information is used to identify speakers with a maximum likelihood based approach. The extracted information can be used to build the video's Table of Content as well as the index table.

UNIQUE OR DISTINGUISHING CHARACTERISTICS RELATIVE TO STATE-OF-THE-ART

While most previous work on speaker identification was carried out in a supervised mode using pure audio data, our approach integrates knowledge from multimedia sources in an unsupervised mode. It not only saves a lot of time for pre-training speaker models, but also is more robust because it can accommodate for speakers' acoustic variations along time by adapting to their newly contributed speech.

<p>APPLICATIONS</p> <ul style="list-style-type: none"> • Video indexing, browsing and retrieval • Video summarization 	<p>RECENT HIGHLIGHTS, LEVEL OF DEVELOPMENT, UPCOMING MILESTONES</p> <ul style="list-style-type: none"> • Commercial break detection algorithm, including integrating video and audio features. • Develop a video skimming system for home video. The system will be implemented based on tempo analysis of background music.
<p>UNDERLYING TECHNOLOGIES</p> <ul style="list-style-type: none"> • Video shot detection algorithm • Face detection and recognition algorithm • Speech clustering • Maximum likelihood-based speaker identification • MAP-based model adaptation 	
<p>LIST OF PUBLICATIONS, REFERENCES, URLs</p> <ol style="list-style-type: none"> 1. Ying Li, S. Narayanan and C.-C. Jay Kuo, "Identification of speakers in movie dialogs using audiovisual cues", <i>ICASSP2002</i>, May 2002. 2. Ying Li and C.-C. Jay Kuo, "Unsupervised real-time speaker identification for daily movies", <i>SPIE Proc. on Internet Multimedia Management Systems III (ITCOM'02)</i>, Vol, 4862, August 2002. 3. Ying Li, S. Narayanan and C.-C. Jay Kuo, "Content-based movie analysis and indexing based on audiovisual cues", Submitted to <i>IEEE Trans. on Circuit and Systems for Video Technology</i>, 2002. 4. Ying Li and C.-C. Jay Kuo, "A robust video scene extraction approach to movie content abstraction", To appear in the <i>International Journal of Imaging Systems and Technology with Special Issue on Multimedia Content Description and Video Compression</i>, 2002. 	

For additional information, please contact the Principal Investigator listed above via email, or contact

Isaac Maya, Ph.D., P.E.
 Director, Industry and Technology Transfer Programs

213-740-2592
imaya@imsc.usc.edu

Ann Spurgeon
 Associate Director of Industry Programs

213-740-4877
aspurgeo@imsc.usc.edu

Integrated Media Systems Center
 3740 McClintock Avenue, Suite 131
 Los Angeles, CA 90089-2561
 213-740-8931 (fax)

For additional information on the Integrated Media Systems Center (IMSC), please visit our web site at <http://imsc.usc.edu>