

Audio Hot Spotting

Qian Hu

781-271-2959 • qian@mitre.org

MITRE Sponsored Research



Problem

Nuggets of information are buried within huge mountains of multi-media data. A system is needed to quickly and automatically identify and retrieve these audio “hot spots.”

The current approach to audio information retrieval of simply combining text-based information retrieval with automatic speech recognition does not meet user needs in real applications.

Background

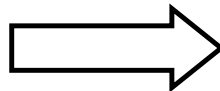
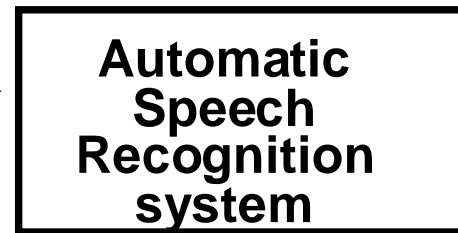
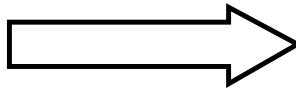
Current Approach: ASR + IR Doesn't Meet the Needs of Critical Applications

".. cross into **Korea**"



Smuggler ID 719

Background surf noise



"... cross into career"

Nothing Found

[TREC 6-9: Spoken Document Retrieval]

Objective

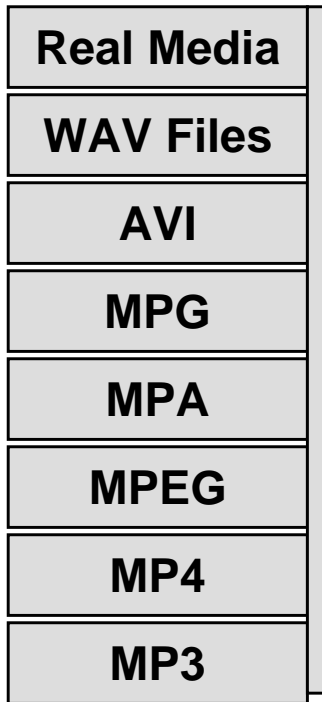
- **Provide prototype hot-spotting capability to enable efficient audio filtering and retrieval**
 - **Determine the limitations of algorithms in existing component technologies**
 - **Leverage, integrate, and extend the best technologies for the hot-spotting problem**
 - **Research and develop new algorithms when COTS and GOTS systems fail to meet the requirements of audio hot-spotting (AHS)**
 - **Research and develop audio-specific query algorithms making use of multiple types of audio information**

Activities

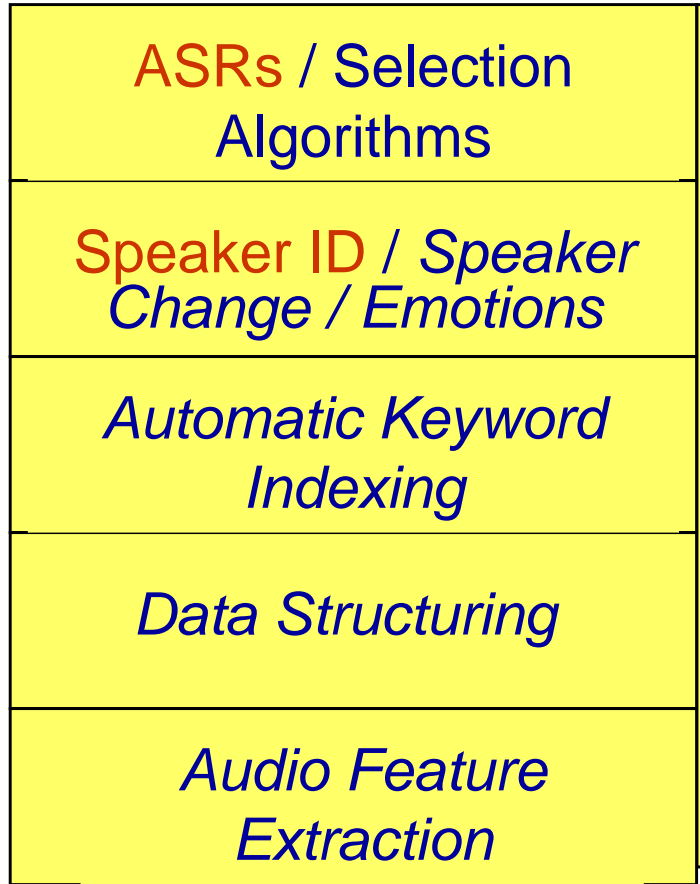
- **Developed AHS evaluation criteria and automated evaluation tool kit**
- **Researched algorithms for detecting speaker change, prosodic features indicating elevated emotions, and overlapping speech**
- **Deployed the prototype for field experimentation**
- **Researched methods to combine word- and phoneme-based ASR systems for error correction and improved retrieval**
- **Conducted AHS experiments on specific domain applications (e.g., air traffic control)**

Highlight

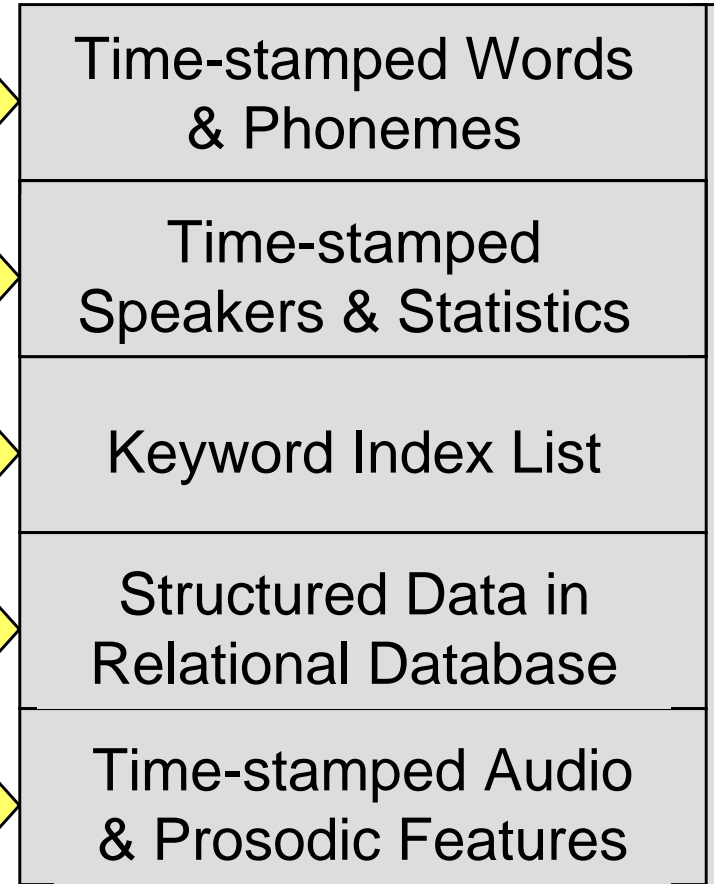
Input



Preprocessing



Output



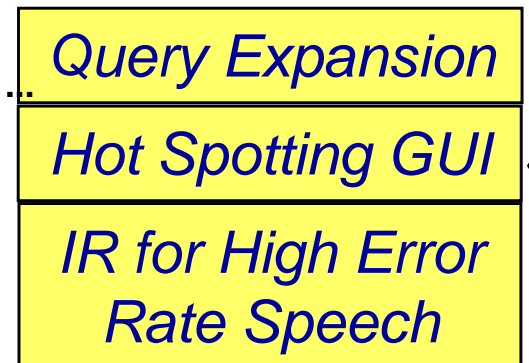
Blue = MITRE

Red = GOTS/COTS

ATF agents located the TNT ...



Identified Selection Available
as Original Media or Text



Demonstration

Built Audio Hot Spotting prototype to query and retrieve key words, speakers, audio, and prosodic effects for English and Spanish

The screenshot shows the 'Audio Hot Spotting' web application in a Microsoft Internet Explorer browser. The address bar shows the URL: http://mm106647-2k.mitre.org/pls/audio/audio_rev14.results_single_combined. The page features a logo with the text 'AUDIO HOT SPOTTING' and a magnifying glass over the word 'SPOTTING'. The interface is divided into several sections:

- Search Results:** Displays 'Word/Phrase: "terrorism"' and 'Speaker: "Marty Faga"'. Below this, it states 'Your search produced 3 hit(s)'.
- Table of Results:** A table with three columns: 'Time', 'Text', and 'Speaker'. The 'Text' column contains text excerpts with the word 'terrorism' highlighted in red.
- Text Excerpt:** A paragraph of text at the bottom of the results section, with 'terrorism' highlighted in red.
- Query Cues:** A vertical yellow box on the left side of the interface, containing the text 'QUERY CUES'.
- Annotations:** Several yellow callout boxes provide additional information: 'Allows Combination of Query Cues' points to the search input fields; 'Provides both multimedia and text retrieval' points to the search results section; 'Keyword' points to the 'terrorism' search term; 'Speaker' points to the 'Marty Faga' search term; and 'Audio Effect' points to the 'Sound effect search' dropdown.
- Media Viewer:** A small window in the top right corner shows a video of a man speaking at a podium.

Impacts

- Presented our work at the 3rd International Workshop on Multimedia Data and Document Engineering
- Demonstrated new ways to explore hidden information from multimedia data such as prosody and background audio cues
- Briefed MITRE sponsors of possibilities of and ways to applying AHS technology
- Worked with industry leaders to improve component technologies

Future Plans

- Audio-specific query algorithms
- Domain-specific applications
- IR & audio mining research
- Explore multiple ASR engines

- Audio feature extraction
- Speaker change detection
- Foreign language support
- Query expansion research

- Integrate and extend component technologies
- Develop a prototype system to **hotspot** speaker and/or keywords

