
THISL progress report - 1999sep

Dan Ellis

International Computer Science Institute, Berkeley CA

<dpwe@icsi.berkeley.edu>

Outline

- 1 Spoken queries recognized
- 2 TREC modspeg processed
- 3 ThisIGUI updated
- 4 FYI: General audio retrieval



1

Spoken query recognition

- 23 TREC query sentences (351 words)
- Pronounced by 9 speakers (from BBC) + 1
- Recognized by RNN and RNN+MSG

Speaker	RNN WER	RNN+MSG WER
nch	16.0%	15.1%
mrm	21.4%	21.9%
cdm	24.5%	21.1%
jp	24.5%	24.8%
mac	25.9%	23.9%
mpj	26.5%	24.2%
ljc	26.5%	25.6%
cmb	35.0%	29.3%
cp	38.7%	34.2%
dpwe	36.2%	37.9%
Avg	27.5%	25.8%

- Pass to IR for evaluation ...



2

TREC modspec processing

- **Modulation-filtered spectrogram (msg) features**
 - helped in BN'98 eval
 - could help for TREC8 data?
 - run msg neural net for probability-combination with RNN
- **Automatic wav→msg→Ina loop**
 - ftp between Sheffield and ICSI
 - 2.5M weight network ran in 0.3xRT on TetraSPERT
 - 900+ files (550 hours) processed in 4 weeks



3

ThisIRGUI updated

Date filters

Speech input

BBC data

click-to-play

The screenshot shows the 'ThisIR demo' window with a menu bar (File, Options) and a toolbar. The main interface includes a search area with 'Enter query: at one stories and bangladesh', date filters for 'Start date: 1995 January 01' and 'End date: 2004 December 31', and a 'Programs' list. Below this is a table of search results for 'stories bangladesh' with columns for Program, Date, Offset, and Context. At the bottom, there is a playback area showing a transcript of a news segment.

Program	Date	Offset	Context
Radio 4: Midnight News	1998nov21	11:35	the homes people in southern bangladesh
Radio 4: Midnight News	1999mar18	24:35	the british high commission in bangladesh
Radio 4: Six O'Clock News	1998may27	23:35	dhaka / many arrive in bangladesh with th
Radio 4: Six O'Clock News	1998sep17	14:47	. how's the floods in bangladesh start to re
BBC1: Six O'Clock News	1998sep17	16:23	is facing the people in bangladesh / becau
BBC1: Nine O'Clock News	1998sep03	05:59	the atlantic / and in bangladesh / tens of t
BBC1: Nine O'Clock News	1998sep17	01:11	/ annie and a mere bangladesh goes down
BBC1: One O'Clock News	1997feb27	00:11	charles among the poor of bangladesh / th
BBC1: Nine O'Clock News	1998sep11	21:11	it's now clear that much bangladesh that i

Annotations on the screenshot include:

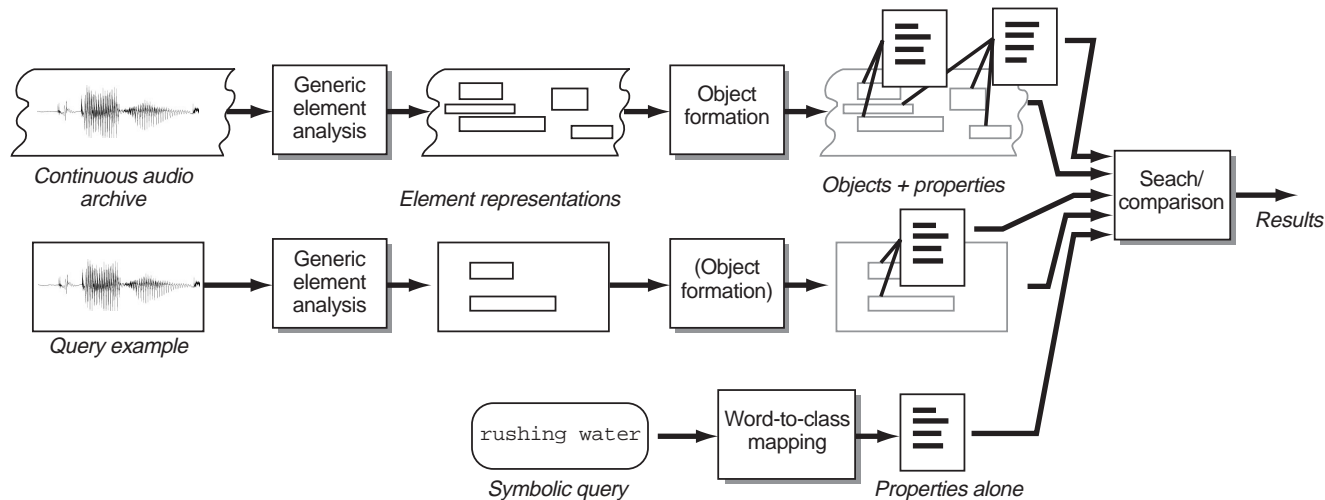
- Date filters:** Points to the start and end date selection fields.
- Speech input:** Points to the 'Record speech' button.
- BBC data:** Points to the search results table.
- click-to-play:** Points to the 'Play speech' button.
- Program filter:** Points to the 'Programs' list.
- thisIR-0.2 & thisIR-0.5 compatible:** A general annotation for the interface.
- Pauses & sentence breaks shown:** Points to the transcript text.
- Whole-show transcripts:** Points to the bottom section of the transcript.



4

General audio mixture retrieval

- **Information retrieval for nonspeech audio**
 - non-word events (laughter, shouting)
 - sound effects on soundtracks (engines, cars)
 - ambiences
- **Approach: auditory elements, learned patterns:**



- **Just-submitted proposal to NSF**

