

Audio Similarity Retrieval Engine

SISAP '10, September 18–19 2010, Istanbul, Turkey

Pavel Jurkas

Milan Stefina

David Novak

Michal Batko

{xjurkas|xstefina|david.novak|batko}@fi.muni.cz

Masaryk University, Brno, Czech Republic

MUFIN
Multi-feature Indexing Network

Overview

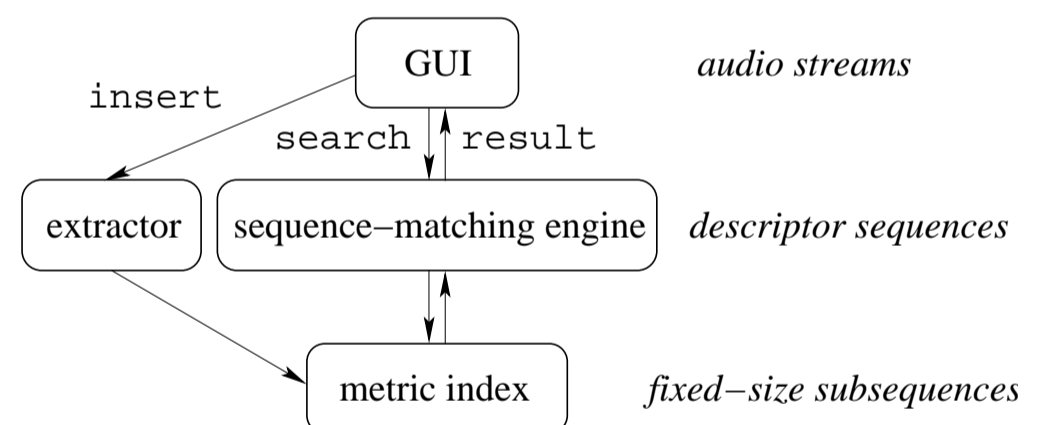
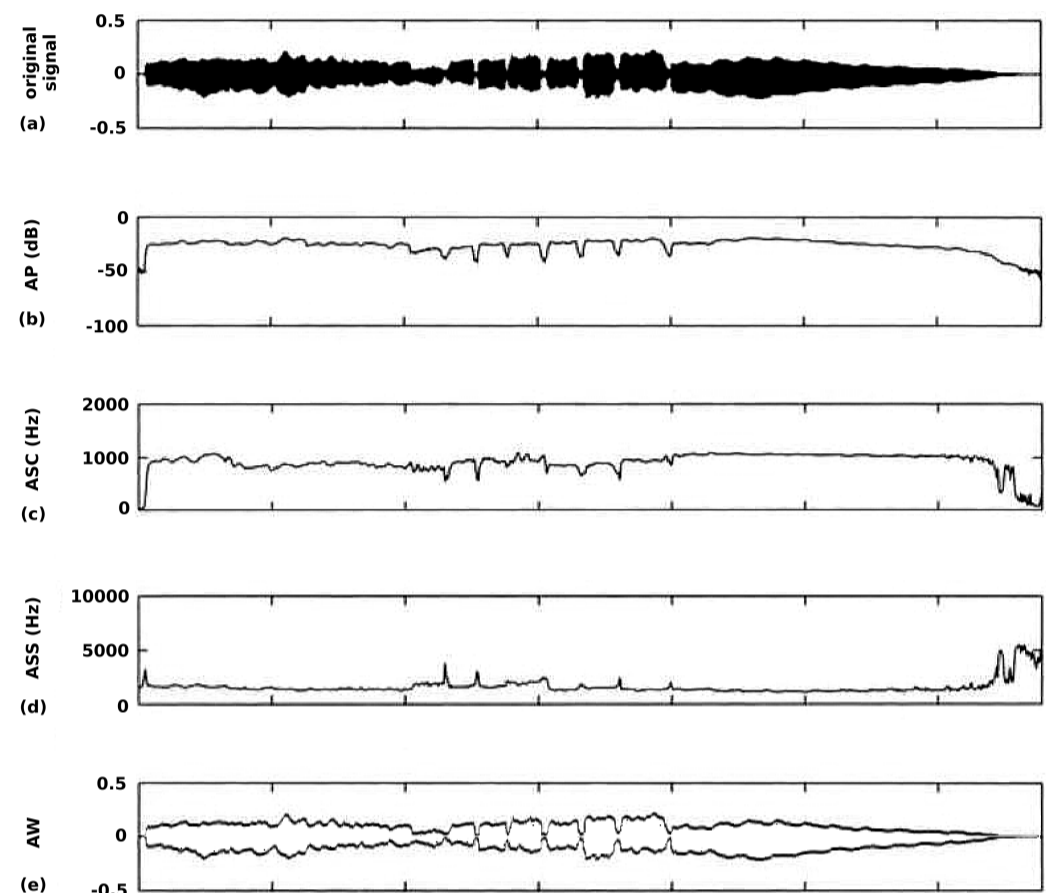
- general engine for similarity search in audio data
- primary indexing based on metric model
- modular architecture

Data

- short streams with various sound samples
- four low-level MPEG-7 Audio descriptors:
 - Audio Power (AP) – power of the signal
 - Audio Spectrum Centroid (ASC) – the central frequency
 - Audio Spectrum Spread (ASS) – frequency spread in the audio signal
 - Audio Waveform (AW) – min & max value

System Overview

- both data & query descriptors are split into fixed-size segments
- metric similarity (L_2) on level of segments
- overall similarity measured for “longest overlapping subsequence”
- segments organized in a metric index
- individual descriptors can be aggregated



Current Demo and Vision

Current Demo Evaluation

- 1,100 short sound samples
 - from *Partners in Rhyme* and *Freesound*
- human-judged evaluation of effectiveness
 - individual descriptors tested
 - not a clear winner for all test cases
- aggregation (1:1:1:1) gives best results
 - see <http://mufin.fi.muni.cz/audio/>

System Potential and Vision

- multiply size of the database
- implement better metric similarity functions
- test different descriptors (high-level)
- add semantics by keywords reranking
- create a universal sequence-matching module
- stream processing

Query: Animals\birdies\59360_dobroide_birdies.01.wav		
AP	Animals\birdies\59360_dobroide_birdies.01.wav	0.0
	Human\girl_Cough\45578_J.Zazvurek_Girl_Cough.wav	0.00136
	Human\Human_voice_Clock\80300_Corsica_S_04.wav	0.00147
	Human\Human_voice_Clock\80301_Corsica_S_05.wav	0.00152
	Human\Human_voice_Clock\80304_Corsica_S_08.wav	0.00153
ASC	Human\Human_voice_Clock\80298_Corsica_S_02.wav	0.00154
	Animals\birdies\59360_dobroide_birdies.01.wav	0.0
	Animals\Snake\snakehiss2.wav	0.03725
	Misc\Sword\1467_lostchocolatelab_10SWORD04.aif	0.04071
	Animals\birdies\59361_dobroide_birdies.02.wav	0.0418
ASS	Misc\Sword\1453_lostchocolatelab_06SWORD06.aif	0.04201
	Misc\Sword\1464_lostchocolatelab_10SWORD01.aif	0.04332
	Animals\birdies\59360_dobroide_birdies.01.wav	0.0
	Natural\AoE2\tf3.wav	0.02073
	Human\Baby_Noises\62070_NoiseCollector_vanessa10.wav	0.02076
AW	Animals\birdies\59373_dobroide_birdies.14.wav	0.02338
	Misc\Sword\1453_lostchocolatelab_06SWORD06.aif	0.02354
	Human\Baby_Noises\62079_NoiseCollector_vanessa8.wav	0.02397
	Animals\birdies\59360_dobroide_birdies.01.wav	0.0
	Human\Human_voice_Clock\80300_Corsica_S_04.wav	0.02293
1:1:1:1	Animals\birdies\59376_dobroide_birdies.17.wav	0.02393
	Human\Human_voice_Clock\80301_Corsica_S_05.wav	0.02439
	House\Metal_Pan\16373_JonathanJansen_Metaal_1.wav	0.026
	House\Metal_Pan\16383_JonathanJansen_Metaal_19.wav	0.026
	Animals\birdies\59360_dobroide_birdies.01.wav	0.0
1:1:1:1	Natural\AoE2\tf3.wav	0.10712
	Animals\birdies\59361_dobroide_birdies.02.wav	0.11945
	Animals\birdies\59370_dobroide_birdies.11.wav	0.11964
	Natural\AoE2\tf3.wav	0.121
	Animals\birdies\59363_dobroide_birdies.04.wav	0.12325