USING LYRICAL DURATION FOR AUTOMATIC LYRICS-TO-AUDIO ALIGNMENT IN CLASSICAL TURKISH MUSIC

Georgi Dzhambazov
@georgi_d_
georgi.dzhambazov@upf.edu

supervisor Dr. Xavier Serra

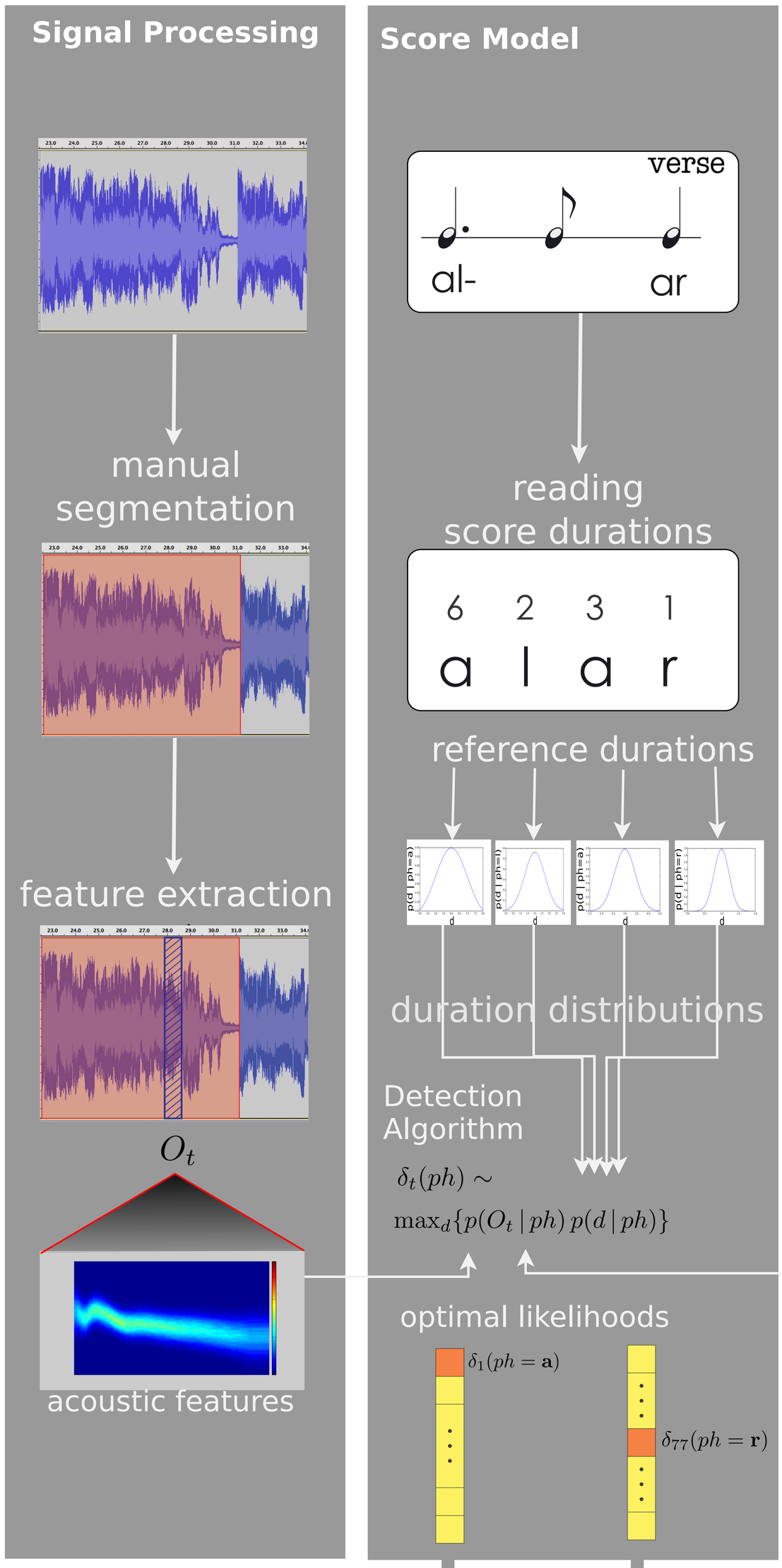
Abstract

- automatic synchronisation of sound and textual lyrics on lyrical phrase level
- extends a speech-to-text alignment approach
- explicitly models sung vocal durations
- evaluated on polyphonic and acapella recordings from classical Turkish music

Motivation

- automatic generation of Karaoke
- enables navigation-by-lyrics playback (available e.g. in Spotify and Deezer)
- lack of reproducible research on the topic
- no work on lyrics-to-audio alignment on non-eurogenetic music

Approach



Dataset

Training Corpus

~500 minutes Turkish speech

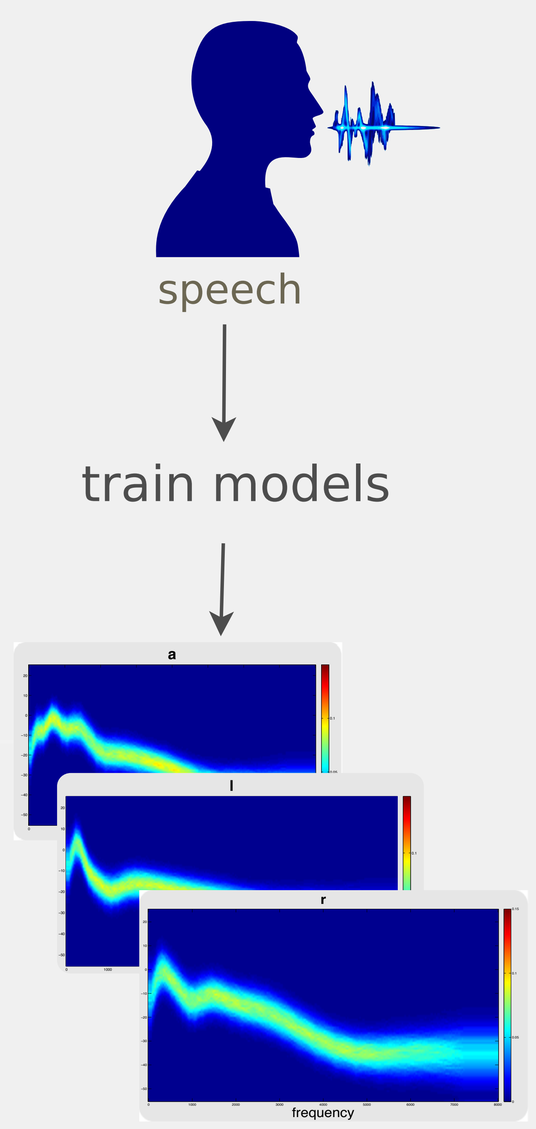
Test Corpus

12 recordings, ~18 mins

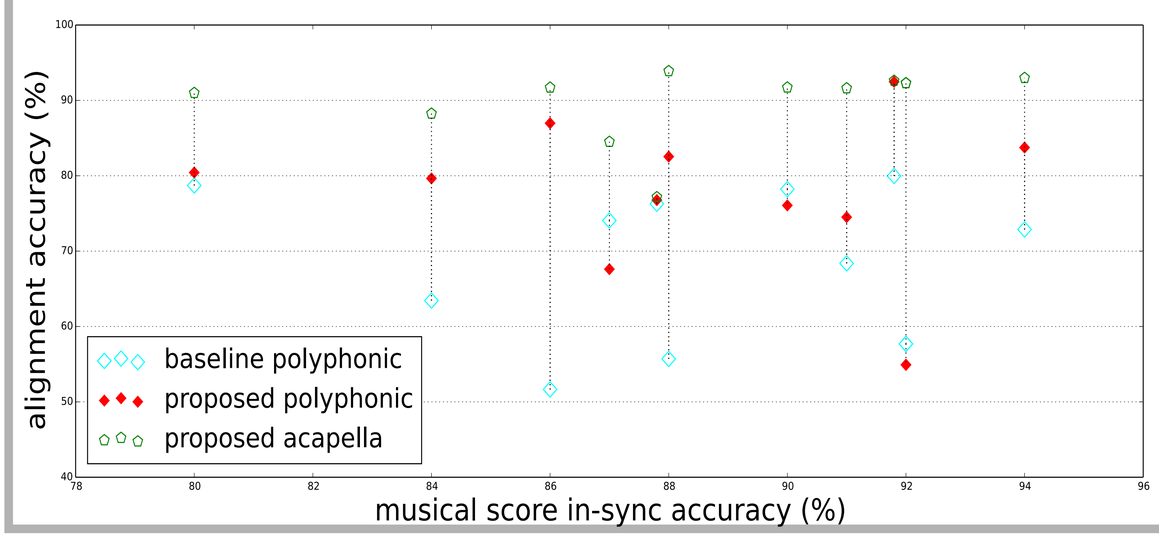
- original from MusicBrainz
- esepecially recorded acapella versions

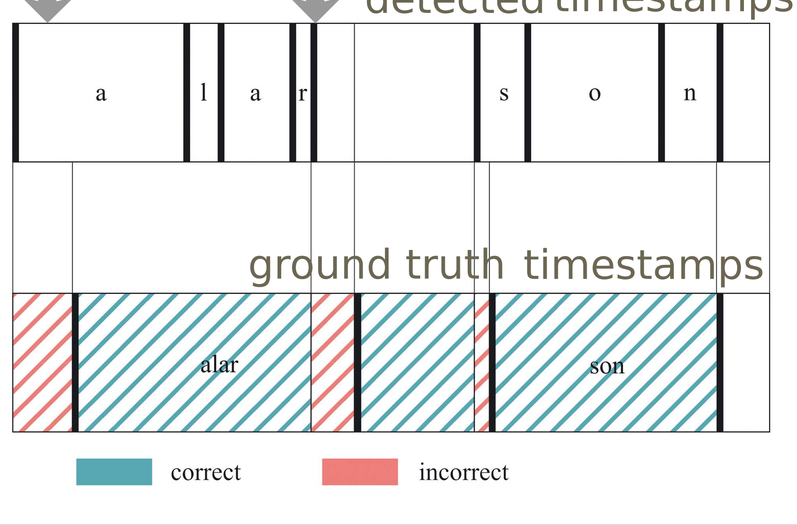
http://dunya.compmusic.upf.edu/makamplayer/
listen to dataset

Acoustic model



Results





https://vimeo.com/121270319
example result

upf. Music Technology Group
compmusic
erc European Research Council Established by the European Commission
upf. Universitat Pompeu Fabra Barcelona
Department of Information and Communication Technologies