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

-
-
- synchronisation of sound and
  rics on lyrical phrase level
  speech-to-text alignment

- models sung vocal durations
  on polyphonic and acapella
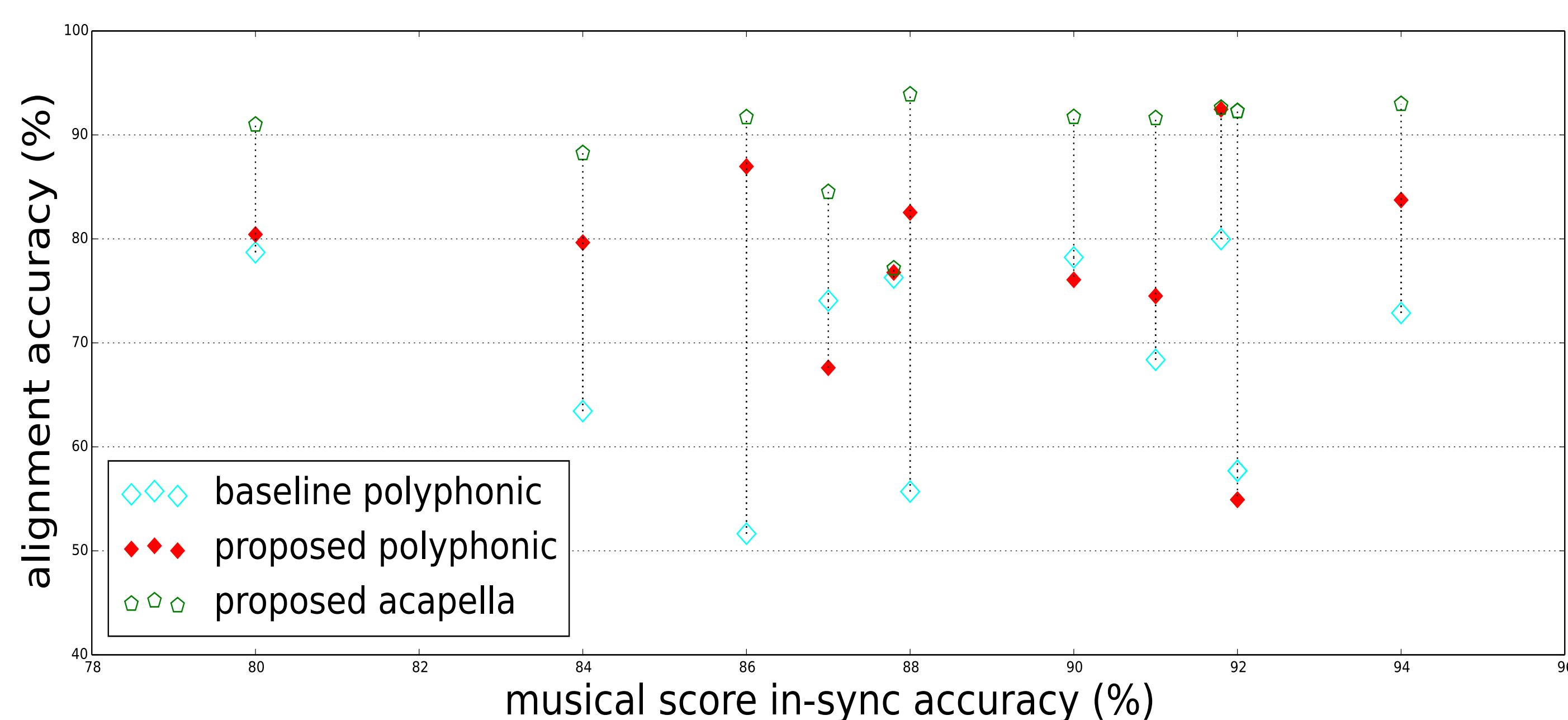 s from classical Turkish music

Motivation

- automatic Karaoke
- playback navigation by lyrics
  currently manual in
  musixMatch and Deezer
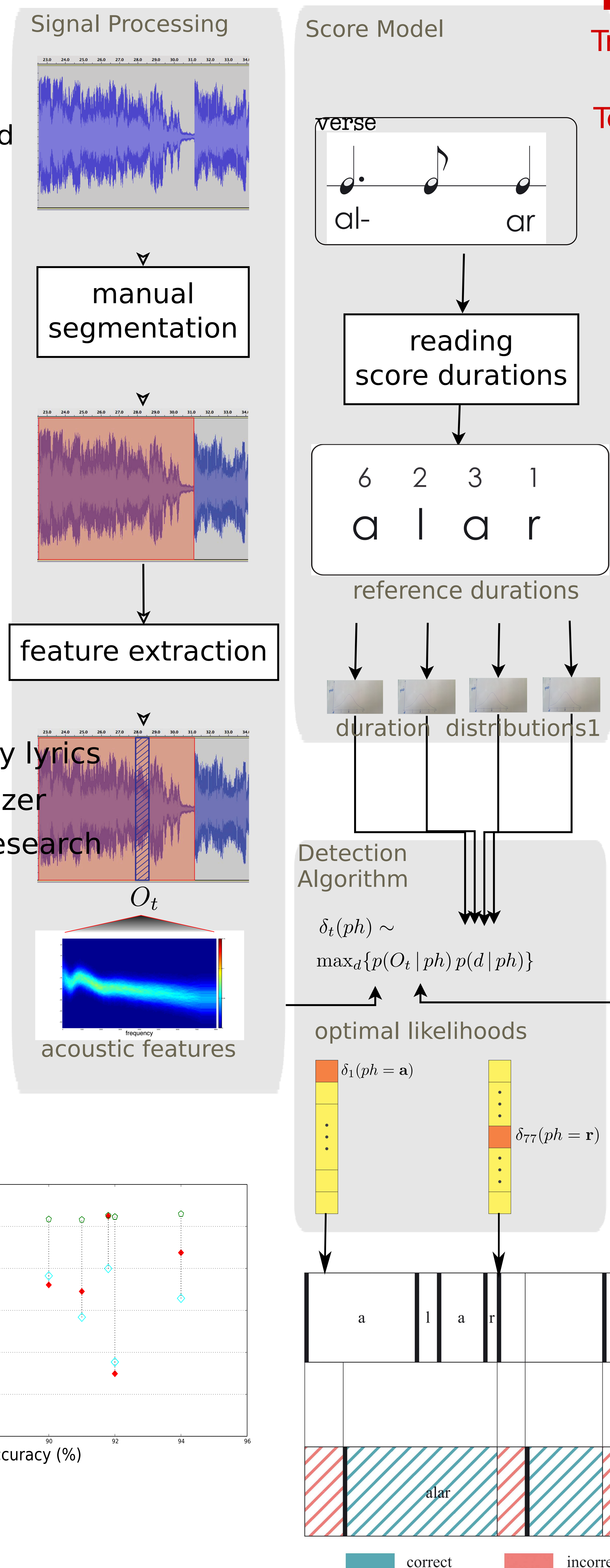- lack of reproducible research
  on this topic

Results

https://vimeo.com/121270319

example result

Approach

Signal Processing

manual segmentation

feature extraction

O_t

acoustic features

Score Model

verse

al-   ar

reading score durations

6 2 3 1
a l a r

reference durations

duration distributions1

Detection Algorithm

$\delta_t(ph) \sim$

$\max_d \{ p(O_t \,|\, ph)\, p(d \,|\, ph) \}$

optimal likelihoods

$\delta_1(ph = \mathbf{a})$

$\delta_{77}(ph = \mathbf{r})$

Acoustic Model

speech

train models

Dataset

a	l	a	r		s	o	n

detected timestamps

alar		son

ground truth timestamps

correct   incorrect

example result

baseline polyphonic	
proposed polyphonic	
proposed acapella	

alignment accuracy (%)

musical score in-sync accuracy (%)