AUTOMATIC LYRICS-TO-AUDIO ALIGNMENT IN CLASSICAL TURKISH MUSIC

Georgi Dzhambazov, Sertan Şentürk, Xavier Serra georgi.dzhambazov@upf.edu

upf. Music Technology Group

compmusic

Abstract

- New work on lyrics-to-audio alignment for classical Turkish music
- Viterbi decoding with phonetic models
- Creation of evaluation dataset
- Evaluation done on phrases of şarkı recordings

Background

şarkı: classical turkish form

- heterophonic nature
- melodic ornamentations: melismas, vibrato

Turkish language

- almost 1-to-1 letter-to-phoneme mapping
- 8 vowels and no diphtongues

Dataset

Training Corpus

~500 minutes turkish speech

Test Corpus

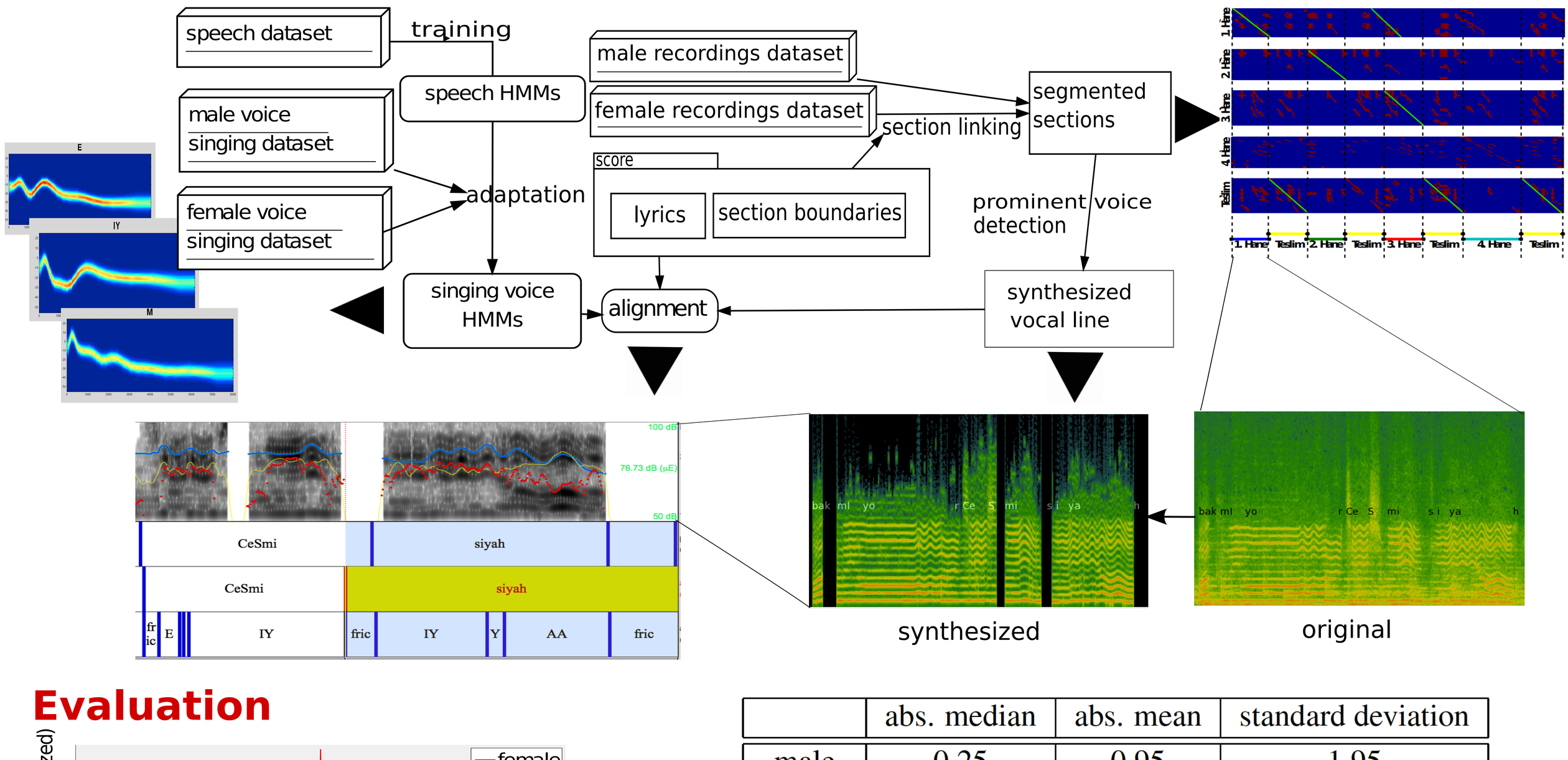
10 şarkı recordings, ~25 mins

	total #sections	section duration	#phrases per section
male	48	8 to 20 seconds	3 to 7
female	41	8 to 35 seconds	3 to 7

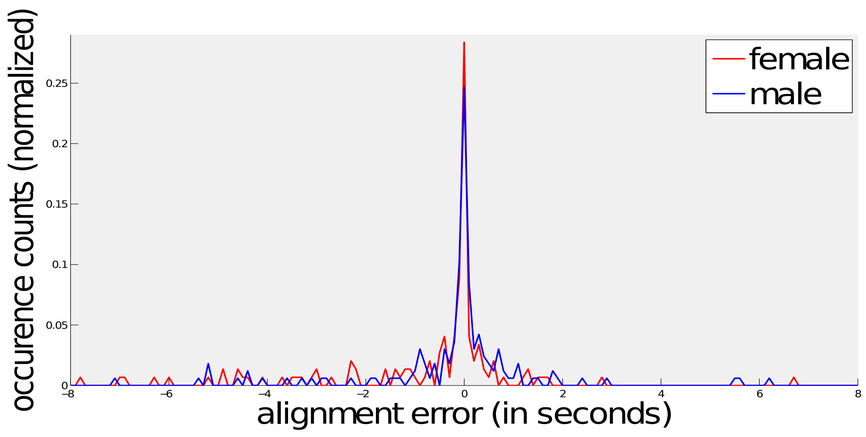
listen to the dataset:

http://dunya.compmusic.upf.edu/makamplayer/

Approach



Evaluation



	abs. median	abs. mean	standard deviation
male	0.25	0.95	1.95
female	0.3	1.61	3.15
total	0.28	1.26	2.63

Reasons for misalignment

- singers omit some consonants
- sections of duration over 20 seconds
- synchronous vibrato and variation in articulation

Conclusions

- alignment erros comparable to state of the art for english language (abs. mean=1.4)
- worse alignment for female due to long sections

Future work

study correlation between vocal descriptors and

- pitch range
- vibrato and melisma