AUDIO-SCORE ALIGNMENT FOR OTTOMAN-TURKISH MAKAM MUSIC

Sertan Şentürk

Universitat Pompeu Fabra – CompMusic Project

sertan.senturk@upf.edu

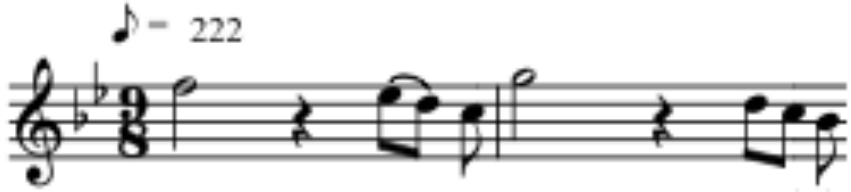
4th International Workshop on Folk Music Analysis

12 Haziran 2014

Music Data Sources

Music can be represented by various data sources

- Audio recordings, scores, videos, lyrics, metadata ...



Outline

Audio-Score Alignment

Methodology

- Tonic Identification
- Section Linking
- Note-Level Alignment

Discussions

- Technical Limitations
- Conceptual Problems

Conclusion

Audio Score Alignment

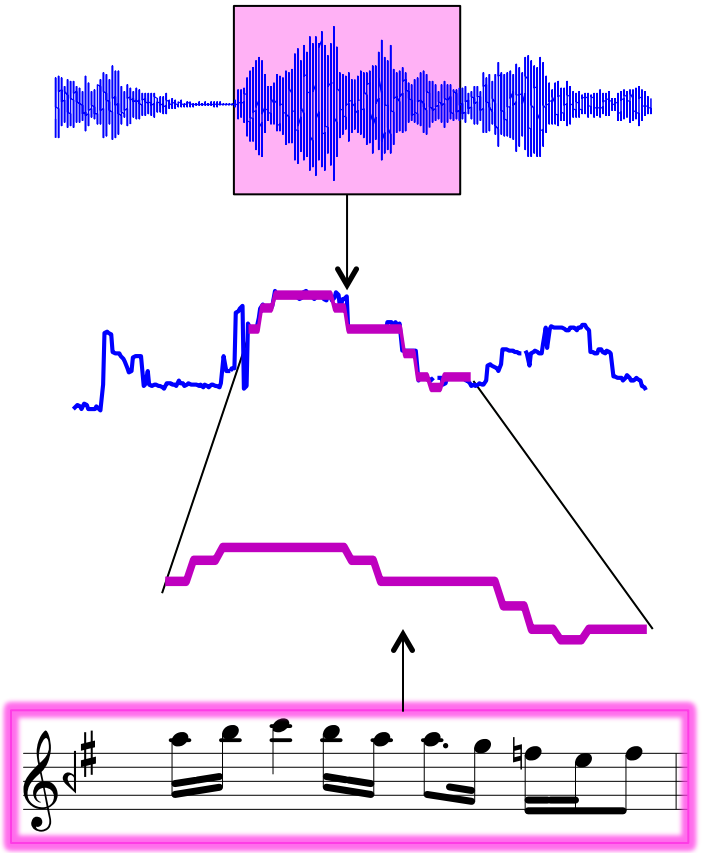
Audio Score Alignment = Synchronisation of the musical events in an audio performance of a music piece with corresponding events in the score of the same piece.

We can take advantage of the complementary aspects aligned audio and scores:

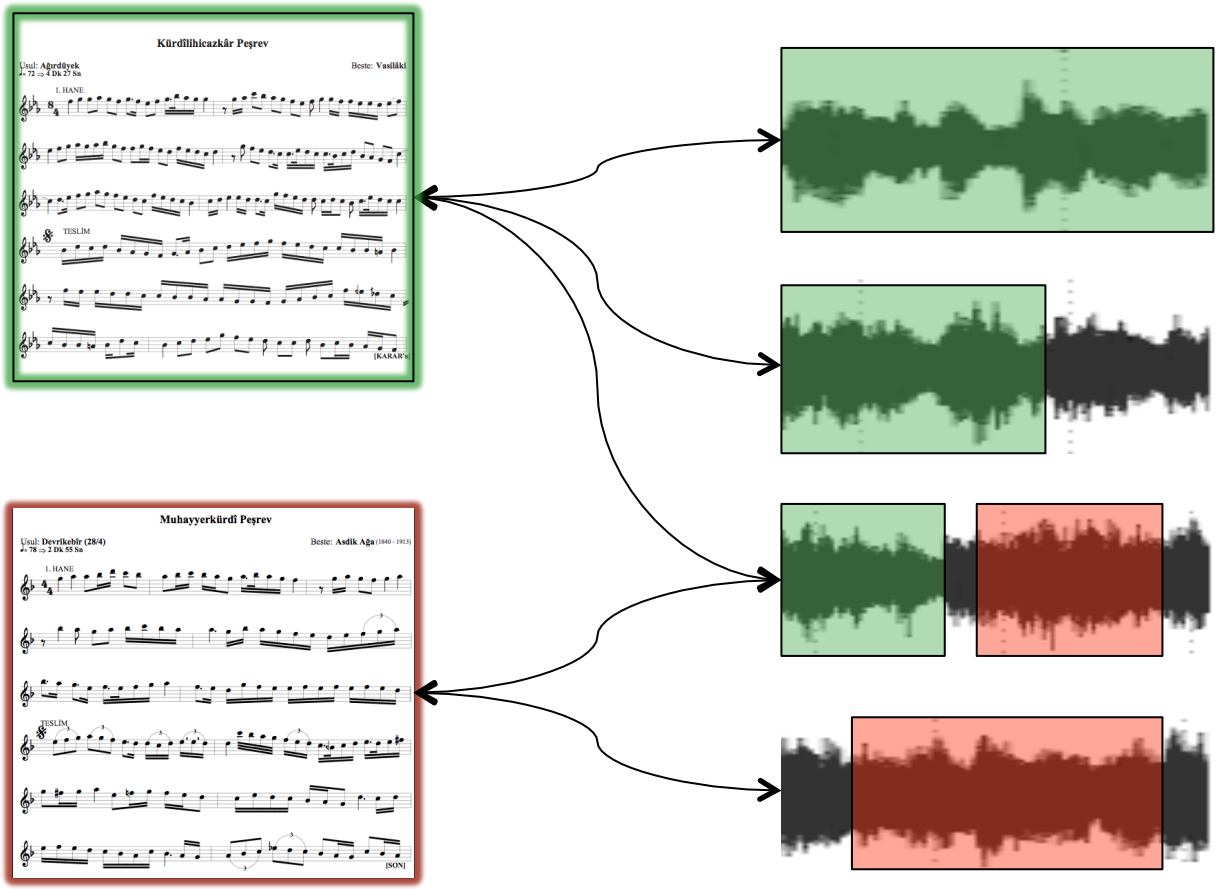
- Version Detection
- Tuning/Intonation Analysis
- Automatic Accompaniment
- Expression Analysis
- Motif Analysis
- Source Separation ...

Fragment Linking

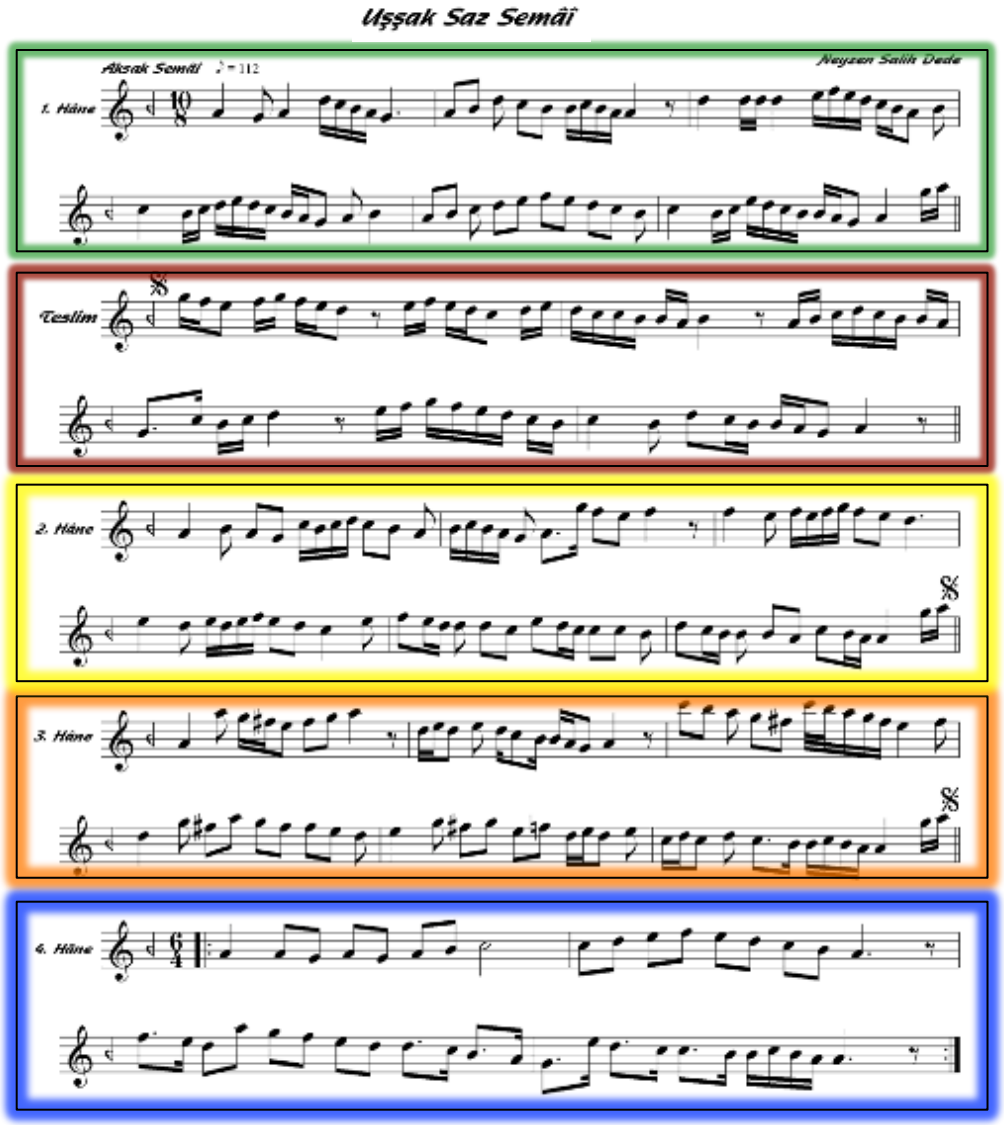
Linking Score and Audio Fragments = Marking the locations of a score fragment within an audio fragment

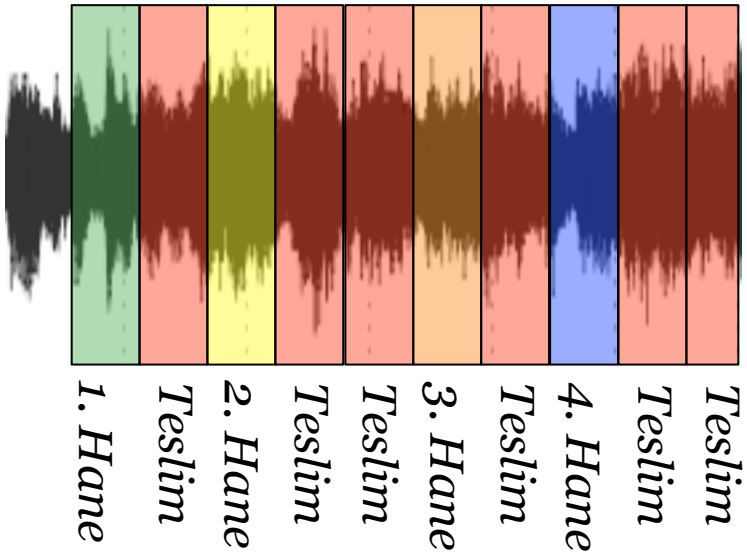


Composition Identification

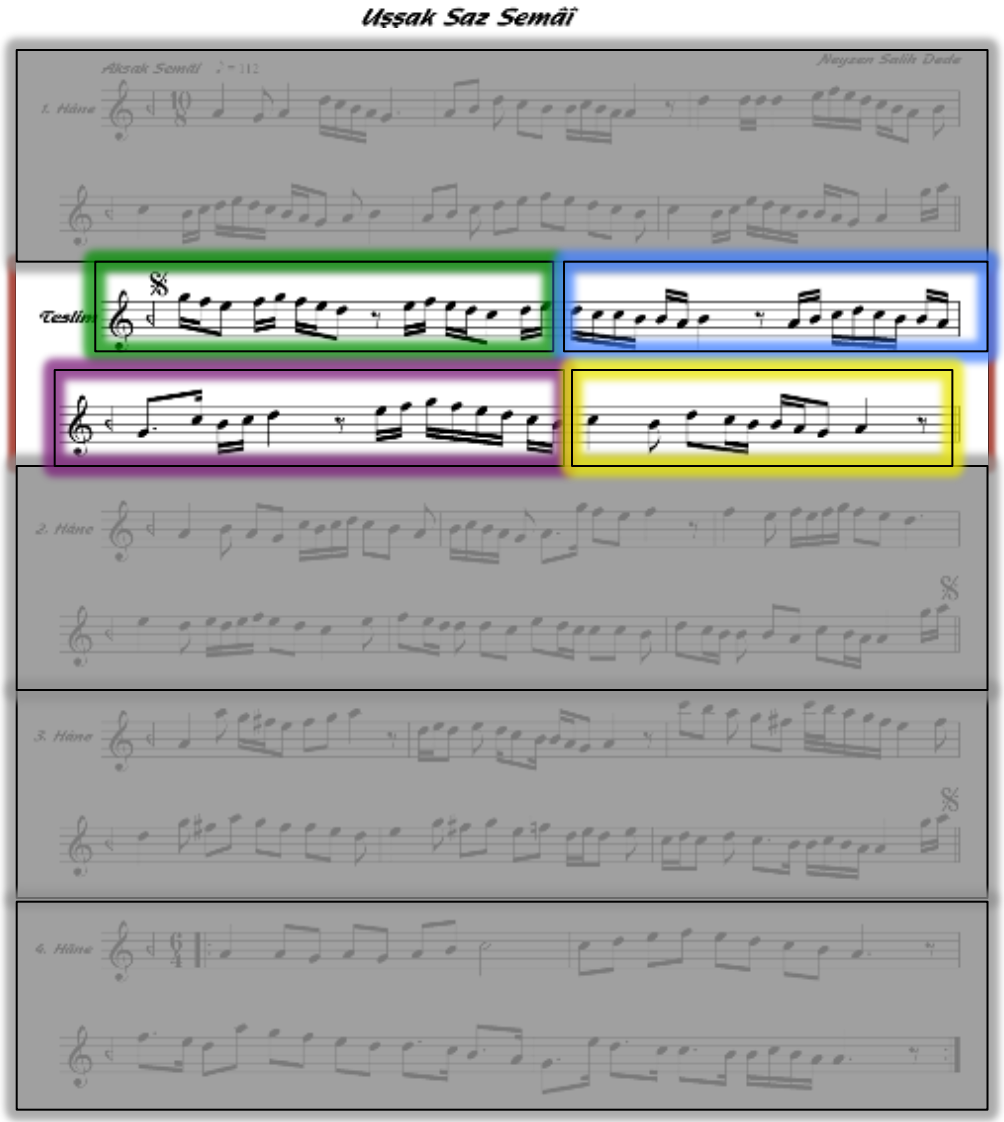


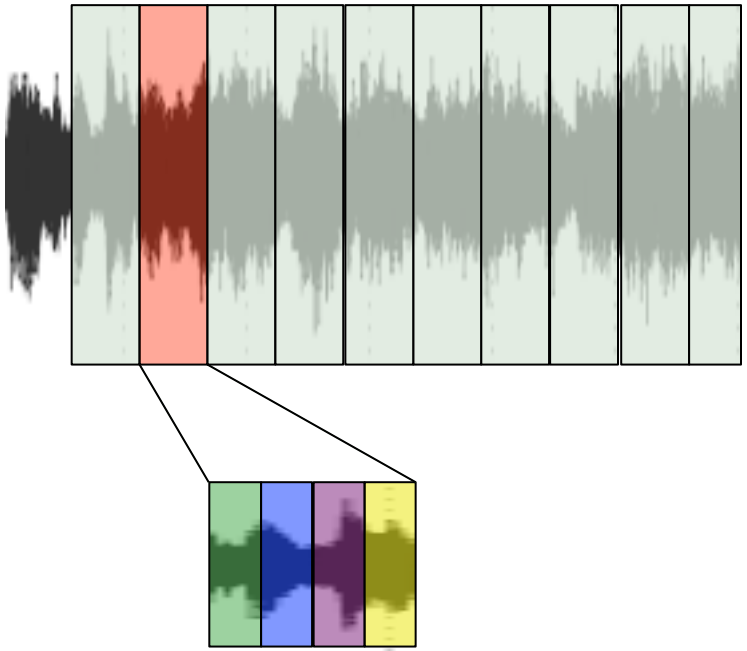
Section Linking



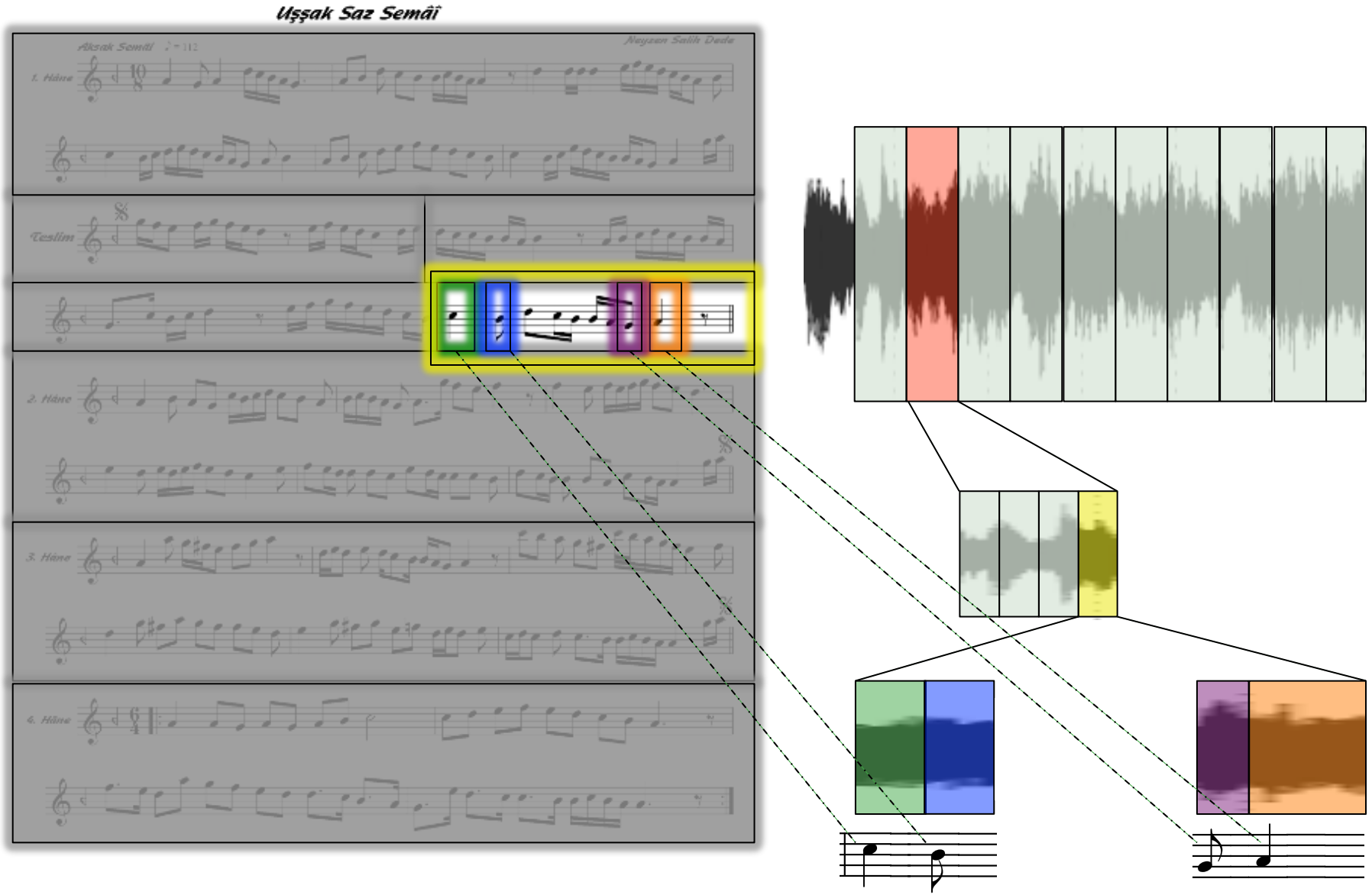


Measure-level Alignment

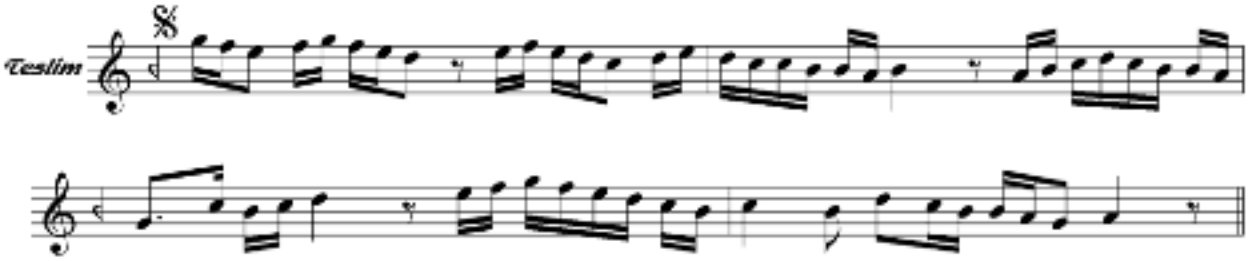




Note-level Alignment



Makam Music of Turkey - Challenges



More than 12 notes in an octave

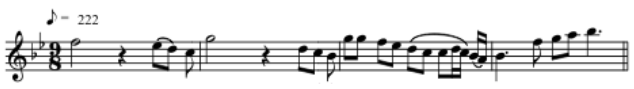
Diverse tuning and intonation

Scores notate simple melody lines

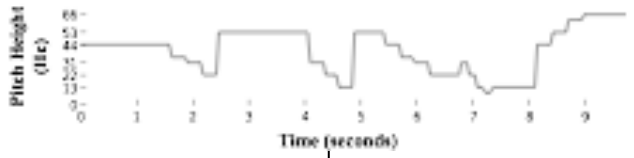
Performers deviate from the score considerably

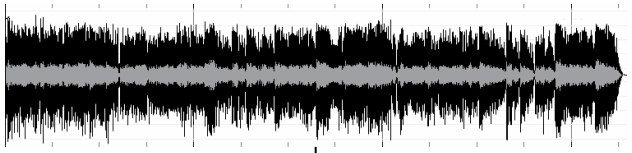
- Non-notated embellishments
- Section repetitions, phrase repetitions, improvisations
- Heterophony

Fragment Linking

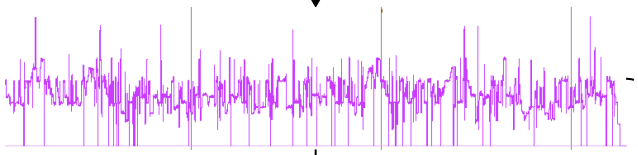


Tonic symbol = 0

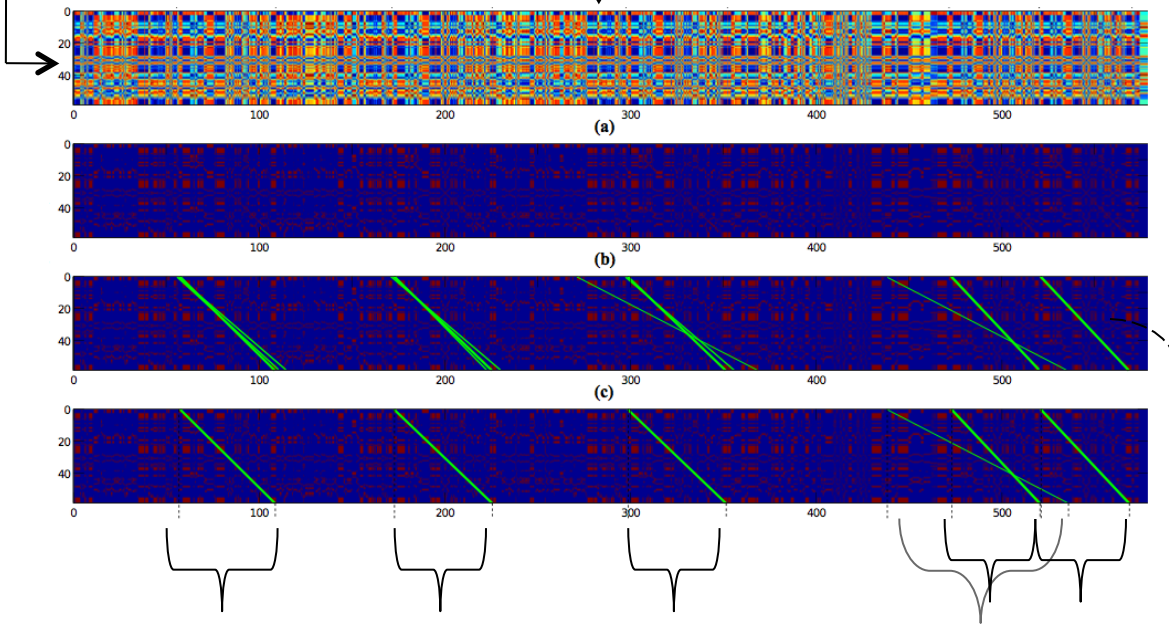




Identified tonic = 0



Essentia (Bogdanov et al., 2013) implementation of **Melodia** (Salamon and Gomez, 2012)



Hough Transform (Duda and Hart, 1972)

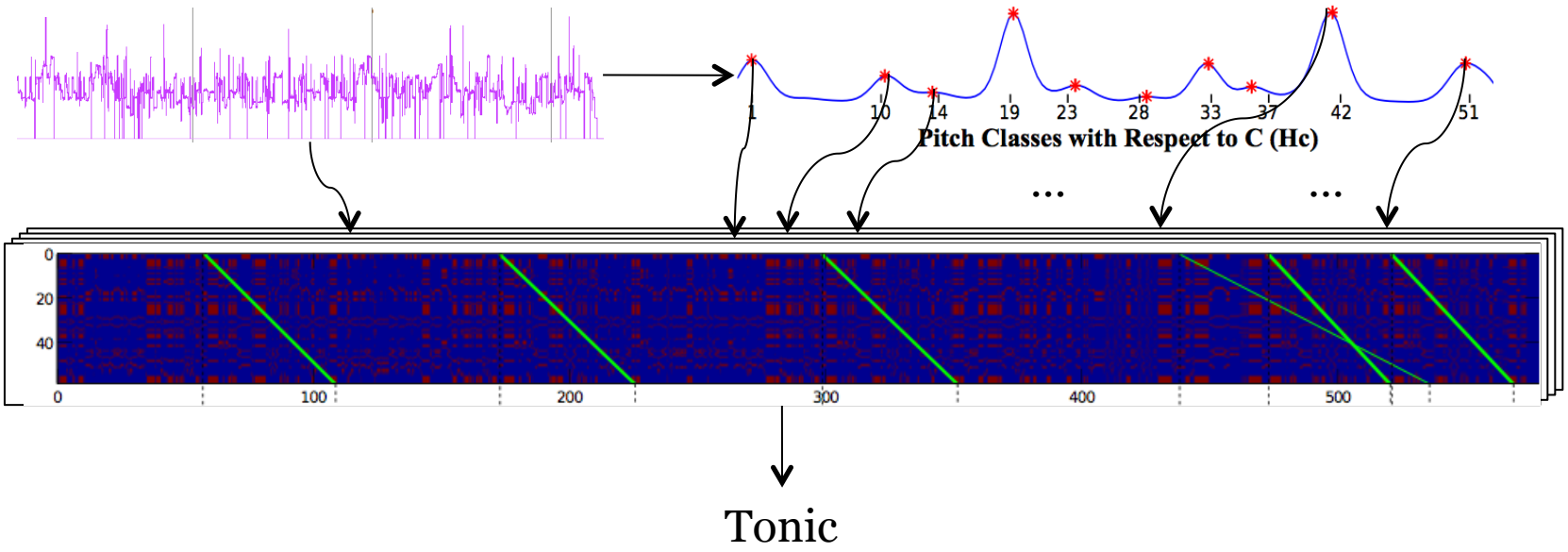
Candidate Locations in Audio

Score Informed Tonic Identification

The tonic is not tuned to a definite frequency.

Number of *ahenk*s (≈ transposition)

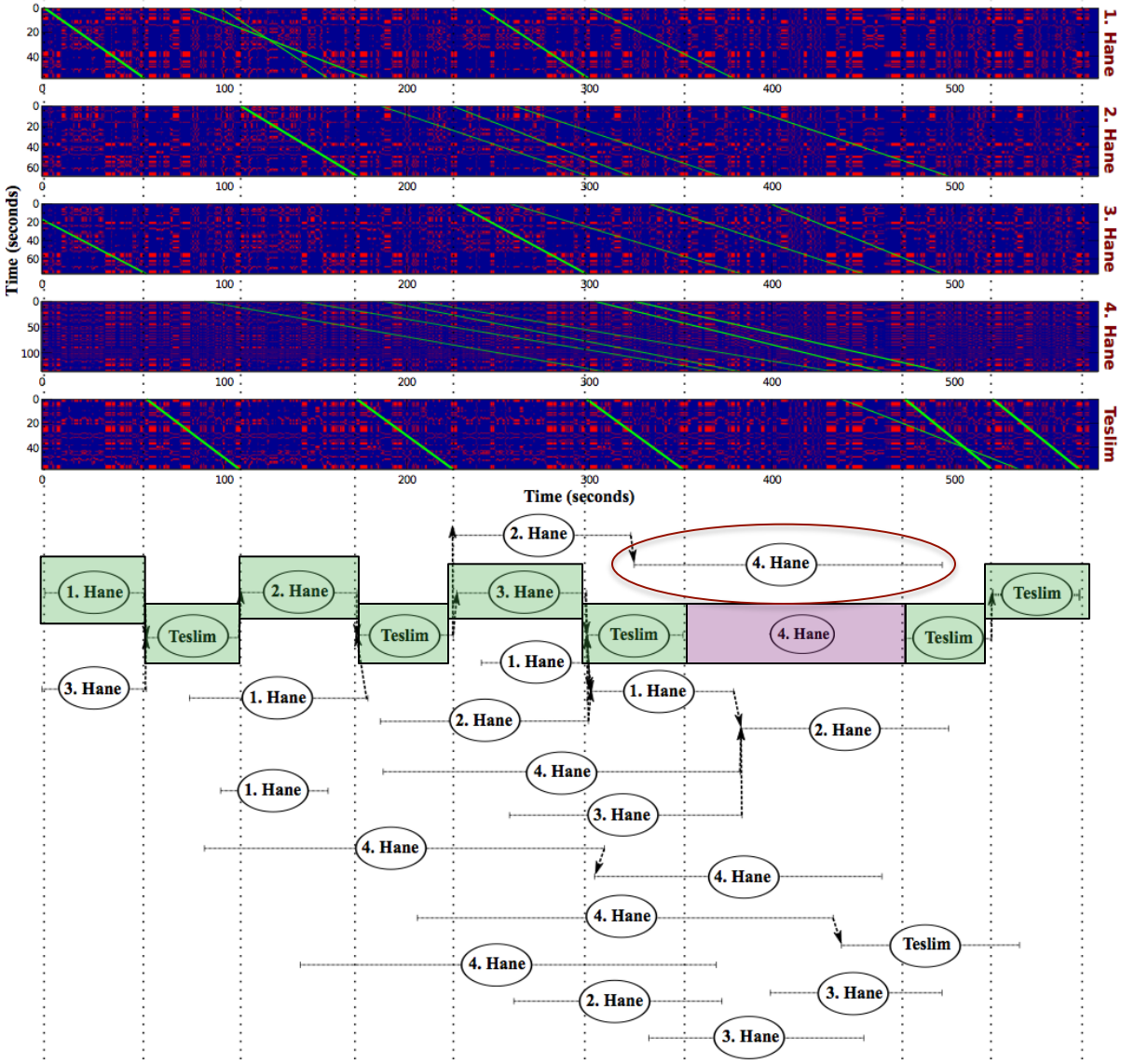
- "Default" transposition: G4 ~ 293Hz (D4)



Section Linking

- Fragments ⟶ Musically relevant structural elements
- The sequence of sections are given in the score
- Section insertions, omissions & repetitions in the performance

Section Linking



Note-level Alignment

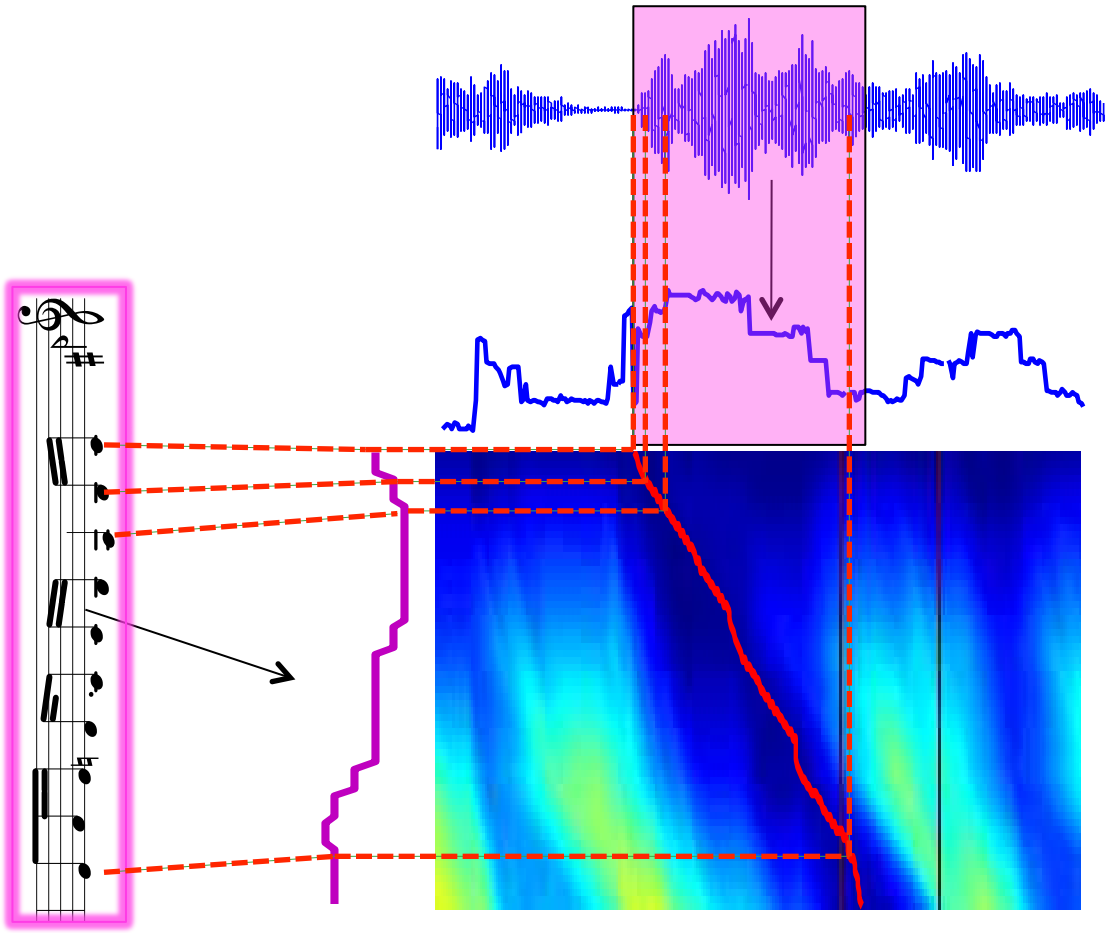
Hough Transform is good at section-level alignment, but not at note-level alignment

- Linear operation ⟶ Cannot model local tempo changes

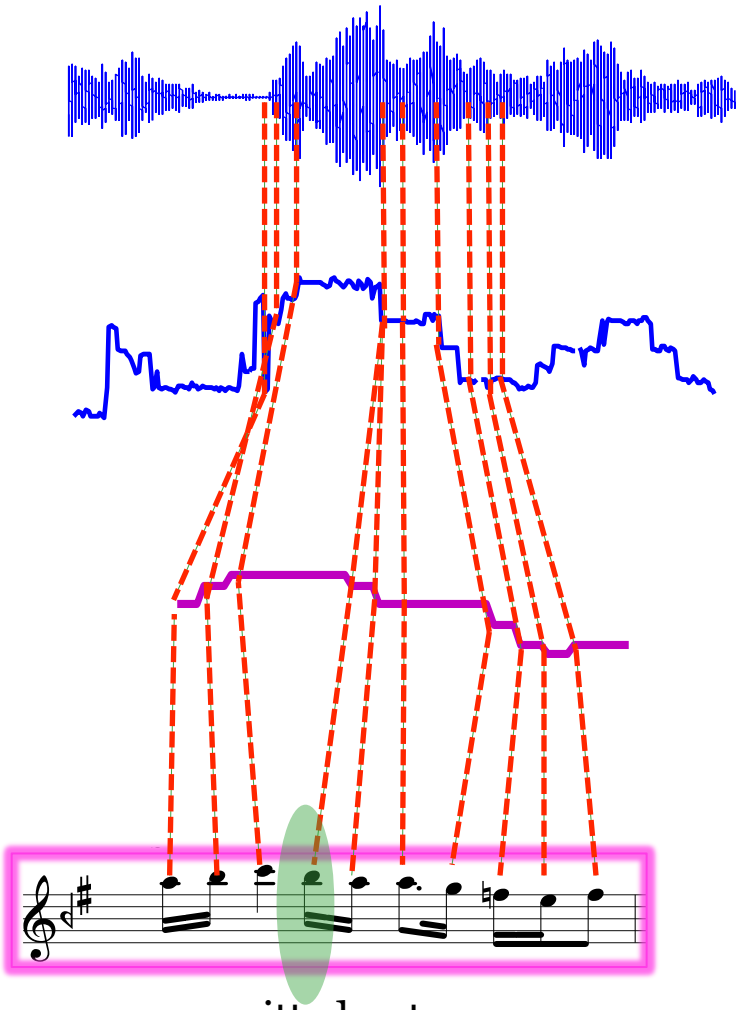
We use DTW to refine the the alignment of the note onsets

- Commonly used for audio-score alignment
- Can take care of local alignment problems
- We use subsequence DTW
  - For Section boundaries

DTW



DTW



Experiments

4 compositions, 6 audio recordings

Scores taken from the SymbTr data collection

The note-level ground truth is from (Benetos & Holzapfel, 2013).

- Transcriptions follow the note sequence in the SymbTr-scores.
- 3896 annotations, 3 insertions, 49 omissions

Evaluation

- A note onset is correct if the time distance between it and the annotation is less than 200ms.

$$P = \frac{t_p}{t_p + f_p}, \quad R = \frac{t_p}{t_p + f_n}, \quad F_1 = 2\,\frac{P\,R}{P + R}$$

Results

100% accuracy in tonic identification

100% F_1 in section linking

66.1% F_1 in note-level alignment

- 2591 notes out of 3896 notes aligned correctly
- 89.2% of the notes are aligned with a margin of ±1 second
- 299 ms average distance, 93 ms median distance and 498 ms standard deviation between the alignment and the annotation

symbTr-score	Audio MBID	Instrumentation	#Anno	t_p	f_p	f_n	$F_1\%$
beyati–pesrev–hafif—-seyfettin_osmanoglu	70a235be-074d-4b9b-8f94-b1860d7be887	ensemble	906	790	116	116	87.2
huseyni–pesrev–muhammes—-lavtaci_andon	8b78115d-f7c1-4eb1-8da0-5edc564f1db3	ensemble	614	482	132	132	78.5
	9442e4cf-0cb3-4cb3-a060-77aa37392501	ney & percussion	302	260	45	42	85.7
rast–pesrev–devrikebir—-giriftzen_asim_bey	31bf3d56-03d8-484e-b63c-ae5ae9a6e733	tanbur	658	374	306	281	56.0
	5c14ad3d-a97a-4e04-99b6-bf27f842f909	ney	673	418	262	255	61.8
segah–pesrev–devrikebir—-yusuf_pasa	e49f33b8-cf8a-4ca9-88cf-9a994dbad1c0	ney & kanun	743	267	490	476	35.6

Technical Limitations

DTW cannot handle portamentos (kaydırma)

- Treats the portamento as an insertion
- Places the note onset after portamento

A trill (or similar embellishments) might cause a note onset to be marked earlier.

The prominent pitch in heterophonic recordings might be harder to track, causing the DTW to lose the track

Since these elements are not (can not be) present in the score representation DTW is expected to fail.

Conceptual Challenges

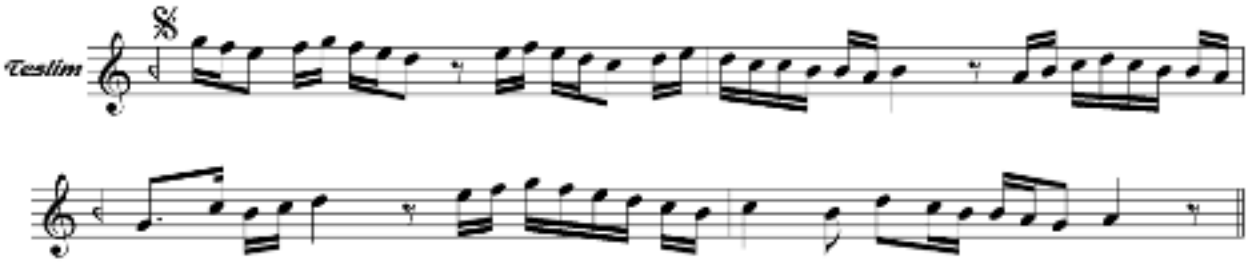
Scores are essentially transcriptions done later

- We are trying to align a transcription to an audio recording
- Do the users need an alignment or a transcription?

Heterophony

- Where should be the note onset considered?

The way how the melodic phrases are written and how it is performed might be substantially different.



Conclusion and More...

Section level alignment is very accurate

- We started experimenting on the phrase level and seem to be as good maybe even better...
- We attempted to use subsequence DTW as a replacement for Hough transform
  - Hough seems to be better than DTW even if it's a simpler method

We have good results from note-level alignment

- There is room to improve
- HMMs might be a good replacement for DTW
- But what to improve
- Need to check with multiple human annotators and do a case study to figure out what the users actually need

Publications

Şentürk, S., Gulati, S., & Serra, X. (submitted). Composition identification in audio-score collections of makam music of Turkey

Şentürk, S., Gulati, S., & Serra, X. (2014). Towards Alignment of Score and Audio Recordings of Ottoman-Turkish Makam Music. In *Proceedings of* the 4th International Workshop on Folk Music Analysis., İstanbul, Turkey (to appear)

Şentürk, S., Holzapfel, A., & Serra, X. (2014). Linking scores and audio recordings in makam music of Turkey. *Journal of New Music Research*, *43*, 34–52.

Şentürk, S., Gulati, S., & Serra, X. (2013). Score informed tonic identification for makam music of Turkey. In *Proceedings of 14th International Society for Music Information Retrieval Conference (ISMIR)*, (pp. 175–180)., Curitiba, Brazil.

Sentürk, S., Holzapfel, A., & Serra, X. (2012). An approach for linking score and audio recordings in Makam music of Turkey. *In Proceedings of the 2nd CompMusic Workshop*; 2012 Jul 12-13; Istanbul, Turkey. Barcelona: Universitat Pompeu Fabra; 2012. p. 95-106.. Universitat Pompeu Fabra.