Composition Identification in Ottoman-Turkish Makam Music Using Transposition-Invariant Partial Audio-Score Alignment

Sertan Şentürk and Xavier Serra

sertan.senturk@upf.edu and xavier.serra@upf.edu

Music Technology Group, UPF, Barcelona, Spain

Motivation

- Automatic audio content analysis & music discovery
- Searching, navigating and accessing large digital music libraries
- A first step in generating linked data between semantic concepts
- May complement "enhanced" listening, music recommendation ...

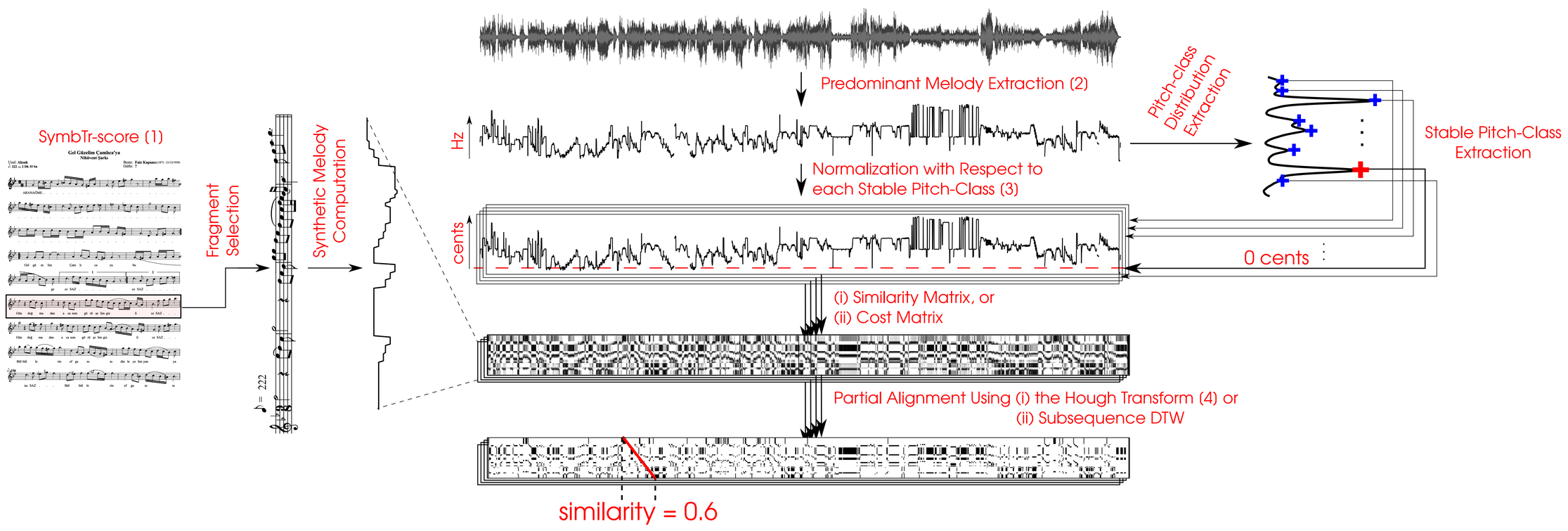
Problem Definition

- **Composition retrieval:** Identification of the composition(s) in an audio recording
- **Performance retrieval:** Identification of the audio recording(s) of a composition

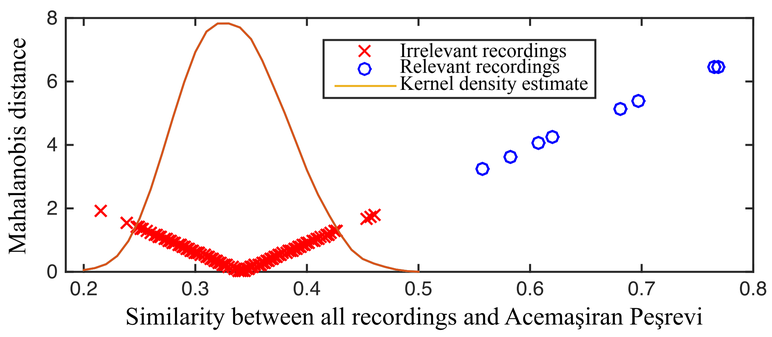
Challenges

- No definite performance tonic frequency (e.g. ~~A4 = 440Hz~~)
- Improvisations within the performances
- Section repetitions, insertions and omissions
- Interpretation differences due to oral propagation
- Heterophony
- Tuning and Intonation
- Embellishments
- Expressive decisions by the performers
- Scores only notate basic, monophonic melodic lines

Transposition Independent Partial Audio-Score Alignment



Document Selection

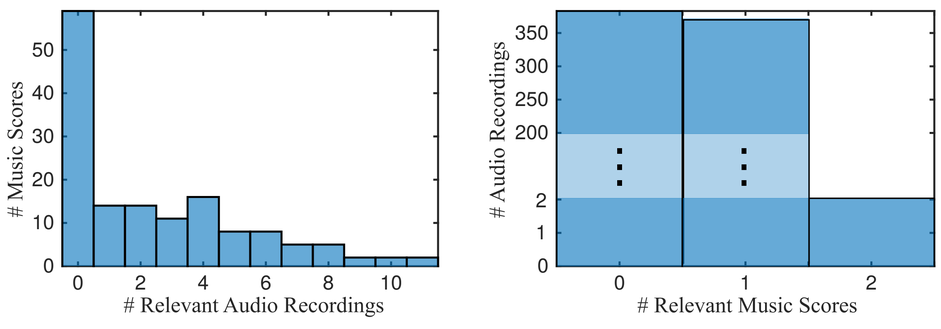


- "Outlier" Detection using Logistic Regression

Experimental Setup

- **4 to 24 seconds** from **the start or the repetition** of the score
- **Hough Transform** or **Subsequence DTW** for partial alignment
- 10 fold cross validation to evaluate document selection

- **Test Dataset:** 743 audio recordings, 146 music scores



Results

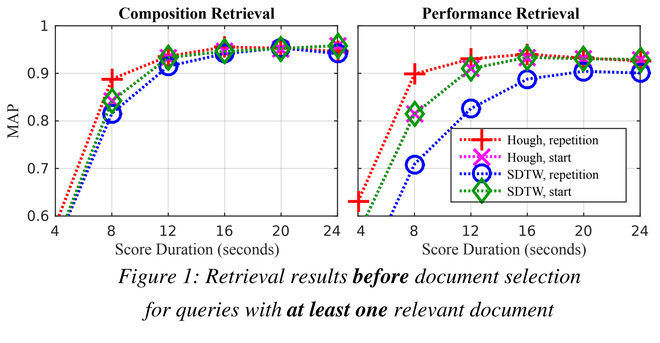


*Figure 1: Retrieval results **before** document selection for queries with **at least one** relevant document*

Methods	Locations	Durations (sec.) 4	8	12	16	20	24
Hough	Start	30	15	2	3	2	2
	Repetition	14	5	0	0	0	0
SDTW	Start	32	6	3	3	3	3
	Repetition	24	3	1	2	3	3

Table 1: Number of Errors in Tonic Identification

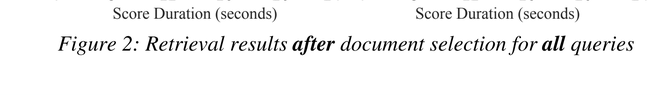


*Figure 2: Retrieval results **after** document selection for **all** queries*

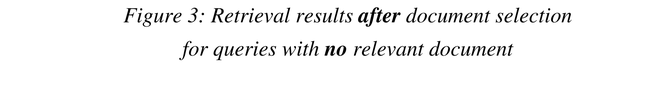


*Figure 3: Retrieval results **after** document selection for queries with **no** relevant document*

Summary

- **~0.95 average MAP** for both composition and performance retrieval
- **Score fragment duration > 8 secs**. Optimal is around **16 secs**.
- Aligning the **start of the score** is sufficient
- Hough transform and SDTW produce similar results; **Hough transform** can be a simpler alternative
- **Document selection is simple and robust** to changes in the score fragment duration & location and the alignment method
- The dataset and the results are **publicly available** via:

  http://compmusic.upf.edu/node/306

[1] Kemal Karaosmanoğlu. A Turkish makam music symbolic database for music information retrieval: SymbTr. In Proceedings of ISMIR, pages 223–228, Porto, Portugal, 2012
[2] Hasan Sercan Atlı, Burak Uyar, Sertan Şentürk, Barış Bozkurt, and Xavier Serra. Audio feature extraction for exploring Turkish makam music. In Proceedings of ATMM, pages 142–153, Ankara, Turkey.
[3] Sertan Şentürk, Sankalp Gulati, and Xavier Serra. Score informed tonic identification for makam music of Turkey. In Proceedings of ISMIR, pages 175–180, Curitiba, Brazil, 2013.
[4] Sertan Şentürk, André Holzapfel, and Xavier Serra. Linking scores and audio recordings in makam music of Turkey. Journal of New Music Research, 43(1):34–52, 2014.

13th Sound and Music Computing Conference
31.08. – 03.09.2016
Hamburg, Germany