Using Dynamic Time Warping to Improve the Classical Music Production Workflow

Smriti Pramanick[1], Eran Egozy[2]
[1]MIT EECS - BS & MEng, [2]MIT Music Technology - Professor of the Practice

Thesis

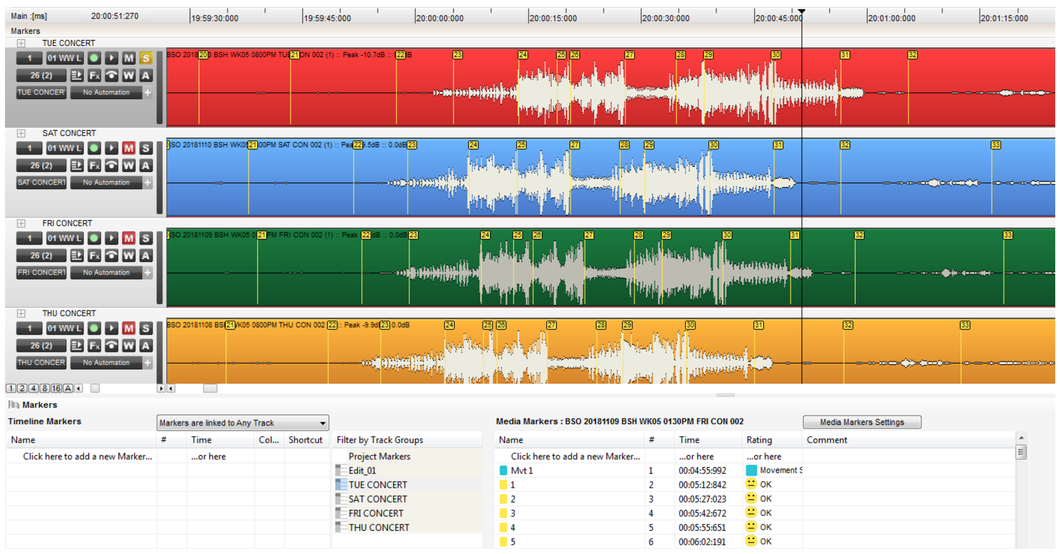
GitHub Repo

Background & Motivation

- Current music production workflow requires a great deal of manual work for the sound engineer
- **Goal: Bring recent advances in Music Information Retrieval (MIR) techniques to music production tools to streamline the process**
- Plan:
  - Investigate the current music production workflow
  - Explore all areas in the workflow that could benefit from digital signal processing and MIR-based tools
  - Build and iterate on these tools
  - Transform the tools into products that are beneficial and easy to use

The Classical Music Production Workflow

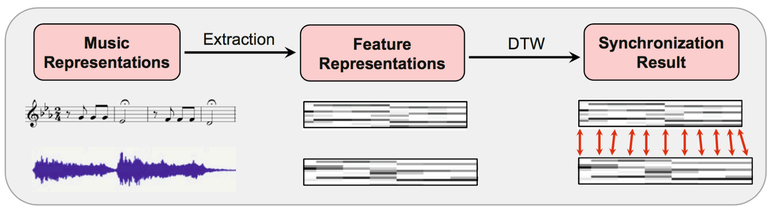
- Comprises recording, editing, mixing, and mastering music
- Tools for producing music:
  - Digital Audio Workstations (DAWs)
  - Virtual Studio Technology (VST) plugins
- **Pain point: identifying the content of several tracks of music while editing and mixing**
  - **No universal solution**
- Related work:
  - Handwritten notes and visual cues: direct markings on the score, looking at waveforms
  - Digital markings: add markings directly to audio recording (in DAW software)
  - Sequoia's Multi-Synchronous Cut: find similar audio in a music project



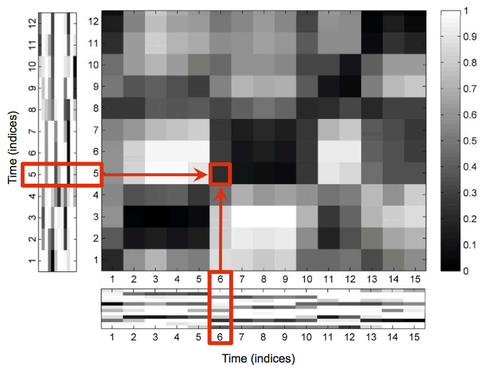
Digital markings in Pyramix DAW

Proposed Solution: Dynamic Time Warping (DTW)

- Reduce manual work by using DTW to help identify relative content of several tracks
- DTW: robust algorithm for synchronizing two recordings of a piece of music

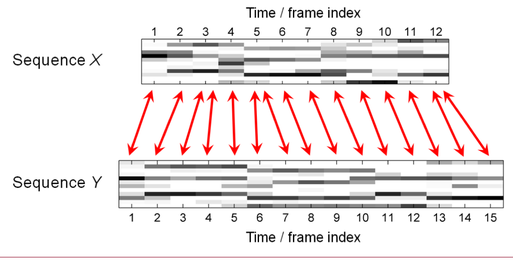


- Goal: find alignment between two sequences (may be nonlinear)
  - Need a cost/similarity metric to compare elements of sequences with each other (we use cosine distance)
  - Calculate cost metric for each pair of elements for two sequences, yields a matrix ("cost matrix")
  - Find alignment with minimal overall cost (optimization problem)

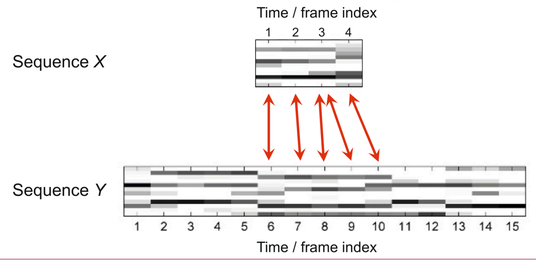


- Applications:

Audio Synchronization:



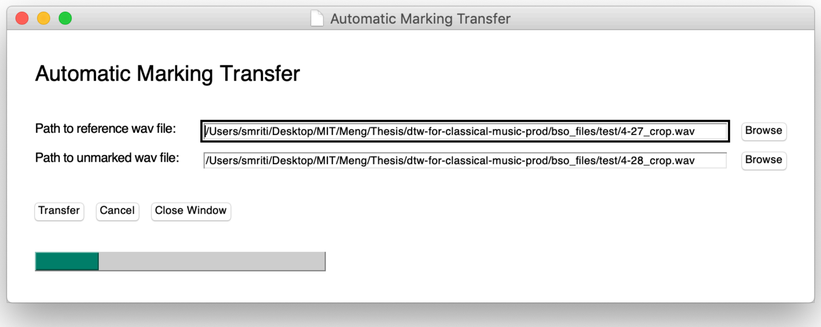
Audio Matching:

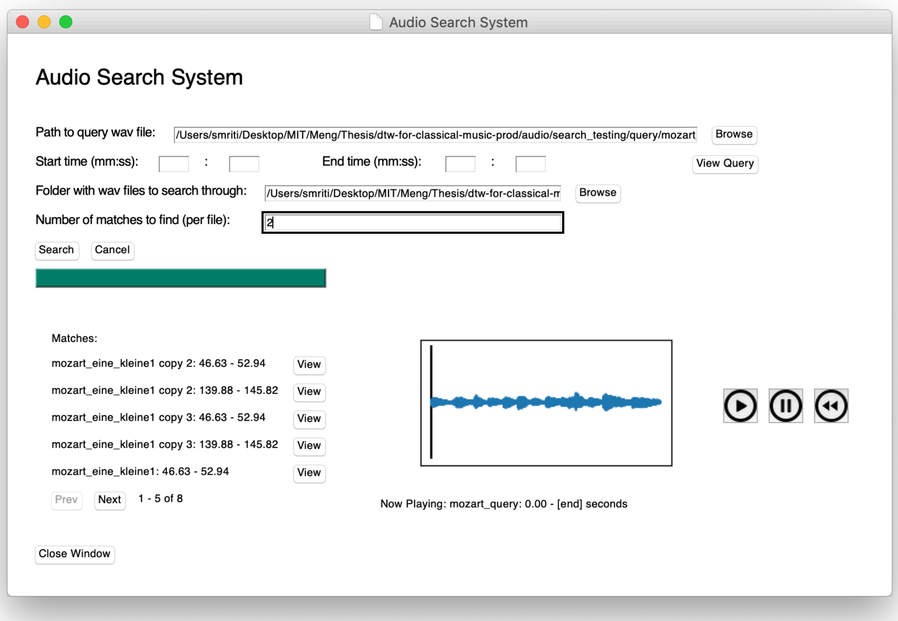


Methodology

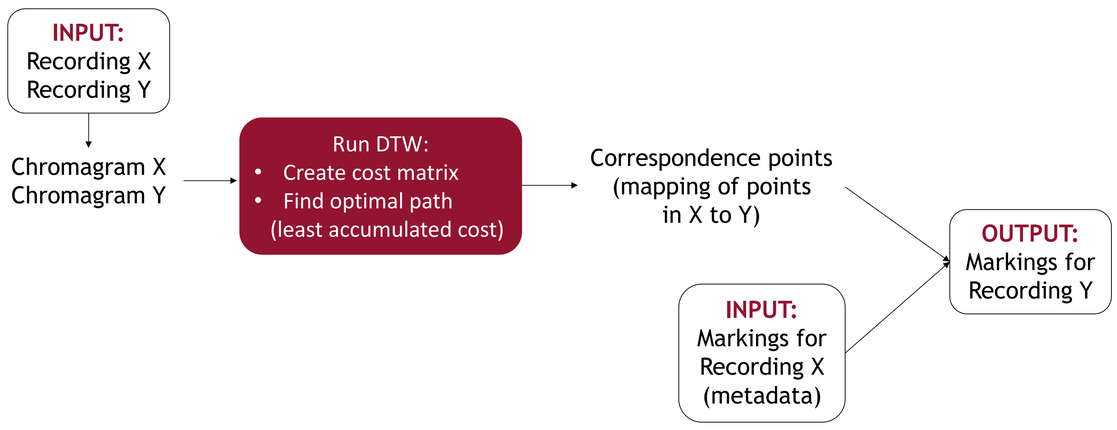
- Collaborate with stakeholders
  - Boston Symphony Orchestra (BSO) sound engineers, to gather requirements
  - Freelance and radio station sound engineers, to identify additional workflows
- Propose, design, implement, and test tools
  - Created successful standalone applications, packaged and distributed as Windows executables

Tools Implemented: Automatic Marking Transfer (AMT) and Audio Search (AS)

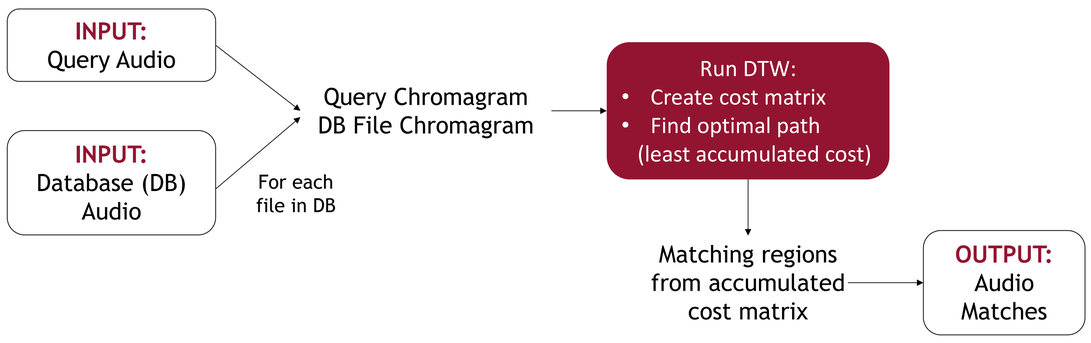




Automatic Marking Transfer (AMT): automatically transfers user-created markings (timed metadata) from one recording of a piece to other recordings of the same piece (which may vary in tempo, timbre, and dynamics)



Audio Search (AS): searches a project for all occurrences of a musical section of a piece that match an audio query snippet (i.e., audio matching)



- Libraries used: PySimpleGUI, LibROSA, NumPy, xml, Pygame, threading, trace, wavfile (scipy.io), PyInstaller
- Code (open source under GNU GPL) at: https://github.com/smritip/dtw-for-classical-music-prod

Evaluation & Results

- Metrics: accuracy, time saved, ease of use and integration into traditional workflow
- Results: initial feedback positive
  - AMT: "performed better than even me and saves time" - BSO sound engineer
  - AS: "extremely useful in speeding up mixing" - radio station sound engineer

Future Directions

- Further Evaluation and Testing
- Integration of AMT & AS into DAWs
- Additional features: rewrite in C++ and use JUCE, Mac applications, richer GUIs, …
- Other applications: genres other than classical music, speech and video, …
- More potential tools: MIR algs. beyond DTW for beat tracking, chord detection, …

Key References:
[1] Muller, Meinard. Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications. Springer, 2015.
[2] Pramanick, Smriti. "Using Dynamic Time Warping to Improve the Classical Music Production Workflow." *MIT Music Technology Lab*, Aug. 2019, musictech.mit.edu/sites/default/files/documents/pramanick_meng.pdf.