



Audio Analysis with DTW

Hashim Hashimov

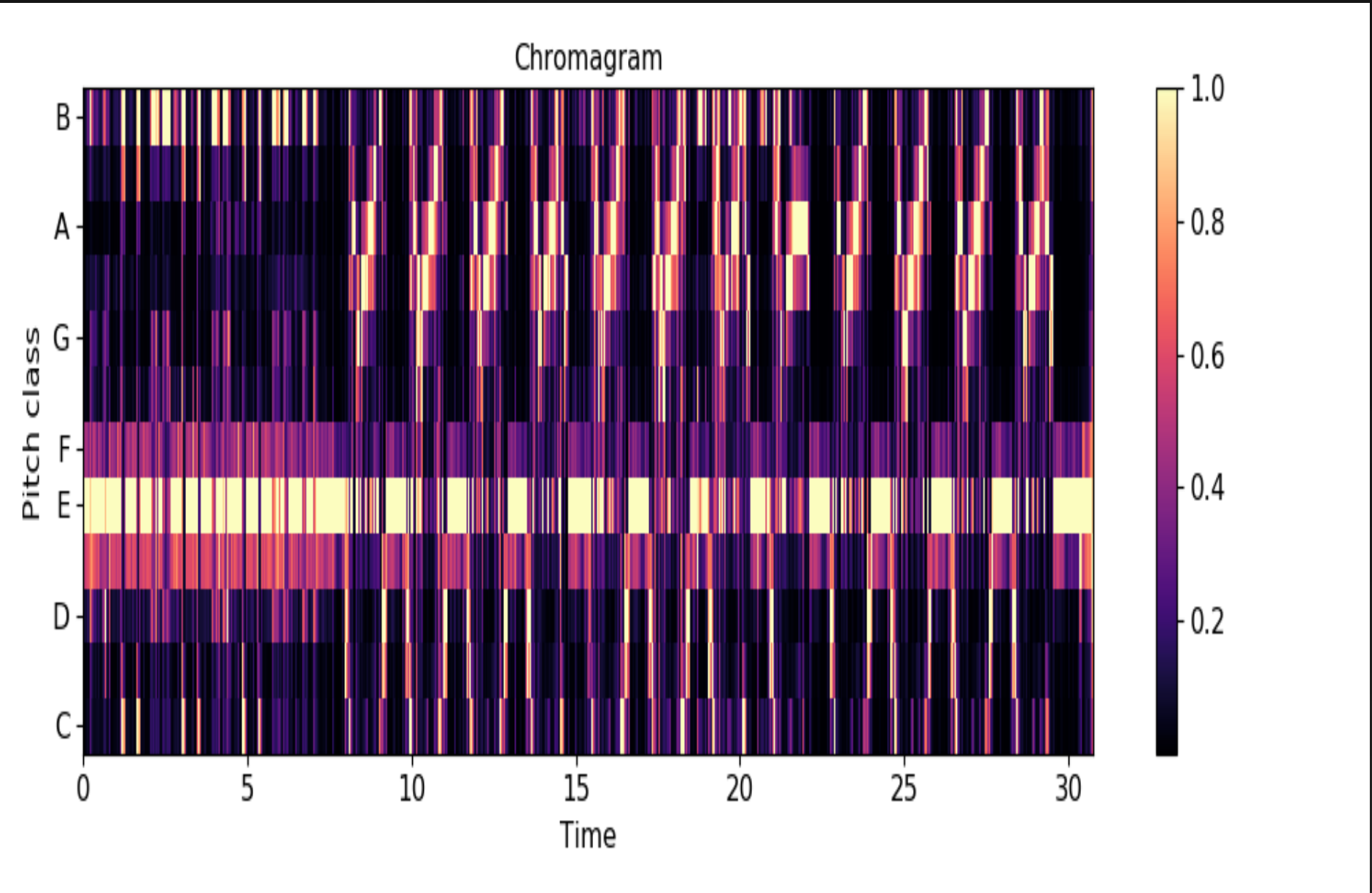
Kayahan Kaya

University of Tartu



Background

Dynamic time warping is an algorithm used to measure similarity between two sequences which may vary in time or speed. **DTW** is useful in many domains such as speech recognition, comparing two audio signals, It's also commonly **used** in data mining **to** measure the distance between two **time-series**.



Extracting Music Features

Before using the audio files in DTW algorithm we extract chroma features of audio file

What is chroma?

Chroma-based features, which are also referred to as “pitch(note) class profiles”, are a powerful tool for analyzing music whose pitches can be meaningfully categorized (often into twelve categories).One main property of chroma features is that they capture harmonic and melodic characteristics of music.

Dataset and Library

We created dataset which contains 496 mp3 files from different genres such as Alternate, Blues, Electronic, Jazz, Rock, Rap, funk soul r&b. Additionally we have pairs of audios which has different speed pace.

We used **librosa** library to extract music features for **DTW** algorihm.

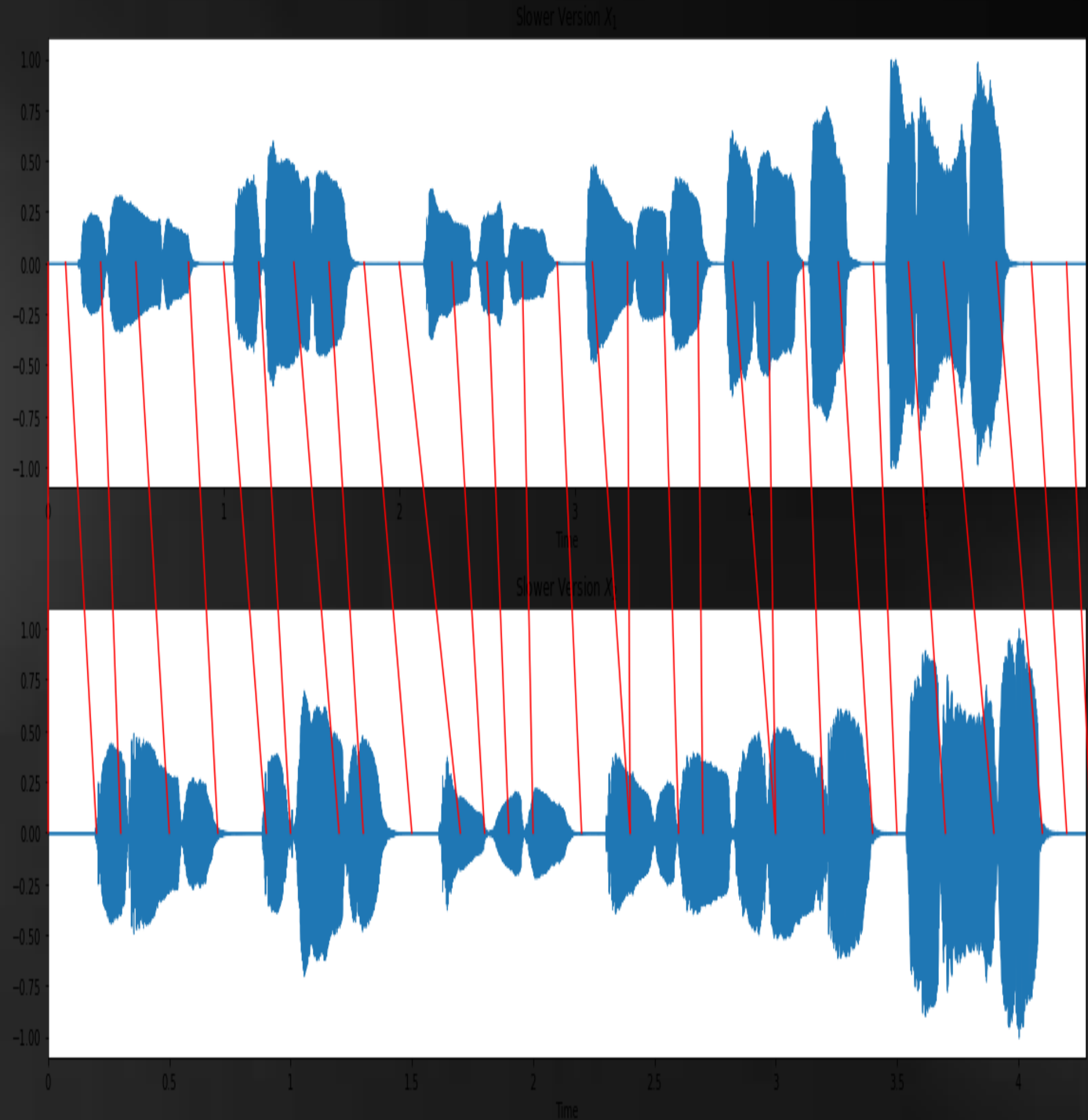
Purpose

We have two main goals for this project;

- Our first aim is to detect similarities between two audio signals by using dynamic time warping.
- Second aim is to test our algorithm with different audio files to analyze the given result and see how accurate it works?

Experimental Result

The picture on the right is the output of our algorithm. As we can clearly see on the graph, Although These signals have different time length our algorithm detects similarities between two signals which is demonstrated by red line.



Conclusion

In the end we tested our algorithm around 500 audio files and the audio files which has different speed pace ;

Here is the results;

- Our algorithm returns distance value **0**, if the audios are completely same .
- Returns distance value between **(10,15)**, if the audios are the same but in different speed .
- Returns distance value greater than **30**, if the two audios are completely different than each other.

STAY POSITIVE, DISCOVER
MUSIC...