EE 624 Speech Technology
Report for Project 1
Dynamic Time Warping

By: Kota Yamini Suguna
Roll number: 194102309
M. Tech Signal Processing and Machine Learning

Objective:

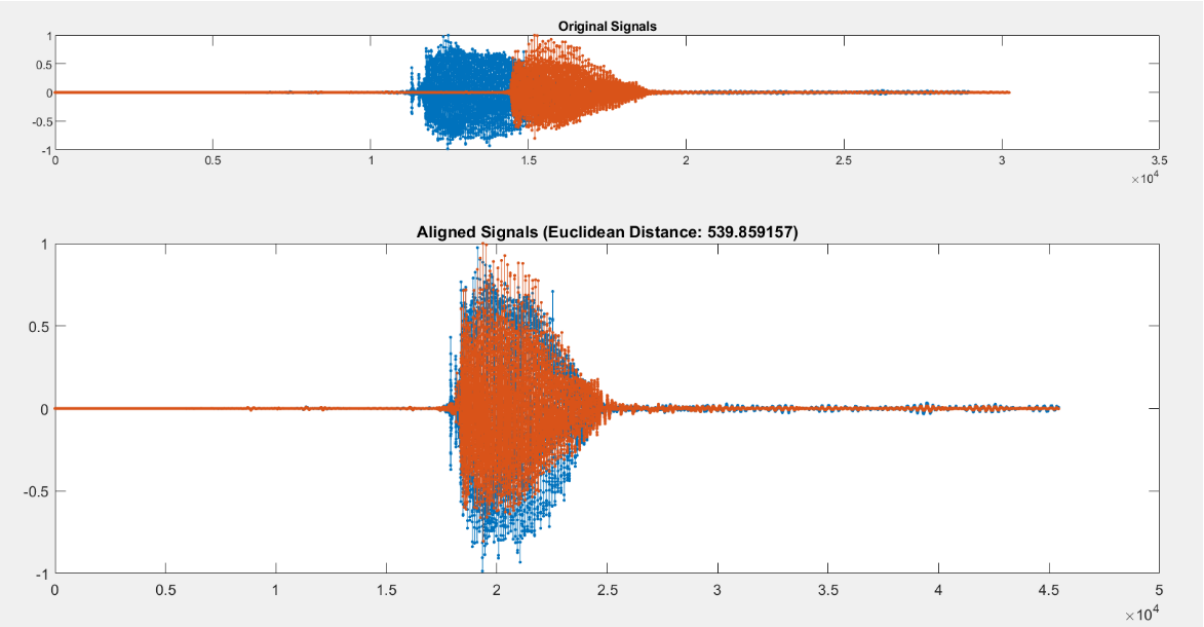
Understand how alignment is done between the audio sequences using Dynamic Time Warping.

Observations:

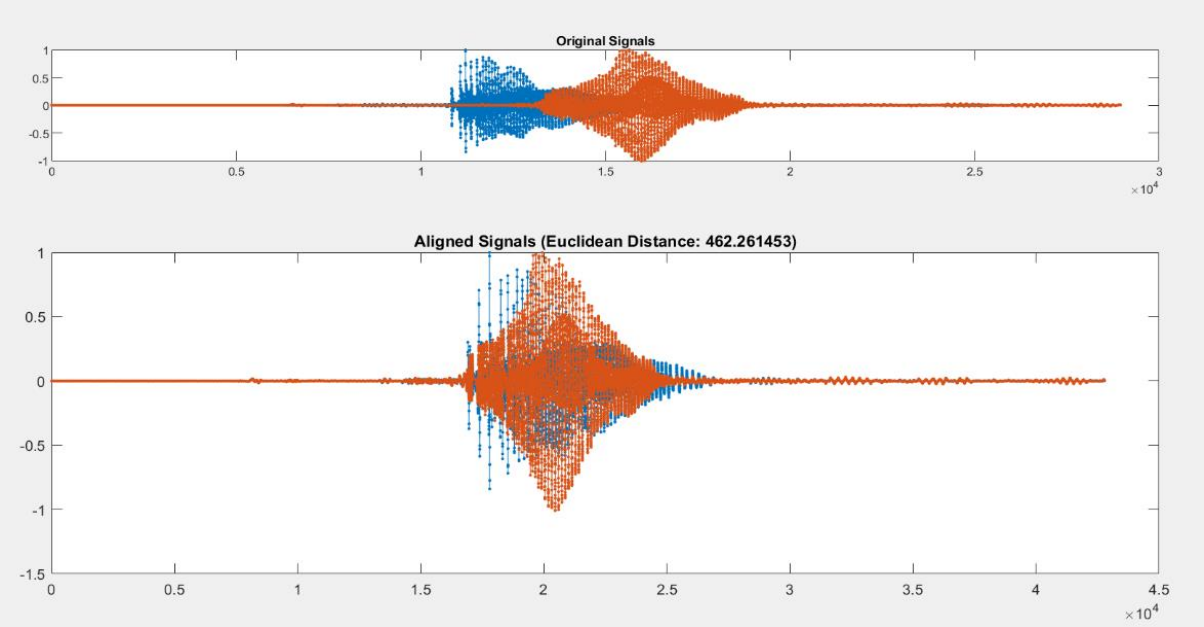
- The Euclidean distances obtained for similar sounding audios (for e.g. same vowels) are comparatively lower (typically less than 500).
- The Euclidean distances obtained for different sounding audios (for e.g. different vowels) are comparatively higher (typically more than 500).
- DTW alignment path plot for similar sounding audios is tending to a straight line.
- Whereas, DTW alignment path for the dissimilar sounding audios has more granularity (more ups and straight paths).
- The basic shape or outline for the normalized audios tends to be same for similar sounding audios.
- The basic shape or outline for the normalized audios of dissimilar sounding audios is visibly different.

Figures:

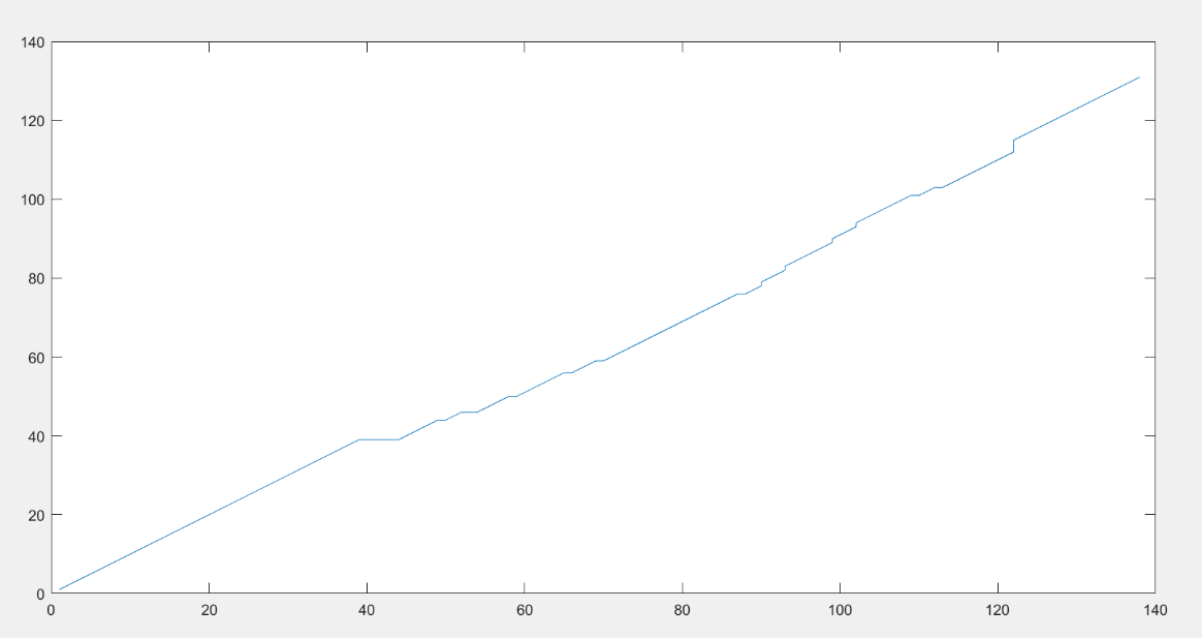
Typical Audio Alignment for Similar Sounds:



Typical Audio Alignment for Dissimilar Sounds:



Typical DTW Alignment Path Plot for Similar Sounds:



Typical DTW Alignment Path Plot for Dissimilar Sounds: