THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

SIPL
Signal and Image Processing Lab

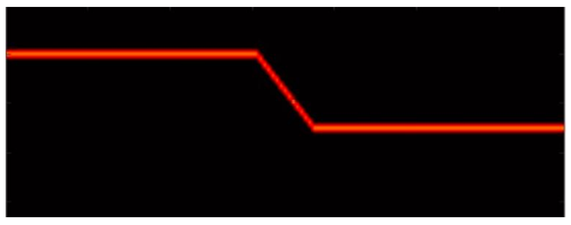
TECHNION Israel Institute of Technology

Music Manipulation using optimal transport

Roy Lazoc and Yair Amar, Supervised by Ori Kats

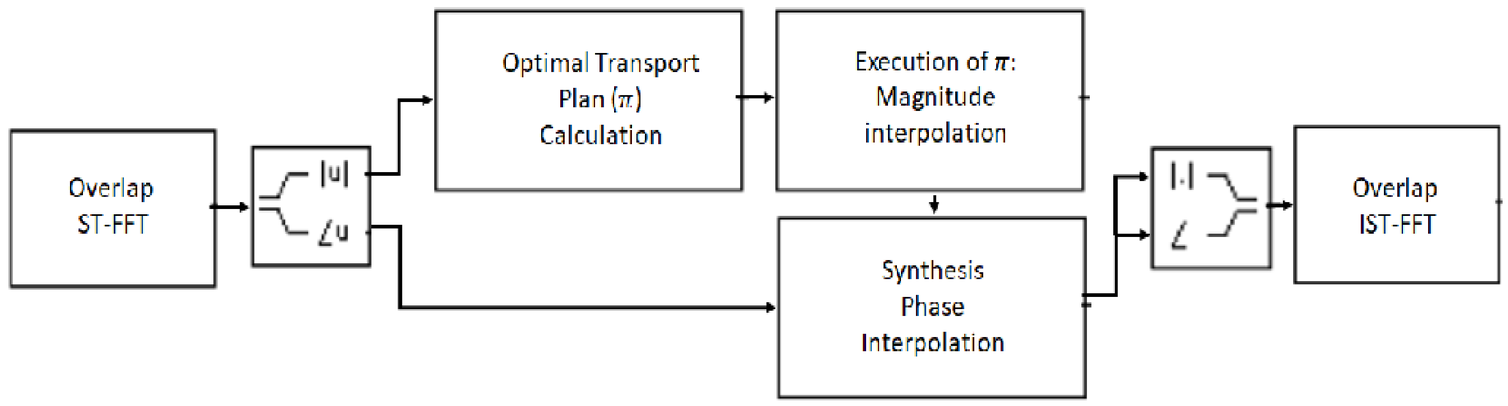
Introduction

- Portamento is a musical term which means pitch – sliding from one note to another.
- Portamento is achievable only with some musical instruments, such as guitar and flute.



Spectrogram of a portamento effect

Goal

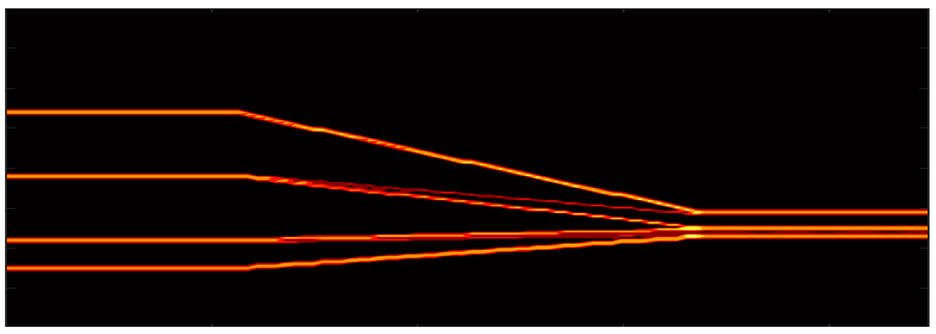


Project's Block Diagram

- Generalization of the Portamento effect for complex signals
- Calculating Optimal Transport plan, and applying on the signals in the freq. domain
- Preserving the audio fidelity over this transformation

Optimal Transport

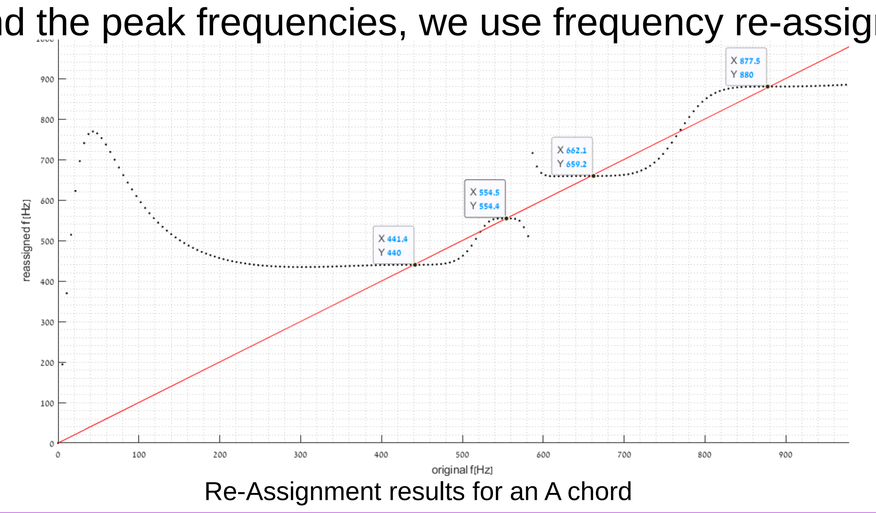
With the OT plan, we transfer mass from one bin of the spectrogram to another, and by that we achieve the desired portamento effect.
The OT plan tells us the amount of energy we need to transfer from each bin in the source spectrogram, to each bin in the target spectrogram.



Optimal Transport between

Frequency Re-Assignment

The OT plan is applied on the magnitude of the spectrogram.
Due to the time-frequency resolution trade-off, we need to calculate manually the phase in each bin & in each time-frame, in order to fix audio artifacts, known as vertical & horizontal incoherence.
In order to find the peak frequencies, we use frequency re-assignment



Re-Assignment results for an A chord

Phase Accumulation

Using the re-assigned frequency, we can find the phase for each frequency bin i , at time-frame t, using the following formula:

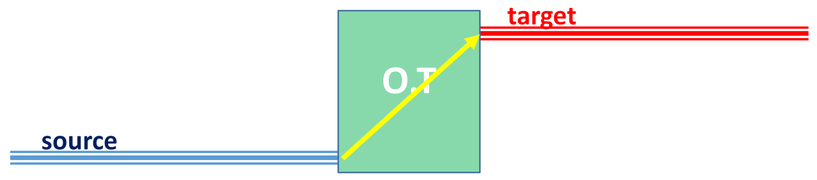
$$\phi_i^t = \phi_i^{t-1} + \hat{\omega}_i{}^{t-1} \cdot \Delta t$$

Where $\hat{\omega}_i{}^{t-1}$ is the reassigned frequency

Interpolation

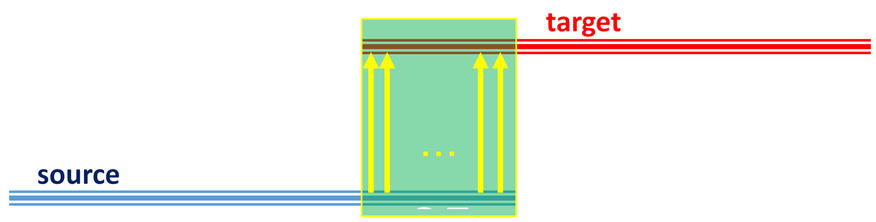
- Original approach – placing each mass π_{ij} in the frequency

$$\omega_{ij} = (1-k) \cdot \hat{\omega}_i^{source}(t_{end}) + k \cdot \hat{\omega}_j^{target}(t_{start})$$



- New approach – at each frame (t_k) calculate new OT plan, between $source(t_k)$, $target(t_k)$. placing each mass π_{ij} in the frequency :

$$\omega_{ij} = (1-k) \cdot \hat{\omega}_i^{source}(t_k) + k \cdot \hat{\omega}_j^{target}(t_k)$$



Conclusions

- We have used OT to achieve the effect of the portamento
- We have used Frequency Reassignment to solve vertical and horizontal incoherence
- We have implemented 2 interpolation methods, in order to achieve a nice and clean result

November 2020