Homework 4 Report

ECE 477 Computer Audition

Meiying(Melissa) Chen

10/14/2019

Problem 1-a
Code: myNMF.m, normW.m

Problem 1-b
Code: myNMF.m, normW.m, istft.m, plotNMF.m, main.m

To run code: evaluate part 1-b in main.m
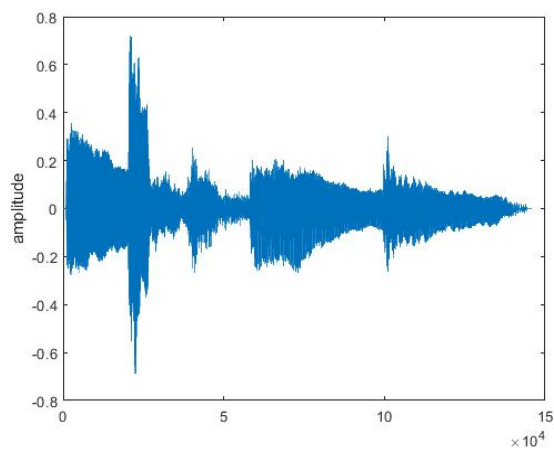
File: piano_recon_r4.wav
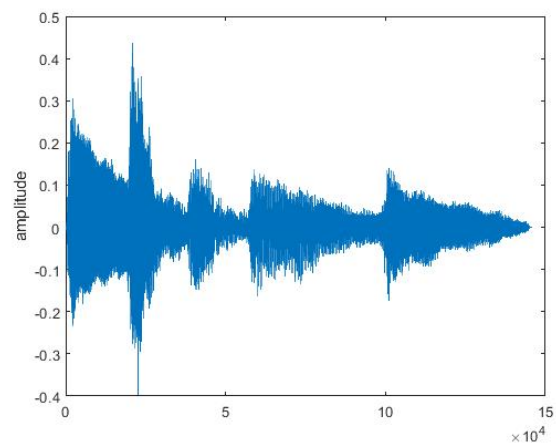

Figure 1 Original signal wave
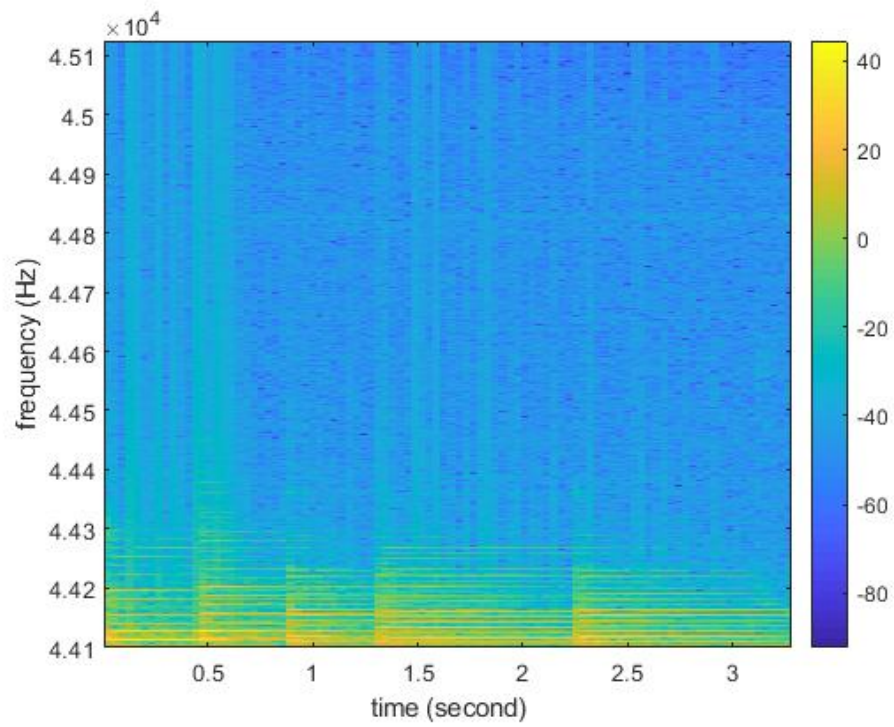

Figure 2 Reconstructed signal wave (r = 4)
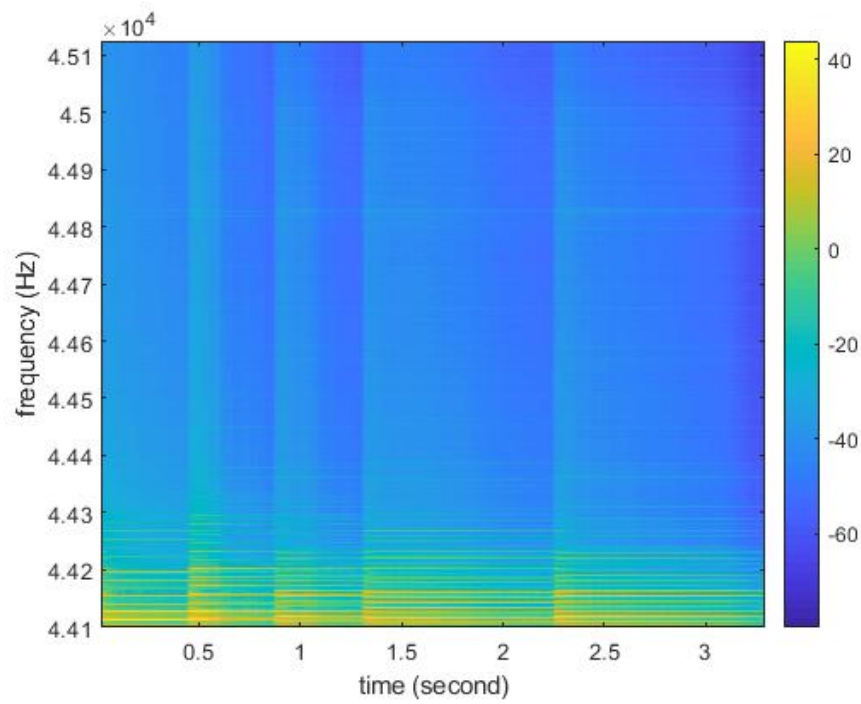
Figure 3 Original signal spectrogram

Figure 4 Reconstructed signal spectrogram (r = 4)
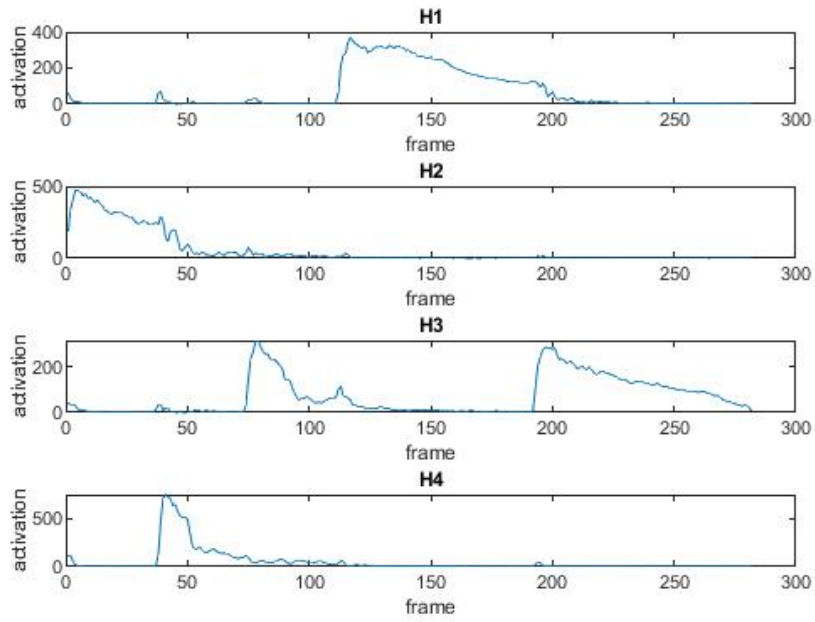
Figure 5 H rows i.e. activation (r = 4).

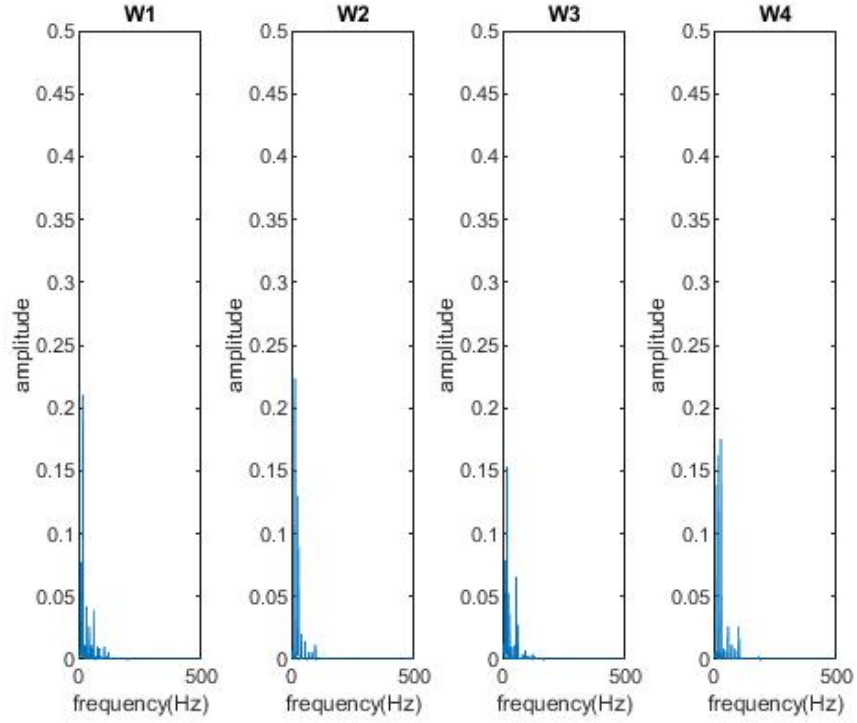*The activation of four notes at each time are clearly separated.*

Figure 6 W columns i.e. dictionary (r = 4)

Each Wi describes a frequency distribution of a unique note.

Problem 1-c

Code: part 1-c in main.m
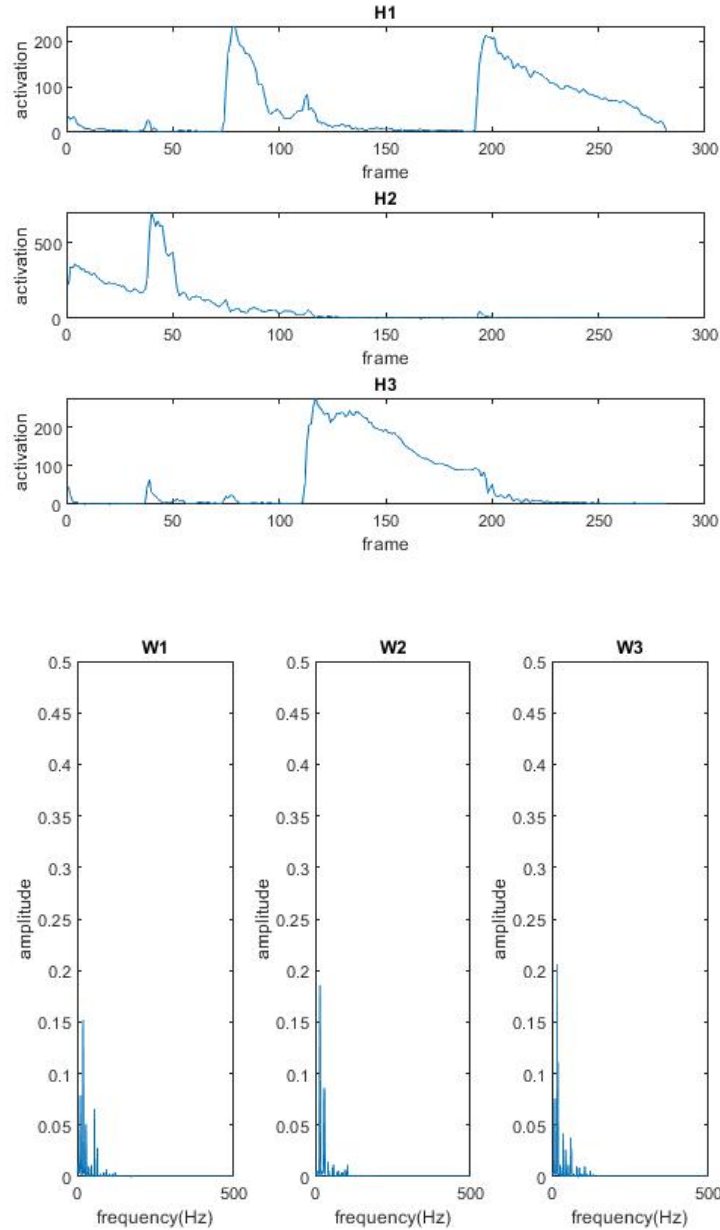
Files: piano_recon_r3.wav, piano_recon_r5.wav

Figure 7 H rows (upper) and W columns (lower) when r = 3

When r is set to 3, which is lower than the number of unique notes in the audio piece, the model does not have enough power to resolve all different notes, some maybe confused. Two out of four notes (W1 and W3) are successfully decomposed but the other two notes are confused. The two confused notes have more similar frequency distributions.

In compare to figure 3&4, we can see H1 in figure 7 is H3 in figure 3, H3 in figure 7 is H1 in figure 3, each denotes an activation of a unique node. And W1, W3 each denotes a frequency distribution accordingly. W2 and H2 are mixed by the other two nodes in figure 3&4.
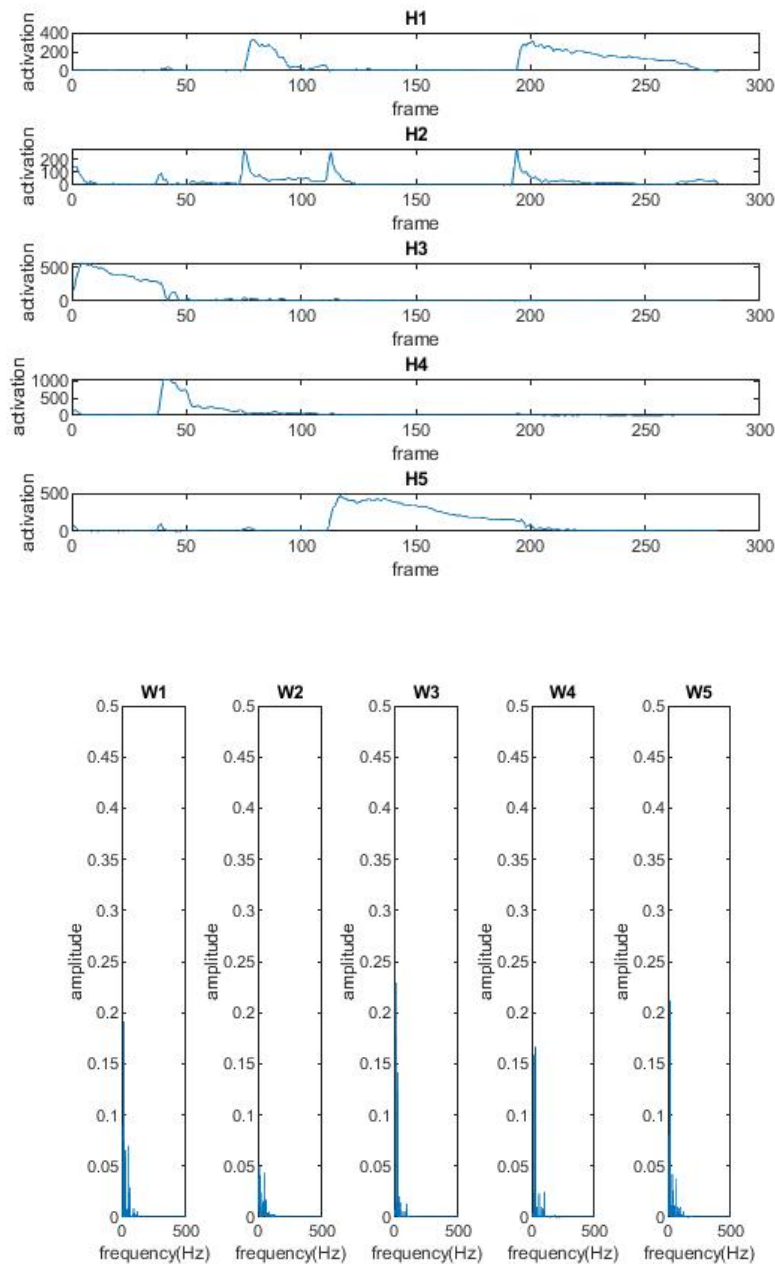

Figure 8 H rows (upper) and W columns (lower) when r = 5

When r is set to 5, which is larger than the number of unique notes, the model over decomposed the signal and outcomes some extra meaningless components (W4, H4).

In figure 8, H1, H3, H4, H5 each denotes an activation of a unique node, and W1, W3, W4, W5 are their frequency distribution accordingly. H2 and W2 are just components generated from the mix of some other nodes.

Problem 1-d
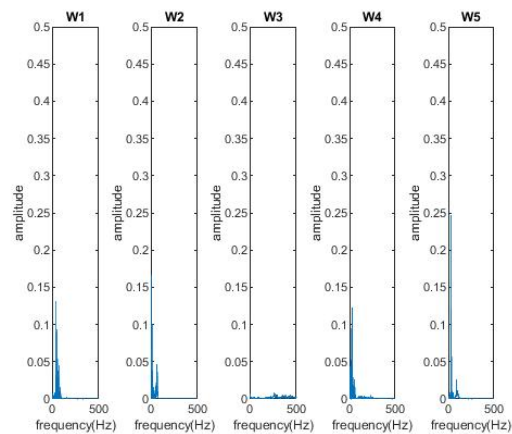Code: part 1-d of main.m

File: speech_train_rec.wav

Figure 9 Several basis in the dictionary of 'speech_train.wav'

Problem 1-e
Code: part 1-e of main.m
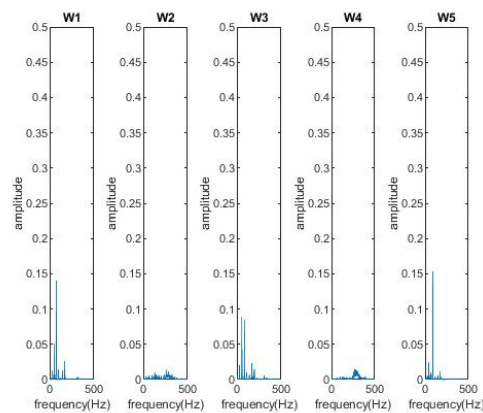
File: noise_train_rec.wav

Figure 9 Several basis in the dictionary of 'noise_train.wav'

Problem 1-f
Code: part 1-f of main.m

File: speech_sep.wav, noise_sep.wav

Problem 1-g
Code: part 1-g of main.m

	SDR	SIR	SAR	perm
Speech	-18.3878	4.8577	-17.1393	1.0000
Noise	-31.0749	-3.4632	-25.9891	2.0000

Problem 2-a
Code: myViterbi.m

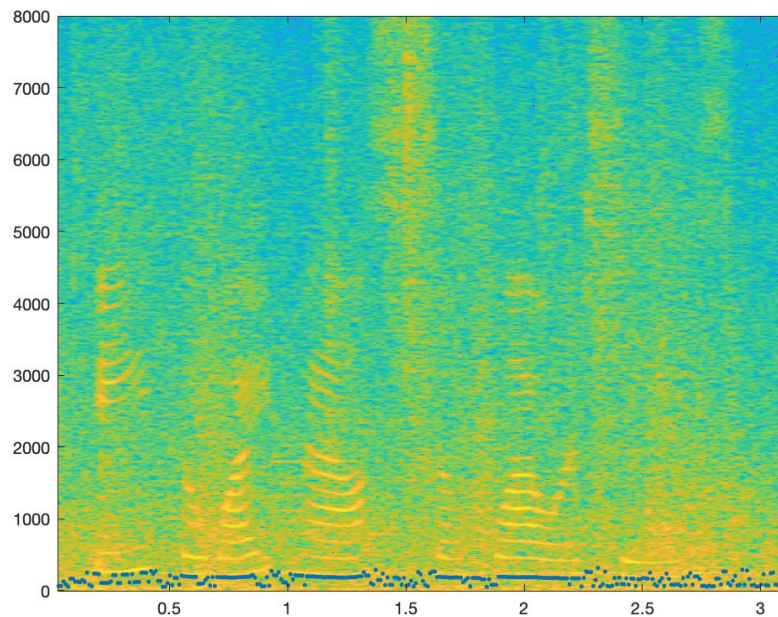## Problem 2-b
Code: part 2-b of main.m

Figure 10 Estimated pitches on magnitude spectrogram using max log-likelihood estimation

Most of the estimation are on the first harmonic places. But the contour of this estimation is not smooth enough.

Problem 2-c
Code: myViterbi.m, viterbi.m, part 2-c of main.m

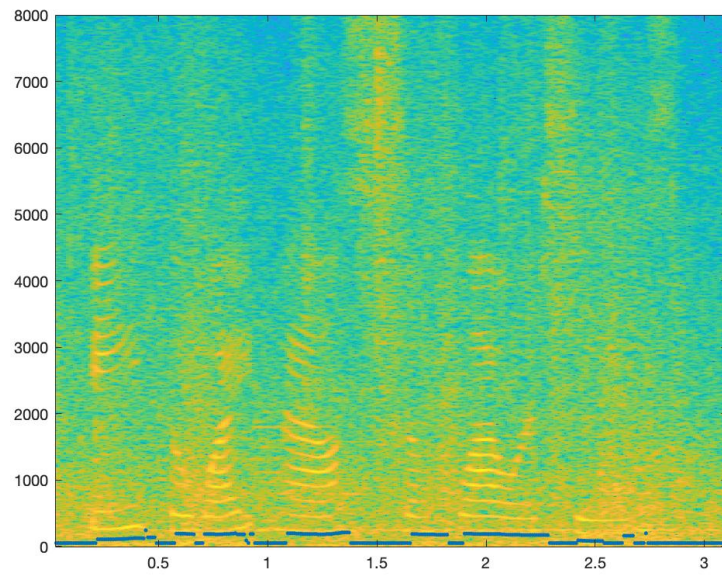The contour of the Viterbi algorithm estimation(below) is much more smoother.

Figure 10 Estimated pitches on magnitude spectrogram using Viterbi algorithm