520.445/645 - Audio signal processing
Fall 2021
Project 1

Due: 21 October 2021, 11:59pm

Assignment

PART 1: This project introduces you to the concept of signal estimation from the magnitude spectrum (i.e. without access to the phase). The paper by *Zhu et al.* (attached) outlines a detailed approach to reconstruct a near artifact-free signal. Your task is to carefully read the paper; then implement your own code that replicates the proposed method (sections II and III of the paper). You are given a variety of audio signals. Your task is to compute the magnitude short-term spectrum of each signal (i.e. throw away the phase) then reconstruct the signal back using the method you developed. Follow the algorithm in sections II and III; and explore which parameter values give you the best signal estimation. Discuss these choices in your report.

Note: The proposed method uses overlap-and-add which is very sensitive to signal normalization. Pay close attention to scaling (normalizing) your windowed signal. Also note that even if you are not able to replicate the SER values shown in the paper, you could suggest other measures of reconstruction fidelity and discuss them in your report.

PART 2: Next, you will examine the benefits of this reconstruction method for Time-Scale Modification (TSM). TSM refers to speeding up or slowing down an audio signal without affecting its pitch or timbre. Section IV of the paper presents a method to achieve TSM. Implement TSM on the audio signals you reconstructed from Part 1. Implement a factor of 2 speeding up and factor of 2 slowing down. You are also given a paper by *Driedger and Muller*, that reviews a variety of other methods to achieve TSM and discusses their strengths and weaknesses. You have to implement a method of your choice described in the paper by *Driedger and Muller* and *compare* its results to TSM based on *Zhu, et al.*. Note, there is a lot of code already available for TSM methods; some are mentioned in the review paper. Make sure you implement your own algorithm and fully understand how it works. No plagiarism!

Report

Write a report that explains what you did and how it worked. Your report should not include your code (up to 5 pages). Include any relevant plots and graphs to highlight your points and how you set the parameters. In your analysis, comment on how the methods you implemented work, and how their performance compares between speech and music signals.

Notes

- Work on the project individually.
- Include your report, your MATLAB code and reconstructed and modified signals in your final submission. Make sure your MATLAB functions are working properly and send the TAs all the functions necessary to properly test your system. If your code crashes for any reason, you will not get credit for that part of the project.
- Email your project to both TAs <mheidar1@jhu.edu> and <dgrant22@jhu.edu>, no later than 21 October 2021, 11:59pm.