Singing Voice Separation System Design for Monaural Polyphonic Music Signals

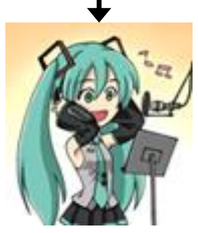
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Established in collaboration with MIT

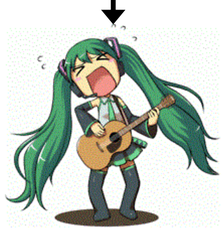
Motivation and Challenge

Given a Monaural Polyphonic Music Signal

Separated Into



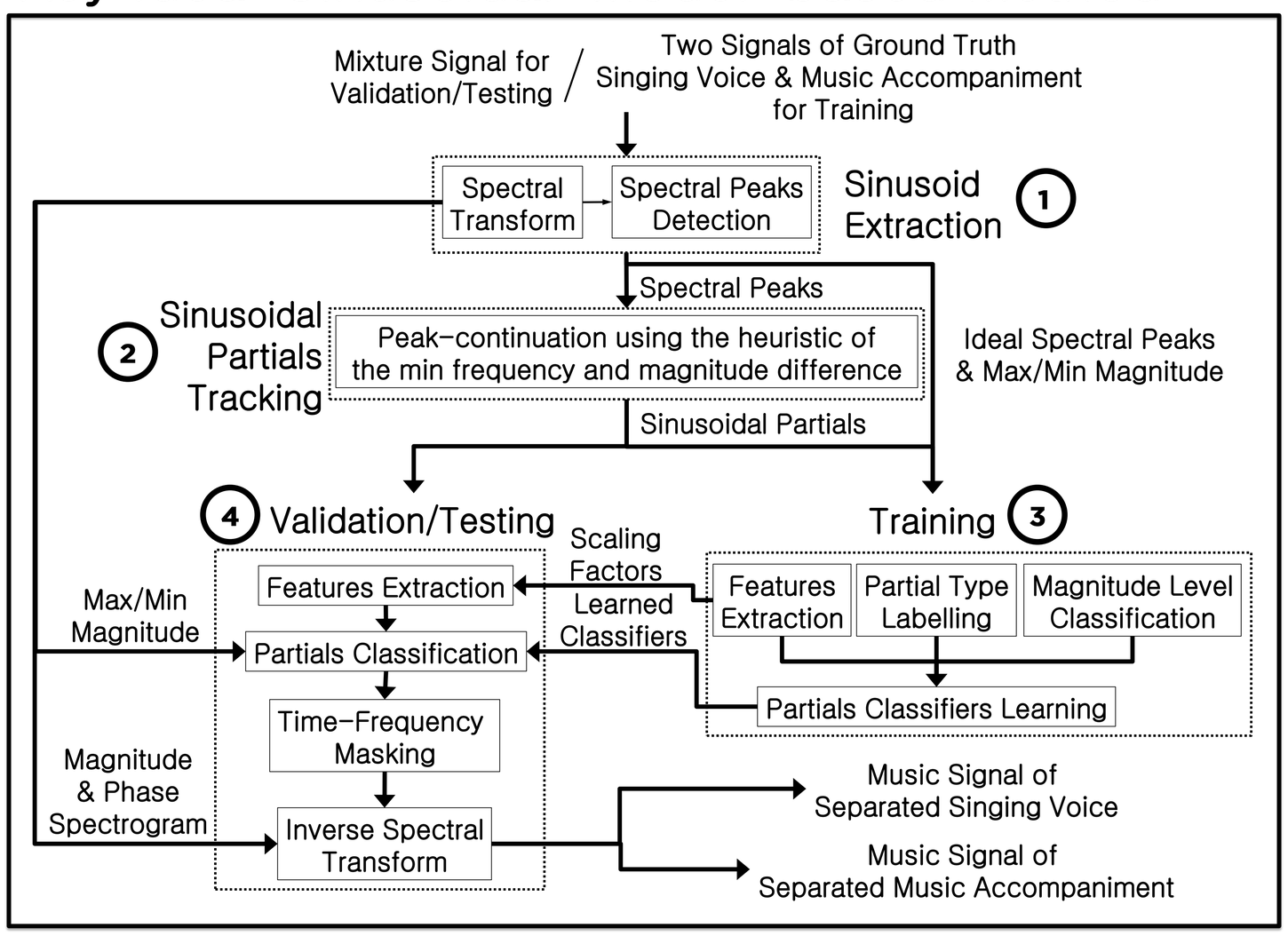
Singing Voice



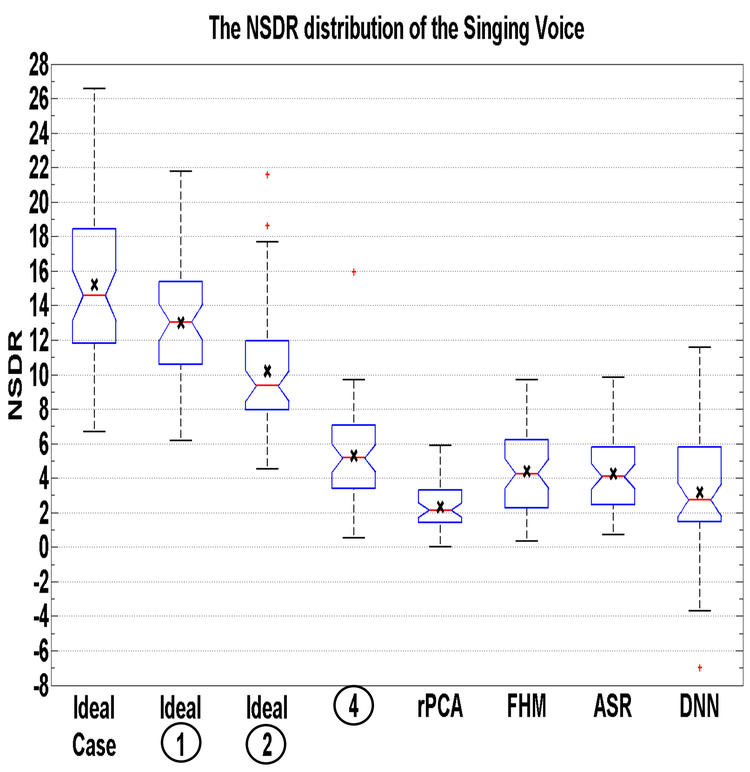
Music Accompaniment

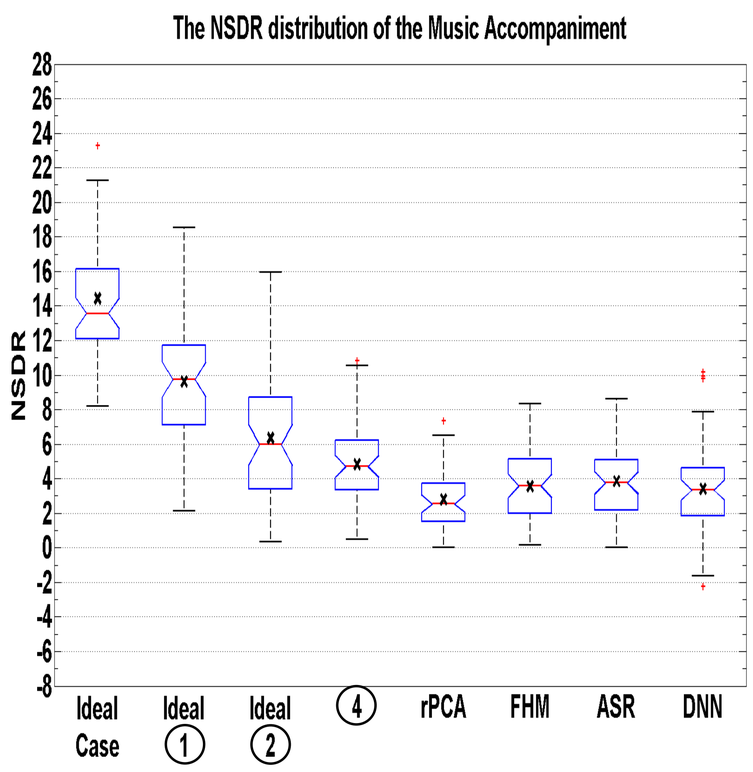
One Equation with 2 Unknown Variables! An ill-posed Problem!

Key Idea: Sinusoidal-Model-based Method

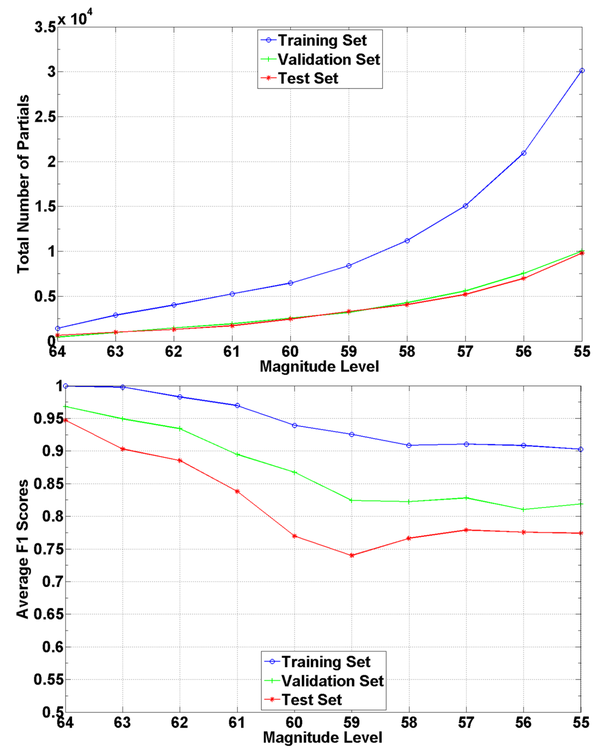


Experimental Results





Usage of Support Vector Machine (SVM)



Remark: (1) The larger the mean NSDR (marked as x), the better the separation quality. (2) rPCA, FHM, ASR & DNN are current state-of-the-art solutions.

Contributions

- To the best of our knowledge, this is the first work in the context of singing voice separation to do the sinusoidal partials classification using partial vibrato and tremolo with the magnitude level treated as the major feature
- The comparative results show that our system outperforms the rPCA-based, melody-based and deep-learning-based singing voice separation systems by 1.0975~3.0264 dB GNSDR gain on the iKala dataset.
- The nature of progressively solving the problem makes the separation result predictable. It also allows us to identify the weaknesses or limitation of the current solution, and then to solve them respectively.
- Our usage of SVM can be served as the reference for improving the partials classification in the future.

Name: Kin Wah Edward, LIN

SUTD PhD Programme (September 2013 Intake)

SUTD Pillar: ISTD

Supervisor: Prof. Simon, LUI