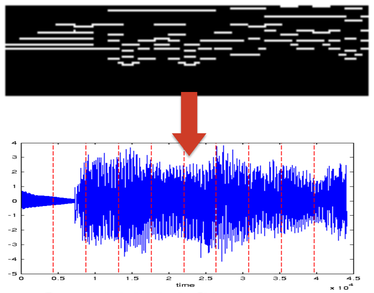
Polyphonic piano transcription

Jean-Baptiste Boin, Mohammad Sadegh Ebrahimi

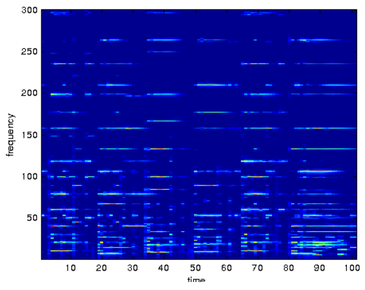
Motivation: The MIDI format gives a list of notes and their duration in a music. It can be used to render a sound file of that music. Our goal is to do the reverse operation, and to detect the notes played in a sound file so that we can eventually output a MIDI file of that music that is as close as possible to the original.

Rendering



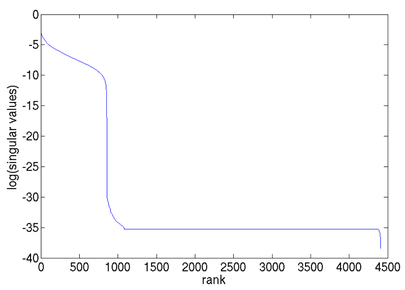
Our WAV files were rendered from MIDI files

Pre-processing



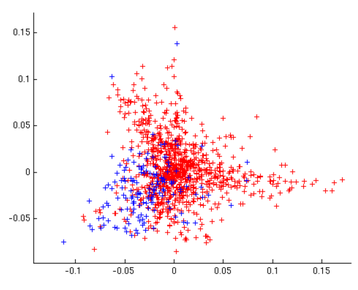
FFT on 100ms segments : 4410 components

PCA



Feature reduction down to 300 components

Sub-sampling

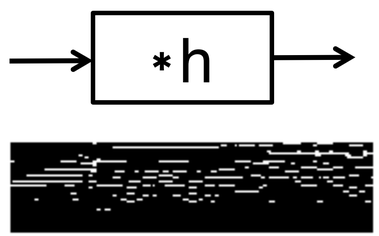


Increase the ratio of positive examples (see below)

Logistic regression

Train a (different)classifier for each note

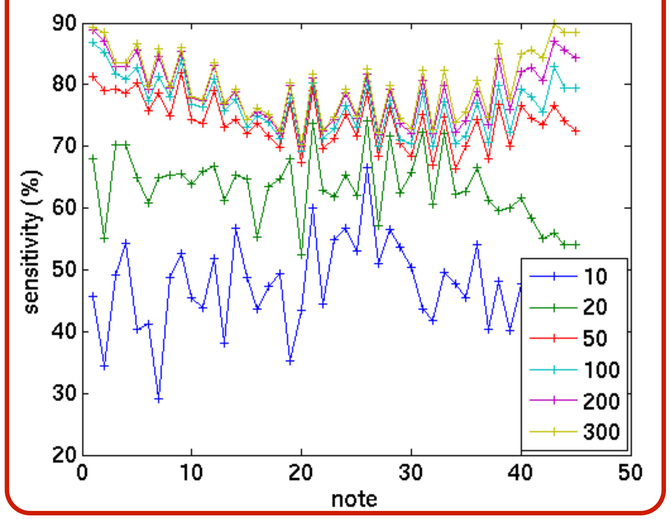
Post-processing*



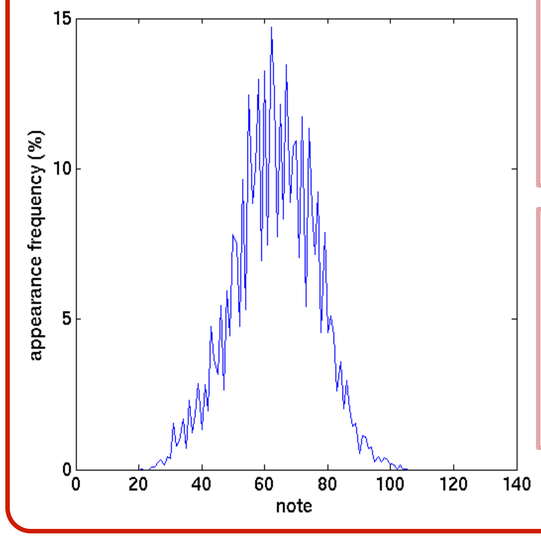
Filtering to get rid of many outliers

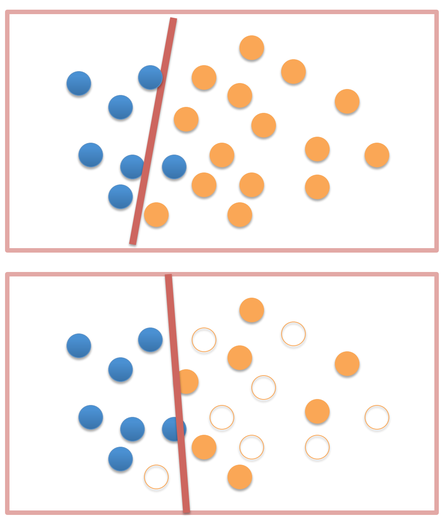
**Optional Step*

PCA



Asymmetric data





Sub-sampling

Results