Real-Time Detection and Visualization of Bad Clarinet Sounds

Aggelos Gkiokas, Kostas Perifanos, Stefanos Nikolaidis
Institute for Language and Speech Processing

Introduction

The growing use of computer software for music teaching led research to a new evolving domain, Music Visualization. Music visualization can serve different purposes in the context of music education. The aim of this work is to provide a real-time music 3D-visualization tool for clarinet sound in the context of music education. The feedback provided in real-time should be short and simple, avoiding to distract the students and helping them to go on despite any errors. Furthermore, the tool must help the student to gain a perception of their progress as the time goes by. Our approach focuses on students from beginning to intermediate level.

Clarinet Bad Sounds

Unstable Tones:

Unstable notes have many causes such as insufficient amount of airflow for the specific tone or not firm embouchure. Instability can be either pitch instability or RMS energy instability. Both can be easily detected by calculating the standard deviation of pitch and RMS-energy within a note (or part of note).

Onset-Offset Discarding:

To cope with different levels of students we substitute standard deviation for each component by a multiple of it by a value \(\alpha\). The lesser \(\alpha\), the better the sound is.

System Overview

Real-Time Audio Recognizer (RTAR) reads streamed data from the microphone as the student performs the musical piece. With a conventional front-end processing scheme RTAR processes a window of 25 ms long every 10 ms with a 60% overlap. For every window it processes, RTAR writes output data to the Audio Buffer and sends a message to the synchronizer module that a new frame is processed. The synchronizer activates the Error-Detection (ED) module. ED reads the Audio Buffer and computes the “Quality” of the sound. Changing the values according to the student’s performance produces a meaningful shape evolving over time:

The Visual Model

The main idea is to represent a note as circle. This circle has four attributes to control (plus the color, a total five). These attributes can be shown in figure below. Changing the values according to the student’s performance produces a meaningful shape evolving over time:

The RMS instability is directly related with the sphere shape, because of the proportional relationship between RMS energy and \(R\). The lesser \(R\), the better the sound is. A Hollow note is represented as a more “flabby”, “sleazy” shape, as shown in Figure 7. \(Ry/Rx\) is decreasing as hollowness increases and attribute \(freq\) is low valued and decreases as hollowness increases. \(dR/Rx\) is proportional to RMS energy.

Pitch and RMS Instability

The RMS instability is directly related with the sphere shape, because of the proportional relationship between RMS energy and \(Rx\). Therefore an RMS unstable note is directly shown.