

Informe Final - Proyecto 2016

PROCESAMIENTO DIGITAL DE SEÑALES DE AUDIO

MAESTRÍA EN INGENIERÍA ELÉCTRICA del *Instituto de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de la República.*

Juan Braga

1 de diciembre de 2016

Extracción de embocadura en Aliento/Arrugas: Introducción

En Aliento/Arrugas de Marcelo Toledo se utiliza como recurso compositivo tres tipos de embocadura para ejecución de la flauta. En la ejecución del instrumento se diferencian por el ángulo que forma la embocadura frente al flujo de aire. Desde la audición se puede percibir las diferencias en el material sonoro generado. Se enlista a continuación los nombres para cada caso, manteniendo su denominación en Inglés (idioma utilizado en la partitura de la obra). Además en la Figura 1 se observa su notación en la partitura de Aliento/Arrugas.

- *Blow Hole Covered*: El flujo de aire ingresa directo al tubo de la flauta, sin generar turbulencia contra el filo de la embocadura.
- *Breathy Embouchure*: Caso intermedio entre las otras dos embocaduras.
- *Normal Embouchure*: Embocadura clásica de la flauta, donde el el flujo de aire frente al filo de la embocadura genera la excitación tonal.

Figure 1 displays three musical systems (a, b, c) illustrating different flute embouchure techniques and their corresponding vocalizations. Each system includes a Flute staff and a Voice staff.

- (a) Blow Hole Covered:** The system is marked "INTENSO E CON FORZA!". The Flute staff shows a melodic line with dynamics ranging from *sfz/p* to *ff*. The Voice staff shows vocalizations like "Jh hak - sh" with dynamics *sfz*, *p*, and *mf*. A box labeled "bhc" (Blow Hole Covered) is present.
- (b) Breathy Embouchure:** The system is marked "Lento". The Flute staff shows a melodic line with dynamics ranging from *mf* to *ff*. The Voice staff shows vocalizations like "Ksh - K" and "Ksh - Ks ts Ks ts Ks" with dynamics *mf*, *f*, and *sfz/p*. A box labeled "bhc" (Breathy Embouchure) is present.
- (c) Normal Embouchure:** The system is marked "LENTO, DELICATO E LONTANO" and "AGITATO". The Flute staff shows a melodic line with dynamics ranging from *ppp* to *ff*. The Voice staff shows vocalizations like "a - o - a - o" and "mf/p" to *ff*. A box labeled "bhc" (Normal Embouchure) is present.

Figura 1: Notación de las embocaduras se observa en la parte superior de los sistemas. (a) *Blow Hole Covered*. (b) *Breathy Embouchure*. (c) *Normal Embouchure*. Fragmentos extraídos de la partitura de Aliento/Arrugas.

Definición del Problema

Se propone la extracción automática del tipo de embocadura a través del análisis computacional de grabaciones de la obra. En este caso se utilizará una estrategia de resolución del tipo de reconocimiento de patrones, en particular de clasificación supervisada.

Datos

Se cuenta con 3 grabaciones de diferentes intérpretes de la obra *Aliento/Arrugas*. Los intérpretes son: Pablo Somma, Emma Resmini y Ulla Suokko. Los archivos de audio se etiquetaron utilizando el software *Sonic Visualizer* dividiendo los fragmentos de audio en 5 clases:

- Silencio.
- Silencio con respiración del intérprete.
- Sonido generado con *Blow Hole Covered*.
- Sonido generado con *Breathy Embouchure*.
- Sonido generado con *Normal Embouchure*.

Extracción de características

Para esto se supone que el grado de tonalidad permite caracterizar el material sonoro generado con cada embocadura. Se enlista a continuación de forma cualitativa el grado de tonalidad con la respectiva embocadura:

- *Blow Hole Covered*: Sonido atonal.
- *Breathy Embouchure*: Sonido semi-tonal.
- *Normal Embouchure*: Sonido tonal.

Se propone entonces la utilización de dos características de audio clásicas para inferir el grado de tonalidad del material sonoro, desde las muestras digitales de audio:

- *Voicing* (Klapuri and Davy, 2007, Chapter 12).
- *Zero-Crossing Rate* (Rabiner and Schafer, 1978, Chapter 4).

Se supone entonces que mediante el computo de estas características es posible distinguir entre clases. A continuación se presentan los resultados y conclusiones sobre la estrategia planteada para la extracción de embocadura a partir del audio.

Resultados

Para el cálculo se utilizan ventanas de análisis de $23ms$ y saltos del 50% del largo de la ventana. Se dejaron afuera del análisis las clases: *Silencio* y *Silencio con respiración del intérprete*.

A continuación en la Figuras 2 y 3 se observa el espacio de características con distinción por clase para los diferentes intérpretes. Se puede observar que la *Normal Embouchure* es separable frente al resto con estos cálculos. No es así el caso de las embocaduras *Blow Hole Covered* y *Breathy Embouchure*.

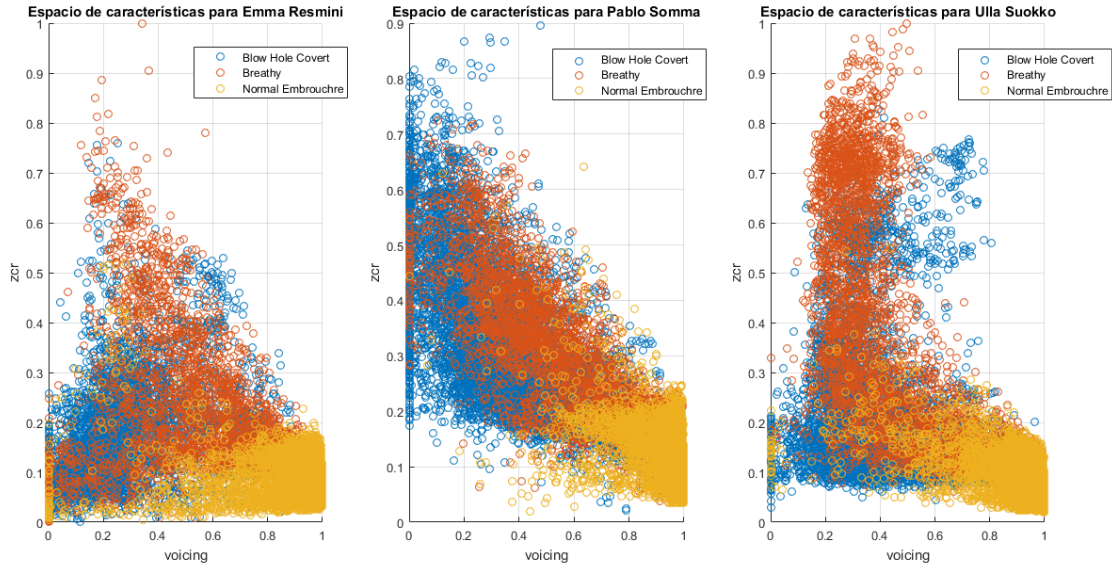


Figura 2: Espacio de características etiquetado por clases para cada uno de las interpretaciones.

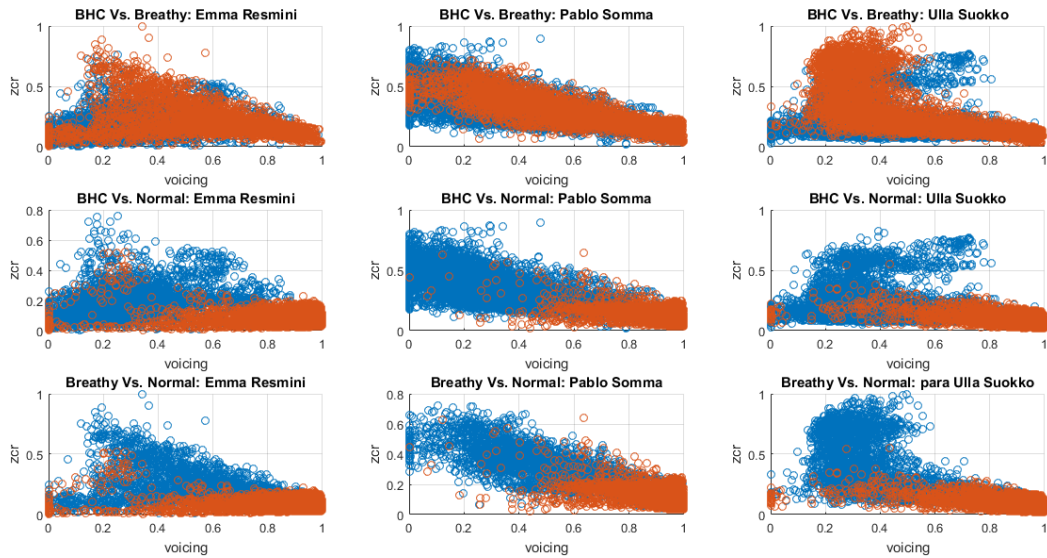


Figura 3: Comparación clase vs. clase en el espacio de características.

Se observa además en las Figuras 4, 5 y 6 el comportamiento temporal de las características *Voicing* y *Zero-Crossing Rate* superpuestas con el espectrograma, para tres fragmentos de la interpretación de Emma Resmini. Corresponden a *Blow Hole Covered*, *Breathy Embouchure* y *Normal Embouchure* respectivamente. Se corrobora lo dicho anteriormente sobre la separabilidad entre clases.

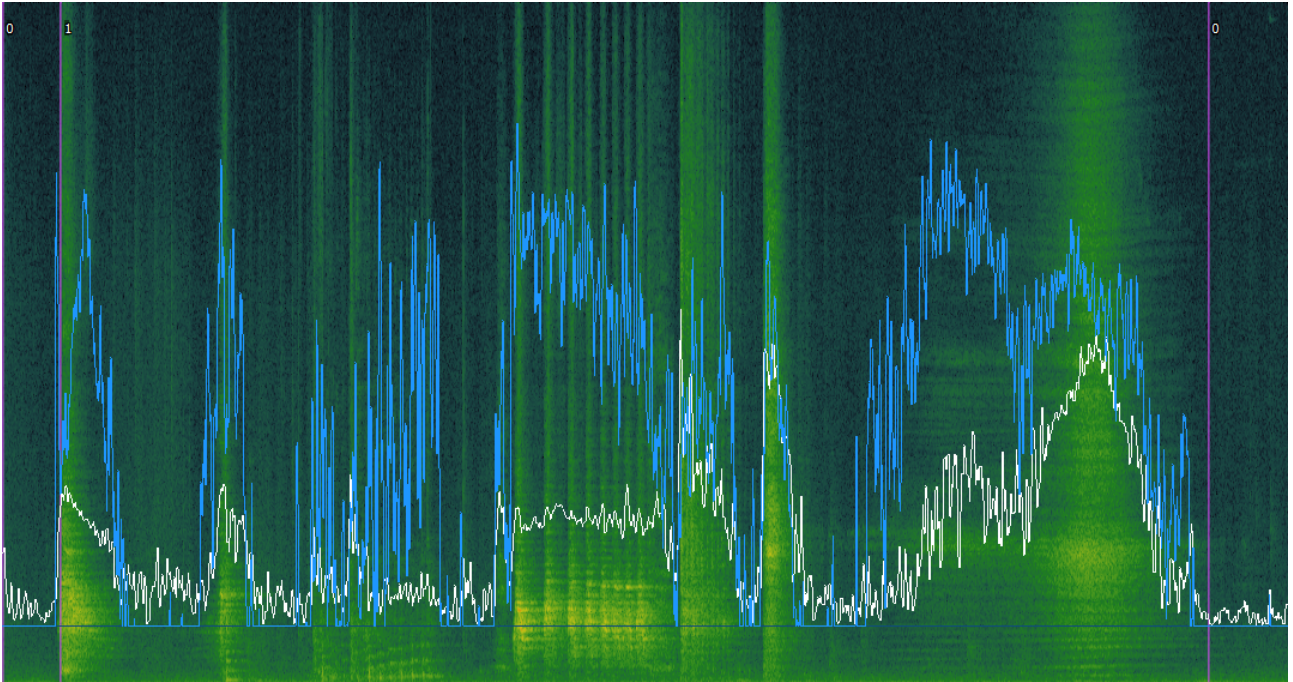


Figura 4: Visualización de un fragmento de la interpretación de Emma Resmini con embocadura *Blow Hole Covered*. Se observa en celeste el *Voicing* y en blanco *Zero-Crossing Rate*.

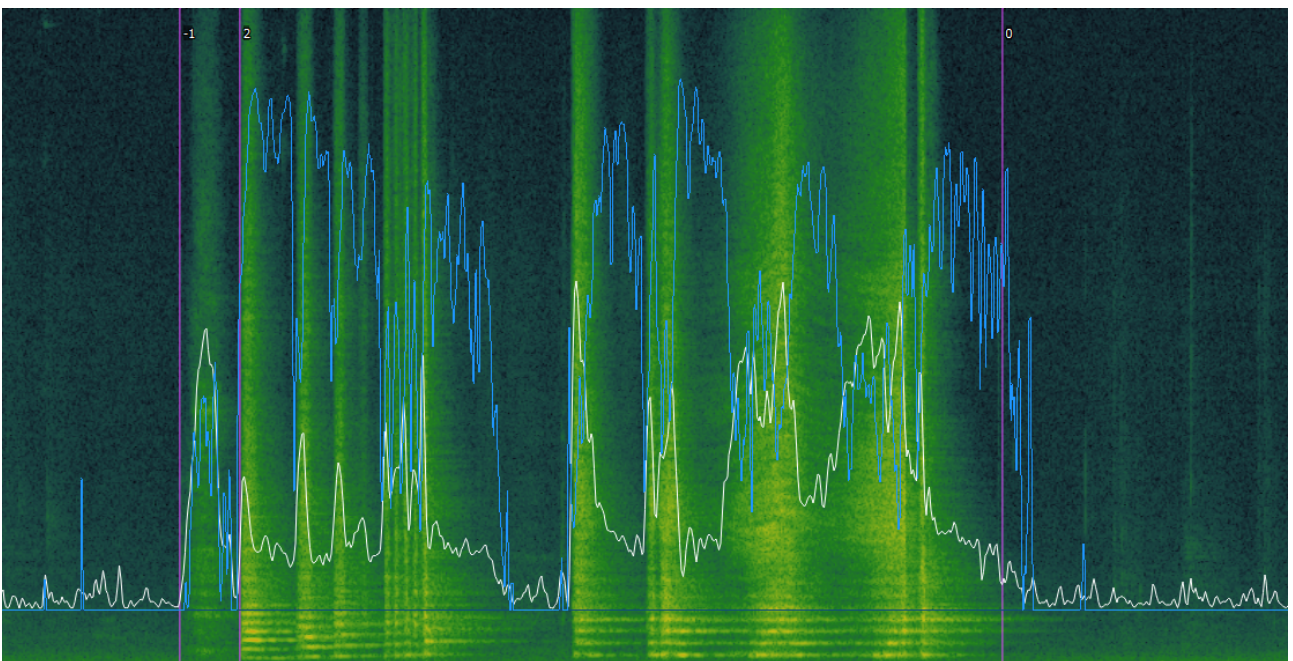


Figura 5: Visualización de un fragmento de la interpretación de Emma Resmini con embocadura *Breathy Embouchure*. Se observa en celeste el *Voicing* y en blanco *Zero-Crossing Rate*.

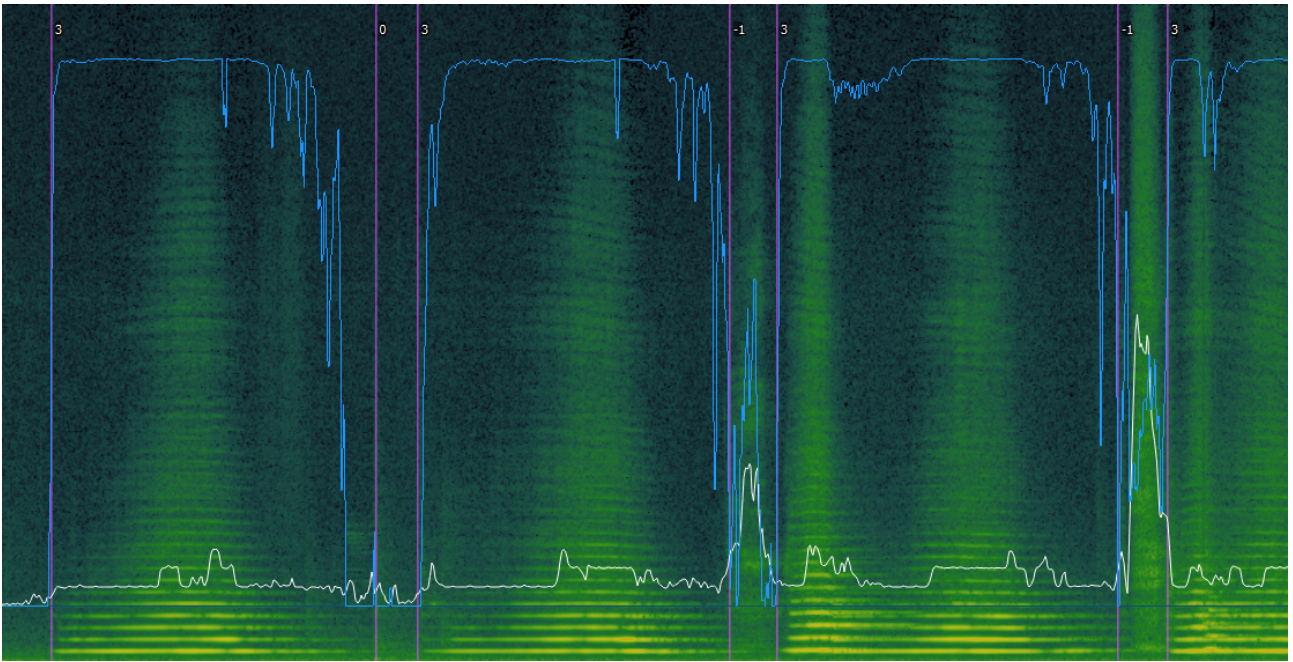


Figura 6: Visualización de un fragmento de la interpretación de Emma Resmini con embocadura *Normal Embouchure*. Se observa en celeste el *Voicing* y en blanco *Zero-Crossing Rate*.

Conclusiones

Las clases definidas por las embocaduras *Breathy Embouchure* y *Blow Hole Covered* no son separables en el espacio de características presentando en el informe. Por el contrario *Normal Embouchure* se distingue del resto.

Referencias

- Klapuri, A. and Davy, M. (2007). *Signal processing methods for music transcription*. Springer Science & Business Media.
- Rabiner, L. and Schafer, R. (1978). *Digital Processing of Speech Signals*. Prentice-Hall signal processing series. Prentice-Hall.