

Extracción de Embocadura en Aliento/Arrugas

Procesamiento Digital de Señales de Audio - Curso 2016

Juan Braga

Instituto de Ingeniería Eléctrica (IIE)
Facultad de Ingeniería (FING)
Universidad de la República (UdelaR)

Diciembre 2016

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Agenda

Introducción

Embocadura

Aliento/Arrugas de Marcelo Toledo

Definición del Problema

Conjunto de Datos

Extracción de Características

Experimentos

Primer experimento

Segundo experimento

Refinamiento de la extracción de características basada en MFCC

Conclusiones

Trabajo a Futuro

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura

Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos

Extracción de
Características

Experimentos

Primer experimento

Segundo experimento

Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Introducción

Flauta Traversa

- ▶ Flauta Traversa perteneciente a la familia de las *Maderas*
- ▶ Excitación periódica de la columna de aire generada por turbulencia de la colisión del flujo de aire contra el bisel de la embocadura.
- ▶ Altura definida por las llaves del instrumento.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Introducción

Técnicas Extendidas

- ▶ Exploración de nuevas sonoridades.
- ▶ Extensión de los límites naturales del instrumento.
- ▶ Técnicas reproducibles y bien definidas, generan que los instrumentistas tengan que expandir sus habilidades con el instrumento.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Introducción

Extracción de contenido musical (MIR por su sigla en inglés)

- ▶ Las alturas y duraciones ya no definen el material sonoro.
- ▶ El repertorio contemporáneo define una área del MIR desafiante y de espectro más amplio.
- ▶ Hay lugar a contribuciones científicas mediante el acercamiento al material sonoro exótico.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

- ▶ Refiere al aparato de producción de la excitación de la columna de aire, en conjunto a la técnica de soplido.
- ▶ Es determinante del material sonoro ejecutado, siendo perceptible de forma auditiva a través de cambios en la Dinámica, Altura y Timbre.
- ▶ Parámetros físicos que determinan la embocadura:
 - ▶ Ángulo de la Flauta.
 - ▶ Apertura de los labios.
 - ▶ Posición de los labios.
 - ▶ Presión de Aire.

Introducción

Embocadura

Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Aliento/Arrugas de Marcelo Toledo

Sobre la pieza

- ▶ Obra para Flauta Traversa solista, compuesta por el argentino Marcelo Toledo.
- ▶ Emplea técnicas extendidas como recurso compositivo:
 - ▶ *Flutter Tonguing* - Aleteo de lengüa.
 - ▶ *Tongue Noises* - Ruidos con la lengüa.
 - ▶ *Percussive Sounds* - Llaveo.
 - ▶ *Microtonal Inflections* - Inflecciones microtonales.
 - ▶ *Multiphonics* - Sonidos multifónicos (cantar y tocar a la vez).
- ▶ Además emplea como recurso expresivo cambios en la embocadura a lo largo de la pieza. Detallando en la partitura la embocadura correspondiente .
- ▶ Tres tipos de embocadura:
 - ▶ *Blow Hole Covert* - Embocadura normal o tradicional.
 - ▶ *Normal Embouchure* - Agüjero cubierto.
 - ▶ *Breathy Embouchure* - Embocadura con aire.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Aliento/Arrugas de Marcelo Toledo

Notación

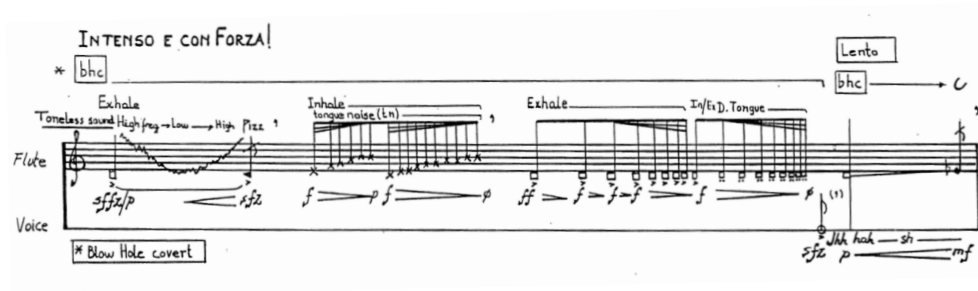


Figura: Notación de *Blow Hole Covert* se observa en la parte superior del sistema.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura

Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos

Extracción de
Características

Experimentos

Primer experimento

Segundo experimento

Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Aliento/Arrugas de Marcelo Toledo

Notación

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura

Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

The image displays a musical score for Flute and Voice. The Flute part is written on a single staff, and the Voice part is written on a single staff. The score is divided into three sections: 'LENTO, DELICATO E LONTANO', 'AGITATO', and 'CON FORZA'. The 'LENTO, DELICATO E LONTANO' section includes a box labeled 'normal embouchure'. The 'AGITATO' section includes a box labeled 'bhc'. The 'CON FORZA' section includes a box labeled 'bhc'. The score includes various dynamics such as *ppp*, *p*, *mf/p*, *f*, *ff*, and *fff*. It also includes articulation marks like accents, slurs, and breath marks. The Voice part includes lyrics: 'a', 'σ', 'a', 'σ'. The Flute part includes various musical notations such as notes, rests, and slurs.

Figura: Notación de *Normal Embouchure* se observa en la parte superior del sistema.

Aliento/Arrugas de Marcelo Toledo

Notación

The image shows a handwritten musical score for Flute and Voice. The Flute part is written on a single staff with a treble clef. It includes dynamic markings such as *mf*, *ff*, *sfz/p*, and *f*, along with crescendo and decrescendo hairpins. Above the staff, there are three boxes labeled "C breathy", "bhc", and "breathy", with lines indicating their duration. The Voice part is written below the Flute staff, consisting of syllables: "Ksh", "K", "Ksh", "K", "In", "Ex", "In", "Ex", "In", "Ex", "s", and "Ksh . Ks ts Ks ts Ks". The syllables are grouped under the "bhc" and "breathy" boxes. The notation is a mix of standard musical symbols and phonetic transcriptions.

Figura: Notación de *Breathy Embouchure* se observa en la parte superior del sistema.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura

Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Definición del Problema

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Hipótesis central

Los cambios de embocadura generan variaciones en el material sonoro perceptibles de forma auditiva.

Se propone...

La extracción automática del tipo de embocadura a través del análisis computacional de la obra.

Estrategia de resolución

Enfoque tipo *Reconocimiento de Patrones*. Se procesa el audio tipo *Bag of Frames* a partir del cómputo de características de audio.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

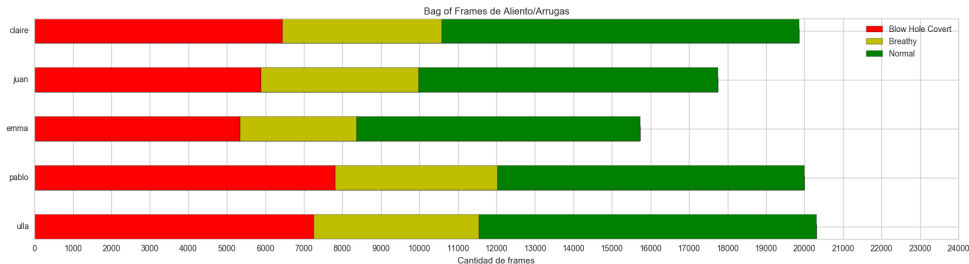
Conjunto de Datos

Grabaciones

Se cuenta con 5 grabaciones etiquetadas manualmente de Aliento/Arrugas por los intérpretes: Ulla Suokko, Pablo Somma, Emma Resmini, Juan Pablo Quinteros y Claire Chase.

5 folds

Cada obra define un conjunto de test.



Extracción de Características

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Mel-Frequency Cepstral Coefficients (MFCC)

Extraen contenido tímbrico de la señal.

Linear Prediction Coefficients (LPC)

Envolvente espectral.

Características Espectrales y Armónicas (CEA)

Roll-off, Centroid, Bandwidth, Zero-Crossing Rate y Voicing

Octave-based Spectral Contrast (SC)

Es una medida de contraste por octavas. Extrae la diferencia entre los picos y los valles.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
**Extracción de
Características**

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Experimentos

Consideraciones generales

- ▶ Se utilizan tres clasificadores distintos para la evaluación de desempeño:
 - ▶ *K-Nearest Neighbors* (10 vecinos)
 - ▶ *Random Forest* (10 árboles).
 - ▶ *Support Vector Machines* (kernel lineal).
- ▶ Se utilizan con parámetros por defecto, ya que no es objetivo del presente trabajo optimizar los clasificadores.
- ▶ Se centran en cero y se escalan a varianza unidad todos los datos, previo al entrenamiento y clasificación.

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Primer experimento

Objetivo

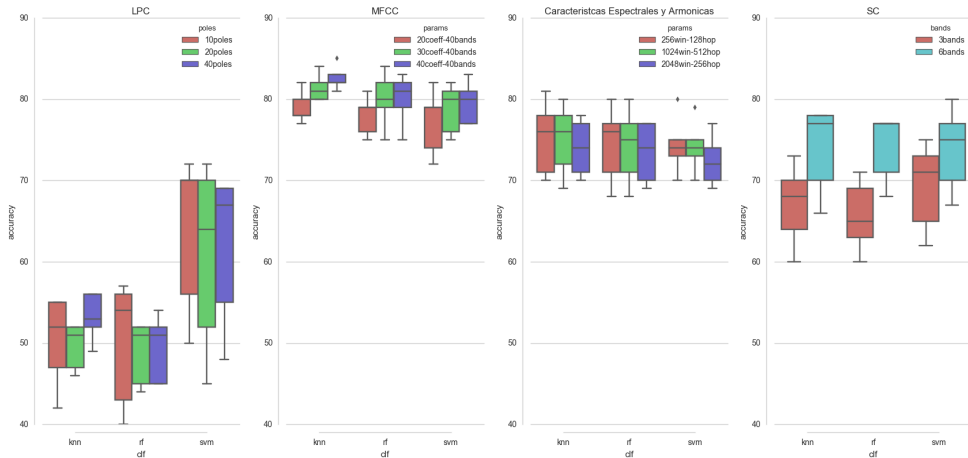
Evaluar el poder de separación de embocaduras de las características propuestas y determinar la de mejor desempeño.

Parámetros de la extracción de características:

- ▶ Características Espectrales y Armónicas: (a) 11ms y 50 % de salto (256-128 muestras), (b) 23ms y 50 % (1024-512 muestras) y por último (c) 46ms y 50 % (2048-1024 muestras).
- ▶ MFCC: Ventana de análisis de 23ms y salto del 50 %, 40 bandas Mel y liftrado: (a) 20 coeficientes, (b) 30 coeficientes y (c) 40 coeficientes.
- ▶ LPC: Ventana de análisis de 23ms y salto del 50 % y cantidad de polos: (a) 10, (b) 20 y (c) 40.
- ▶ SC: Ventana de análisis de 23ms y salto del 50 % y numero de bandas: (a) 3 y (b) 6.

Primer experimento

Resultados



Extracción de
Embocadura en
Aliento/Arrugas de

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Evaluación de resultados

- ▶ *LPC*: El peor desempeño. Además la variación del número de polos no muestra variaciones significativas.
- ▶ *SC*: Desempeño medio. Existe una mejora notoria al aumentar las bandas.
- ▶ *CEA*: Desempeño medio. La variación de la ventana de análisis no genera cambios apreciables. Alguna estrategia de decorrelación podría mejorar los resultados.
- ▶ *MFCC*: El mejor desempeño. Contrapartida el número de dimensiones del vector de características.
- ▶ *k-NN*: Salvo para *LPC* mejor desempeño entre clasificadores.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

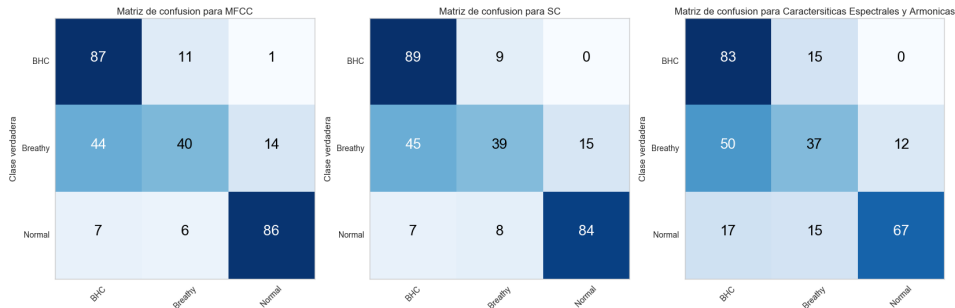
Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Matrices de confusión



- ▶ Independientemente de los *features* las clases *Blow Hole Covert* y *Normal Embouchure* se separan frente al resto.
- ▶ *Breathy* se confunde principalmente con *Normal Embouchure*

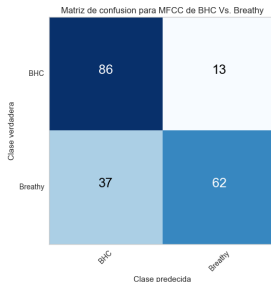
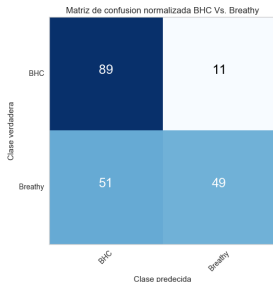
Segundo Experimento

Objetivo

Evaluar el desempeño de *MFCC* en un problema reducido: *Blow Hole Covert* Vs. *Breathy Embouchure*.

- Por simpleza se trabaja solamente con 40 bandas *Mel* y 20 coeficientes.

Matrices de confusión



Segundo Experimento

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Evaluación de resultados

- ▶ La confusión de *Breathy* sigue siendo considerable con un 37 %.
- ▶ Existe una mejoría con respecto al problema de tres clases.
- ▶ Una estrategia de clasificadores en cascada mejoraría el desempeño general, a cambio de costo computacional.

Refinamiento de la extracción de características basada en MFCC

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del
Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

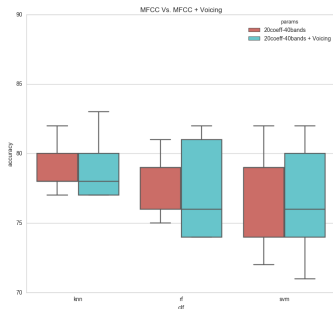
Primera estrategia

MFCC extrae aspectos tímbricos de la señal, por lo que a priori no tiene en cuenta la periodicidad de la señal. Se propone agregar *Voicing* como una dimensión más del vector de características, con el objetivo de mejorar la confusión de la clase *Normal Embouchure*.

- Confusión de *Normal Embouchure* contra el resto (para MFCC 40 bandas y 20 coeficientes): 13 %.

Refinamiento de la extracción de características basada en MFCC

Resultados



Matriz de confusión para MFCC+Voicing

	BHC	Breathy	Normal
BHC	88	10	1
Breathy	41	43	15
Normal	5	6	88

Clase verdadera

Clase predicha

- No hay mejoras significativas ni en el *accuracy* ni la confusión (11 %).

Refinamiento de la extracción de características basada en MFCC

Extracción de
Embocadura en
Aliento/Arrugas

Juan Braga

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
**Refinamiento de la
extracción de
características basada
en MFCC**

Conclusiones

Trabajo a Futuro

Preguntas

Segunda estrategia

Variación del largo de ventana de análisis, con salto del 50 % en todos los casos:

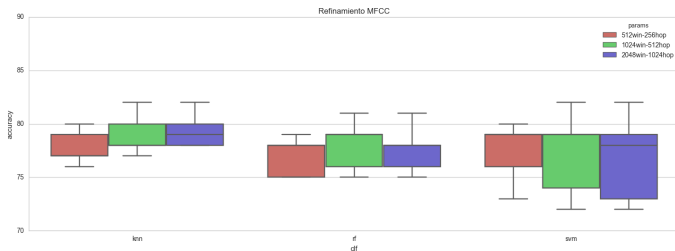
- ▶ 11ms
- ▶ 23ms
- ▶ 46ms

Refinamiento de la extracción de características basada en MFCC

Extracción de Embocadura en Aliento/Arrugas

Juan Braga

Resultados



- ▶ No existen cambios relevantes al variar la ventana de análisis.
- ▶ El mejor desempeño esta dado por 40 bandas *Mel* y 40 coeficientes. Resultado curioso.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del
Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

Extracción de Embocadura en Aliento/Arrugas

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Conjunto de Datos
Extracción de
Características

- Primer experimento
- Segundo experimento
- Refinamiento de la extracción de características basada en MECC

Trabajo a Futuro

Preguntas

- ▶ Se realiza satisfactoriamente la separación de las clases: *Normal Embouchure* y *Blow Hole Covert* frente al resto. Sin embargo existe gran confusión entre las clases *Blow Hole Covert* y *Breathy Embouchure*. Algún feature diseñado con un estudio más profundo de la naturaleza acústica de estas clases podría desambiguar el problema.
- ▶ La ambigüedad sugiere que el enfoque basado en el aspecto tímbrico a partir de la envolvente espectral no es suficiente. Queda planteada la posibilidad de que la información relevante para la separación de estas clases se encuentre la excitación.
- ▶ Se debe tener en cuenta por otro lado, que existen pasajes de transición, inicios y finales de frases musicales en que las clases se vuelven ambigüas y se debe tener en cuenta al momento de etiquetado.
- ▶ En adición a lo anterior, la estrategia de *Bag of Frames* es exigente y descarta la redundancia temporal como la información a priori del problema.

Introducción

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Definición del Problema

Conjunto de Datos
Extracción de
Características

Experimentos

Primer experimento
Segundo experimento
Refinamiento de la
extracción de
características basada
en MFCC

Conclusiones

Trabajo a Futuro

Preguntas

- ▶ Enfocar la resolución del problema desde la estimación de la excitación generada por el intérprete, y evaluar el desempeño con esta estrategia.
- ▶ Segmentación automática del audio en fragmentos de actividad de la flauta y silencios, como primer etapa del sistema de extracción de embocadura, previo a la etapa de clasificación.
- ▶ Salir del bag of frames, para utilizar la redundancia temporal y la información a priori del problema.

Preguntas

Extracción de Embocadura en Aliento/Arrugas

Juan Braga

Embocadura
Aliento/Arrugas de
Marcelo Toledo

Conjunto de Datos
Extracción de
Características

- Primer experimento
- Segundo experimento
- Refinamiento de la extracción de características basada en MFCC

Trabajo a Futuro

Preguntas

