

Un modelo de Markov para la segmentación automática de señales de audio

Rafael de Jesús Robledo Juárez

`rrobledo@cimat.mx`

Asesor: Dr. Salvador Ruíz Correa

Departamento de Ciencias de la Computación
Centro de Investigación en Matemáticas, Guanajuato



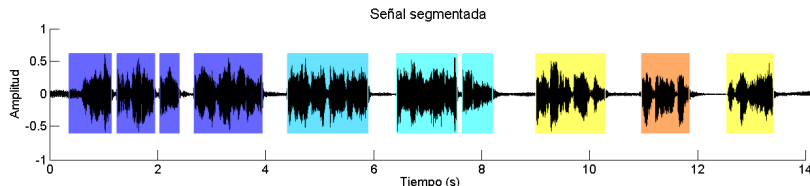
CIMAT

Presentación de Avance de Tesis, 2013



Objetivo

- Utilizar un modelo del tipo HMM (*Hidden Markov Model* por sus siglas en Inglés) por medio del cual sea posible particionar una secuencia de audio en regiones homogéneas. Cada región debe corresponder con la identidad de un hablante.
- A esta tarea se le conoce con el nombre de "*speaker diarisation*".



- Los segmentos obtenidos se etiquetarán de acuerdo a cada persona diferente que se identifique.



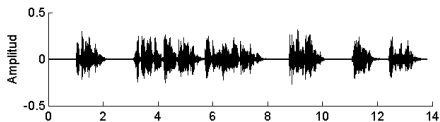
Motivación

- ▶ La identificación de las personas que participan en una grabación de audio, es una etapa importante en aplicaciones tales como transcripción automática y reconocimiento de voz.
- ▶ Un paso importante para la automatización de este proceso, consiste en poder segmentar la señal de audio sin necesidad de conocimiento a priori sobre el número o género de las personas involucradas en la grabación.

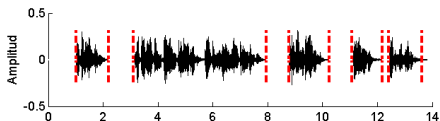


Metodología

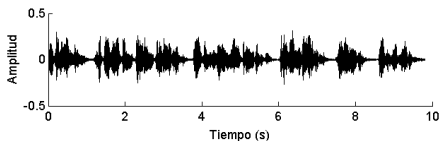
Procesamiento de señal (1)



► Señal original



► Identificación de silencios

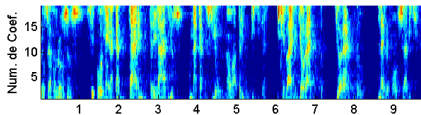


► Señal de audio recortada

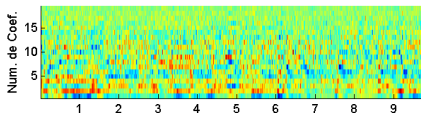


Metodología

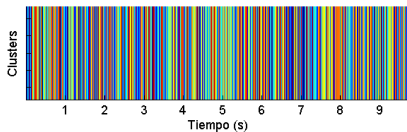
Procesamiento de señal (2)



► FFT + Banco de filtros + log



► MFCC

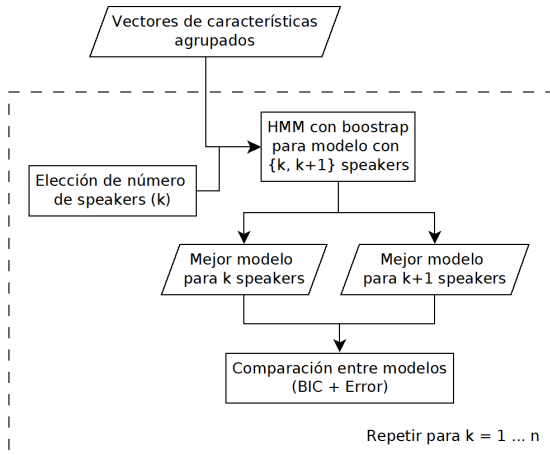


► k-means



Metodología

Modelo

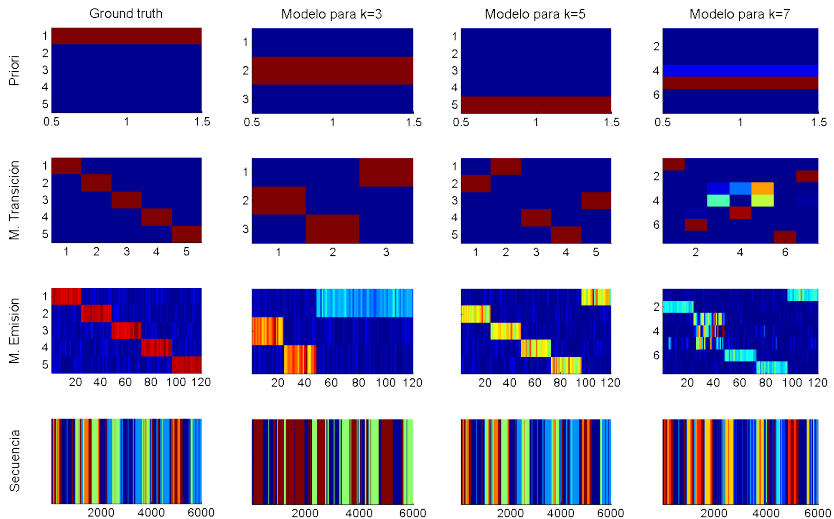


Avances

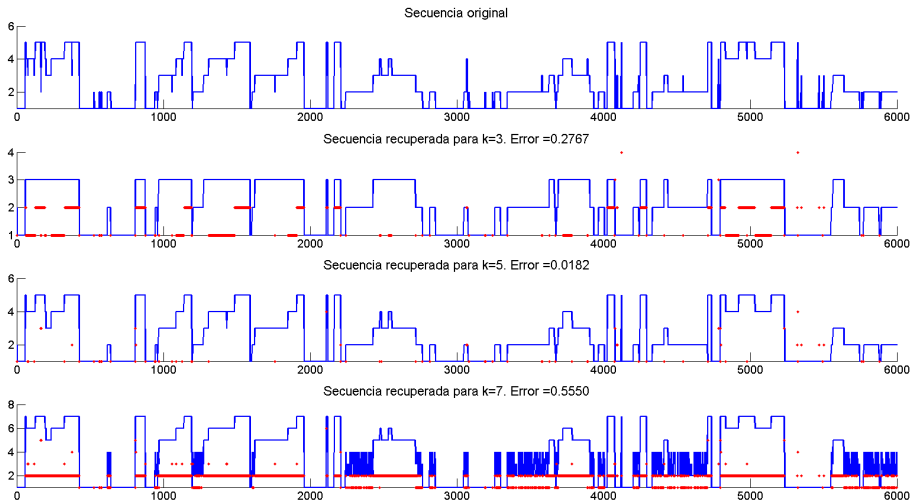
- ▶ Implementación completa del sistema.
- ▶ Resultados con datos sintéticos.
- ▶ Pruebas con datos reales.



Pruebas con datos sintéticos



Pruebas con datos sintéticos



Actividades

Febrero

Marzo

Abril

Mayo

Junio

Julio

Pruebas con
datos reales

Escritura y revisión

Comparación y
análisis de resultados



Material de consulta



L.R. Rabiner, B.H. Juang
Fundamentals of speech recognition
Pearson Education India, 2008.



C. M. Bishop.
Pattern Recognition and Machine Learning.
Springer, 2006.



L.R. Rabiner
A tutorial on hidden Markov models and selected applications in speech recognition
Proceedings of the IEEE, 1989.



T. Rydén.
Versus Markov chain Monte Carlo for Estimation of Hidden Markov Models: A Computational Perspective
Bayesian Analysis (2008) 3, Number 4, p. 659-688



E.B. Fox, E.B. Sudderth, M.I. Jordan, A.S. Willsky.
A sticky HDP-HMM with application to Speaker diarization.
Annals of Applied Statistics, 2011.

