

Procesamiento Digital de Audio

Profesor:
Caleb Rascón (caleb@unam.mx)

Objetivo: Este curso expondrá al(la) alumno(a) a temas de procesamiento de señales de audio, donde se cubrirán desde los conceptos teóricos hasta los aspectos de implementación. La intención es que el(la) alumno(la), al terminar este curso, le sea posible crear, de una manera eficiente, software que analice, procese, y regrese resultados de señales de audio en línea. Aunque se espera que el(la) alumno(a) tenga bases previas de Señales y Sistemas, así como de programación, el curso se llevará a cabo de tal manera que dichas bases serán refinadas y repasadas.

Temario:

1. Introducción al Curso (3 horas)

- (a) Terminología.
- (b) Motivación.
- (c) Aplicaciones (enfoque a Audición Robótica).

Objetivo: Comprensión de los beneficios del curso.

2. Preparación para Implementar Software de Audio (3 horas)

- (a) Repaso de Estructura de Datos (relevantes a Audio).
- (b) Instalación e Introducción a JACK Audio Connection Kit.
- (c) Uso de Baudline

Objetivo: Compilar exitosamente un simple programa de JACK.

3. Captura y Análisis de Señales de Audio (9 horas)

- (a) Conceptos Generales de Captura de Audio.
- (b) Transformada de Fourier.
- (c) La Biblioteca FFTW3 para Hacer la Transformada de Fourier en C.
- (d) Filtrado básico (paso-banda, paso-alto, paso-bajo).
- (e) Firmas Espectrales y Cancelación de Ruido
- (f) Detección de Actividad de Voz

Objetivo: (Obligatorio) Programa de JACK en el que se desfase (utilizando FFT) una señal de audio de un micrófono, y se filtre en tiempo real hacia una bocina.

4. Procesamiento de Múltiples Señales Concurrentes (3 horas)

- (a) Arreglos de Micrófonos.
- (b) Arreglos Multi-dimensionales de Micrófonos.
- (c) Obtención y prueba del Hardware.
- (d) Introducción al corpus AIRA.

Objetivo: Tener las bases de arreglos de micrófonos para los siguientes temas.

5. Cálculo de Dirección de Arribo (9 horas)

- (a) De una fuente: Correlación Cruzada.
- (b) Modelo de Campo Lejano y Cercano.
- (c) Reverberación.
- (d) Transformada de Fase (PHAT).
- (e) La Biblioteca Eigen para Manipular Matrices y hacer la Eigendescomposición en C++.
- (f) Cálculo de Múltiples Direcciones de Arribo: Método MUSIC.
- (g) Cálculo de Múltiples Direcciones de Arribo: Beamforming.

6. Separación de Fuentes (9 horas)

- (a) Por Análisis Estadístico Avanzado:
 - i. Principal Component Analysis.
 - ii. Independent Component Analysis.
 - iii. Non-negative Matrix Factorization.
- (b) Por Beamforming:
 - i. Delay-and-Sum (Tiempo).
 - ii. Delay-and-Sum (Frecuencia).
 - iii. Minimum Variance Distortionless Response.
 - iv. Linear Constraint Minimum Variance.
 - v. Generalized Side-lobe Canceler.
 - vi. Phase-based Frequency Masking.
- (c) Solución Híbrida: Geometric Source Separation.

Objetivo: Obtener las bases para el Proyecto Final.

7. Proyecto Final.

Deseable: *Máxima calificación: 10.* Implementación de una solución que calcule la dirección de arribo de una o varias señales de audio y las separe en línea. Entre más sofisticado y mejor justificado, más posibilidad de mayor puntaje.

Mínimo: *Máxima calificación: 9.* Implementación de sólo una parte de la solución completa:

- * Cálculo de la dirección de arribo de una o varias señales de audio.
- * O, separación en línea de una o varias fuentes en el ambiente, asumiendo que su dirección es conocida.

Entre más sofisticado y mejor justificado, más posibilidad de mayor puntaje.

Demostración en vivo en clase, así como una presentación del(la) alumno(a) del método y motivación del algoritmo.

Objetivo: Aprender los beneficios y problemas de cada método.

Esquema de calificación:

Por proyecto final.

Página del Curso: <http://calebrascon.info/PDA>

Bibliografía:

1. DeLian Wang
Computational Auditory Scene Analysis: Principles, Algorithms, and Applications
Edited by: DeLian Wang and Guy J. Brown,
IEEE Press, 2006
2. *Signals and Systems*
http://upload.wikimedia.org/wikipedia/commons/6/67/Signals_and_Systems.pdf
3. Paul Davis
JACK Audio Connection Kit
<http://jackaudio.org/>

Lecturas:

1. C. Rascon, G. Fuentes, I. V. Meza
Lightweight multi-DOA tracking of mobile speech sources.
EURASIP Journal on Audio, Speech, and Music Processing. 2015(11)
<http://asmp.eurasipjournals.com/content/2015/1/11/>
2. C. Rascon, I. V. Meza.
Localization of sound sources in robotics: A review.
Robotics and Autonomous Systems, 96(C), 2017
<https://www.sciencedirect.com/science/article/pii/S0921889016304742>
3. C. Rascon.
A corpus-based evaluation of beamforming techniques and phase-based frequency masking.
Sensors, 21(15), 2021
<https://www.mdpi.com/1424-8220/21/15/5005>