

Propuesta de Trabajo Fin de Grado en Física

Tutor/a: Jose Luis Bernier Villamor

Departamento y Área de Conocimiento:
Arquitectura y Tecnología de
Computadores

Cotutor/a:

Departamento y Área de Conocimiento:

Título del Trabajo: Estudio de algoritmos para la extracción de voz de una fuente de sonido monoaural

Tipología del Trabajo:

(Segun punto 3 de las
Directrices del TFG
aprobadas por Comisión
Docente el 10/12/14)

(Marcar
con X)

1. Revisión bibliográfica	X	4. Elaboración de nuevas prácticas de laboratorio	
2. Estudio de casos teórico-prácticos		5. Elaboración de un proyecto	
3. Trabajos experimentales		6. Trabajo relacionado con prácticas externas	

Breve descripción del trabajo:

Un problema habitual que el ser humano realiza con facilidad es la extracción o aislamiento de una voz en situaciones de ruido o cuando se encuentra mezclada con otros sonidos. Así pues, somos capaces de centrar nuestra atención en una conversación aislando la voz del resto de sonidos. Este tema ha recibido bastante atención últimamente y se han propuesto diversos algoritmos basados en distintos modelos y estrategias, por ejemplo, en modelos de convolución, modelos basados en las características físicas de la voz humana o en Deep Learning.

Para este trabajo se propone la implementación y estudio de algunos de los algoritmos propuestos, de cara a comparar sus resultados y prestaciones, así como estudiar en qué situaciones proporcionan mejores soluciones.

Hay que resaltar que, a diferencia del problema de la separación de fuentes ampliamente estudiado en las décadas anteriores donde se trata de separar n fuentes a partir de n señales mezcladas, en este caso se trata de separar sonidos que se encuentran mezclados en una única señal.

Objetivos planteados:

1. Revisión del estado del arte
2. Implementación de algoritmos propuestos
3. Elección de un banco de pruebas
4. Experimentación con los algoritmos
5. Análisis de los resultados

Metodología:

Se empezará revisando la bibliografía reciente y estudiando algunos conceptos relacionados con la Física del sonido y las características de la voz humana. A partir de la bibliografía existente, se implementarán algunos de los algoritmos que se han propuesto para resolver el problema, usando el lenguaje C++ y la plataforma Octave o Matlab.

Se escogerá un banco de pruebas estándar (benchmark) con el cual se probarán los algoritmos implementados.

Se revisarán los resultados de cara a comparar la bondad de los algoritmos estudiados.

Campus Fuentenueva
Avda. Fuentenueva s/n
18071 Granada

Tfno. +34 958242902
fjruiz@ugr.es

Comisión Docente de Físicas
Facultad de Ciencias

Firmado por: FERNANDO JOSE ROJAS RUIZ Secretario/a de Departamento

Sello de tiempo: 16/05/2018 10:50:13 Página: 1 / 2



cfP6wllk8uS1BJA4HsbZvn5CKCJ3NmbA

Bibliografía:

- Sayeh Mirzaei a, Hugo Van Hamme a, Yaser Norouzi b. *Under-determined reverberant audio source separation using Bayesian Non-negative Matrix Factorization*. Speech Communication 81 (2016) 129–137. Elsevier.
- Andrew J.R. Simpson, Gerard Roma, Mark D. Plumbley. *Deep Karaoke: Extracting Vocals from Musical Mixtures Using a Convolutional Deep Neural Network*
- Po-Sen Huang, Minje Kim, Mark Hasegawa-Johnson, Paris Smaragdis. *Joint Optimization of Masks and Deep Recurrent Neural Networks for Monaural Source Separation*. IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 23, NO. 12, DECEMBER 2015
- Pritish Chandna, Marius Miron, Jordi Janer, and Emilia Gómez. *Monoaural Audio Source Separation Using Deep Convolutional Neural Networks*. Low Latency Online Monoaural Source Separation
- Yukara Ikemiya, Katsutoshi Itoyama, and Kazuyoshi Yoshii. *Singing Voice Separation and Vocal F0 Estimation Based on Mutual Combination of Robust Principal Component Analysis and Subharmonic Summation*. IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 24, NO. 11, NOVEMBER 2016
- Bahar Khalighinejad, Tasha Nagamine, Ashesh Mehta, Nima Mesgarani. *NAPLIB: AN OPEN SOURCE TOOLBOX FOR REAL-TIME AND OFFLINE NEURAL ACOUSTIC PROCESSING*
- Guy Wolf, Stéphane Mallat, and Shihab Shamma. *Rigid Motion Model for Audio Source Separation*. IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 64, NO. 7, APRIL 1, 2016

A rellenar sólo en el caso que el alumno sea quien realice la propuesta de TFG
Alumno/a propuesto/a:

Granada, 14 de Mayo de 2018

Sello del Departamento

Campus Fuentenueva
Avda. Fuentenueva s/n
18071 Granada

Tfno. +34 958242902
fjruiz@ugr.es

Comisión Docente de Físicas
Facultad de Ciencias

Firmado por: FERNANDO JOSE ROJAS RUIZ Secretario/a de Departamento

Sello de tiempo: 16/05/2018 10:50:13 Página: 2 / 2



cfP6wllk8uS1BJA4HsbZvn5CKCJ3NmbA

La integridad de este documento se puede verificar en la dirección <https://sede.ugr.es/verifirma/pfinicio.jsp> introduciendo el código de verificación que aparece debajo del código de barras.