



## Guía docente

### 230622 - DSAP - Procesado Digital de Audio y Voz

Última modificación: 29/04/2020

**Unidad responsable:** Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona  
**Unidad que imparte:** 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

**Titulación:** MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN (Plan 2013). (Asignatura optativa).  
MÁSTER UNIVERSITARIO EN TECNOLOGÍAS AVANZADAS DE TELECOMUNICACIÓN (Plan 2019).  
(Asignatura optativa).

**Curso:** 2020      **Créditos ECTS:** 5.0      **Idiomas:** Inglés

#### PROFESORADO

---

**Profesorado responsable:** Climent Nadeu

**Otros:**

#### CAPACIDADES PREVIAS

---

Conocimientos avanzados de Señales, Sistemas y Transformadas  
Conocimientos básicos de Probabilidad, Variables Aleatorias y Procesos Estocásticos  
Experiencia de programación en Matlab  
Recomendado:  
- Conocimientos básicos de aprendizaje automático  
- Lenguaje Python

#### REQUISITOS

---

Dos asignaturas del área de Señales, Sistemas y Transformadas  
Al menos una asignatura de Probabilidad, Variables Aleatorias y Procesos Estocásticos

#### COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

---

**Específicas:**

1. Capacidad para aplicar métodos de la teoría de la información, la modulación adaptativa y codificación de canal, así como técnicas avanzadas de procesamiento digital de señal a los sistemas de comunicaciones y audiovisuales.

**Transversales:**

2. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

3. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad, y valorar de forma crítica los resultados de dicha gestión.

4. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.



## METODOLOGÍAS DOCENTES

- Clases expositivas (50%)
- Clases de aplicación (con Matlab o similar) (50%)
- Trabajo en equipo: proyecto de asignatura, clases de aplicación
- Trabajo individual: preparación y finalización (fuera del aula) de las actividades de aplicación

## OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

Objetivos de aprendizaje:

Adquirir la comprensión y ser competente en un conjunto relevante de conceptos y técnicas del campo del procesamiento digital de audio y su aplicación a los problemas que aparecen en aplicaciones reales. Es especial, se considerarán las señales y aplicaciones de habla y música.

Resultados de aprendizaje:

La capacidad de procesar digitalmente señales de audio, en un contexto orientado a aplicación, a fin de realizar análisis, modelado, extracción de información, limpieza, modificación, y síntesis / generación.

## HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo grande	39,0	31.20
Horas aprendizaje autónomo	86,0	68.80

**Dedicación total:** 125 h

## CONTENIDOS

### Introducción

**Descripción:**

Presentación de la asignatura  
La diversidad de las señales de audio  
Características del habla y la música. El modelo de producción  
Audición y modelado auditivo  
STFT y espectrograma  
La transformada de Fourier a corto plazo

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h

### Modelado y representación de señales de audio

**Descripción:**

Modelado con sólo polos basado en producción  
Determinación del tono para habla y música  
Codificación basada en LPC que se usa en telefonía móvil

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h



### Mejora de señales de voz y audio

**Descripción:**

Reducción de ruido: substracción espectral, filtrado de Wiener, redes neuronales  
Separación ciega de fuentes: ICA, NMF  
Cancelación: eco, interferencias

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h

### Reconocimiento y detección de audio y habla

**Descripción:**

Enfoques basados en ajuste dinámico de patrones  
Enfoques estadísticos y de aprendizaje profundo  
Detección de actividad de audio  
Aplicación al reconocimiento de habla y de sonidos

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h

### Procesado multimicrófono de audio

**Descripción:**

Acústica de salas  
Conformación de haz con agrupación de micrófonos  
Localization de fuente acústica y seguimiento

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h

### Análisis y síntesis de señales de audio

**Descripción:**

Análisis y síntesis localizadas de señales (casi)periódicas. Modificación de la escala temporal y el tono  
Síntesis espacial de audio con funciones HRTF

**Dedicación:** 12h

Grupo grande/Teoría: 6h  
Aprendizaje autónomo: 6h

### Realización y presentación del proyecto

**Descripción:**

Diseño, implementación i prueba de un sistema de proc. de audio para una aplicación específica  
Presentación oral de 1) propuesta, 2) revisión crítica, y 3) realización y conclusiones  
Escritura de un informe tipo artículo en congreso

**Dedicación:** 55h

Grupo grande/Teoría: 4h  
Aprendizaje autónomo: 51h



## ACTIVIDADES

---

### Ejercicios prácticos con Matlab o similar, aprox. 50% del tiempo en el aula

**Dedicación:** 20h  
Grupo grande/Teoría: 20h

### Tests cortos al final de cada tema

**Dedicación:** 1h  
Grupo grande/Teoría: 1h

### Proyecto de asignatura

**Descripción:**

Realización y presentación de un proyecto en equipo, que incluye trabajo experimental en procesamiento de audio, y se presenta tanto de forma oral como escrita.

**Dedicación:** 60h  
Grupo grande/Teoría: 60h

### Presentaciones orales

**Descripción:**

Presentación oral del proyecto en tres etapas: propuesta, revisión y final.

**Dedicación:** 0h 40m  
Grupo grande/Teoría: 0h 40m

## SISTEMA DE CALIFICACIÓN

---

Asistencia/participación en clase (10%)  
Tests cortos cada dos semanas (30%)  
Proyecto de asignatura (60%)

## BIBLIOGRAFÍA

---

**Básica:**

- Quatieri, T.F. Discrete-time speech signal processing: principles and practice. Upper Saddle River, NJ: Prentice Hall, 2002. ISBN 013242942X.
- Gold, B.; Morgan, N.; Ellis, D. Speech and audio signal processing: processing and perception of speech and music. 2nd rev. ed. Wiley-Blackwell, 2011. ISBN 978-0-470-19536-9.
- Dutoit, T.; Marqués, F.; Rabiner, L.R. Applied signal processing: a MATLAB-based proof of concept [en línea]. New York ; London: Springer, 2009 [Consulta: 10/05/2020]. Disponible a: <http://dx.doi.org/10.1007/978-0-387-74535-0>. ISBN 978-0-38774534-3.

**Complementaria:**

- Rabiner, L.R.; Schafer, R.W. Theory and applications of digital speech processing. Prentice Hall, 2010. ISBN 9780136034285.
- Huang, Y.A.; Benesty, J. (eds.). Audio signal processing for next-generation multimedia communication systems [en línea]. New York: Kluwer Academic Publishing, 2004 [Consulta: 23/07/2013]. Disponible a: <http://link.springer.com/book/10.1007/b117685/page/1>. ISBN 1402077688.



## RECURSOS

---

### Material audiovisual:

- Slides. Diapositivas usadas en clase

### Material informático:

- Codi programes. Código programas en Matlab o similar

### Otros recursos:

Presentaciones usadas en las clases

Enunciados y programas del trabajo práctico