

ANX-PR/CL/001-01
GUÍA DE APRENDIZAJE

ASIGNATURA

Tratamiento digital de voz y audio

CURSO ACADÉMICO - SEMESTRE

2016-17 - Primer semestre

Datos Descriptivos

Nombre de la Asignatura	Tratamiento digital de voz y audio
Titulación	09TT - Grado en Ingeniería de Tecnologías y Servicios de Telecomunicacion
Centro responsable de la titulación	Escuela Técnica Superior de Ingenieros de Telecomunicacion
Semestre/s de impartición	Séptimo semestre
Módulo	Mod tecnol esp sonido e imagen
Materia	Tecnología específica sonido e imagen
Carácter	Optativa
Código UPM	95000070
Nombre en inglés	Speech and audio signal processing

Datos Generales

Créditos	6	Curso	4
Curso Académico	2016-17	Período de impartición	Septiembre-Enero
Idioma de impartición	Castellano	Otros idiomas de impartición	

Requisitos Previos Obligatorios

Asignaturas Previas Requeridas

El plan de estudios Grado en Ingeniería de Tecnologías y Servicios de Telecomunicacion no tiene definidas asignaturas previas superadas para esta asignatura.

Otros Requisitos

El plan de estudios Grado en Ingeniería de Tecnologías y Servicios de Telecomunicacion no tiene definidos otros requisitos para esta asignatura.

Conocimientos Previos

Asignaturas Previas Recomendadas

Señales y sistemas

Señales aleatorias

Tratamiento digital de señales

Otros Conocimientos Previos Recomendados

El coordinador de la asignatura no ha definido otros conocimientos previos recomendados.



CAMPUS
DE EXCELENCIA
INTERNACIONAL

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros de Telecomunicación
PROCESO DE COORDINACIÓN DE LAS ENSEÑANZAS

ANX-PR/CL/001-01: GUÍA DE APRENDIZAJE



Código PR/CL/001

Competencias

CE-SI1 - Capacidad de construir, explotar y gestionar servicios y aplicaciones de telecomunicaciones, entendidas éstas como sistemas de captación, tratamiento analógico y digital, codificación, transporte, representación, procesado, almacenamiento, reproducción, gestión y presentación de servicios audiovisuales e información multimedia.

CE-SI5 - Capacidad para crear, codificar, gestionar, difundir y distribuir contenidos multimedia, atendiendo a criterios de usabilidad y accesibilidad de los servicios audiovisuales, de difusión e interactivos

CG1 - Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio

CG10 - Creatividad

CG12 - Organización y planificación

CG2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio

CG3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética

CG4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado

CG5 - Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

CG7 - Trabajo en equipo

CG8 - Comunicación oral y escrita

CG9 - Uso de Tecnologías de la Información y de las Comunicaciones

Resultados de Aprendizaje

RA523 - Capacidad para aplicar diferentes técnicas de tratamiento digital de señales de voz y audio en servicios y aplicaciones de telecomunicaciones: codificación, síntesis, reconocimiento y sistemas interactivos

RA82 - Capacidad para crear, codificar, gestionar, difundir y distribuir contenidos multimedia, atendiendo a criterios de usabilidad y accesibilidad de los servicios audiovisuales, de difusión e interactivos.

RA78 - Capacidad de construir, explotar y gestionar servicios y aplicaciones de captación, tratamiento analógico y digital, codificación, transporte, representación, procesado, almacenamiento, reproducción, gestión y presentación de servicios audiovisuales e información multimedia.

Profesorado

Profesorado

Nombre	Despacho	e-mail	Tutorías
Hernandez Gomez, Luis Alfonso (Coordinador/a)	C-330	luisalfonso.hernandez@upm.es	M - 15:00 - 18:00 X - 15:00 - 18:00
Lopez Gonzalo, Eduardo	C-330	eduardo.lopez@upm.es	L - 11:00 - 14:00 X - 10:00 - 13:00

Nota.- Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

Descripción de la Asignatura

Esta asignatura ofrece una perspectiva global de las diferentes técnicas del tratamiento digital de señales de voz y audio y su utilización en servicios y aplicaciones de telecomunicaciones: codificación, síntesis, reconocimiento y sistemas interactivos.

La asignatura propone una metodología de trabajo esencialmente práctica que parte del estudio de las características básicas de las señales de voz y audio, asociadas a sus mecanismos de producción y percepción. Se estudian las principales técnicas de análisis en tiempo y en frecuencia de señales de voz y audio, y se plantea su aplicación a sistemas de codificación, síntesis y reconocimiento. Las técnicas de codificación de voz y audio se analizan en el contexto de los diferentes servicios y aplicaciones de telecomunicaciones fijos, móviles e Internet, y se presentan las técnicas objetivas de medida de calidad de voz y audio. También se proporciona una visión global de las tecnologías y ámbitos de aplicación de sistemas de reconocimiento y síntesis de voz y audio, así como el diseño de sistemas interactivos y su evaluación desde perspectivas de la experiencia de uso y accesibilidad.

Temario

1. Caracterización, Síntesis y Percepción de señales de voz y audio
 - 1.1. Introducción al tratamiento de voz y audio y sus aplicaciones
 - 1.2. Caracterización de señales de voz y audio
 - 1.2.1. Producción de voz y audio
 - 1.2.2. Caracterización mediante representaciones tiempo-frecuencia
 - 1.3. Modelos básicos de síntesis de voz y audio
 - 1.4. Revisión del mecanismo de percepción de sonidos
 - 1.4.1. Sistema auditivo periférico
 - 1.4.2. Enmascaramiento tiempo y frecuencia
 - 1.4.3. Modelos psicoacústicos
2. Técnicas de tratamiento de señales de voz y audio
 - 2.1. Análisis localizado en tiempo
 - 2.1.1. Energía localizada, tasa de cruces por cero y detección de actividad vocal
 - 2.2. Análisis localizado en frecuencia y bancos de filtros
 - 2.3. Autocorrelación, Predicción Lineal
 - 2.4. Análisis armónico, pitch
3. Codificación de Voz y Audio
 - 3.1. Esquemas de codificación de voz: ITU, AMR, G729, iLBC, iSAC, SILK
 - 3.2. Esquemas de codificación de audio: MPEG, AAC, Dolby,...
 - 3.3. Convergencia codificación de voz y audio: Full-HD Voice, VoLTE, Opus, EVS, HE-AAC

4. Introducción al reconocimiento de voz y audio y aplicaciones interactivas
 - 4.1. Reconocimiento de voz y audio. habla, idioma, locutor, biometría, voces patológicas biometría, eventos acústicos.
 - 4.2. Sistemas de diálogo con interacción hablada y audio - Diseño, experiencia de usuario y accesibilidad
5. Practicas de Tratamiento Digital de Voz y Audio
 - 5.1. Caracterización acústica y perceptual, y síntesis en tiempo y frecuencia de señales de voz y audio
 - 5.2. Técnicas básicas de codificación de voz y audio

Cronograma

Horas totales: 62 horas

Horas presenciales: 62 horas (39.7%)

Peso total de actividades de evaluación continua:
100%

Peso total de actividades de evaluación sólo prueba final:
100%

Semana	Actividad Presencial en Aula	Actividad Presencial en Laboratorio	Otra Actividad Presencial	Actividades Evaluación
Semana 1	<p>Tema 1.1 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Tema 1.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
Semana 2	<p>Tema 1.3 Duración: 01:30 LM: Actividad del tipo Lección Magistral</p> <p>Tema 1.4 Duración: 00:30 LM: Actividad del tipo Lección Magistral</p>	<p>Practica 1: Caracterización voz y audio Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
Semana 3	<p>Tema 2.1 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Tema 2.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
Semana 4		<p>Practica 1: Representaciones tiempo-frecuencia Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
Semana 5		<p>Practica 1: Síntesis tiempo y frecuencia Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		<p>Evaluación Temas 1 y 2 (hasta 2.1) Duración: 04:00 Pl: Técnica del tipo Presentación Individual Evaluación continua Actividad presencial</p>
Semana 6	<p>Temas 2.3 y 2.4 Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
Semana 7	<p>Tema 3.1 Duración: 04:00 LM: Actividad del tipo Lección Magistral</p>	<p>Practica 2: Estimación de frecuencia fundamental y detección de actividad Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
Semana 8		<p>Práctica 2: Codificación de voz basada en LPC Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		

Semana 9	Tema 3.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral	Practica 2: Tecnicas coffificación audio Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 10	Tema 3.3 Duración: 02:00 LM: Actividad del tipo Lección Magistral	Practica 2: Tecnicas coffificación audio Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 11	Tema 4.1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			Evaluación hasta Tema 3 (incluido) Duración: 02:00 PI: Técnica del tipo Presentación Individual Evaluación continua Actividad presencial
Semana 12	Tema 4.2 Duración: 02:00 LM: Actividad del tipo Lección Magistral	Laboratorio: actividades Proyecto Final Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 13		Laboratorio: actividades Proyecto Final Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio		
Semana 14				Presentación Proyecto Final Duración: 04:00 PI: Técnica del tipo Presentación Individual Evaluación continua Actividad presencial
Semana 15				
Semana 16				
Semana 17				Examen Final Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial Evaluación actividad práctica Duración: 02:00 PI: Técnica del tipo Presentación Individual Evaluación sólo prueba final Actividad presencial Examen Final Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación sólo prueba final Actividad presencial

Nota.- El cronograma sigue una planificación teórica de la asignatura que puede sufrir modificaciones durante el curso.

Nota 2.- Para poder calcular correctamente la dedicación de un alumno, la duración de las actividades que se repiten en el tiempo (por ejemplo, subgrupos de prácticas") únicamente se indican la primera vez que se definen.

Actividades de Evaluación

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
5	Evaluación Temas 1 y 2 (hasta 2.1)	04:00	Evaluación continua	PI: Técnica del tipo Presentación Individual	Sí	20%	3.5 / 10	CG9, CG5, CG4, CG3, CG2, CG1, CG10, CG12, CE-SI1, CG7, CG8
11	Evaluación hasta Tema 3 (incluido)	02:00	Evaluación continua	PI: Técnica del tipo Presentación Individual	Sí	20%	3.5 / 10	CG7, CE-SI5, CG8, CG9, CG5, CG4, CG3, CG2, CG1, CG10, CG12, CE-SI1
14	Presentación Proyecto Final	04:00	Evaluación continua	PI: Técnica del tipo Presentación Individual	Sí	20%	3.5 / 10	CG7, CE-SI5, CG8, CG9, CG5, CG4, CG3, CG2, CG1, CG10, CG12, CE-SI1
17	Examen Final	02:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	40%	3.5 / 10	CE-SI5, CG8, CG5, CG3, CG2, CG1, CG12, CE-SI1
17	Evaluación actividad práctica	02:00	Evaluación sólo prueba final	PI: Técnica del tipo Presentación Individual	Sí	50%	3.5 / 10	CG7, CE-SI5, CG8, CG9, CG5, CG4, CG3, CG2, CG1, CG10, CG12, CE-SI1
17	Examen Final	02:00	Evaluación sólo prueba final	EX: Técnica del tipo Examen Escrito	Sí	50%	3.5 / 10	CE-SI5, CG8, CG5, CG3, CG2, CG1, CG12, CE-SI1

Criterios de Evaluación

La forma de calificación más adecuada para esta asignatura es la evaluación continua, pero, en cumplimiento de la Normativa de Evaluación de la UPM, los alumnos que lo deseen podrán ser evaluados mediante un único examen final, siempre y cuando lo comuniquen al Director del Departamento de Señales, Sistemas y Radiocomunicaciones en plazo y forma (hasta el último día de clase del segundo semestre, y mediante solicitud presentada en el registro de la ETSIT).

La evaluación continua supondrá ir realizando actividades de laboratorio (Práctica 1, Práctica 2 y Proyecto Final) que serán evaluadas mediante presentaciones orales individuales en clase (aunque las actividades se realicen en grupo). Al final de curso, junto a la presentación del Proyecto Final de la asignatura, se realizará un examen escrito. La nota final será la suma de 20% (Práctica 1) + 20% (Práctica 2) + 20% (Proyecto Final) + 40% Examen Final. Siendo necesario que en cada prueba se supere una calificación mínima de 3.5 puntos.

Los alumnos que no sigan la evaluación continua deberán presentarse a un Examen Final escrito y elaborar un proyecto de actividades prácticas que deberán exponer de forma oral el mismo día del examen escrito. La nota final será 50% la calificación del Examen Final + 50% de la actividad práctica que presenten. Siendo nuevamente necesario que en cada prueba se supere una calificación mínima de 3.5 puntos.

La evaluación en la convocatoria extraordinaria se realizará mediante un Examen Final escrito y la presentación oral de una actividad práctica. La nota final será 50% la calificación del Examen Final + 50% de la actividad práctica siempre que en cada prueba se supere una calificación mínima de 3.5 puntos.

Recursos Didácticos

Descripción	Tipo	Observaciones
Presentaciones de la asignatura	Recursos web	Disponible en Moodle UPM
Speech and Audio Signal Processing, B. Gold, N. Morgan, and D. Ellis, 2nd edition, Wiley Press 2011.	Bibliografía	Libro de referencia para la asignatura
Spoken Language Processing, X. Huang, A. Acero, and H. Won, Prentice Hall, 2001	Bibliografía	Referencias para tratamiento de habla
Introduction to Data Compression, Khalid Sayood, Fourth Edition, The Morgan Kaufmann Series in Multimedia Information and Systems, 2012	Bibliografía	Referencia para codificación de voz y audio
Laboratorio A-202-L ETSIT-UPM	Equipamiento	Sala de trabajo para las actividades prácticas de la asignatura