

Tratamiento digital de voz y audio

Guía de Aprendizaje – Información al estudiante

1.Datos Descriptivos

Asignatura	Tratamiento digital de voz y audio
Materia	M12 TECNOLOGÍA ESPECÍFICA DE SONIDO E IMAGEN
Departamento responsable	Señales, Sistemas y Radiocomunicaciones
Créditos ECTS	6
Carácter	Específico
Titulación	Graduado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Curso	4º
Especialidad	Sonido e Imagen
Curso académico	2014-2015
Semestre en que se imparte	Primero
Idioma en que se imparte	Castellano
Página Web	http://moodle.upm.es/titulaciones/oficiales

2. Profesorado

NOMBRE Y APELLIDO	DESPACHO	Correo electrónico
Luis Hernández Gómez (Coordinador)	C-330	luisalfonso.hernandez@upm.es
Eduardo López Gonzalo	C-330	eduardo.lopez@upm.es

3. Conocimientos previos requeridos para poder seguir con normalidad la asignatura

Asignaturas superadas	N/A
Otros resultados de aprendizaje necesarios	Conocimientos de análisis y tratamiento de señales y sistemas impartidos en las asignaturas de segundo curso Señales y Sistemas y Señales Aleatorias, así como conocimiento de técnicas y manejo de herramientas de programación de tratamiento digital impartidas en la asignatura de tercer curso Tratamiento Digital de Señales.

4. Objetivos de Aprendizaje.

COMPETENCIAS ASIGNADAS A LA ASIGNATURA Y SU NIVEL DE ADQUISICIÓN		
Código	Competencia	Nivel
CG1- CG13	Todas las asignaturas del Plan de Estudios contribuyen en mayor o menor medida a la consecución de las competencias generales del perfil de egreso	1
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.	3
CE-SI2	Capacidad de analizar, especificar, realizar y mantener sistemas, equipos, cabeceras e instalaciones de televisión, audio y vídeo, tanto en entornos fijos como móviles.	2
CE-SI3	Capacidad para realizar proyectos de locales e instalaciones destinados a la producción y grabación de señales de audio y vídeo.	2
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.	3

LEYENDA: Nivel de adquisición 1: Básico
 Nivel de adquisición 2: Medio
 Nivel de adquisición 3: Avanzado



POLITÉCNICA



RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA			
Código	Resultado de aprendizaje	Competencias asociadas	Nivel de adquisición
RA1	Adquirir conocimientos sobre las diferentes áreas del tratamiento digital de señales de voz y audio y su utilización en servicios y aplicaciones de telecomunicaciones: codificación, síntesis, reconocimiento y sistemas interactivos.	CESI1, 2, 3, 5	3
RA2	Conocer las características básicas de las señales de voz y audio, asociadas a sus mecanismos de producción y percepción, comprendiendo las técnicas de síntesis de voz y audio, y sabiendo utilizar herramientas de captura, grabación, reproducción y análisis tiempo-frecuencia.	CESI1, 2, 3, 5	3
RA3	Conocer las técnicas de análisis localizado tanto en tiempo como en frecuencia de señales de voz y audio, sabiendo utilizar herramientas para su implementación y comprendiendo su aplicación a sistemas de codificación, síntesis y reconocimiento.	CESI1, 2, 3, 5	3
RA4	Métodos de codificación de fuente y codificación perceptual, y su aplicación en esquemas de codificadores de voz y audio para diferentes servicios y aplicaciones de telecomunicaciones fijos, móviles e Internet, conociendo también las principales técnicas objetivas de medida de calidad de voz y audio.	CESI1, 2, 3, 5	2
RA5	Métodos y ámbitos de aplicación de sistemas de reconocimiento de voz y audio, y sistemas interactivos, conociendo tanto sus principios de diseño como metodologías de evaluación desde perspectivas de usabilidad y accesibilidad	CESI1, 2, 3, 5	1

LEYENDA: Nivel de adquisición 1: Conocimiento descriptivo
 Nivel de adquisición 2: Comprensión/Aplicación
 Nivel de adquisición 3: Análisis/Síntesis/Implementación

5. Sistema de evaluación de la asignatura

INDICADORES DE LOGRO		
Ref	Indicador	Relacionado con RA
I1	Poseer una perspectiva global de las diferentes áreas del tratamiento digital de señales de voz y audio y su utilización en servicios y aplicaciones de telecomunicaciones: codificación, síntesis, reconocimiento y sistemas interactivos.	RA1
I2	Conocer las características básicas de las señales de voz y audio, su captura y grabación, sabiendo asociar sus mecanismos de producción al análisis basado en representaciones tiempo-frecuencia.	RA2
I3	Conocer los principales mecanismos de generación de señales de voz y audio a partir de modelos básicos de síntesis de voz y audio.	RA2
I4	Conocer los procesos de percepción de sonidos, enmascaramiento de sonidos tanto en el dominio del tiempo como en frecuencia, y la definición de modelos psicoacústicos.	RA2
I5	Dominar técnicas de análisis de señales de voz y audio en el tiempo, conociendo los principios del análisis localizado y la utilización de técnicas específicas como: detección de actividad, control de dinámica, igualación y efectos de sonido.	RA3
I6	Dominar técnicas de análisis de voz y audio en dominios transformados: transformadas localizadas y bancos de filtros, análisis de autocorrelación, Predicción Lineal, y análisis de la estructura armónica y pitch.	RA3
I7	Dominar los conceptos básicos de codificación de fuente y codificación perceptual y conocer su aplicación al diseño de codificadores de señales de voz y audio.	RA4
I8	Conocer los esquemas básicos de los principales codificadores de voz y audio y su utilización en diferentes servicios y aplicaciones de telecomunicaciones sobre entornos de telefonía fija, móvil e Internet. Conociendo también las principales técnicas objetivas de medida de calidad de voz y audio.	RA4, RA1
I9	Conocer los principales mecanismos de extracción de información sobre señales de voz y audio a través de técnicas de reconocimiento de voz y audio, siendo capaz de identificar sus ámbitos de aplicación más relevantes.	RA5, RA1



POLITÉCNICA



INDICADORES DE LOGRO		
Ref	Indicador	Relaciona- do con RA
I10	Conocer la integración de técnicas de síntesis y reconocimiento para el desarrollo de sistemas con interacción por voz y audio, siendo capaz de abordar su diseño y evaluación desde perspectivas tanto de usabilidad como de accesibilidad.	RA5, RA1

EVALUACION SUMATIVA			
Breve descripción de las actividades evaluables	Momento	Lugar	Peso en la calif.
Evaluación Tema 1 + Práctica 1	Semana 5	Aula	25
Evaluación Tema 2 + Práctica 2	Semana 10	Aula	25
Evaluación Tema 3 + Práctica 3	Semana 14	Aula	25
Evaluación Tema 4 + Proyecto Final	Convocatoria oficial	Franja horaria especial	25
			Total: 100%

CRITERIOS DE CALIFICACIÓN

Los alumnos serán evaluados, por defecto, mediante evaluación continua. La calificación de la asignatura se realizará del siguiente modo:

NOTA FINAL = 25% Examen Tema 1 y Práctica 1+ 25% Examen Tema 2 y Práctica 2 + 25% Examen Tema 3 y Práctica 3 +25% Examen Tema 4 +Proyecto final

En cada uno de los exámenes se evaluará tanto la parte teórica como la parte práctica correspondiente a las actividades realizadas en el laboratorio, cada una de estas partes supondrá un 50% de la nota del examen. Para la parte teórica se exigirá una nota mínima de 4.0 sobre 10.0 para poder hacer media con la parte práctica.

El proyecto final será una actividad libre que podrá ser tanto de carácter práctico como teórico. Los alumnos deberán plantear y desarrollar este proyecto a lo largo de todo el curso y presentar sus resultados al final del mismo.

Los alumnos que no hayan superado alguno de los exámenes podrán presentarse a una segunda convocatoria en la fecha de la exposición del proyecto final de la asignatura.

La asistencia al Laboratorio es obligatoria tanto para los alumnos que elijan evaluación continua como para los que se presenten sólo examen final. 2 faltas de asistencia sin justificar suponen el suspenso del Laboratorio.

En cumplimiento de la Normativa de Evaluación de la Universidad Politécnica de Madrid, los alumnos que lo deseen serán evaluados mediante un único examen final siempre y cuando lo comuniquen al Director del Departamento de Señales, Sistemas y Radiocomunicaciones mediante solicitud presentada en el registro de la Escuela Técnica Superior de Ingenieros de Telecomunicación antes del día 21 de octubre de 2013. Esta opción supone la renuncia a la evaluación continua.

6. Contenidos y Actividades de Aprendizaje

CONTENIDOS ESPECÍFICOS		
Bloque / Tema / Capítulo	Apartado	Indicadores Relacionados
Tema 1: Caracterización, Síntesis y Percepción de señales de voz y audio	1.1 Introducción al tratamiento de voz y audio y sus aplicaciones	I1
	1.2 Caracterización de señales de voz y audio - Producción de voz y audio - Caracterización mediante representaciones tiempo-frecuencia	I2
	1.3 Modelos básicos de síntesis de voz y audio	I3
	1.4 Percepción de sonidos - Sistema auditivo periférico - Enmascaramiento tiempo y frecuencia - Modelos psicoacústicos	I4
Tema 2: Técnicas de tratamiento de señales de voz y audio	2.1. Análisis localizado en tiempo: - Energía localizada, tasa de cruces por cero y detección de actividad vocal.	I5
	2.2 Análisis localizado en frecuencia y bancos de filtros	I6
	2.3 Autocorrelación, Predicción Lineal	
	2.4. Análisis armónico, pitch.	



POLITÉCNICA



CONTENIDOS ESPECÍFICOS

Bloque / Tema / Capítulo	Apartado	Indicadores Relacionados
Tema 3: Codificación de Voz y Audio	3.1. Codificación de fuente y codificación perceptual	17
	3.2 Esquemas básicos de codificación: - Codificadores de voz: ITU, 3GPP, Internet - Codificadores de audio: ISO MPEG, Dolby.	18
	3.3 Medidas objetivas de calidad	
Tema 4: Reconocimiento de voz y audio y aplicaciones interactivas	4.1 Reconocimiento de voz y audio. habla, idioma, locutor, biometría, voces patológicas biometría, eventos acústicos.	19
	4.2 Sistemas de diálogo con interacción hablada y audio - Diseño, evaluación de usabilidad y accesibilidad	110
Laboratorio de Tratamiento Digital de Voz y Audio	Práctica 1: Caracterización, Síntesis y Percepción de señales de voz y audio.	
	Introducción a la captura, grabación y análisis tiempo-frecuencia de señales de voz y audio, estudiando su asociación con los mecanismos de producción. Se realizarán experimentos simples de síntesis por concatenación de forma de onda y de percepción de sonidos.	11 a 14
	Práctica 2: Técnicas de análisis localizado: detección de actividad, efectos, autocorrelación y pitch. Realización de diversos ejercicios dirigidos a desarrollar algoritmos específicos de análisis localizado: detección de actividad, efectos acústicos, y estimación de pitch a partir de la función de autocorrelación.	15, 16
	Práctica 3: Análisis en frecuencia y Predicción Lineal Se realizarán ejercicios de análisis/síntesis de señales de voz y audio en el dominio transformado y mediante uso de Predicción Lineal. Se experimentará tanto considerando su aplicación al ámbito de la codificación, como a la identificación de patrones simples de voz y audio como aproximación a técnicas de reconocimiento.	16 a 19

7. Breve descripción de las modalidades organizativas utilizadas y de los métodos de enseñanza empleados

CLASES DE TEORIA	Se exponen en ella los conceptos y herramientas del tratamiento digital de señales de voz y audio
CLASES DE PROBLEMAS	Durante las clases presenciales, además de los desarrollos teóricos, se resolverán ejercicios prácticos entresacados entre los propuestos para que el alumno resuelva como trabajo personal
PRÁCTICAS	En el laboratorio se desarrollan de forma práctica los conceptos de análisis y tratamiento digital de señales de voz y audio vistos en las clases presenciales. El alumno debe llevar preparados los conceptos teóricos a implementar en la práctica correspondiente
TRABAJOS AUTONOMOS	Estarán relacionados con las prácticas de laboratorio y especialmente con el proyecto final de la asignatura.
TRABAJOS EN GRUPO	Las prácticas de laboratorio y el trabajo final podrán realizarse en grupos de 2 personas
TUTORÍAS	Las tutorías se ajustarán a la normativa vigente

8. Recursos didácticos

RECURSOS DIDÁCTICOS	
BIBLIOGRAFÍA	'Speech and Audio Signal Processing', B. Gold, N. Morgan, and D. Ellis, 2nd edition, Wiley Press 2011.
	'Spoken Language Processing', X. Huang, A. Acero, and H. Won , Prentice Hall, 2001
	'Introduction to Data Compression', Khalid Sayood, Fourth Edition, The Morgan Kaufmann Series in Multimedia Information and Systems, 2012
RECURSOS WEB	Página web de la asignatura http://moodle.upm.es/titulaciones/oficiales
EQUIPAMIENTO	Laboratorio de señales A.202-L
	Aula : Asignada por Jefatura de Estudios
	Sala de trabajo en grupo: Laboratorio A.202-L

9. Cronograma de trabajo de la asignatura

Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 1 (6 h)	- Tema 1,1 (1 h) - Tema 1.2. (3 h)	•	Estudio de teoría y solución de problemas propuestos (2 h)	•	•	•
Semana 2 (8 h)	- Tema 1.3 (2 h) - Tema 1.4 (2 h)	•	Estudio de teoría y solución de problemas propuestos (4 h)	•	•	•
Semana 3 (8 h)	- Tema 1.4 (1.5 h) - Presentación Práctica 1 (0.5 h)	•	- Estudio de teoría y solución de problemas propuestos (2 h) - Preparación de la practica (2 h)	•	•	•
Semana 4 (9 h)	•	• Práctica 1 (4 h)	- Preparación de la práctica (3 h) - Preparación examen parcial (2 h)	• Trabajo en el laboratorio (4 h)	•	•



POLITÉCNICA



Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 5 (9 h)	Evaluación Tema 1 y Práctica 1 (2 h)	Práctica 1 (2 h)	- Preparación examen parcial (3 h) - Preparación examen practica (2 h)	Trabajo en el laboratorio (2 h)	Evaluación Tema 1 y Práctica 1	
Semana 6 (8 h)	Tema 2.1. (4 h)		Estudio de teoría y solución de problemas propuestos (4 h)			
Semana 7 (9 h)	Tema 2.2 (4 h)		Estudio de teoría y solución de problemas propuestos (5 h)			
Semana 8 (9 h)	- Tema 2.3 (2 h) - Tema 2.4 (1.5 h) - Presentación Práctica 2 (0.5 h)		- Estudio de teoría, solución de problemas propuestos (3 h) - Preparación de la practica (2 h)			
Semana 9 (9 h)		Práctica 2 (4 h)	- Preparación de la práctica (3 h) - Preparación examen parcial (2 h)	Trabajo en el laboratorio (4 h)		
Semana 10 (9 h)	Evaluación Tema 2 y Práctica 2 (2 h)	Práctica 2 (2 h)	- Preparación examen parcial (3 h) - Preparación examen practica (2 h)	Trabajo en el laboratorio (2 h)	Evaluación Tema 2 y Práctica 2	



POLITÉCNICA



Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 11 (9 h)	Tema 3.1 (4 h)		Estudio de teoría y solución de problemas propuestos (5 h)			
Semana 12 (10 h)	- Tema 3.2 (2 h) - Tema 3.3 (1.5 h) - Presentación Práctica 2 y Proyecto Final (0.5 h)		- Estudio de teoría, solución de problemas propuestos (3 h) - Preparación de la practica (2 h) - Preparación proyecto final (1 h)			
Semana 13 (10 h)		Práctica 3 (4 h)	- Preparación de la práctica (3 h) - Preparación examen parcial (2 h) - Preparación proyecto final (1 h)	Trabajo en el laboratorio (4 h)		
Semana 14 (10 h)	Evaluación Tema 3 y Práctica 3 (2 h)	Práctica 3 (2 h)	- Preparación examen parcial (3 h) - Preparación examen practica (2 h) - Preparación proyecto final (1 h)	Trabajo en el laboratorio (2 h)	Evaluación Tema 3 y Práctica 3	

Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 15 (9 h)	- Tema 4.1. (2 h) - Tema 4.2 (2 h)		- Estudio de teoría y solución de problemas propuestos (2 h) - Preparación proyecto final (3 h)	•	•	•
Periodo hasta el día del examen en enero (9,5 h)	•	•	- Preparación del examen (3 h) - Preparación proyecto final (5 h)	•	• Evaluación Final (Tema 4 y Proyecto Final) (1,5 h)	•

Nota: Para cada actividad se especifica la dedicación en horas que implica para el alumno.

Las semanas reseñadas lo son de docencia efectiva (no las semanas de calendario)