



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

GUÍA DOCENTE DE TRATAMIENTO DE SEÑALES DE VOZ Y AUDIO

La presente guía docente corresponde a la asignatura Tratamiento de Señales de Voz y Audio (TSVA), aprobada para el curso lectivo 2017-2018 en Junta de Centro y publicada en su versión definitiva en la página web de la Escuela Politécnica Superior. La guía docente de TSVA aprobada y publicada antes del periodo de matrícula tiene el carácter de contrato con el estudiante.



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

ASIGNATURA

TRATAMIENTO DE SEÑALES DE VOZ Y AUDIO (TSVA)

1.1. Código

18499 del Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

1.2. Materia

Sonido e Imagen

1.3. Tipo

Formación de Tecnología Específica

1.4. Nivel

Grado

1.5. Curso

4º

1.6. Semestre

1º

1.7. Número de créditos

6 créditos ECTS

1.8. Requisitos previos

Tratamiento de Señales de Voz y Audio forma parte de la *Materia 3.2 (Sonido e Imagen)* del módulo de *Formación de Tecnología Específica en Sonido e Imagen* del plan de estudios del Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación.

Esta materia está desglosada en ocho asignaturas semestrales (*Sistemas y Servicios de Audio y Vídeo, Ingeniería Acústica, Tratamiento de Señales Multimedia, Tecnologías de Vídeo, Tecnologías de Audio, Tratamiento de Señales Visuales, Tratamiento de Señales de Voz y Audio, Televisión Digital*) que presentan desde



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

fundamentos a técnicas avanzadas en análisis, tratamiento y diseño de señales y sistemas de audio y video.

La asignatura de *Tratamiento de Señales de Voz y Audio* precisa de conocimientos básicos presentados con anterioridad por un lado en las asignaturas de *Sistemas Lineales* (2º curso, 1º semestre), *Tratamiento Digital de Señales* (3º curso, 1º semestre), y *Tratamiento de Señales Multimedia* (3º curso, 2º semestre); y por otro en las asignaturas *Física General* (1º curso, 1º semestre) y *Álgebra Lineal* (1º curso, 1º semestre).

1.9. Requisitos mínimos de asistencia a las sesiones presenciales

La asistencia a las sesiones de teoría se considera especialmente relevante para la consecución de los objetivos previstos en la asignatura y para participar en las pruebas de evaluación continua. Sin embargo no se imponen requisitos mínimos de asistencia a las sesiones de teoría para participar en la evaluación continua.

La asistencia a las sesiones de prácticas es obligatoria. Sólo se permitirá faltar por motivos justificados y debidamente documentados a una sesión de prácticas. En este caso, el trabajo de la sesión tendrá que recuperarse en el plazo de una semana, en la forma que se acuerde con el profesor de prácticas. En caso de no recuperarse en el plazo dado, y aunque haya sido justificada la ausencia, la práctica será contada como falta de asistencia y puntuada con cero puntos. La falta de asistencia a dos o más sesiones supone la calificación de NO APTO en prácticas, que conlleva la no superación de la asignatura.

Con objeto de crear un adecuado ambiente de trabajo, no se permitirá acceder a la sala de prácticas 10 minutos después de que comience la sesión ni se podrá abandonarlo, salvo por causa justificada o con permiso del profesor, antes de que finalice.

1.10. Datos del equipo docente

Nota: se debe añadir @uam.es a todas las direcciones de correo electrónico.

Profesor de teoría:

Dr. Javier Ortega García (Coordinador)
Departamento de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Despacho - Módulo: C-202; Edificio C - 2ª Planta
Teléfono: +34 914972248
Correo electrónico: javier.ortega



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

Página web: <http://atvs.ii.uam.es/listpeople.do#Ortega-Garcia>
Horario de atención al alumnado: Petición de cita previa por correo electrónico.

Profesor de prácticas:

Dr. Javier González Domínguez
Departamento de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Despacho - Módulo: C-213; Edificio C - 2ª Planta
Teléfono: +34 914977558
Correo electrónico: javier.gonzalez
Página web: <http://atvs.ii.uam.es/listpeople.do#Gonzalez-Dominguez>
Horario de atención al alumnado: Petición de cita previa por correo electrónico.

1.11. Objetivos del curso

TSVA es una asignatura centrada en el ámbito de las denominadas tecnologías de voz y audio, las técnicas de procesado de señal que sobre éstas se pueden aplicar y los esquemas y aplicaciones que a partir de estas técnicas se derivan.

La asignatura tiene cinco bloques, que marcan el itinerario a seguir desde el conocimiento inicial del tipo de señales hasta las aplicaciones avanzadas derivadas.

El primer bloque, Producción y Percepción de Voz y Audio, presenta el mecanismo humano de producción de voz, así como la producción de audio y señales musicales; desde la acústica de su producción, hasta el tipo de señales que se generan en los diferentes dominios bajo estudio; se abordarán asimismo los detalles del mecanismo de percepción y la fisiología de la audición, ahondando en los modelos de producción que emulan y representan estos complejos mecanismos.

El segundo bloque está dedicado a los Fundamentos del Tratamiento Digital de Voz y Audio; en él, se presentarán las técnicas básicas de tratamiento de este tipo de señales, así como las técnicas avanzadas que en los dominios del tiempo y la frecuencia se pueden emplear. Se explicará la transformación cepstral, esencial en el proceso de extracción paramétrica de información, y se desarrollará el análisis de predicción lineal, para terminar el bloque centrados en técnicas de estimación, modelado y clasificación de patrones de señales de voz y audio.

El tercer bloque se centra en la problemática del Reconocimiento Automático de Habla, pasando por sus primeras etapas de categorización lingüística, para así ver cómo implementar esquemas de reconocimiento a distintos niveles de complejidad (palabras aisladas, conectadas, secuencias, *word spotting*); se tratará la problemática del modelado acústico y la estimación estadística de modelos, para así



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

dar una visión global tanto de la perspectiva de reconocimiento como la de comprensión de habla (*speech understanding*).

Por último, el cuarto bloque abordará las problemáticas de síntesis automática de voz y audio, y codificación de dichas señales; se tratarán en un mismo bloque por la interrelación existente entre ambas problemáticas; así, comenzaremos estudiando la problemática de síntesis básica de habla y detección de frecuencia fundamental (*pitch*), para llegar a esquemas complejos de conversión texto a voz (*text-to-speech*). Se abordará el modelo básico del *vocoder*, para así establecer los diferentes criterios y procedimientos que dan lugar a *vocoders* de regímenes binarios diferenciados, para terminar el bloque centrándonos en codificación perceptual, tanto de señales de voz como de audio.

El quinto y último bloque nos permitirá tratar Otras Aplicaciones del Procesado de Voz y Audio, como son; el análisis y la síntesis de señales musicales, la búsqueda de información musical (*music information retrieval*), las técnicas de transformación de voz, el reconocimiento de locutores y las técnicas de diarización.

Paralelamente al desarrollo teórico de la asignatura, se abordará el contenido de carácter práctico. Se propondrá la realización de proyectos de procesado de señales de audio y vídeo para la generación de aplicaciones abiertas que permitan el análisis de técnicas y procedimientos explicados en teoría. De esta forma, se pretende reforzar la teoría y profundizar en el desarrollo de las herramientas computacionales necesarias para ayudar tanto en el análisis como en el diseño de señales y sistemas de voz y audio en sus distintas aplicaciones de interés.

Las **competencias** de tecnología específica correspondientes a Sonido e Imagen (SI) que se abordan con esta asignatura son las siguientes:

- 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.
- 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.

Los objetivos que se pretenden alcanzar con esta asignatura son:

OBJETIVOS GENERALES	
G1	Conocer la naturaleza de la señal de voz, así como las características de las señales musicales. Comprender su mecanismo de producción, su transmisión en canales acústicos, electroacústicos y eléctricos; así como la fisiología del oído humano y la percepción subjetiva de dicha señales.
G2	Conocer la técnicas de procesado de señal que son aplicables a las señales de voz y audio, en los dominios temporal, frecuencial y cepstral; conocer las técnicas de



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

	eventanado y análisis localizado, así como el análisis de predicción lineal; familiarizarse con técnicas básicas de reconocimiento de patrones aplicables a este ámbito.
G3	Conocer en profundidad la problemática del reconocimiento automático de habla; sus aproximaciones, las técnicas más empleadas, su rendimiento actual y sus limitaciones y retos de futuro.
G4	Conocer las técnicas de producción artificial, síntesis y conversión texto a voz de las señales de voz; conocer y comprender las técnicas de compresión y codificación de voz y audio a diferentes regímenes binarios, en su doble vertiente de tasa binaria vs. calidad perceptual.
G5	Conocer diferentes aplicaciones de las técnicas anteriores que tienen gran importancia en el ámbito del tratamiento de voz y audio, como son la síntesis y la búsqueda musical, así como la transformación de voz, y el reconocimiento y diarización de locutores.

OBJETIVOS ESPECIFICOS POR TEMA	
TEMA 1.- Producción y Percepción de Voz y Audio	
1.1.	Conocer el proceso de producción - emisión - recepción en la comunicación hablada
1.2.	Capacidad para caracterizar las señales de voz y audio.
1.3.	Modelar el mecanismo de producción de las señales de voz y audio.
1.4.	Conocer el mecanismo de audición humano, la fisiología del oído y sus características psicoacústicas
TEMA 2.- Fundamentos del Tratamiento Digital de Voz y Audio	
2.1.	Conocer las técnicas básicas de análisis y tratamiento digital de procesado de señales de voz y audio.
2.2.	Conocer las técnicas avanzadas de análisis en tiempo y frecuencia sobre las señales.
2.3.	Conocer la transformación homomórfica sobre las señales de voz, y la extracción de parámetros ceptrales.
2.4.	Conocer y aplicar la técnica de análisis de predicción lineal sobre la señal de voz
2.5.	Conocer las técnicas de modelado y clasificación de patrones usadas en este ámbito
TEMA 3.- Reconocimiento Automático de Habla	
3.1.	Conocer el proceso de categorización lingüística propia del reconocimiento de habla.
3.2.	Abordar el reconocimiento de palabras aisladas, conectadas, y la detección de palabras clave.
3.3.	Conocer las técnicas de modelado acústico y de estimación estadística propias del reconocimiento de habla.
3.4.	Abordar el proceso completo de reconocimiento y de comprensión automática de habla.
TEMA 4.- Síntesis y Codificación de Voz y Audio	
4.1.	Conocer las técnicas de producción sintética de habla y de detección de pitch.
4.2.	Profundizar en el proceso completo de conversión texto a voz.
4.3.	Conocer las diferentes propuestas de vocoders que dan lugar a regímenes binarios diferenciados.
4.4.	Ahondar en las singularidades de la codificación perceptual de señales de audio.



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

TEMA 5.- Otras Aplicaciones del Procesado de Voz y Audio	
5.1.	Conocer las técnicas de análisis y síntesis de audio.
5.2.	Conocer las técnicas de búsqueda de información musical.
5.3.	Conocer las técnicas de transformación de la señal vocal.
5.4.	Conocer las problemáticas de reconocimiento de locutores y de diarización.

1.12. Contenidos del programa

Programa Sintético

PARTE I: Producción y Percepción de Voz y Audio
PARTE II: Fundamentos del Tratamiento Digital de Voz y Audio
PARTE III: Reconocimiento Automático de Habla
PARTE IV: Síntesis y Codificación de Voz y Audio
PARTE V: Otras Aplicaciones del Procesado de Voz y Audio

Programa Detallado

1. Producción y Percepción de Voz y Audio

- 1.1 Mecanismo de Comunicación Hablada
- 1.2 Caracterización de las Señales de Voz y Audio
- 1.3 Modelos de Producción de Voz y Audio
- 1.4 Sistema Auditivo: Fisiología de la Audición y Psicoacústica

2. Fundamentos del Tratamiento Digital de Voz y Audio

- 2.1 Técnicas Básicas de Procesado de Señales de Voz y Audio
- 2.2 Análisis Avanzado en Tiempo y Frecuencia de las Señales de Voz y Audio
- 2.3 Transformación Homomórfica: el Dominio Cepstral
- 2.4 Extracción Paramétrica y Análisis de Predicción Lineal
- 2.5 Técnicas de Clasificación de Patrones

3. Reconocimiento Automático de Habla

- 3.1 Extracción de Categorías Lingüísticas para Reconocimiento
- 3.2 Reconocimiento de Palabras Aisladas, Secuencias y *Word Spotting*
- 3.3 Modelado Acústico y Estimación Probabilística
- 3.4 Reconocimiento y Comprensión de Habla

4. Síntesis y Codificación de Voz y Audio

- 4.1 Síntesis de Habla y Detección de Pitch
- 4.2 Conversión Texto a Voz



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

4.3 Vocoders a Diferentes Regímenes Binarios

4.4 Codificación Perceptual de Audio

5. Otras Aplicaciones del Procesado de Voz y Audio

5.1 Análisis y Síntesis de Señales Musicales

5.2 Búsqueda de Información Musical

5.3 Transformación de Voz

5.4 Reconocimiento y Diarización de Locutores

Prácticas de laboratorio

P1: Análisis y visualización tiempo/frecuencia de señales de voz

P2: Desarrollo de un segmentador voz/no voz

P3: Segmentación voz sorda/sonora

P4: Detección y seguimiento de pitch

P5: Análisis de predicción lineal

P6: Realización de un codificador/sintetizador de voz basado en LPC

1.13. Referencias de consulta

Bibliografía:

Por la diversidad de asuntos cubiertos por la asignatura, no existe un libro de texto único que coincida en extensión y nivel de profundidad con la asignatura en su totalidad. No obstante, el texto básico se adapta de forma significativa al temario de la asignatura, por lo que se considera un instrumento esencial para el estudio y la profundización en la misma.

La bibliografía complementaria propuesta, completa el temario en profundidad y extensión, por lo que el coordinador de la asignatura indicará en qué aspecto de la asignatura conviene usar cada uno de los libros propuestos.

Texto básico:



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

- B. Gold, N. Morgan, D. Ellis, “Speech and Audio Signal Processing - Processing and Perception of Speech and Music”, Wiley, 2nd edition, 2011.
 - *Accesible on-line desde la Biblioteca de la EPS-UAM*

Textos complementarios:

- J. Benesty, M. M. Sondhi, Y. Huang (eds.), “Handbook of Speech Processing”, Springer, 2008.
- X. Huang, A. Acero, H.-W. Hon, “Spoken Language Processing”, Prentice Hall, 2001.
- L. Rabiner, R. Schafer, “Theory and Applications of Digital Speech Processing”, Prentice Hall, 2nd edition, 2010.
- D. O’Shaughnessy, “Speech Communications - Human and Machine”, IEEE Press, 2nd edition, 2000.
- J. R. Deller, J. H. L. Hansen, J. Proakis, “Discrete-Time Processing of Speech Signals”, IEEE Press, 2nd edition, 2000.
- D. G. Childers, “Speech Processing and Synthesis Toolboxes”, John Wiley & Sons, 2000.
- S. Furui, “Digital Speech Processing, Synthesis, and Recognition”, CRC Press, 2nd edition, 2000.

Transparencias:

Dado que ninguno de los textos anteriores cubre todo el contenido de la asignatura, se facilitarán transparencias de la asignatura, cubriendo todos los puntos del temario. No obstante, las transparencias son sólo el guión de clase, y se hará necesario completarlas tomando apuntes de clase, así como haciendo uso de la bibliografía (básica y complementaria) especificada.

2. Métodos docentes

Las clases presenciales programadas en esta asignatura están orientadas a las explicaciones teóricas, a la resolución de cuestiones y problemas, a la realización de la parte práctica y a las actividades de evaluación continua.

La actividad presencial se divide, de acuerdo con el horario de la asignatura, en tres horas semanales en el aula y dos horas semanales de actividad docente para la parte práctica.

Actividad en el aula:



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

La actividad en el aula se encuentra repartida en tres aspectos: explicaciones teóricas, resolución de cuestiones y problemas, y actividades de evaluación continua.

La evaluación continua consistirá en la realización de dos exámenes parciales, de dos horas de duración cada uno; adicionalmente se podrán proponer diferentes pruebas breves, a realizar en el horario de teoría de la asignatura.

Las pruebas breves consistirán en problemas o cuestiones a la finalización de cada tema, y/o la resolución de ejercicios y/o problemas fuera del aula.

El primer parcial se realizará durante una de las sesiones de dos horas del horario de laboratorio, tras finalizar el primer bloque de la asignatura (Partes I y II); mientras que el segundo de ellos se realizará coincidiendo con el examen final.

Aparte de las sesiones especiales de evaluación, el resto de sesiones en el aula se dividirán aproximadamente, en cómputo global, en un 80% dedicado a las explicaciones teóricas y un 20% dedicado a la realización de ejemplos y problemas.

Las explicaciones teóricas resumirán los conceptos más importantes de cada tema, pero se considera esencial que los estudiantes tomen apuntes en relación con las explicaciones dadas en clase por el profesor de teoría, así como que profundicen posteriormente en las diferentes temáticas explicadas, haciendo uso de la bibliografía básica y complementaria propuesta.

Actividad de carácter práctico:

Las actividades de carácter práctico tendrán la consideración de 'proyectos', en el sentido de que el profesor propondrá en el laboratorio la realización de un trabajo en equipo no circunscrito a cada sesión semanal de laboratorio, sino que abarcará varias semanas de forma intensiva; consistirá en la realización de aplicaciones que implementen diversos de los aspectos de análisis y procesado vistos en clase de teoría. En este sentido, las prácticas están planificadas para ser completadas de forma autónoma por el equipo de trabajo fuera del horario estricto de laboratorio. Las aplicaciones desarrolladas deberán ser documentadas y presentadas de forma oral y pública, con intervención individual de cada uno de los miembros del equipo; presentación que dará lugar en sí misma a una calificación individualizada del trabajo realizado.

3. Tiempo de trabajo del estudiante

		Horas	%	Horas	%
Presencial	Clases - Explicación teoría	36	24.0%	78	52%
	Clases - Resolución de cuestiones/problemas	4	2.6%		
	Clases - Parte Práctica	24	16.0%		



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

	Presentaciones y Pruebas de evaluación continua	14	9.3%		
No presencial	Estudio semanal regulado	29	19.3%	72	48%
	Realización de parte práctica	12	8.0%		
	Preparación exámenes	31	20.7%		
Carga total de horas de trabajo: 25 horas x 6 ECTS		150	100%	150	100%

(*) Incluye la convocatoria ordinaria y la extraordinaria

4. Métodos de evaluación y porcentaje en la calificación final

La evaluación de la asignatura, o nota final (NF), dependerá de la nota de teoría (TE) y de la nota de prácticas (PR), en la siguiente proporción:

$$NF = 0.6*TE + 0.4*PR$$

Ambas partes, TE y PR se puntúan sobre 10 puntos. Es necesario obtener una calificación mínima de 5 puntos en TE y PR para poder aplicar la expresión anterior. Si no se cumple esta condición, la calificación numérica que se hará constar en actas será:

$$NF = 0.6*\text{Mín}(5, TE) + 0.4*\text{Mín}(5, PR)$$

Nota de teoría, TE:

La nota de teoría será el resultado de uno de los dos procesos de evaluación que se describen:

- 1- Evaluación continua (TE-C): realización de dos exámenes parciales (EP1 y EP2, puntuados cada uno sobre 10 puntos) que dan lugar a dos notas de evaluación continua que se corresponden con cada mitad de la asignatura (EC1, EC2, puntuadas cada una sobre 10 puntos). Podrá solicitarse adicionalmente la realización de pruebas breves a la finalización de cada tema, o la resolución de ejercicios y/o problemas fuera del aula (pruebas breves, PB1 a PBn, puntuadas cada una sobre 10 puntos).
- 2- Evaluación única (TE-U): la realización de una prueba o examen final (EF, puntuado sobre 10 puntos) en la convocatoria ordinaria o en la convocatoria extraordinaria de la asignatura.

Todas las pruebas de evaluación se realizarán sin libros ni apuntes, y no se permitirá el uso de calculadoras programables, teléfonos móviles, ni dispositivo electrónico



alguno, salvo calculadora científica básica (no programable y sin memoria). Consistirán en la resolución de ejercicios teórico/prácticos similares a los realizados, más la contestación a una serie de cuestiones y/o desarrollos teóricos sobre conceptos explicados en clase.

La *evaluación continua* será el proceso asumido por defecto.

$$PB12 = \text{media}(PB1, PB2, \dots, PBi);$$

siendo PB1, PB2, ..., PBi la serie de pruebas breves que se hubieran llevado a cabo con anterioridad a la realización del examen parcial EP1.

Si la nota del primer parcial $EP1 \geq 3.0$, se puede subir nota con la media de las pruebas breves mediante:

$$EC1 = \max(EP1, 0.35 \cdot PB12 + 0.65 \cdot EP1)$$

Los contenidos del primer parcial se liberan para el final si $EC1 \geq 5.0$

El examen final constará de dos partes correspondientes a cada uno de los parciales (EF1 y EF2). Los que hayan liberado el primer parcial, conservarán su nota EC1 y no estarán obligados a realizar esta parte en el examen final. No obstante, podrán en todo caso presentarse a ella para subir su nota, siendo en este caso la nota final de la primera parte la mayor de las dos obtenidas (EC1 y EF1).

Para el examen final se considerarán adicionalmente las pruebas breves, donde:

$$PB12 = \text{media}(PB1, PB2, \dots, PBi);$$

siendo PB1, PB2, ..., PBi la serie de pruebas breves que se hubieran llevado a cabo con anterioridad a la realización del examen parcial EP1; y:

$$PB34 = \text{media}(PBi+1, PBi+2, \dots, PBn);$$

siendo PBi+1, PBi+2, ..., PBn la serie de pruebas breves que se hubieran llevado a cabo con posterioridad a la realización del examen parcial EP1.

Si la nota obtenida en los contenidos del primer parcial $EF1 \geq 3.0$, se puede subir nota con la media de las pruebas breves mediante:

$$EC1 = \max(EF1, 0.35 \cdot PB12 + 0.65 \cdot EF1)$$

Si la nota en los contenidos del segundo parcial $EP2 \geq 3.0$, se puede subir nota con la media de las pruebas breves mediante:

$$EC2 = \max(EF2, 0.35 \cdot PB34 + 0.65 \cdot EF2)$$



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

Para realizar media entre EC1 y EC2, la nota mínima en cada uno de ellos es 3.0. En ese caso, la nota final de teoría mediante evaluación continua sería:

$$TE-C = 0.5 \cdot EC1 + 0.5 \cdot EC2$$

Caso de no poder realizar media, la nota de teoría de evaluación continua será:

$$TE-C = 0.5 \cdot \text{Mín}(5, EC1) + 0.5 \cdot \text{Mín}(5, EC2)$$

La falta de asistencia a cualquiera de las pruebas de evaluación continua significará ser calificado con cero puntos en esa prueba.

La *evaluación única* es el proceso excepcional dirigido a estudiantes que o bien no siguen el proceso de evaluación continua ($TE-C=0$), o bien, habiéndolo seguido, optan por presentarse a un examen final (EF) para aprobar o aumentar su nota, que constará de dos partes (EF1 y EF2) correspondientes a cada uno de los dos parciales.

Para realizar media entre EF1 y EF2, la nota mínima en cada uno de ellos es 3.0. Si es así:

$$EF = 0.5 \cdot EF1 + 0.5 \cdot EF2$$

Caso de no poder realizar media, la nota del examen final será:

$$EF = 0.5 \cdot \text{Mín}(5, TE) + 0.5 \cdot \text{Mín}(5, PR)$$

La calificación final de teoría se obtendrá según:

$$TE = \text{Max}(EF, TE-C)$$

La calificación de teoría sólo se conserva para la convocatoria extraordinaria del mismo curso académico, salvo si es superior a 7 puntos, en cuyo caso se conserva indefinidamente.

Nota de prácticas, PR:

La nota de prácticas será el resultado de uno de los dos procesos de evaluación que se describen:

- 1- Evaluación continua (PR-C): la valoración promedio de los resultados de cada parte del contenido práctico (NL1), obtenidas en forma de evaluación de memorias de prácticas y notas de clase, y la realización de una prueba final de evaluación práctica (NL2), puntuadas cada una sobre 10 puntos.



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

- 2- Evaluación única (PR-U): la realización de una prueba o examen final de prácticas (PRF) planificado en la convocatoria ordinaria o en la convocatoria extraordinaria de la asignatura.

La **evaluación continua** será el proceso asumido por defecto. Sólo se podrá evaluar el laboratorio mediante evaluación única en circunstancias excepcionales debidamente justificadas.

La asistencia y realización de la parte práctica de la asignatura es obligatoria y la evaluación de la misma constará de dos partes, calificadas como NL1 y NL2, cada una de ellas sobre 10 puntos. La primera de ellas (NL1) se desarrollará de forma continuada en cada parte del contenido práctico, donde el profesor asignará una calificación que valorará, de forma global, los resultados obtenidos en cada parte. Así, NL1 será el promedio de las notas obtenidas en cada parte del contenido práctico.

NL2 será el resultado de un ejercicio de evaluación que se desarrollará durante la última sesión de carácter práctico.

La nota final de prácticas por evaluación continua, PR-C, se obtendrá realizando la ponderación de NL1 y NL2 siempre que $NL2 \geq 3.0$, mediante:

$$PR-C = 0.7 \cdot NL1 + 0.3 \cdot NL2$$

Caso de no poder realizar media, la nota será:

$$PR-C = 0.7 \cdot \text{Mín}(5, NL1) + 0.3 \cdot \text{Mín}(5, NL2)$$

La **evaluación única** es el proceso excepcional dirigido a estudiantes que por causas excepcionales debidamente justificadas no siguen el proceso de evaluación continua (PR-C=0), o bien, habiéndolo seguido, optan por presentarse a una prueba final de prácticas (PRF) para aprobar o aumentar su nota. También es la única posibilidad de evaluación en convocatorias extraordinarias.

Si por motivos de asistencia (ver apartado 1.9) un estudiante es declarado NO APTO en prácticas, su nota de prácticas será 0.

La calificación de la parte práctica sólo se conserva para la convocatoria extraordinaria del mismo curso académico, salvo si es superior a 7 puntos, en cuyo caso se conserva indefinidamente.

Evaluación de la parte práctica en convocatorias extraordinarias: por motivos de organización, ha de solicitarse expresamente al coordinador de la asignatura con anterioridad de al menos 15 días respecto a la fecha del examen extraordinario.

ATENCIÓN: Cualquier copia descubierta que se haya realizado a lo largo del curso, tanto en cualquiera de las actividades de teoría desarrolladas, como en cualquiera de



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

los apartados de las prácticas, serán penalizadas con rigurosidad. Cualquier copia o plagio será penalizado siguiendo las normativas de la UAM y de la EPS.

Cronograma

Actividad en el aula:

El siguiente cronograma indica la distribución orientativa de contenido programada para la actividad en el aula, incluyendo la programación de las pruebas de evaluación sobre dicho contenido. El cronograma está planificado para 14 semanas. Las dos últimas semanas se planifican para actividades de resolución de problemas para que dé tiempo a terminar el último tema de teoría con anterioridad a la finalización del contenido práctico de la asignatura, y para también amoldarse a las distintas disponibilidades horarias de los distintos cursos académicos.

Semana	Contenido	Horas presenciales	Horas no presenciales
1	- Presentación y motivación de la asignatura, descripción del programa, normativa y los métodos de evaluación. - 1.1. Mecanismo de Comunicación Hablada - 1.2. Caracterización de las Señales de Voz y Audio	3	2 Trabajo del estudiante: Lectura de las normativas de teoría y prácticas. Estudio del material propuesto. Revisión de los problemas y ejemplos de clase.
2	- 1.3. Modelos de Producción de Voz y Audio - 1.4. Sistema Auditivo: Fisiología de la Audición y Psicoacústica	5	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
3	- 2.1. Técnicas Básicas de Procesado de Señales de Voz y Audio - 2.2. Análisis Avanzado en Tiempo y Frecuencia de las Señales de Voz y Audio	5	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
4	- 2.3. Transformación Homomórfica: el Dominio Cepstral - 2.4 Extracción Paramétrica y Análisis de Predicción Lineal	5	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase. Preparación evaluación continua.
5	- 2.4 Extracción Paramétrica y Análisis de Predicción Lineal	5	3 Trabajo del estudiante: Estudio de material



Asignatura: Tratamiento de Señales de Voz y Audio
 Código: 18499
 Centro: Escuela Politécnica Superior
 Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
 Nivel: Grado
 Tipo: Formación de Tecnología Específica
 Nº de créditos: 6

Semana	Contenido	Horas presenciales	Horas no presenciales
	- 2.5 Técnicas de Clasificación de Patrones		propuesto. Preparación evaluación continua.
6	- Primera Prueba de Evaluación Continua (EC1)	2	
6	- 3.1. Extracción de Categorías Lingüísticas para Reconocimiento - 3.2. Reconocimiento de Palabras Aisladas, Secuencias y <i>Word Spotting</i>	3	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
7	- 3.3. Modelado Acústico y Estimación Probabilística - 3.4. Reconocimiento y Comprensión de Habla - 4.1. Síntesis de Habla y Detección de Pitch - 4.2. Conversión Texto a Voz	5	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
8	- 4.3. Vocoders a Diferentes Regímenes binarios - 4.4. Codificación Perceptual de Audio - 5.1. Análisis y Síntesis de Señales Musicales	5	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
9	- 5.2. Búsqueda de Información Musical - 5.3. Transformación de Voz - 5.4. Reconocimiento y Diarización de Locutores	4	3 Trabajo del estudiante: Estudio de material propuesto. Revisión y realización de los problemas y ejemplos de clase.
10	- S1: Desarrollo proyecto / aplicación práctica	4	3 Trabajo del estudiante: Desarrollo e implementación de aplicación práctica.
11	- S2: Desarrollo proyecto / aplicación práctica	5	3 Trabajo del estudiante: Desarrollo e implementación de aplicación práctica.
12	- S3: Desarrollo proyecto / aplicación práctica	5	3 Trabajo del estudiante: Desarrollo e implementación de aplicación práctica.
13	- S4: Desarrollo proyecto / aplicación práctica	5	3 Trabajo del estudiante: Desarrollo e implementación de aplicación práctica.



Asignatura: Tratamiento de Señales de Voz y Audio
Código: 18499
Centro: Escuela Politécnica Superior
Titulación: Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación
Nivel: Grado
Tipo: Formación de Tecnología Específica
Nº de créditos: 6

Semana	Contenido	Horas presenciales	Horas no presenciales
14	- S5: Desarrollo proyecto / aplicación práctica	5	3 Trabajo del estudiante: Desarrollo e implementación de aplicación práctica.
15 y 16			11 Preparación evaluación continua.
N/A	<i>Entrega del Proyecto / Aplicación Desarrollada (PR1)</i>	2	
N/A	<i>Segunda Prueba de Evaluación Continua (EC2) (coincidiendo con examen final ordinario)</i>	3	
N/A			20 Preparación examen extraordinario
N/A	<i>Presentación del Proyecto / Aplicación Desarrollada (PR2)</i>	4	
N/A	<i>Examen extraordinario de teoría</i>	3	
	TOTAL HORAS	78	72