



DATOS IDENTIFICATIVOS

Procesado Digital Avanzado de Voz

Asignatura	Procesado Digital Avanzado de Voz			
Código	V05M038V01108			
Titulación	Máster Universitario en Teoría de la Señal y Comunicaciones.			
Descriptor	Creditos ECTS	Seleccione	Curso	Cuatrimestre
	5	OP	1	1c
Lengua Impartición	Castellano Inglés			
Departamento	Teoría de la señal y comunicaciones			
Coordinador/a	Rodríguez Banga, Eduardo			
Profesorado	Rodríguez Banga, Eduardo			
Correo-e	erbang@uvigo.es			
Web	http://fatic.uvigo.es			
Descripción general	Esta asignatura proporciona al alumno una visión completa y actual de las técnicas de procesado avanzado de la señal de voz. Es un curso pensado para introducir nuevos conceptos y técnicas y hacer reflexionar al alumno sobre ellos, mostrándole, por ejemplo, como un mismo procedimiento puede ser empleado con distintos fines dentro del procesado de voz. Los primeros tres temas del curso se centran fundamentalmente en el análisis y codificación de la señal de voz. Por supuesto, se aborda la descripción y comparación de los estándares más relevantes, ya sea por su importancia histórica o por su amplia utilización en la actualidad, lo que proporciona al alumno la necesaria conexión entre los modelos teóricos y su aplicación práctica. El cuarto tema se dedica al estudio de las técnicas de síntesis de voz, una línea de investigación en la que durante los últimos años se han realizado importantes avances. Aunque en un sistema de conversión texto-voz tiene gran importancia el procesado lingüístico y prosódico, en este curso nos centraremos fundamentalmente en las técnicas de generación de la señal sintética. Finalmente, se hace una breve introducción a los codificadores de voz contemplados en los estándares MPEG-audio y a la codificación unificada de voz y audio.			

Competencias de titulación

Código	
A10	(*)describir las técnicas más actuales para codificación, síntesis y reconocimiento de voz
A11	(*)saber utilizar el concepto de análisis mediante síntesis para explicar sistemas de la naturaleza que generan señales medibles
B1	(*)Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios o multidisciplinares relacionados con el campo de estudio
B5	(*)Que los estudiantes adquieran habilidades de aprendizaje que les permitan actualizar sus conocimientos de un modo autónomo, consciente y crítico
B6	(*)demostrar su capacidad de analizar y definir propuestas de sistemas, modelos, especificaciones y algoritmos
B8	(*)transmitir el conocimiento adquirido redactando un informe con la extensión adecuada y al nivel exigido por el destinatario del mismo
B10	(*)analizar resultados experimentales, determinar su validez y emitir juicios razonados sobre su alcance
B11	(*)definir, realizar y ejecutar modelos de simulación en un lenguaje de programación de alto nivel como el Matlab o de bajo nivel como el C/C++
B12	(*)saber manejar paquetes de software de simulación de sistemas de procesado de señal y comunicaciones
B14	(*)juzgar críticamente pero de forma positiva los razonamientos de sus compañeros en los foros de la herramienta e-learning y permitir que los demás juzguen los suyos, sacando así provecho de la puesta en común
B16	(*)demostrar su capacidad para aprender nuevos conceptos, metodologías y técnicas en el campo del procesado de señal y comunicaciones de forma autónoma
B18	(*)tener iniciativa y creatividad en la propuesta de soluciones sistémicas y algorítmicas alternativas a las estándar

Competencias de materia

Resultados previstos en la materia	Tipología	Resultados de Formación y Aprendizaje
(*)Saber describir las técnicas avanzadas de procesado de voz para codificación y síntesis.	saber	A10 A11 B1 B5 B6 B8 B10 B11 B12 B14 B16 B18
(*)Manejar los estándares actuales de codificación de voz	saber hacer	A10 A11 B1 B5 B6 B8 B10 B11 B12 B14 B16 B18

Contenidos

Tema

1. Análisis tiempo-frecuencia de la señal de voz.	1.1 Producción y percepción 1.2 Análisis localizado. 1.3 Predicción lineal. 1.4 Medidas de distorsión espectral 1.5 Cuantificación vectorial.
2. Modelado y codificación de la señal de voz (I)	2.1 Estructura y propiedades deseables de un codificador de voz. 2.2 Clasificación de los codificadores de voz 2.3 Vocoder LPC. Estándar FS-1015 2.4 Codificación multipulso. Estándar GSM 6.10 RPE-LTP. 2.5 Codificación excitada por código (CELP, VSELP, ACELP,...). Estándares relacionados (FS1016, TIA IS54, ETSI GSM 6.20, ITU-T G.728, ITU-T G.729, ...) 2.6 Codificación MELP. Estándar MIL-STD-3005
3. Modelado y codificación de la señal de voz (II)	3.1 Modelado de la señal de voz mediante suma de sinusoides. 3.2 Modelado de la señal de voz mediante suma de una componente armónica y otra ruidosa. 3.3 Codificador MBE. Estándar INMARSAT-M. 3.4 Codificadores paramétricos relacionados.
4. Síntesis de voz.	4.1 Introducción a la conversión texto-voz. Etapas. 4.2 Clasificación de los sintetizadores de voz. 4.3 Síntesis por formantes. 4.4 Síntesis por concatenación de unidades. Algoritmos de modificación prosódica. 4.5 Síntesis mediante selección de unidades. 4.6 Síntesis mediante HMMs
5. Codificación de voz y audio	5.1 Codificación perceptual de señales de audio. 5.2 Codificación paramétrica de voz y audio.

Planificación

	Horas en clase	Horas fuera de clase	Horas totales
Sesión magistral	25	0	25
Estudio de casos/análisis de situaciones	12	22	34
Resolución de problemas y/o ejercicios	2	31	33
Foros de discusión	5	5	10
Pruebas de respuesta corta	3	0	3
Informes/memorias de prácticas	0	20	20

*Los datos que aparecen en la tabla de planificación son de carácter orientativo, considerando la heterogeneidad de alumnado

Metodologías	
	Descripción
Sesión magistral	Se pondrán a disposición del alumno los apuntes y material asociado de cada uno de los temas del curso.
Estudio de casos/análisis de situaciones	Se proporciona al alumno diversos artículos e información de estándares que el alumno debe analizar críticamente, teniendo en cuenta técnicas alternativas que se podrían haber considerado.
Resolución de problemas y/o ejercicios	En el material de cada tema se proponen una serie de ejercicios que el alumno debe resolver. La mayoría de ellos se resuelven utilizando Matlab. Aunque el primer tema los ejercicios son básicos, en temas posteriores son más complejos, estando algunos de ellos basados en los artículos e información de estándares puestos a disposición del alumno.
Foros de discusión	Además de como medio de comunicación entre profesor y alumnos, en los foros se pretende promover la discusión crítica entre los propios alumnos. Así, por ejemplo, éste será el cauce para que los alumnos consulten con sus compañeros las dudas surgidas en la realización de los distintos ejercicios. La misión del profesor en los foros será la de encauzar las discusiones por un camino adecuado y la de proponer nuevos temas de discusión.

Atención personalizada	
Metodologías	Descripción
Sesión magistral	La vía habitual para que el profesor aclare las dudas de los alumnos será a través de los distintos foros. Las consultas de carácter privado se podrán realizar vía correo electrónico. Asimismo, se contempla la posibilidad de consultas vía Skype (o similar) previa cita.
Resolución de problemas y/o ejercicios	La vía habitual para que el profesor aclare las dudas de los alumnos será a través de los distintos foros. Las consultas de carácter privado se podrán realizar vía correo electrónico. Asimismo, se contempla la posibilidad de consultas vía Skype (o similar) previa cita.
Foros de discusión	La vía habitual para que el profesor aclare las dudas de los alumnos será a través de los distintos foros. Las consultas de carácter privado se podrán realizar vía correo electrónico. Asimismo, se contempla la posibilidad de consultas vía Skype (o similar) previa cita.
Estudio de casos/análisis de situaciones	La vía habitual para que el profesor aclare las dudas de los alumnos será a través de los distintos foros. Las consultas de carácter privado se podrán realizar vía correo electrónico. Asimismo, se contempla la posibilidad de consultas vía Skype (o similar) previa cita.
Pruebas	Descripción
Informes/memorias de prácticas	La vía habitual para que el profesor aclare las dudas de los alumnos será a través de los distintos foros. Las consultas de carácter privado se podrán realizar vía correo electrónico. Asimismo, se contempla la posibilidad de consultas vía Skype (o similar) previa cita.

Evaluación		
	Descripción	Calificación
Foros de discusión	Se evaluará la participación del alumno en los foros de discusión en base a la relevancia de las preguntas planteadas y de las respuestas a preguntas de otros alumnos y del profesor.	10
Pruebas de respuesta corta	Se realizarán diversas pruebas cortas escritas u orales (vía Skype o similar) para comprobar la autoría de los informes de las prácticas y el grado de comprensión de los conceptos básicos adquiridos por el alumno. Inicialmente están previstas tres, realizadas a continuación de los temas uno, tres y cinco.	20
Informes/memorias de prácticas	El alumno presentará una memoria de prácticas por tema de la asignatura que contendrá la resolución de los ejercicios propuestos y la interpretación de los resultados obtenidos. Adjuntará también el código Matlab empleado. La evaluación se realizará en función de la correcta resolución de los ejercicios, la comprensión de los conceptos que demuestre en la memoria escrita y la organización de la misma. En caso de mutuo acuerdo entre profesor y alumno, la memoria escrita podrá ser sustituida por un video o presentación demostrativa.	70

Otros comentarios sobre la Evaluación

En caso de no superar la asignatura mediante los actos de evaluación definidos en la guía docente para la primera oportunidad, el coordinador de la asignatura comunicará al alumno en los quince días siguientes a la finalización de las actividades académicas del cuatrimestre correspondiente qué actos de evaluación tiene que realizar para superar la

asignatura en la segunda oportunidad.

Fuentes de información

A.M. Kondoz, **Digital Speech. Coding for Low Bit Rate Communication Systems**, Second Edition.2004,
J.R Deller, J.G. Proakis, J.H. Hansen, **Discrete-Time Processing of Speech signals**, Reprint Edition 1999,
Wai C. Chu, **Speech Coding Algorithms: Foundation and Evolution of Standardized Coders**, 2003,
Speech Communications. Human and Machine, **Douglas O'Shaughnessy**, Second Edition. 2000,
An Introduction to Text-to-Speech Synthesis, **T. Dutoit**, 1997,
Text to Speech Synthesis : New Paradigms and Advances, **Shrikanth Narayanan, Abeer Alwan**, 2004,
Introduction to Digital Audio Coding and Standards, **Marina Bosi, Richard E. Goldberg**, 2002,
Text-to-Speech Synthesis, **Paul Taylor**, 2009,
Speech Coding and Synthesis, **W. Bastiaan Kleijn, Kuldip K. Paliwal**, 1995,
Audio signal processing and coding, **Andreas Spanias, Ted Painter, Venkatraman Atti**, 2007,
Springer Handbook of Speech Processing, **Benesty, Jacob; Sondhi, M. M.; Huang, Yiteng (Eds.)**, 2008,

Recomendaciones

Asignaturas que continúan el temario

Trabajo Fin de Máster/V05M038V01207