

PROGRAMA DE CURSO

Código	Nombre			
EL 7022	Procesamiento de voz en telecomunicaciones			
Nombre en Inglés				
Speech processing in telecom				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar	Horas de Trabajo Personal
6	10	4.0	0	6.0
Requisitos			Carácter del Curso	
Para alumnos de ICE y de ICC: EL 3005			Electivo para ICE Postgrado en el DIE Electivo en ICC	
Resultados de Aprendizaje				
<p>Al final del curso el alumno debe ser capaz de:</p> <p>R1. Comprender los principios básicos acerca de las técnicas de procesamiento de voz utilizadas en codificación, reconocimiento, biometría y síntesis.</p> <p>R2. Comparar y criticar los diferentes modelos y técnicas de modo que pueda analizar los requerimientos de aplicaciones reales en el ámbito de redes convergentes: servicios telefónicos y educación.</p> <p>R3. Analizar los protocolos y ejemplos de plataformas tecnológicas que posibilitan implementar servicios de VoIP y ToIP.</p> <p>R4. Diseñar e implementar un sistema de codificación de voz LPC.</p> <p>R5. Diseñar e implementar por lo menos un sistema basado en alguna de las otras tecnologías abordadas en el curso.</p>				

Metodología Docente	Evaluación General
Clases expositivas con énfasis en: a) el diseño e implementación de proyectos; y b) la investigación bibliográfica como trabajo personal.	<p>La evaluación se basa en tres informes y presentaciones correspondientes de los proyectos e investigación realizados a lo largo del semestre. Se sigue la ponderación que se plantea a continuación:</p> $NF = (NC1 + NC2 + NC3 + NE) / 4$ <p>Donde NC_i corresponde a la nota de un par informe/presentación y NE es la nota del examen.</p>

Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Introducción a procesamiento de voz y a técnicas asociadas	5
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
1.1 El proceso del habla 1.2 Voz como señal acústica 1.3 Codificación 1.4 Reconocimiento de voz 1.5 Verificación de locutor 1.6 Síntesis de Voz 1.7 Procesamiento digital de señales (muestreo, FFT, filtrado, enventanado, etc.) 1.8 Análisis LPC (Linear Prediction Coding). 1.9 Cuantización Vectorial 1.10 Reconocimiento de patrones y clasificación bayesiana. 1.11 Modelos para la producción y percepción de la voz.	Comprender los principio básicos acerca de las técnicas de procesamiento de voz utilizadas en codificación, reconocimiento, biometría y síntesis.	[1], caps.1 a 4 [2], caps.1 a 6

Número	Nombre de la Unidad	Duración en Semanas
2	Codificación y síntesis de voz	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
2.1 Codificación de voz 2.1.1 Codificación por forma de onda (PCM y ADPCM). 2.1.2 Codificación por modelo de producción de la voz (vocoder LPC, RECLP, CELP, GSM y MELP). 2.2 Síntesis de voz o TTS (text-to-speech) 2.2.1 Síntesis por reglas 2.2.2 Síntesis con fonemas dependientes de contexto. 2.2.3 Síntesis de prosodia	-Comparar y criticar los diferentes modelos y técnicas de codificación y síntesis de voz. -Analizar los requerimientos para codificación y síntesis de voz en aplicaciones reales en el ámbito de redes convergentes	[1], cap. 4 [2], cap. 4 y 9 [3] cap. 7.

Número	Nombre de la Unidad	Duración en Semanas	
3	Reconocimiento de voz y biometría	6	
Contenidos		Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
3.1 Reconocimiento automático de voz 3.1.1 DTW 3.1.2 HMM 3.1.3 Algoritmo de Viterbi 3.1.4 Algoritmos Forward y Backward 3.1.5 Algoritmo de Baum-Welch basado en EM 3.1.6 Palabra aislada v/s palabra continua 3.1.7 Dependencia e independencia del locutor 3.1.8 Efecto de canal en reconocimiento de voz 3.1.9 Ejemplos de aplicaciones en telefonía y educación 3.2 Biometría por voz 3.2.1 Verificación e identificación de locutor 3.2.2 Dependencia e independencia de texto 3.2.3 Efecto de canal en verificación de locutor 3.2.4 Dependencia de la cantidad de datos de entrenamiento. 3.2.5 Ejemplo de aplicaciones en telefonía.		-Comparar y criticar los diferentes modelos y técnicas de reconocimiento de voz y del hablante. -Analizar los requerimientos para reconocimiento de voz y del hablante en aplicaciones reales en el ámbito de redes convergentes	[1] cap. 5. [2] caps. 10 y 11.

Número	Nombre de la Unidad	Duración en Semanas	
4	Tecnología de voz en redes convergentes	2	
Contenidos		Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
4.1 Modelo de capas de referencia TCP/IP 4.2 Arquitectura cliente-servidor y VoIP. 4.3 Aplicaciones en redes con procesamiento centralizado 4.4 Arquitecturas de ToIP		-Analizar los protocolos y ejemplos de plataformas tecnológicas que posibilitan implementar servicios de VoIP y ToIP.	[4] caps. 3 a 5.

Bibliografía

- [1] John R. Deller, et al. "Discrete-Time Processing of Speech Signals". (IEEE Press Classic Reissue), 1999.
- [2] Douglas O'Shaughnessy. "Speech Communications : Human and Machine". 1999)
- [3] Rappaport, T. "Wireless Communications, principles & practice", 2nd Edition, Prentice Hall, 2001.
- [4] Shyam Chakraborty, Janne Peisa, Tomas Frankkila, and Per Synnergren. "IMS IMS Multimedia Telephony over Cellular Systems: VoIP Evolution in a Converged Telecommunication World". Wiley 2007.

BIBLIOGRAFÍA COMPLEMENTARIA

- [5] Frederick Jelinek "Statistical Methods for Speech Recognition". MIT Press. (Language, Speech, and Communication), 1999.
- [6] Tanenbaum, A. "Computer Networks". Prentice Hall, 4^o Ed., 2002.
- [7] Olivier Hersent, Jean-Pierre Petit, David Gurle. "IP Telephony: Deploying Voice-over-IP Protocols (Hardcover)". Wiley, 2005.
- [8] Néstor Becerra Yoma, Apuntes de clase, 2010.
- [9] IEEE Transactions on Acoustics, Speech and Language Processing, de 1990 a 2010
- [10] Speech Communications (Elsevier), de 1990 a 2010.
- [11] Artículos de otras revistas especializados y de congresos internacionales.

Vigencia desde:	
Elaborado por:	Néstor Becerra Yoma,