

Programa de Procesamiento Digital de Señales de Voz



Código:

Identificación y características de la Actividad Curricular

Carrera/s: Ingeniería Electrónica

Plan de Estudios: 2014

Carácter: Electiva

Bloque/Campo: Tecnologías Aplicadas

Área: Sistemas Digitales

Régimen de cursado: Cuatrimestral

Cuatrimestre: 10º [ECA]

Carga horaria: 64 hs / 4 hs semanales

Formato Curricular: ECE

Escuela: Ingeniería Electrónica

Departamento: Control

Docentes responsables: Dr. Juan C. Gómez, Ing. Federico Miyara, Dr. Lucas Terissi, Dr. Gonzalo Sad

Programa Sintético

Procesos de Producción y de Percepción de Señales de Voz – Representación de señales de voz en dominios temporal y frecuencial – Diseño de filtros digitales – Técnicas de análisis en los dominios temporal y frecuencial – Procesamiento en tiempo real usando DSP – Codificación de señales de voz – Reconocimiento Automático de Habla: Modelo Ocultos de Markov, Técnicas de Reconocimiento de Patrones – Identificación/Verificación de locutor – Implementación en tiempo real de sistemas de reconocimiento automático de habla (palabra aislada).

Asignaturas Relacionadas

Previas: Sistemas y Señales II - Probabilidad y Estadística

Simultáneas Recomendadas:

Posteriores:

Vigencia desde 2018

.....
Firma Profesor

.....
Fecha

.....
Firma Aprob. Escuela

.....
Fecha

Con el Aval del Consejo Asesor:

Características Generales

La importancia que siempre ha tenido la voz en el proceso de comunicación humana se ve, en nuestros días, incrementada por el rápido avance tecnológico. La enorme cantidad de posibilidades que la tecnología digital, basada en el desarrollo de microprocesadores cada vez más potentes, ofrece, hace que las aplicaciones de procesamiento digital de señales se multipliquen. Entre estas aplicaciones, las que afectan a la señal de voz han permitido disponer de un conjunto de servicios que hasta hace algunos años eran impensables. Redes de integración de voz y datos, diálogo hombre-máquina, síntesis a partir de texto, identificación/verificación de locutores, son algunos ejemplos de los logros alcanzados por el procesado digital de señales de voz. En este Curso se pretende dar un panorama general de las técnicas modernas de procesamiento digital de señales de voz, y de las aplicaciones más difundidas.

Objetivos

Al concluir el Curso el alumno deberá estar en condiciones de:

- Explicar las nociones básicas de los mecanismos de producción de voz y de percepción de sonidos por el oído humano.
- Explicar y aplicar las formas más difundidas de representación de voz en los dominios temporal y frecuencial, y las técnicas modernas de análisis de voz en esos dominios.
- Explicar y aplicar los principales métodos de síntesis y codificación de voz.
- Describir en términos generales la problemática del Reconocimiento automático de voz así como los distintos paradigmas propuestos para la implementación de sistemas de reconocimiento de voz.
- Describir en forma general la problemática de la Identificación/ Verificación de locutor basada en procesamiento de voz.
- Implementar en DSP aplicaciones básicas de filtrado y reconocimiento de habla/locutor en tiempo real.

Contenido Temático

- 1. Proceso de Producción de Señales de Voz**
 - i. Nociones básicas de Fonética castellana.
 - ii. Teoría acústica de producción de sonidos.
 - iii. Modelos prácticos del tracto vocal para análisis y síntesis de voz. Modelo lineal del tracto vocal: Modelo LPC (Linear Predictive Coding).
- 2. Proceso de Percepción de Señales de Voz**
 - i. Anatomía y fisiología del oído.
 - ii. Mecanismo de Percepción de sonidos.
 - iii. Aplicación del conocimiento sobre percepción en el procesamiento de voz.
- 3. Representación de Voz en los dominios temporal y frecuencial**
 - i. Oscilogramas
 - ii. Espectros
 - iii. Espectrogramas
- 4. Diseño de Filtros Digitales**
 - i. Sistemas Lineales Estacionarios como filtros selectivos en frecuencia
 - ii. Características de los filtros ideales
 - iii. Características de los filtros prácticos
 - iv. Diseño de Filtros FIR: Diseño de filtros FIR de fase lineal por el método de Muestreo en Frecuencia; Diseño de filtros FIR de fase lineal usando ventanas; Diseño de filtros FIR inversos por el método de Mínimos Cuadrados
 - v. Diseño de Filtros Digitales IIR a partir de filtros analógicos mediante la aproximación de derivadas
- 5. Técnicas de Análisis en los dominios temporal y frecuencial**
 - i. Análisis localizado (medidas temporales) -Extracción de característica
 - ii. Análisis espectral
 - iii. Cepstrum
 - iv. Estima de la frecuencia fundamental
 - v. Análisis con Bancos de Filtros
 - vi. Análisis con Codificación Predictor lineal
- 6. Procesamiento en Tiempo Real usando DSP**
 - i. Principales características del C6713 DSP Starter Kit (DSK).
 - a. El Procesador Digital de Señales TMS320C6713
 - b. Code Composer Studio

- c. Ejemplos de Programación
- d. Entrada y Salida al DSK
- e. Arquitectura y set de instrucciones del Procesador C6x
- ii. Implementación de Filtros FIR
- iii. Implementación de Filtros IIR
- iv. Aplicaciones en Síntesis de voz con modelo LPC
- v. Aplicación en generación de efectos de audio (Eco, reverberación, etc.)
- 7. Codificación de la Voz**
 - i. Cuantificación
 - ii. Codificación de forma de onda en el tiempo.
 - iii. Cuantificación escalar / vectorial
 - iv. Codificación en el dominio de la frecuencia.
 - v. Codificadores híbridos.
 - vi. Codificadores de baja y muy baja velocidad.
- 8. Reconocimiento de Voz**
 - i. Paradigmas de Reconocimiento de Voz
 - a. Técnicas de Comparación de patrones - Medidas de Distorsión Espectral - Incorporación de características espectrales dinámicas - Normalización.
 - b. Enfoque acústico-fonético
 - c. Enfoque de Inteligencia Artificial
 - ii. Diseño e implementación de sistemas de reconocimiento de voz
 - iii. Teoría e Implementación de Modelos Ocultos de Markov (HMM) - Uso para reconocimiento de palabra aislada.
- 9. Identificación/Verificación de Locutor.**
- 10. Implementación en Tiempo Real de Sistemas de Reconocimiento de habla/locutor**

Modalidades de enseñanza-aprendizaje

La actividad consiste en el dictado de clases teórico-prácticas de carácter magistral en aula equipada con computadoras y proyector de video. Se presentan los temas teóricos y se realizan implementaciones de los algoritmos en Matlab, que los estudiantes pueden reproducir en sus computadoras. Se realizan tres Trabajos Prácticos y un Trabajo Integrador.

Actividades de Formación Práctica

Se realizarán tres trabajos Prácticos y un Trabajo Integrador.

- **Trabajo Práctico 1:** Introducción al modelado, predicción y síntesis de voz.
- **Trabajo Práctico 2:** Diseño de Filtros Digitales
- **Trabajo Práctico 3:** Reconocimiento automáticos de vocales
- **Trabajo Integrador:** Reconocimiento automático de dígitos aislados

Evaluación

La evaluación consistirá en la defensa del Informe de los Trabajos Prácticos y del Trabajo Integrador.

Distribución de la carga horaria

Presenciales

Teoría		30 hs
Práctica	Experimental de laboratorio	20 hs
	Experimental de Campo	hs
	Resolución de Problemas y Ejercicios	hs
	Problemas abiertos de ingeniería	14 hs
	Actividades de Proyecto y Diseño	hs
	Práctica Profesional Supervisada	hs

	Total	64 hs
Evaluaciones		6 hs
Dedicadas por el alumno fuera de clase		
	Preparación Teórica	10 hs
	Preparación Práctica	10 hs
	Elaboración y redacción de informes, trabajos, presentaciones, etc.	10 hs

Bibliografía básica

- Rabiner, L.R. & Schafer, R.W. (1978): *Digital Processing of Speech Signals*, Prentice Hall, Englewood Cliffs, N.J..
- Rabiner, Lawrence & Juang, Biing-Hwang (1993): *Fundamentals of Speech Recognition*, Prentice Hall, Englewood Cliff, N.J.
- Huang, X. D. & Ariki, Y. & Jack, M. A. (1990): *Hidden Markov Models for Speech Recognition* Edinburgh University Press, Edinburgh, UK.
- John R. Deller & John H.L. Hansen & John G. Proakis (2000): *Discrete-Time Processing of Speech Signal*, IEEE Press, N.J..
- Sadaoki Furui (2001): *Digital Speech Processing, Synthesis, and Recognition, Second Edition*, Marcel Dekker, N.Y..
- F. Miyara & J.C. Gómez (2011): ProDiVoz: Notas de Clase, Presentaciones Power-Point y Apuntes cubriendo todos los temas del programa de la asignatura, Depto. Electrónica y Depto, Control, FCEIA. Material disponible en: <http://www.fceia.unr.edu.ar/prodivoz>
- F. Miyara & J.C. Gómez (2011): ProDiVoz: Guía de Trabajos Prácticos Nros. 1, 2, 3 y Trabajo Práctico Final, Depto. Electrónica y Depto, Control, FCEIA. Material disponible en: <http://www.fceia.unr.edu.ar/prodivoz>

Bibliografía complementaria

- Sanjit K. Mitra (2006): *Procesamiento de Señales digitales - Un Enfoque basado en Computadora*, Tercera Edición, McGraw Hill Interamericana, México DF.
- Rulph Chassaing & Donald Reay (2008): *Digital Signal Processing and Applications with the TMS320C6713 and TMS320C6416 DSK*, Second Edition, John Wiley & Sons, Inc., N.J..
- Proakis, John G. & Manolakis, Dimitris G. (1998): *Tratamiento digital de señales - Principios, algoritmos, y aplicaciones*, Tercera Edición, Prentice Hall, Madrid. Traducido de *Digital Signal Processing - Principles, Algorithms, and Applications*, Prentice Hall, Inc.
- Madisetti, Vijay K. & Williams, Douglas B. (ed.)(1998): *The Digital Signal Processing Handbook, Section X: Speech Processing*, CRC Press and IEEE Press.
- A. V. Oppenheim and Ronald W. Schafer (2004): From Frequency to Quefrequency: A History of the Cepstrum, *IEEE Signal Processing Magazine*, pp. 95-99, 106, September 2004.
- F. J. Harris (1978): On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform, *Proceedings of the IEEE*, pp. 51-83, Vol. 66, No. 1, January 1978.
- L. Rabiner (1989): A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, Vol. 77, No. 2, pp: 257-286, 1989.
- A. J. Viterbi (2006): A Personal History of the Viterbi Algorithm, *IEEE Signal Processing Magazine*, pp: 120-122+142, July 2006.

Recursos web y otros recursos

Sitio Web de la Asignatura: <http://www.fceia.unr.edu.ar/prodivoz> conteniendo:

- Notas de Clase
- Presentaciones PowerPoint
- Guías de Trabajos Prácticos
- Scripts Matlab

Cronograma de actividades

Semana	Unidad	Tema	Actividad
1	1		Clases teórico-prácticas
2	2		Clases teórico-prácticas
3	2		Clases teórico-prácticas
4	3		Clases teórico-prácticas, TP1
5	4		Clases teórico-prácticas
6	4		Clases teórico-prácticas, TP2
7	5		Clases teórico-prácticas
8	6		Clases teórico-prácticas
9	7		Clases teórico-prácticas
10	8		Clases teórico-prácticas
11	8		Clases teórico-prácticas, TP3
12	8		Clases teórico-prácticas
13	9		Clases teórico-prácticas
14	9		Clases teórico-prácticas, TP Int.
15	10		Clases teórico-prácticas
16	10		Clases teórico-prácticas