

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE INGENIERÍA



PROGRAMA DE ESTUDIO

**PROCESAMIENTO DIGITAL DE VOZ**

**0752**

**8°, 9°**

**06**

Asignatura

Clave

Semestre

Créditos

**Ingeniería Eléctrica**

**Ingeniería en Computación**

**Ingeniería en Computación**

División

Departamento

Carrera en que se imparte

**Asignatura:**

Obligatoria

Optativa

de elección

**Horas:**

Teóricas

Prácticas

**Total (horas):**

Semana

16 Semanas

Aprobado:  
Consejo Técnico de la Facultad  
Consejo Académico del Área de las Ciencias  
Físico Matemáticas y de las Ingenierías

Fecha:  
25 de febrero, 17 de marzo y 16 de junio de 2005  
11 de agosto de 2005

**Modalidad:** Curso

**Asignatura obligatoria antecedente:** Ninguna.

**Asignatura obligatoria consecuyente:** Ninguna.

**Objetivo(s) del curso:**

El alumno realizará el tratamiento automático de las principales aplicaciones de voz a través de las técnicas básicas en codificación, síntesis y reconocimiento de voz, así como reducción de ruido para voz.

**Temario**

NÚM.	NOMBRE	HORAS
1.	Aspectos básicos sobre sonidos y voz	3.0
2.	Análisis	12.0
3.	Codificación y síntesis	15.0
4.	Reconocimiento	12.0
5.	Ensanchamiento	6.0
		48.0
	Prácticas de laboratorio	0.0
	Total	48.0



## 1 Aspectos básicos sobre sonidos y voz

**Objetivo:** El alumno conocerá algunos aspectos básicos de acústica; de la fisiología de la generación de voz y de la percepción de sonidos, con énfasis en su modelado; y de fonética acústica.

### Contenido:

- 1.1 Generación de sonidos.
- 1.2 Generación de voz.
- 1.3 Percepción de sonidos y voz.
- 1.4 Fonética acústica.

## 2 Análisis

**Objetivo:** El alumno aplicará las principales técnicas de tiempo y de frecuencia para el tratamiento de señales de voz.

### Contenido:

- 2.1 Filtrado digital y preénfasis.
- 2.2 Muestreo, submuestreo y sobremuestreo.
- 2.3 Ventanas rectangulares, de Hamming, Bartlett, de Hanning y Blackman.
- 2.4 Parámetros en el dominio del tiempo: Cruces por cero, energía en tiempo corto y correlación en tiempo corto.
- 2.5 Parámetros en el dominio de la frecuencia: Transformada corta de Fourier, generación de espectogramas.

## 3 Codificación y síntesis

**Objetivo:** El alumno analizará y aplicará diferentes técnicas de codificación básicas de voz, en diferentes ambientes de cómputo.

### Contenido:

- 3.1 Codificación escalar uniforme, logarítmica y adaptable.
- 3.2 Técnicas de codificación DM, DPCM, ADM y ADPCM.
- 3.3 Predicción lineal, métodos LPC, coeficientes PARCOR, LPC cepstral, Mel cepstral, PLP, y CELP.
- 3.4 Cuantización vectorial (VQ), métodos de agrupamiento: Simple, maximin, k-medias, LBG e Isodata.
- 3.5 Técnicas actuales de síntesis de voz.
- 3.6 Técnicas de síntesis basadas en difonemas.

## 4 Reconocimiento

**Objetivo:** El alumno analizará las principales técnicas de reconocimiento de comandos y palabras conectadas, así como los principios de reconocimiento de palabras continuas.



**Contenido:**

- 4.1 Ajuste dinámico en el tiempo (DTW) y su aplicación a reconocimiento de palabras aisladas.
- 4.2 Aplicación de VQ multiseccionada en tiempo al reconocimiento de palabras aisladas.
- 4.3 Modelos ocultos de Markov (HMM) y su aplicación al reconocimiento de palabras aisladas.
- 4.4 Método de un solo paso para reconocimiento de palabras conectadas.

**5 Ensanchamiento**

**Objetivo:** El alumno conocerá las técnicas actuales de reducción de ruido y aplicará la técnica LMS a problemas de contaminación de voz por diferentes tipos de ruido.

**Contenido:**

- 5.1 Eliminación de ruido. Revisión de técnicas actuales.
- 5.2 Métodos adaptables LMS de eliminación de ruido.

**Bibliografía básica:**

**Temas para los que se recomienda:**

RABINER, L.R.; SCHAFER, R.W. <i>Digital Processing of Speech</i> Englewood Cliffs Prentice-Hall, 1978	<b>Todos</b>
OPPENHEIM ALAN V.; SCHAFER, RONALD W. <i>Discrete-Time Signal Processing</i> 2nd edition New Jersey Prentice-Hall, 1999	<b>Todos</b>
RABINER, L.R.; JUANG. <i>Fundamentals of Speech Recognition</i> Englewood Cliffs Prentice Hall, 1993	<b>Todos</b>
PARSONS, THOMAS W. <i>Voice and Speech Processing</i> McGraw-Hill, 1987	<b>Todos</b>



**Bibliografía complementaria:**

OWENS, F.J. **Todos**  
*Signal Processing of Speech*  
 New York  
 McGraw-Hill, 1993

SHAUGHENESSY, DOUGLAS **Todos**  
*Speech Communication, Human and Machine*  
 Quebec  
 Addison Wesley, 1999

DELLER, J.R., PROAKIS, J.G., HANSEN, J.H. **Todos**  
*Discrete-Time Processing of Speech Signals*  
 Wiley-IEEE Press, 1999

ROWDEN, CHRIS **Todos**  
*Speech Processing*  
 McGraw-Hill, 1992

WIDROW, B.; STEARNS, S.D. **Todos**  
*Adaptive Signal Processing*  
 Prentice-Hall, 1985

**Sugerencias didácticas:**

Exposición oral	<input checked="" type="checkbox"/>	Lecturas obligatorias	<input checked="" type="checkbox"/>
Exposición audiovisual	<input checked="" type="checkbox"/>	Trabajos de investigación	<input checked="" type="checkbox"/>
Ejercicios dentro de clase	<input checked="" type="checkbox"/>	Prácticas de taller o laboratorio	<input checked="" type="checkbox"/>
Ejercicios fuera del aula	<input checked="" type="checkbox"/>	Prácticas de campo	<input type="checkbox"/>
Seminarios	<input type="checkbox"/>	Otras	<input type="checkbox"/>

**Forma de evaluar:**

Exámenes parciales	<input checked="" type="checkbox"/>	Participación en clase	<input type="checkbox"/>
Exámenes finales	<input checked="" type="checkbox"/>	Asistencias a prácticas	<input checked="" type="checkbox"/>
Trabajos y tareas fuera del aula	<input checked="" type="checkbox"/>	Otras	<input type="checkbox"/>

**Perfil profesiográfico de quienes pueden impartir la asignatura**

Investigadores reconocidos del área. Ingeniero en computación familiarizado con el procesamiento digital de señales y voz.