

EM 756 PROCESAMIENTO DE VOZ

8 UD

REQUISITOS: EL 55A, AD.

CARÁCTER: Electivo de la Carrera de Ingeniería Civil Eléctrica

OBJETIVOS:

Entregar a los alumnos los conocimientos básicos acerca de las técnicas utilizadas en codificación, reconocimiento y síntesis de voz.

CONTENIDOS:

1. Introducción

- 1.1 El proceso del habla
- 1.2 Voz como señal acústica
- 1.3 Conceptos de Codificación
- 1.4 Reconocimiento de voz
- 1.5 Verificación de locutor
- 1.6 Síntesis de Voz

2. Procesamiento Digital de Señales

- 2.1 Muestreo de la señal
- 2.2 Análisis de señales discretas
- 2.3 Filtros digitales
- 2.4 DFT y FFT
- 2.5 Ventana de Hamming
- 2.6 LPC

3. Modelo de producción y de percepción de la señal de voz

- 3.1 Filtro del tracto vocal
- 3.2 La frecuencia fundamental y las cuerdas vocales
- 3.3 Formantes
- 3.4 Escala Mel
- 3.5 Enmascaramiento

4. Codificación de la Voz

- 4.1 PCM
- 4.2 ADPCM
- 4.3 Vocoder LPC
- 4.4 CELP
- 4.5 RELP
- 4.6 Normas ITU

5. Reconocimiento de Voz

5.1 Introducción

5.2 DTW

5.3 HMM

5.3.1 Algoritmo de Viterbi

5.3.2 Algoritmo Forward y Backward

5.3.3 Algoritmo de Baum-Welch

5.4 Palabra aislada v/s palabra continua

5.5 Dependencia e independencia del locutor

5.6 Redes Neuronales

6. Verificación de Locutor

6.1 Introducción

6.2 Dependencia e independencia del texto

6.3 Estado del arte

7. Síntesis de Voz

7.1 Introducción

7.2 Síntesis por reglas

7.3 Síntesis con fonemas dependientes del contexto

7.4 Prosodia

ACTIVIDADES:

Sesiones semanales de clase de cátedra y desarrollo de aplicaciones mediante programación de algoritmos.

EVALUACION:

- 2 Controles
- 1 Trabajo computacional
- Lecturas personales de artículos, tesis y libros

N.C. = Nota de controles

T.C. = Trabajo Computacional

N.F.=Nota Final= $0.5*N.C. + 0.5*T.C.$

Donde N.C. y N.E. deben ser mayores o iguales a 4.0 separadamente.

BIBLIOGRAFIA

[1] John R. Deller, et al. "Discrete-Time Processing of Speech Signals". (IEEE Press Classic Reissue), 1999.

[2] Frederick Jelinek "Statistical Methods for Speech Recognition". MIT Press. (Language, Speech, and Communication), 1999

[3] Douglas O'Shaughnessy. "Speech Communications : Human and Machine". 1999)

[4] Néstor Becerra Yoma. "Estudio y aplicación de los métodos determinístico y estocástico para reconocimiento de voz". Tesis de Magister, Facultad de Ingeniería Eléctrica, UNICAMP, Brasil, 1993.

[5] Artículos de periódicos especializados y de congresos internacionales.

RESUMEN DE CONTENIDOS

Introducción a los fenómenos del lenguaje y del habla, y a los mecanismos fisiológicos que producen la señal acústica de la voz. Las técnicas de codificación se explican en el contexto del modelamiento del aparato fonador. Se discute también el estado del arte de aplicaciones como reconocimiento de voz, reconocimiento de locutor y síntesis de voz. A seguir se realiza una breve introducción a los fundamentos básicos de procesamiento digital de señales que permita al alumno entender e implementar algoritmos de procesamiento de voz. Se discuten los mecanismos y procesos asociados con la producción y percepción del habla, así como modelos que intentan llevar en consideración el conocimiento fisiológico y psico-lingüístico en sistemas prácticos. Esquemas de codificación/compresión de voz se discuten en detalle y normas ITU relacionadas con las comunicaciones móviles y redes de paquetes son analizadas. El problema de reconocimiento de voz se trata en el ámbito algorítmico y de los modelamientos determinístico (DTW), estocástico (HMM) y por redes neuronales. La aplicabilidad y complejidad de estas técnicas se abordan considerando el diseño de sistemas con aplicación práctica. La problemática de verificación de locutor se discute considerando los mismos algoritmos empleados en reconocimiento de voz dando énfasis a las limitaciones intrínsecas y a las posibles aplicaciones comerciales. Finalmente, las principales técnicas de *Text-to-speech* o síntesis de voz empleadas hoy en día se discuten y comparan considerando criterios de complejidad y naturalidad de la señal acústica generada.