

Procesamiento de audio

Helenca Duxans Barrobés
Marta Ruiz Costa-jussà

PID_00154788

Material docente de la UOC

Helena Duxans Barrobés**Marta Ruiz Costa-jussà**

El encargo y la creación de este material docente han sido coordinados por el profesor: David García Solórzano (2012)

Primera edición: febrero 2012
© Helena Duxans Barrobés, Marta Ruiz Costa-jussà
Todos los derechos reservados
© de esta edición, FUOC, 2012
Av. Tibidabo, 39-43, 08035 Barcelona
Diseño: Manel Andreu
Realización editorial: Eureka Media, SL
Depósito legal: B-3.154-2012



Los textos e imágenes publicados en esta obra están sujetos –excepto que se indique lo contrario– a una licencia de Reconocimiento-NoComercial-SinObraDerivada (BY-NC-ND) v.3.0 España de Creative Commons. Podéis copiarlos, distribuirlos y transmitirlos públicamente siempre que citéis el autor y la fuente (FUOC. Fundació para la Universitat Oberta de Catalunya), no hagáis de ellos un uso comercial y ni obra derivada. La licencia completa se puede consultar en <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

Introducción

Esta asignatura es una continuación de *Señales y sistemas I y II*, y al mismo tiempo se presenta como un complemento perfecto de la asignatura *Procesamiento de imagen*, por lo que dota a los estudiantes de unos conocimientos transversales del ámbito del procesamiento de la señal. Por lo tanto, una asignatura como esta es fundamental para cualquier ingeniero del itinerario de Audiovisuales del grado de Tecnologías de Telecomunicación, y que, probablemente, utilizará el conocimiento adquirido en esta asignatura en algún momento de su vida profesional.

El procesamiento de audio incluye diferentes aplicaciones tecnológicas, como la reproducción de música en alta fidelidad, el reconocimiento del habla o la síntesis del habla. Esta asignatura, en concreto, pretende combinar una explicación teórica de los principales conceptos del procesamiento de audio y ver sus aplicaciones más relevantes, con una vertiente práctica que permita lograr mejor algunos de los conceptos expuestos.

Finalmente, hemos de comentar que esta asignatura es una de las tres de audio que hay en el itinerario de Audiovisuales, junto a las de *Acústica* y *Electroacústica*. Por lo tanto, al acabar el grado, el estudiante tendrá un amplio conocimiento del ámbito del audio, que irá desde tratar la señal de audio (lo veremos en esta asignatura) y condicionar espacios (*Acústica*) hasta conocer el funcionamiento de varios dispositivos de captación y transmisión de la señal acústica (*Electroacústica*).

Estos apuntes pretenden ser una introducción experimental al procesamiento de audio. Están organizados en tres bloques principales y cada bloque está dividido en módulos. Cada módulo contiene teoría, ejemplos, resúmenes y ejercicios. Asimismo, os presentamos continuamente bibliografía de referencia por si queréis ampliar conocimientos.

El primer bloque presenta, en el primer módulo, conceptos ya vistos sobre señales y sistemas, y en el segundo módulo, explica el diseño de filtros.

El **primer módulo** hace un repaso de los principales conceptos de señales y sistemas que usaremos en esta asignatura. Básicamente incluye los contenidos siguientes: la definición y clasificación de señal y sistema; la caracterización de señales y sistemas mediante la transformada de Fourier y la transformada Z; así como la conversión entre el dominio analógico y discreto.

El **segundo módulo** explica el diseño de filtros. El objetivo principal del módulo es explicar la **teoría básica de diseño de filtros digitales y analógicos**. Concretamente, se pretende dar al estudiante la capacidad de aprender a utili-

zar cualquier **software** de diseño de filtros y que así pueda diseñar sus propios filtros dadas unas especificaciones concretas. Asimismo, se combina la teoría y la práctica del **diseño de filtros** con el estudio de diferentes utilidades de los filtros digitales y analógicos en el área de procesamiento de la señal de audio.

A continuación, entramos en el bloque 2, que se centra en desarrollar conceptos de audio. Por lo tanto, es el bloque más largo de los tres. Este bloque incluye los módulos 3, 4 y 5.

El **tercer módulo** lleva a cabo una introducción a la acústica. Se desarrollan en él conceptos como qué es el sonido y cómo lo percibimos, medimos, se genera y se propaga. Concretamente, estudiaremos cómo funciona la acústica fisiológica, es decir, de qué modo percibe el sonido el oído humano, cómo emite el sonido el sistema de fonación humana y cuál es la directividad de la voz humana. Dentro de las herramientas de percepción del sonido definiremos las curvas isofónicas, los filtros de ponderación y el nivel de ruido de fondo, y dentro de la medición del sonido veremos el sonómetro. También analizaremos la tipología de fuentes sonoras existentes y el comportamiento que tienen en diferentes tipos de recintos. Por último, veremos cómo se propaga el sonido y qué son la reverberación y la inteligibilidad del sonido.

El **cuarto módulo** presenta el proceso de codificación que sigue la señal de audio para almacenarse o transmitirse digitalmente. Veremos la cuantificación digital y algunos conceptos que están relacionados, como el sobremuestreo, el tramado (*dithering*) y la conformación del ruido (*noise shaping*). También conoceremos las codificaciones principales que actualmente se utilizan para el audio y distinguiremos entre codificadores de audio generalistas y codificadores específicos para la voz. En este módulo veremos cómo se utiliza la psicoacústica, es decir, las características de la percepción del sonido en el oído humano, para mejorar la eficiencia de los codificadores. Finalmente, la última parte del cuarto módulo está dedicada a presentar los formatos de ficheros de audio más utilizados en el mundo digital, para almacenamiento o transmisión, sobre todo en el ámbito de redes IP, como Internet.

En el **quinto módulo** conoceremos un ejemplo real del modo como se aplica la teoría vista en los módulos anteriores en el campo del procesamiento de audio: la creación de efectos sonoros. Después de presentar qué es un efecto digital de audio haremos una pequeña introducción a los efectos digitales de audio más habituales, sobre todo en el mundo de la producción musical, proporcionando –junto a las bases teóricas del procesamiento digital de la señal de los sistemas que los generan– ejemplos y referenciando creaciones musicales que puedan resultaros fáciles de encontrar.

El tercer bloque de la asignatura presenta dos aplicaciones de procesamiento de audio, pero centradas en el ámbito del habla: el reconocimiento y la síntesis del habla. Este bloque está formado por los módulos 6, 7 y 8.

El **sexto módulo** es un módulo introductorio a las características del habla que nos permitirá tener los conocimientos básicos para entender los fundamentos de las técnicas presentadas en los módulos siguientes. Antes de nada, veremos cómo se produce la voz en el sistema fonador humano y cuáles son las características acústicas de la voz. Por último, haremos el paso entre la acústica y la fonética para presentar cómo se clasifican fonéticamente todos los sonidos que somos capaces de articular.

En el **séptimo módulo** introducimos el concepto de reconocimiento automático del habla y haremos un repaso de las técnicas más importantes. Concretamente, analizaremos dos de las etapas de los reconocedores: la extracción de características acústicas sobre las que se basará el reconocimiento y el algoritmo de reconocimiento. Veremos cómo se puede transformar el reconocimiento en un problema de búsqueda, utilizando modelos acústicos y modelos de lenguaje valorados previamente para reducir su complejidad. Por último, conoceremos las herramientas que nos permiten medir el grado de “bueno” de un reconocedor.

El **octavo módulo** está dedicado a la síntesis del habla. Concretamente, veremos los convertidores de texto a voz, cuya finalidad es transformar en voz cualquier texto escrito. Dentro de todas las técnicas que hay para sintetizar voz, este módulo focaliza el tema en los sistemas de síntesis por concatenación, resaltando la importancia que tiene el corpus (o base de datos), tanto en la selección de los segmentos que se deben concatenar como en la calidad final de la voz sintetizada. Como en el módulo séptimo, proporcionaremos herramientas para medir la calidad de los convertidores de texto a voz para poderlos comparar entre sí.

Actividades

Conceptos de señales y sistemas

1. Dibujad una señal analógica, una señal discreta en tiempo y una señal digital.
2. ¿Qué puede ayudar a mejorar la reconstrucción ideal, si nos fijamos en el convertidor A/D?

Diseño y análisis de filtros en procesamiento de audio

3. ¿Cuál es la respuesta impulsional del filtro paso bajo ideal? Razonad por qué no es realizable un filtro paso bajo ideal.
4. Haced un filtro paso bajo de orden 2 siguiendo el esquema del ejemplo [SLPF], con coeficiente $b_2 = 1$. ¿Qué $y[n]$ se obtiene?
5. Hemos visto el diseño de dos grandes tipos de filtros: FIR e IIR. En esta actividad queremos comparar estas dos variantes de filtros. La comparativa la llevaremos a cabo mediante la herramienta del FDAtool ayudándonos del ejemplo del apartado 5.2.2.

En primer lugar, se pide que diseñéis con FDAtool un filtro paso banda con banda de paso entre 4 kHz y 8 kHz, y atenuación de 6 dB en las frecuencias de 2 kHz y 10 kHz. Comparad diferentes métodos FIR e IIR. ¿Cuál es el método que requiere un orden más bajo para cumplir las especificaciones? ¿Cuál en este orden? Visualizad la respuesta impulsional.

En segundo lugar, se pide que defináis unas especificaciones para un filtro paso alto. Utilizad nuevamente FDAtool para diseñar un filtro IIR. Experimentad qué aproximación (Butterworth, Txebyxev, inversa de Txebyxev o Cauer) cumple las especificaciones con el mínimo orden y por qué.

6. Demostrad que la respuesta impulsional de un filtro IIR tiene un número infinito de muestras diferentes de cero.

Introducción a la acústica

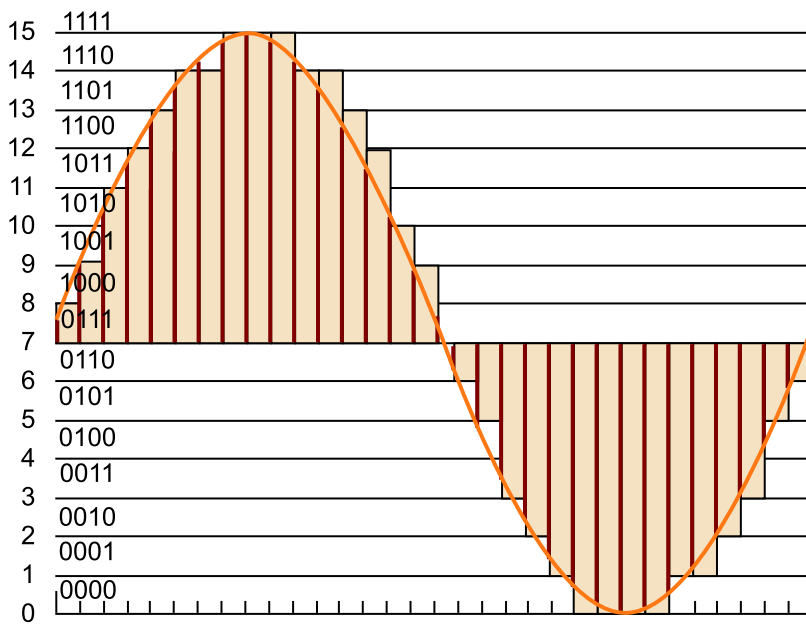
7. En el diseño acústico de teatros, auditorios o cines, es muy importante tener en cuenta los coeficientes de absorción de los materiales (como butacas, cortinas o moqueta) que se ponen en las salas. Buscad en sengpielaudio.com y poned un ejemplo de material muy absorbente (con un $\alpha > 0,7$), un material medianamente absorbente ($0,2 < \alpha < 0,7$) y un material poco absorbente ($\alpha < 0,2$). Considerad como frecuencia 1 kHz.
8. Los fenómenos de reverberación y eco dependen del solapamiento del sonido directo y reflejado. Vamos a analizar estos conceptos con un ejemplo práctico. Si la velocidad del sonido es de 350 m/s y el objeto reflector introduce un camino extra de 10 m respecto al camino directo, ¿qué retraso tendrá la señal reflejada respecto a la señal directa? ¿Se puede considerar eco o será reverberación? ¿Qué sucede si el objeto reflector se encuentra a 50 m de la fuente? Realizad la comprobación con Audacity.
9. Sabemos que hay diferentes formas de calcular el tiempo de reverberación. Vamos a analizarlas y compararlas con un ejemplo práctico. Tenemos una sala rectangular de $3 \times 4 \times 5$ m (altura \times anchura \times profundidad). El techo es de madera; las paredes, de vidrio, y el suelo, de parqué. Calculad el tiempo de reverberación para las frecuencias de 125 Hz, 250 Hz y 500 Hz, según Sabine, Eyring y Millington. Comentad si los tiempos que se obtienen son iguales o diferentes y por qué.

La tabla siguiente muestra los coeficientes de absorción de los diferentes materiales utilizados:

Coefficiente de absorción	125 Hz	250 Hz	500 Hz
Madera	0,15	0,11	0,10
Vidrio	0,18	0,06	0,04
Parqué	0,04	0,04	0,07

Codificación del audio

10. Escribid la secuencia de 0 y 1 de la señal codificada de la siguiente figura:



11. ¿Qué cambios introduce la cuantificación logarítmica en un cuantificador uniforme para que este cuantificador uniforme sea no uniforme?

12. Pensad cómo afecta aumentar el número de bits por muestra y la frecuencia de muestreo en la cantidad de memoria necesaria para almacenar el audio codificado. Haced estas comparaciones para calidad CD (16 bits por muestra y 44,1 kHz) y para 24 bits por muestra y 96 kHz:

a) 30 minutos de música estéreo.

b) Y si lo guardáramos en formato MP3, con una tasa de bits de 128 kbps, ¿cuál sería el factor de compresión?

Efectos digitales de la señal de audio

13. Identificad los efectos siguientes:

Audio original	Audio con efecto	Tipo de efecto
1	1	
2	2	
3	3	

14. ¿Cuál es el diagrama de bloques de un sistema digital que introduce tres ecos en un audio original y que aplica un trémolo solo al audio original?

15. Cread un efecto sonoro con Audacity. Abrid Audacity y grabad una palabra cualquiera. Seleccionad la señal grabada y añadid el efecto eco. Podéis jugar con el factor de atenuación (inverso de la ganancia) y el tiempo de retraso (siempre mayor de 50 milisegundos) para ver cómo se modifica la señal. Ahora añadid un eco con un retraso de menos de 50 milisegundos. ¿Notáis alguna diferencia?

Reconocimiento automático del habla

16. Los reconocedores de palabras clave, denominados también *word spotting*, se diseñan para detectar solo las palabras que el usuario indica al sistema, y por lo tanto ignora el resto de la voz. Entre las posibles implementaciones de estos reconocedores existe toda una familia que se basa en hacer modelos HMM de toda la palabra clave que se ha de detectar. Estos modelos, junto con modelos *filler* o *garbage*, que modelan el resto de la voz que no es una palabra clave, se utilizan en el bloque de descodificación para encontrar una transcripción del tipo siguiente: palabra clave, *filler* + palabra clave, *filler* + palabra clave + *filler* o palabra

clave + *filler*. Indicad los pasos necesarios para introducir una palabra clave nueva en este tipo de sistemas.

17. Calculad el porcentaje de error de inserción, sustitución y eliminación para la transcripción siguiente. ¿Cuál es el valor de la WER?

- Transcripción de referencia: "a la reunión asistieron diez personas".
- Transcripción automática: "en la reunión dijeron sí diez personas".

18. Buscad dos reconocedores del habla que incluyan el castellano entre los idiomas que reconocen.

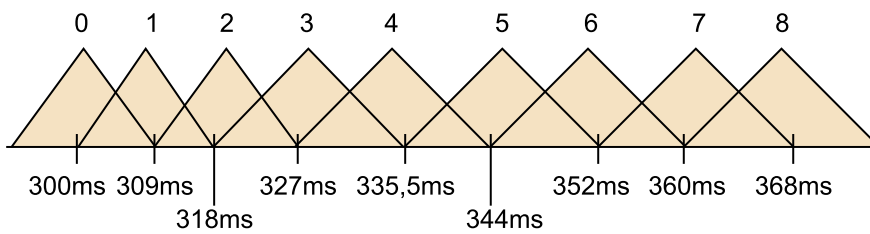
Síntesis del habla

19. Identificad qué bloques de un TTS basado en concatenación acústica hay que modificar para introducir una voz en un idioma nuevo. ¿Y si se quiere introducir una voz nueva, pero en un idioma de los que ya existen?

20. Buscad los puntos de síntesis y la secuencia de ventanas que se han de concatenar para la señal de voz siguiente, si se quiere hacer una modificación de velocidad constante de 1,4 (es decir, una modificación de duración de 1/1,4) y una modificación de altura tonal de 1,1.

Instantes de análisis: 300 ms, 309 ms, 318 ms, 327 ms, 335,5 ms, 344 ms, 352 ms, 360 ms, 368 ms y 368 ms.

Ventanas:



Solucionario

Conceptos de señales y sistemas

2. Por ejemplo, utilizar un cuantificador de más bits.

Diseño y análisis de filtros en procesamiento de audio

3. Es una sinc centrada en su origen. No es realizable porque la respuesta impulsional es infinita.

4. $y[n] = \{1,3,6,9,12,15,18,21,24,27\}$ ($v(n) = [0;0]$, $b = [1,1,1]$, $M = 2$)

5. IIR elíptico de orden 4.

6. Cauer requiere el mínimo orden porque tiene un comportamiento con rizado de amplitud constante en las bandas de paso y atenuada.

Introducción a la acústica

7. Material muy absorbente: fibra de vidrio.

Material medianamente absorbente: alfombra sobre cemento.

Material poco absorbente: mármol.

8. En el primer caso, el retraso es de 28 ms; por lo tanto, se considera reverberación.

En el segundo caso, el retraso es de 142 ms; por lo tanto, se considera eco.

9.

TR (s)	125	250	500
Sabine	0,77	1,64	1,82
Eyring	0,73	1,58	1,76
Millington	0,71	1,57	1,75

Codificación del audio

10.

0111 1001 1011 1100 1101 1110 1110 1111 1111 1111 1110 1101 1100 1010 1001 0111
0110 0101 0011 0010 0001 0000 0000 0000 0000 0000 0001 0001 0010 0011 0101 0110

11. El cuantificador logarítmico incorpora previamente al cuantificador uniforme una etapa de compresión y a la salida del cuantificador uniforme añade una expansión logarítmica que modifica la señal.

12.

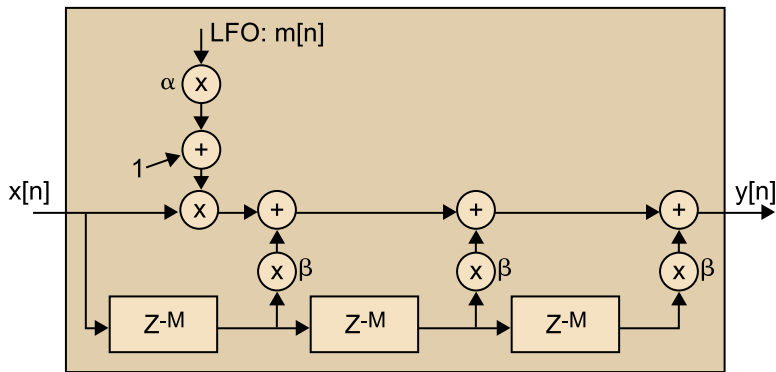
a) Calidad CD: 2.540.160.000 bits; para la otra codificación: 4.147.200.000 bits.

b) 230.400.000 bits, factor de compresión para calidad CD 11,025; para la otra codificación: 18.

Efectos digitales de la señal de audio

13. *Wah-wah*, trémolo, eco.

14.



15. Cuando el tiempo de retraso es más pequeño de 50 milisegundos, el efecto que se percibe es una reverberación.

Reconocimiento automático del habla

16.

1.º Generar la transcripción fonética de la palabra clave.

2.º Incluir la palabra clave en el diccionario del modelo de lenguaje.

3.º Generar el HMM de la palabra clave. Si podemos hacer grabaciones nuevas: hacer múltiples grabaciones con diferentes usuarios de la palabra clave, procesar el audio y extraer los MFCC y entrenar un único HMM. Si no podemos hacer grabaciones nuevas: generar el HMM concatenando los HMM de los trifenemas que forman la palabra clave.

4.º Incluir el HMM generado en el modelo acústico del reconocedor de palabras clave.

5.º Opcionalmente, volver a valorar el HMM *filler* o los HMM *filler*.

Todos los cambios en el reconocedor de palabras clave se llevan a cabo en la fase de entrenamiento. La fase de ejecución del sistema continúa igual.

17.

Porcentaje de error de inserción: $1/7 * 100 = 14,29\%$

Porcentaje de error de sustitución: $1/7 * 100 = 14,29\%$

Porcentaje de error de omisión: $1/7 * 100 = 14,29\%$

WER = 42,86%

18.

Veamos algún ejemplo:

Nuance

Loquendo

Microsoft

IBM

Verbio

Síntesis del habla

19. Para una voz nueva en un idioma nuevo:

Todo el módulo de procesamiento de lenguaje natural: el analizador morfosintáctico, el transcriptor fonético y el generador prosódico. El corpus de unidades acústicas (diseño del corpus + grabación + etiquetado).

Para una voz nueva en un idioma que ya existe:

El corpus de unidades acústicas (grabación + etiquetado).

20. Primer instante de síntesis: 300 ms.

Primera ventana: 0.

Segundo instante de síntesis: $300 \text{ ms} + (309 \text{ } 300)/1,1 = 308,18 \text{ ms}$.

Instante virtual central de la segunda ventana: $308,18 + (308,18 \text{ } 300) * 1,4 = 311,45 \text{ ms}$.

Segunda ventana: 1 (ventana con instante central más cercano a 311,45 ms).

Tercer instante de síntesis: $308,18 + (318 \text{ } 309)/1,1 = 316,36 \text{ ms}$.

Instante virtual central de la tercera ventana: $316,36 + (316,36 \text{ } 308,18) * 1,4 = 327,81 \text{ ms}$.

Tercera ventana: 3 (ventana con instante central más cercano a 327,81 ms).

Cuarto instante de síntesis: $316,36 + (335,5 \text{ } 327)/1,1 = 324,09 \text{ ms}$.

Instante virtual central de la cuarta ventana: $324,09 + (324,09 \text{ } 316,36) * 1,4 = 334,91 \text{ ms}$.

Cuarta ventana: 4 (ventana con instante central más cercano a 334,91 ms).

Quinto instante de síntesis: $324,09 + (344 \text{ } 335,5)/1,1 = 331,82 \text{ ms}$.

Instante virtual central de la quinta ventana: $331,82 + (331,82 \text{ } 324,09) * 1,4 = 342,64 \text{ ms}$.

Quinta ventana: 5 (ventana con instante central más cercano a 342,64 ms).

Sexto instante de síntesis: $331,82 + (352 \text{ } 344)/1,1 = 339,09 \text{ ms}$.

Instante virtual central de la sexta ventana: $339,09 + (339,09 \text{ } 331,82) * 1,4 = 349,27 \text{ ms}$.

Sexta ventana: 6 (ventana con instante central más cercano a 349,27 ms).

Séptimo instante de síntesis: $339,09 + (360 \text{ } 352)/1,1 = 346,36 \text{ ms}$.

Instante virtual central de la séptima ventana: $346,36 + (346,36 \text{ } 339,09) * 1,4 = 356,54 \text{ ms}$.

Séptima ventana: 7 (ventana con instante central más cercano a 356,54 ms).

Contenidos

Módulo didáctico 1

Conceptos de señales y sistemas

Marta Ruiz Costa-jussà y Helenca Duxans Barrobés

1. Señales y sistemas
2. Transformación del dominio temporal al dominio frecuencial

Módulo didáctico 2

Diseño y análisis de filtros en procesamiento de audio

Marta Ruiz Costa-jussà y Helenca Duxans Barrobés

1. Concepto y tipos de filtros. ¿Por qué se debe aprender cómo funciona un filtro?
2. Conceptos básicos para el diseño de filtros digitales
3. Filtros reales: plantilla de especificación de un filtro
4. Diseño de filtros digitales
5. De la teoría a la práctica

Módulo didáctico 3

Introducción a la acústica

Marta Ruiz Costa-jussà y Helenca Duxans Barrobés

1. Definición y caracterización del sonido y del ruido
2. Percepción humana del sonido. Fenómenos sonoros
3. Fuentes sonoras y propagación del sonido

Módulo didáctico 4

Codificación del audio

Marta Ruiz Costa-jussà y Helenca Duxans Barrobés

1. Introducción al audio digital
2. Cuantificación
3. Cuantificación inversa
4. Procesos del audio digital
5. Clasificación de los codificadores de audio
6. Codificadores de forma de onda
7. Codificadores perceptivos
8. Codificaciones específicas para voz
9. Formatos de ficheros de audio

Módulo didáctico 5

Efectos digitales de la señal de audio

Helenca Duxans Barrobés y Marta Ruiz Costa-jussà

1. Introducción y clasificación de los efectos digitales de audio
2. Efectos sonoros basados en retardadores
3. Efectos sonoros basados en moduladores
4. Efectos sonoros basados en sistemas lineales

5. Efectos sonoros basados en sistemas no lineales
6. Otros tipos de efectos

Módulo didáctico 6

Introducción al habla

Helena Duxans Barrobés y Marta Ruiz Costa-jussà

1. Introducción a las tecnologías del habla
2. La producción de la voz en tres pasos
3. Propiedades acústicas de la señal de voz
4. Clasificación fonética de los sonidos
5. Unidades acústicas utilizadas en las tecnologías del habla

Módulo didáctico 7

Reconocimiento automático del habla

Helena Duxans Barrobés y Marta Ruiz Costa-jussà

1. Introducción al reconocimiento automático del habla
2. Aplicaciones de los reconocedores automáticos del habla
3. Funcionamiento básico de los reconocedores
4. El módulo de extracción de características
5. El módulo de decodificación
6. Técnicas de adaptación
7. Evaluación de la transcripción automática

Módulo didáctico 8

Síntesis del habla

Helena Duxans Barrobés y Marta Ruiz Costa-jussà

1. Introducción a la síntesis del habla
2. Aplicaciones de los convertidores de texto a voz
3. Los convertidores de texto a voz
4. Síntesis por concatenación
5. Modificaciones prosódicas
6. Medidas de calidad de la voz sintetizada

Bibliografía

Carrión Isbert, A. (1998). *Diseño acústico de espacios arquitectónicos*. Barcelona: Edicions UPC.

Cremer, L.; Muller, H. A. (1982). *Principles and Applications of Room Acoustics*(vol. 1). Londres: Applied Science Publishers.

ETSI SE 202 050 V1.1.1 (2002-2010). "Speech processing, transmission and quality aspects (STQ); distributed speech recognition; advanced frente-end feature extraction algorithm; compression algorithms" (ref. DES/STQ-00008)

Franco Contadini, M. (2010). "Oversampling with averaging to increase ADC resolution".

González, M. D. (2006). "Comparación de filtros FIR de fase lineal, por Método Óptimo y de Ventanas".

Iosu, D. (1999). "Análisis de Fourier".

Irizar Picón, A. (2002). "Tratamiento Digital de Señal".

Kompis, M.; Dillier, N. (1993). "Simulating transfer functions in a reverberant room including source directivity and head-shadow effects". *JASA*(núm. 93, pág. 2779-2787).

Lamba, D. (2010). "Audio Signal Filtering".

Mariño, J. B. y otros (1999). *Tratamiento digital de la señal: una introducción experimental*. Barcelona: Edicions UPC.

McClellan, J. H.; Parks, T. W. (2005). "A personal history of the Parks-McClellan algorithm". *Signal Processing Magazine, IEEE* (vol. 22, núm. 2, marzo, pág. 82-86). Atlanta, GA, EE. UU.: Georgia Institute of Technology.

Molina, R. (2008). "Cuantificación Escalar".

Moreno, A. (2003). "Cuantificación". Universitat Politècnica de Catalunya.

Wagner, B.; Barr, M. (2007). "Filtres FIR i IIR".

Proakis, J. G.; Manolakis, D. G. (2007). *Tratamiento digital de señales*. Madrid: Pearson Prentice Hall.

Smith, J. O. (2008, octubre). "Spectral Audio Signal Processing".

Smith, J. O. (2011). . "Amplitude response".

wikipedia.org (2012). "Filtros analógicos".

Zawistowski, Th.; Shah, P. (2005). "An Introduction to Sampling Theory".