

Guia docent

230622 - DSAP - Processament Digital d'Àudio i Veu

Última modificació: 24/05/2024

Unitat responsable: Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona

Unitat que imparteix: 739 - TSC - Departament de Teoria del Senyal i Comunicacions.

Titulació: MÀSTER UNIVERSITARI EN ENGINYERIA DE TELECOMUNICACIÓ (Pla 2013). (Assignatura optativa).
MÀSTER UNIVERSITARI EN TECNOLOGIES AVANÇADES DE TELECOMUNICACIÓ (Pla 2019). (Assignatura optativa).

Curs: 2024

Crèdits ECTS: 5.0

Idiomes: Anglès

PROFESSORAT

Professorat responsable: MARTA TOLOS RIGUEIRO

Altres: Primer quadrimestre:
MARTA TOLOS RIGUEIRO - 10

CAPACITATS PRÈVIES

Haver cursat almenys dues assignatures de l'àrea de Senyals, Sistemes i Transformades
Haver cursat almenys una assignatura de l'àrea Probabilitat, Variables Aleatòries i Processos Estocàstics
Experiència de programació en Matlab
Recomanable:
- Coneixements bàsics d'aprenentatge automàtic
- Llenguatge Python

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

1. Capacitat per aplicar mètodes de la teoria de la informació, la modulació adaptativa i codificació de canal, així com tècniques avançades de processat digital del senyal als sistemes de comunicacions i audiovisuals.

Transversals:

2. TREBALL EN EQUIP: Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o duent a terme tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, tot assumint compromisos considerant els recursos disponibles.

3. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

4. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

METODOLOGIES DOCENTS

- Classes expositives (50%)
- Classes d'aplicació (amb Matlab o similar) (50%)
- Treball en equip: projecte d'assignatura i altres
- Treball individual: tasques per fer a casa, relacionades amb les aplicacions

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

Objectius d'aprenentatge:

Adquirir la comprensió i ser competent en un conjunt relevant de conceptes i tècniques del camp del processament digital d'àudio i la seva aplicació als problemes que apareixen en aplicacions reals. Especialment, es consideraran els senyals i les aplicacions de parla i música.

Resultats d'aprenentatge:

La capacitat de processar digitalment els senyals d'àudio, en un context orientat a aplicació, per tal de fer anàlisi, modelatge, extracció d'informació, neteja, modificació, i síntesi/generació.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores aprenentatge autònom	86,0	68.80
Hores grup gran	39,0	31.20

Dedicació total: 125 h

CONTINGUTS

Introducció

Descripció:

Presentació de l'assignatura
La diversitat dels senyals d'àudio
Característiques de la parla i la música. El model de producció
Audició i modelat auditiu
STFT i espectrograma
La transformada de Fourier a curt termini

Dedicació: 12h

Grup gran/Teoria: 6h

Aprenentatge autònom: 6h

Modelatge i representació de senyals d'àudio

Descripció:

Modelat només amb pols basat en producció
Determinació del to per a parla i música
Codificació basada en LPC que es fa servir en telefonia mòbil

Dedicació: 8h

Grup gran/Teoria: 4h

Aprenentatge autònom: 4h

Millora de senyals de veu i àudio

Descripció:

Reducció de soroll: substracció espectral, filtrat de Wiener, xarxes neuronals (aprenentatge profund)
Cancel·lació: eco, interferències
Separació cega de fonts: NMF

Dedicació: 16h

Grup gran/Teoria: 8h

Aprenentatge autònom: 8h

Reconeixement i detecció d'àudio i parla

Descripció:

Enfocaments basats en ajustament dinàmic de patrons
Enfocaments estadístics i d'aprenentatge profund
Detecció d'activitat d'àudio
Aplicació al reconeixement de parla i sons

Dedicació: 16h

Grup gran/Teoria: 8h

Aprenentatge autònom: 8h

Processament multimicròfon d'àudio

Descripció:

Acústica de sales
Conformació de feix amb agrupació de micròfons
Localització de font acústica i seguiment

Dedicació: 12h

Grup gran/Teoria: 6h

Aprenentatge autònom: 6h

Anàlisi i síntesi de senyals d'àudio

Descripció:

Anàlisi i síntesi localitzades de senyals (quasi)periòdics. Modificació de l'escala temporal i el to
Síntesi espacial d'àudio amb funcions HRTF

Objectius específics:**Dedicació:** 8h

Grup gran/Teoria: 4h

Aprenentatge autònom: 4h

Realització i presentació del projecte

Descripció:

Disseny, implementació i prova d'un sistema de proc. d'àudio per a una aplicació específica
Presentació oral de 1) proposta, 2) revisió crítica, i 3) realització i conclusions.
Espectura d'un informe tipus article de congrés

Dedicació: 48h

Grup gran/Teoria: 4h

Aprenentatge autònom: 44h

ACTIVITATS

Tasques per fer a casa i tests curts

Dedicació: 25h

Aprenentatge autònom: 25h

Projecte d'assignatura

Descripció:

Realització i presentació d'un projecte en equip, que inclou treball experimental en processament d'àudio, i es presenta tant de forma oral com escrita.

Dedicació: 60h

Aprenentatge autònom: 60h

Presentacions orals

Descripció:

Presentació oral del projecte en tres etapes: proposta, revisió i final. Presentacions de treballs menors.

Dedicació: 1h

Aprenentatge autònom: 1h

SISTEMA DE QUALIFICACIÓ

Assistència/participació a classe (10%)

Tasques a casa, petits projectes, i tests (30%)

Projecte d'assignatura (60%)

BIBLIOGRAFIA

Bàsica:

- Quatieri, T.F. Discrete-time speech signal processing: principles and practice. New Delhi: Prentice Hall, 2006. ISBN 9788177587463.
- Gold, B.; Morgan, N.; Ellis, D. Speech and audio signal processing: processing and perception of speech and music [en línia]. 2nd ed. Wiley-Blackwell, 2011 [Consulta: 30/06/2022]. Disponible a: <https://onlinelibrary-wiley-com.recursos.biblioteca.upc.edu/doi/book/10.1002/9781118142882>. ISBN 978-0-470-19536-9.
- Dutoit, T.; Marqués, F. Applied signal processing: a MATLAB-based proof of concept [en línia]. New York ; London: Springer, 2009 [Consulta: 30/06/2022]. Disponible a: <https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/978-0-387-74535-0>. ISBN 9780387745343.
- Loizou, P.C. Speech enhancement: theory and practice. 2nd ed. Boca Raton: CRC Press, 2013. ISBN 9781466504219.

Complementària:

- Rabiner, L.R.; Schafer, R.W. Theory and applications of digital speech processing. London: Pearson, 2010. ISBN 9780137050857.
- Huang, Y.A.; Benesty, J. (eds.). Audio signal processing for next-generation multimedia communication systems [en línia]. New York: Kluwer Academic Publishing, 2004 [Consulta: 30/06/2022]. Disponible a: <https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b117685>. ISBN 1402077688.

RECURSOS

Material audiovisual:

- Slides. Diapositives usades a les classes

Material informàtic:

- Codi programes. Codi programes en Matlab o similar

Altres recursos:

Presentacions usades en les classes
Enunciats i programes del treball pràctic