



Pla Docent de l'Assignatura

Guia Docent

Nom de l'assignatura: Processament de la Parla

Curs acadèmic: 2012-2013

Curs: Segon

Trimestre: Tercer

Estudis: Grau en Enginyeria en Informàtica (Optativa), Grau en Enginyeria en Telemàtica (Optativa) i Grau en Enginyeria en Sistemes Audiovisuals (Obligatòria)

Codi assignatura: 21480-21762-21610

Nombre de crèdits ECTS: 4

Nombre total d'hores de dedicació: 100 hores

Llengua o llengües de docència: Català (explicacions), anglès (material) Elijà un element.

Professorat: Emilia Gómez, Waldo Nogueira, Martí Umbert

Professorat responsable: Emilia Gómez

1. Dades descriptives de l'assignatura

Nom de l'assignatura: Processament de la Parla

Curs acadèmic: 2012-2013

Curs: Segon

Trimestre: Tercer

Estudis: Grau en Enginyeria en Informàtica (Optativa), Grau en Enginyeria en Telemàtica (Optativa) i Grau en Enginyeria en Sistemes Audiovisuals (Obligatòria)

Codi assignatura: 21480-21762-21610

Nombre de crèdits ECTS: 4

Nombre total d'hores de dedicació: 100 hores

Llengua o llengües de docència: Català (explicacions), anglès (material) *Elija un elemento.*

Professorat: Emilia Gómez, Waldo Nogueira, Martí Umbert

Professorat responsable: Emilia Gómez

2. Presentació de l'assignatura

Aquest és un curs intermitj al processament digital del senyals sonores, pensat per estudiants de Enginyeria en Sistemes Audiovisuals.

El curs està construït sobre la base de assignatures prèvies de la titulació, principalment Enginyeria Acústica i Senyals i Sistemes (segón curs, Enginyeria en Sistemes Audiovisuals).

Ens centrarem en l'estudi de les principals tècniques d'anàlisi, descripció, síntesi i processament de senyals de veu.

3. Competències a assolir a l'assignatura

Competències¹ a treballar a l'assignatura segons l'indicat en el pla d'estudis del grau.

Competències transversals	Competències específiques
<i>Instrumentals</i>	<i>Competències Específiques Professionals</i>
G1. Capacitat d'anàlisi i síntesi	Elija un elemento.
G2. Capacitat d'organització i planificació	<i>Competències Específiques de Formació Bàsica</i>
G3. Capacitat per aplicar els coneixements a l'anàlisi de situacions i la resolució de problemes	B4-INF. Capacitat per analitzar funcions de variable complexa.
G4. Habilitat en la cerca i la gestió	B7-INF. Conèixer les transformades de Fourier per a senyals analògics i digitals i els elements bàsics de la teoria del mostreig de senyals.

¹ **Competències transversals:** Aquelles que es requereixen en l'exercici de qualsevol titulació o carrera (comunicació verbal i escrita, pensament analític i sistèmic, resolució de problemes, creativitat, etc.). Es classifiquen en:

- **Instrumentals:** Inclouen habilitats cognitives, metodològiques, tecnològiques i lingüístiques. (Ex: capacitat d'organitzar i planificar, capacitat de comunicar-se amb propietat de forma oral i escrita en català, castellà i/o anglès, tant davant d'audiències expertes com inexpertes).
- **Interpersonals:** Es defineixen com habilitats que tendeixen a facilitar els processos d'interacció social i cooperació. (Ex: capacitat per a treballar en grup, expressió del compromís ètic/social).
- **Sistèmiques o integradores:** Suposen una combinació de comprensió, sensibilitat i coneixement que permeten veure com s'agrupen i s'estableixen relacions entre les parts d'un tot. Aquestes competències requereixen, com a base, l'adquisició prèvia de competències instrumentals i interpersonals. (Ex: capacitat d'adaptar-se a nous contextos d'aprenentatge)

Competències específiques: Es relacionen amb els coneixements i pràctiques concretes del grau. (Ex: capacitat de descriure, programar, validar i optimitzar protocols i interfaces de comunicació en els diferents nivell d'una arquitectura de xarxes)

de la informació G5. Habilitat en la presa de decisions Elija un elemento de la lista Capacitat de comunicar-se en context acadèmic de forma oral i escrita en català, castellà i anglès <i>Interpersonals</i> G8. Capacitat de treball en equip Elija un elemento. Capacitat de treball en contextos internacionals i interdisciplinaris. <i>Sistèmiques</i> G11. Capacitat d'aplicar amb flexibilitat i creativitat els coneixements adquirits i d'adaptar-los a contextos i situacions noves G12. Capacitat per progressar en els processos de formació i aprenentatge de manera autònoma i contínua Elija un elemento.	B8-INF. Capacitat per resoldre sistemes lineals i invariants i les funcions i transformades relacionades. B7-T. Comprendre i utilitzar els principis de la probabilitat, els B4-A. Capacitat per analitzar funcions de variable complexa. B7-A. Conèixer les transformades de Fourier per a senyals analògics i digitals. Conèixer els elements bàsics de la teoria del mostreig de senyals. Capacitat per resoldre sistemes lineals i invariants i les funcions i transformades relacionades. B9-A. Adquirir els coneixements bàsics de la física de la propagació del so i la seva relació amb els mètodes de processament de l'assenyalis àudio. Elija un elemento. <i>Competències Específiques d'Enginyeria en Informàtica</i> Elija un elemento. <i>Competències Específiques Comunes a la branca de Telecomunicació</i> Elija un elemento. <i>Competències de tecnologia específica: Enginyeria Telemàtica</i> Elija un elemento. <i>Competències de tecnologia específica: Sistemes Audiovisuals</i> AU1. Capacitat de construir, explotar i gestionar serveis i aplicacions de telecomunicacions, enteses aquestes com a sistemes de captació, tractament analògic i digital, codificació, transport, representació, processament, emmagatzematge, reproducció, gestió i presentació de
---	--

	serveis audiovisuals i informació multimèdia. AU3. Capacitat de seleccionar un sistema, component o procés de l'àmbit del so i imatge per complir les especificacions requerides. AU4. Adquirir el coneixement de les tècniques de processament d'àudio necessàries per al desenvolupament de sistemes de processament acústic. AU5. Coneixement de les tècniques de tractament del senyal utilitzades en l'anàlisi, síntesi, codificació i reconeixement del parla. Desenvolupament d'aplicacions informàtiques per al tractament del parla. AU6. Coneixements de les tècniques de tractament de senyals d'àudio i música. Desenvolupament d'aplicacions informàtiques basades en el processament de música. AU22. Entendre i conèixer els principis matemàtics en els quals es basen els sistemes de codificació d'àudio i música. Entendre i conèixer els avantatges i inconvenients de cadascun d'aquests sistemes de codificació d'àudio i música. Elija un elemento.
--	---

4. Continguts

En aquesta assignatura es vol que l'alumne desenvolupi una metodologia per a analitzar, codificar, reconèixer i sintetitzar senyals de veu (principalment de parla) utilitzant tècniques de processament del senyal. Més concretament es pretén que l'estudiant adquereixi competències teòriques i pràctiques relacionades amb:

- Fonaments acústics, fisiològics i perceptuals de la veu.
- Fonaments de l'anàlisi digital de senyals de veu.
- Mètodes de modelat i processament de senyals de veu.

- Us de programari de processament de veu.
- Implementació, amb un llenguatge de programació, d'algorismes de processament del senyal pel processament de la veu.

Aquests conceptes s'estructuren en els següents **Blocs de Contingut**:

Bloc 1. Introducció:

- Cadena de generació/percepció de la parla.
- Fonaments acústics del so.
- Aplicacions del processament digital de la parla.

Bloc 2. Fonaments:

- Fonaments acústics de la veu: mecanismes de producció de la veu, parla vs cant, classificació dels sons de la parla, representació fonètica.
- Fonaments perceptuals: alçada, intensitat i timbre.

Bloc 3. Anàlisi espectral de senyals de veu.

- Anàlisi espectral de senyals de veu: STFT, espectrograma, anàlisi multiresolució.
- Extracció de paràmetres: energia, ZCR, ST-ACF, pitch.

Bloc 4. Models de veu: models basats en la percepció.

- Models físics vs espectrals vs basats en formants.
- Fonaments de percepció de la parla.
- Transformació espectral de senyals de veu.

Bloc 5. Models basats en la producció de la parla. Anàlisi Linear Predictiu (LPC).

Bloc 6. Síntesi Texte-Parla (*Text-to-Speech Synthesis*).

Bloc 7. Reconeixement de parla (*Automatic Speech Recognition*).

- Anàlisi cepstrum.
- Models Ocults de Markov.

5. Avaluació del nivell d'assoliment de les competències

Criteris generals d'avaluació

L'avaluació es fa per a cada una de les tres activitats que constitueixen l'assignatura: classes de teoria, seminaris i laboratoris. Essent:

- T: l'avaluació de la teoria mitjançant:
 - Control de coneixements al mig del trimestre (30% de T). No recuperable.
 - Examen final (70% de T). L'examen final avalua tota la matèria de l'assignatura i és recuperable al juliol. Inclou preguntes associades a les pràctiques. Recuperable.
 - Es requereix un mínim de 5/10 a la nota de Teoria per a aprovar l'assignatura.

- L: l'avaluació dels laboratoris amb l'entrega de les memòries de pràctiques (35% de L) (de forma individual o en parelles) i una entrevista de pràctiques al llarg del trimestre (5% de L). No recuperable. Es requereix un mínim de 5/10 a la nota de Laboratoris per a aprovar l'assignatura.
- S: l'avaluació dels seminaris amb la participació en aquests. No recuperable.

La nota final s'obté fent la mitjana ponderada de la següent manera:

$$\text{Nota Final} = 0,5 * T + 0,4 * L + 0,1 * S$$

6. Bibliografia i recursos didàctics

6.1. Fonts d'informació per a l'aprenentatge. Bibliografia bàsica (suport paper i electrònic)

- Quatieri, T. F. 2001. Discrete-Time Speech Signal Processing: Principles and Practice. Prentice Hall.
- Rabiner, L. R. and R. W. Schafer. 2007. Introduction to Digital Speech Processing. Foundations and Trends in Signals Processing, Vol. 1, Nos. 1-2, 2007.

6.2. Fonts d'informació per a l'aprenentatge. Bibliografia complementària (suport paper i electrònic)

- Rabiner, L. R. and R. W. Schafer. 1978. Digital Signal Processing of Speech Signals. Prentice Hall.
- O'Shaughnessy, D. 1999. Speech communications: human and machine. Wiley, John & Sons.
- Rabiner, L. R. and B. H. Juang. 1993. Fundamentals of Speech Recognition. Prentice Hall.
- Park, Sung-won. Linear Predictive Speech Processing.
- Park, Sung-won. Discrete Wavelet Transform.
- Spanias, Andreas. 1994. "Speech Coding: A Tutorial Review". Proceedings of the IEEE.
- Pan, Davis. 1995. "A Tutorial on MPEG/Audio Compression". IEEE Multimedia Journal.
- Rabiner, Lawrence. 1989. "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition". Proceedings of the IEEE.

6.3. Recursos didàctics. Material docent de l'assignatura

- Per a cada sessió de teoria hi ha disponible uns apunts a la Web de l'assignatura.
- Per a cada sessió de seminari hi ha disponible una col·lecció d'activitats a la Web de l'assignatura.
- Per a cada sessió de pràctiques hi ha disponible el enunciat de la pràctica a la Web de l'assignatura.

6.4. Programari

- PRAAT <http://www.fon.hum.uva.nl/praat/>
- Octave <http://www.gnu.org/software/octave/>
- MATLAB

7. Metodologia

7.1. Enfocament metodològic de l'assignatura

El procés habitual d'aprenentatge per cada un dels temes està compost per una sessió de teoria, un seminari i una part d'una sessió de laboratori. Cada bloc comença amb una sessió de teoria en la que es presenten certs fonaments teórico-pràctics. Aquesta activitat es realitza en grup gran. L'estudiant ha de complementar aquesta activitat amb una lectura detinguda dels seus propis apunts i del material addicional que el professor hagi proporcionat. Per exemple, una sessió de teoria de 2 hores, convenientment aprofitades, requerirà un treball addicional fora de l'aula d'1 hora per part de l'estudiant.

Posteriorment es realitzarà un seminari centrat en la resolució d'exercicis o problemes per posar en pràctica els conceptes i tècniques presentades a la sessió de teoria. L'objectiu és que l'estudiant consolidi els fonaments per tal que posteriorment pugui resoldre problemes de major complexitat. Aquesta activitat es realitza en grups petits de seminari on tots els estudiant han de participar activament en la resolució dels problemes.

El següent pas en el procés d'aprenentatge és el laboratori o sessió de pràctiques. En ell es proposen uns problemes pràctics que requereixen un disseny previ de la solució a implementar i que han d'integrar diferents conceptes i tècniques. Aquesta activitat es pot realitzar per parelles, i es pressuposa que continua fora de l'aula.

L'últim pas en el procés d'aprenentatge de cada bloc de continguts és el de la realització del control de coneixements i l'examen final per comprovar si l'estudiant ha adquirit les competències demanades.

Set.	Tema	Sessions presencials			Hores d'estudi
		Grup gran (2h)	Grup petit (1h)	Grup mitjà (2h)	
1	1. Introducció 2. Fonaments	2	1		5
2	3. Anàlisi espectral	1	1		5
3	3. Anàlisi espectral			1	5
4	4. Models perceptuals	1		1	5
5	5. Models producció	1	1	1	5
6	6. <i>Text-To-Speech Synthesis</i>	1	1		7 (control de coneixements)
7	7. Anàlisi cepstral	1	1	1	5
8	8. Reconeixement	1	1		6
9	8. Reconeixement	1	1	1	6
10	Recapitulació		1		8
	Preparació de l'examen final				7

	Total	18	8	10	64
--	-------	----	---	----	----

Teoria: es corresponen amb els blocs de continguts esmentats. 18 hores (9 sessions de 2 hores).

- Teoria 1: Introducció.
- Teoria 2: Producció i classificació de sons de parla.
- Teoria 3: Anàlisi espectral.
- Teoria 4: Models de la parla.
- Teoria 5: Models basats en la producció de veu (LPC).
- Teoria 6: Síntesi de la Parla: *Text-To-Speech*.
- Teoria 7: Reconeixement de parla: Anàlisi cepstrum, Sistemes de reconeixement automàtic.
- Teoria 8: Reconeixement de parla: Models ocults de markov (I).
- Teoria 9: Reconeixement de parla: Models ocults de markov (II).

Seminaris: 8 sessions de 1 hora.

- Seminari 1: Acústica de la veu.
- Seminari 2: Anàlisi espectral.
- Seminari 3: LPC.
- Seminari 4: Control de coneixements.
- Seminari 5: Cepstrum.
- Seminari 6: Reconeixement de la parla.
- Seminari 7: Transformacions de la parla.
- Seminari 8: Revisió de coneixements.

Laboratoris: 5 sessions de 2 hores.

- Laboratori 1: Enregistrament i anàlisi espectral de veu.
- Laboratori 2: Model espectral.
- Laboratori 3: Anàlisi i síntesi de veu.
- Laboratori 4: Anàlisi Cepstrum.
- Laboratori 5: Reconeixement de la Parla.

8. Programació d'activitats

Programació de sessions presencials

Professors: [Emilia Gómez](#), [Waldo Nogueira](#), [Martí Umbert](#)

	Dimarts 14.30-16.30	Dijous 16.30-18.30	Divendres 18.30-20.30
1 08-12 abr	9/04 Teoria 1: Introduction	11/04 Teoria 2: Production and Classification of Speech Sounds	12/04 Seminari 1: Voice Acoustics (1 h S101, 1h S102)
2 15-19 abr	16/04 Seminari 1: Voice Acoustics (1 h S103, 1h S104)	18/04 Teoria 3: Short-Time Analysis of Speech	19/4 Seminari 2: Spectral Analysis (1 h S103, 1h S104)

3 22-26 abr	23/04 FESTIU	25/04 Seminari 2: Spectral Analysis (1 h S101, 1h S102) Lab 1: Speech Analysis (2h G1)	26/04 Lab 1: Speech Analysis (2h G2)
4 29 abr- 3maig	30/04 Teoria 4: Voice Models	2/05 Lab 2: Sinusoidal Models (2h G2)	3/05 Lab 2: Sinusoidal Models (2h G1)
5 6-10 maig	7/05 Teoria 5: Source-filter Models	9/05 Seminari 3: LPC (1 h S104, 1h S103) Lab 3: LPC (2h G1)	10/05 Seminari 3: LPC (1 h S102, 1h S101) Lab 3: LPC (2h G2)
6 13-17 maig	14/05 Teoria 6: Text-to-Speech Synthesis	16/05 Seminari 4: Examen (1 h S102, 1h S101)	17/05 Seminari 4: Examen (1 h S104, 1h S103)
7 20-24 maig	21/05 Teoria 7: Cepstral Analysis	23/05 Seminari 5: Cepstrum (1 h S101, 1h S102) Lab 4: Cepstrum (2h G2)	24/05 Seminari 5: Cepstrum (1 h S103, 1h S104) Lab 4: Cepstrum (2h G1)
8 27-31 maig	28/05 Teoria 8: HMMs and Automatic Speech Recognition	30/05 Seminari 6: Automatic Speech Recognition (1 h S103, 1h S104)	31/05 Seminari 6: Automatic Speech Recognition (1 h S101, 1h S102)
9 3-7 juny	4/06 Teoria 9: HMMs and Automatic Speech Recognition	6/06 Seminari 7: Transformation (1 h S104, 1h S103) Lab 5: Speech Recognition (2h G1)	7/06 Seminari 7: Transformation (1 h S102, 1h S101) Lab 5: Speech Recognition (2h G2)
10 10-14 juny	11/06 Seminari 8: Simulacre d'examen i dubtes (1 h S101, 1h S102)	13/06 Seminari 8: Simulacre d'examen i dubtes (1 h S103, 1h S104)	14/06

- Llistat d'activitats (amb termini de lliurament i avaluació)
 - A cada seminari hi hauran petits exercicis que es lliuraran avant del seminari següent.
 - A cada laboratori hi haurà una memòria de pràctiques a lliurar avant del laboratori següent.

- iii. La sol·lució al control de coneixements es donará al final del seminari.

Activitat	Data enunciat	Data lliurament	Data de lliurament de resultats
Laboratori 1	25 d'Abril	1 de Maig	13 de Maig
Laboratori 2	2 de Maig	8 de Maig	20 de Maig
Laboratori 3	9 de Maig	15 de Maig	27 de Maig
Laboratori 4	23 de Maig	29 de Maig	10 de Juny
Laboratori 5	6 de Juny	11 de Juny	15 de Juny
Control	16/17 de Maig	16/17 de Maig	24 de Maig