

FISA DISCIPLINEI

Denumirea disciplinei	Prelucrarea Semnalului Vocal
Domeniul de studiu	Inginerie electronica si telecomunicatii
Specializarea	Tehnologii si sisteme de telecomunicatii
Codul disciplinei	51325308-2
Titularul disciplinei	Prof.dr.ing. Mircea Giurgiu – Mircea.Giurgiu@com.utcluj.ro
Colaboratori	
Catedra	Comunicatii
Facultatea	Electronica, Telecomunicatii si Tehnologia Informatiei

Sem.	Tipul disciplinei	Curs			Aplicații			Stud. Ind.	TOTAL	Credit	Forma de verificare		
		[ore/săpt.]			[ore/sem.]								
		S	L	P	S	L	P						
8	Ing. din domeniu	2	-	1	1	28	-	14	14	64	120	4	Examen

Competențe dobândite:

Cunoștințe teoretice, (Ce trebuie sa cunoască)

- Sa cunoasca modelarea numerica a semnalului vocal si metodele de analiza in domeniul timp / frecventa / cepstrum a semnalului vocal;
- Sa cunoasca procedurile de extragerea parametrilor din unda acustica (LPC, PARCOR, LSF) si sistemele de procesare avansata a semnalului vocal (TESPAR, Wavelet);
- Sa cunoasca standardele si aplicatiile de comunicatii pentru codarea in domeniul timp (PCM, ADPCM, IMA-ADPCM) si in domeniul frecventa (MELP, MBE); Standardele MPEG pentru codarea semnalelor audio (Layer I, II, III, Algoritmul AC-3); Sistemele de codare folosind analiza prin sinteza (MPE, RPE-LTP, CELP si ACELP);
- Sa cunoasca algoritmi de compresie la debit redus a semnalului vocal (FS1060, GSM) si aplicatiile pentru comunicatiile mobile (codecul GSM si GSM la rata injumatatita) si pentru telefonia VoIP;
- Sa cunoasca metodele de reducere a zgomotului si a ecoului din semnalul vocal; Conceptele si metodele de implementare a cuantizarii vectoriale pentru compresia de semnale;
- Sa cunoasca principiile sintezei din text a semnalului vocal: metode de sinteza, controlul automat al parametrilor semnalului in tehnicile PSOLA si TD-PSOLA;
- Sa cunoasca principiile recunoasterii automate a vorbirii: metode (alinierea dinamica, modelarea HMM, retele neuronale de tip MLP), exemple de aplicatii.

Deprinderi dobândite: (Ce știe să facă)

După parcurgerea disciplinei studenții vor fi capabili:

- Sa proiecteze un sistem complet de achizitie si redare de semnal vocal, conectabil la un sistem de calcul
- Sa identifice si sa interpreteze principalele caracteristici ale unei vocale, impreuna cu parametrii specifici (frecventa fundamentala, formanti, distributie statistica a amplitudinii, etc)
- Sa fie capabili sa implementeze principalele metode si algoritmi de analiza in domeniul timp / frecventa / cepstral si sa extraga informatii relevante pe care sa le interpreteze corect
- Sa interpreteze cu usurinta formele de unda si spectrogramele extrase din semnal
- Sa cunoasca foarte bine principiile si sa implementeze diferitele tipuri de codecuri pentru semnalul vocal: LPC, ADPCM, in subbezi, etc. de asa maniera incat rezultatele sa poata fi incorporate in aplicatii specifice de telecomunicatii
- Sa fie capabili sa aplice corect (studiu, proiectare, implementare, testare) algoritmi de filtrare adaptiva a semnalului vocal, in vederea reducerii zgomotului si eliminarii ecoului
- Sa poata lua decizii corecte privind selectarea transformatei Wavelet pentru: estimarea frecventei fundamentale si pentru reducerea zgomotului din semnalul audio.

Abilități dobândite: (Ce echipamente, instrumente știe să mânăiască)

După parcurgerea disciplinei studenții vor fi capabili:

- Sa opereze cu usurinta in mediul de programare Matlab, pentru implementarea aplicatiilor de prelucrare de semnal vocal si sa poata realiza diferite experimente in acest mediu, in conformitate cu lucrarile de laborator
- Sa stie sa testeze orice sistem de compresie de semnal audio si sa-i stabileasca performantele intr-o maniera corecta.
- Sa stie sa utilizeze placile de interfatare cu linia telefonica din familia Dialogic 4JT, precum si un set redus de functii CTAPI pentru comanda acestora
- Sa fie capabil sa proiecteze si sa selecteze tehnologia potrivita pentru o arhitectura complexa, in sisteme de informare automata.

FISA DISCIPLINEI

Cerințe prealabile (Dacă este cazul)	
•	Cunoștințe de teoria semnalelor și sistemelor: semnale analogice/numerice, caracterizarea semnalelor în domeniul s și în z , transformata Fourier a semnalelor, funcții de transfer și caracteristica de transfer, stabilitatea sistemelor.
•	Cunoștințe despre prelucrarea numerică a semnalelor: reprezentarea numerică a semnalelor, filtre numerice, filtre FIR și IIR.
•	Cunoștințe de teoria informației: informație, entropie, informație mutuală, sisteme pentru codarea informației, modelarea surselor de informație cu lanțuri Markov, principii ale compresiei informației.
•	Cunoștințe de tehnici de comunicație: caracterizarea canalelor de comunicație, protecția la erori, modularea semnalelor digitale, sisteme simple de codare a formelor de undă.
•	Cunoștințe de programare: fundamentele programării în C/C++, Matlab la nivel de bază, programare obiectuală.

A. Curs (titlul cursurilor + programa analitică)	
1	Caracteristici ale semnalului vocal (caracteristici acustice, caracteristici statistice, modele de producere a semnalului vocal)
2	Modelarea numerică a semnalului vocal (reprezentarea numerică a semnalului vocal, restructurii privind esanționarea și cuantizarea, refacerea semnalului vocal)
3	Metode de analiză în domeniul timp (amplitudinea maximă, energia, numărul de treceri prin zero, autocorelația, funcția AMDF, metoda TESPARG, estimarea frecvenței fundamentale în domeniul timp)
4	Metode de analiză în domeniul frecvență (analiză prin transformata Fourier rapidă, analiză și interpretarea subbenzilor, alte metode de analiză)
5	Analiză cepstrală și analiză prin transformata wavelet (analiză homomorfică, principiile estimării frecvenței fundamentale și a funcției de transfer a tractului vocal, coeficienții cepstrali)
6	Tehnici de codare în domeniul timp și aplicații în VoIP (PCM, APCM, DPCM, ADPCM și ADPCM de bandă largă, sisteme de tip VoIP)
7	Sisteme de codare în subbenzi (schema de codare și decodare în subbenzi, rolul blocurilor, legătura cu FFT, tipuri de codare vocale folosind analiză în subbenzi)
8	Codarea semnalului vocal în standardul MPEG (principiile codării, efectul de mascare, MPEG1: Layer 1, Layer 2, Layer 3, MPEG2, MP3, algoritmul AC3, compresia semnalului vocal pentru aplicații multimedia)
9	Sisteme de codare folosind analiză prin sinteză (principiul codării folosind analiză prin sinteză, definirea blocurilor componente, exemple de codare și performanțele acestora).
10	Codarea semnalului vocal în GSM (schema bloc, rolul componentelor, codorul-decodorul, protecția la erori și imunitatea mesajelor în GSM, blocul VAD, codarea la rată înjumătățită)
11	Compresia semnalului vocal prin cuantizare vectorială (principiile cuantizării vectoriale, rata codorului, proiectarea dicționarului, algoritmi Lloy și LBG, tipuri de cuantizare vectoriale)
12	Principiile sintezei din text a semnalului vocal (schema bloc, analiză lingvistică, sinteză de semnal, analiză și generarea automată a prozodiei, algoritmi de modificare a frecvenței fundamentale: PSOLA/TD-PSOLA).
13	Tehnici de bază în recunoașterea automată a vorbirii (modele pentru recunoașterea automată a vorbirii: DTW, ANN-MLP, HMM, modelarea statistică, antrenarea și testarea sistemelor, exemple).
14	Sinteza cursului și pregătire pentru examen

B1. Aplicații – LUCRARI (lista lucrări, teme de seminar, conținutul proiectului de an)	
Lucrări de laborator (1h / săptămână)	
1	Algoritm robust pentru detectia liniste/vorbire pe baza energiei și a numărului de treceri prin zero
2	Analiză semnalului vocal prin funcția de autocorelație și funcția AMDF. Estimarea frecvenței fundamentale
3	Analiză spectrală a semnalului vocal și estimarea frecvenței fundamentale prin analiză cepstrală.
4	Analiză prin predicție liniară LPC. Sinteza de semnal din coeficienții de predicție.
5	Sisteme de codare a semnalului vocal în subbenzi și codarea IMA-ADPCM
6	Implementarea unui codor GSM RPE-LTP și evaluarea performanțelor
7	Metode de modificare a frecvenței fundamentale: PSOLA și TD-PSOLA.

FISA DISCIPLINEI

Exemple de proiecte de an (1h / saptamana)		
<ul style="list-style-type: none"> - Filtre adaptive utilizate in prelucrarea semnalului vocal - Studiul compresiei semnalului vocal prin transformata Wavelet - Reducerea zgomotului din semnalul vocal folosind transformata Wavelet - Software educational pentru codarea TESPAP - Aplicarea retelelor neuronale pentru predictia parametrilor acustici la sinteza TTS - Studii privind variatia duratei pauzelor in vorbire (elaborarea unui model) - Aplicarea operatorului TEO in analiza semnalului vocal - Codarea MELP - Estimarea robusta a frecventei fundamentale folosind Transformata Wavelet - Codorul ACELP, ITU G. 729 - Algoritmul AC-3 utilizat in codarea MPEG-3 - Chat vocal in Internet folosind componente Macromedia Flash comunicator - Aplicatie TTS cu tehnologii Microsoft Speech Server - Metode de control a parametrilor acustici la sinteza text-vorbire - Recunosaterea automata a vocalelor folosind TESPAP si Rețele Neuronale - Modelul Fujisaki pentru controlul parametrilor acustici la sinteza TTS - Studiul standardului VXML - Implementarea unui codor CELP, etc. 		
B2. Sala laborator (Sala/suprafata, adresa): Sala 508 / 48m², Observator		
Echipament	Descriere echipament	Anul achizitiei
Retea de calculatoare (11 buc)	<ul style="list-style-type: none"> • Calculatoare PC- Procesor Pentium IV, frecventa 1,8 ... 3 GHz (in functie de anul achizitiei), Memorie: tipic 256 MB, HDD: 20GB, Multimedia: placi de sunet si casti pentru fiecare sistem. • Servere: Linux, Windows 2003 Server. 	2003 - 2006
Software:	<ul style="list-style-type: none"> • Mediul de programare Matlab 7.0, • Software pentru editare/prelucrare fisiere audio: Soundforge 7.0, Goldwave, Speech Filling System – produs de University College of London, • Microsoft Speech Server, • PRAAT – mediu distribuit de Universitatea din Delft, • HTK – Hidden Markov Models Toolkit, • VOICEBOX – biblioteca de functii Matlab, MAD – Matlab Auditory Demonstrator produs la Sheffield Univ., • Alte produse software educationale, dezvoltate in laborator 	2003-2006.
Placa de conectare la linia telefonica:	<ul style="list-style-type: none"> • Intel Dialogic JCT. 	2005
Instrumente de vizualizare	<ul style="list-style-type: none"> • Osciloscop 	1996
Placi de achizitie de semnal	<ul style="list-style-type: none"> • Placa bazata pe MC3417, • Sistem de dezvoltare cu procesor de semnal TMS320C50 	1994

C. Studiul individual (tematica studiilor bibliografice, materiale de sinteză, proiecte, aplicații etc.)						
<ul style="list-style-type: none"> - Studiu individual pe baza: materialului de curs, lucrarilor de laborator si documentare pentru elaborarea proiectului de an. - Realizarea unor sinteze cu rezultate stiintifice si experimentale privind lucrarile de laborator - Documentare individuala in Internet sau folosind resursele din laborator (carti, articole stiintifice, proiecte din anii precedenti, CD si benzi video). Colaborare cu colegii pentru rezolvarea in echipa a unor experimente sau pentru realizarea proiectului. 						
Structura studiului individual	Studiu materiale curs	Rezolvări teme, lab., proiecte	Pregătire aplicații	Timp alocat examenărilor	Studiu bibliografic suplimentar	Total ore pregătire individuală
Nr. ore	20	26	7	3	8	64

D. Strategii si metode de predare
Curs: prezentare interactiva, dialogul, demonstratia, descoperirea
Laborator / Proiect: Demonstratia, experimentul, implementarea si experimentarea, descoperirea.

FISA DISCIPLINEI

Bibliografie (Cursuri, îndrumătoare de lucrări, proiect, culegeri de probleme)
In biblioteca UTC-N: <ol style="list-style-type: none">1. M. Giurgiu, <i>Compresia semnalului vocal in aplicatii multimedia</i>, Ed. Risoprint Cluj-Napoca, 2003.2. M. Giurgiu, <i>Sinteza din text a semnalului vocal. Vol I.</i>, Editura Risoprint, Cluj-Napoca, 2006.3. G. Stolojanu, <i>Prelucarea numerica a semnalului vocal</i>, Ed Militara, 19844. M. Draganescu, <i>Analiza si sinteza semnalului vocal</i>, Ed. Academiei, 19865. S. Furui, <i>Advances in Speech Signal processing</i>, Marcel Dekker, 19956. D.G. Childers – <i>Speech Processing and Synthesis Toolboxes</i>, Wiley Publ., 1999
Materiale didactice virtuale: <ol style="list-style-type: none">7. M. Giurgiu - <i>Compresia semnalului vocal in aplicatii multimedia</i>, www.concorde.utcluj.ro8. O serie de resurse din Internet, recomandate de catre grupul de specialisti din European Language and Speech Network
In alte biblioteci (in laborator): <ol style="list-style-type: none">9. K. Ponting, <i>Computational Models for Speech Pattern Processing</i>, Springer Verlag, 199710. S. Furui, <i>Digital Speech processing, Synthesis and Recognition</i>, Marcel Dekker, 198911. E. Keller, <i>Fundamentals of Speech Synthesis and Speech Recognition</i>, John Willey & Sons, 199412. A. M. Kondoz, <i>Digital Speech Coding for Low Bit Rate Communications Systems</i>, John Willey, 1995 L13. L.R. Rabiner, <i>Digital Processing of Speech Signals</i>, Prentice Hall, 197814. P. Papamichalis, <i>Practical Approaches to Speech Coding</i>, Prentice Hall, 1995

Modul de examinare și atribuire a notei	
Modul de examinare	Examenul constă dintr-un test grilă (1 oră) și o evaluare scrisă a cunoștințelor teoretice + probleme (2 ore); După cursul 7 se poate susține un examen parțial (1,5 ore). Rezultatele experimentale de la un laborator se corectează și se notează în săptămâna imediat următoare laboratorului. Nu se admit întârzieri. Proiectul se desfășoară și se verifică în etape precizate la preluarea acestuia. Predarea proiectului este planificată în ultima săptămână de școală.
Componentele notei	Laborator (nota L); Proiect (nota P); Examen (nota E)
Formula de calcul a notei	$N=0,6E+0,2L+0,2P$;

Responsabil de disciplina:

Prof.dr.ing. Mircea GIURGIU