SYLLABUS

Discipline name	Digital Speech Processing
Profile	Electronics and Telecommunications Engineering
Specialization	Telecommunications Technologies and Systems
Code	51325309-2
Course leader	Professor Mircea Giurgiu, Ph.D – Mircea.Giurgiu@com.utcluj.ro
Collaborators	Eng. Adriana Stan, Ph.D. student, Adriana.Stan@com.utcluj.ro
Department	Communications
Faculty	Electronics, Telecommunications and Information Technology

Sem.	Type of discipline	Course	Applications			Course	Applications			Ind. study	TOTAL	Credits	Form of assessment
		[hours/week]				[hours/semester]							
			S	L	P		S	L	P				
8	**Speciality, Optional**	**2**	**-**	**1**	**1**	**28**	**-**	**14**	**14**	**94**	**150**	**5**	**V**

Acquired competences :
Acquired skills (what the student is able to do):
• to design a speech acquisition system connected to the PC • to identify and to interpret the main features of the speech signal, including statistical distribution • to implement the algorithms for speech processing in time, frequency, domain and to interpret the results • to be able to interpret the processing flows for various speech coders: LPC, ADPCM, subband, RPE-LTP, etc. • to be able to apply the most suitable schemes for adaptive filtering and noise enhancement of speech • to be able to consider the appropriate decisions the speech processing flows in any telecom application.
Acquired abilities (what type of equipment/ instruments/ software the student is able to handle):
• to handle the specific tools for speech processing: CoolEdit, PRAAT, SFS, VoiceBox, WaveSurfer, MAD. • to easily program in Matlab (or similar signal processing environments) the specific applications for speech processing and to run dedicated experiments that involves data analysis and interpretation • to test most of the speech codecs with the aim to take appropriate implementation decisions in specific telecom environments (fixed lines, optical fibers, VoIP, mobile communications, etc) and other speech applications. • to use specific Matlab toolboxes dedicated to speech processing: spectral analysis, wavelet analysis, entropic coding of speech. • to implement and to experiment specific algorithms for: speech recognition, speech synthesis, speech coding.

Prerequisites (if necessary):
Knowledge about: a) signals and systems: analog/digital signals, Fourier transform transfer functions, system stability; b) digital signal processing: digital filters, FIR, IIR; c) information theory: entropy, mutual information, Markov modeling of information sources, information coding/compression; d) digital communications: errors, modulations; e) programming: fundaments of Matlab, C/C++, OOP.

A. Course/Lecture (course/lecture titles)	
1	Introduction. Features of the speech signal (acoustics, statistics, models for speech productions).
2	Digital modeling of speech.
3	Methods for speech analysis in time domain (energy, zero crossing, autocorelation, AMDF, TESPAR).
4	Methods for speech analysis in frequency domain (Fourier, subband).
5	Cepstral analysis. Speech analysis using wavelet transform.
6	Techniques for speech coding in time domain and applications for VoIP (PCM, ADPCM, Delta)
7	Subband speech coding.
8	Speech coding in MPEG standard: MPEG1 (Layer 1,2,3), MPEG2, MP3, AC3 algorithm.
9	Speech coding using analysis by synthesis technique (MPE, RPE-LTP, CELP).
10	Speech coding in GSM. Half rate coder and the VAD.
11	Speech compression by Vector Quantization. The algorithms: Lloyd, LBG.
12	The principles of Text to Speech Synthesis (TTS). Automatic control of intonation (PSOLA/TD-PSOLA).
13	Techniques for Automatic Speech Recognition: DTW, ANN-MLP, HMM.
14	Synthesis of the course and preparation for exam.

B. Applications – Laboratory (list of laboratories), **Seminar** (contents), **Project** (project contents)	
Laboratory (1 L)	
1	Robust algorithm for speech endpoint detection using the energy and the zero crossing rates.
2	Estimation of fundamental frequency of speech by autocorrelation and AMDF methods.
3	Spectral analysis of speech by FFT . Fundamental frequency estimation by cepstral analysis.

4	Linear Predictive Analysis. Speech synthesis from LPC coefficients.
5	Speech coding by using ADPCM and IMA-ADPCM (performance evaluation of the methods).
6	Implementation of a GSM RPE-LTP speech coder and evaluation of the performances.
7	Implementation and evaluation of the methods to change the F0 (PSOLA si TD-PSOLA).
Project (1 P)	
1	Project planning (Presentation of the project themes, the students select the project, project schedule)
2	Analysis and documentation on scientific and technical background
3	Implementation. Designing the software and speech processing models.
4	Implementation. Report I.
5	Implementation. Report II.
6	Experiments and interpretation of results
7	Project submission.

C. Individual study (reference study contents, synthesis materials, projects, applications etc.)

Reports with the results obtained in the laboratory, scientific documentation for the project, project implementation, team-working to produce the experimental results.

Individual study structure	Course study	Problem solving, laboratory, project	Applications preparation	Examination time	Additional reference study	Total no. of individual study hours
Hours	28	42	7	3	14	94

References (Textbooks, courses, laboratory manual, exercise book)

UTC-N Library:
1. R. Martin, sa, *Advances in digital speech transmission,* Chichester, West Sussex, 2008.
2. M. Giurgiu, *Compresia semnalului vocal in aplicatii multimedia,* Ed. Risoprint Cluj-Napoca, 2003.
3. M. Giurgiu, *Sinteza din text a semnalului vocal. Vol I.,* Editura Risoprint, Cluj-Napoca, 2006.
4. S. Furui, *Digital Speech Processing, Synthesis and Recognition*, New York, 2001.
5. D.G, Childers – *Speech Processing and Synthesis Toolboxes,* Wiley Publ., 2000
6. F. Jelinkek – *Statistical methods for speech recognition*, London, 1999.
7. H. Dybkjaer - *Designing interactive speech systems : from first ideas to user testing,* Berlin, 1998.
8. R Boite, M Kunt – *Traitement de la parole,* Presse Politechnique Romandes, Lausanne, 1987.
9. G. Stolojanu, *Prelucarea numerica a semnalului vocal,* Ed Militara, 1984
10. M. Draganescu, *Analiza si sinteza semnalului vocal*, Ed. Academiei, 1986
11. E. Pupu, P. Pop – *Prelucrarea Numerica a Semnalului Vocal*, Ed. Risoprint, Cluj, 2004

Online materials:
12. M. Giurgiu - *Compresia semnalului vocal in aplicatii multimedia*, www.concorde.utcluj.ro
13. O serie de resurse din Internet, recomandate de catre grupul de specialisti din European Language and Speech Network

Other libraries (or available in the lab):
14. K. Ponting, *Computational Models for Speech Pattern Processing,* Springer Verlag, 1997
15. S. Furui, *Digital Speech processing, Synthesis and Recognition,* Marcel Dekker, 1989
16. E. Keller, *Fundamentals of Speech Synthesis and Speech Recognition*, John Willey & Sons, 1994
17. A. M. Kondoz, *Digital Speech Coding for Low Bit Rate Communications Systems*, John Willey, 1995 L
18. L.R. Rabiner, *Digital Processing of Speech Signals,* Prentice Hall, 1978
19. P. Papamichalis, *Practical Approaches to Speech Coding*, Prentice Hall, 1995

Final evaluation	
Evaluation method	Final examination (FE). Laboratory activity and experimental reports (LB). Project (P)
Mark components	Final Exam (0…6), Project (0…2), Laboratory activity and reports (0..2).
Mark computation	M = LB + P + FE.

Course leader,

Professor Mircea GIURGIU, Ph.D.