SYLLABUS

1. Data about the program of study

1.1	Institution	The Technical University of Cluj-Napoca
1.2	Faculty	Faculty of Electronics, Telecommunications and Information Technology
1.3	Department	Communications
1.4	Field of study	Electronics and Telecommunications Engineering
1.5	Cycle of study	Bachelor of Science
1.6	Program of study/Qualification	Telecommunications Technologies and Systems
1.7	Form of education	Full time
1.8	Subject code	TST-E53.20

2. Data about the subject

2.1	Subject name			Digital Speech Processing			
2.2	Subject area			Telecommunications			
2.3	Course responsible/lecturer			Prof. Mircea Giurgiu, Ph.D.			
2.4	Teachers in charge of seminars			Adriana Stan, PhD			
2.5 Year of study	4	2.6 Semester	8	2.7 Assessment	VP	2.8 Subject category	DS/DOP

3. Estimated total time

3.1 Number of hours per week	4	3.2 of which, course:	2	3.3 applications:	2
3.4 Total hours in the curriculum	78	3.5 of which, course:	28	3.6 applications:	28

Individual study	hours
Manual, lecture material and notes, bibliography	8
Supplementary study in the library, online and in the field	2
Preparation for seminars/laboratory works, homework, reports, portfolios, essays	8
Tutoring	2
Exams and tests	2
Other activities	0

3.7	Total hours of individual study	22
3.8	Total hours per semester	78
3.9	Number of credit points	3

4. Pre-requisites (where appropriate)

4.1	Curriculum	Digital Signal Processing, Information Theory
4.2	Competence	Basic computer programming skills

5. Requirements (where appropriate)

5.1	For the course	Lecture room with video-projector
5.2	For the applications	LAN in the lab room with Internet connection, Matlab environment, Python tools for speech processing, speech databases, Audacity tools, HTK toolkit, PRAAT tools, VoiceBox Toolkit.

6. Specific competences

Professional competences	C4. To design, implement and operate data, voice, video and multimedia services, based on the understanding and application of fundamental concepts from the field of communications and information transmission. C5. To select, install, configure and exploit fixed and mobile telecommunications equipment. To equip a site with common telecommunications networks.
Cross competences	N.A.

7. Discipline objectives (as results from the *key competences gained*)

7.1	General objective	To develop knowledge and skills on the methods and the algorithms used for speech processing in time, frequency or cepstral domain by handling specific software tools.
7.2	Specific objectives	• to know the main features of the speech signal and the production models • to handle specific software tools for speech processing • to know the concepts and the methods applied for temporal, spectral and statistical representation if the speech signal • to understand the concepts and the processing flows for speech coding in time / spectral domain or hybrid coding • to be able to design and to implement specific algorithms in a software application used for speech processing (extract the parameters, interpret the results) • to be skilled in the use and the application of automatic speech classification methods employing dedicated software libraries.

8. Contents

8.1. Lecture (syllabus)		Teaching methods	Notes
1.	Introduction. Features of the speech signal (acoustics, statistics, models for speech productions).	PPT presentations, practical demos, interactive discussions and debates, problem solving.	NA
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.		

Bibliography:

1. Gopi E, *Digital Speech Processing Using Matlab,* Springer, 2014
2. Ramakrishnan S, *Modern Speech Recognition Approaches with Case Studies,* InTech, 2012.
3. R. Martin, et al, *Advances in digital speech transmission,* Chichester, West Sussex, 2008.
4. M. Giurgiu, *Compresia semnalului vocal in aplicatii multimedia*, Ed. Risoprint Cluj-Napoca, 2003.
5. M. Giurgiu, *Sinteza din text a semnalului vocal. Vol I.,* Editura Risoprint, Cluj-Napoca, 2006.
6. S. Furui, *Digital Speech Processing, Synthesis and Recognition*, New York, 2001.
7. D.G, Childers – *Speech Processing and Synthesis Toolboxes,* Wiley Publ., 2000
8. F. Jelinkek – *Statistical methods for speech recognition,* London, 1999.
9. H. Dybkjaer - *Designing interactive speech systems : from first ideas to user testing,* Berlin, 1998.
10. R Boite, M Kunt – *Traitement de la parole,* Presse Politechnique Romandes, Lausanne, 1987.
11. G. Stolojanu, *Prelucarea numerica a semnalului vocal,* Ed Militara, 1984
12. M. Draganescu, *Analiza si sinteza semnalului vocal*, Ed. Academiei, 1986
13. E. Lupu, P. Pop – *Prelucrarea Numerica a Semnalului Vocal,* Ed. Risoprint, Cluj, 2004

8.2. Applications: Laboratory (L) & Project (P)		Teaching methods	Notes
1.	L1: Robust algorithm for speech endpoint detection using the energy and the zero crossing rates.	Individual hands on activities, experiments, following demos, problem-based and project-based learning.	NA
2.	L2: Estimation of fundamental frequency of speech by autocorrelation and AMDF methods.		
3.	L3: Spectral analysis of speech by FFT. Fundamental frequency estimation by cepstral analysis.		
4.	L4: Linear Predictive Analysis. Speech synthesis from LPC coefficients.		
5.	L5: Speech coding by using ADPCM and IMA-ADPCM (performance evaluation of the methods).		
6.	L6: Implementation of a GSM RPE-LTP speech coder and evaluation of the performances.		
7.	L7: Implementation and evaluation of the methods to change the F0 (PSOLA si TD-PSOLA).		
8.	P1: Project planning (Presentation of the project themes, the students select the project, project schedule)		
9.	P2: Analysis and documentation on scientific and technical background		
10.	P3: Implementation. Designing the software and speech processing models.		
11.	P4: Implementation. Report I.		
12.	P5: Implementation. Report II.		
13.	P6: Experiments and interpretation of results		
14.	P7: Project submission.		

Bibliography:

1.M. Giurgiu, *Sinteza din text a semnalului vocal. Vol I.,* Editura Risoprint, Cluj-Napoca, 2006.
2.Gopi E, *Digital Speech Processing Using Matlab,* Springer, 2014
3.S. Furui, *Digital Speech Processing, Synthesis and Recognition*, New York, 2001.
4. ***, HTK Handbook, Cambridge University, 2008.

9. Bridging course contents with the expectations of the representatives of the community, professional associations and employers in the field

Competences acquired will be used in the following COR occupations (Electronics Engineer; Telecommunications Engineer; Electronics Design Engineer; System and Computer Design Engineer; Communications Design Engineer) or in the new occupations proposed to be included in COR (Sale Support Engineer; Multimedia Applications Developer; Network Engineer; Communications Systems Test Engineer; Project Manager; Traffic Engineer; Communications Systems Consultant).

10. Evaluation

Activity type	10.1 Assessment criteria	10.2 Assessment methods	10.3 Weight in the final grade
Course	Student performance and deep of knowledge against the defined learning outcomes	Written examination (knowledge and problem solving skills)	50%
Applications	Performance in accuracy and originality of problem solving, experiment running and presentation of results.	Running the experiment, solving the problems, intermediary evaluation, individual work, laboratory reports	50%
10.4 Minimum standard of performance			
To implement algorithms for speech processing in time, frequency or cepstral domain by handling specific software tools with a practical view on speech analysis, coding at low bit rates or automatic recognition.			

Date of filling in
01.10.2014

Course responsible
Professor
Mircea GIURGIU, PhD

Teachers in charge of applications
Adriana STAN, PhD

Date of approval
in the department
01.10.2014

Head of Communications
Department
Professor Virgil DOBROTA, PhD