SYLLABUS

1. Data about the program of study

1.1	Institution	The Technical University of Cluj-Napoca
1.2	Faculty	Electronics, Telecommunications and Information Technology
1.3	Department	Telecommunications Technologies and Systems
1.4	Field of study	Electronics and Telecommunications Engineering
1.5	Cycle of study	Master
1.6	Program of study/Qualification	Multimedia Technologies
1.7	Form of education	Full time
1.8	Subject code	TC-E16.30

2. Data about the subject

2.1	Subject name					Speech Analysis, synthesis and recognition					
2.2	Subject area					Electronics and Telecommunications Engineering					
2.3	Course responsible/lecturer					Professor Eugen LUPU, Ph.D. Eugen.Lupu@com.utcluj.ro					
2.4	Teachers in charge of applications					Professor Eugen LUPU, Ph.D. Eugen.Lupu@com.utcluj.ro					
2.5	Year of study	II	2.6	Semester	1	2.7	Assessment	Exam	2.8	Subject category	DS/DO

3. Estimated total time

Year/ Sem.	Subject name	No. of weeks	Course [hours/ week]	Applications [hours/ week]			Course [hours/ semester]	Applications [hours/ semester]			Indiv. study	TOTAL	Credits
				S	L	P		S	L	P			
II/1	Speech Analysis, synthesis and recognition	14	2		1		28		14		58	100	4

3.1	Number of hours per week	3	3.2	of which, course	2	3.3	applications	1
3.4	Total hours in the curriculum	42	3.5	of which, course	28	3.6	applications	14

Individual study	Hours
Manual, lecture material and notes, bibliography	20
Supplementary study in the library, online and in the field	10
Preparation for seminars/laboratory works, homework, reports, portfolios, essays	22
Tutoring	3
Exams and tests	3
Other activities	

3.7	Total hours of individual study	58
3.8	Total hours per semester	100
3.9	Number of credit points	4

4. Pre-requisites (where appropriate)

4.1	Curriculum	NA
4.2	Competence	Signal processing, MATLAB

5. Requirements (where appropriate)

5.1	For the course	Cluj-Napoca
5.2	For the applications	Cluj-Napoca

6. Specific competences

Professional competences	
Cross competences	

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

7.1	General objectives	Developing the competences regarding the speech signal and the typical application connected to speech..
7.2	Specific objectives	1. Items on the speech signal production, acquisition and modeling. 2. Speech parameters/features of interest depending on the application 3. Speech analysis, synthesis and recognition 4. Notions of biometry, speaker recognition

8. Contents

8.1. Lecture (syllabus)		Teaching methods	Notes
1	Introduction on the speech. Phonation mechanism. Acoustic parameter s speech signal (SS). SS objective and subjective features. SS production models (Fant). SS classification among signals.	Presentation, heuristic conversation, exemplification, problem presentation, teaching exercise, case study, formative evaluation	Use of .ppt presentation, projector, blackboard
2	SS processing. Analysis, synthesis, compression, recognition. SS Download and sampling SV. Acoustic transducers. Features. Sampling and reconstruction SV. Windowing. Preemphasis of SS.		
3	SS analysis in the time domain. Average amplitude / maximum. Amplitudine.ZCR density function. Energy. Teager energy. The		

	fundamental frequency (FF). Methods of obtaining the FF.		
4	SS analysis in the frequency domain. Short-time Fourier transform- STFT. FFT algorithms (TFR). Spectrogram. Analysis of SS by digital filters bank.		
5	Cepstral Analysis. FFT cepstrum. Getting the smoothed spectrum and fundamental frequency-FF from cepstral coefficients. LPC analysis of SS. Determination of LPC model parameters. (voiced/ unvoiced decision, FF, Gain, filter coefficients). LPC processing steps for speech recognition.		
6	Perceptual Analysis: PLP Analysis. MEL Cepstral Analysis.MFCC coefficients		
7	Speech synthesis. Introduction. Speech Synthesis systems and speech synthesizer. Methods commonly used in the synthesis schemes. Elements of prosody. Summary generation systems using direct-channel synthesizer. Speech synthesis systems based on models. Formants based synthesis. LPC Synthesis. Speech synthesis systems by simulating the vocal tract. Speech synthesis based on rules.		
8	Text-to-speech synthesis overview. Text and linguistic analysis. Speech synthesis.		
9	Basics on automatic speech/ speaker recognition Recognition Systems. Metric in acoustic space. Distances (Euclidean, LPC, cepstral). Methods used in the recognition of SS. Dynamic time warping method (DTW).		
10	Vector quantization. Algorithms for determining the dictionary (threshold algoritm, k-means, Linde-Buzo-Gray). Stochastic methods for recognition. Hidden Markov Models (HMM). Elements of HMM. Problems solved for HMM: Assessing total probability of observation. The problem of discovering the optimal sequence of hidden states. Training problem.		
11	TESPAR method. Presentation and applications. Defining TESPAR alphabet of symbols. TESPAR matrices. TESPAR DZ.		
12	Introduction in Biometrics. Biometric checkers. Voice - authentication key. Multimodal biometrics.		
13	Speaker recognition (SR). The characteristics of the speaker. Parameters used in speaker recognition. Taxonomy SR systems. SR text-dependent systems. text independent SR Systems.		
14	Review-exam topics.		
8.2. Applications (lab)		Teaching methods	Notes
1	Acquisition, recording and viewing voice signal using Matlab or other tools. FFT spectrogram.	Didactic and experimental proof, didactic exercise, team work	Use of laboratory instrumentation, experimental boards, computers, magnetic board
2	Determining FF using AMDF or CLIP method.		
3	LPC analysis. Formants determination using the LPC smoothed spectrum.		
4	Summary comparative multilingual TTS system. Evaluation LHS TTS system. Text-to-speech synthesis (English-Italian-Spanish-German).		

	Text-to-Speech Synthesis in Romanian. The ROMVOX. Evaluation.		
5	Speech recognition. DTW method.		
6	CSLU dialog system "Speech toolkit" . Applications Development using RAD system of "Speech toolkit".		
7	Speaker recognition system based on the method TESPAR.		

References:

1. Benesty J,Sondhi M, Huang Y. Handbook of Speech Processing, Springer 2009
2. Vinay K. Ingle, J.G. Proakis Digital signal processing using MATLAB, Brooks/Cole 2000
3. Rabiner, L.R., Juang, B.H. “Fundamentals of speech recognition “ Prentice-Hall, 1993
4. Boite, R.Kunt, M."Traitement de la parole" Presses Polytechnique Romandes ,1991
5. Furui, S. "Digital speech processing, synthesis and recognition", Ed. Marcel Dekker 2001
6. Lupu E., Pop G.P. “Prelucrarea numerica a semnalului vocal, Ed. RISOPRINT, 2004
7. Giurgiu M., Peev Luciana Sinteza din text a semnalului vocal Ed. RISOPRINT 2006

SLIDES : http://users.utcluj.ro/~elupu/Curs/

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. Evaluations

Activity type	10.1	Assessment criteria	10.2	Assessment methods	10.3	Weight in the final grade
Course		Written test with 15-20 questions and problems		Written test (E=50%) +		E = 50%
Applications		Laboratory work and Project developed during the semester		Laboratory reports + Project defended at the end of semester		L=20% ; P = 30%
10.4 Minimum standard of performance						
L, P ≥ 5 and Ex ≥ 4.5 and N= 0.5E+0.2L+0.3P ≥ 5						

Course responsible
Professor Eugen LUPU, Ph.D.

Teachers in charge of applications
Professor Eugen LUPU Ph.D.

Date of approval in the department
01.10.2018

Head of Communications Department
Professor Virgil DOBROTA,PhD