SYLLABUS

1. Data about the program of study

1.1 Institution	Technical University of Cluj-Napoca
1.2 Faculty	Faculty of Electronics, Telecommunications and information Technology
1.3 Department	Communications
1.4 Field of study	Electronic Engineering, Telecommunications and Information Technologies
1.5 Cycle of study	Bachelor of Science
1.6 Program of study / Qualification	Telecommunications Technologies and Systems/ Engineer
1.7 Form of education	Full time
1.8 Subject code	TST-E54.20

2. Data about the subject

2.1 Subject name	Speech Processing						
2.2 Subject area	Theoretical area Methodological area Analytic area						
2.3 Course responsible	Prof. Mircea Giurgiu, Ph.D – Mircea.Giurgiu@com.utcluj.ro						
2.4 Teacher in charge with the laboratory / project	Alexandra Drobut, Ph.D student – Alexandra.Drobut@com.utcluj.ro						
2.5 Year of study	4	2.6 Semester	8	2.7 Assessment	VP	2.8 Subject category	DS/DO

3. Estimated total time

3.1 Number of hours per week	5	of which: 3.2 course	2	3.3 laboratory & project	3
3.4 To Total hours in the curriculum	70	of which: 3.5 course	28	3.6 laboratory & project	42

Distribution of time	hours
Manual, lecture material and notes, bibliography	25
Supplementary study in the library, online specialized platforms and in the field	13
Preparation for seminars / laboratories, homework, reports, portfolios and essays	10
Tutoring	2
Exams and tests	3
Other activities: project demonstration	2

3.7 Total hours of individual study	55
3.8 Total hours per semester	125
3.9 Number of credit points	5

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 laboratories / projects	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, Deep Neural Networks toolkits, 2 systems with GPU cards

6. Specific competences

Professional competences	**C4. Design, implementation and operation of data, voice, video and multimedia services. This is based on the understanding and the application of fundamental concepts in telecommunications and transmission of information** C4.2 Solving practical problems using general knowledge of multimedia techniques C4.3 Explanation and interpretation of the main requirements and specific approach techniques for data, voice, video, multimedia transmissions C4.3 Solving practical problems using general knowledge of multimedia techniques C4.4 Use of the main specific parameters in evaluations based on the concept of quality of service in communications C4.5 Development of simple communications services **C5. Selecting, installing, configuring and operating fixed or mobile telecommunications equipment. Equipping a site with usual telecommunications networks**
Transversal 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, and 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. 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. S. Sen, A. Dutta, N. Dey, Audio processing and speech recognition – concepts, techniques and research overviews, Springer, 2019

8.2 Laboratory	Teaching methods	Notes
1. Introduction on using Python and Jupiter Notebooks.	Individual hands on activities, experiments, following demos, problem-based and project-based learning.	NA
2. Short time analysis of the speech signal.		
3. Analysis and evaluation of the speech production models.		
4. Speech endpoint detection using the energy and the zero crossing rates. (Test #1)		
5. Estimation of the fundamental frequency using time domain analysis.		
6. Spectral analysis using FFT		
7. Cepstral analysis and MFCC (Test #2)		
8.3. Project		
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. Setting up the development environment (software tools)		
4. Speech data collection and preliminary analysis		
5. Designing of the application		
6. Implementation of the project – first trials		
7. Intermediary evaluation of project progress (report)		
8. Project implementation	Project design and software implementation (Python / Matlab / Java / C++ /	
9. Project implementation		
10. Preliminary testing and evaluation		
11. Project improvements		
12. Experiments and interpretation of results (I)		
13. Writing the final report		
14. Project submission.		

Bibliography

1. M. Giurgiu, Sinteza din text a semnalului vocal. Vol I., Editura Risoprint, Cluj-Napoca, 2006.
2. E. Gopi, 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.
5. S. Sen, A. Dutta, N. Dey, Audio processing and speech recognition – concepts, techniques and research overviews, Springer, 2019

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

The subject is oriented towards the development of practical applications involving speech processing for telecommunications purposes (speech coding at low bit rates, speech enhancement) or for development of human-machine interfaces such as IVRs. The contents are aligned with the requirements of the IT industry and meet the expectations of important local software development companies to implement speech technology interfaces on mobile devices or as web-based services, other small and medium size enterprises.

10. Evaluation

Activity type	10.1 Assessment criteria	10.2 Assessment methods	10.3 Weight in the final grade
10.4 Course	The level of acquired theoretical knowledge and practical skills	Written examination (knowledge and problem solving skills)	50%
10.5 Laboratory / Project	The level of acquired knowledge and abilities	Running the experiment, solving the problems, 2 intermediary laboratory tests, individual work, laboratory reports, 2 project evaluations. Labs (2 pts), Project (3pts)	50%
10.6 Minimum standard of performance			

Quality level:

Minimum knowledge:

- ✓ to know different models for speech production and the main features of the speech signal
- ✓ to know the methods for speech analysis in time, frequency and cepstral domains
- ✓ to know the speech coding at low bit rates in time, frequency and parametric domains
- ✓ to know the principles for automatic speech recognition and text to speech synthesis

Minimum competences:

- ✓ to be able to handle different speech processing algorithms in Python / Matlab
- ✓ to implement Python algorithms for speech processing in time, frequency in cepstral domains
- ✓ to use available speech processing toolkits to extract the speech feature
- ✓ to implement experimental projects by using basic machine learning algorithms for speech processing (speaker recognition, emotion recognition, speech recognition using DeepSpeech toolkit, text to speech synthesis using the Tacotron 2 / DC TTS toolkits).

Quantitative level:

- ✓ to properly execute the laboratory activities and to implement a successful the project
- ✓ to pass the laboratory tests
- ✓ overall mark is calculated as: 0,2 * Laboratory + 0,3 * Project + 0,5 * FinalExam

Date of filling in:	**Responsible**	**Title Surname NAME**	**Signature**
27.09.2021	Course	Prof. Mircea Giurgiu, Ph.D	
	Applications	Alexandra Drobut, Ph.D student	

Date of approval in the Department of Communications 27.09.2021	Head of Communications Department Prof. Virgil DOBROTA, Ph.D.
Date of approval in the Council of Faculty of Electronics, Telecommunications and Information Technology 27.09.2021	Dean Prof. Gabriel OLTEAN, Ph.D.