# 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<br>Methodological area<br>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 | Prof. Mircea GIURGIU, Ph.D. – Mircea.Giurgiu@com.utcluj.ro<br>Eng. 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 | 4 | 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**<br>C4.2 Solving practical problems using general knowledge of multimedia techniques<br>C4.3 Explanation and interpretation of the main requirements and specific approach techniques for data, voice, video, multimedia transmissions<br>C4.3 Solving practical problems using general knowledge of multimedia techniques<br>C4.4 Use of the main specific parameters in evaluations based on the concept of quality of service in communications<br>C4.5 Development of simple communications services<br>**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<br>• to handle specific software tools for speech processing<br>• to know the concepts and the methods applied for temporal, spectral and statistical representation if the speech signal<br>• to understand the concepts and the processing flows for speech coding in time / spectral domain or hybrid coding<br>• to be able to design and to implement specific algorithms in a software application used for speech processing (extract the parameters, interpret the results)<br>• 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. 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. 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). | | |
| **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. 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. | | |

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.
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, intermediary evaluation, individual work, laboratory reports, project evaluation | 50% |
| 10.6 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: | Responsible | Title Surname NAME | Signature |
|---|---|---|---|
| 20.06.2023 | Course | Professor Mircea GIURGIU Ph.D. | |
| | Applications | Professor Mircea GIURGIU Ph.D. | |
| | | Eng. Alexandra DROBUT, Ph.D. student | |

| | |
|---|---|
| Date of approval in the Council of the Communications Department<br>11.07.2023 | Head of Communications Department<br>Prof. Virgil DOBROTA, Ph.D. |
| Date of approval in the Council of the Faculty of Electronics, Telecommunications and Information Technology<br>12.07.2023 | Dean<br>Prof. Ovidiu POP, Ph.D. |