POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

COURSE DESCRIPTION CARD - SYLLABUS

Course name
Digital technology of audio and speech [S2EiT2E-TIT>CTDiM]

Course

Field of study
Electronics and Telecommunications

Year/Semester
2/3

Area of study (specialization)
Information and Communication Technologies

Profile of study
general academic

Level of study
second-cycle

Course offered in
English

Form of study
full-time

Requirements
elective

Number of hours

Lecture
30

Laboratory classes
30

Other
0

Tutorials
0

Projects/seminars
0

Number of credit points

5,00

Coordinators

dr hab. inż. Olgierd Stankiewicz prof. PP
olgierd.stankiewicz@put.poznan.pl

Lecturers

Prerequisites

Has a structured and mathematically grounded knowledge of one-dimensional signal theory necessary to understand the representation and analysis of signals in the time and frequency domain. Has a structured, mathematically based detailed knowledge of basic methods of digital signal processing. Can obtain information from literature and databases and other sources in Polish or English; can integrate information obtained, interpret it, reach conclusions and justify opinions. Can solve basic problems in electronics and telecommunications with the use of mathematical analysis, algebra and probability calculus. Can solve tasks related to the analysis of signals in time and frequency domain. Can perform basic calculation algorithms using programming languages (e.g. Matlab, C). Knows the limitations of his/her knowledge and skills and understands the necessity of further training. Is aware of the necessity of a professional approach to solving technical problems and taking responsibility for the technical solutions he/she proposes. Is able to participate in team projects.

Course objective

Extending the knowledge on human auditory perception, understanding the underlying physics, and the limitations thereof. A deeper insight into most important digital processing techniques for speech and audio, and their applications in current telecommunications. Acquiring general knowledge on data compression techniques for speech and audio. Introduction to audio signal enhancement techniques. Introduction to speech recognition and synthesis, as well as musical sound synthesis. Gaining the ability to select an appropriate data compression technique for speech and audio in wireless telephony, packet data transmission systems, DVB, internet TV, home cinema, and professional systems. Understanding the limitations of speech recognition and synthesis techniques as well as their possible applications in telecommunication systems and electronic services.

Course-related learning outcomes

Knowledge:
Deep knowledge on construction and principles of audio parts in telecommunication systems.
Structured and theoretically underpinned knowledge on compression techniques for speech and wideband audio, including applications of perceptual compression in communications via data transport networks.

Skills:
Ability to analyze the operation of a multimedia system for its audio transmission path, as well as its limitations, and to gain the maximum use of the data transmission conditions offered.
Ability to implement the audio path in VoIP and programmable radio, as well as to take into account the peculiarities of audio representation and perception, and how they impact the requirements of such systems. Ability to select appropriate tools for enhancement of an audio signal distorted or interfered by noise.

Social competences:
Ability of self-learning (textbooks, computer programs).
Understanding the non-technical factors of engineering.
Knowing the responsibility for the electronic and telecommunication systems being designed.

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Learning outcomes presented above are verified as follows:
Lecture: written exam
The written exam is a set of 6-10 problematic questions, for which desciptory answer is expected.
Each answer is ranked from 0 to 1 points (fractinal points also possible).
The exam is passed if the number of attained points is greater than 50%. More than 50% indicates the knowledge above of satisfactory level.
The course issues of which the questions are prepared, are sent to students by e-mail using the university"s e-mail system.
Laboratories:
Activity during classes, reports from particular activities. Laboratory project realized individually/in small groups.

Programme content

Human auditory perception (the structure of the organ of hearing, the physical characteristics of the wave vs perceptual attributes of a sound, critical bands, limitations of perception, masking (simultaneous and temporal).
Frequency domain and time-requency representations of audio signals (STFT, the spectrogram, the uncertainty principle, filterbanks, cosine transform, MDCT, linear and quadratic time-frequency distributions, the uncertainty principle in time and frequency, Gabor transform, wavelet transform, subband analysis, filterbanks, cosine transform, MDCT, linear and quadratic time-frequency distributions, Wigner-Ville distributions, cepstrum analysis)
Spectral modeling of audio signals (linear AR, MA, and ARMA models, linear prediction, cepstrum, sinusoidal modeling and sinusoids+noise model)
Audio signal enhancement for distorted and noised signals (adaptive filtering, prediction-based

smoothing, LMS and RLS filters, narrowband and wideband noise supression, spectral subtraction, nonlinear filters, masking and mitigation of gaps, reconstruction of missing frames in MPEG data)
Speech coding (lossy compression techniques: ADPSM, LPC, CELP, ACELP, RPE, AMR, vector quantization, sinusoidal coding, ITU-T standards and recommendations)
Perceptual coding of wideband audio (the peculiarities of subband and transform coding, fundamentals of perceptual coding, compression schemes of MPEG-1 and 2 layer 1, 2, and 3, MPEG-2 AAC, SBR bandwidth extension, parametric stereo coding, MPEG-4 AAC-HE technique, MPEG-USAC, parametric (model based) audio coding).
Sound synthesis (additive and subtractive methods, sampling, granular synthesis, frequency modulation, waveshaping, physical modelling, applications in contemporary multimedia systems)
Speech recognition and synthesis (speech components and properties important for recognition and synthesis, signal analysis and feature selection based on LP, MFCC, spectral moments, formant frequencies; classification methods: feature sets, pattern recognition; synthesis based on vocoder methods, spectral models, and physical models, corpus-based synthesis; text phonetization problem, implementation of text-to-speech systems).

Course topics

none

Teaching methods

Lecture: multimedia presentation with examples presented on the blackboard.
Laboratories: implementation of projects on computers (individual or in groups of few people).
Examples illustrated on screen/blackboard.

Bibliography

Basic
J. Watkinson, The Art of Digital Audio, Focal Press, 2001
Audio Signal Processing and Coding, A. Spanias, T. Painter, V. Atti, Wiley, 2007
DAFX, Digital Audio Effects, Udo Zoelzer (red.), Wiley, 2002
Additional

Breakdown of average student's workload

	Hours	ECTS
Total workload	125	5,00
Classes requiring direct contact with the teacher	70	3,00
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	55	2,00