

Télécommunications

Speech and Image processing

IDENTIFICATION

CODE : TC-4-S1-EC-TIP
ECTS : 2.0

HOURS

Lectures : 0.0 h
Seminars : 24.0 h
Laboratory : 8.0 h
Project : 0.0 h
Teacher-student
contact : 32.0 h
Personal work : 0.0 h
Total : 32.0 h

ASSESSMENT METHOD

TEACHING AIDS

TEACHING LANGUAGE

French

CONTACT

M. BASKURT Atilla
atilla.baskurt@insa-lyon.fr

AIMS

The aim of this course is the presentation of basic techniques used in speech processing, image processing and video processing. The compression standards are also studied in detail. This course is illustrated with several real studies, which are related to image analysis, image retrieval, face recognition, video content analysis and human action recognition in videos.

This module begins with a first part that formalizes the signal theory and presents the necessary mathematical developments for a better understanding of signal tools such as Fourier Transform or Wavelet Transform. The Hilbert space and the orthogonal projection theorem are the main concepts of this framework.

This EC is part of the teaching unit Signal and Image [TC-4-S1-S & I] and contributes to the following skills:

C1 Specify, design and model transmission and signal / image / data processing systems

Capacity: Model a continuous or discrete signal with an approximate signal

Capacity: Model the speech system and the auditory system to analyze and synthesize the voice and audio

Capacity: Given a problematic (segmentation or denoising or identification or other), design a chain of image processing to solve this problem

Knowledge: Main Characteristics of Representation Spaces of Signals and Images

Knowledge: Optimal approximation of signals and images

Knowledge: Understanding wavelet transform and multi-resolution analysis

Knowledge: Basic Processing Tools for Images

Knowledge: Hilbert Space, Fourier Transform, Cosine Transform, Wavelet Transform,

A1 Analyze a real or virtual system (or problem)

Capacity: Analyze (understand the processing chain, identify different bricks) compression standards of the speech, audio, images and videos

Capacity: Design a compression chain of the speech, audio, images and videos

Knowledge: Optimal approximation of signals and images

Knowledge: Quantization and coding techniques for voice, audio, 2D image and video

Knowledge: Speech, audio, picture and video compression standards

C5 Implement, realize, develop, deploy systems for transmission and processing of signals / images / data

Capacity: Implement multiresolution analysis techniques for signals and images

Capacity: Realize an image processing software chain using tools such as MATLAB or STARLING

Knowledge: Methodologies for analysis and synthesis of speech and audio

Knowledge: Quantization and coding techniques for speech, audio, image and video

Knowledge: Methodologies for Big Data Analysis ("Machine Learning" and "Deep Learning")

A5 Process data

Capacity: Given an image or video database, implement data analysis techniques ("Machine Learning" learning and "Deep Learning" deep learning) to classify or identify visual objects [Artificial Intelligence]

In addition, it requires the following skills:

A4 Design a system that meets a set of specifications

B2 Work, learn, evolve autonomously

B3 Interact with others, work in a team

A2 Exploit a model of a real or virtual system

A3 Implement an experimental approach

CONTENT

INSA LYON

Campus LyonTech La Doua

20, avenue Albert Einstein - 69621 Villeurbanne cedex - France

Phone +33 [0]4 72 43 83 83 - Fax +33 [0]4 72 43 85 00

www.insa-lyon.fr

Part I

- Hilbert space, Fourier transform in Hilbert space
- Orthogonality, projection theorem and orthonormal basis of Hilbert space - Matrix, function approximation
- Wavelet transform, multiresolution analysis

Part II : Speech, Audio, Image and Video Processing Remind

- Digital filtering and ARMA modeling
- Scalar and Vector Quantization
- Entropy Coding (Huffman, Arithmetic, LZW)
- Speech: analysis, synthesis, coding
- Audio: compression standards
- Introduction to image processing: preprocessing, linear and non-linear filtering
- Image representation: Fourier, Cosinus et wavelets
- Image compression and JPEG standard
- Video compression and MPEG1, MPEG2, H.264/AVC standards
- Introduction to image segmentation
- Introduction to image watermarking
- Video sequence analysis: 2D interest point (SIFT), motion object detection and tracking

Part III: Data Analysis and Multimedia Applications

- Introduction to Data analysis tools: Principal Component Analysis (PCA)
- Multimedia indexing and retrieval: non-supervised classification and image retrieval from databases
- Face recognition in images: statistical and neuronal supervised classification
- Video: content analysis, moving object detection and tracking methods, automatic recognition of objects in motion and human actions, video compression standards

BIBLIOGRAPHY

- [1] Pierre Brezmaud, Mathematical principles of signal processing, Fourier and wavelet analysis, Springer, ISBN 0-387-95338-8, 269 p., 2002.
- [2] Todd K. Moon, Wynn C. Stirling, Mathematical methods and algorithms, Prentice-Hall, ISBN 0-201-36186-8, 937 p., 2000.
- [3] Steéphane Mallat, A wavelet tour of signal processing, Academic Press, New York, 1998.
- [4] Digital Image Processing: http://users.dcc.uchile.cl/~jsaavedr/libros/dip_gw.pdf

PRE-REQUISITE

Signal processing

INSA LYON

Campus LyonTech La Doua

20, avenue Albert Einstein - 69621 Villeurbanne cedex - France
Phone +33 [0]4 72 43 83 83 - Fax +33 [0]4 72 43 85 00

www.insa-lyon.fr