**Nazwa przedmiotu:**

Image and Speech Recognition

**Koordynator przedmiotu:**

Prof. dr hab. inż. Włodzimierz Kasprzak

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Computer Science and Information Systems

**Grupa przedmiotów:**

Obligatory

**Kod przedmiotu:**

1120-INSZI-MSA-0116

**Semestr nominalny:**

2 / rok ak. 2022/2023

**Liczba punktów ECTS:**

5

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

.

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

.

**Język prowadzenia zajęć:**

angielski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

.

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład: | 30h |
| Ćwiczenia: | 15h |
| Laboratorium: | 15h |
| Projekt: | 0h |
| Lekcje komputerowe: | 0h |

**Wymagania wstępne:**

Students are expected to have the following background:
- knowledge of basic computer science principles and skills, at a level sufficient to write a reasonably non-trivial computer program, preferably in one of the languages: C/C++, Java, C# or Pascal/Delphi,
- familiarity with basic mathematical analysis, linear algebra and probability theory.

**Limit liczby studentów:**

Bez limitu

**Cel przedmiotu:**

The goal is to learn about basic methods and algorithms in digital image- and speech-analysis. The students will be able to design image and speech analysis programs dealing with pattern (image or speech) processing, pattern segmentation and object (or word) recognition.

**Treści kształcenia:**

We start with the principles of pattern recognition theory - pattern classification, recognition, understanding. Basic feature space transformations PCA, LDA and ICA are introduced and different classifiers are presented, e.g. a linear discriminate classifier, SVM, Bayes-, k-NN- and MLP-classifiers. Basic unsupervised learning approaches for pattern clustering and vector quantization are shown too.
The second part covers basic image recognition solutions: image processing, segmentation and object recognition. The topics of image formation and low-level processing are presented, like camera auto-calibration, color spaces and compression schemas, and the processes of image binarization, normalization and filtering. Among image segmentation methods we introduce algorithms for edge following and line segment detection, Hough transforms, homogeneous region detection, texture- and shape description. Approaches to model-based sequence and object recognition incude: dynamic programming search, heuristic matching strategies, graph search and MAP estimation.
The third part deals with speech recognition. It starts with basic signal processing in the time and frequency domain (speech source detection, basic frequency estimation, noise elimination, windowed Fourier transform, FFT). Then basic feature detection approaches are presented, leading to the MFCC and LPC-based features. We illustrate the phonetic model of speech by spectrograms for different phoneme types, e.g. monophtongs, diphthongs, approximants, nasals, fricatives, and we introduce the tri-phone model. The spoken word recognition problem is solved by the use of Hidden Markow Models for word modeling with the Baum-Welch training and Viterbi search methods. Finally we introduce the N-gram-based modelling of sentences and token passing search for sentence recognition.

**Metody oceny:**

There is a continuous assessment method applied in this course. The points are collected during the semester time, a they come from two tests (2 x 30 pts.), covering the lecture and exercise material, and from a project work assessed in the project time (40 pts.).

**Egzamin:**

nie

**Literatura:**

Recommended reading:
1. W. Kasprzak, Image and speech recognition (in Polish: Rozpoznawanie obrazów i sygnałów mowy), WUT publishing house, Warszawa, 2009.
2. R. Duda, P. Hart, D. Stork, Pattern classification, 2nd edition, John Wiley & Sons, New York, 2001.
3. H. Niemann, Klassifikation von Mustern, 2nd edition, Springer, 2003.
4. R. C. Gonzales, P. C. Wintz, Digital image processing, Addison-Wesley, Reading, MA, 1987.
5. I. Pitas, Digital image processing algorithms and applications, John Wiley, New York, 2000.
6. L. Rabiner, B.-H. Juang, Fundamentals of speech recognition, Prentice Hall, New York, 1993.
7. J. Benesty, M.M. Sondhi, Y. Huang (eds), Handbook of speech processing, Springer, Berlin, 2008.
Software:
The exercise part is illustrated by programs in Matlab. The project work explores two open source program libraries (e.g. OpenCV and Marf), that contain implementations of basic tasks in image and speech analysis as well as pattern classification.

**Witryna www przedmiotu:**

brak

**Uwagi:**

.

## Charakterystyki przedmiotowe

### Profil ogólnoakademicki - wiedza

**Charakterystyka W01:**

Has extensive knowledge of mathematics regarding the use of linear programming and the linear and non-linear optimization for pattern classification and pattern recognition

Weryfikacja:

tests

**Powiązane charakterystyki kierunkowe:** I2\_W01

**Powiązane charakterystyki obszarowe:**

**Charakterystyka W02:**

Knows advanced machine learning and computational intelligence methods, especially methods of learning the pattern classifiers and of pattern recognition methods (applied for digital image- and speech recognition)

Weryfikacja:

tests

**Powiązane charakterystyki kierunkowe:** I2AI\_W02, I2AI\_W04

**Powiązane charakterystyki obszarowe:**

**Charakterystyka W03:**

Has knowledge of the advanced algorithmics, data structures and methods of creating algorithms in particular required in systems of pattern recognition

Weryfikacja:

tests

**Powiązane charakterystyki kierunkowe:** I2AI\_W04, I2\_W02

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - umiejętności

**Charakterystyka U01:**

Has the ability to collect, select and critically to interpret the technical information, and the ability to formulate opinions, ideas, problems and solutions

Weryfikacja:

tests

**Powiązane charakterystyki kierunkowe:** I2\_U02, I2\_U08

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U02:**

Can design efficient algorithms and justify their accuracy

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2\_U03

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U03:**

Can formulate a hypothesis regarding technical and scientific topics in the field of Computer Science

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2AI\_U03

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U04:**

Speaks English in sufficient detail to enable seamless communication on professional issues

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2\_U08, I2\_U10

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U05:**

Is able to work independently and in a team

Weryfikacja:

tests

**Powiązane charakterystyki kierunkowe:** I2\_U11

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U06:**

Is able to define the implementation stages and practically to carry out a complex information technology undertaking

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2AI\_U08, I2\_U13

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U07:**

Can design information systems using artificial neural networks for function approximation or vector quantization

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2AI\_U01, I2AI\_U07, I2\_U07

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U08:**

Has the ability to perform word processing in signals (to process inscriptions in images and spoken words in audio signals)

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2AI\_U07, I2\_U07

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - kompetencje społeczne

**Charakterystyka K01:**

Is able to think and act creatively

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2\_K04

**Powiązane charakterystyki obszarowe:**

**Charakterystyka K02:**

Is aware of the responsibility for the tasks performed jointly within a teamwork

Weryfikacja:

graded project tasks

**Powiązane charakterystyki kierunkowe:** I2\_K05

**Powiązane charakterystyki obszarowe:**