**Nazwa przedmiotu:**

Rozpoznawanie obrazów i sygnałów mowy

**Koordynator przedmiotu:**

prof.nzw. Włodzimierz Kasprzak

**Status przedmiotu:**

Fakultatywny dowolnego wyboru

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Automatyka i Robotyka

**Grupa przedmiotów:**

Przedmioty techniczne - zaawansowane

**Kod przedmiotu:**

ROSM

**Semestr nominalny:**

2 / rok ak. 2012/2013

**Liczba punktów ECTS:**

4

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

Bilans nakładu pracy studenta (100 godz.):
- udział w wykładach: 15 x 2 godz. = 30 godz.,
- udział w zajęciach projektowych: 15 x 1 godz. = 15 godz.,
- przygotowanie do wykładów (przejrzenie materiałów z wykładu i dodatkowej literatury (10 godz.), oraz próby rozwiązania zadań zawartych w podręczniku (5 godz.): 1 x 15 = 15 godz.
- realizacja zadania projektowego: 15 godz.,
- udział w konsultacjach: 2 godz.,
- przygotowanie do egzaminu (rozwiązanie zadań przedegzaminacyjnych) oraz obecność na egzaminie: 20 godz. + 3 godz. = 23 godz.

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

Liczba godzin zajęć wymagających bezpośredniego udziału nauczycieli akademickich wynosi: 30 + 15 + 2 + 3 = 50 godz., co odpowiada ok. 2 punktom ECTS.

**Język prowadzenia zajęć:**

polski

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

Liczba godzin pracy związanych z zajęciami o charakterze praktycznym wynosi: 15 + 15 + 5 + 15 = 50 godz., co odpowiada ok. 2 punktom ECTS.

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

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

**Wymagania wstępne:**

**Limit liczby studentów:**

30

**Cel przedmiotu:**

Celem przedmiotu jest zapoznanie z podstawami teorii klasyfikacji wzorców i zasadniczymi metodami komputerowego rozpoznawania obrazów cyfrowych i sygnałów mowy. W szczególności w zakresie rozpoznawania obrazów omawiane są zagadnienia: cyfrowej reprezentacji i segmentacji obrazów oraz rozpoznawania obiektów 2D i 3D. Rozpoznawanie mowy obejmuje zagadnienia: reprezentacji cyfrowego sygnału, analizy akustyczno-fonetycznej i rozpoznawania słów lub zdań mówionych. Studenci nabywają umiejętności praktycznego posługiwania się metodami rozpoznawania rozwiązując zadania i projektując systemy komputerowego rozpoznawania obrazu lub sygnału mowy.

**Treści kształcenia:**

Materiał podzielony jest na trzy części: rozpoznawanie wzorców, rozpoznawanie obrazów i rozpoznawanie sygnałów mowy.
W pierwszej części przedstawiono pojęcie wzorca, rodzaje wzorców, procesy klasyfikacji prostych wzorców i rozpoznawania złożonych wzorców. Następnie omówiono zagadnienia transformacji przestrzeni reprezentacji (cech) metodami analizy składowych głównych, linowej analizy dyskryminacyjnej i ananlizy składowych niezależnych.
Na koniec części pierwszej przedstawiono główne rodzaje klasyfikatorów (według funkcji potencjałów, Bayesa, k-NN, SVM, MLP), łaczenie klasyfikatorów (np. boosting) i podstawy klasteryzacji (k-średnich, X-średnich, EM).
W zakresie rozpoznawania obrazów omawiana jest najpierw problematyka reprezentacji obrazów, auto-kalibracji kamery, progowania i normowania obrazów.
Następnie przedstawiono podstawowe sposoby filtracji wstępnej obrazów i wyznaczania funkcji momentowych dla obszarów obrazu. Kolejne metody obejmują segmentację obrazu - wyznaczanie segmentów krawędziowych, linii prostych i krzywych, obszarów jednorodnych, cech tekstur i konturów.
Zakończenie części drugiej stanowią metody rozpoznawania obiektów 2D i 3D w obrazach.
W zakresie rozpoznawania mowy, prezentowane są zagadnienia: reprezentacji cyfrowego sygnału mowy w dziedzinie czasu i częstotliwości, przetwarzania i detekcji sygnału mowy w sygnale akustycznym, wyznaczania cech Mel-cepstralnych i według liniowej predykjcji, modelowania fonetycznego sygnału mowy, tworzenia statystycznego modelu HMM dla rozpoznawania słów i zdań.

**Metody oceny:**

Sprawdzanie założonych efektów kształcenia realizowane jest przez:
- ocenę wiedzy i umiejętności związanych z realizacją zadań projektowych – dokonywaną przez nauczyciela akademickiego, prowadzącego zajęcia z projektu, ocenę właściwego wyboru i zrozumienia metod oraz poprawności działania zaprojektowanych programów i sprawozdań z realizacji poszczególnych projektów,
- ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym o charakterze problemowym (na egzaminie student może korzystać z materiałów dydaktycznych).

**Egzamin:**

tak

**Literatura:**

1. W. Kasprzak: Rozpoznawanie obrazów i sygnałów mowy. Warszawa, 2009, Oficyna Wydawnicza PW.
2. W. Malina, M. Smiatacz: Metody cyfrowego przetwarzania obrazów. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2005.
3. W. Skarbek: Metody reprezentacji obrazów cyfrowych. Akademicka Oficyna Wydawnicza PLJ, W-wa 1993.
4. R. Duda, P. Hart, D. Stork: Pattern Classification. 2nd edition, John Wiley, 2001.
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 Heidelberg, 2008.
8. H. Niemann. Klassifikation von Mustern. 2nd edtition, Springer, Berlin, 2003.
9. D. Paulus, J. Hornegger: Applied Pattern Recognition. A Practical Introduction to Image and Speech Processing in C++. Vieweg, Braunschweig, 3d edition, 2001.

**Witryna www przedmiotu:**

http://studia.elka.pw.edu.pl/pub/ROSM.A/index.html

**Uwagi:**

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt W1:**

Po zakończeniu, student: - posiada rozszerzoną i pogłębioną wiedzę z matematyki, w tym z zakresu systemów stochastycznych; - ma teoretycznie podbudowaną wiedzę z zakresu analizy danych; - ma szczegółową wiedzę w zakresie metod klasyfikacji wzorców, rozpoznawania obrazów i sygnałów mowy.

Weryfikacja:

Egzamin pisemny z zakresu wykładu. Projekt inżynierski demonstrujący znajomość metod i metodyki stosowanej dla rozwiązania wybranego zagadnienia rozpoznawania wzorców.

**Powiązane efekty kierunkowe:** K\_W01, K\_W02, K\_W04, K\_W06

**Powiązane efekty obszarowe:** T2A\_W01, T2A\_W02, T2A\_W07, T2A\_W02, T2A\_W07, T2A\_W01, T2A\_W03, T2A\_W04, T2A\_W07

### Profil ogólnoakademicki - umiejętności

**Efekt U1:**

potrafi posługiwać się technikami i narzędziami informacyjno-komunikacyjnymi właściwymi do realizacji zadań z zakresu rozpoznawania obrazów i sygnałów mowy; - potrafi planować i wykonywać symulacje komputerowe; - potrafi formułować i rozwiązywać proste problemy badawcze; - potrafi ocenić przydatność metod i narzędzi; - potrafi zrealizować projekt inżynierski na podstawie specyfikacji.

Weryfikacja:

Ocena realizacji projektu inżynierskiego, w tym sposobu analizy problemu i implementacji oraz dokumentacji.

**Powiązane efekty kierunkowe:** K\_U07, K\_U08, K\_U09, K\_U11, K\_U17, K\_U18

**Powiązane efekty obszarowe:** T2A\_U07, T2A\_U08, T2A\_U09, T2A\_U11, T2A\_U18, T2A\_U19

### Profil ogólnoakademicki - kompetencje społeczne

**Efekt K1:**

Po ukończeniu, student potrafi myśleć i działać w sposób kreatywny.

Weryfikacja:

Ocena poprawności własnych ścieżek rozwiązań zadań egzaminacyjnych. Ocena innowacyjności rozwiązania w projekcie inżynierskim.

**Powiązane efekty kierunkowe:** K\_K01

**Powiązane efekty obszarowe:** T2A\_K06