**Nazwa przedmiotu:**

Zaawansowane przetwarzanie danych multimedialnych

**Koordynator przedmiotu:**

Rajmund Kożuszek

**Status przedmiotu:**

Fakultatywny dowolnego wyboru

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Informatyka

**Grupa przedmiotów:**

Przedmioty techniczne - podstawowe

**Kod przedmiotu:**

ZPDM

**Semestr nominalny:**

4 / rok ak. 2021/2022

**Liczba punktów ECTS:**

4

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

1. liczba godzin kontaktowych – 64 godz., w tym
obecność na wykładach 30 godz.,
obecność na laboratorium 15 godz.,
obecność na egzaminie 2 godz.
spotkania projektowe 15 godz.
konsultacje 2 godz.
2. praca własna studenta – 61 godz., w tym
realizacja projektu 25 godz.,
przygotowanie do laboratorium 20 godz.,
przygotowanie do egzaminu 16 godz.
Łączny nakład pracy studenta wynosi 125 godz., co odpowiada 5 pkt. ECTS.

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

2,56 pkt. ECTS, co odpowiada 79 godz. kontaktowym

**Język prowadzenia zajęć:**

polski

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

2,20 pkt. ECTS, co odpowiada 55 godz. zajęć praktycznych

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

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

**Wymagania wstępne:**

Osoby uczęszczające na przedmiot powinny przede wszystkim mieć wiedzę związaną z podstawami cyfrowego przetwarzania sygnałów dźwiękowych i obrazów. Wymagana będzie również umiejętność programowania w języku Python lub korzystania ze środowiska Matlab.

**Limit liczby studentów:**

32

**Cel przedmiotu:**

Celem przedmiotu jest omówienie i analiza zaawansowanych zagadnień związanych z cyfrowym przetwarzaniem danych multimedialnych, obejmujących przede wszystkim przetwarzanie sygnałów muzyki, mowy i obrazu z uwzględnieniem akustycznych i optycznych zjawisk fizycznych, ograniczeń percepcji słuchowej i wzrokowej oraz efektów psychoakustycznych. Poruszane zagadnienia będą przedstawiane w formie interaktywnej z uwzględnieniem możliwości wykorzystania metod uczenia maszynowego (w tym sieci głębokich) w analizie, przetwarzaniu i syntezie sygnałów multimedialnych.

**Treści kształcenia:**

Wykład (30h):
1. Zagadnienia wstępne (2h):
• Sygnały oraz systemy foniczne i wizyjne, przetwarzanie analogowo-cyfrowe (próbkowanie, nadpróbkowanie, kwantyzacja, kształtowanie szumu), modulacje cyfrowe, przetworniki foniczne konwencjonalne (PCM) i sigma-delta (SDM).
• Współczesne systemy przetwarzania danych multimedialnych.
2. Kompresja, kodowanie i transmisja dźwięku (3h):
• Algorytmy kompresji stratnej z wykorzystaniem transformacji MDCT, modulacji ADPCM, SBC i innych sygnałów audio.
• Algorytmy kompresji bezstratnej.
• Standardy przewodowej i bezprzewodowej transmisji dźwięku.
3. Filtracja cyfrowa (2h).
• Filtry o stałej częstotliwości próbkowania.
• Zespoły filtrów i filtry o zmiennej częstotliwości próbkowania.
• Filtry adaptacyjne.
4. Cyfrowe efekty dźwiękowe (4h):
• Algorytmy przetwarzania dźwięku w dziedzinie amplitudy (procesory dynamiki sygnału, bramki szumów).
• Algorytmy przetwarzania dźwięku w dziedzinie czasu (echo, opóźnienia, pogłos, flanger, chorus).
• Algorytmy przetwarzania dźwięku w dziedzinie częstotliwości (korekcja charakterystyk częstotliwościowych, filtracja, redukcja szumów i zniekształceń).
5. Lokalizacja i śledzenie źródeł dźwięku, detekcja, śledzenie i rozpoznawanie obiektów, metody akwizycji i analizy obrazu (4h).
6. Algorytmy syntezy i rozpoznawania w przetwarzaniu danych multimedialnych (3h).
7. Ekstrakcja i analiza cech sygnałów dźwiękowych (2h).
8. Analiza semantyczna obrazu (2h).
9. Uczenie maszynowe w przetwarzaniu sygnałów fonicznych, obrazów i sekwencji wizyjnych (8h):
• Wspomaganie algorytmów ASR, detekcja, klasyfikacja sygnałów dźwiękowych, rozpoznawanie mówców, języka naturalnego, tłumaczenie.
• Detekcja, analiza, klasyfikacja i rozpoznawanie sygnałów muzycznych oraz środowiskowych.
• Wspomaganie przetwarzania danych multimedialnych z przetworników analogowo-cyfrowych i wizyjnych w lokalizacji i śledzeniu źródeł dźwięku i obiektów.
• Detekcja i klasyfikacja obiektów obrazach statycznych i sekwencjach obrazów.
Laboratorium (15h):
Ćwiczenia laboratoryjne są zorganizowane w formie pięciu bloków tematycznych po trzy godziny zajęć i są realizowane w zespołach 2 osobowych w grupach laboratoryjnych 8 osobowych. Treści zadań laboratoryjnych obejmują: (1) badanie algorytmów kompresji i kodowania dźwięku oraz implementację algorytmów automatycznego rozpoznawania mowy (ASR), (2) implementację wybranych cyfrowych efektów dźwiękowych, (3) przetwarzanie sygnałów z macierzy mikrofonowych i głośnikowych oraz badanie algorytmów lokalizacji i śledzenia, (4) detekcję i klasyfikację obiektów w sekwencji obrazów oraz (5) ekstrakcję i analizę cech danych multimedialnych. Student zapoznaje się również z technikami projektowania i programowania algorytmów niezbędnych w badaniach związanych z analizą danych multimedialnych.
Projekt (15h):
Projekt jest definiowany w formie wymagań i parametrów, jakie musi spełniać aplikacja i jest realizowany w grupach projektowych złożonych z 4-6 osób. Grupa projektowa wybiera implementację aplikacji z aktualnej listy projektów lub proponuje własną aplikację do akceptacji prowadzącego zajęcia projektowe. Treści zadań projektowych dotyczą zagadnień uczenia maszynowego w przetwarzaniu sygnałów fonicznych, obrazów i sekwencji wizyjnych.

**Metody oceny:**

Zajęcia laboratoryjne są prowadzone w grupach 2 osobowych z podziałem na zespoły laboratoryjne liczące 8 osób (5 terminów po 3 godziny).
Projekt jest realizowany w grupach liczących od 4 do 6 osób. Spotkanie projektowe w sumie dla każdej grupy to 3 godziny.

**Egzamin:**

tak

**Literatura:**

1. LERCH, Alexander. An introduction to audio content analysis: Applications in signal processing and music informatics. Wiley-IEEE Press, 2012.
2. HILL, Paul. Audio and Speech Processing with MATLAB. CRC Press, 2018.
3. LI, Francis F.; COX, Trevor J. Digital Signal Processing in Audio and Acoustical Engineering. CRC Press, 2019.
4. ZÖLZER, Udo. Digital audio signal processing. New York: Wiley, 2008.
5. ZIELIŃSKI, T. P.; KOROHODA, P.; RUMIAN, R. Cyfrowe przetwarzanie sygnałów w telekomunikacji. Wydawnictwo Naukowe PWN, 2014.
6. DOWNEY, Allen B. Think DSP: digital signal processing in Python. " O'Reilly Media, Inc.", 2016.

**Witryna www przedmiotu:**

https://usosweb.usos.pw.edu.pl/kontroler.php?\_action=katalog2/przedmioty/pokazPrzedmiot&prz\_kod=103A-TLTBM-MSP-ZPDM

**Uwagi:**

(-)

## Charakterystyki przedmiotowe