

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Przetwarzanie i rozpoznawanie dźwięku		Kod 1010512321010510327
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 1 / 2
Ścieżka obieralności/specjalność Inteligentne systemy wspomagania decyzji	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obligatoryjny
Stopień studiów: II stopień	Forma studiów (stacjonarna/niestacjonarna) stacjonarna	
Godziny Wykłady: 30 Ćwiczenia: - Laboratoria: 30 Projekty/seminaria: -		Liczba punktów 4
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) kierunkowy		(ogólnouczelniany, z innego kierunku) z danego kierunku
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne		Podział ECTS (liczba i %) 4 100%
Odpowiedzialny za przedmiot / wykładowca:		
<p>dr inż. Ewa Łukasik email: ewa.lukasik@cs.put.poznan.pl tel. 61 6652922 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań</p>		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien mieć podstawową wiedzę z matematyki i uczenia maszynowego oraz podstaw cyfrowego przetwarzania sygnałów.
2	Umiejętności:	Powinien mieć umiejętność rozwiązywania podstawowych problemów z dziedziny zastosowań informatyki oraz umiejętność pozyskiwania informacji ze wskazanych źródeł.
3	Kompetencje społeczne	Powinien również rozumieć konieczność poszerzania swoich kompetencji i mieć gotowość do podjęcia współpracy w ramach zespołu. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
Cel przedmiotu:		
<ol style="list-style-type: none"> Przekazanie studentom podstawowej wiedzy w analizie sygnałów dźwiękowych (środowiskowych, audiofonicznych i mowy) Zapoznanie studentów z podstawowymi metodami i algorytmami ekstrakcji cech z sygnałów dźwiękowych dla celów wyszukiwania informacji, klasyfikacji, rozpoznawania i identyfikacji. Zapoznanie studentów z podstawowymi metodami rozpoznawania mowy i mówców oraz syntezy mowy. Rozwijanie u studentów umiejętności definiowania nowych obszarów zastosowań informatyki oraz rozwiązywania prostych problemów związanych z implementacją poznanych algorytmów w nowych zastosowaniach. 		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		
<ol style="list-style-type: none"> Ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie algorytmów i metod przetwarzania i rozpoznawania dźwięku pod kątem zastosowań m.in. w zakresie wspomagana decyzji, baz danych i sztucznej inteligencji. - [K_W4] Ma podbudowaną teoretycznie szczegółową wiedzę w zakresie algorytmów ekstrakcji cech i klasyfikacji dźwięku (środowiskowego, mowy, audiofonicznego) - [K_W5] Ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w informatyce związanych z eksplozją danych multimedialnych, a szczególnie związanych z dźwiękiem w różnych dziedzinach pokrewnych (cyfrowa dystrybucja, wyszukiwanie, komunikacja człowieka z komputerem, sonifikacja) - [K_W6] Ma podstawową wiedzę o cyklu życia systemów przetwarzania i rozpoznawania dźwięku - [K_W7] Zna podstawowe metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich w zakresie przetwarzania i rozpoznawania dźwięku - [K_W8] 		
Umiejętności:		

<ol style="list-style-type: none">1. Potrafi pozyskiwać informacje z literatury (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie. - [K_U1]2. Potrafi określić kierunki dalszego uczenia się i zrealizować proces samokształcenia. - [K_U5]3. Potrafi wykorzystać do formułowania i rozwiązywania zadań inżynierskich i prostych problemów badawczych metody analityczne, symulacyjne oraz eksperymentalne. - [K_U9]4. Potrafi - przy formułowaniu i rozwiązywaniu zadań inżynierskich - integrować wiedzę z różnych obszarów informatyki i dziedzin pokrewnych oraz zastosować podejście systemowe, uwzględniające także aspekty pozatechniczne. - [K_U10]5. Potrafi formułować i testować hipotezy związane z problemami inżynierskimi i prostymi problemami badawczymi w zakresie przetwarzania i rozpoznawania dźwięku. - [K_U12]6. Potrafi ocenić przydatność i możliwość wykorzystania nowych osiągnięć (metod i narzędzi) oraz nowych produktów informatycznych w zakresie przetwarzania i rozpoznawania dźwięku - [K_U13]7. potrafi ocenić złożoność obliczeniową algorytmów i problemów związanych z przetwarzaniem i rozpoznawaniem dźwięku - [K_U16]8. potrafi - stosując m.in. koncepcyjnie nowe metody - rozwiązywać nietypowe zadania informatyczne oraz zadania zawierające komponent badawczy - [K_U25]
Kompetencje społeczne:
<ol style="list-style-type: none">1. Rozumie, że w informatyce, a zwłaszcza w przetwarzaniu danych, technologie i umiejętności związane z implementacją metod kompresji danych bardzo szybko stają się przestarzałe. - [K_K1]2. Zna możliwości dalszego kształcenia się w zakresie algorytmów kompresji i kodowania danych i ich zastosowania w różnych dziedzinach - [K_K3]3. Zna przykłady i rozumie przyczyny wadliwie działających systemów informatycznych, które doprowadziły do poważnych strat finansowych, społecznych lub też do poważnej utraty zdrowia, a nawet życia. - [K_K4]4. Potrafi odpowiednio określić priorytety służące realizacji określonego przez siebie lub innych zadania. - [K_K6]

Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
 - na podstawie odpowiedzi na pytania dotyczące materiału omówionego na wykładach.
- b) w zakresie laboratoriów / ćwiczeń:
 - na podstawie oceny bieżącego postępu realizacji zadań.

Ocena podsumowująca:

- a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym różnej charakterystyce problemów do rozwiązania: 40% pytań dotyczy podstawowej wiedzy i jest przedstawiona w postaci testowej (pytania testowe wielokrotnego wyboru, treść do uzupełnienia), 40% pytań stanowią proste zadania algorytmiczne lub interpretacyjne, natomiast pozostałe 20% pytań to zadania problemowe o większej złożoności; liczba pytań na egzaminie to ok. 10; wszystkie pytania są podobnie punktowane, łącznie można otrzymać 100 punktów; zaliczenie kolokwium jest od 50 punktów; na ostateczną ocenę składa się w 60% ocena z kolokwium pisemnego i w 40% ocena z laboratorium.
 - omówienie wyników egzaminu.
- b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę przygotowania studenta do poszczególnych sesji zajęć laboratoryjnych (punktowany sprawdzian wejściowy)
 - ocenę ciągłą umiejętności związanych z realizacją ćwiczeń laboratoryjnych, które sprawdzane są ustnie i także punktowane.
 - ocenę sprawozdań, które wymagane są do niektórych ćwiczeń laboratoryjnych (łączna punktacja wynikająca z trzech powyższych ocen stanowi 60% oceny końcowej)
 - 40% oceny końcowej z laboratorium student uzyskuje na podstawie przygotowania (30%) i prezentacji (10%) zadania domowego - projektu dotyczącego zagadnień rozszerzających materiał realizowany na zajęciach. Obejmuje ono eksplorację zastosowań metod przetwarzania i rozpoznawania dźwięku poznanych na zajęciach, metod nietypowych, algorytmów zaimplementowanych w nietypowych środowiskach (mobilne systemy operacyjne, GPU), bądź najnowszych rozwiązań i standardów.
 - Jest możliwość uzyskania dodatkowych punktów za uwagi związane z udoskonaleniem i udoskonalanie materiałów dydaktycznych oraz za wskazywanie trudności percepcyjnych studentów umożliwiające bieżące doskonalenie procesu dydaktycznego.

Treści programowe

Wykłady wygłaszane są w 3-godzinnych w ciągu 10 spotkań.

- Powtórzenie i ugruntowanie i rozszerzenie wiadomości dotyczących interpretacji i cyfrowej analizy sygnałów w odniesieniu do sygnałów dźwiękowych; szeregi Fouriera, widmo amplitudowe i fazowe, moc, energia RMS sygnału, splot cyfrowy, twierdzenie o próbkowaniu, aliasing, przepróbowanie w górę i w dół, funkcje okna i ich wpływ na reprezentację widma, reprezentacje czasowo-częstotliwościowe, spektrogram, filtry cyfrowe o skończonej i nieskończonej odpowiedzi impulsowej.
- Sposób i model generowania sygnału mowy i jego cyfrowa implementacja za pomocą metod liniowej predykcji. Schemat analiza-synteza. Charakterystyka fonetyczna i widmowa typowych fonemów języka polskiego i angielskiego. Wizualne rozpoznawanie typowych głosek na podstawie ich widma i spektrogramu.
- Metody generowania obwiedni widma sygnału na podstawie analizy cepstralnej, deskryptorów MPEG-7 oraz na podstawie biegunów filtru uzyskanego za pomocą analizy LPC
- Percepcyjne skale częstotliwości: barkowa, melowa i ERB, współczynniki cepstralne w skali melowej i ich wykorzystanie w charakterze atrybutów sygnału mowy. Preemfaza sygnału mowy
- Dynamiczne dopasowanie czasowe sygnału mowy (DTW - Dynamic Time Warping) dla celów synchronizacji czasowej wypowiedzi. Rozpoznawanie minimalno odległościowe
- Ukryte modele Markowe (HMM - Hidden Markov Models) jako podstawowa metoda rozpoznawania mowy. Przykłady systemów
- Mieszanie rozkładów Gaussa (GMM - Gaussian Mixture Models) i ich zastosowanie w rozpoznawaniu mówców
- Metody syntezy mowy. Przykładowe systemy.
- Przegląd metod ekstrakcji cech dla sygnałów audiofonicznych i dźwięków środowiskowych (MPEG-7, parametry statystyczne, chroma). Metody wyszukiwania informacji muzycznej. Automatyczna transkrypcja muzyki monofonicznej i polifonicznej (ekstrakcja częstotliwości podstawowej, rozpoznawanie tonacji, rozpoznawanie akordów, metrum i rytmu, instrumentów muzycznych, gatunku muzycznego). Metody odseparowywania źródeł dźwięku. Rozpoznawanie śpiewu ptaków.
- Generowanie efektów dźwiękowych. Złudzenia akustyczne.

Zajęcia laboratoryjne prowadzone są w formie piętnastu dwugodzinnych ćwiczeń, odbywających się w laboratorium. Ćwiczenia realizowane są indywidualnie. W wyjątkowych wypadkach zadania projektowe mogą być wykonywane w zespołach 2-osobowych. Pierwsze zajęcia są poświęcone omówieniu programu ćwiczeń i przedstawieniu środowiska dydaktycznego, po czym następuje 11 ćwiczeń laboratoryjnych ściśle związanych z ilustrującą tematykę poruszaną na wykładach. Ostatnie 3 zajęcia są przeznaczone na prezentację projektów/eksperymentów/badań/studiów zagadnień rozszerzających tematykę przedmiotu o nowe zagadnienia, zastosowania i implementacje. Regularne ćwiczenia laboratoryjne podzielone są na 4 bloki tematyczne:

1. Zestaw ćwiczeń laboratoryjnych ugruntowujących wiedzę podstawową z dziedziny cyfrowego przetwarzania sygnału dźwiękowego i narzędzi wizualizacji ściśle związane z materiałem prezentowanym na zajęciach
2. Analiza sygnału mowy: określanie częstotliwości podstawowej, charakterystyki częstotliwościowej poszczególnych fonemów, identyfikacja charakterystycznych formantów, analiza LPC i analiza cepstralna. Analiza charakterystycznych cech mówców. Dynamiczne dopasowanie czasowe (DTW) i rozpoznawanie mowy
3. Charakterystyka sygnałów muzycznych. Wykorzystanie dostępnych bibliotek programowych do realizacji zadań związanych z automatyczną transkrypcją nagrań muzycznych (wykrywanie częstotliwości podstawowej, tonacji, akordów, analiza chromatyczna, badanie podobieństwa utworów muzycznych)
4. Badanie interaktywnych interfejsów umożliwiających aktywne słuchanie muzyki oraz wyszukiwanie melodii na podstawie przykładu melodycznego bądź rytmicznego)

Metody dydaktyczne:

1. wykład: prezentacja multimedialna, prezentacja ilustrowana przykładami podawanymi na tablicy, pokaz multimedialny, demonstracja.
2. ćwiczenia laboratoryjne: ćwiczenia praktyczne i wykonywanie eksperymentów w środowisku dydaktycznym, dyskusja, demonstracja.
3. prezentacje studentów relacjonujące wykonane samodzielnie eksperymenty, badania, studia literaturowe, własne programy realizujące wybrane metody przetwarzania i rozpoznawania dźwięku, metody i algorytmy stanowiące rozszerzenie materiału obowiązkowego, interesujące zastosowania.

Literatura podstawowa:

1. A. Czyżewski, Dźwięk cyfrowy, EXIT, W-wa, 1998
2. T. Zieliński, Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, WKŁ, Wyd. 2 (dodruk), 2009
3. R. Tadeusiewicz, Sygnał mowy, WKŁ, Warszawa 1988, <http://winntbg.bg.agh.edu.pl/skrypty/4/>

Literatura uzupełniająca:

1. A. Klapuri (Ed.), Signal Processing Methods for Music Transcription, Springer, 2006
2. B. Kostek, Perception-Based Data Processing in Acoustics: Applications to Music Information Retrieval and Psychophysiology of Hearing, Springer, 2010
3. B. Kostek, Soft Computing in Acoustics: Applications of Neural Networks, Fuzzy Logic and Rough Sets to Musical Acoustics, Springer, 1999.
4. IEEE Transactions on Audio, Speech, and Language Processing, <http://www.signalprocessingsociety.org/publications/periodicals/taslp/>

Bilans nakładu pracy przeciętnego studenta

Czynność		Czas (godz.)
1. Udział w zajęciach laboratoryjnych/ćwiczeniach		30
2. Dokończenie (w ramach pracy własnej) zadań z ćwiczeń laboratoryjnych lub przygotowanie sprawozdania:		5
3. Zadanie domowe: realizacja projektu/eksperymentu + przygotowanie prezentacji		5
4. Udział w konsultacjach związanych z realizacją procesu kształcenia (częściowo mogą być realizowane drogą elektroniczną)		6
5. Bieżące przygotowanie do zajęć:		10
6. Udział w wykładach:		30
7. Zapoznanie się ze wskazaną literaturą i materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.), 50 stron		5
8. Przygotowanie do egzaminu		8
9. Udział w egzaminie		2
10. Omówienie wyników egzaminu		2
Obciążenie pracą studenta		
forma aktywności	godzin	ECTS
Łączny nakład pracy	107	4
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	74	3
Zajęcia o charakterze praktycznym	50	2