**Nazwa przedmiotu:**

Podstawy teoretyczne multimediów

**Koordynator przedmiotu:**

dr hab. inż. Przemysław Dymarski, dr inż. Artur Janicki, doc. dr Sławomir Kula, mgr inż. Marcin Golański

**Status przedmiotu:**

Fakultatywny ograniczonego wyboru

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Telekomunikacja

**Grupa przedmiotów:**

Przedmioty techniczne - zaawansowane

**Kod przedmiotu:**

PTMT

**Semestr nominalny:**

4 / rok ak. 2012/2013

**Liczba punktów ECTS:**

4

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

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

**Język prowadzenia zajęć:**

polski

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

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

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

**Wymagania wstępne:**

Podstawy przetwarzania sygnałów i telekomunikacji

**Limit liczby studentów:**

70

**Cel przedmiotu:**

Student zapoznaje się z podstawami teoretycznymi przetwarzania sygnałów multimedialnych, w tym z metodami kompresji bezstratnej i stratnej w zastosowaniu do mowy, szerokopasmowych sygnałów akustycznych, obrazów nieruchomych i sekwencji wideo. Nabiera umiejętności w zakresie wykorzystania technik kompresji w usługach telekomunikacyjnych, jak VoIP, wideokonferencja, VoD, IPTV. Ponadto zapoznaje się z technikami znakowania wodnego sygnałów multimedialnych, rozpoznawania mowy i mówcy.

**Treści kształcenia:**

Treść wykładu (w nawiasie liczba godzin):

Wprowadzenie w techniki multimedialne . Standardy multimedialne i standardy kompresji. Wykorzystanie multimediów w telekomunikacji. Wideokonferencja jako przykład techniki multimedialnej. (1)
Sygnały występujące w systemach multimedialnych i ich modele matematyczne. Właściwości sygnału mowy, sygnałów akustycznych, obrazów nieruchomych, sekwencji wideo. Modele sygnałów: procesy stochastyczne, autokorelacja, gęstość mocy, procesy gaussowskie, procesy Markowa. Metody oceny jakości sygnałów i usług telekomunikacyjnych opartych na transmisji dźwięku i obrazu. (2)
Podstawowe wiadomości z teorii informacji (repetytorium). Ilość informacji, entropia, twierdzenie Shannona o kodowaniu źródła, twierdzenie Shannona o pojemności informacyjnej kanału. Kody przedrostkowe, kodowanie bezstratne - kod Huffmana, metody słownikowe. (2)
Kwantowanie sygnałów. Kwantyzator równomierny, nierównomierny, algorytm Lloyda projektowania kwantyzatora, entropia na wyjściu kwantyzatora, kwantowanie optymalne przy danej entropii. Kwantyzatory z adaptacją. Wykorzystanie w standardzie telefonicznym PCM (ITU-T G.711) i ADPCM (G.726) – (2)
Kwantowanie wektorowe. Kwantyzator wektorowy jako optymalny (w sensie granicznym) koder źródła informacji. Teoria wysokiej rozdzielczości. Algorytm LBG projektowania kwantyzatora wektorowego. Kwantyzatory wektorowe typu kształt-wzmocnienie. (2)
Liniowa predykcja sygnału. Metoda autokorelacyjna i kowariancyjna. Dekompozycja Choleskiego i algorytm Levinsona-Durbina. Zysk predykcji, struktury predyktora, stabilność filtru predykcyjnego. Kodowanie współczynników predyktora. Interpretacja liniowej predykcji w dziedzinie częstotliwości; modele procesów AR, MA, ARMA. (3)
Kodowanie różnicowe. Struktura kodera DPCM i ADPCM. Metody adaptacji predyktora. Zastosowanie w telefonii w pasmach 3.4kHz (G.726) i 7kHz (G.722). - (1)
Kodery predykcyjno-wektorowe (CELP). Zastosowanie liniowej predykcji do dekompozycji "słownika" kwantyzatora wektorowego. Predykcja długookresowa sygnałów quasi-periodycznych (np. sygnału mowy dźwięcznej). Zastosowanie w standardach telefonii klasycznej (G.728, G.729 , G.723.1) i ruchomej (ETSI GSM-HR, GSM-EFR, GSM-AMR), a także wideotelefonii (H.324). – (2)
Kodowanie dźwięku w dziedzinie częstotliwości. Kodery subpasmowe i kodery transformaty. Banki filtrów umożliwiające dokładną rekonstrukcję. Zjawiska psychoakustyczne i ich wykorzystanie w kompresji sygnałów audio. Maskowanie szumu sygnałem użytecznym, wyznaczanie progów maskowania. Kształtowanie widma szumu kwantyzacji. Algorytmy adaptacyjnego rozdziału bitów z wykorzystaniem progów maskowania. Zastosowanie w kodowaniu sygnałów fonicznych: standardy MPEG, G.722.1, ATRAC, AAC. (3)
Kodowanie obrazu nieruchomego i ruchomego. Wykorzystanie transformat dwuwymiarowych. Standard kodowania obrazów nieruchomych JPEG. Estymacja i kompensacja ruchu w kodowaniu sekwencji wideo. Standardy kompresji sekwencji wideo MPEG-1, MPEG-2, MPEG-4, w szczególności H.264. (2)
Zabezpieczanie plików dźwiękowych – “znaki wodne”. Ukrywanie sygnału transmisji danych w sygnale akustycznym. Wykorzystanie zjawiska maskowania “znaku wodnego” sygnałem audio. (2)
Metody syntezy tekstowej mowy. Wstępne przetwarzanie ciągu znaków alfanumerycznych. Jednostki akustyczne używane w syntezie. Synteza wg reguł, synteza połączeniowa, synteza synchronizowana tonem krtaniowym. (2)
Rozpoznawanie mowy. Wykorzystanie parametrów czasowych, częstotliwościowych i cepstralnych. Dynamiczna normalizacja czasowa. Ukryte modele Markowa (ciągłe i dyskretne). Model Bakisa. Kryteria decyzyjne. (2)
Zaawansowane techniki przetwarzania sygnału mowy. Rozpoznawanie (identyfikacja i weryfikacja) mówcy. Rozpoznawanie stanu emocjonalnego mówcy. Wizyjna synteza mowy. Transformacja głosu. Poprawianie jakości sygnału mowy. Algorytmy niwelowania strat pakietów w sieciach VoIP. Sądowe zastosowania przetwarzania sygnału mowy. (4)

Laboratorium:
 Laboratorium obejmuje 6 dwugodzinnych ćwiczeń, w trakcie których studenci wykonują w zespołach dwuosobowych pomiary i obliczenia.
1. Kwantowanie wektorowe sygnałów audio i obrazów nieruchomych. Projektowanie słownika: uogólniony algorytm Lloyda i algorytm LBG. Symulacja kwantyzatora optymalnego i suboptymalnego (kształt- wzmocnienie).

2. Kodery CELP i wokodery predykcyjne. Symulacja komputerowa kodera predykcyjnego o pobudzeniu stochastycznym CELP - badanie szumu kwantyzacji w funkcji wybranych parametrów. Symulacja wokodera predykcyjnego - analiza zniekształceń w procesie kodowania.

3. Kompresja sygnałów w dziedzinie częstotliwości - kodery subpasmowe. Symulacja koderów MPEG- Audio o różnych przepływnościach binarnych, badanie efektu maskowania szumu kwantyzacji. Porównanie algorytmów rozdziału bitów między podpasma.

4. Symulacja kodera sekwencji wideo H.264. Obserwacja różnych rodzajów zniekształceń. Wpływ kompensacji ruchu na jakość pracy kodera.

5. Synteza tekstowa mowy. Zapoznanie się z konkatenacyjną metodą syntezy na bazie różnych jednostek akustycznych.

6. Rozpoznawanie mowy. Badanie mocy dystynktywnej wybranych parametrów sygnału w procesie rozpoznawania głosek i izolowanych słów.

**Metody oceny:**

Sposoby weryfikacji zakładanych efektów kształcenia:
 Egzamin: 60% (musi być ocena pozytywna)
 Oceny z ćwiczeń laboratoryjnych: 40% (średnia z 6 ocen pozytywnych)

**Egzamin:**

tak

**Literatura:**

N.S.Jayant, P.Noll "Digital coding of waveforms"
A.Gersho, R.M.Gray "Vector quantization and signal compression"
R.Tadeusiewicz "Sygnał mowy"
Cz.Basztura "Jak rozmawiać z komputerem"
W.Skarbek ”MULTIMEDIA – algorytmy i standardy kompresji”
A.Drozdek “Wprowadzenie do kompresji danych”
A.M. Kondoz “Digital speech”
L.Hanzo, F.Clare, A.Somerville, J.P.Woodward: “Voice compression and communications”
K.Sayood “Kompresja danych – wprowadzenie”, Wyd. RM, W-wa 2002
P.Vary, R.Martin „Digital speech transmission”, Wiley 2005
K.K. Parhi, T. Nishitani « Digital Signal Processing for Multimedia Systems » , Marcel Dekker, New York, 1999
T.P.Zieliński « Cyfrowe przetwarzanie sygnałów – od teorii do zastosowań », WKiŁ, Warszawa 2005
W.Kasprzak “Rozpoznawanie obrazów i sygnałów mowy”
N. Netravali, B. Haskell "Digital pictures: representation, compression, and standards"
M.Ghanbari „Standard codecs – image compression to advanced video coding”

**Witryna www przedmiotu:**

www.tele.pw.edu.pl/ptmt/

**Uwagi:**

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt Wpisz opis:**

Orientuje się w zaawansowanych technikach przetwarzania mowy, jak synteza z tekstu, rozpoznawanie, weryfikacja mówcy.

Weryfikacja:

sprawozdania z ćwiczeń laboratoryjnych, egzamin

**Powiązane efekty kierunkowe:** K\_W07, K\_W10

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

### Profil ogólnoakademicki - umiejętności

**Efekt Wpisz opis:**

Potrafi ocenić jakość sygnału mowy, muzyki i obrazu ruchomego

Weryfikacja:

laboratorium

**Powiązane efekty kierunkowe:** K\_U07, K\_U08, K\_U10

**Powiązane efekty obszarowe:** T2A\_U07, T2A\_U08, T2A\_U09, T2A\_U12, T2A\_U13, T2A\_U07, T2A\_U09, T2A\_U15

### Profil ogólnoakademicki - kompetencje społeczne

**Efekt Wpisz opis:**

Realizacja ćwiczeń laboratoryjnych w zespołach

Weryfikacja:

sprawozdania z ćwiczeń, obserwacja funkcjonowania studenta w trakcie realizacji ćwiczenia

**Powiązane efekty kierunkowe:** K\_K01

**Powiązane efekty obszarowe:** T2A\_K06