

MATERIAŁY POMOCNICZE DO WYKŁADU Z KOMPTEROWYCH SYSTEMÓW EDYCJI DŹWIĘKU

2. Formaty i parametry plików dźwiękowych oraz ich graficzna reprezentacja (1 godzina wykładu)

S3. Plik dźwiękowy – definicje

Plik dźwiękowy (*ang. sound file*) to zbiór danych w postaci binarnej, które odwzorowują dowolny sygnał audio w taki sposób, aby mógł on być rozpoznawany i odtwarzany za pomocą komputera.

Istnieje wiele **formatów** plików dźwiękowych, z których każdy reprezentuje jeden określony sposób rozmieszczenia dyskretnych danych odwzorowujących sygnał audio. Formaty te wykorzystywane są przez odpowiednie programy komputerowe do czytania i zapisu informacji dźwiękowych:

IFF/8SVX; VOC; SND/AU; VOX; MOV; AIFF; WAV; AVI;
RAW;
MPEG; ASF; RM (RA, RV); OGG; WMA i WMV;
MOD; MID; KAR

S4. Tworzenie plików dźwiękowych

Przetwarzanie A/C analogowego sygnału audio na cyfrowy następuje przez doprowadzenie go do wejść kodera PCM (*ang. Pulse Code Modulation* – modulacja kodowo-impulsowa).

S5. Schematy kodowania plików dźwiękowych

PCM (*Pulse Code Modulation*) – bez kompresji;

ADPCM (*Adaptive Differential PCM* – adaptacyjna różnicowa modulacja kodowo-impulsowa) – zapis różnic pomiędzy wartościami próbek;

μ -law, A-law – opracowane dla systemów telefonicznych – logarytmiczna skala kwantyzacji;

TrueSpeech – olbrzymi stopień kompresji do 93 %, zamiast zapisu próbek analizowana jest i zapisywana charakterystyka głosu mówcy;

MPEG (*The Motion Picture Experts Group*) – kodowanie percepcyjne poprzez maskowanie częstotliwości i poziomu dźwięku.

S.6-8. Formaty plików dźwiękowych

Pliki WAV i RIFF (Resource Interchange File Format);

PCM – “surowe” dane dźwiękowe PCM;

SND – rodzimy format systemów operacyjnych Macintosha;

AIFF (*Audio Interchange File Format*) i AIFC – dźwiękowy format plików Macintosha;

AU (inaczej Sun Audio lub NeXT) – format systemów Unix/Linux opracowany przez firmę Sun Microsystems;

MP3 – MPEG Layer 3;

TXT – tekstowe dane dźwiękowe

Przykład pliku tekstowego:

```
SAMPLES: 1582
BITPERSAMLE: 16
CHANNELS: 2
SAMPLERATE: 22050
NORMALIZED: FALSE
164 <tab> -1372
492 <tab> -876
....
```

S.10. Transformata Fouriera

DFT – dyskretna transformata Fouriera;

$$Y_k = \sum_{j=0}^{n-1} X_j e^{-2\pi j k \sqrt{-1} / n}$$

FFT Cooleya i Tuckeya (rozmiar danych – potęga 2) – szybka transformata Fouriera;

Efektywna implementacja FFT Real Laurenta de Sorasa;

Okienkowanie danych – okno Hamminga:

$$W_n = 0.54 - 0.46 \cos\left(\frac{2n\pi}{N-1}\right)$$

S.11. Transformata Fouriera

Aliasing – niejednoznaczność postaci sygnału w dziedzinie częstotliwości;

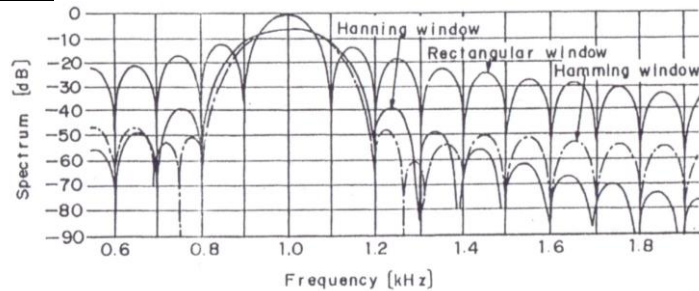
S.12. Transformata Fouriera

Zniekształcenia sygnału;

S.13. Okienkowanie danych

Stosowane nakładki: 25 ÷ 75 %

S.14. Przeciek widma



S.15-18. Wizualizacja widma - sposoby

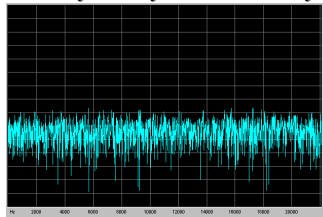
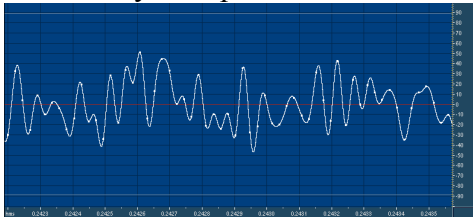
ETF (*Energy Time-Frequency*);

CSD (Cumulative SPECTral Density);

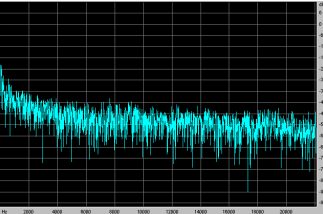
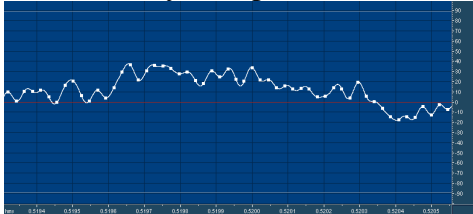
Spektrogram

S.23. Graficzna reprezentacja plików

Szum biały – amplituda widma nie zmienia się z częstotliwością – ch-ka płaska;



Szum różowy – amplituda widma $\sim 1/f$



Szum brązowy – amplituda widma $\sim 1/f^2$

