

Przemysław Dymarski
Instytut Telekomunikacji Politechniki Warszawskiej

ZAAWANSOWANE TECHNIKI PRZETWARZANIA SYGNAŁÓW W TELEKOMUNIKACJI

Ćwiczenie 4 Transformaty w kompresji sygnałów fonicznych

Uwagi ogólne

- Program symulacyjny nosi nazwę tran.exe. Posiada rozbudowany system pomocy.
- Sygnałem wejściowym i wyjściowym są pliki .wav lub .pcm. Przetwarza się sygnał monofoniczny o częstotliwości próbkowania 44100 Hz.
- Program umożliwia obserwację przebiegu czasowego sygnałów przed i po kompresji (jeśli kwantuje się współczynniki transformaty). Jest możliwość dokonania odsłuchu sygnałów, oraz obliczenia SNR - stosunku mocy sygnału do mocy szumu kwantyzacji.
- Można obserwować wykres współczynników transformaty (DFT, DCT, ELT) przed i po kwantyzacji, oraz odczytywać wartości liczbowe współczynników.
- Kwantowanie może odbywać się w całym paśmie, lub podpasmach częstotliwości. Kwantyzatory są adaptacyjne, z adaptacją "w przód" (można wybrać liczbę bitów przeznaczoną na zakodowanie zakresu pracy kwantyzatora). Rozdział bitów pomiędzy kwantyzatory może być sterowany "ręcznie" (wpisuje się liczbę bitów dla danego podpasma), albo automatycznie. W tym ostatnim przypadku wykorzystuje się kryterium energetyczne (minimalizacja mocy błędu kwantowania) lub psychoakustyczne (wykorzystanie maskowania błędu kwantowania sygnałem akustycznym). Dla "psychoakustycznego" rozdziału bitów wykorzystuje się 25 podpasm o nierównej szerokości (skala "barków").

Zadania do wykonania

1. Porównanie dyskretnej transformaty Fouriera (DFT, FFT) i transformaty cosinusoidalnej (DCT)

Do analizy wybierz sygnał sinusoidalny 2000 Hz (plik sin2000.wav). Wybierz transformatę DFT (w programie pod nazwą FFT - Fast Fourier Transform) o długości okna 256 próbek, a następnie 1024 próbki. Wyłącz kwantyzację! Obejrzyj wykres współczynników DFT i wartości współczynników (program wyprowadza wartości bezwzględne zespolonych współczynników DFT). Które współczynniki mają wartość maksymalną? Dlaczego jest para takich współczynników? Określ częstotliwości odpowiadające tym największym współczynnikom.

Powtórz eksperyment dla DCT. Czym różni się DCT od DFT? Który współczynnik niesie teraz największą energię? Określ częstotliwość odpowiadającą temu największemu współczynnikowi.

Obserwuj wykres DCT dla kolejnych ramek sygnału sinusoidalnego. Dlaczego wykres się zmienia z ramki na ramkę? Czy to samo zjawisko występuje dla modułów współczynników DFT? Która transformata, DFT czy DCT, jest lepszym analizatorem częstotliwości?

2. Kwantowanie w podzakresach częstotliwości

Do symulacji wykorzystaj wybraną frazę muzyczną. Wybierz DCT o długości ramki 512 próbek, a liczbę bitów na ramkę =2048, lub najbliższą tej wartości. Oznacza to kodowanie ze średnią przepływnością 4 bity na próbkę sygnału.

a) Wykorzystanie adaptacji kwantyzatorów

Włącz kwantyzację, lecz wyłącz analizę psychoakustyczną i alokację bitów (bity będą rozdzielane "ręcznie"). Ustaw 5 bitów na kwantowanie zakresu pracy kwantyzatorów operujących w podpasmach (wart. max.). Rozdzielając bity po równo między podpasma, przeprowadź symulację:

- kwantowanie transformaty w całym paśmie (liczba podpasm =1, 2048 bitów w tym podpaśmie)
- liczba podpasm =4, po 512 bitów w każdym podpaśmie
- liczba podpasm =16, po 128 bitów w każdym podpaśmie

Przeprowadź odsłuchy i zanotuj SNR. Co jest przyczyną poprawy jakości przy wykorzystaniu dużej liczby podpasm?

b) Wykorzystanie alokacji bitów

Włącz kwantyzację i alokację bitów, lecz wyłącz analizę psychoakustyczną. Ustaw 2048 bitów na ramkę i 5 bitów na kwantowanie zakresu pracy kwantyzatorów operujących w podpasmach (wart. max.). Wykonaj symulację kodera 1-, 4-, i 16-pasmowego z automatyczną alokacją bitów pomiędzy podpasma (energetyczne kryterium rozdziału bitów). Jak uzasadnisz poprawę jakości sygnału w porównaniu z równomiernym rozdziałem bitów (punkt a)?

c) Wykorzystanie analizy psychoakustycznej

Włącz kwantyzację, alokację bitów i analizę psychoakustyczną. Ustaw 2080 bitów na ramkę i 5 bitów na kwantowanie zakresu pracy kwantyzatorów operujących w podpasmach (wart. max.). Wykonaj symulację kodera z automatyczną alokacją bitów pomiędzy 25 podpasm (psychoakustyczne kryterium rozdziału bitów). Zaobserwuj kształt krzywej maskowania. Która metoda alokacji bitów, energetyczna czy psychoakustyczna, oferuje większe SNR? Dlaczego stosowany jest "psychoakustyczny" rozdział bitów?

3. Transformata cosinusoidalna "z nakładaniem": ELT, MDCT

Dla transformaty ELT przetwarzającej 1024 próbki sygnału audio w 256 próbek transformaty zaobserwuj (na sygnale audio) jakie jest położenie sąsiednich okien na skali czasu. Przy wyłączonej kwantyzacji sprawdź odwracalność transformacji prostej i odwrotnej. Wyjaśnij jak obliczana jest transformata prosta i odwrotna ELT.

Przeprowadź kwantyzację sygnału audio w 16 podpasmach ELT z alokacją bitów. Ustaw 256, 512, 1024 i 2048 bitów na ramkę ELT (licząc 256 próbek transformaty przy długości okna czasowego 1024 próbki). Jaka jest zależność SNR w dB od rozdzielczości b w bitach na próbkę?