

Sprawozdanie 1

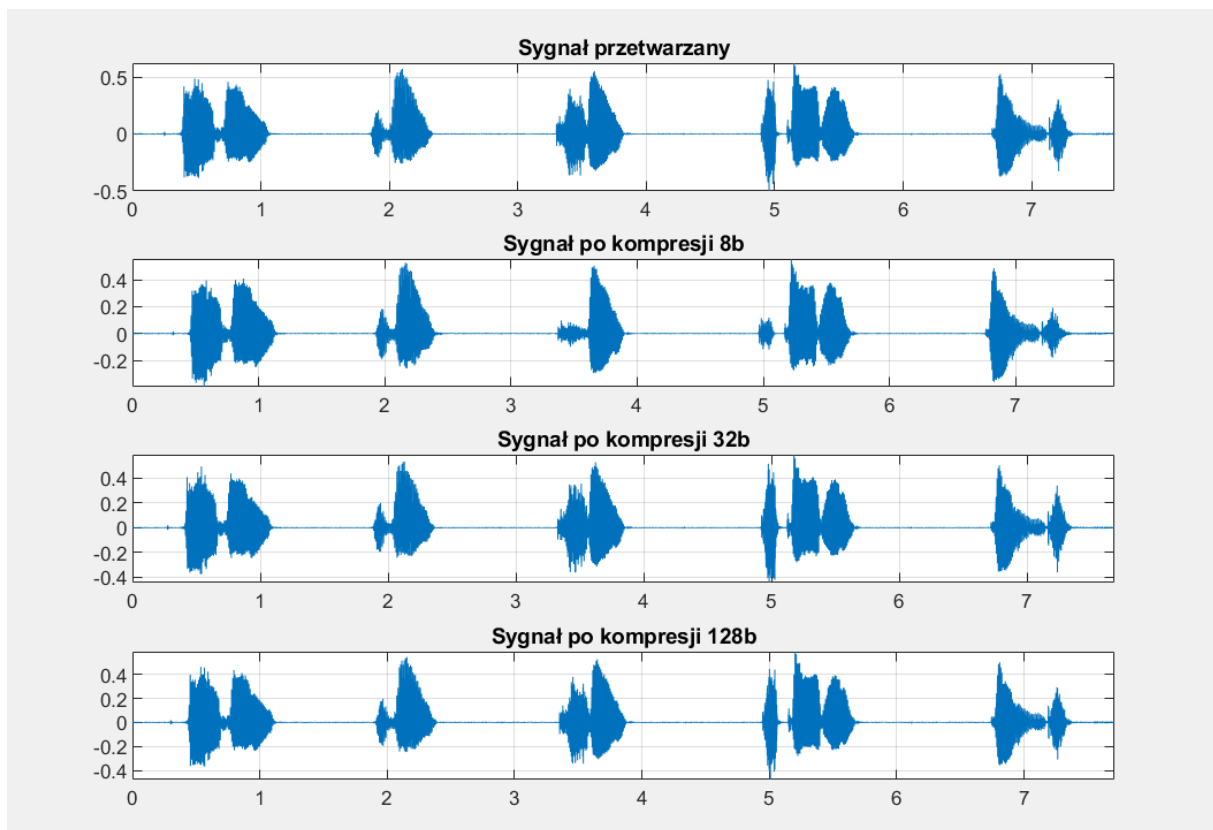
Kompresja dźwięku

1. Kodowanie MP3

Program do zadania znajduje się w pliku lab1.m

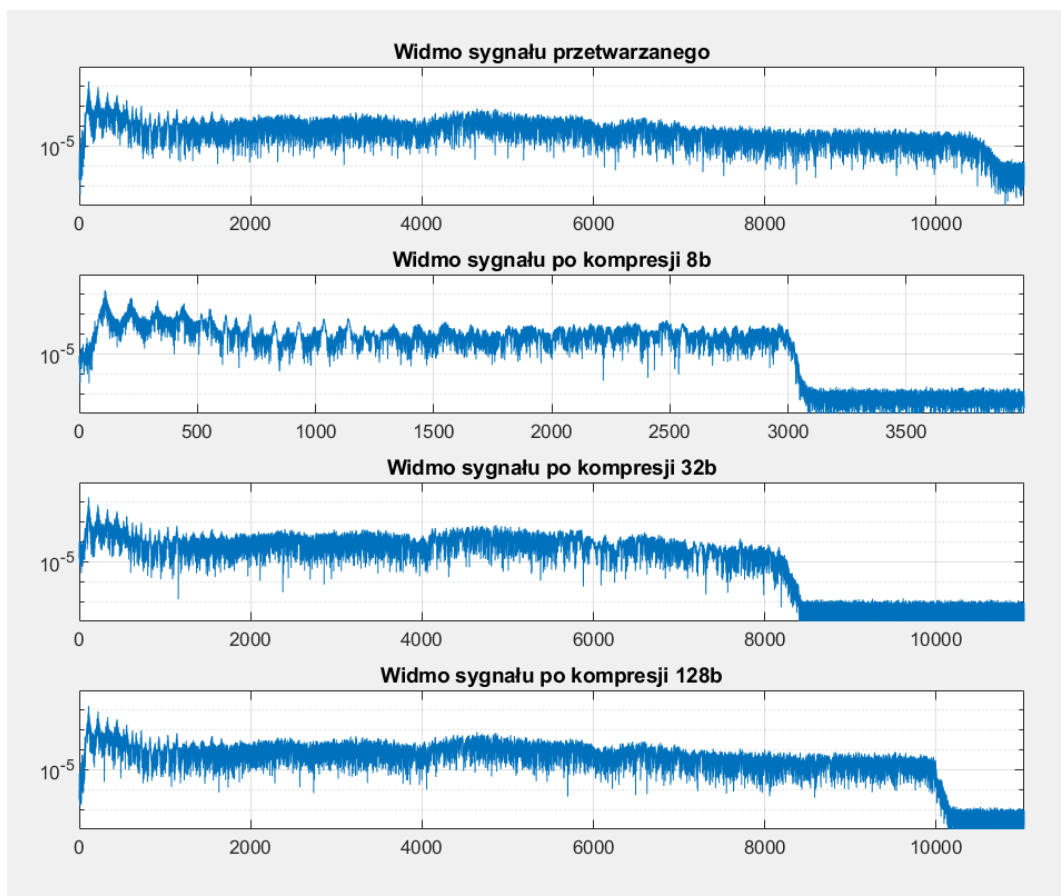
Wpływ stopnia kompresji CBR na nagranie mowy

Został wykorzystany do tego plik 123.wav. Dla kompresji 128b i 32b nie widać na wykresach znaczącej różnicy w porównaniu z sygnałem przetwarzanym. Po odsłuchaniu nagrań w skompresowanych plikach 128b i 32b nie słychać praktycznie żadnej różnicy w porównaniu z oryginalnym plikiem. Dla kompresji 8b już widać różnicę na wykresie przy wyrazach „trzy, cztery, pięć”, zostały usunięte wysokie częstotliwości. Po odsłuchaniu nagranie jest znacząco gorsze, dźwięk jest przytłumiony ale nadal jesteśmy w stanie zrozumieć przekazywaną treść

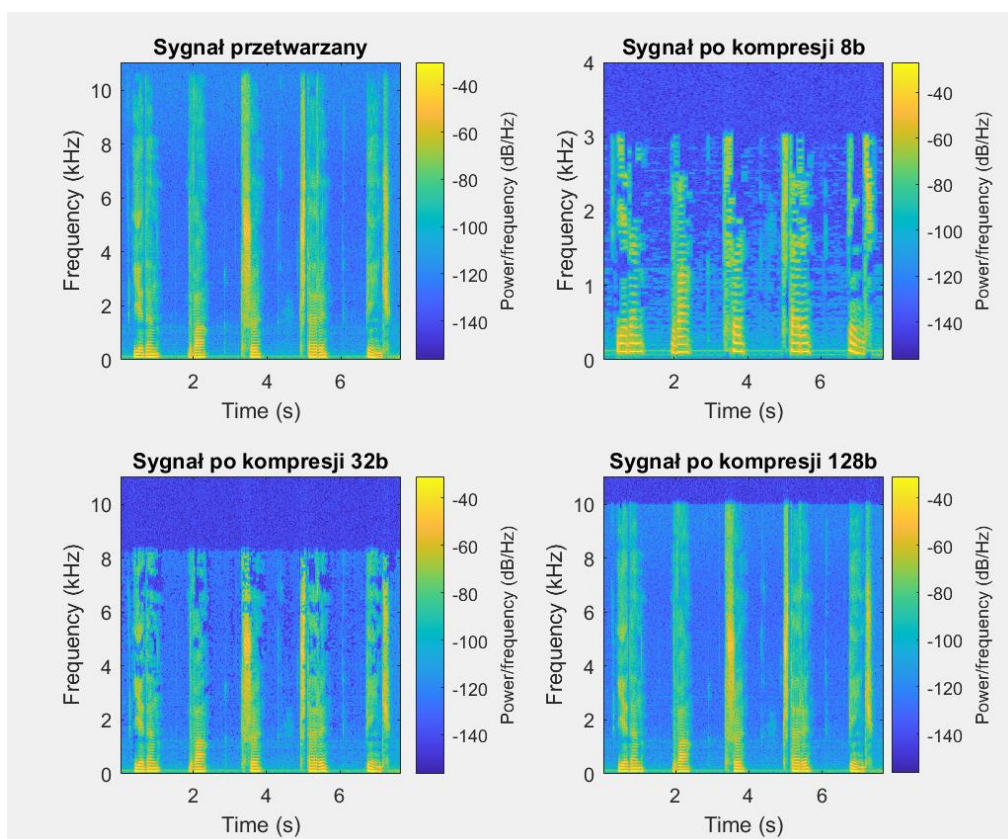


Na widmach sygnału można zauważyć:

- Dla kompresji 128b - zastosował zastosowany filtr dolnoprzepustowy dla ok. 10,1 kHz
- Dla kompresji 32b – zastosował zastosowany filtr dolnoprzepustowy dla ok. 8,2 kHz
- Dla kompresji 8b - zastosował zastosowany filtr dolnoprzepustowy dla ok. 3,2 kHz



Na spektrogramach można zauważyć że kompresja 128b nieznacznie odbiega, natomiast kompresja 32 b i 8 b są już mocno zniekształcone.

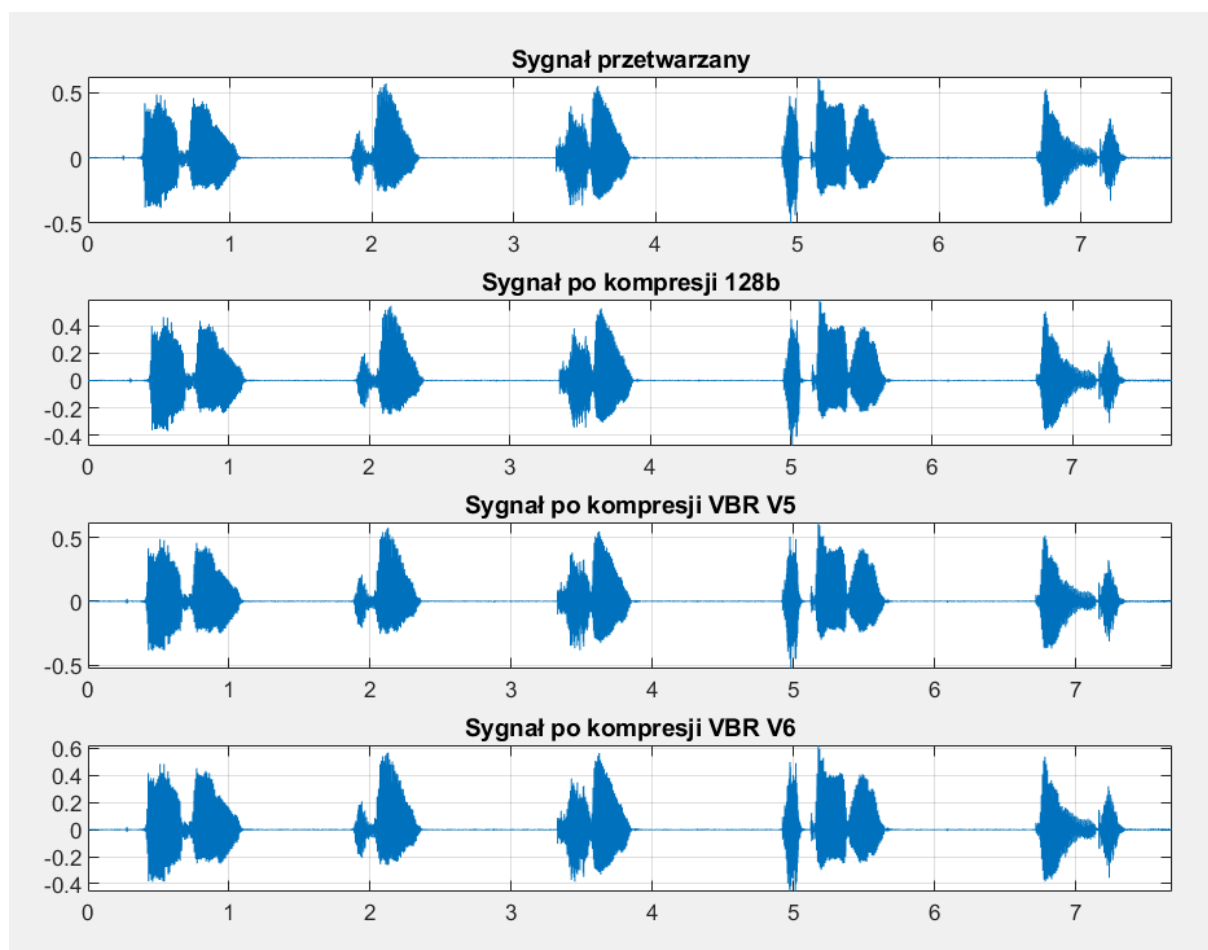


Nazwa pliku	Kompresja	Rozmiar
123_128b.mp3	CBR 128b	120KB
123_32b.mp3	CBR 32b	30,1KB
123_8b.mp3	CBR 8b	7,66KB

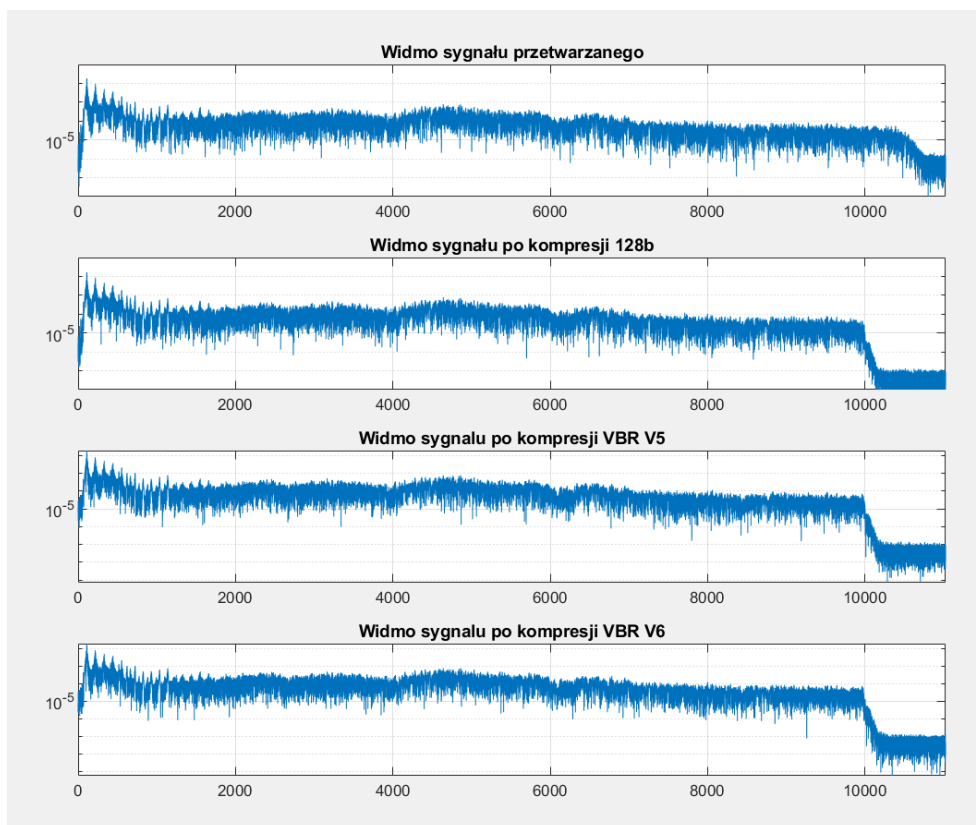
Porównanie kompresji 128kb/s z VBR -V5 i -V6 na nagranie mowy

Program do zadania znajduje się w pliku lab12.m

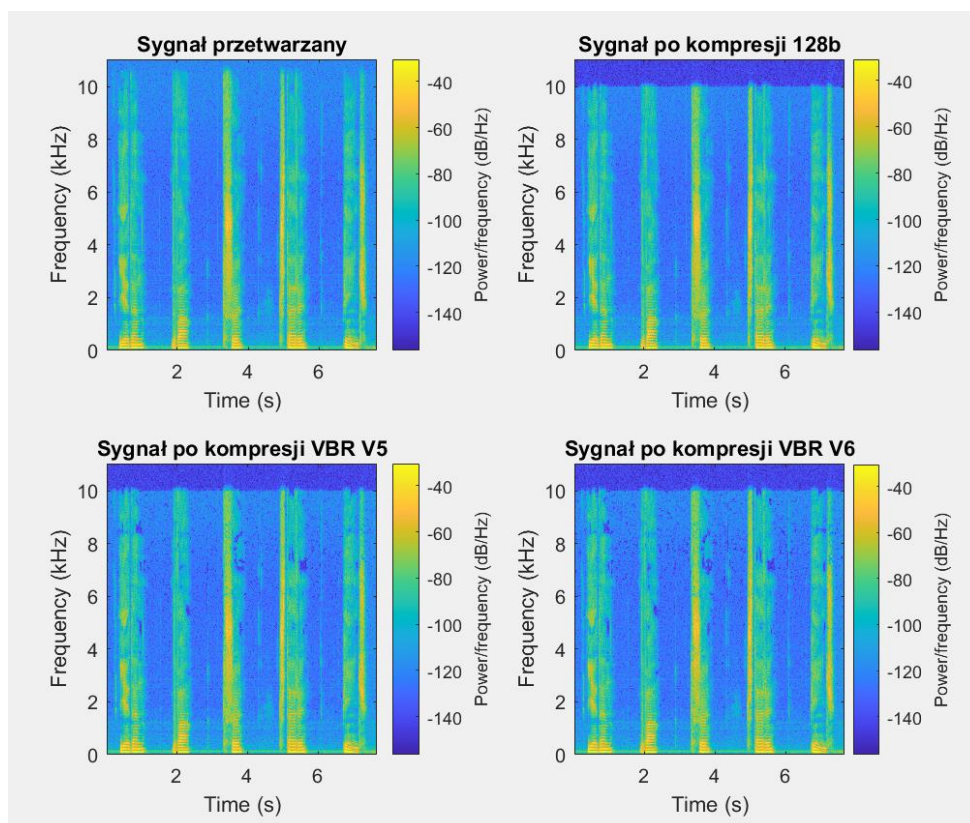
Dla wszystkich kompresji nie widać na wykresach znaczącej różnicy w porównaniu z sygnałem przetwarzanym. Po odsłuchaniu wszystkich plików nie słychać znaczącej różnicy.



Na widmach sygnału widać że zostały zachowane podobne częstotliwości. Dla kompresji 128b oraz kompresji VBR zastosowały zastosowane filtry dolnoprzepustowe dla ok. 10,1 kHz.



Na spektrogramy dla kompresji 128 b i VBR są do siebie podobne i nie odbiegają znacząco do sygnału przetwarzanego.

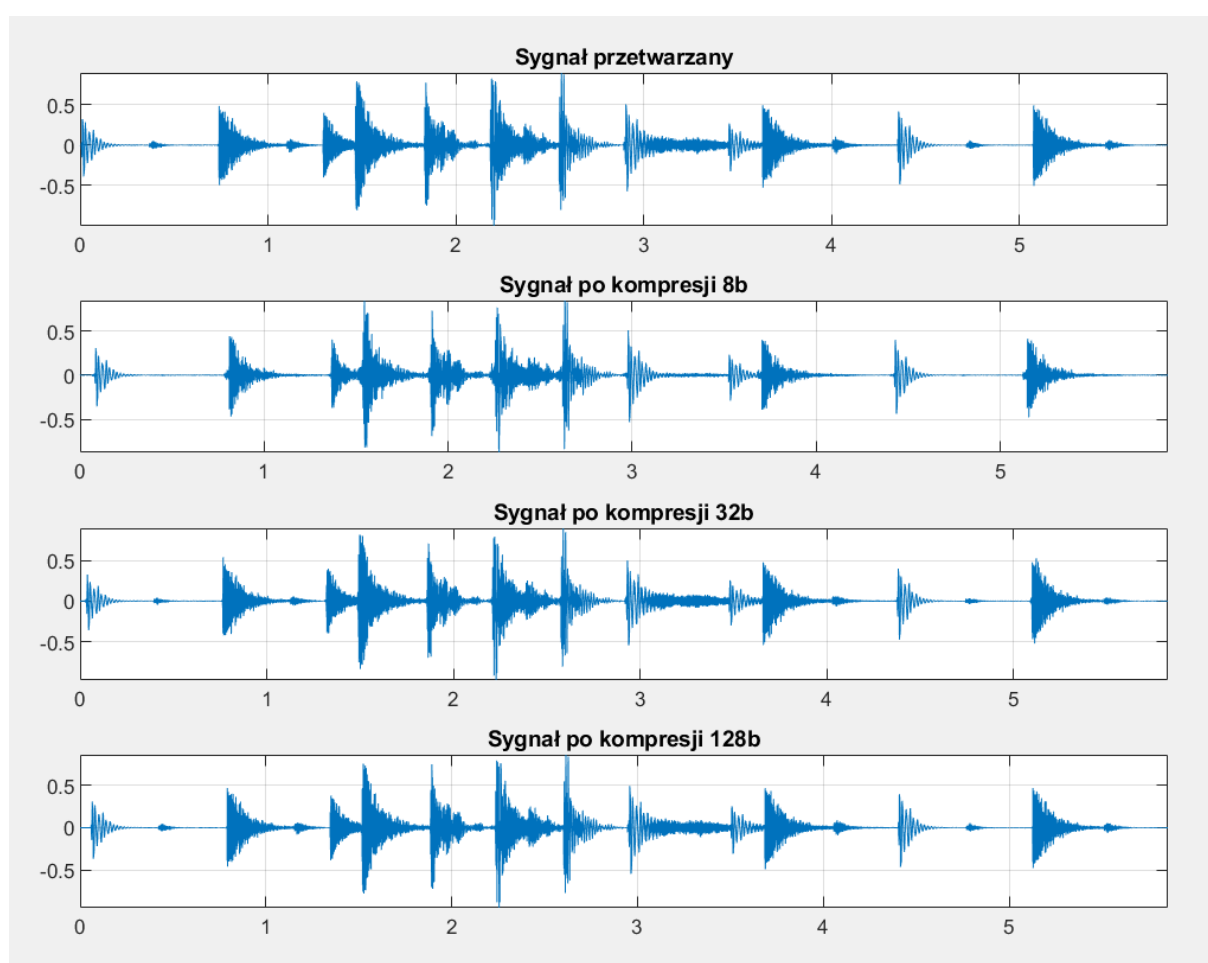


Nazwa pliku	Kompresja	Rozmiar
123_128b.mp3	CBR 128b	120KB
123_v5.mp3	VBR 5	52,5KB
123_v6.mp3	VBR 6	49,1KB

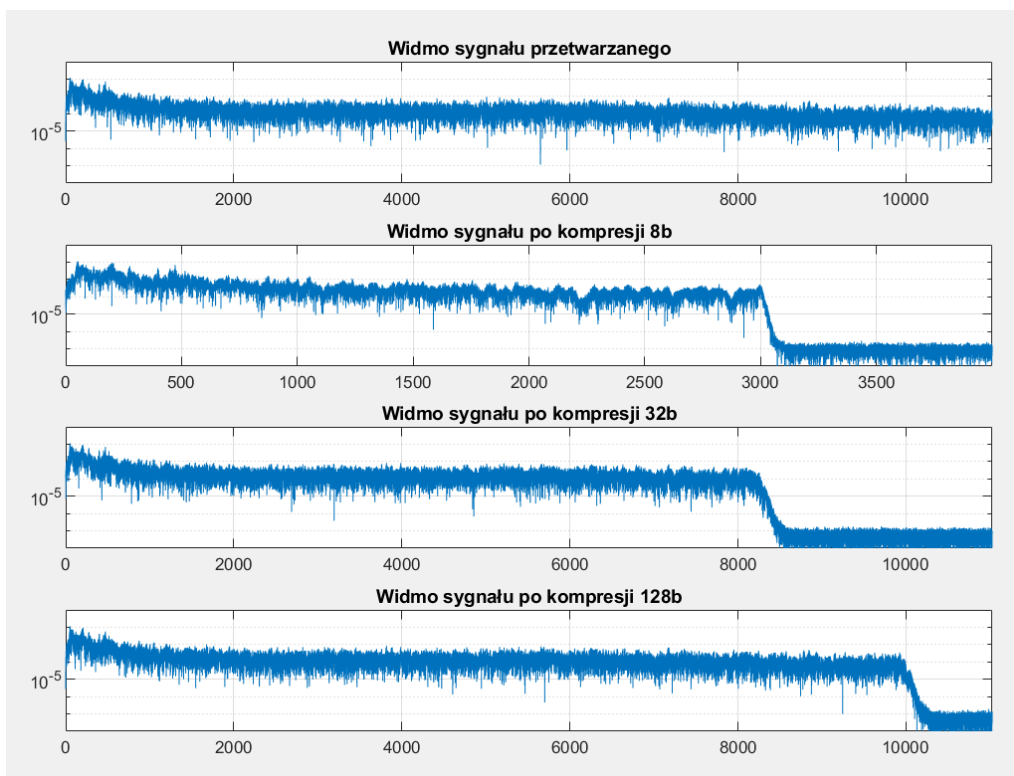
Wpływ stopnia kompresji CBR na nagranie muzyki

Program do zadania znajduje się w pliku lab1.m, zostały zmienione jedynie ścieżki do plików.

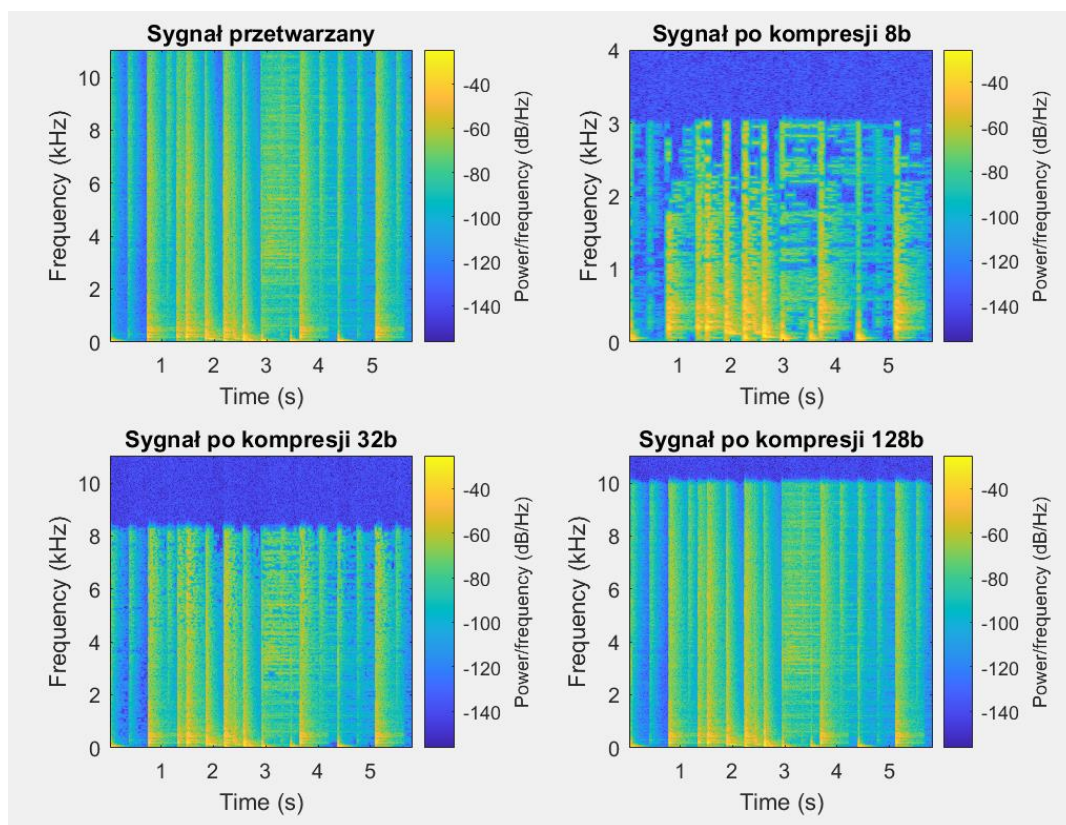
Dla kompresji 128b i 32b nie widać na wykresach znaczącej różnicy w porównaniu z sygnałem przetwarzanym, natomiast dla kompresji 8b już widać różnicę na wykresie. Po odsłuchaniu nagrania z kompresją 128b dźwięk nieznacznie obiega od oryginału, dla kompresji 32b dźwięk już jest wyraźnie przytłumiony a dla kompresji 8b dźwięk już jest bardzo przytłumiony i nagranie nie nadaje się do odsłuchania muzyki.



Na widmach sygnału można zauważyć że zostały zastosowane takie same filtry jak w przypadku nagrania z mową.



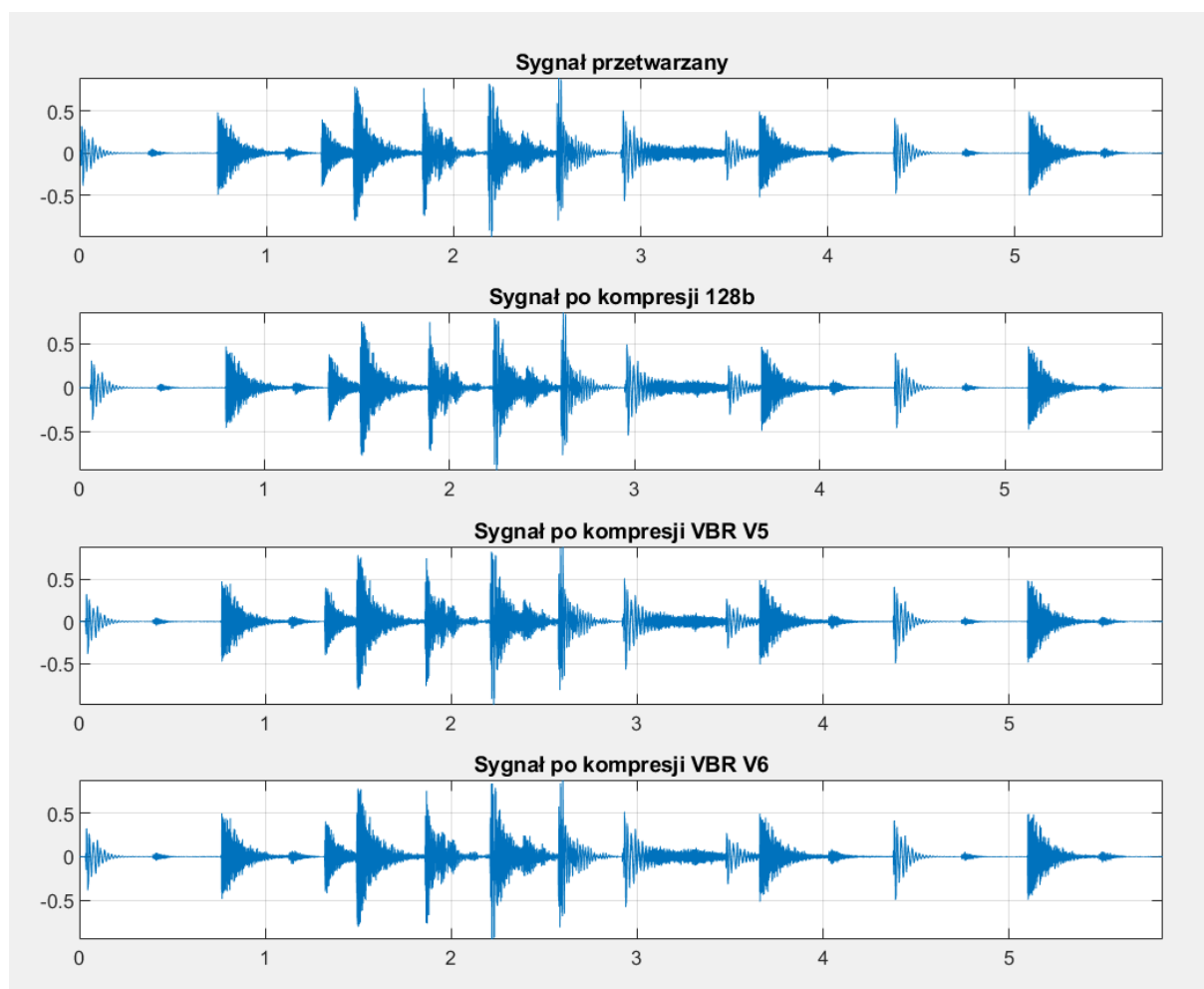
Można zauważyć że spektrogramy zostały podobnie zniekształcone jak w przypadku nagrania z mową.



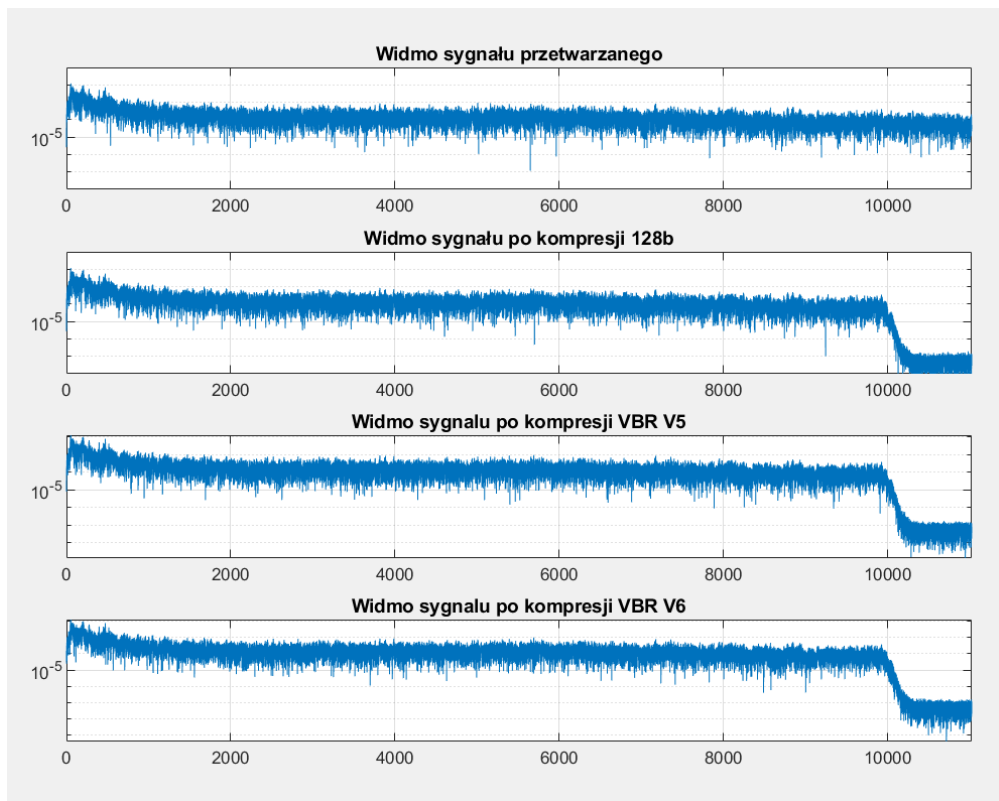
Nazwa pliku	Kompresja	Rozmiar
muzyka_128b.mp3	CBR 128b	91,8KB
muzyka_32b.mp3	CBR 32b	22,8KB
muzyka_8b.mp3	CBR 8b	5,83KB

Porównanie kompresji 128kb/s z VBR -V5 i -V6 na nagranie muzyki

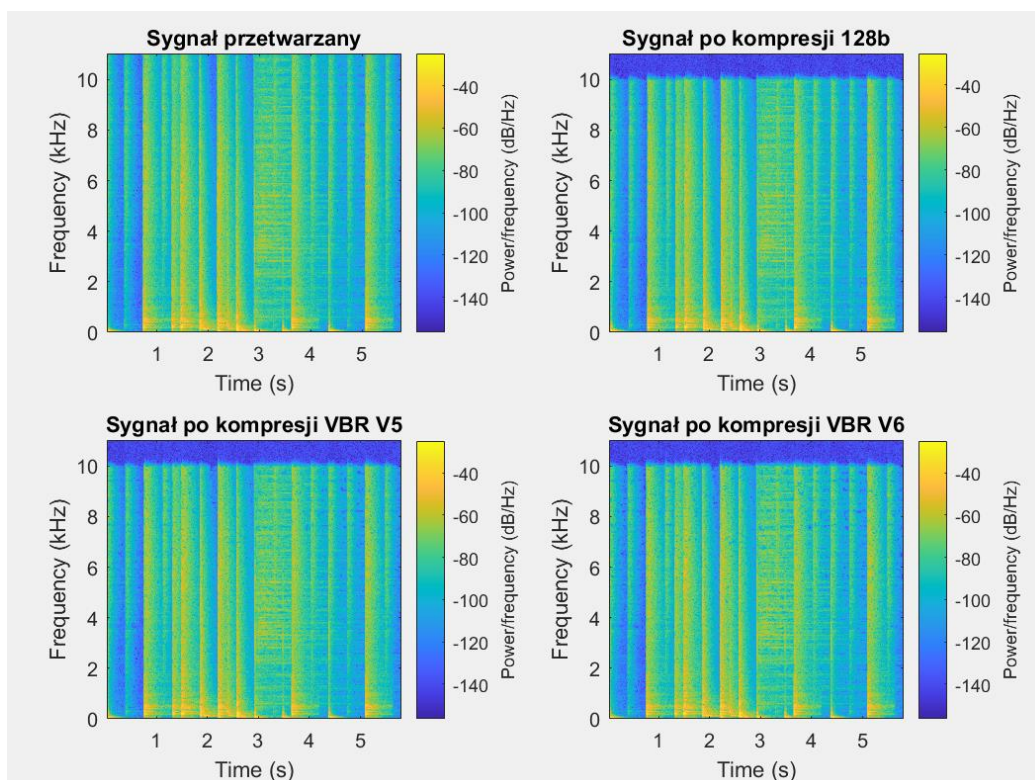
Dla kompresji 128b i VBR nie widać na wykresach znaczącej różnicy w porównaniu z sygnałem przetwarzanym. Po odsłuchaniu nagrań z kompresją dźwięk praktycznie nie różni się od oryginału.



Można zauważyć że zostały zastosowane takie same filtry jak w przypadku nagrania z mową.



Można zauważyć że spektrogramy zostały podobnie zniekształcone jak w przypadku nagrania z mową. Największe częstotliwości zostały wycięte, jednak nie wpływa na odbiór gdyż maksymalne częstotliwości jakie słyszy człowiek to 20kHz



Nazwa pliku	Kompresja	Rozmiar
muzyka_128b.mp3	CBR 128b	91,8KB
muzyka_v5.mp3	VBR 5	39,3KB
muzyka_v6.mp3	VBR 6	36,7KB

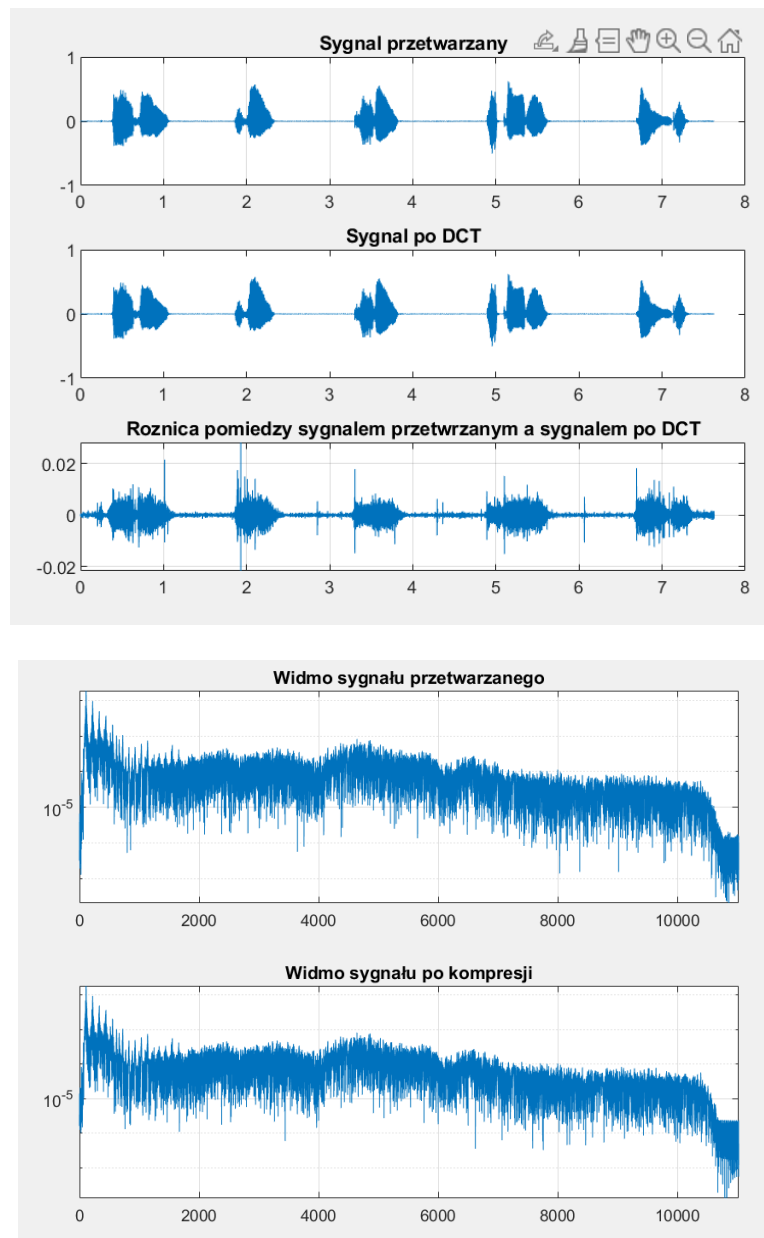
Wnioski

Łatwiej jest skompresować mowę niż muzykę, dlatego że nie zależy nam na jakości nagrania a na przekazaniu treści. Do kompresji lepiej się sprawdza tryb VBR, pliki są znacząco mniejsze oraz utrzymuje stałą jakość sygnału wyjściowego.

2. Kompresja z transformacją DCT

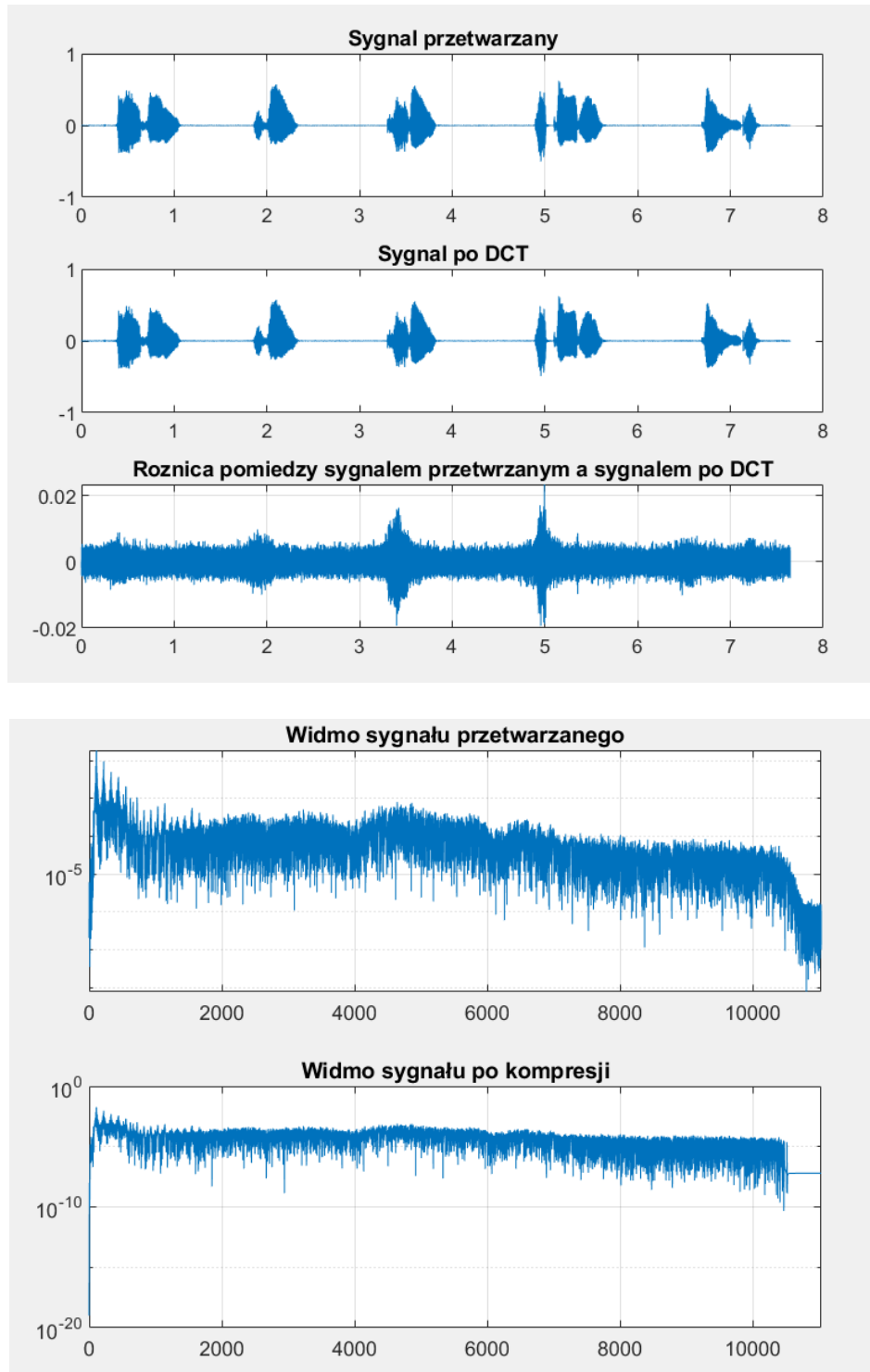
Użyte programy znajdują się w plikach `dct_dzwiek_ramki.m` i `dct_dzwiek_calosc.m`

Analiza pliku podzielonego na ramki z mową

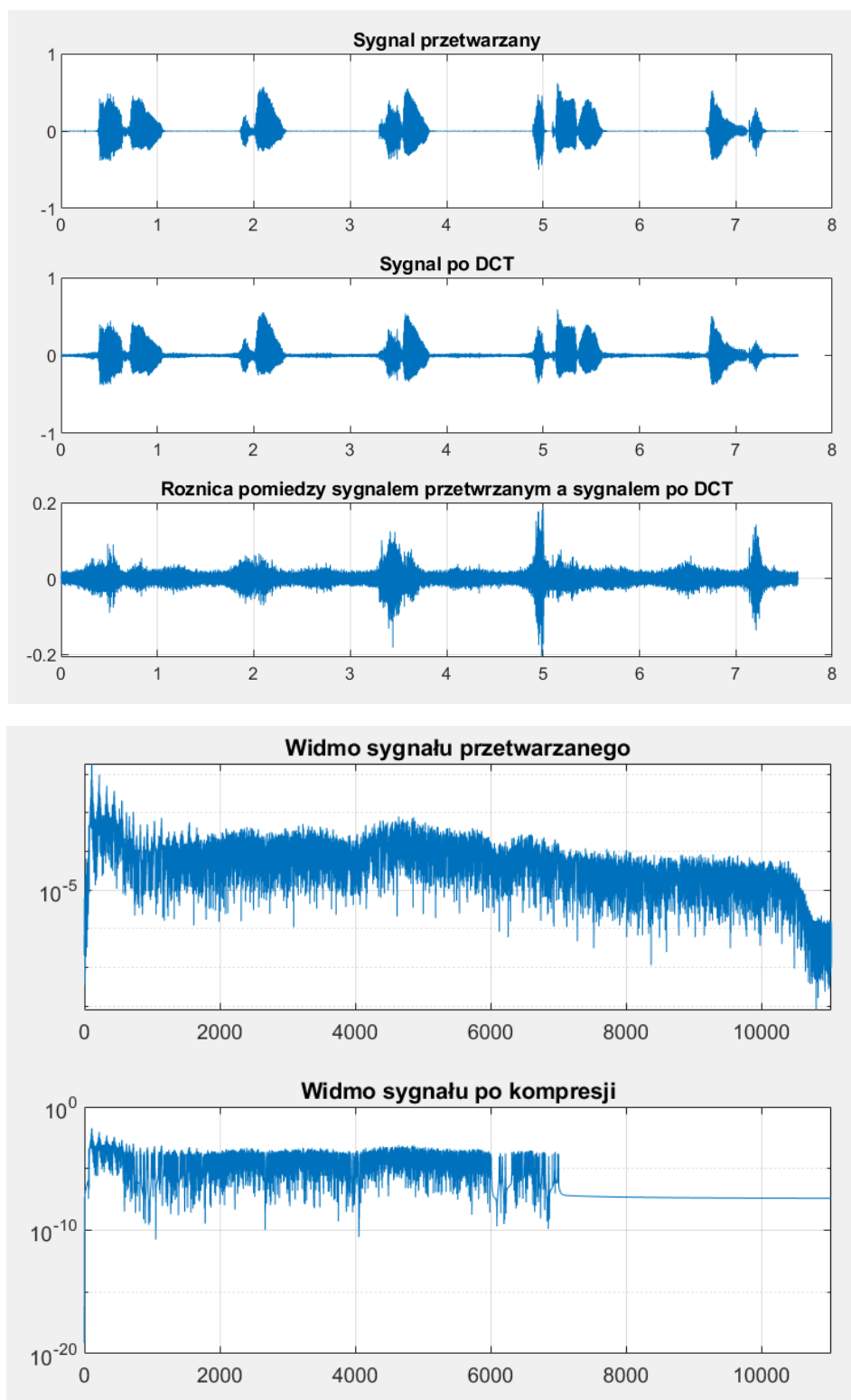


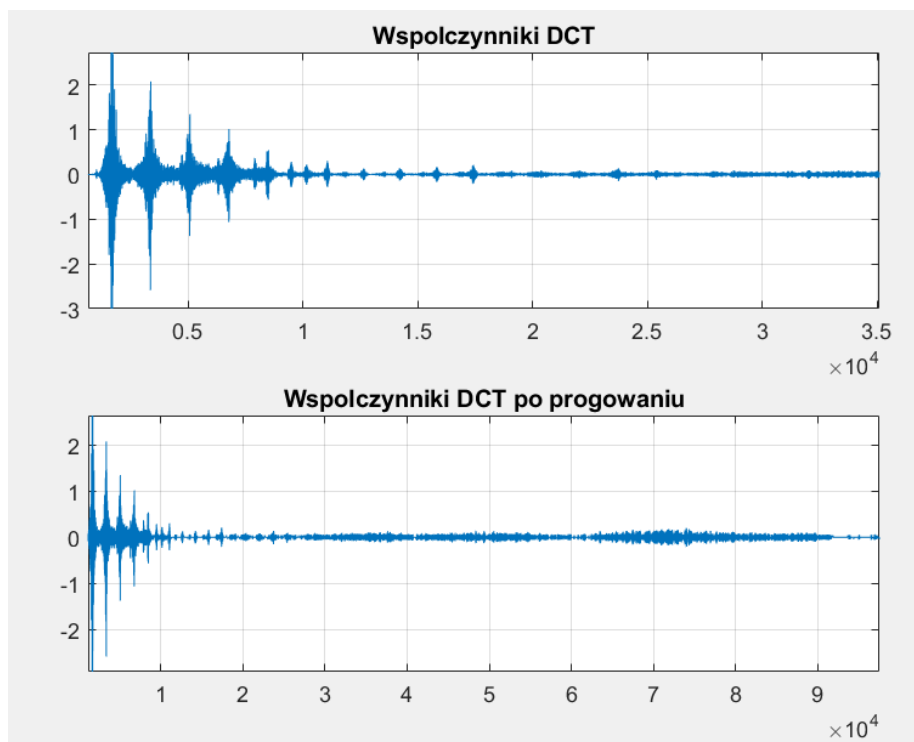
Dla wartości progowania 0.0065 wyzerowaniu uległo 85.8 procent współczynników DCT. Można zauważyć ze wykresach że największe różnice są dla słów „trzy, cztery pięć”, po odsłuchaniu nie słychać różnic z plikiem pierwotnym.

Analiza całego pliku z mową

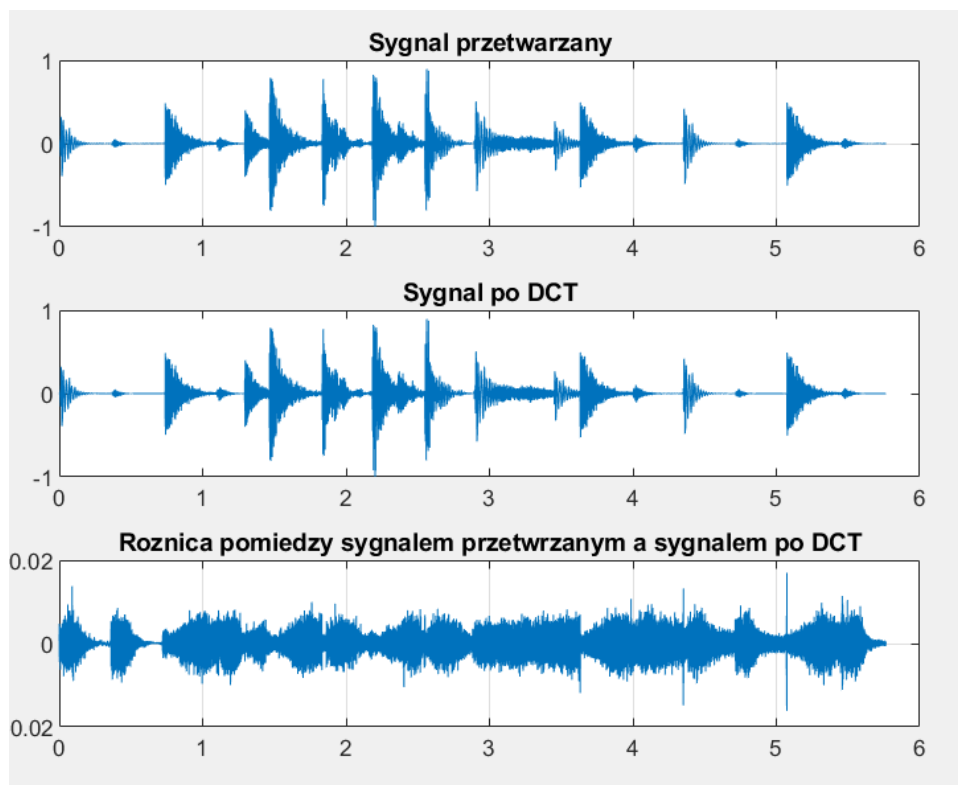


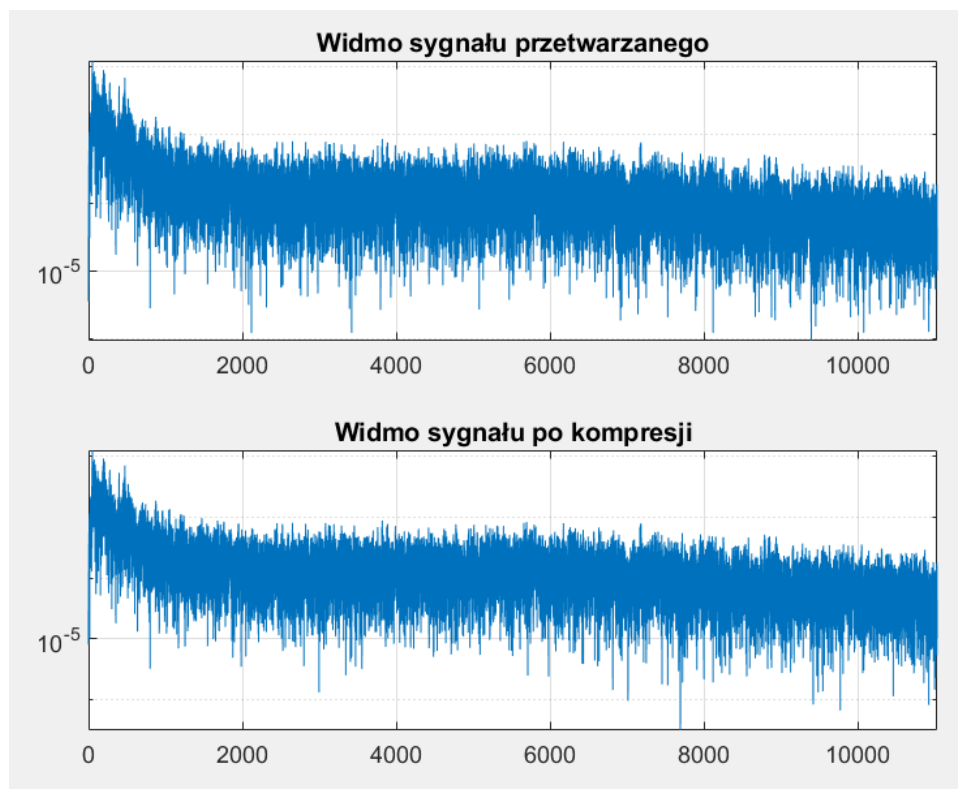
Współczynnik kompresji dla progowania 0.0065 wyniósł 38,4%. Następnie znalazłam współczynnik progowania 0.039 dla którego kompresja wynosi 85%. Na przybliżeniu widać wyraźnie różnicę pomiędzy współczynnikami. Po odsłuchaniu plik jest wyraźnie gorszej jakości, słychać na nim szum.



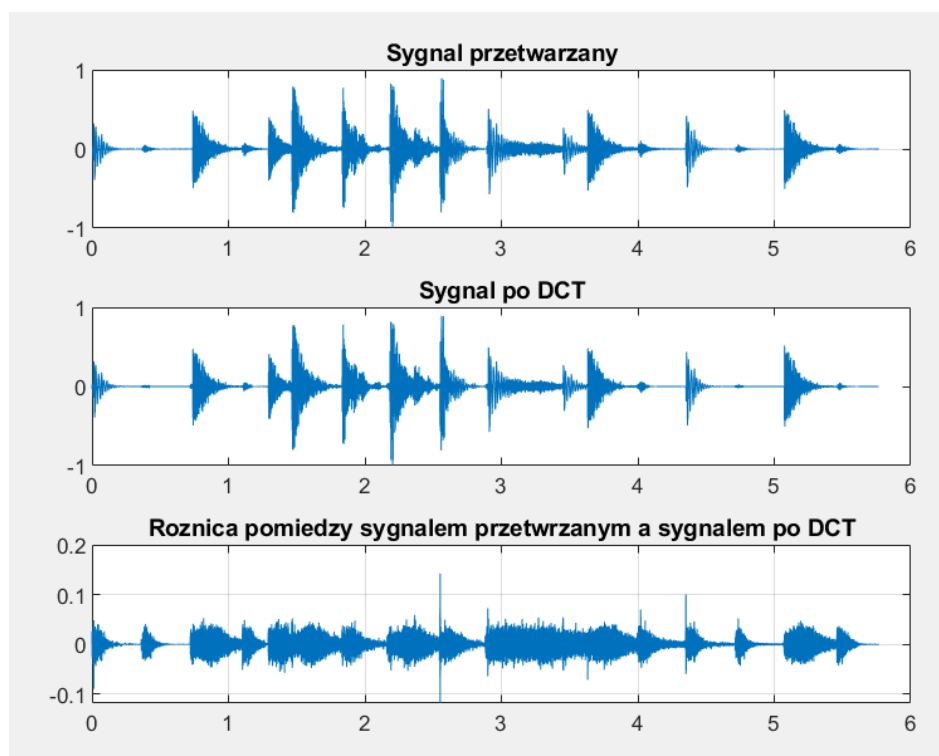


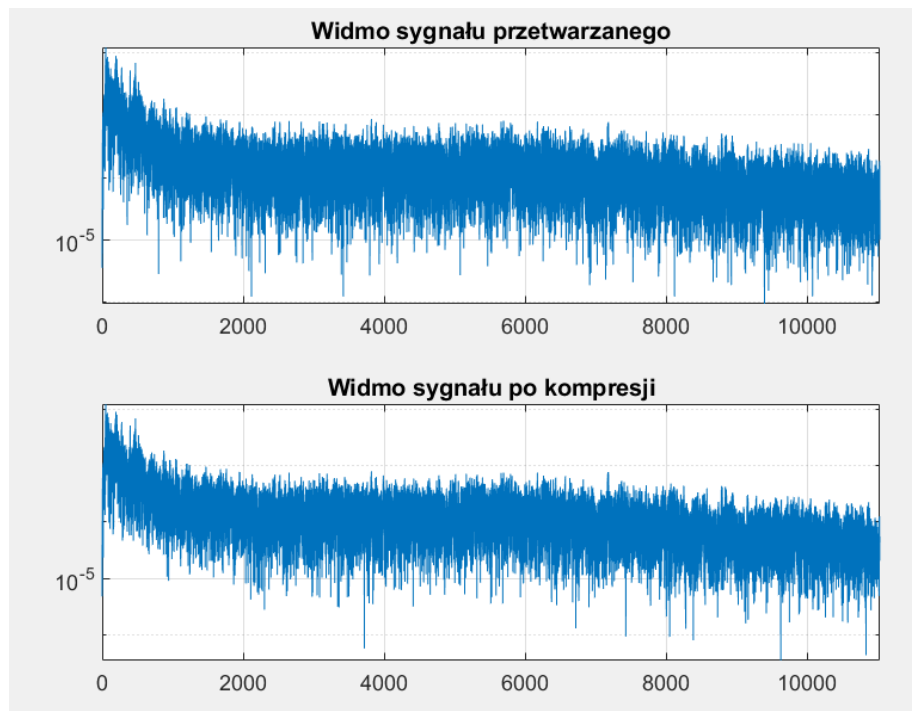
Analiza pliku podzielonego na ramki z muzyka



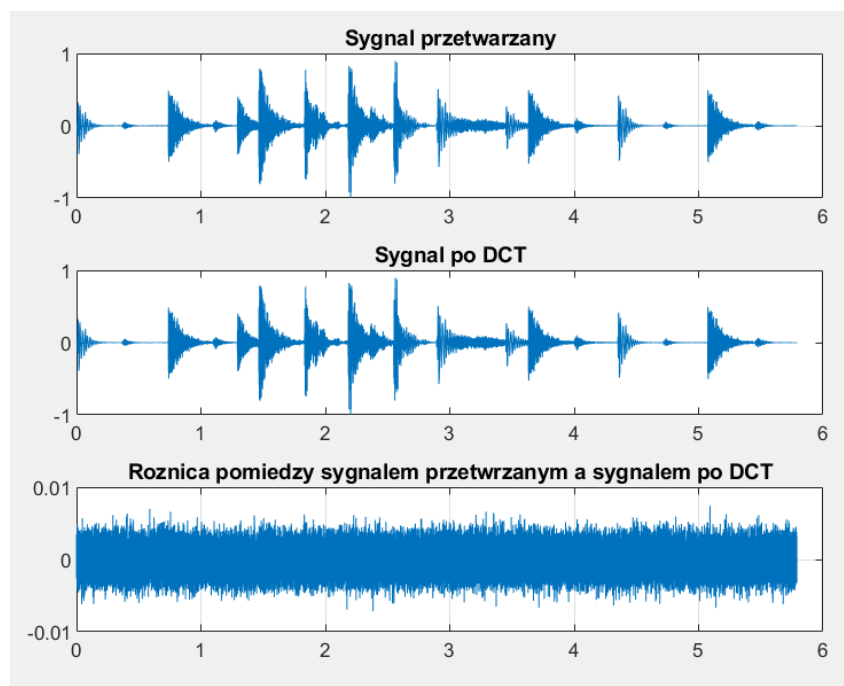


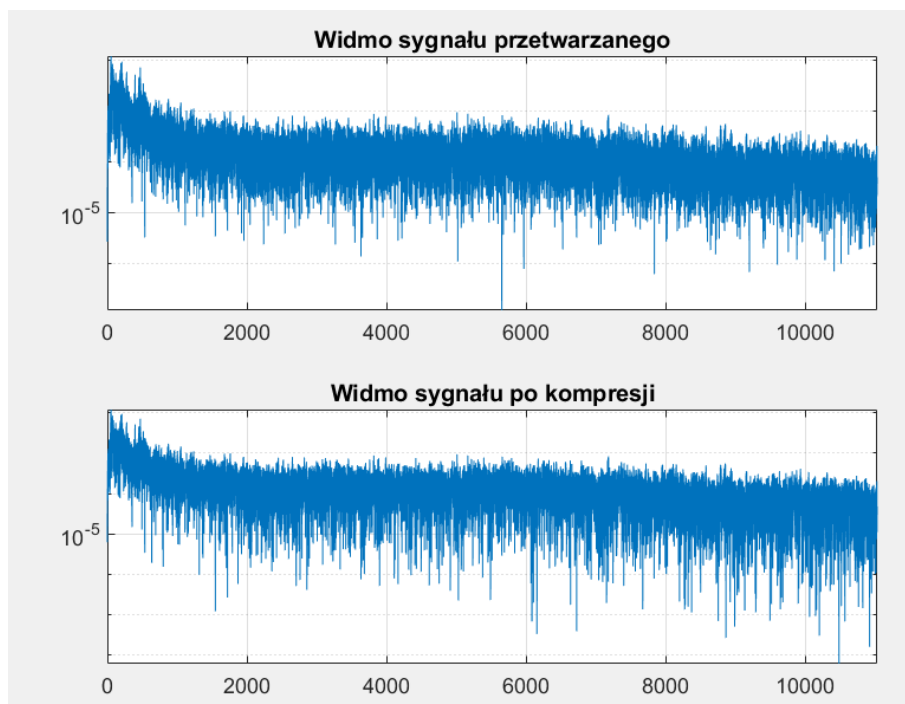
Przy progowaniu 0.0065 wyzerowaniu uległo 57.4 procent współczynników DCT. Następnie została znaleziona wartość progowania 0.034 dla której kompresja wynosi 85%.



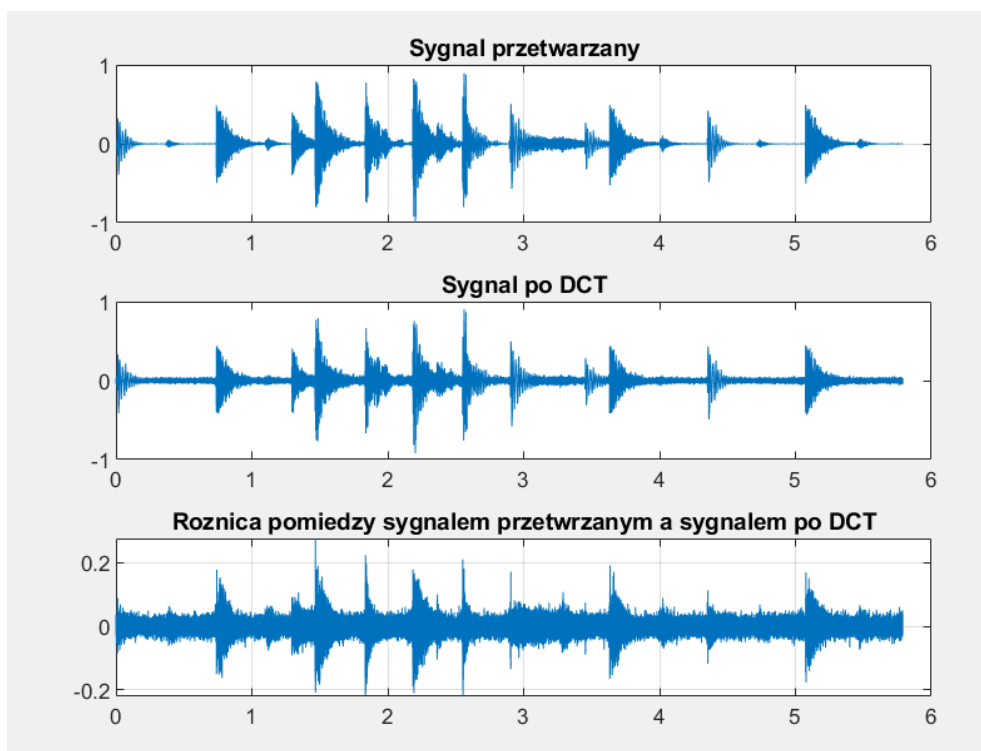


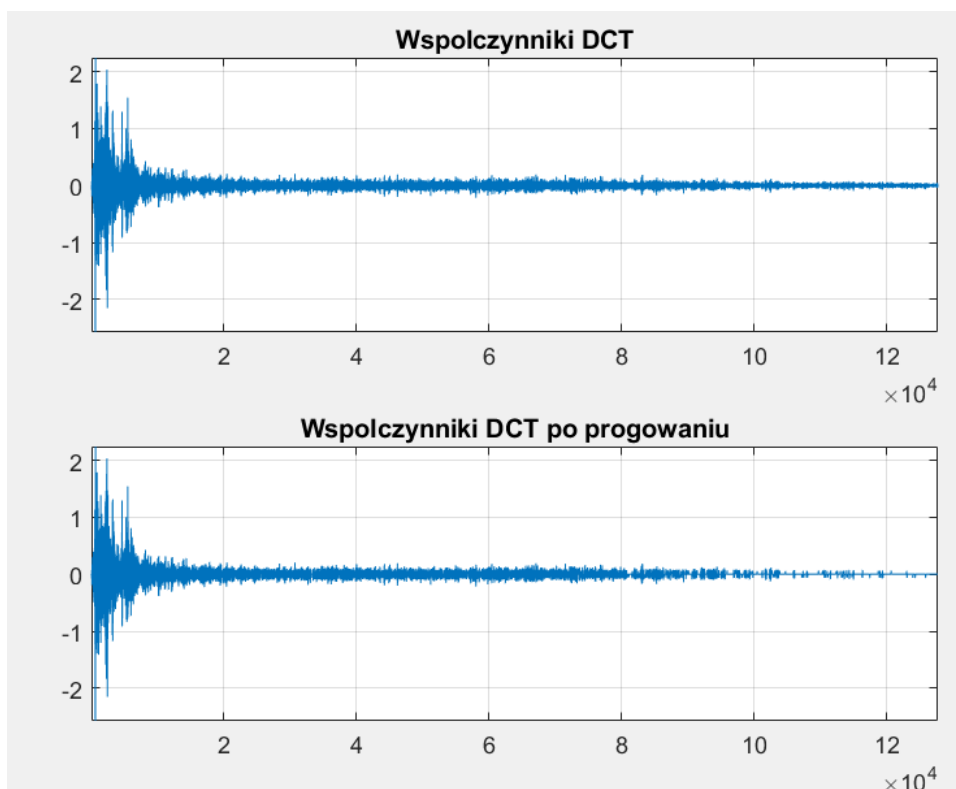
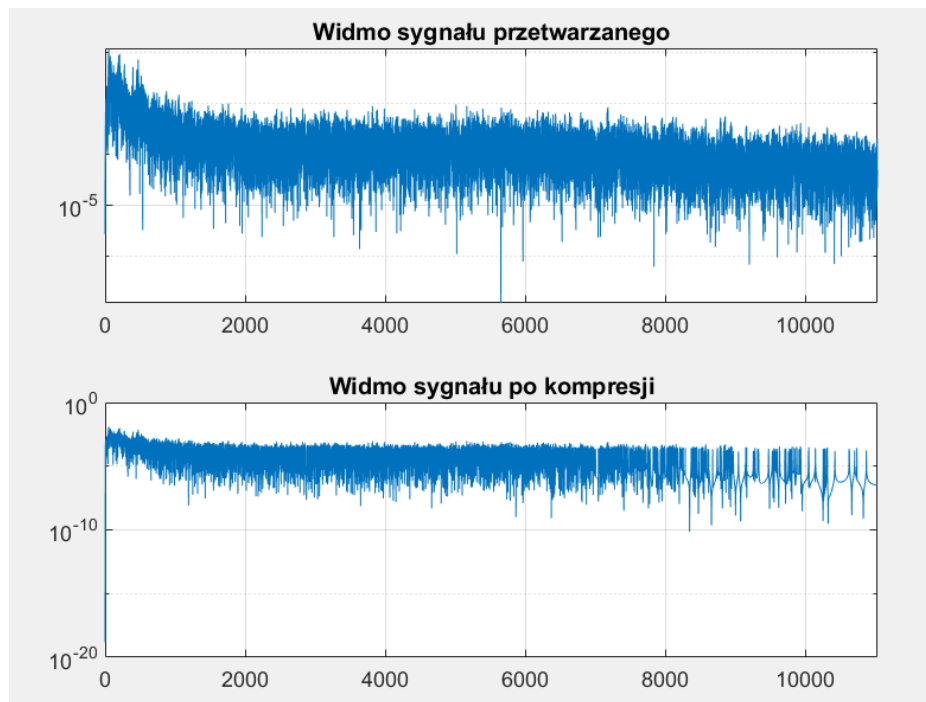
Analiza całego pliku z muzyką





Powyżej są wykresy dla progowania 0.0065. Następnie została znaleziona wartość progowania 0.055 dla której kompresja wynosi 85%. Widać znaczące różnice na wykresie, po odsłuchaniu plik jest wyraźnie zaszumiony i przytłumiony.





Wnioski

DCT z podziałem na ramki daje lepsze rezultaty oraz są mniejsze wartości na wykresie różnicy dla dźwięku i mowy. DCT daje dobre rezultaty dla głosu natomiast dla dźwięku są one akceptowalne.