

Data oddania: _____

Ocena: _____

Paweł Musiał 178726
Łukasz Michalski 178724

Zadanie 3:

Analiza częstotliwości podstawowej dźwięku.*

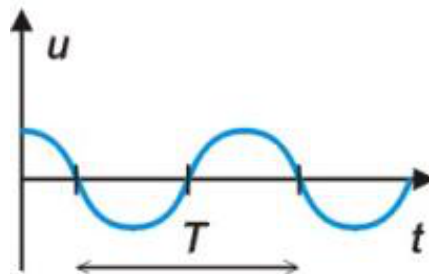
1. Cel

Celem zadania było stworzenie aplikacji realizującej poniższe element :

- wyznaczenie częstotliwości podstawowej dźwięku przy pomocy metod :
 - metody w dziedzinie czasu : metoda AMDF ²
 - metody z dziedzinie częstotliwości : analiza cepstralna
- odtworzenie sekwencji dźwięków na podstawie uzyskanych częstotliwości podstawowych.

2. Wprowadzenie

Dźwięk możemy spostrzegać na dwa sposoby : subiektywny i obiektywny. W ujęciu obiektywnym rozpatrywany jest jako fala mechaniczna w ośrodku sprężystym, której parametrami są amplituda i częstotliwość, możemy słyszany dźwięk opisać równaniem $A\sin(t)$.



Rysunek 1: Fala dźwiękowa.

Opisując nasze subiektywne odczucie dźwięku opisujemy go parametrami takimi jak : wysokość(zależy od częstotliwości), głośność(zależy od natężenia),

* SVN: https://serce.ics.p.lodz.pl/svn/labs/poid/at_sr0830/lmpm@!!NR_REWIZJI!

!

² ang. average magnitude difference function

barwa(zależy od widma fali akustycznej). Dźwięk w najprostszym przypadku może być pojedynczą sinusoidą(ton) o ustalonych parametrach amplitudy i częstotliwości, jednak w naturze mamy do czynienia z reguły w dźwiękami złożonymi(wielotonami). Wielotonem nazywamy sygnał będącym złożeniem sygnałów sinusoidalnych, mogą być one harmoniczne, którego składowe są krotnościami tonu podstawowego oraz nieharmoniczne gdzie taka zależność nie zachodzi, sygnały te są niezależne - w ludzkiej percepcji postrzegany jako szum.

Analizę dźwięku możemy przeprowadzać w dziedzinie czasu, sygnałe takim jaki bezpośrednio potrafimy zarejestrować np. mikrofonem. Po dyskretyzacji sygnału do postaci cyfrowej dysponujemy ciągiem reprezentującym w pewnym przybliżeniu przebieg, w najprostszym ujęciu - sinusoidy(pojedynczego tonu). Posiadając tą wiedzę możemy zastosować szereg metod pozwalających na określenie częstotliwości dźwięku. Znacznym utrudnieniem dla tych metod będzie istnienie harmonicznych które mogą być mylące.

Innym podejściem, równie bardzo naturalnym jest transformacja sygnału z dziedziny czasu do dziedziny częstotliwości. Dzięki tej transformacji mamy bezpośrednio informację o częstotliwościach bazowych przetwarzanego sygnału. Jednak minusem badania sygnału w dziedzinie częstotliwości jest jego wrażliwość na szum, dlatego też dobrze poddać wstępnej filtracji z celu wyeliminowania szumu bądź przynajmniej jego zminimalizowaniu.

2.1. Metoda w dziedzinie czasu : AMDF

Metoda ta bazuje na porównaniu sygnału oryginalnego z opóźnionym.

$$d(m) = \sum_{n=0}^{N-1} \|x(n) - x(n+m)\| \quad (1)$$

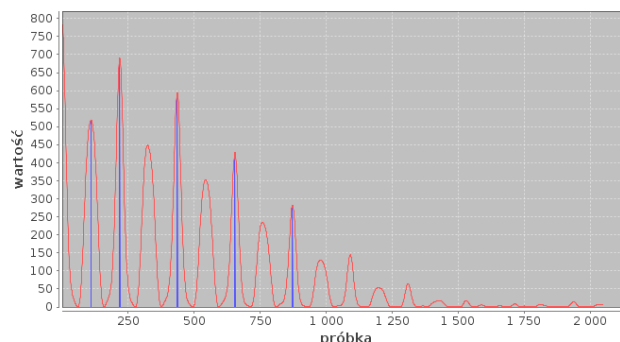
Położenie pierwszego minimum funkcji $d(m)$ dla $m \neq 0$ wyznacza długość okresu najniższej składowej częstotliwościowej sygnału.

W najprostszym przypadku analizy pojedynczego tonu, metoda ta będzie sprawdzać się bezbłędnie, ponieważ porównując sygnał przesunięty z oryginalnym w momencie gdy wartość będzie równa 0 lub jej bardzo bliska, oznaczać to będzie jednakowość tych sygnałów lub znaczny stopień podobieństwa. Dlatego pierwsze minimum oznaczać będzie okres sinusoidy podstawowej sygnału. Pierwszym problemem może być szum, którego występowanie skutkować będzie spłyceciem minimów funkcji $d(m)$, jednak nadal będą one jednoznacznie określały okres częstotliwości dźwięku. W przypadku wielotonu spodziewać się możemy niejednoznaczności minimów, ponieważ składowe harmoniczne mogą się „znosić”, lokalnie powodując minima - fałszywe, funkcji $d(m)$.

2.2. Metoda w dziedzinie częstotliwości : analiza cepstralna

Analiza cepstralna opiera się na obliczeniu cepstrum sygnału wejściowego. Czyli widma amplitudowego, widma amplitudowego sygnału wejściowego.

$$C(x) = \|FFT(\ln(\|FFT(x)\|))\| \quad (2)$$



Rysunek 2: Przykładowy wykres cepstrum.

Jak wspomniano na wstępie metoda ta jest wrażliwa na szum, wynika to z charakteru wykonywanej transformacji. Również jak w przypadku analizy dźwięku w dziedzinie czasu, składowe harmoniczne wielotonu będą powodować problemy w rozpoznawaniu częstotliwości podstawowej dźwięku. Jak widać na powyższym rysunku. Obok maksimów będących kolejnymi wielokrotnościami częstotliwości bazowej (maksymalny skok) występują pośrednie skoki, które w newralgicznych przypadkach mogą mieć wartość wyższą od skoku odpowiadającego częstotliwości bazowej, dlatego też zadanie to również nie zawsze da prawidłowy wynik.

3. Opis implementacji

3.1. Metoda w dziedzinie czasu : AMDF

3.2. Metoda w dziedzinie częstotliwości : analiza cepstralna

Na podstawie analizy sygnałów wejściowych jak i odpowiadających im cepstrum, w drodze uzyskania jak najlepiej sprawdzającej się metody - ostatecznie otrzymaliśmy taki oto algorytm :

1. uśrednienie sygnału w celu zniwelowania wpływu szumu
2. ustawienie rozmiaru okna Hamminga na 4096
3. obliczenie transformaty fouriera aktualnie przetwarzanego fragmentu sygnału dyskretnego i wybranie widma mocy $X = \|FFT(X)\|$
4. obliczenie logarytmu naturalnego kwadratu widma mocy sygnału $X = \log(\|X\| * \|X\|)$
5. obliczenie transformaty fouriera połowy otrzymanego sygnału (widmo jest symetryczne wystarczy jego połowa) i wybranie widma mocy $C(X) = \|FFT(X)\|$
6. obliczenie kwadratu mocy widma $X = C(X) * C(X)$
7. wybranie skoków najwyższych tak otrzymanego sygnału, nie tylko co do wartości, ale również najwyższych pod względem wysokości zbocza
8. wyznaczenie okresu dźwięku podstawowego na podstawie odległości pomiędzy dwoma kolejnymi najwyższymi skokami

Ostatnie kroki są poprawką na podstawową wersję algorytmu niwelującą pewne niepożądane przypadki, tzn niezgodność maksymalnego lokalizacji maksymalnego skoku z oczekiwaną lokalizacją częstotliwości bazowej.

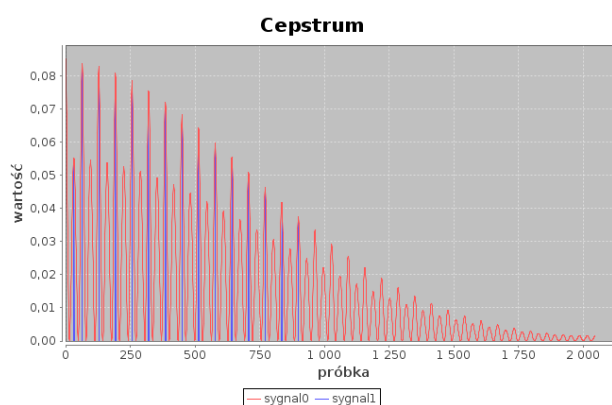
4. Wyniki

4.1. Metoda w dziedzinie czasu : AMDF

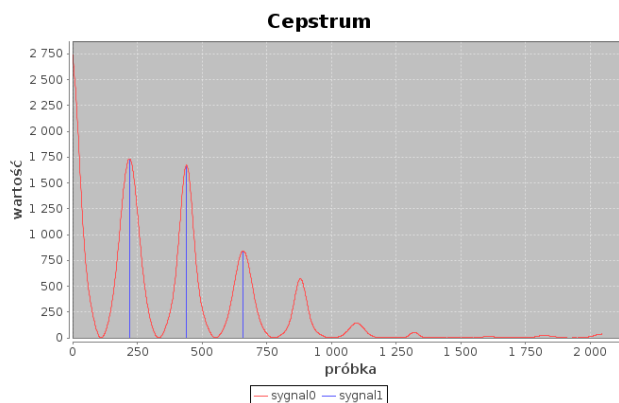
4.2. Metoda w dziedzinie częstotliwości : analiza cepstralna

| lokalizacja przewidywana | lokalizacja wyznaczona | częstotliwość przewidywana[Hz] | częstotliwość wyznaczona[Hz] |
|--------------------------|------------------------|--------------------------------|------------------------------|
| 96 | 97 | 455,00 | 454,64 |
| 43 | 43 | 1025,00 | 1025,58 |
| 64 | 64 | 683,00 | 689,06 |
| 28 | 29 | 1537,00 | 1520,69 |
| 490 | 490 | 90,00 | 90,00 |
| 38 | 38 | 1139,00 | 1160,53 |
| 196 | 196 | 225,00 | 225,00 |
| 130 | 131 | 337,00 | 336,64 |
| 551 | 276 | 80,00 | 159,78 |
| 48 | 97 | 911,00 | 454,64 |
| 119 | 120 | 369,00 | 367,50 |
| 24 | 25 | 1779,00 | 1764,00 |
| 159 | 81 | 276,00 | 544,44 |
| 49 | 49 | 887,00 | 900,00 |

Tablica 1: Wyznaczanie częstotliwości podstawowej w wybranych dźwiękach przykładowych.



Rysunek 3: Wykres cepstrum dźwięku o częstotliwości 698Hz.



Rysunek 4: Wykres cepstrum dźwięku o częstotliwości 100Hz.

5. Dyskusja

5.1. Metoda w dziedzinie czasu : AMDF

5.2. Metoda w dziedzinie częstotliwości : analiza cepstralna

Jak widać w tabeli 1 metoda oparte o analize cepstralną radzi sobie dobrze z dźwiękami przykładowymi, również zadowalający - subiektywny, wynik otrzymaliśmy analizując sekwencje gry na wiolonczeli, gorzej poradziła sobie niestety z odtwarzaniem sekwencji gry na pianinie. Zauważalny problem który czasami występuje, jest identyfikacja jako częstotliwości podstawowych, wyższych lub niższych harmonicznnych. Problem ten jest widoczny na rysunku 4. Pierwszy skok - maksymalny, w punkcie 220 wypada w połowie prawidłowej lokalizacji czyli 441. W zebranych danych w tabeli widoczny jest przypadek odwrotny czyli zidentyfikowanie częstotliwości dwa razy niższej od oczekiwanej.

6. Wnioski

Obie metody radzą sobie dobrze z powierzonym im zadaniem. Występują jednak przypadki które należałoby dokładniej przeanalizować w celu poprawy wyników odnajdywania częstotliwości podstawowej dźwięku. Napotkane problemy które mają wpływ na wyniki metod to złożoność dźwięku z wielu harmonicznnych oraz szum który wydaje być się zjawiskiem naturalnym i nieodłącznym dla realnych przypadków.

Metoda oparta o analizę w dziedzinie częstotliwości wydaje się być dokładniejszą i mniej podatną na charakterystykę przetwarzanego dźwięku, czyli rodzaj i natężenie szumu i złożoność dźwięku.

Analiza cepstralna mimo, że z założenia powinna dawać wyniki łatwiejsze do interpretacji, to ze względu na zróżnicowanie sygnałów wejściowych nie jest jednak to zadaniem trywialnym. Zauważone i opisane problemy w sekcji dyskusji wskazują na problematyczność w interpretacji cepstrum. Również to, co nie było ujęte w dyskusji, zależność wyników tej metody od wielkości okna analizy. Pomijając przypadki oczywiste gdzie okno jest za małe z stosunku do okresu badanego dźwięku, przekształcenie sygnału do postaci cepstrum daje różne wyniki z zależności od wielkości okna.

Literatura

- [1] Fast Fourier Transform (FFT) <http://www.cmlab.csie.ntu.edu.tw/cml/dsp/training/coding/transform/fft.html>
- [2] The Radix 2 Decimation In Frequency (DIF) Algorithm. <http://www.engineeringproductivitytools.com/stuff/T0001/PT03.HTM>
- [3] !!!JAK COŚ ZNAJDZIESZ TO DODAJ!!! :P