

Ograniczenia ekstrakcji częstotliwości podstawowej dźwięku fortepianu w oparciu o analizę cepstrum

Limit the extraction of fundamental frequency sound of the piano based on cepstrum analysis

dr inż. Arkadiusz Rajs

Uniwersytet Technologiczno-Przyrodniczy im. Jana i Jędrzeja Śniadeckich w Bydgoszczy,
Wydział Telekomunikacji, Informatyki i Elektrotechniki,
Zakład Teletransmisji, Al. Prof. S. Kaliskiego 7, 85-790 Bydgoszcz, arajs@utp.edu.pl.

dr hab. Katarzyna Rajs, prof. nadzw.

Akademia Muzyczna im. Feliksa Nowowiejskiego w Bydgoszczy
ul. Słowackiego 7, 85-008 Bydgoszcz, k.rajs@wp.pl

Abstrakt

Artykuł ten opisuje sposób wykorzystania algorytmu ekstrakcji częstotliwości podstawowej sygnału opartego na analizie cepstrum do określenia wysokości pojedynczych dźwięków instrumentów muzycznych, a dokładniej - dźwięku fortepianu. Przedstawia szczegółowo tą metodę oraz wskazuje problemy związane z dokładnością tego algorytmu dla różnych wysokości dźwięku.

Abstract

This article describes how to use the algorithm for the extraction of fundamental frequency signal based on cepstrum analysis to determine the amount of the individual sounds of musical instruments, and more-the sound of the piano. Shows in detail the method and indicates problems with the accuracy of this algorithm for various.

Słowa kluczowe: analiza cepstrum, przetwarzanie sygnałów

Key words: cepstrum analysis, signal processing

1. Wprowadzenie

Algorytmy ekstrakcji częstotliwości podstawowej sygnału znajdują szerokie zastosowanie – w technice studyjnej, sterowaniu urządzeniami, przy parametryzacji sygnałów. Możliwość określenia wysokości dźwięków instrumentu muzycznego może być również wykorzystana w zapisie melodii, konwersji na format MIDI lub automatycznym strojeniu instrumentu, itp. Algorytmy ekstrakcji częstotliwości dzielone są na trzy grupy: oparte na analizie w dziedzinie czasu, częstotliwości oraz hybrydowe (oparte na analizie czasowo-częstotliwościowej). Jedną z prostych i często stosowanych metod czasowych jest metoda oparta na analizie sygnału cepstrum.

Celem niniejszej pracy jest przedstawienie metodyki wydzielenia częstotliwości przy pomocy cepstrum sygnału oraz omówienie ograniczeń tej metody w stosunku do dźwięku fortepianu.

2. Metodyka ekstrakcji częstotliwości przy pomocy cepstrum sygnału

Jedną z prostych metod określenia częstotliwości podstawowej jest metoda wykorzystująca cepstrum sygnału. Jest ona oparta na obliczaniu odwrotnej transformaty Fouriera logarytmu widma amplitudowego analizowanej ramki sygnału, wg wzoru (1):

$$c_r = \sum_{n=1}^m \ln(X_n) \cdot \cos\left(\frac{r \cdot n\pi}{m}\right) \quad (1)$$

Następnie należy dokonać analizy maksimumów w dziedzinie cepstrum. Maksimum funkcji występuje dla zerowego opóźnienia ($m=0$). Położenie następnego maksimum daje estymatę okresu. Dla sygnałów dyskretnych długość okresu (w próbkach) można oszacować na podstawie maksimum lokalnego korzystając z zależności:

$$\hat{T}_1 = \max_m c_r(m) \quad , \text{ dla } r > 0 \quad (2)$$

3. Analiza dokładności metody ekstrakcji częstotliwości podstawowej dźwięku fortepianu

3.1. Zakres dźwięku fortepianu

W roku 1855 w Wiedniu, wobec istnienia dowolności strojenia instrumentów, międzynarodowa komisja normalizacyjna ustaliła, że dźwięk a^1 (z oktawy razkreślnej), według którego ustala się wysokość strojenia wszystkich instrumentów (również fortepianu), ma częstotliwość 440 Hz [3]. Fortepian jest instrumentem muzycznym, który posiada najbardziej rozległą skalę dźwięków

spośród wszystkich instrumentów akustycznych. Obecną skalę dźwięków fortepianu (liczba dźwięków przez wieki nieznacznie się zmieniała) rozpoczyna najniższy ton A subkontra, a kończy najwyższy - c pięciokreślne, co daje w rezultacie 88 klawiszy [4]. Współczesny fortepian zazwyczaj strojony jest według tzw. stroju równomiernie temperowanego. Taki strój ma ważną właściwość - stosunek częstotliwości interwału oktawy jest równy 2:1, a każda oktawa temperowana składa się z 12 równych półtonów. Z tych dwóch właściwości można wyznaczyć współczynnik zwany mnożnikiem temperacji. Jest to iloraz częstotliwości każdego kolejnego i poprzedniego półtonu, który wynosi:

$$\sqrt[12]{2} \approx 1,059463 \quad (3)$$

Znając mnożnik temperacji i liczbę półtonów tworzących chromatyczną, dwunastostopniową gamę oraz wiedząc, że a^1 stroi się na 440 Hz, można znaleźć częstotliwości wszystkich dźwięków temperowanej gamy chromatycznej.

W celu określenia zakresu częstotliwości generowanego przez fortepian można posłużyć się wzorem (3), pamiętając jednak, że najniższy możliwy dźwięk to A subkontra. Wystarczy skorzystać z właściwości stroju równomiernie temperowanego, która mówi, iż skok wysokości dźwięku o oktawę daje dwukrotny wzrost lub zmniejszenie o połowę częstotliwości podstawowej. Odległość pomiędzy a^1 a A subkontra to 4 oktawy, częstotliwość najniższego dźwięku fortepianu będzie więc wynosiła:

$$f_{A\text{subkontra}} = \frac{440 \text{ Hz} (f_{a^1})}{2^4} = 27,5 \text{ Hz}$$

Najwyższy dźwięk fortepianu to c pięciokreślne (c^5). Częstotliwość dźwięku c^2 (odległego od a^1 o 3 półtony) to:

$$f_{c^2} = f_{a^1} \cdot (\sqrt[12]{2})^3 = f_{a^1} \cdot 2^{\frac{3}{12}} \approx 523,2511 \text{ Hz}$$

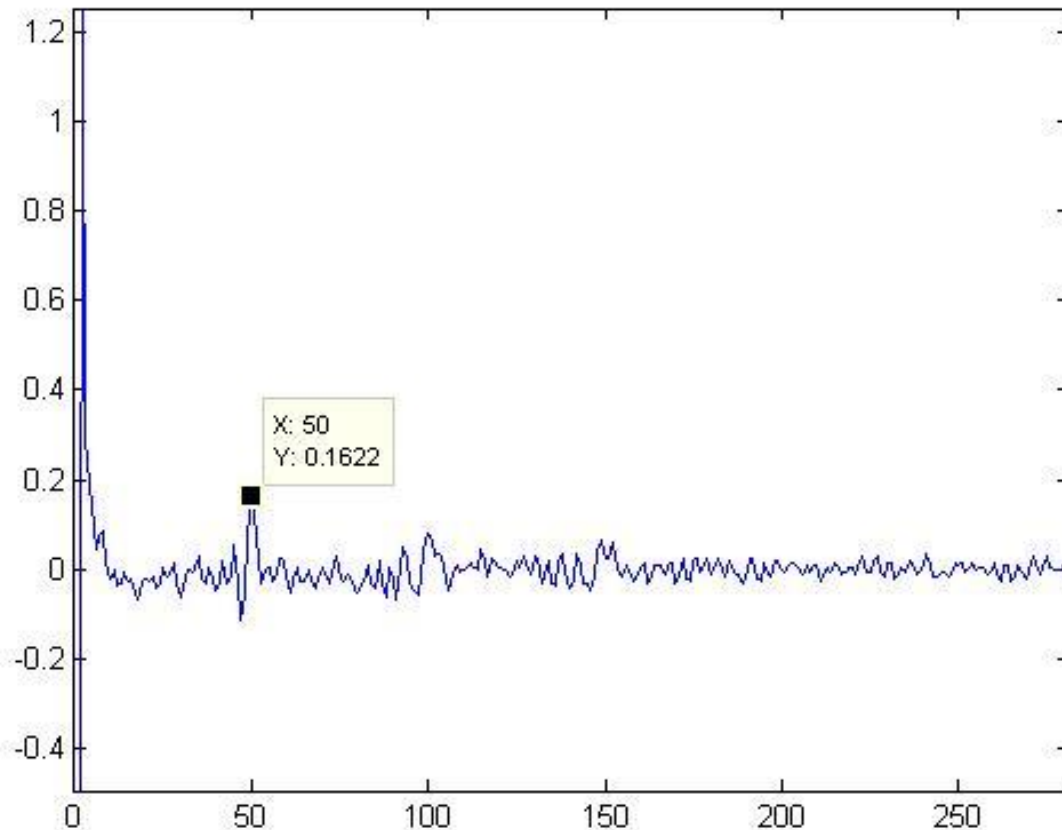
Wiedząc, że odległość najwyższego dźwięku c^5 od c^2 to 3 oktawy częstotliwość najwyższego dźwięku fortepianu wynosi:

$$f_{c^5} = 523,2511 \text{ Hz} \cdot 2^3 \approx 4186,0090 \text{ Hz}$$

3.2. Wpływ długości analizowanego sygnału na dokładność wyznaczanej częstotliwości

Biorąc pod uwagę najniższy i najwyższy dźwięk fortepianu można określić maksymalną i minimalną współrzędną maksimum lokalnego wyznaczonego za pomocą zależności (2). Tak więc jeden okres analizowanego sygnału może zawierać (w przypadku częstotliwości próbkowania 44100 Hz) 1604 próbek ($44100 \text{ Hz}/27,5 \text{ Hz}=1603,64$). Z kolei górną granicą jest 11 próbek. Przy niższych częstotliwościach próbkowania teoretyczna ilość potrzebnych próbek będzie odpowiednio mniejsza (dla $f_p=22050 \text{ Hz}$, zakres ilości próbek między 802 a 6 próbek). Jednak dla prawidłowego rozpoznania częstotliwości generowanego tonu potrzebne jest pobranie do analizy kilku okresów do kilkunastu okresów. Wynika z tego, że analizować trzeba kilka tysięcy próbek. Przykładowy

analizowany w oparciu o 4000 próbek dźwięku pokazuje Rys. 1. Na rysunku zaznaczono położenie maksimum lokalnego. Analizie poddano zdyskretyzowany z częstotliwością 22050 Hz dźwięku A razkreślne (440Hz).



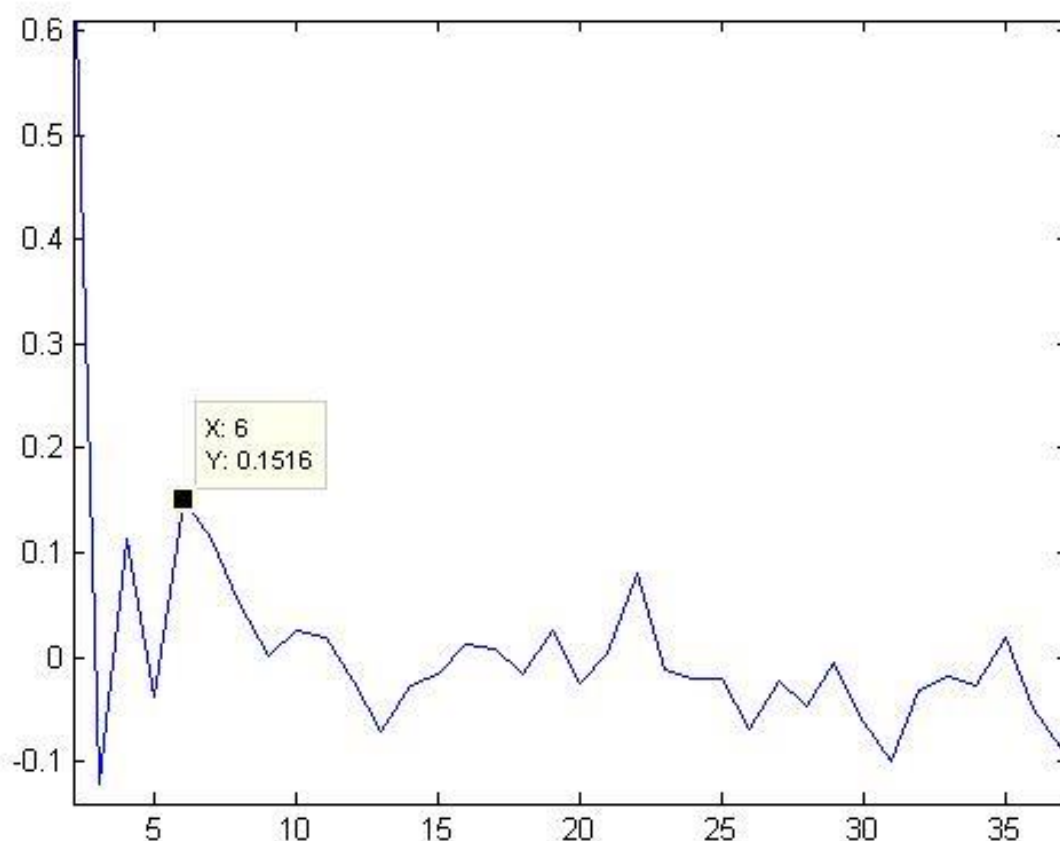
Rys. 1. Wykres cepstrum sygnału uzyskany dla 4000 próbek, $f_p=22050$ Hz

3.3. Wpływ wysokości dźwięku na położenie lokalnego maksimum

Drugim ważnym ograniczeniem omawianej metody jest rozdzielczość wynikająca z całkowitej wartości liczby próbek. Można ją wyrazić zależnością:

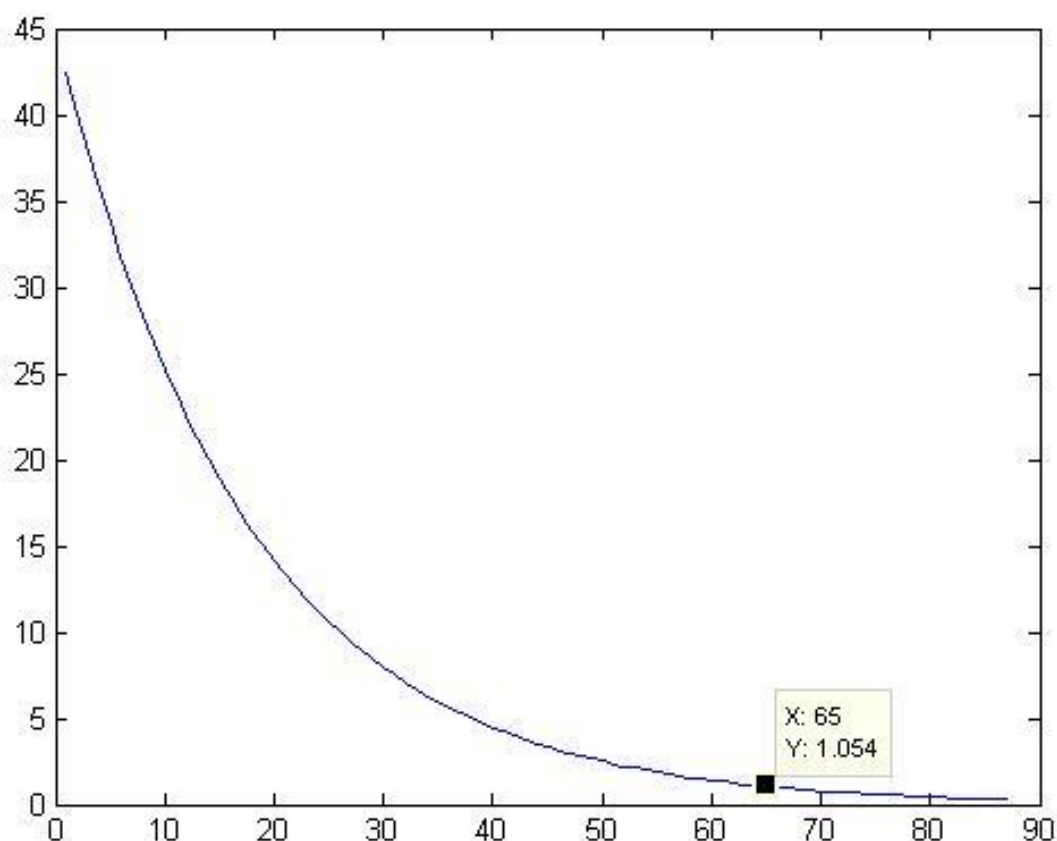
$$r(i) = f_p \left(\frac{1}{f_i} - \frac{1}{f(i+1)} \right) \quad (4)$$

Ma ona wpływ na dokładność analizy dźwięków wysokich. Różnica częstotliwości pomiędzy c^5 (ok. 4186 Hz) a h^5 (ok. 3951 Hz) wynosi 235 Hz, natomiast współrzędna maksimum wyrażona w liczbie próbek wyznaczona według zależności (2) to 6 (dla $f_p=22050$ Hz). Pokazano to na Rys.2, gdzie analizie poddano dźwięk c^5 . Wynika stąd, że różnica wyznaczonej częstotliwości dla $n=6$ i $n=7$ próbek wyniesie 525 Hz, czyli 3 sąsiadujące ze sobą najwyższe dźwięki nie mogą zostać rozpoznane prawidłowo. Wraz z malejącą częstotliwością dźwięków wzrasta rozdzielczość tak aby dla sąsiednich dźwięków c^5 i cis^5 (klawisze 64 i 65 licząc od najniższego dźwięku) przekroczyć wartość odstępu równą jednej próbce.



Rys. 2. Wykres cepstrum uzyskany dla 4000 próbek, $f_p=22050$ Hz, dźwięk c⁵

Zależność pomiędzy kolejnymi dźwiękami (klawisze numerowane od najniższego dźwięku do najwyższego) a wartością współrzędnej maksimum lokalnego (ilość próbek) dla częstotliwości próbkowania $f_p=22050$ Hz przedstawiona jest na Rys. 3. Przy wzroście częstotliwości próbkowania sygnału rozdzielczość rośnie wprost proporcjonalnie (zależność (4)).



Rys. 3. Wykres różnicy ilości próbek pomiędzy kolejnymi dźwiękami (klawiszami) fortepianu dla $f_p=22050$ Hz

4. Wnioski

Wykorzystanie metody opartej na sygnale cepstrum do wyznaczania częstotliwości podstawowej dźwięku fortepianu jest możliwe pod warunkiem pobrania do analizy dużej ilości próbek (kilka, kilkanaście okresów sygnału). Metoda jest prosta i nie wymaga skomplikowanych arytmetycznych obliczeń (w porównaniu do analizy widmowej). Ograniczenia dokładności ekstrakcji częstotliwości (głównie dotyczące wysokich dźwięków fortepianu) mogą być minimalizowane poprzez zwiększenie częstotliwości próbkowania analizowanego sygnału.

Literatura

- [1] Otnes R.K., Enochson L., *Analiza numeryczna szeregów czasowych*. Wydawnictwo Naukowo-Techniczne, Warszawa, 1978.
- [2] Bendat J.S., Piersol A.G., *Metody analizy i pomiaru sygnałów losowych*, PWN, Warszawa, 1976.
- [3] Fischer, J. C., *Piano Tuning*, Dover, N.Y., 1975.
- [4] Drobner M., *Instrumentoznawstwo i akustyka*. PWN, Kraków, 1975.
- [5] Rajs A. *The extract restriction of a piano sound basic frequency based on autocorrelation signal's analysis*, *Відбір і обробка інформації* № 25 (101), Львів, 2006. s.34-37.