



Politechnika Gdańska
WYDZIAŁ ELEKTRONIKI
TELEKOMUNIKACJI I
INFORMATYKI



Katedra:	Systemów Multimedialnych
Imię i nazwisko dyplomanta:	Piotr Pastuszek
Nr albumu:	93936
Rodzaj studiów:	magisterskie
Specjalność :	Inżynieria Dźwięku i Obrazu
Temat pracy:	Śledzenie podnośnych sygnałów fonicznych
Kierujący pracą:	prof. zw. dr hab. inż. Andrzej Czyżewski
Konsultant:	mgr inż. Przemysław Maziewski

Zakres pracy:

Celem pracy jest przeanalizowanie istniejących algorytmów do śledzenia charakterystyki pasożytniczych zmian częstotliwości (1) w zniekształconych nagraniach fonicznych, a następnie stworzenie własnego algorytmu bazującego na śledzeniu tonu MTS. W części teoretycznej ~~dotatkowo~~ przedstawione jest zjawisko zniekształcenia ~~drżenia~~ (ang. wow). W części praktycznej omówiona jest implementacja stworzonego algorytmu w środowisku MATLAB.

Abstract:

The aim of this work is to analyze existing algorithms for estimating the pitch variation curve of distorted audio recordings, and then to propose an algorithm for the MTS pilot tone tracking. Additionally, in the theoretical part, “Wow” distortion phenomenon is described. In practical part of this work, implementation of created algorithm in MATLAB environment is presented.

Data wydania tematu:	01.03.2007
Data złożenia pracy:	

Kierujący pracą

Kierownik Katedry

Spis treści

1. Wprowadzenie.....	3
2. Podstawy teoretyczne.....	6
2.1 Transformacja Fouriera.....	6
2.1.1 FFT.....	7
2.1.2 STFT.....	8
2.1.3 Widmo sygnału.....	9
2.1.3.1 Przecieki widma.....	10
2.1.3.2 Okna czasowe.....	11
2.1.3.3 Momenty widmowe.....	14
2.1.4 Spektrogram.....	14
2.2 Nośniki magnetyczne.....	19
2.2.1 Prąd podkładu.....	21
2.3 Standard NTSC.....	23
2.3.1 System MTS.....	23
3. Zniekształcenie Wow	25
3.1 Przyczyny powstawania zniekształcenia.....	25
3.2 Wybrane metody estymacji charakterystyki zniekształcenia....	27
3.2.1 Metoda śledzenia składowych tonalnych.....	28
3.2.2 Metoda autokorelacyjna.....	36
3.2.1 Adaptacyjne badanie środka ciężkości widma.....	41
3.2.4 Metody związane ze śledzeniem podnośnych.....	44
3.2.4.1 Przydźwięk sieciowy.....	45
3.2.4.2 Wysokoczęstotliwościowy prąd podkładu.....	49
4. Realizacja algorytmów do śledzenia podnośnych.....	52
4.1 Wysokoczęstotliwościowy prąd podkładu.....	52
4.2. Pilot MTS.....	56
4.2.1 Wprowadzone modyfikacje.....	58
4.2.2. Opis algorytmu.....	59
4.2.3 Eksperyment i wyniki.....	65
5. Wnioski i podsumowanie.....	68
6. Bibliografia.....	69
7. Załączniki.....	71

1. Wprowadzenie

Istotą zdarzeń dźwiękowych jest ich chwilowość i niepowtarzalność. O ile drugą z tych cech można kompensować w pewnym zakresie poprzez naukę ich generowania (mowa w zasadzie o dźwiękach o charakterze muzycznym), to specyficzna ulotność wrażeń psychoakustycznych, jakie powodują jest związana z działaniem mózgu, a konkretniej – jego pamięci. To dlatego ludzie od wieków starali się znaleźć sposób, by dźwięk utrwalić, zupełnie analogicznie do tego, jak malarze utrwalają wrażenia wizualne.

Pierwsze próby zapisania dźwięku dotyczyły **niejako** automatycznych instrumentów muzycznych (których przedstawicielem jest dobrze znana wszystkim pozytywka). Wraz z nadejściem epoki Renesansu przeżyły one swój rozkwit. Zasada działania była prosta – wprawiany w ruch za pomocą ludzkiej ręki mechanizm kołowy z przytwierdzonymi doń szpilkami potraçał różne dzwoneczki, w zależności od układu tychże szpilek. Ważne było, że dawało to okresowo powtarzaną sekwencję dźwięków – można było więc trwale zapisać pewien muzyczny układ (oczywiście względnie prosty).

Na przełomie XIII i IX wieku pewien francuski wynalazca imieniem Jacques de Vaucanson znalazł zastosowanie kart perforowanych (które były pierwszym zastosowanym w historii informatyki nośnikiem danych) przy konstrukcji automatycznych instrumentów muzycznych [34].

W 1857 roku we Francji został opatentowany „fonautograf” (ang. *phonautograph*) [36][29] – urządzenie Leona Scotta de Martinville'a, które potrafiło zapisać na zaciemnionym papierze przebieg wychyleń membrany wprawionej w drgania falą akustyczną. Odtworzenie w ten sposób zapisanego przebiegu dźwięku nie było jednak możliwe. Wynalazek ten zainspirował jednak T. Edisona [35] do stworzenia podobnego urządzenia, umożliwiającego także odtwarzanie dźwięku (zapisanego na wałku lub na płycie), znanego dziś jako fonograf, który został opatentowany pod koniec 1877 roku. Wtedy właśnie pojawiło się pierwsze nagranie ludzkiego głosu („*Marry had a little lamb*”) zapisanego na cienkiej folii cynowej [47].

Fonograf szybko znalazł swoich następców – najpierw grafon, potem gramofon (lata 1885 i 1887) [35][29]. Nikt jeszcze nie myślał o jakości i wierności nagrań – wielkim sukcesem była sama możliwość nagrywania dźwięków i odtwarzania ich w rozpoznawalnej

1. Wprowadzenie

formie [30]. Dalszy rozwój urządzeń i materiałów do zapisu dźwięku polegał na zwiększaniu czasu trwania nagrań oraz łatwości samego ich dokonywania. **Ogromne** zniekształcenia były wnoszone już na etapie zamiany fali akustycznej na elektryczną (niedoskonałe mikrofony węglowe), podobnie nieliniowość i niedokładność zapisu na ówczesnych nośnikach. W 1935 roku miała miejsce publiczna demonstracja działania pierwszego magnetofonu. W międzyczasie powstały mikrofony dynamiczne i pojemnościowe, które także wpłynęły pozytywnie na jakość ówczesnie rejestrowanego dźwięku. O ile jednak na sam etap rejestracji dźwięku wpływ miały coraz lepsze przetworniki elektroakustyczne, to na trwałość i dokładność nagrań kluczowy wpływ miała technika cyfrowa. Cyfrowe formaty zapisu, cyfrowy tor foniczny – wszystko to pozwalało na uczynienie urządzeń bardziej przezroczystymi dla dźwięku, nie zmieniającymi oryginalnej postaci sygnału. **Powoli zaczęło się pojawiać pojęcie korekcji dźwięku, polepszania go.**

Dziś ~~już~~ celem rejestracji dźwięku jest nie tyle samo jego utrwalenie, ale zrobienie tego w sposób jak najwierniejszy i jak najtrwalszy. Mimo sukcesywnego zbliżania się do doskonałości, rejestracja dźwięku nadal pozostaje podatna na wszelkie zakłócenia, zapis podlega zniekształceniom, a materiał przy nim wykorzystywany podlega ~~całkowicie naturalnemu~~ zużyciu.

Pośród wielu możliwych zakłóceń i zniekształceń ~~analogowo zapisywanego sygnału~~ fonicznego występuje zjawisko modulacji częstotliwości. Dzielimy je na wolno- i szybkozmienne, a w fachowej terminologii określane jest odpowiednio jako ~~drżenie i kołysanie~~ (ang. *wow and flutter*). Z pierwszym z nich związana jest ściśle tematyka niniejszej pracy. **W dalszej jej części wykorzystywana będzie angielska nazwa zniekształcenia – „wow”.**

Drugi rozdział poświęcony jest w całości wstępnym zagadnieniom teoretycznym, które przewijają się w dalszej **jej** części. Dotyczą one zarówno matematycznych kwestii związanych z transformacją Fouriera, widmem sygnału, jak i kwestii technicznych dotyczących magnetycznego zapisu i standardu telewizji **NTSC**.

Trzeci rozdział pracy dotyczy zniekształcenia wow. Opisane jest samo zjawisko, przyczyny jego powstawania oraz wybrane sposoby **detekcji** jego charakterystyki. Metody detekcji są specyficzne w stosunku do nośnika, na którym doszło do powstania zniekształcenia. Ogólnie zagadnienie detekcji i eliminacji wow zawiera się w znalezieniu krzywej ~~zmiany~~ częstotliwości (*pitch variation curve* zwana dalej w skrócie PVC) dla zniekształconego nagrania, a następnie odpowiednim przepróbkowaniu sygnału celem

1. Wprowadzenie

skompensowania niechcianej modulacji częstotliwości.

Właściwa (praktyczna) część pracy zawarta zostanie w rozdziale czwartym. W pierwszym podrozdziale przedstawiona zostanie szczegółowo metoda określania PVC na podstawie obserwacji wysokoczęstotliwościowego prądu podkładu stosowanego w nośnikach magnetycznych. Metoda ta stanowi punkt wyjścia dla głównego zadania niniejszej pracy – stworzenia algorytmu do śledzenia tonu pilota MTS (Multichannel Television Sound) obecnego w stereofonicznych nagraniach fonicznych w telewizji standardu NTSC. Tworzeniu i omówieniu tej metody poświęcony jest drugi podrozdział rozdziału czwartego.

Ostatni z podrozdziałów części praktycznej przedstawia efekty działania stworzonego algorytmu.

W podsumowaniu (rozdział piąty) znajdują się wnioski z przeprowadzonej pracy badawczej oraz możliwości dalszego usprawnienia użytej metody wykrywania zniekształceń. Tam też znajdzie się ocena opracowanego algorytmu i przedyskutowane zostaną jego ograniczenia.

W bibliografii znajdują się pozycje źródłowe wykorzystane przy pisaniu pracy, a załączniki obejmują kody źródłowe i dokumentacje stworzonych algorytmów.

2. Podstawy teoretyczne

Przed przejściem do technicznej części pracy, należy przybliżyć niektóre ze zjawisk i pojęć występujących w dalszej jej części. Osobny rozdział ma na celu uporządkowanie pojedynczych aspektów wiedzy z różnych tematów, których niniejsza praca dotyczy. Szczegółowy charakter zagadnienia, jakim jest śledzenie **podnośnych sygnałów fonicznych** narzuca opis zarówno wiedzy ogólnej, jak i bardzo specyficznej. W takiej też kolejności omówione zostaną wszelkie kwestie teoretyczne.

2.1 Transformacja Fouriera

Fundamentalną operacją pozwalającą badać sygnały okresowe jest transformacja Fouriera. Nazwę swą zawdzięcza francuskiemu matematykowi, **Josephowi Fourierowi** (Rys. 1), który żyjąc na przełomie XIII i XIX wieku dowiódł, że każdą funkcję okresową da się przedstawić jako (nieskończony ~~najczęściej~~) szereg podstawowych funkcji trygonometrycznych z odpowiednimi współczynnikami zwanymi dalej współczynnikami Fouriera. Twierdzenie to było częścią opublikowanej w 1922 roku pracy „*Théorie analytique de la chaleur*”, która dała podstawy ^Aanalizy sygnałów.



Rysunek 1: Jean Baptiste Joseph Fourier (1768 - 1830), francuski matematyk

2. Podstawy teoretyczne

Ciągła transformacja Fouriera (FT) to operacja całkowania sygnału w dziedzinie pulsacji zespolonej [45]. Pulsacja to wielkość proporcjonalna do częstotliwości, wyrażająca się wzorem:

$$\omega = 2\pi f \quad (1)$$

Sama transformacja Fouriera wyraża się wzorem:

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (2)$$

gdzie $\hat{f}(\omega)$ to transformata funkcji w dziedzinie czasu $f(t)$,
a j oznacza jednostkę urojoną (czyli $\sqrt{j} = -1$)

Jak widać z powyższego wzoru, operacje dokonywane są na funkcji dziedziny czasu, a więc na sygnałach analogowych. Gdy chcemy przejść w dziedzinę czasu dyskretnego, czyli *de facto* w dziedzinę próbek, musimy skorzystać z dyskretniej transformacji Fouriera (DFT) [39]. Wyraża się ona wzorem:

$$A_n = \sum_{k=0}^{N-1} a_k e^{-j \frac{2\pi}{N} kn}, \quad 0 \leq k \leq N-1 \quad (3)$$

gdzie a_k to wartość k -tej próbki sygnału,
 N to całkowita liczba próbek,
 n to numer harmonicznej,
a j oznacza jednostkę urojoną

2.1.1 FFT

Ze względu na dużą złożoność obliczeniową DFT wynoszącą $O(N^2)$ operacji, w informatyce zamiast DFT stosuje się FFT, czyli szybką transformację Fouriera, której złożoność obliczeniowa wynosi $O(N \log(N))$. Warunkiem koniecznym do przeprowadzenia obliczeń w oparciu o FFT jest ilość próbek N będąca potęgą liczby 2. W praktyce cyfrowego przetwarzania sygnałów audio najczęściej spotyka się z wykładnikiem w okolicach 10, co przekłada się na $N = 2^{10} = 1024$ próbek [22].

2. Podstawy teoretyczne

Operacje FFT na komputerach implementowane są z wykorzystaniem szeroko rozpowszechnionej biblioteki FFTW (Fastest Fourier Transform in the West) [40]. Jest to najszybsza i niezależna od sprzętu biblioteka, która znalazła zastosowanie w wielu komercyjnych (takich jak MATLAB czy programy audio), jak i niekomercyjnych aplikacjach wykorzystujących bądź umożliwiających cyfrowe przetwarzanie sygnałów.

2.1.2 STFT

Rozdzielczość czasowo-częstotliwościowa, jaką dysponujemy wynika bezpośrednio z ilości próbek i częstotliwości próbkowania analizowanego fragmentu. Jeśli poddamy dyskretną transformację Fouriera fragment o długości N próbek, gdy częstotliwość próbkowania wynosi f_s , uzyskamy ciąg N współczynników widmowych. Reprezentują one częstotliwości od 0 do f_s , przy czym interesująca dla nas jest tylko ich połowa, jako że druga jest jej lustrzanym odbiciem, przy założeniu, że analizujemy sygnał rzeczywisty (nie zespolony). Częstotliwość równa $\frac{f_s}{2}$ określana jest jako częstotliwość Nyquista (nazwa pochodzi od twierdzenia Nyquista o próbkowaniu, zgodnie z którym najwyższą możliwą do próbkowania częstotliwością jest połowa częstotliwości próbkowania). Mając zatem do dyspozycji $\frac{N}{2}$ przedziałów częstotliwości od 0 do $\frac{f_s}{2}$, szerokość pojedynczego przedziału można przedstawić za pomocą wzoru:

$$B = \frac{\frac{f_s}{2}}{\frac{N}{2} - 1} \stackrel{N \gg 1}{\approx} \frac{f_s}{N} \text{ [Hz]} \quad (4)$$

Mamy zatem dwie zmienne wpływające bezpośrednio na rozdzielczość częstotliwościową transformaty – są nimi częstotliwość próbkowania f_s oraz ilość próbek N . Aby zmniejszyć szerokość pojedynczego pasma, możemy zmniejszyć częstotliwość próbkowania. Spowoduje to jednak faktyczne wydłużenie czasowe fragmentu sygnału poddawanego analizie, przy założeniu, że ilość próbek pozostaje stała. Z drugiej strony,

2. Podstawy teoretyczne

możemy zwiększyć ilość próbek, co przy stałej częstotliwości próbkowania spowoduje, tak jak w poprzednim przypadku, wydłużenie czasowe pobranego fragmentu sygnału. To z kolei powoduje pogorszenie rozdzielczości czasowej **analizy FFT**.

Z tego powodu dla przebiegów szybkozmiennych aby dostarczyć algorytmowi odpowiednią liczbę próbek, stosuje się tzw. okienkowanie, tj. wycinanie z sygnału fragmentów (okien) odpowiedniej długości. Jeśli nie stosujemy przy tym wag wynikające z położenia próbki względem długości wycinanego fragmentu, mówimy wówczas o oknie prostokątnym. Wpływ okien na rozdzielczość częstotliwościową analizy zostanie opisany w podrozdziale 2.1.3.2 – ~~Okna czasowe~~.

Mianem STFT określa się krótkookresową transformację Fouriera (Short-Term/Time Fourier Transform) [31]. Wyraża się ona wzorem:

$$X_{STFT}[n] = \sum_{k=0}^{\infty} x[k] \cdot w[k-m] \cdot e^{-j \frac{2\pi k}{N} n} \quad (5)$$

gdzie $x[]$ oznacza spróbkowany sygnał,
 $w[]$ oznacza funkcję okna,
 m oznacza przesunięcie początku okna
 N oznacza długość okna w próbkach
 n oznacza numer ~~harmonicznej~~

Od zwykłego DFT zapis różni się tym, że wydziela z sygnału fragment przy pomocy funkcji okna, która poza przedziałem swojej określoności ma wartość zerową, jednak dokonuje operacji transformacji Fouriera **z całości sygnału**. Operację taką nazywa się **wypełnianiem zerami (ang. zero padding)**, a pozwala ona zwiększyć rozdzielczość częstotliwościową przy niezminionej rozdzielczości czasowej analizy.

2.1.3 Widmo sygnału

W wyniku operacji FFT na sygnale dyskretnym, uzyskujemy szereg zespolonych współczynników widmowych, których moduły tworzą widmo amplitudowe, argumenty zaś – widmo fazowe.

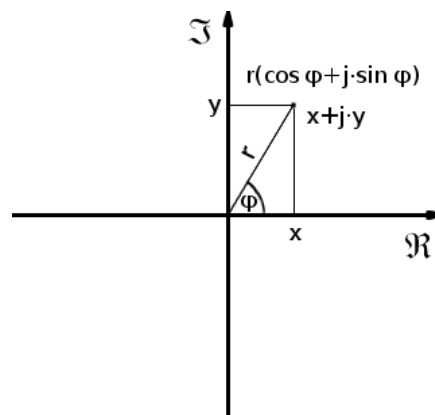
Dla przypomnienia, mając liczbę zespoloną w zapisie algebraicznym $x + j \cdot y$,

2. Podstawy teoretyczne

jej moduł wynosi $|x + j \cdot y| = \sqrt{x^2 + y^2}$ (6)

faza zaś - $\arg(x + j \cdot y) = \begin{cases} \arctan(\frac{y}{x}), & \text{dla } x > 0 \\ \arctan(\frac{y}{x}) + \pi, & \text{dla } x < 0 \end{cases}$ (7)

W zapisie trygonometrycznym liczby zespolonej mamy bezpośrednio widoczny moduł i fazę. Dla $r(\cos(\varphi) + j \sin(\varphi))$, r stanowi moduł, φ – fazę. Obrazuje to Rys. 2.



Rysunek 2: Układ współrzędnych zespolonych

2.1.3.1 Przecieki widma

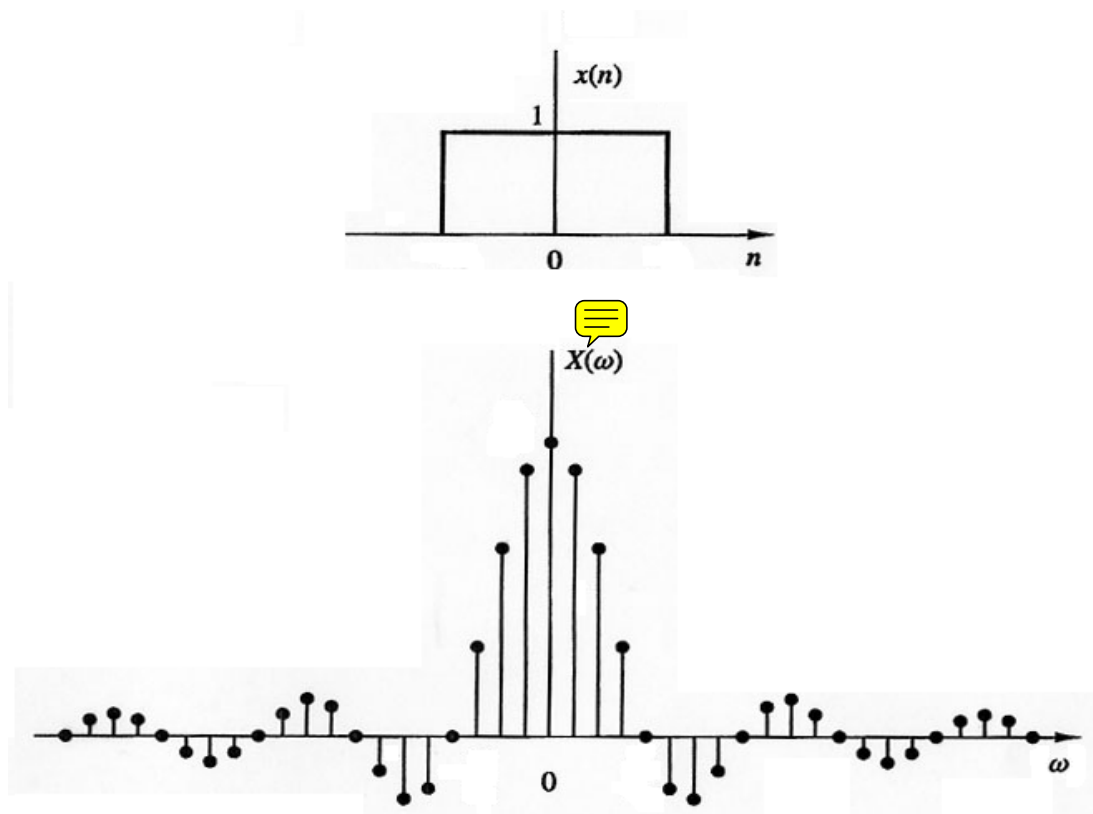
Widmo amplitudowe mówi nam o amplitudach poszczególnych składowych częstotliwościowych sygnału. Im większa ilość próbek branych do analizy, tym większa jej rozdzielczość – zależność jest wprost proporcjonalna do połowy ilości próbek N . W związku ze skończoną ich ilością, mamy do czynienia z tzw. przeciekami widma. Jest to zjawisko polegające na powstawaniu nieistniejących prążków w widmie w bliskim sąsiedztwie częstotliwości składowych faktycznie występujących. Dzieje się tak również, gdy poddajemy analizie fragment sygnału o długości nie będącej wielokrotnością jego okresu, ~~co w przypadku sygnałów rzeczywistych występuje zawsze~~. Z matematycznego punktu widzenia, przecieki widma powstają w wyniku mnożenia sygnału z oknem prostokątnym, ~~a więc ciągiem jedynek o długości odpowiadającej długości tego okna~~. W dziedzinie widma

2. Podstawy teoretyczne

odpowiada to splotowi widm obu sygnałów – pojedynczego prążka danej częstotliwości sygnału okresowego i widma sygnału prostokątnego, czyli funkcji sinc, definiowanej jako

$$\text{sinc}(x) = \frac{\sin(x)}{x} \quad (8)$$

Jako że operujemy na widmie dyskretnym, funkcja sinc zostaje spróbkowana, co obserwujemy jako „wyciek” prążka podstawowego na sąsiednie częstotliwości [14]. Zjawisko to zobrazone jest na Rys. 3.



Rysunek 3: Dyskretna funkcja sinc (u dołu) jako transformata Fouriera funkcji prostokątnej (u góry)

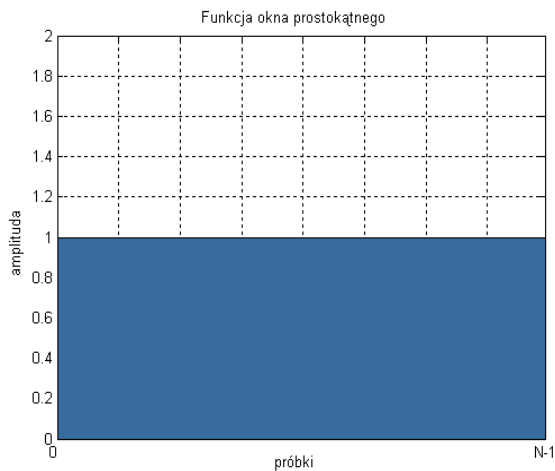
2.1.3.2 Okna czasowe

Niechciane prążki zwane także listkami bocznymi mogą utrudniać, a nawet uniemożliwiać detekcję składowych sygnału o niewielkim poziomie. W celu zminimalizowania ich wpływu stosuje się okna czasowe o specjalnych charakterystykach.

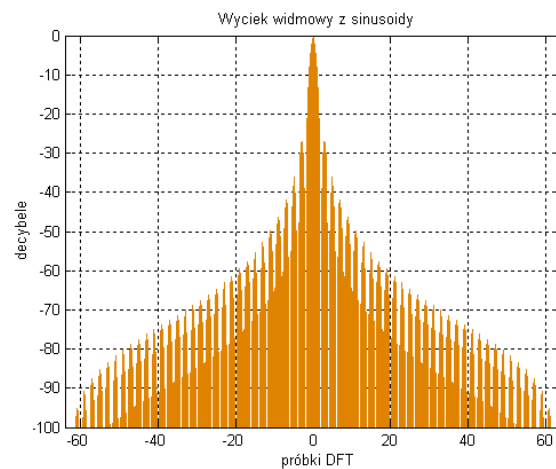
~~Transformata Fouriera~~ okien innych niż prostokątne daje funkcję zbliżoną do sinc, jednak o

2. Podstawy teoretyczne

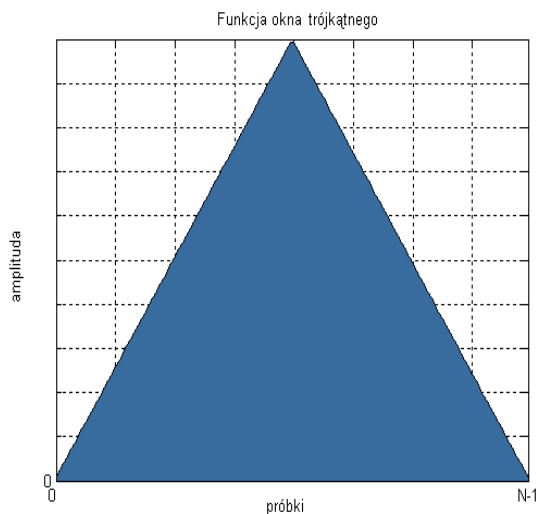
zdecydowanie korzystniejszych parametrach. Dostępnych jest wiele rodzajów okien o nazwach pochodzących najczęściej od nazwiska twórcy. Każde okno określa się wzorem matematycznym zwanym funkcją okna $w[n]$. Poniżej zamieszczone zostały wybrane funkcje okien (Rys. 4,6,8,10) wraz ze zobrazowanym odpowiadającym im wyciekim pojedynczego prążka częstotliwości (Rys. 5,7,9,11). Tab. 1 przedstawia funkcje przedstawionych okien.



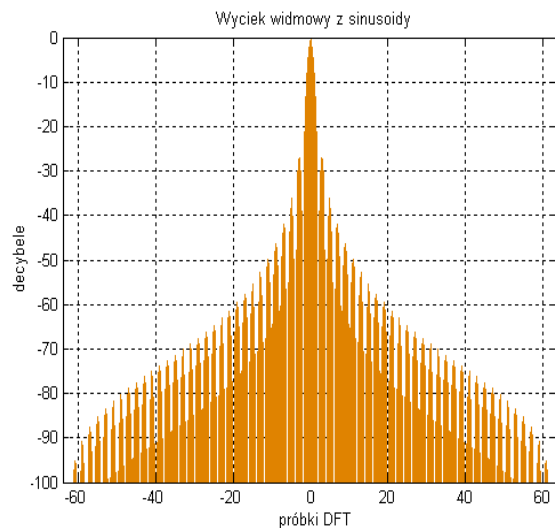
Rysunek 4: Okno prostokątne



Rysunek 5: ~~Odpowiedź impulsowa okna prostokątnego~~

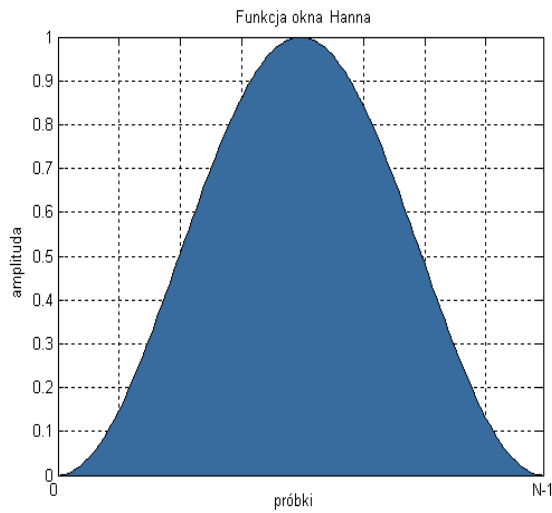


Rysunek 6: Okno trójkątne

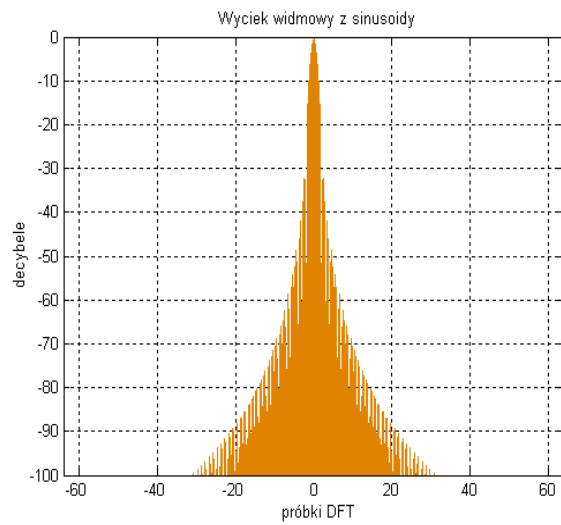


Rysunek 7: Odpowiedź impulsowa okna trójkątnego

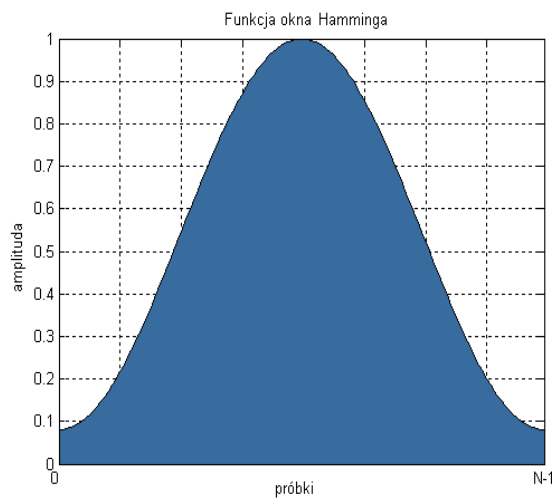
2. Podstawy teoretyczne



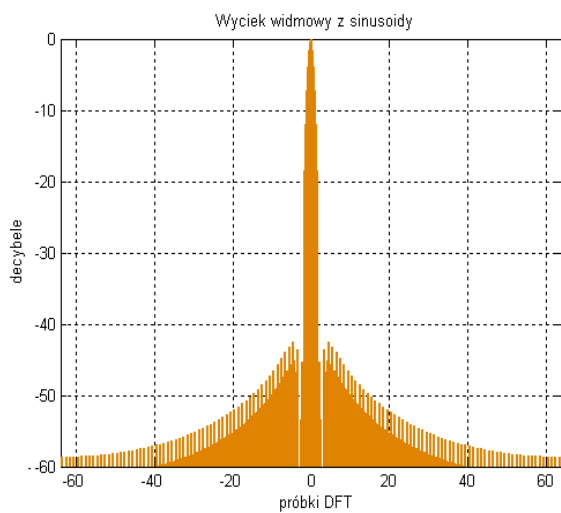
Rysunek 8: Okno Hanna



Rysunek 9: Odpowiedź impulsowa okna Hanna



Rysunek 10: Okno Hamminga



Rysunek 11: Odpowiedź impulsowa okna Hamminga

Tabela 1: Funkcje wybranych okien (N – długość okna; n – numer próbki)

Okno prostokątne	$w(n) = 1$
Okno trójkątne	$w(n) = \frac{N}{2} - \left n - \frac{N-1}{2} \right $
Okno Hanna	$w(n) = 0,5 \cdot \left(1 - \cos\left(\frac{2\pi n}{N-1}\right) \right)$
Okno Hamminga	$w(n) = 0,583836 - 0,46164 \cdot \cos\left(\frac{2\pi n}{N-1}\right)$

2. Podstawy teoretyczne

2.1.3.3 Momenty widmowe

Moment widmowy m-tego rzędu [11] definiuje się jako:

$$M(m) = \sum_{k=0}^{\infty} |G(k)| \cdot [f_k]^m \quad (9)$$

gdzie $G(k)$ oznacza wartość widma mocy dla k-tego pasma częstotliwości,

f_k oznacza częstotliwość środkową k-tego pasma

Definiujemy także moment unormowany m-tego rzędu, który wyraża się wzorem:

$$M_u(m) = \frac{M(m)}{M(0)} = \sum_{k=0}^{\infty} \frac{|G(k)| \cdot [f_k]^m}{|G(k)|} \quad (10)$$

Różne momenty widmowe przydają się do liczbowego określania parametrów związanych z widmem sygnału (środek ciężkości, szerokość, symetria itp.). Przykładowo moment zerowego rzędu ma interpretację mocy sygnału (wzór 11), zaś unormowany moment rzędu pierwszego (wzór 12) jest środkiem ciężkości widma (średnią ważoną wszystkich pasm częstotliwości).

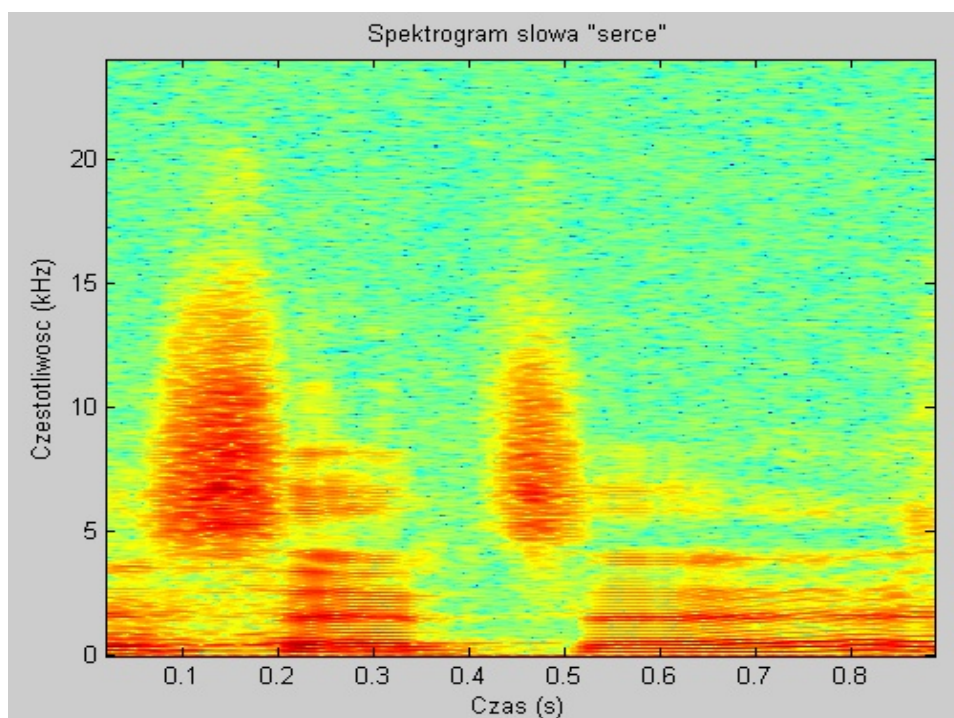
$$M(0) = \sum_{k=0}^{\infty} |G(k)| \quad (11)$$

$$M_u(1) = \sum_{k=0}^{\infty} \frac{|G(k)| \cdot [f_k]}{|G(k)|} \quad (12)$$

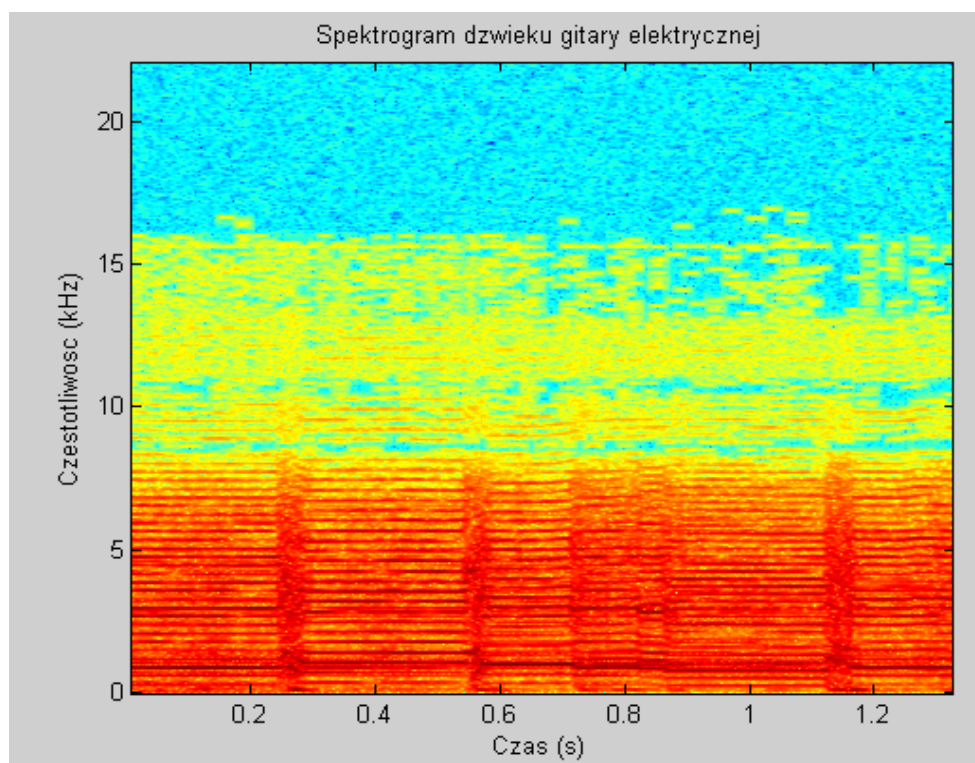
2.1.4 Spektrogram

Spektrogramem nazywamy czasowo-częstotliwościowy wykres podniesionego do drugiej potęgi widma amplitudowego sygnału, gdzie kolor odpowiada amplitudzie określonego pasma częstotliwości. Na osi poziomej jest czas, na osi pionowej – częstotliwość, ~~także~~ możemy mówić o wykresie obrazującym przebieg zmian widma sygnału w czasie. Przykładowe spektrogramy znajduje się na Rys. 12 i 13.

2. Podstawy teoretyczne



Rysunek 12: Przykładowy spektrogram sygnału mowy



Rysunek 13: Przykładowy spektrogram dźwięku przesterowanej gitary elektrycznej

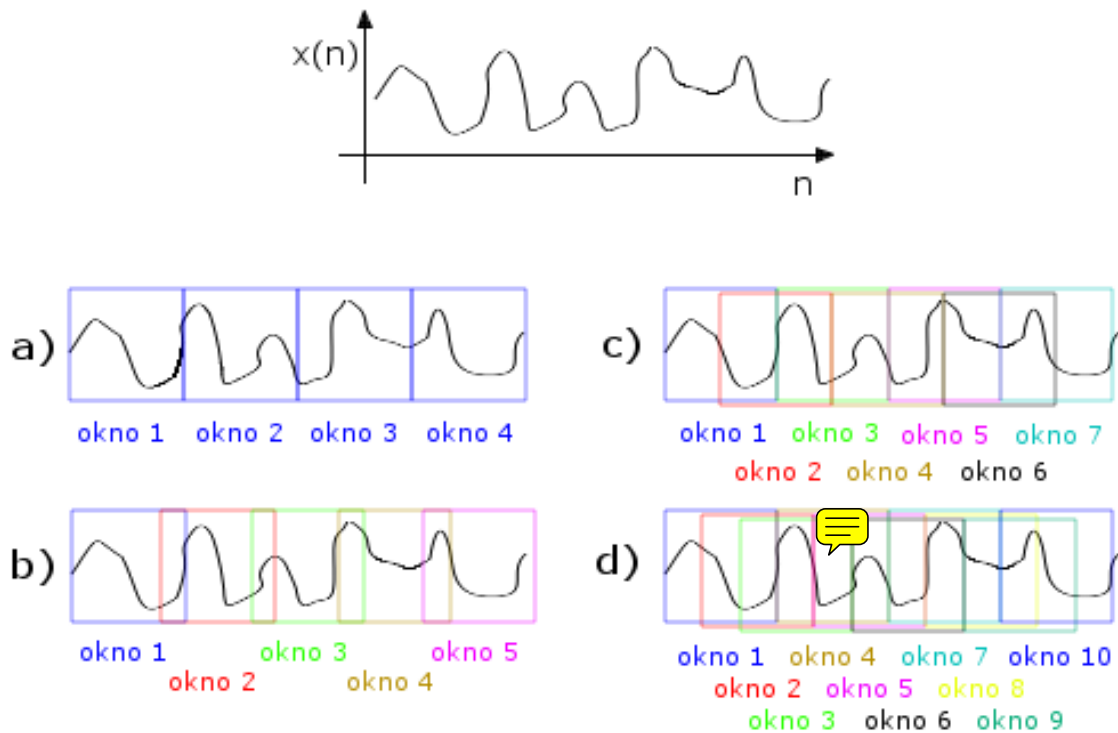
2. Podstawy teoretyczne

Na ostateczny kształt spektrogramu ma wpływ kilka parametrów dotyczących samego algorytmu obliczeń, jak i wizualizacji. Do pierwszej kategorii zaliczamy rodzaj i długość okna oraz wielkość nakładkowania, do drugich zaś – zakres kolorów i krzywa ich odwzorowania, przy czym te ostatnie dobierane są najczęściej automatycznie (wykresy z Rys. 12 i 13 powstały w środowisku MATLAB).

Wpływ rodzaju stosowanego okna na uzyskiwane widmo sygnału został omówiony w podrozdziale 2.1.3.2 – Okna czasowe. Wpływ długości okna ma bezpośrednie przełożenie na rozdzielczość czasową i częstotliwościową spektrogramu. Im dłuższe okno tym wyższa rozdzielczość częstotliwościowa, jednak tym gorsza rozdzielczość czasowa. Oznacza to, że zmiany składowych częstotliwościowych w czasie zaowocują rozmyciem widma. Jeśli wziąć pod uwagę fakt, że w rzeczywistych sygnałach akustycznych nie możemy liczyć na to, że będziemy mieli do czynienia z lokalnie stacjonarnymi sygnałami okresowymi, powstaje pewne ograniczenie co do maksymalnej długości fragmentu sygnału stosowanego do analizy. Jednocześnie zbyt mała jego długość przy oczywistej zalecie dobrej rozdzielczości czasowej w sensie nadążania za zmianami widma, da efekt zgrubnego jedynie oszacowania częstotliwości składowych sygnału.

Dodatkowo każda „porcja” sygnału podawana do algorytmu trwa określoną ilość czasu. Im dłuższe okno, tym rzadziej przeprowadzana analiza, czyli tym gorsza rozdzielczość czasowa (można tu mówić o pewnym próbkowaniu widm sygnału). Zatem by zwiększyć częstość czasową analizy FFT przy jednoczesnym niezmiennianiu długości okna, dokonuje się operacji zakładkowania, czyli częściowego nachodzenia na siebie kolejnych okien sygnału branych do analizy (Rys. 14). Im większe zakładkowanie tym dłuższa macierz spektrogramu, co przy zachowaniu proporcji zwiększa nam jego gęstość czasową.

2. Podstawy teoretyczne



Rysunek 14: Ilustracja działania zakładkowania okien:

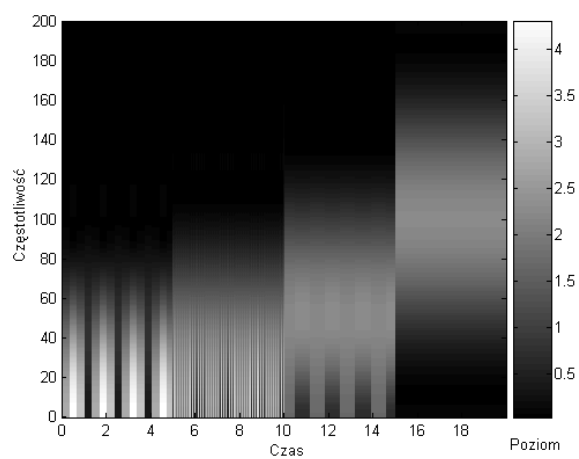
a) brak zakładkowania; b) zakładkowanie 20%; c) zakładkowanie 50%; d) zakładkowanie 66%

Nie pozostawia więc wątpliwości fakt, jak ważną kwestią jest odpowiednie dostosowanie parametrów przy tworzeniu (i późniejszej analizie) ~~spektrogramów~~.

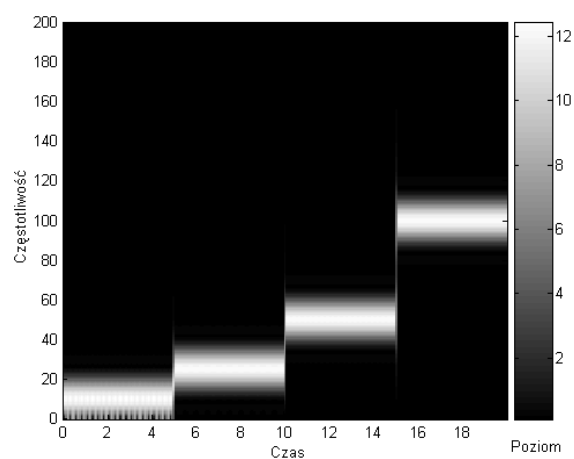
Jako przykład przedstawiono spektrogramy sygnału $x(t)$ dla różnych długości okna czasowego (Rys. 15-18). Sygnał $x(t)$ opisany jest wzorem:

$$x(t) = \begin{cases} \cos(2\pi 10t) & \text{dla } 0 \leq t < 0.1s \\ \cos(2\pi 25t) & \text{dla } 0.1 \leq t < 0.2s \\ \cos(2\pi 50t) & \text{dla } 0.2 \leq t < 0.3s \\ \cos(2\pi 100t) & \text{dla } 0.3 \leq t < 0.4s \end{cases} \quad (13)$$

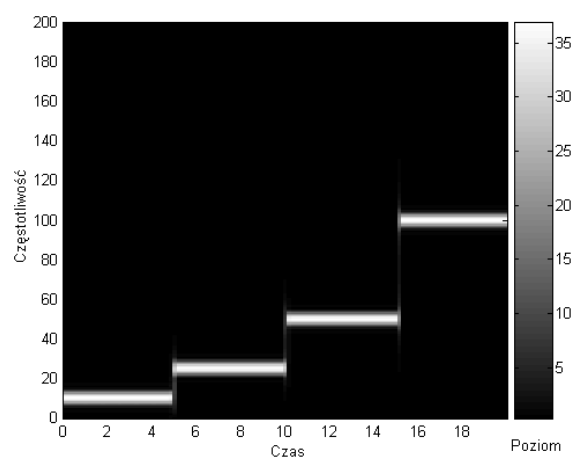
2. Podstawy teoretyczne



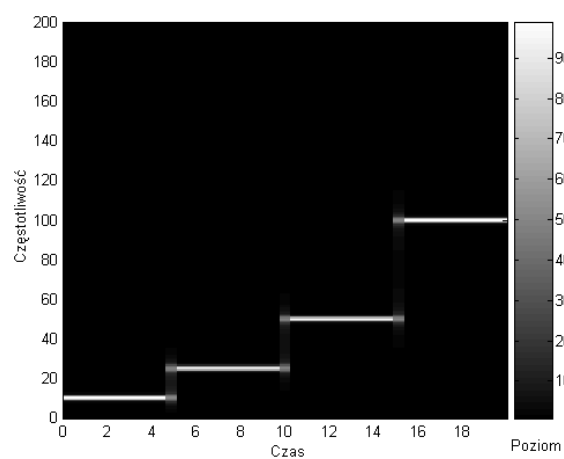
Rysunek 15: Spektrogram $x(t)$ dla $T=25\text{ms}$



Rysunek 16: Spektrogram $x(t)$ dla $T=125\text{ms}$



Rysunek 17: Spektrogram $x(t)$ dla $T=375\text{ms}$



Rysunek 18: Spektrogram $x(t)$ dla $T=1\text{s}$

2. Podstawy teoretyczne

2.2 Nośniki magnetyczne

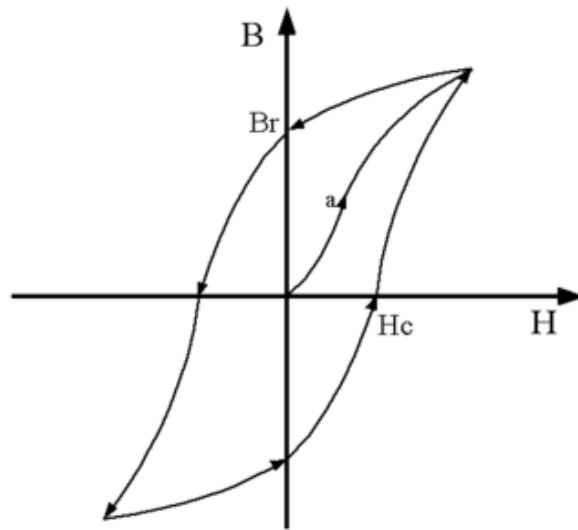
Pierwszym masowym medium dźwiękowym była z pewnością płyta gramofonowa. W pewnym momencie została jednak wyparta przez znacznie poręczniejsze taśmy magnetyczne. Dźwięk nie był na nich zapisywany mechanicznie, a poprzez wykorzystanie rozmieszczonych gęsto na ferromagnetycznej taśmie domen magnetycznych, których odpowiednie ukierunkowanie przy użyciu silnego pola głowicy zapisującej dawało efekt trwałego zapisu danych (taśmy nie służyły bowiem tylko do zapisu dźwięku – wraz z dyskami magnetycznymi przysłużyły się upowszechnieniu pierwszych komputerów osobistych). Wykorzystuje się tutaj zjawisko istnienia remanencji, czyli niezerowej wartości indukcji magnetycznej pozostającej po usunięciu zewnętrznego pola. Jest to jeden z dwóch czynników charakteryzujących histerezę magnetyczną [21].

Ogólnie między indukcją magnetyczną $B[T]$, a natężeniem pola magnetycznego

$H\left[\frac{A}{m}\right]$ istnieje zależność określona równaniem $B=\mu H$ (14), gdzie μ to przenikalność

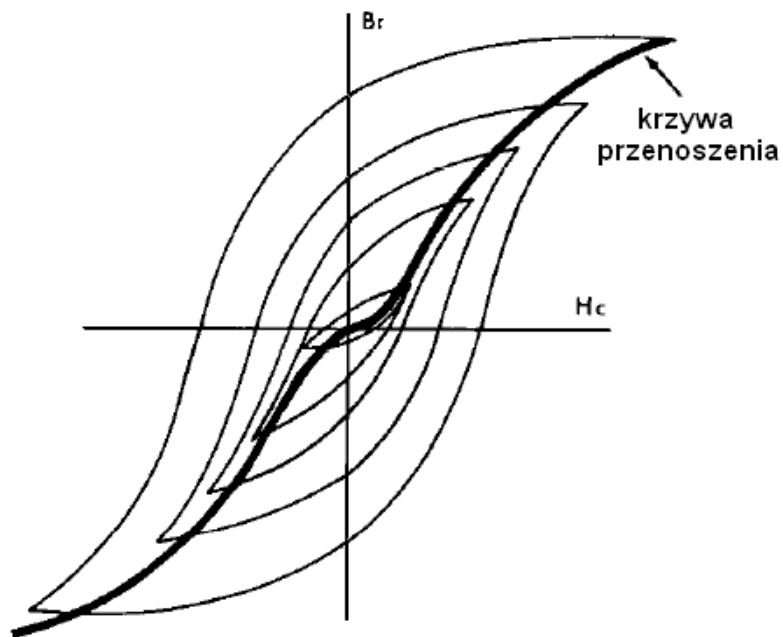
magnetyczna ośrodka wyrażona w $\left[\frac{H}{m}\right]$. Istotną ~~jednak~~ cechą wszystkich procesów związanych z magnetyzacją jest silna nieliniowość, a dodatkowo zależność stanu układu od stanów poprzednich (definicja samej histerezy). Krzywa histerezy (Rys. 19) opisuje zależność między natężeniem zewnętrznego pola magnetycznego, a indukcją magnetyczną ośrodka, na jakie ono działa. Drugą z cech procesów magnetycznych wspomnianych powyżej pokazują strzałki zaznaczone na krzywej. Widać, że po początkowym ~~namagnesowaniu~~ (fragment oznaczony literą a), a następnie wyłączeniu zewnętrznego pola, wartość indukcji nie spada do zera, a utrzymuje się na poziomie remanencji B_r .

2. Podstawy teoretyczne



Rysunek 19: Zjawisko histerezy magnetycznej

Remanencja magnetyczna B_r pozwala zatem na utrwalenie wektora natężenia pola magnetycznego w postaci indukcji magnetycznej. Oznacza to, że modulując wektor natężenia pola sygnałem akustycznym, uzyskujemy trwały zapis na taśmie magnetycznej w postaci odpowiadającej natężeniu pola indukcji. Funkcję przenoszenia wyznacza się z rodziny pętli histerezy dla rosnącego poziomu natężenia pola (Rys. 20).

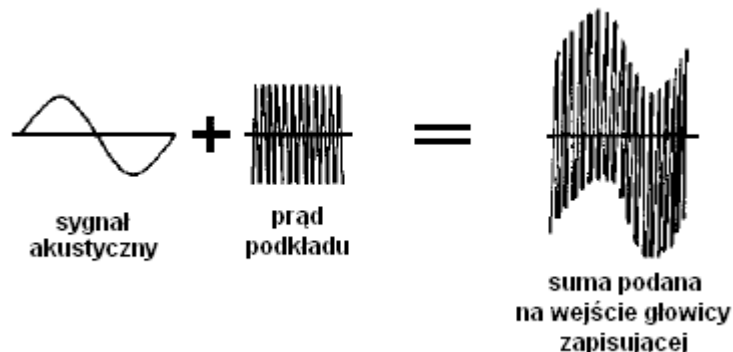


Rysunek 20: Rodzina pętli histerezy dla rosnącego natężenia pola wyznaczająca krzywą przenoszenia ośrodka magnetycznego

2. Podstawy teoretyczne

2.2.1 Prąd podkładu

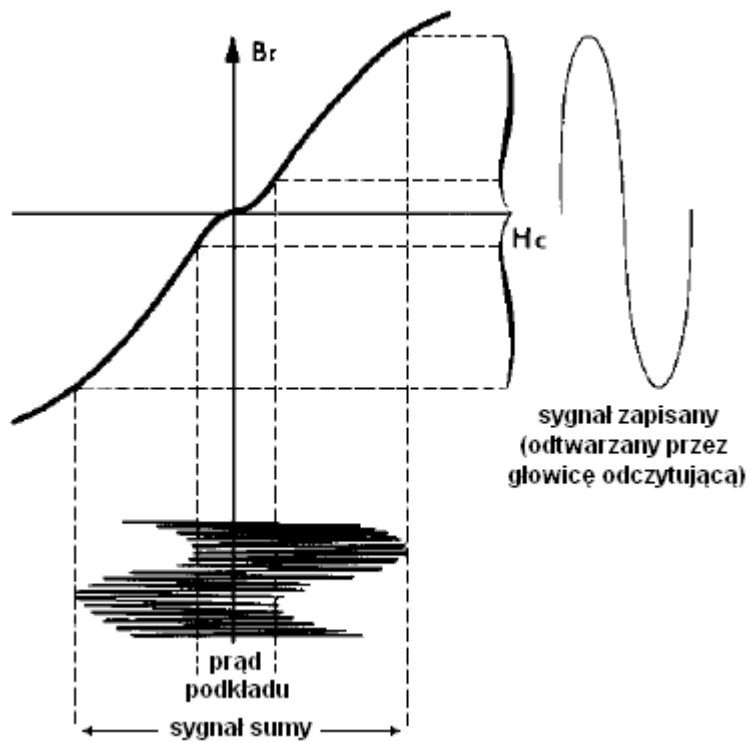
Jak widać z Rys. 20, krzywa przenoszenia taśmy magnetycznej jest dla małych wartości natężenia pola magnetycznego silnie nieliniowa. Podobnie obserwuje się zakres nasycenia ośrodka, kiedy to zwiększanie natężenia zewnętrznego pola magnetycznego nie przekłada się na zwiększoną indukcję. Istnieje zatem pewien wąski zakres liniowego przenoszenia układu głowica-taśma. Aby tylko ten zakres był wykorzystany podczas zapisu, dodaje się do sygnału pewien wysokoczęstotliwościowy (30÷300 kHz) ton zwany prądem podkładu (Rys. 21), który nie wpływa na percepcję sygnału (jest wysoko poza pasmem akustycznym, a ponadto nie jest odtwarzany przez głowicę odczytującą), jednak poprawia właściwości magnetyczne nagrania (Rys. 22) [23][37]. Sama częstotliwość tonu jest różna i zależy od magnetofonu, na którym dokonuje się zapisu. Nie ma to potem wpływu na odczyt na innym sprzęcie – ton prądu podkładu nie jest wykorzystywany podczas odtwarzania!



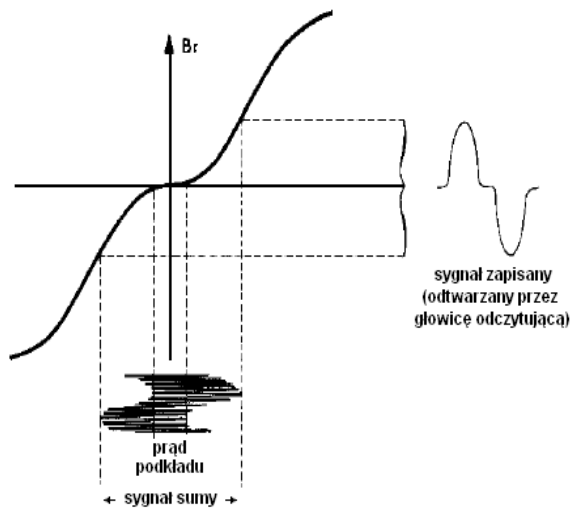
Rysunek 21: Dodawanie prądu podkładu do nagrania

Poziom prądu podkładu powinien być odpowiednio dobrany, żeby w pełni wykorzystać liniowy zakres funkcji przenoszenia. W przypadku złego doboru mogą pojawić się zniekształcenia nieliniowe (dodatkowe harmoniczne, przesterowania). Obrazują to Rys. 23 i 24.

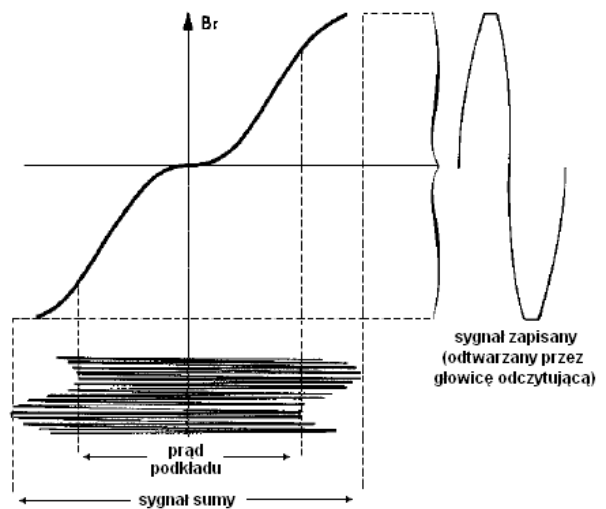
2. Podstawy teoretyczne



Rysunek 22: Krzywa przenoszenia taśmy magnetycznej z uwzględnieniem prądu podkładu w sygnale



Rysunek 23: Zbyt mały prąd podkładu



Rysunek 24: Zbyt duży prąd podkładu

2. Podstawy teoretyczne

2.3 Standard NTSC

NTSC to powstały w 1953 roku amerykański standard kodowania obrazu telewizji analogowej. Jego nazwa wzięła się od instytucji, która zatwierdziła standard – National Television Standard Committee. Jest on wykorzystywany w Ameryce, Azji wschodniej i na Pacyfiku. Dla porównania, w Europie standardem telewizyjnym jest PAL (we Francji zbliżony do niego SECAM).

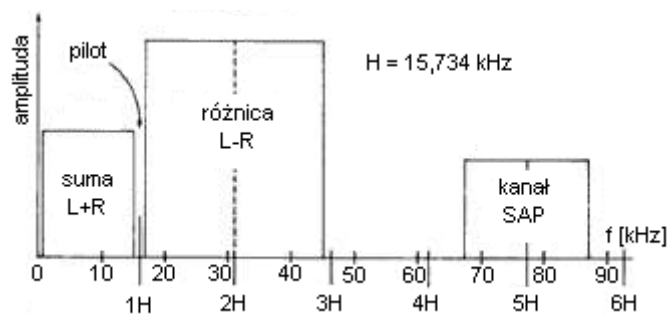
W NTSC przesyłanych jest 525 linii obrazu przy 29,97 ramkach na sekundę. Ilość ta wzięła się stąd, że w Stanach Zjednoczonych sieć energetyczna pracuje na częstotliwości 60 Hz, stąd aby uniknąć zakłóceń przydźwięku, wybrana została częstotliwość odświeżania bliska sieci, czyli 59,94 Hz. Składowe chrominancji (kolor czerwony i niebieski) przesyłane są na jednej podnośnej poprzez amplitudową modulację kwadraturową. Powoduje to nierzadkie problemy z odwzorowaniem koloru, jeśli tylko sygnał zostanie zakłócony i jego faza zmieni się. Problemy te nie występują w systemie PAL, gdzie faza wpływa na nasycenie, nie barwę odbieranego obrazu [28].

2.3.1 System MTS

MTS (Multichannel Television Sound) to system przesyłania dźwięku stereofonicznego w standardzie telewizyjnym NTSC. Pozwala on na przesłanie dźwięku w pasmie 50 Hz – 15 kHz w wersji stereofonicznej z możliwością dołączenia dodatkowych kanałów audio (np. innych wersji językowych) [27]. Jako podstawową częstotliwość podnośnej wykorzystuje się częstotliwość odchyłania poziomego obrazu, 15,734 kHz. W związku z tym faktem, o kodowaniu dźwięku w systemie MTS świadczy obecność pilota na tej częstotliwości w sygnale fonicznym. Oryginalna zawartość audio jest filtrowana w podanym wyżej przedziale pasma, tak by nie interferować z dodawanym pilotem, który podczas transmisji traktowany jest jako zwykły sygnał akustyczny, a podlega odfiltrowaniu dopiero w odbiorniku telewizyjnym. Widmo pojedynczego kanału przedstawione jest schematycznie na Rys. 25. Jak widać, zachowana jest kompatybilność wstecz – w podstawowym pasmie znajduje się sygnał sumy mono obu kanałów stereo. Sygnał różnicowy zajmuje podwojone pasmo ze względu na zastosowanie systemu odszumiania dbx [26]. Gdyby nie dbx, taki sposób kodowania byłby identyczny ze zwykłym stereofonicznym FM.

2. Podstawy teoretyczne

Kanał SAP (Second Audio Program) to dodatkowy kanał, w którym może znaleźć się np. inna ścieżka językowa [25].



Rysunek 25: Widmo pojedynczego stereofonicznego kanału audio zakodowanego w systemie MTS

3. Zniekształcenie Wow

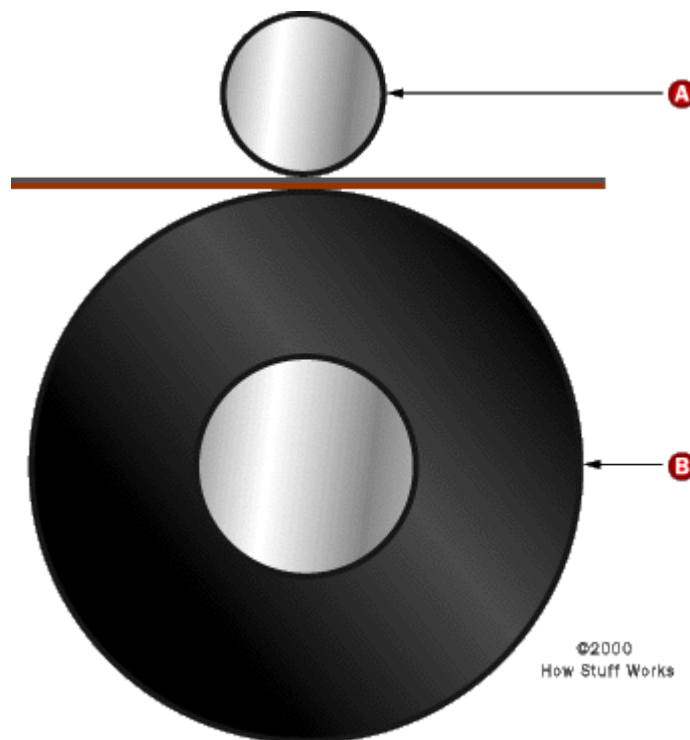
Nieodłącznym elementem toru fonicznego były i będą mechaniczne urządzenia rejestrujące, a także same nośniki dźwięku. O ile dziś większość zakłóceń i zniekształceń powstaje w elektrycznej części toru, gdzie sygnał dźwiękowy występuje pod postacią fali elektromagnetycznej, to niegdyś podstawowym problemem były czynniki fizyczne. Przesunięcie horyzontalne szczeliny przy naświetlaniu taśmy podczas zapisu optycznego dźwięku powodowało w najlepszym przypadku powstanie przesterowań [20], a już w niewielkim stopniu niewłaściwe ustawienie skosu głowicy magnetycznej powoduje efekt niechcianej filtracji dźwięku [3]. Wszelkiego rodzaju brak precyzji czy stabilności pracy urządzeń i wykonania nośników powoduje niechciany efekt różnego rodzaju filtracji, modulacji i zmian dynamiki. Do zniekształceń, które czasem ciężko zidentyfikować (dla niewprawnego ucha mogą być nieuchwytnie) należą **wow i flutter** (dla przypomnienia, polskie drżenie i kołysanie dźwięku). Ich występowanie w parze jest podyktowane **tą samą naturą samego zniekształcenia – jest to modulacja częstotliwości**. Rozróżnia je tylko częstotliwość tej modulacji. **Przyjęło się, że ze zniekształceniem wow mamy do czynienia, gdy wynosi ona od 0.1÷0.5 do 4÷10 Hz, powyżej zaś, do około 100 Hz występuje flutter. Częstotliwości graniczne są różne w zależności od źródeł**. Jeśli chodzi o tę dolną, to powszechnie przyjętą minimalną częstotliwością jest 0.1 Hz, z tym że niektóre źródła [24] wyróżniają jeszcze zniekształcenie **plywania dźwięku** (ang. *drift*) – bardzo powolne zmiany częstotliwości do 0.5 Hz. Z granicą górną efektu wow jest już nieco inaczej. Sama **Wikipedia nie jest spójna**. Angielska jej wersja [33] podaje, że podział na drżenie i kołysanie jest dość anachroniczny, ale że częstotliwością graniczną były 4 Hz (i częstotliwość ta, w wypadku modulacji jest najbardziej nieprzyjemna w percepcji), z kolei holenderska edycja [38] za granicę podaje 10 Hz. Wspominany już standard AES [24] z kolei podaje wartość 6 Hz. W literaturze [24] pojawia się jeszcze termin *scrape flutter* na oznaczenie modulacji powyżej 100 Hz, jednak percypowany jest on jako zwykły szum.

3.1 Przyczyny powstawania zniekształcenia

Na wstępie trzeba zaznaczyć, że zniekształcenie ~~drżenia~~ powstaje tylko podczas pracy

3. Zniekształcenie Wow

z dźwiękiem analogowym i dotyka on nagrań gramofonowych i magnetycznych. W urządzeniach winę za wprowadzenie zniekształcenia mają najczęściej silniki do przesuwu taśmy/obrotu płyty. Ich nierównomierna praca owocuje powstaniem aperiodycznych najczęściej modulacji. Dopuszczalne normy nierównomierności przesuwu taśm dla przykładu wynoszą $\pm 0.5\%$ dla zwykłych magnetofonów kasetowych i $\pm 0.2\%$ dla ich profesjonalnych odpowiedników [19]. Biorąc pod uwagę fakt, że typowy nowoczesny sprzęt ma wartość nierównomierności wynoszącą około 0.08% [32], widać że przyczyny powstawania problemu modulacji częstotliwości zostały w obecnych czasach znacznie zredukowane. Nieprzekroczenie tych wartości owocuje modulacją o na tyle niewielkiej głębokości, że nie jest ona szkodliwa. Trudno bowiem mówić o zapewnieniu idealnej stabilności pracy urządzeń mechanicznych takich jak silnik elektryczny, jeśli wpływ na prędkość jego pracy mogą mieć czynniki i siły zewnętrzne. Oprócz tego, na przesuw taśmy mają elementy mechaniczne takie jak rolka i wałek przesuwu (Rys. 26). Należy zauważyć, że nie mogą one być nigdy idealnie okrągłe ani zamontowane dokładnie pod kątem prostym, dlatego też nie obracają się one ze stałą prędkością liniową. Przy każdym ich obrocie powstają zniekształcenia częstotliwości – dla rolki są one wolnozmienne (wow), dla wałka – szybkozmienne (flutter).



Rysunek 26: Wałek (A) i rolka (B) w magnetofonie

3. Zniekształcenie Wow

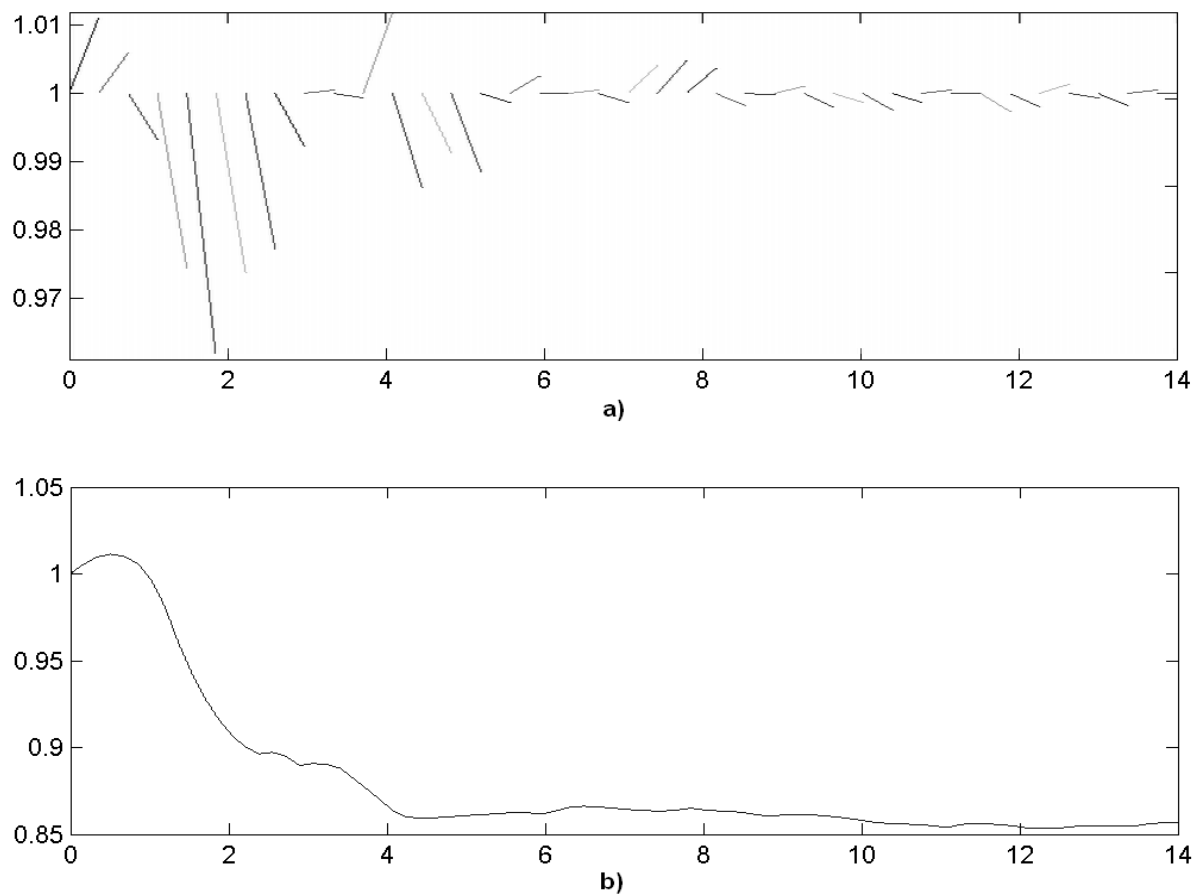
Nieidealności nośników także przekładają się na zakłócenia. W przypadku kaset są to najczęściej nieidealnie okrągłe rolki, na które nawija się taśma, ale także wady konstrukcji powodujące zwiększone tarcie, które spowalnia przesuw taśmy. Jeśli mamy z nim do czynienia na elementach obrotowych kasety, może ono pojawiać się okresowo (raz na pełen obrót) i owocować ~~fluktuacją~~ częstotliwości. Dodatkowo zużyta (bądź wadliwa) perforacja taśmy filmowej może być źródłem cyklicznych zaburzeń prędkości przesuwu (naciąganie i rozluźnianie taśmy przez traktor).

W przypadku płyt gramofonowych, prędkość odczytu zależy od **położenia rowka względem środka** – im jest on dalszy, tym prędkość liniowa większa. Związane jest to ze stałą prędkością kątową mechanizmu napędzającego. Jeśli zatem ~~dziura~~ w płycie nie znajduje się idealnie na jej środku, będziemy mieli do czynienia ze spowalnianiem (gdy odległość się zmniejszy) i z przyspieszaniem (gdy się zwiększy) [32].

3.2 Wybrane metody estymacji charakterystyki zniekształcenia

~~Powyższe~~ przyczyny pokazują, jak różny może być charakter wprowadzanych modulacji. Ich głębokość, a także i ~~rodzaj~~ jest bardzo trudno do zamodelowania. Ponadto wszelkie hipotetyczne modele miałyby bardzo wąskie zastosowanie, zważywszy na to, że wynikowe zniekształcenie może być sumą pomniejszych zniekształceń wprowadzonych przez różne elementy toru, które dodatkowo miałyby na siebie wpływ. Zniekształcenie estymujemy zatem poprzez analizę nie toru, lecz samego sygnału. Wykorzystujemy do tego krzywą ~~zmian~~ ~~częstotliwości~~ określaną skrótem PVC (~~Pitch Variation Curve~~). Informuje ~~nas~~ ona o stosunku częstotliwości aktualnej do nominalnej, zatem wartość 1 oznacza brak zniekształcenia, 0.5 oznacza częstotliwość zmniejszoną dwukrotnie, 3 – trzykrotnie zwiększoną. ~~Często~~ uzyskuje się ją obserwując lokalne zmiany częstotliwości (tj. badając zmiany częstotliwości sygnału co pewien krok i budując na ich podstawie wynikową PVC). **Rozróżniamy zatem lokalną i globalną PVC**, z tym że interesująca dla nas jest ostateczna, czyli druga z nich. **Różnica między nimi jest taka, że lokalna krzywa opisuje zmiany względne, krzywa globalna – bezwzględne.** ~~Z tego powodu przejście między nimi jest proste – wersja globalna jest kumulowaną wersją krzywej lokalnej, zakładając że pierwszą jej wartością będzie 1. Innymi słowy, każda kolejna wartość dodawana jest do sumy poprzednich. Na Rys. 27 zobrazowane są różnice między nimi.~~

3. Zniekształcenie Wow



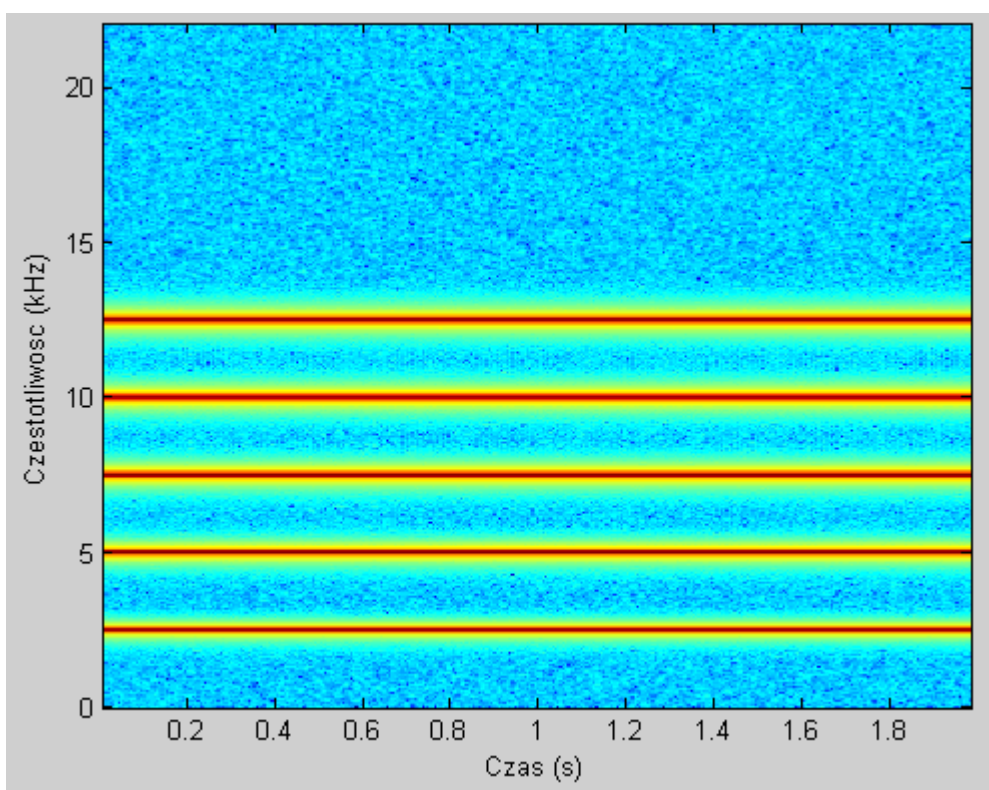
Rysunek 27: Krzywe PVC (na osi czasu): a) lokalna; b) globalna [5]

3.2.1 Metoda śledzenia składowych tonalnych

~~Podstawowym sposobem estymacji charakterystyki zniekształcenia częstotliwości jest test odsłuchowy.~~ Nie jesteśmy jednak w stanie na jego podstawie pozyskać żadnych liczbowych informacji na temat zniekształcenia. Odpowiednikiem tego intuicyjnego podejścia jest obserwacja sygnału w dziedzinie czasowo-częstotliwościowej, tj. przez wykorzystanie spektrogramu sygnału. W przypadku dźwięków muzycznych, tzn. mających komponenty harmoniczne (tonalne), możemy zauważyć zniekształcone fragmenty jako oscylacje poszczególnych komponentów (proporcjonalne do częstotliwości danej składowej). Spektrogram dźwięku harmonicznego o stałej częstotliwości w ogólności składa się z n równoległych linii prostych (ich liczba to liczba harmonicznnych tworzących sygnał). W

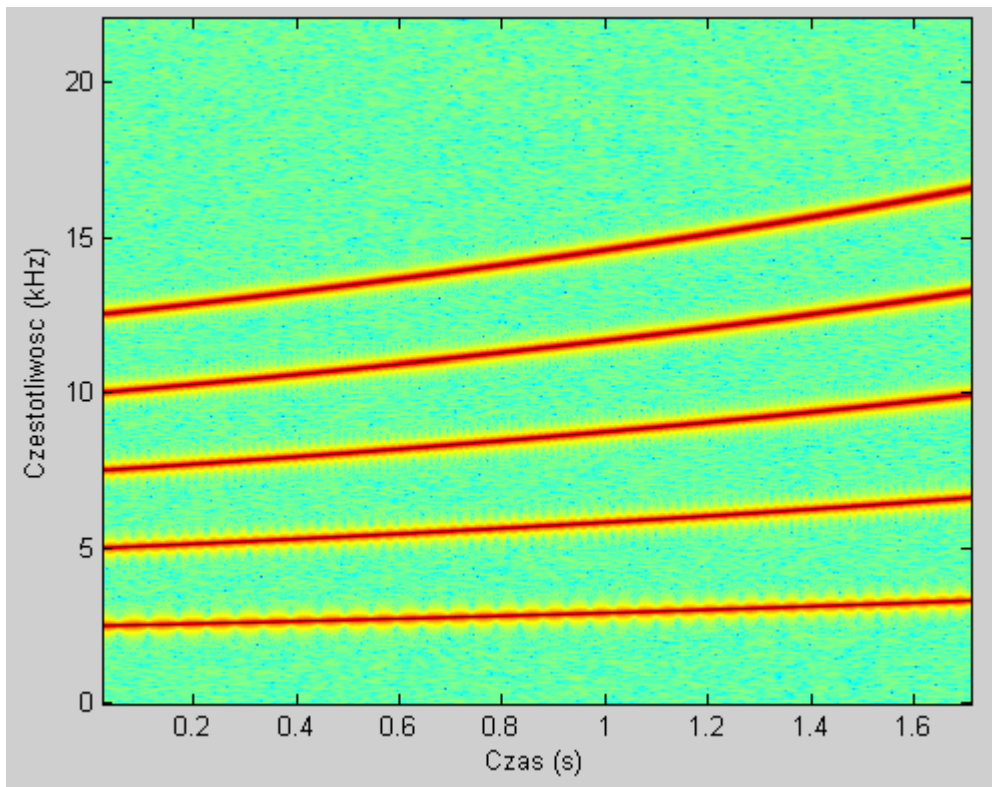
3. Zniekształcenie Wow

przypadku gdy mamy do czynienia z modulacją częstotliwości, każda ścieżka reprezentująca pojedynczą harmoniczną sygnału, proporcjonalnie zmienia swoje położenie. Wykresy na Rys. 28-30 pokazują spektrogramy podstawowego tonu o częstotliwości 2500 Hz wraz z 4 harmonicznymi (odpowiednio 5, 7.5, 10, 12.5 kHz) w przypadku braku modulacji (Rys. 28), przy monotonicznie narastającej wysokości dźwięku (na skali półtonów – stąd odchylenie od linii prostej widoczne szczególnie w przypadku najwyższej składowej pod koniec trwania fragmentu – Rys. 29) oraz przy stałej modulacji o głębokości 200 Hz i częstotliwości 6 Hz (Rys. 30) .

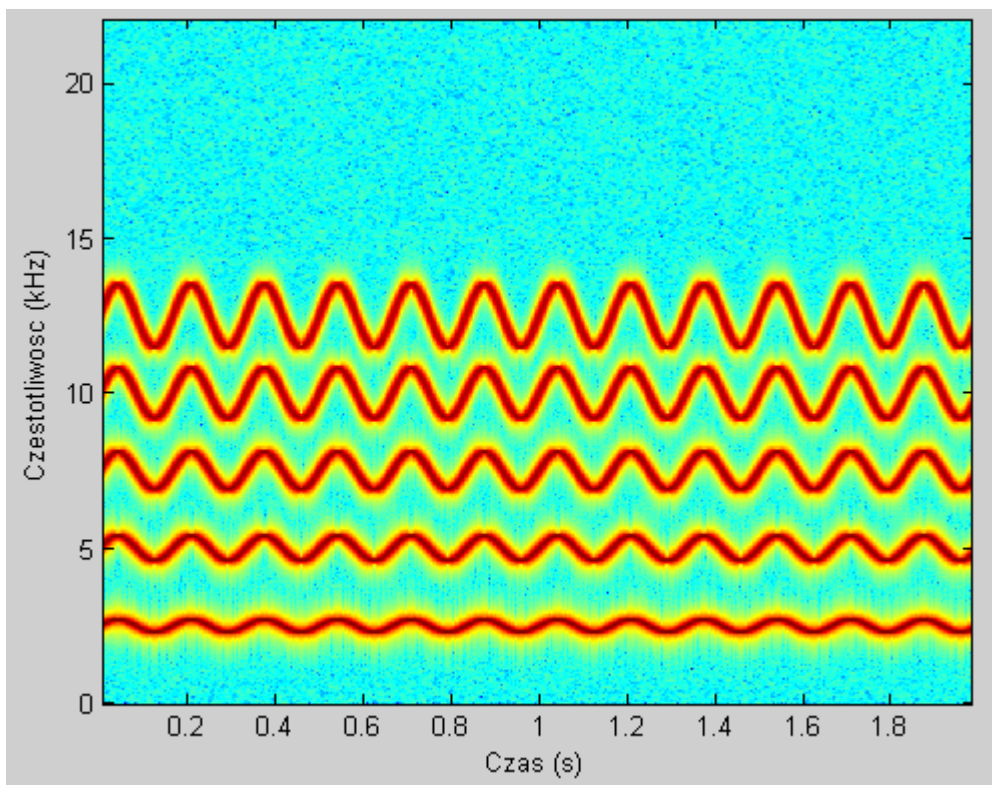


Rysunek 28: Ton 2,5 kHz z czterema harmonicznymi bez modulacji częstotliwości

3. Zniekształcenie Wow



Rysunek 29: Ton 2,5 kHz z czterema harmonicznymi o narastającej w czasie częstotliwości

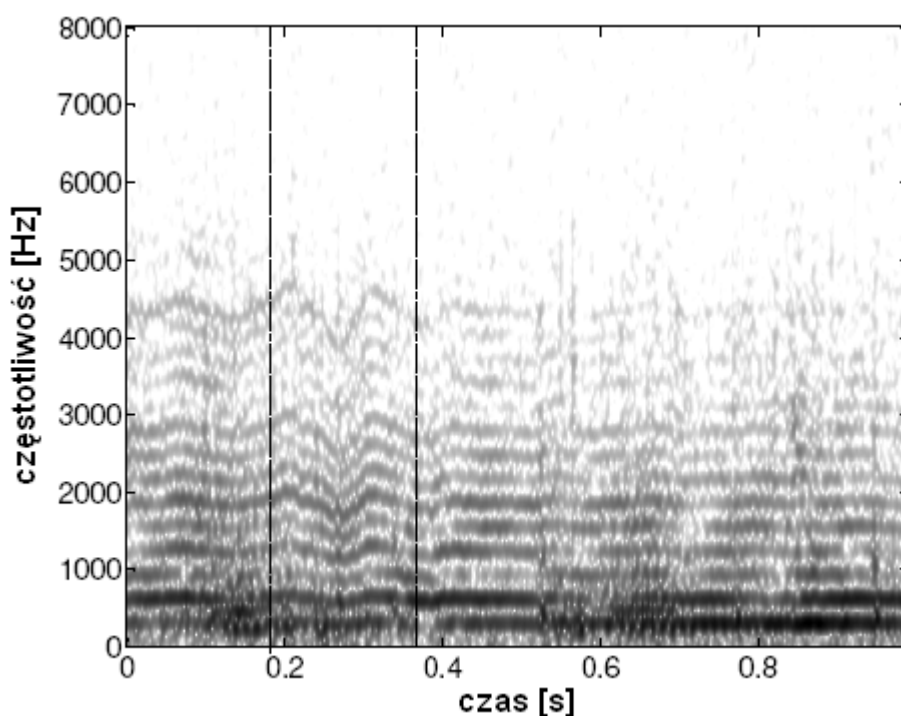


Rysunek 30: Ton 2,5 kHz z czterema harmonicznymi ze stałą modulacji częstotliwości

3. Zniekształcenie Wow

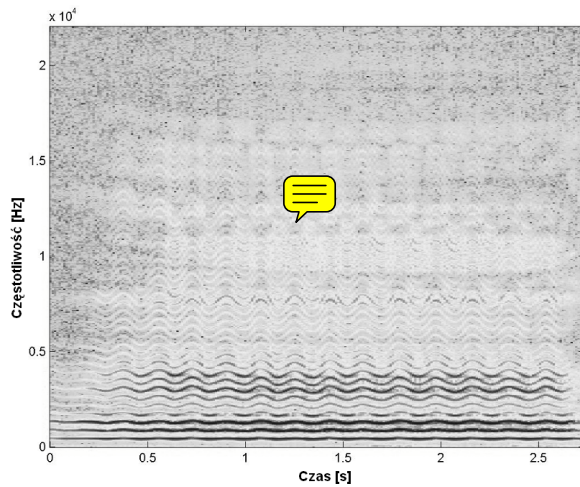
Widoczny na wykresie z Rys. 29 efekt „koralików” dla podstawowej składowej tonalnej wynika z powolnych zmian częstotliwości, które prowadzą do okresowych wycieków widmowych – wtedy, kiedy częstotliwość składowej znajduje się między sąsiadującymi pasmami częstotliwości uzyskanymi przy pomocy FFT. Jeśli porównać przebieg tak zniekształconej ścieżki harmonicznego do fali stojącej, to węzły reprezentują nam miejsca precyzyjnego określenia częstotliwości, strzałki zaś – momenty największego wycieku.

Spektrogram harmonicznego dźwięku naturalnego znajduje się na Rys. 31. **Liniami pionowymi został na nim oznaczony fragment ze zniekształceniem wow.** Należy tutaj zwrócić uwagę, że występowanie zamierzonych efektów modulujących częstotliwość typu vibrato może być na poziomie spektrogramu widziane jako zniekształcenie. Z tego powodu ważna jest ocena słuchowa eksperta, która pozwala wyeliminować błędy w rodzaju fałszywego rozpoznania. Zostały także stworzone algorytmy [6] do automatycznego wykrywania naturalnych zmian częstotliwości. Bazują one na tym, że w przypadku efektów typu vibrato stosowanych w śpiewie, nie wszystkie formanty podlegają modulacji. Widać to na Rys. 32 i 33 na kolejnej stronie.

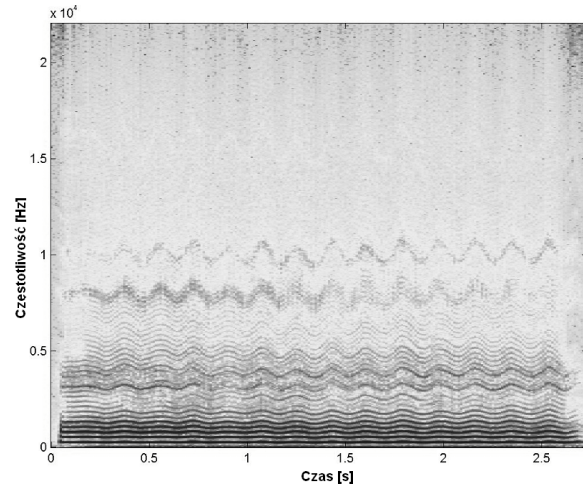


Rysunek 31: Spektrogram fragmentu nagrania ze zniekształceniem wow (0,2 - 0,4 s)

3. Zniekształcenie Wow

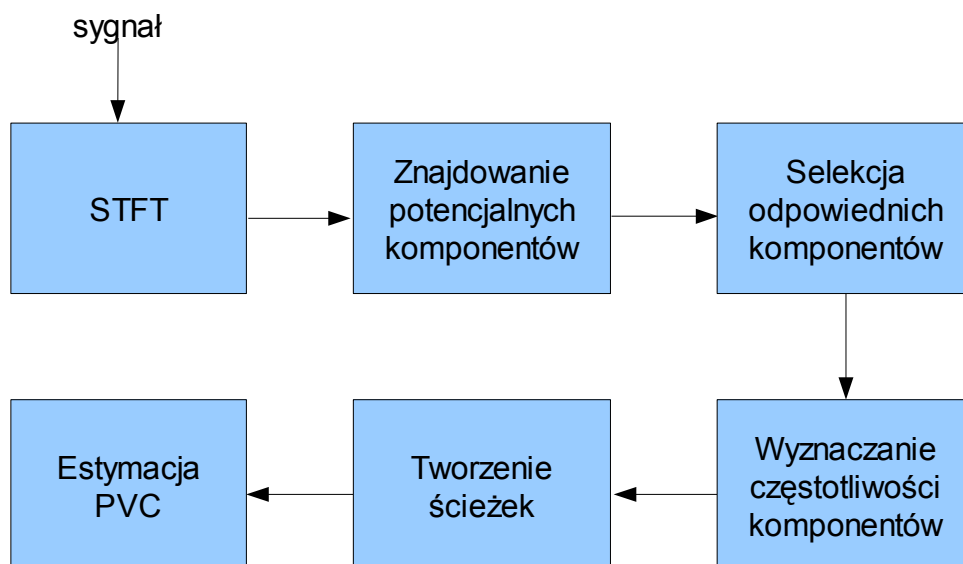


Rysunek 32: Spektrogram samogłoski 'a' z efektem vibrato



Rysunek 33: Spektrogram samogłoski 'a' ze zniekształceniem wow

Ze spektrogramów widać, że w sytuacji sprzyjającej, komponenty tonalne odznaczają się względnie wysokim poziomem na tle szumu. Zagadnienie ich śledzenia polega na znajdowaniu komponentów i łączeniu ich w ścieżki, za pomocą których dokonuje się estymacji PVC, która z kolei służy do usunięcia zniekształcenia. Podstawy metody stworzyli McAulay i Quatrieri [16], dlatego w literaturze nazywa się ją analizą MQ. Podstawowy schemat blokowy omawianej metody przedstawiony jest na Rys. 34.



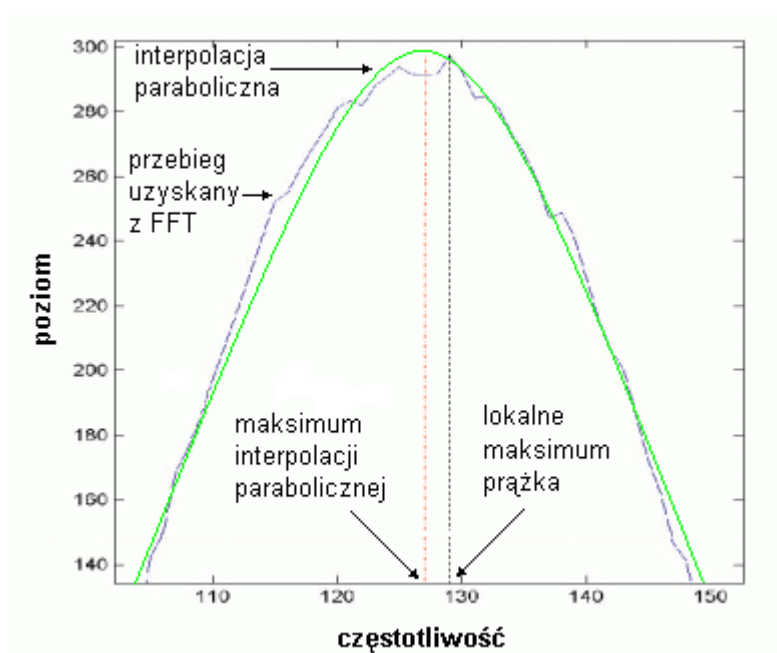
Rysunek 34: Schemat blokowy algorytmu do śledzenia komponentów tonalnych

3. Zniekształcenie Wow

Sygnał na wejściu algorytmu zostaje poddany operacji STFT w celu uzyskania jego spektrogramu. Na podstawie analizy kolejnych widm znajduje się potencjalne komponenty tonalne. Brane są pod uwagę lokalne maksima widma. Aby odróżnić prążki główne (wyznaczające faktyczną częstotliwość komponentu) do prążków bocznych (powstałych jako efekt wycieku widma) i składowych szumowych, stosuje się różne kryteria. Tym najprostszym i najbardziej intuicyjnym jest kryterium energetyczne – mianowicie tylko składowe o odpowiednim poziomie energii zostają zakwalifikowane do dalszej analizy. Sprawdza się ono jednak tylko w przypadku sygnałów o znacznym odstępnie sygnału od szumu, gdzie harmoniczne są wyraźne, czyli w sytuacji rzadko spotykanej w zaszumionych i na różne sposoby zniekształconych nagraniach archiwalnych. Często proponowane metody [12][15][18] choć bardziej skomplikowane, nie dają zadowalających rezultatów na próbkach rzeczywistych. Łukasz Litwic i Przemysław Maziewski z Katedry Systemów Multimedialnych (KSM) wydziału Elektroniki, Telekomunikacji i Informatyki (ETI) Politechniki Gdańskiej (PG) przedstawili metodę bazującą na 3 niezależnych kryteriach: na mierze podobieństwa sinusoidalnego (SLM – Sinusoidal Likeness Measure), pomiarze fazy i progowaniu poziomu względnego amplitudy [13].

Wyznaczanie częstotliwości poszczególnych komponentów jest o tyle ważne, że od tego zależeć będzie później jakość poprawy zniekształconego nagrania. Zastosowanie STFT zamiast zwykłego DFT powoduje zwiększenie rozdzielczości (spowodowane jest to stosowaniem wypełniania zerami), jednak przedział częstotliwości, dla którego występuje maksimum energii może często być przypadkowy, a różnica jednego indeksu wprowadza skokową zmianę kilkudziesięciu lub więcej Herców. Przykładowo dla okna o długości 1024 próbki i częstotliwości próbkowania 44100 Hz, rozdzielczość transformaty to ponad 43 Hz. Stosuje się zatem różne metody dalszego zwiększania szacowania częstotliwości przy zachowaniu stałej rozdzielczości częstotliwościowej. Najbardziej spotykaną metodą jest interpolacja paraboliczna (Rys. 35). Polega ona na braniu pod uwagę sąsiedztwa prążka, dla którego występuje maksimum energii w widmie. Na jego podstawie interpoluje się maksimum leżące już gdzieś między przedziałami, co przekłada się na wzrost dokładności obliczeń. Najczęściej bierze się pod uwagę po jednej dodatkowej próbie z każdej strony ze względu na niewielką złożoność obliczeniową metody (jedno równanie). Na rysunku poniżej pokazano celowość stosowania metody i różnicę między maksimum danego prążka a faktyczną częstotliwością, jaką reprezentuje.

3. Zniekształcenie Wow



Rysunek 35: Ilustracja działania interpolacji parabolicznej w analizie częstotliwości prądka

Inną z metod do lepszej estymacji częstotliwości danego komponentu tonalnego jest metoda przestawiania czasowo-częstotliwościowego (ang. *time-frequency reassignment*) [1][9].

Gdy już określone zostaną częstotliwości wszystkich zakwalifikowanych komponentów tonalnych, tworzone są na ich podstawie ścieżki. Ogólna koncepcja polega na śledzeniu zmian częstotliwości w czasie poszczególnych komponentów zaczynając od najsilniejszego (jako najbardziej znaczące w widmie i najmniej podatne na zakłócenia trajektorii), kończąc zaś na najsłabszych. Sposoby, w jaki jest to dokonywane może być różne, ~~jednak nieznacznie~~ [16][13]. Rozróżnia się 4 podstawowe stany ścieżek. Stan pierwszy to moment narodzin, czyli wykrycie komponentu tonalnego nie pasującego do żadnej z poprzednich ścieżek (bądź po prostu pierwszego komponentu w analizie). Kolejny stan to normalne śledzenie zmian komponentu. Trzeci stan to moment zaniku komponentu (nazywany w literaturze stanem „zombie”), kiedy to z powodu permanentnego lub tymczasowego zaniku w kolejnym kroku nie zostaje znaleziony komponent będący kontynuacją istniejącej ścieżki. W przypadku znalezienia odpowiedniego komponentu w kolejnych oknach, brakujące fragmenty zostają interpolowane. W przypadku dłuższego braku kontynuacji, stan „zombie” przechodzi w stan śmierci ścieżki.

Podobnie jak w przypadku szukania pojedynczych komponentów, także stworzone już ścieżki

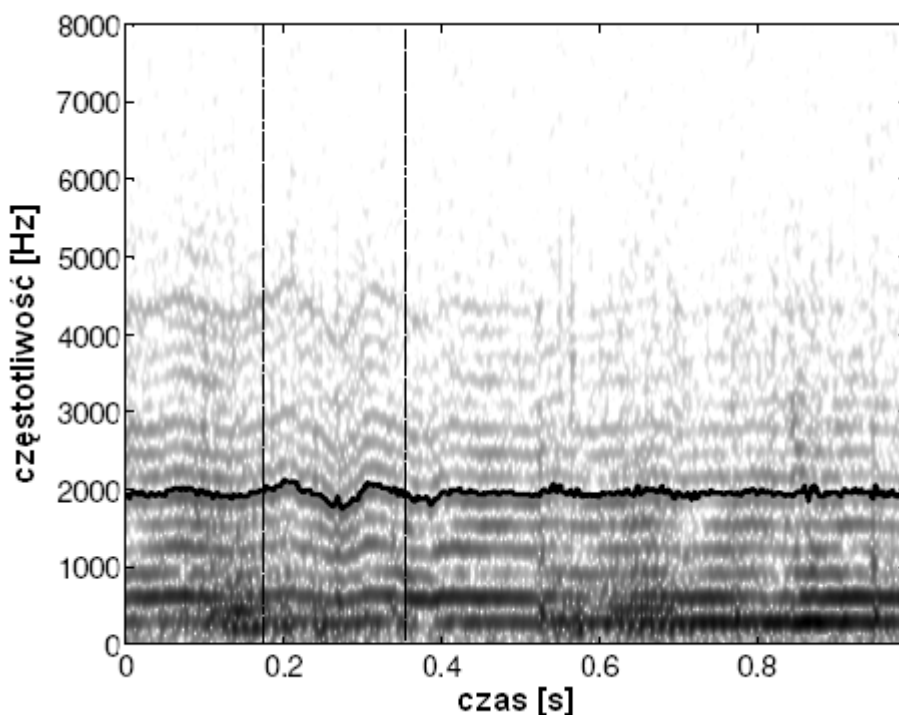
3. Zniekształcenie Wow

podlegają selekcji. Te, które są zbyt krótkie lub wybrakowane (częste stany „zombie”) przestają być brane pod uwagę ze względu na swoją domniemaną niewielką przydatność.

Ostatnim etapem jest stworzenie krzywej PVC na podstawie wyselekcjonowanych ścieżek tonalnych. Może się to odbywać poprzez wybranie jednej (w opinii eksperta lub w wyniku działania algorytmu), jednak sposób ten jest mało bezpieczny ze względu na duże prawdopodobieństwo wystąpienia przypadkowego błędu estymacji częstotliwości, co zmniejszy efektywność usuwania zniekształceń. Dlatego można stosować analizę zbiorczą, tj. stworzyć lokalne krzywe PVC dla każdej z zakwalifikowanych ścieżek, a następnie za pomocą filtracji medianowej i uśredniania uzyskać obarczony statystycznie mniejszym błędem przebieg globalnej PVC, która bezpośrednio posłuży do usunięcia zniekształcenia wow [6].

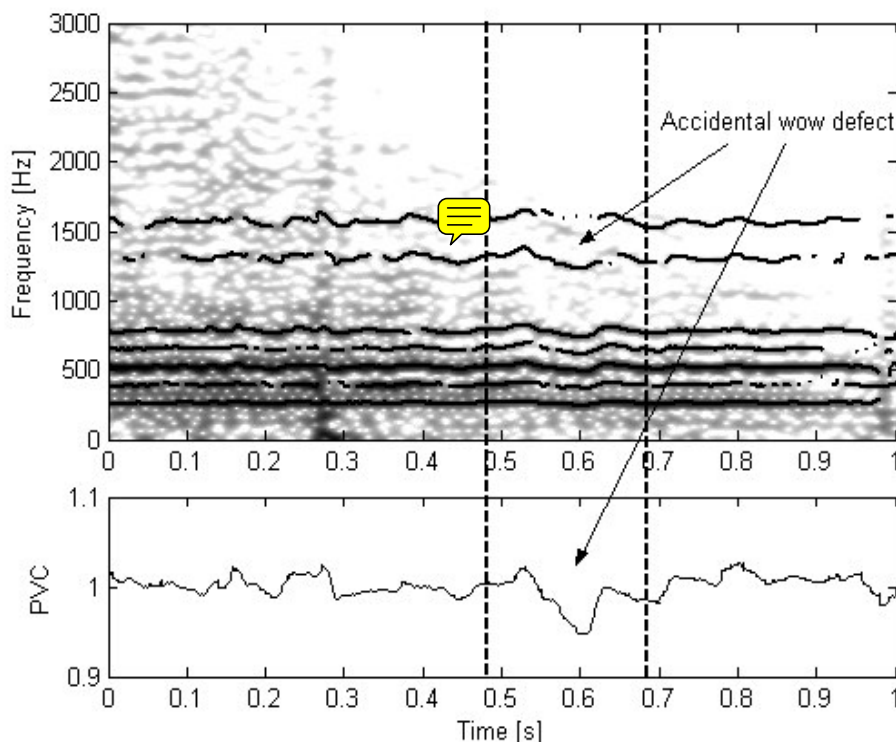
Na Rys. 36 przedstawiono spektrogram z Rys. 31 wraz z zaznaczoną wybraną ścieżką na podstawie której obliczona została krzywa PVC.

Dalej na Rys. 37 znajduje się spektrogram z wyróżnionymi ścieżkami, które zostały zakwalifikowane do dalszych obliczeń i na podstawie których pozyskana została globalna PVC.



Rysunek 36: Spektrogram fragmentu nagrania z zaznaczoną ścieżką komponentu tonalnego

3. Zniekształcenie Wow

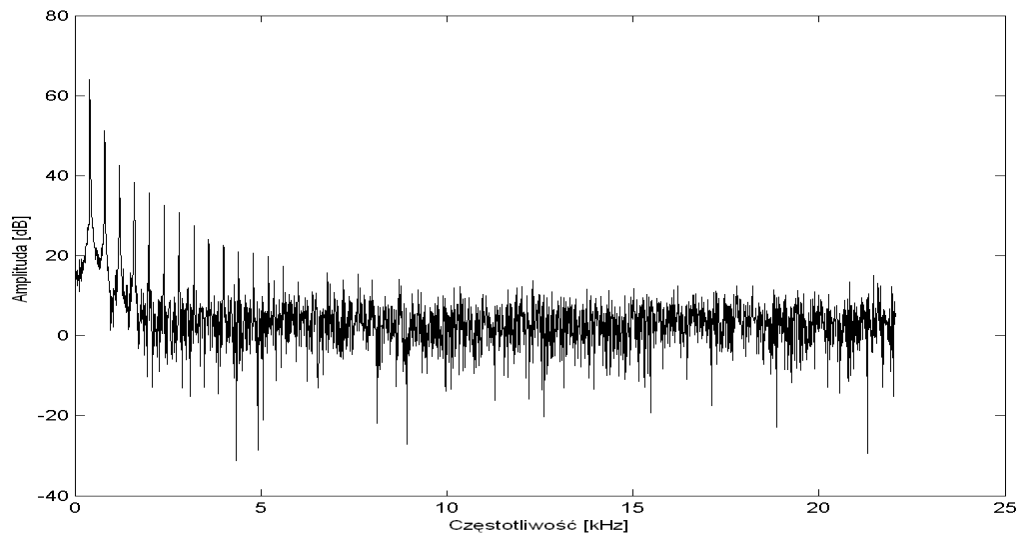


Rysunek 37: Spektrogram z wybranymi ścieżkami (na górze) i obliczoną na ich podstawie krzywą PVC (na dole)

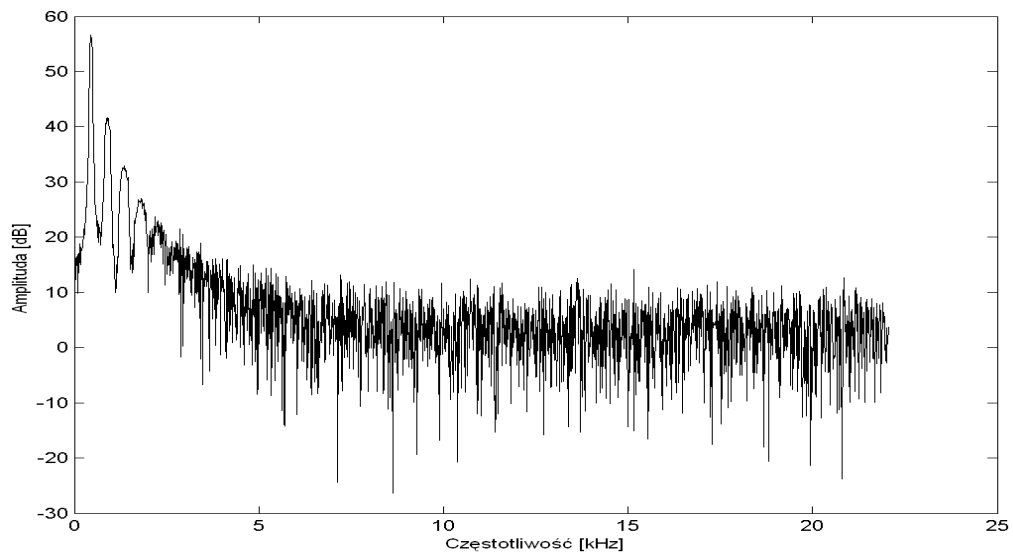
3.2.2 Metoda autokorelacyjna

Jak zostało poprzednio wykazane, osiągnięcie zadowalającej rozdzielczości częstotliwościowo-czasowej może przysporzyć problemy w wypadku modulacji częstotliwości sygnału (z którą mamy do czynienia w przypadku zniekształcenia wow). Problemy te stają się tym bardziej widoczne, im większa jest częstotliwość modulująca. Występuje potrzeba stosowania krótkich okien czasowych, co zaburza nam dokładność oceny częstotliwości chwilowej danego komponentu spektralnego. Na Rys. 38-40 przedstawiono wykresy widma STFT trzech sygnałów harmoniczných: o stałej częstotliwości podstawowej równej 400 Hz (Rys. 38), o częstotliwości podstawowej zmiennej w zakresie 400÷500 Hz (Rys. 39) oraz zmiennej w zakresie 400÷600 Hz (Rys. 40). Widać z nich wyraźnie, że o ile częstotliwość podstawowa jest w każdym przypadku wyraźna, to kolejne harmoniczne ulegają rozmyciu i nie jest możliwe rozróżnienie ich od szumu.

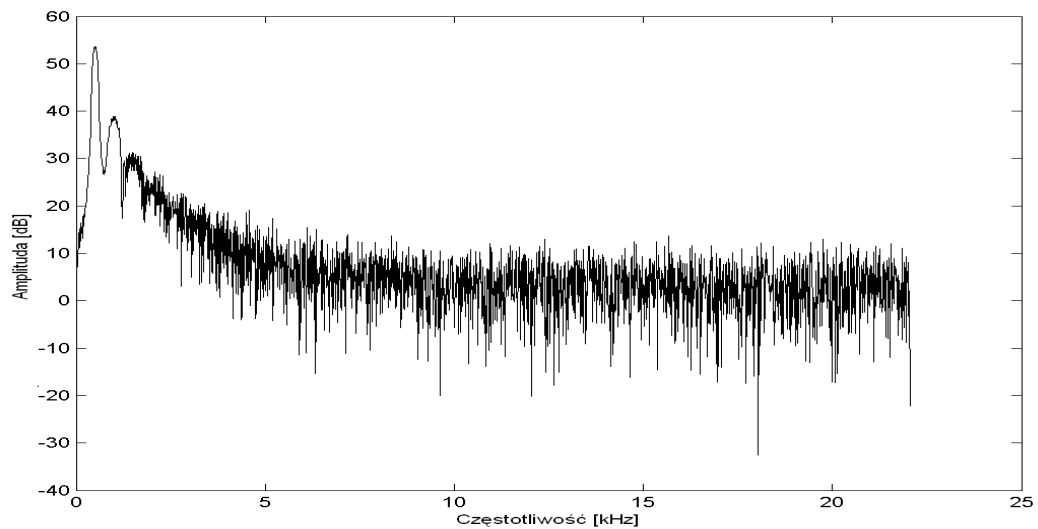
3. Zniekształcenie Wow



Rysunek 38: Widmo sygnału harmonicznego o $f_p=400\text{Hz}$



Rysunek 39: Widmo sygnału harmonicznego o $f_p=400\div 500\text{Hz}$



Rysunek 40: Widmo sygnału harmonicznego o $f_p=400\div 600\text{Hz}$

3. Zniekształcenie Wow

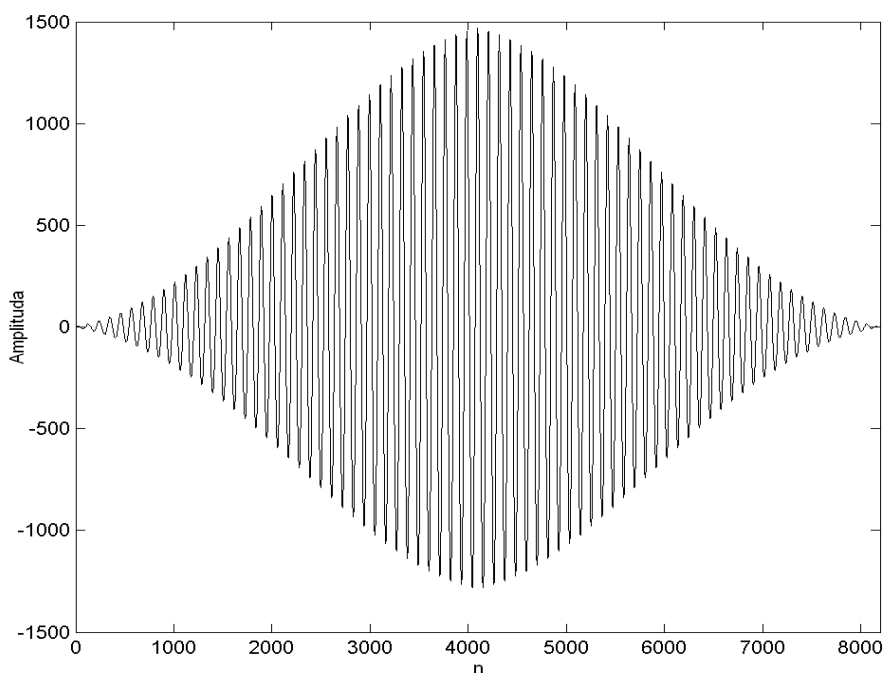
Dziedzina widma wydaje się być zatem kłopotliwa przy ocenianiu szybkich zmian częstotliwości składowych harmonicznego sygnału. Do jednej z podstawowych metod określania częstotliwości podstawowej w dziedzinie czasu należy metoda autokorelacyjna. Polega ona na analizie funkcji autokorelacji wydzielonego okna czasowego sygnału. Funkcja autokorelacji dana jest wzorem:

$$r[n] = \sum_{m=1}^M x[m] \cdot x[m+n] \quad (15)$$

gdzie: M to długość okna czasowego (w próbkach),

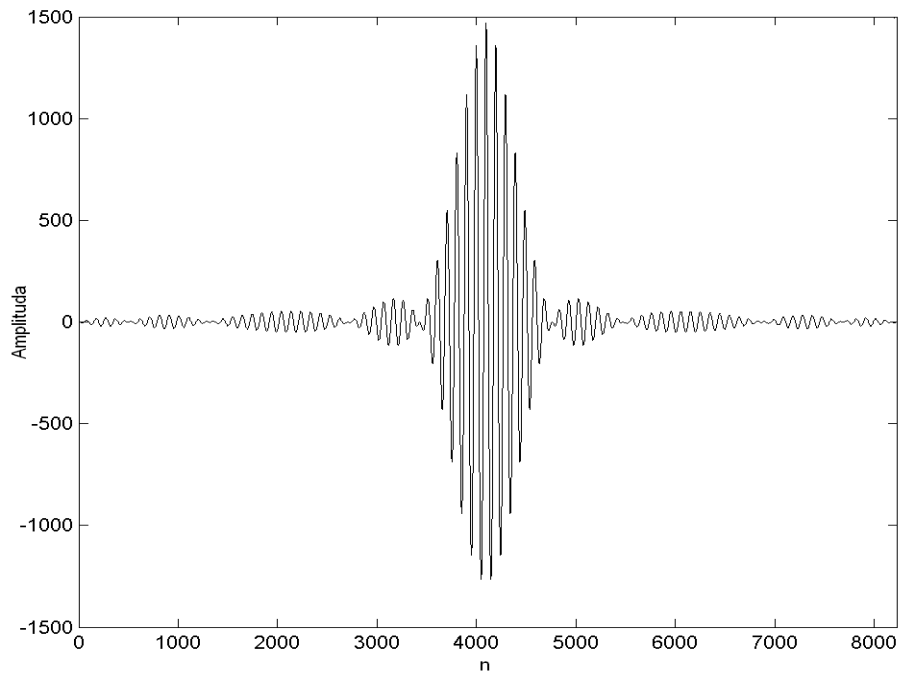
$x[]$ to badany sygnał

Dla fragmentów sygnałów przedstawionych na spektrogramach z Rys. 38-40 obliczona została funkcja autokorelacji, a jej wykresy przedstawiono na Rys. 41-43.

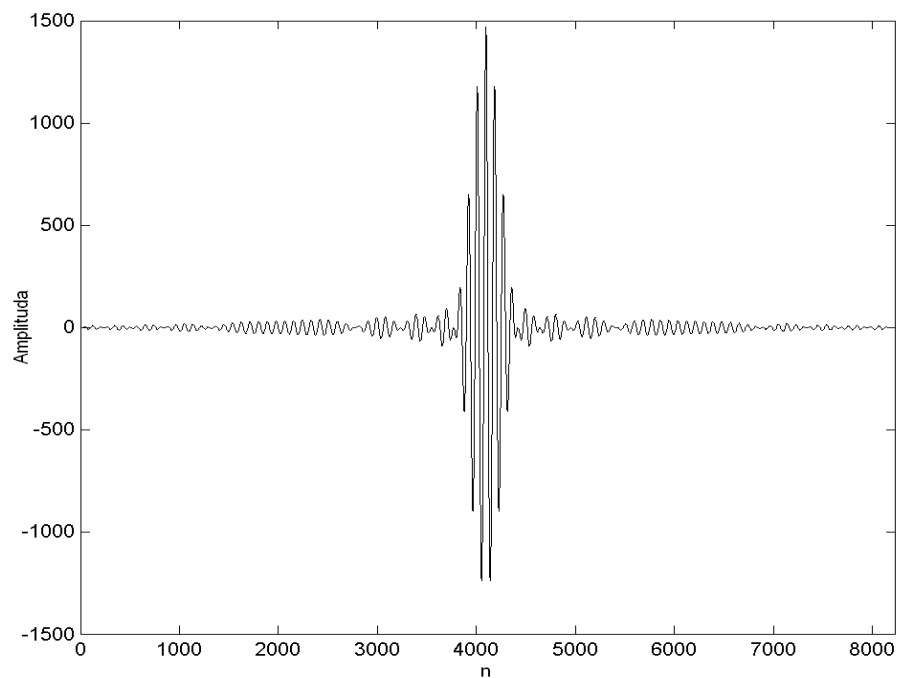


Rysunek 41: Funkcja autokorelacji sygnału harmonicznego o $f_p=400\text{Hz}$

3. Zniekształcenie Wow



Rysunek 42: Funkcja autokorelacji sygnału harmonicznego o $f_p=400\div 500\text{Hz}$



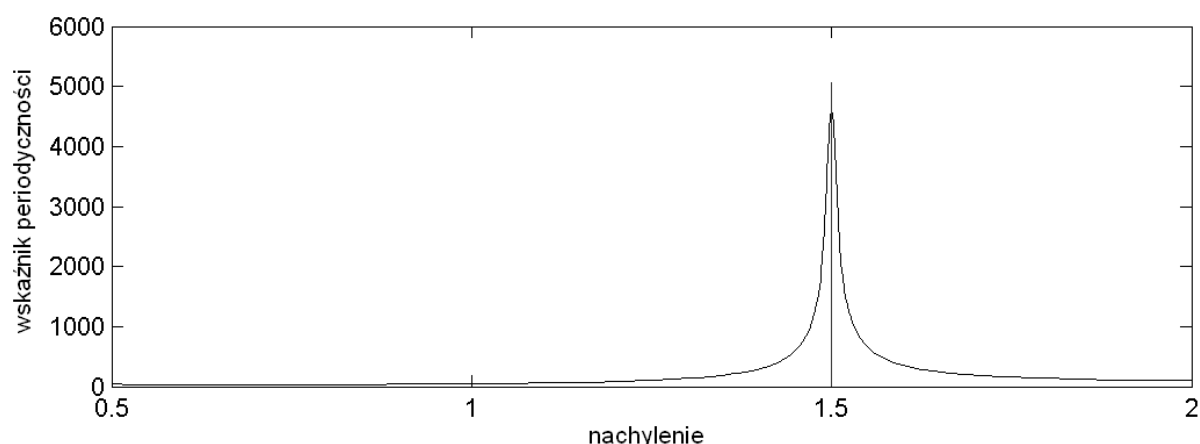
Rysunek 43: Funkcja autokorelacji sygnału harmonicznego o $f_p=400\div 600\text{Hz}$

Ujawnia się tutaj pewna ciekawa cecha autokorelacji sygnałów o zmodulowanej częstotliwości. Mianowicie im głębsza modulacja, tym różnica między lokalnymi maksimami i minimami funkcji autokorelacji opada szybciej przy jednoczesnym niezmiennym poziomie maksimum globalnego (prążek centralny). Na tej podstawie można wprowadzić parametr

3. Zniekształcenie Wow

wskaźnika periodyczności p , którego najprostszą realizacją byłaby znormalizowana w stosunku do maksimum globalnego wartość średnia modułu funkcji autokorelacji. W ten sposób sygnały idealnie harmoniczne będą miały wysokie p , podczas gdy sygnały których okresowość została zaburzona – wprost przeciwnie.

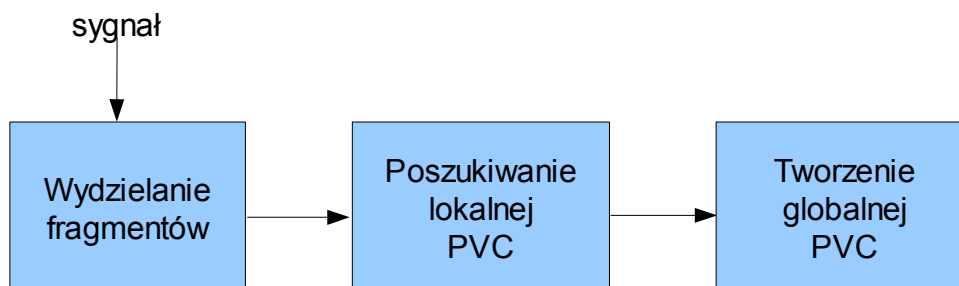
Mając do dyspozycji parametr określający periodyczność sygnału, możemy określić krzywą PVC dla danego nagrania robiąc to niejako „od końca”. Weźmy dwa sąsiadujące krótkie fragmenty sygnału, dla których mamy do czynienia z różnicą częstotliwości pomiędzy jednym a drugim. Poszukujemy lokalnej PVC która określa nam stopień zmiany tej częstotliwości. Załóżmy, że jest ona znana – wtedy przepróbkowanie sygnału zgodnie z lokalną PVC zaowocuje kompensacją zniekształcenia częstotliwości pomiędzy jednym a drugim fragmentem. W takim wypadku wskaźnik periodyczności fragmentu powstałego z dwóch analizowanych fragmentów będzie miał maksymalną możliwą wartość (w każdym innym przypadku periodyczność będzie mniejsza). Jeśli zatem nie znamy lokalnej zmiany częstotliwości, możemy iteracyjnie sprawdzać wskaźnik periodyczności p poszukując jego maksimum dla przepróbkowanego fragmentu w oparciu o różne nachylenia zbocza lokalnej PVC. W ten sposób odnajdujemy przebieg zmian częstotliwości dla każdej pary fragmentów sygnału. Na podstawie zmian lokalnych, określamy zmiany globalne (w sposób, jaki został podany we wstępie do niniejszego podrozdziału). Wykres na Rys. 44 ilustruje przebieg zmian wskaźnika periodyczności w zależności od założonej zmiany częstotliwości (w zakresie od 0,5 do 2, czyli odpowiednio częstotliwości dwa razy mniejszej i dwa razy większej).



Rysunek 44: Przykład zależności periodyczności od nachylenia PVC dla prawdziwej wartości 1,5

3. Zniekształcenie Wow

Ogólny schemat algorytmu autokorelacyjnego pozyskiwania PVC przedstawiony został na Rys. 45:



Rysunek 45: Schemat blokowy algorytmu metody autokorelacyjnej

3.2.1 Adaptacyjne badanie środka ciężkości widma

W pewnych przypadkach mamy do czynienia z wyraźnie zarysowanym komponentem widmowym, który to pojedynczy nawet może stanowić podstawę do estymacji PVC. Metoda adaptacyjnego badania środka ciężkości widma (unormowany moment pierwszego rzędu – patrz podrozdział 2.1.3.3 – Momenty widmowe) pozwala w efektywny sposób śledzić zmiany częstotliwości pojedynczego, ale wyraźnego komponentu. Jako że percypowane zmiany częstotliwości lepiej odzwierciedla skala nie liniowa, a logarytmiczna (zgodnie z prawem Webera-Fechnera [42]), operujemy na zmodyfikowanym wzorze na środek ciężkości widma. Dodatkowo aby ograniczyć analizę do najbliższego sąsiedztwa danej składowej widmowej (aby pozostałe, inaczej zmieniające się komponenty, nie wpływały na zmiany środka ciężkości), wprowadzamy granice sumowania (będące niejako wprowadzeniem filtru środkowoprzepustowego), także wersja, której używamy do obliczeń przedstawia się następująco:

$$M_u(1)_{oktawy} = \sum_{k=k_d}^{k_g} \frac{|G(k)| \cdot [\log_2(f_k)]}{|G(k)|} \quad (16)$$

gdzie: f_k oznacza częstotliwość środkową k-tego pasma,

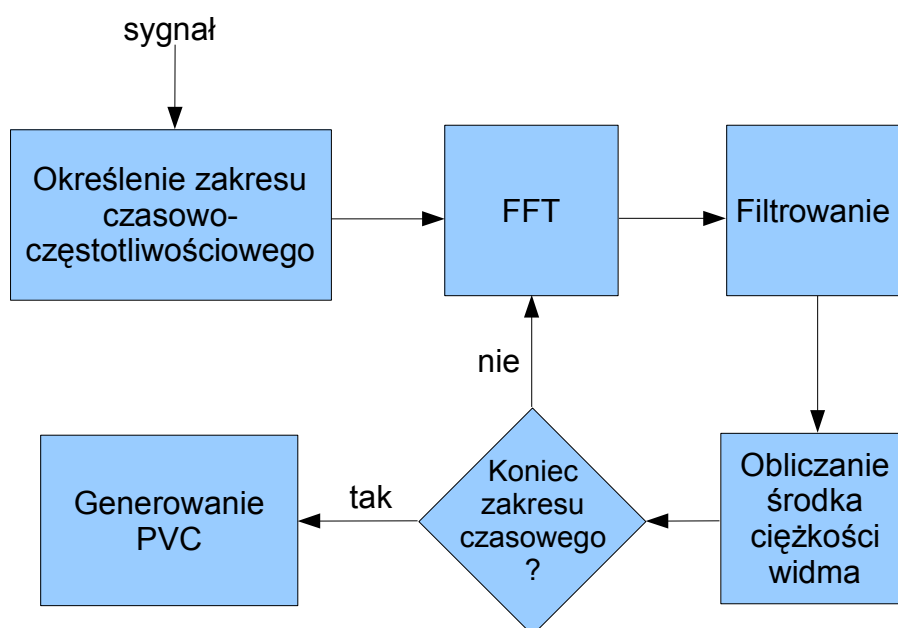
$G(k)$ oznacza wartość widma mocy dla k-tego pasma częstotliwości

k_d to numer pasma częstotliwości dolnej,

k_g to numer pasma częstotliwości górnej,

3. Zniekształcenie Wow

Należy zauważyć, że częstotliwość górna i dolna analizowanego zakresu zależy bezpośrednio od środka ciężkości widma w poprzednim kroku, a więc można powiedzieć, że algorytm jest adaptacyjny. W celu zredukowania wpływu innych komponentów, które mogą dostać się do i tak już zawężonego pasma częstotliwości, możemy dodatkowo stosować filtry wagowe (korzystając z funkcji okien lub tworząc nowe, prostsze). Ogólny schemat tej metody znajduje się na Rys. 46.

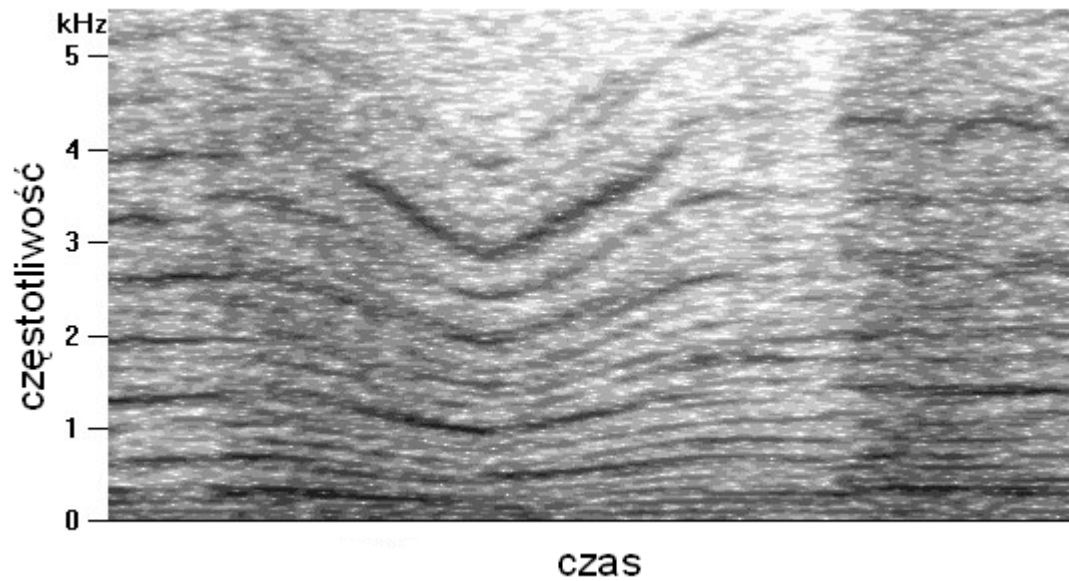


Rysunek 46: Schemat blokowy algorytmu metody autokorelacyjnej

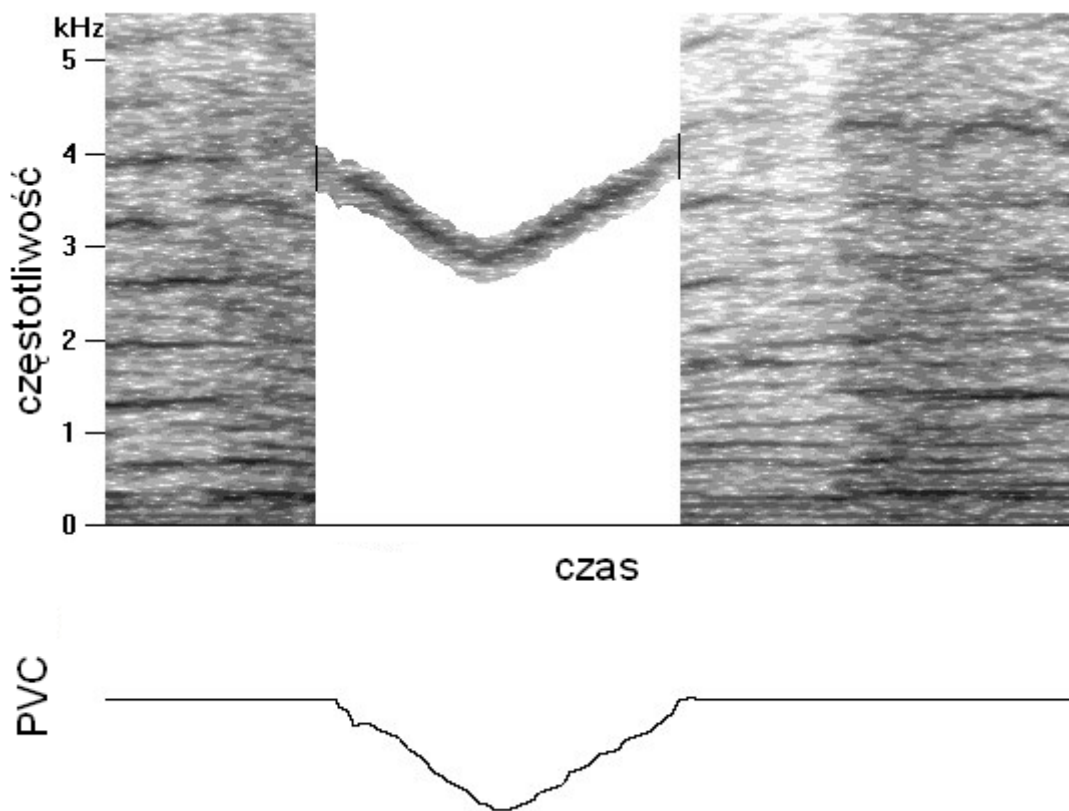
Pierwszy z bloków, tj. „Określenie zakresu czasowo-częstotliwościowego” zdradza pewną istotną cechę (trudną do rozpatrzenia w kategoriach wada-zaleta) algorytmu – konieczność obsługi operatora. Jak widać na spektrogramach z Rys. 47 i 48, nie zawsze cała długość ścieżki nadaje się do analizy, a co więcej – nie zawsze cała długość nagrania jest zniekształcona. Przy korzystaniu z metody należy zatem ręcznie podać zakres czasowy, a także szerokość filtru ŚP, który zostanie użyty przy obliczaniu środka ciężkości widma. Wartości należy podać raz, a wtedy algorytm bez dalszej ingerencji znajduje ścieżkę komponentu w podanym zakresie.

Metoda zatem stanowi doskonałe narzędzie do pół-automatycznej korekty lub tworzenia charakterystyk PVC dla pojedynczo występujących zakłóceń.

3. Zniekształcenie Wow



Rysunek 47: Spektrogram fragmentu nagrania z pojedynczym zniekształceniem wow



Rysunek 48: Spektrogram nagrania z Rys. 47 wybranym komponentem tonalnym (u góry) i wyznaczoną na jego podstawie krzywą PVC (u dołu)

3.2.4 Metody związane ze śledzeniem podnośnych

Poprzednia metoda adaptacyjnego badania środka ciężkości widma polegała na śledzeniu zmian częstotliwości pewnej silnej składowej tonalnej w sygnale. Jej istnienie było przypadkowe i zależało ściśle od rodzaju sygnału, z jakim mieliśmy do czynienia. Powoduje to ograniczenie jej stosowalności do wybranych przypadków, co czyni tę metodę mało uniwersalną do szerszego i automatycznego stosowania. Idąc jednak za tym pomysłem, powstała idea, by próbować śledzić zmiany pewnych stałych, obecnych w nagraniach tonów, co do których istniałaby pewność ich obecności w nagraniu. Zanim opisane zostaną metody śledzenia kilku wybranych tonów tego typu, przedstawione zostaną wymagania i założenia związane ze stosowaniem tej metody. Są nimi:

- dany ton został dodany na etapie nagrywania i podlegał tym samym procesom zniekształcenia, co sygnał użyteczny,
- częstotliwość danego tonu była stała w czasie nagrywania,
- w dalszej części procesu nagrywania nie został dodany kolejny ton tego typu uniemożliwiający określenie pierwotnej charakterystyki zniekształcenia,
- dokładna (teoretyczna) częstotliwość danego tonu nie musi być znana

Pierwsze założenie stanowi podwalinę pod grupę metod związanych ze śledzeniem podnośnych sygnałów. Przekłada się ono na twierdzenie, że modulacje zawartego w nagraniu tonu dodanego na etapie nagrania, niebędącego sygnałem nagrywanym odzwierciedlają modulacje całego nagrania, czyli estymacja PVC na podstawie zmian częstotliwości takiego tonu odzwierciedla PVC całego nagrania i może służyć jego rekonstrukcji.

Drugie z założeń jest o tyle ważne z powodu bezpośredniego przełożenia wahań częstotliwości danego tonu na globalne PVC całego nagrania. Jeśli sam ton był niestabilny, objawi się to zafałszowanym odwzorowaniem zniekształceń, co spowoduje znaczne zmniejszenie efektywności rekonstrukcji, a nawet wprowadzi zniekształcenie tam, gdzie go nie było. Niedopuszczalne zatem jest śledzenie tonów, co do których nie ma pewności, że są stabilne.

Trzecie z założeń dotyczy sytuacji, w których po nagraniu zostałby dodany kolejny ton tego samego typu maskujący użyteczny dla nas, zniekształcony identycznie do reszty nagrania, ton.

3. Zniekształcenie Wow

Czwarte z założeń mówi o tym, że do poprawnej rekonstrukcji nie jest potrzebna znajomość częstotliwości nominalnej tonu, w oparciu o który stworzona została charakterystyka krzywej PVC. Ze względów oczywistych wiedza o tej częstotliwości jest przydatna z uwagi na samo poszukiwanie tonu w widmie sygnału (jest to szczególnie ważne w metodzie śledzenia wysokoczęstotliwościowego prądu podkładu omówionej w podrozdziale 3.2.4.2 – Wysokoczęstotliwościowy prąd podkładu w dalszej części pracy), a także jeśli chcemy wyeliminować dryft częstotliwości, tj. modulację o stałej i niezmiennej w czasie wartości. Drugi przypadek dotyczy sytuacji, w których urządzenia rejestrujące nie pracowały z odpowiednią prędkością, która jednak pozostała niezmienna w czasie. Jeśli takie odstrojenie częstotliwości jest delikatne, staje się trudne do wykrycia przy zwykłym odsłuchu, jednak porównanie nominalnej częstotliwości badanego tonu z częstotliwością wykrytą we fragmencie wolnym od wow pozwala znajdować i korygować także tego typu zniekształcenia. W wypadku przeciwnym (gdy nominalna częstotliwość śledzonego tonu nie jest znana), opieramy się na tworzeniu lokalnej PVC, na podstawie której odwzorowujemy PVC globalną, co nie usuwa dryftu, gdyż zakłada że częstotliwość śledzonego tonu na początku nagrania była poprawna (pierwszą wartością PVC jest 1).

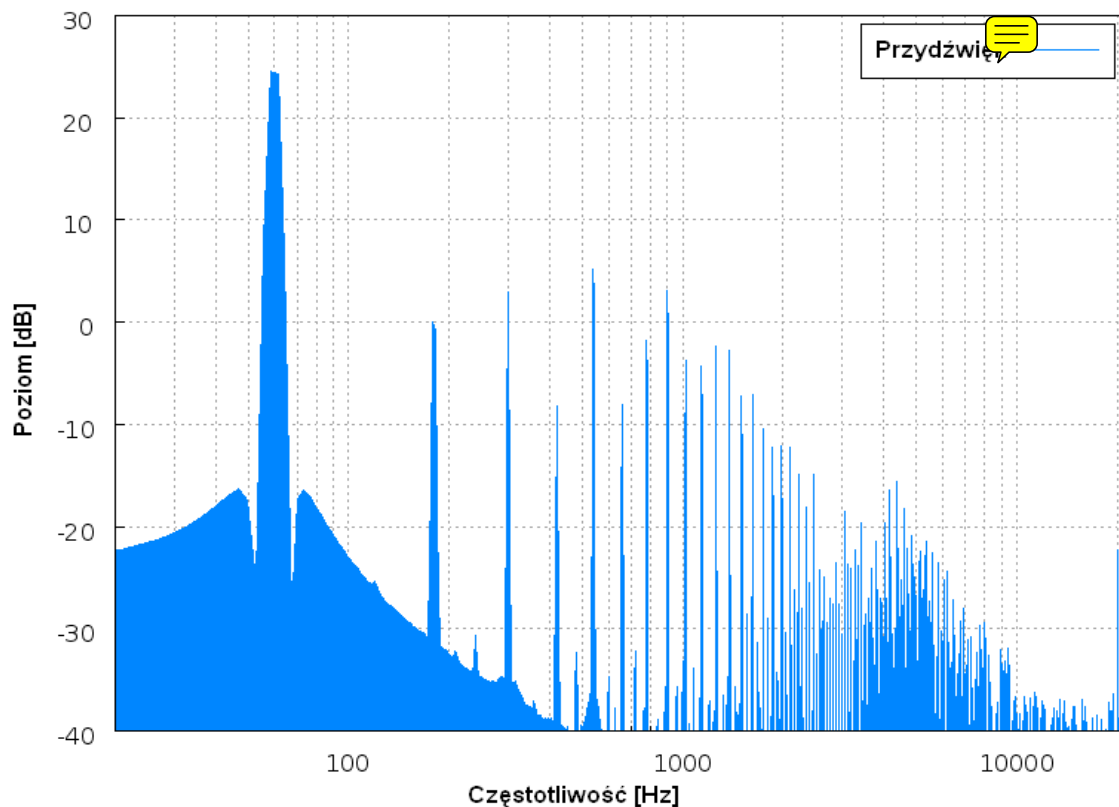
3.2.4.1 Przydźwięk sieciowy

Jedno z najbardziej uciążliwych zakłóceń dostających się podczas nagrywania i odtwarzania dźwięku jest przydźwięk sieciowy (Rys. 49). Jak sama nazwa wskazuje – pochodzi od sieci energetycznej. W zależności od kraju napięcie przemienne występujące w instalacji elektrycznej ma częstotliwość 50 lub 60 Hz. Wraz z rozwojem technik studyjnych, zwiększona została izolacyjność toru fonicznego, a co za tym idzie – zredukowana została szkodliwość przydźwięku. W starych nagraniach analogowych jest on jednak wyraźnie zarysowany, co obok oczywistych wad (jego harmoniczne są słyszalne jako tzw. brum [43]), stanowi niepodważalną zaletę – umożliwia estymację PVC.

Przydźwięk powstaje gdy linie zasilania interferują z liniami toru audio. Dzieje się tak najczęściej poprzez słabe ekranowanie kabli (pierwszej lub drugiej grupy) lub stosowanie kabli niesymetrycznych. Oprócz tego same urządzenia mogą wносить zakłócenie, jako że stanowią one elementy łączące obie linie. Powstawać mogą także pętle masy, kiedy to dwa różne urządzenia połączone są do dwóch różnych linii zasilania, w których występuje różnica

3. Zniekształcenie Wow

faz (~~a jest to bardzo częsty przypadek~~).



Rysunek 49: Widmo przydźwięku sieciowego o częstotliwości 60 Hz

Niska częstotliwość podstawowa tonu sieciowego oraz jego stosunkowo niski poziom (na Rys. 49 amplituda jest zawyżona w celach obrazowych) powodują, że jest to sygnał trudny do analizy za pomocą tradycyjnych metod częstotliwościowych (~~FFT, nawet STFT~~). Przykładowo w oknie o długości 1024 próbek, zmieści się jedynie jeden okres fali tonu podstawowego przydźwięku, a jest to zdecydowanie za mało, żeby dobrze oszacować jego częstotliwość. Zarazem zwiększając długość okna zmniejsza się precyzja czasowa, a co za tym idzie nadążanie za zmianami częstotliwości – objawia się to „rozmywaniem” prążków FFT.

W związku z powyższym, występuje konieczność zastosowania innej metody do określania częstotliwości tonu przydźwięku. Nadaje się do tego dobrze modelowanie **AR** ~~czyli inaczej autoregresywne~~, które doskonale sprawdza się w warunkach małego odstepu sygnału od szumu (SNR). Polega ono na traktowaniu sygnału jako odpowiedzi filtru o

3. Zniekształcenie Wow

nieskończonej odpowiedzi impulsowej (NOI) na ~~nasylenie~~ szumem białym. Model AR rzędu p może być zapisany jako:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + Z_t \quad (17)$$

gdzie X_t to sygnał w chwili t ,

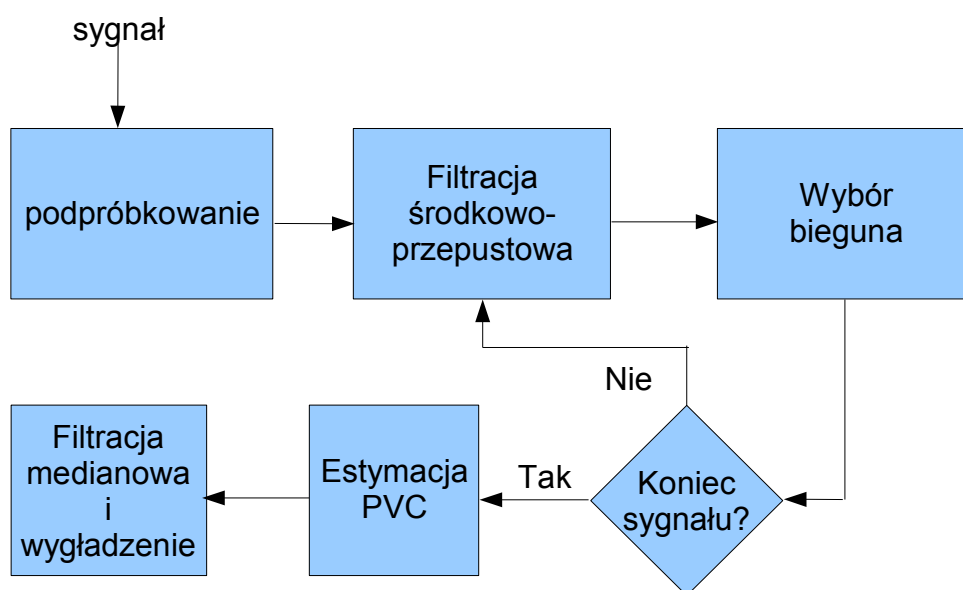
Z_t to sygnał losowy,

$\alpha_1, \dots, \alpha_p$ to współczynniki filtru,

p to rząd filtru

Znanych jest kilka metod uzyskiwania poszukiwanych współczynników filtru – najmniejszych kwadratów (LS), Yule-Walkera (YW) czy Burga [2][17]. Metody te pozwalają na osiągnięcie bardzo dobrej rozdzielczości częstotliwościowej nawet przy bardzo krótkich fragmentach sygnału. Jedynie metoda YW ma istotny mankament, gdyż nie sprawdza się w przypadku współczynników bliskich jedności (przekładających się na bieguny położone zbyt blisko okręgu jednostkowego) [8]. Oprócz parametrów filtru, ważną rolę odgrywa jego rząd. Im jest wyższy, tym więcej biegunów ma filtr, a więc tym więcej lokalnych maksimów (przekładających się na komponenty tonalne) w widmie możemy określić.

Ogólny schemat działania algorytmu śledzenia przydźwięku sieciowego wykorzystującego modelowanie AR przedstawiony został na Rys. 50.



Rysunek 50: Schemat blokowy algorytmu metody śledzenia przydźwięku sieciowego

3. Zniekształcenie Wow

Pamiętając o tym, że poziom przydźwięku sieciowego w nagraniu jest niewielki w stosunku do jego części użytecznej, uzasadnione jest stosowanie filtrów o dużym rzędzie, żeby nie potraktowały interesującego nas tonu jako tła szumowe. Z drugiej ~~jednak~~ strony zbytne zwiększanie rzędu filtru może spowodować uwypuklenie składowych rzeczywiście szumowych i utrudnić lokalizację szukanych komponentów. Dodatkowo zwiększanie rzędu filtru jest równoznaczne ze zwiększaniem złożoności obliczeniowej algorytmu. W celu optymalizacji rzędu filtru można stosować kryterium Akaike'a [17], jednak możliwe jest także jego ręczne wprowadzenie lub nawet odgórne przyjęcie dla danego typu sygnałów. W przypadku analizy częstotliwości przydźwięku sieciowego, **nie jest interesująca zawartość nagrania**, a więc cały zakres widma. Dlatego, by ograniczyć ilość współczynników AR, stosuje się podpróbkowanie sygnału do częstotliwości będącej czwartą wielokrotnością częstotliwości przydźwięku (200 lub 240 Hz). Następnie sygnał filtruje się pasmowo, zostawiając najbliższe sąsiedztwo tonu sieciowego (podejście znane z metody adaptacyjnego badania środka ciężkości widma), **przy czym wybór częstotliwości środkowej filtru, jak i jego dobroci (szerokości) działa adaptacyjnie, tj. częstotliwość środkowa zmienia się wraz ze zmieniającą się częstotliwością przydźwięku, a szerokość pasma zależy od przewidywanej głębokości modulacji wow.**

Blok opisany jako „Wybór bieguna” w praktyce realizowany jest na dwa sposoby. Pierwszy z nich zakłada wybór bieguna najbliższego nominalnej częstotliwości tonu przydźwięku. Drugi z kolei działa adaptacyjnie, tj. polega na wyborze bieguna najbliższego ostatnio zarejestrowanej częstotliwości tonu przydźwięku. Oba podejścia mają swoje wady i zalety, ~~oba mają swoje zastosowania. Typowo~~ **pierwszy sposób stosuje się przy zniekształceniach o stosunkowo dużej częstotliwości modulacji (blisko 10 Hz) i niskiej głębokości – innymi słowy kiedy szukany komponent szybko oscyluje wokół znanej częstotliwości, ale zmiany są niewielkie. Drugi ze sposobów sprawdza się lepiej w sytuacji odwrotnej – kiedy mamy do czynienia z modulacją głęboką, ale o niedużej częstotliwości. W tym przypadku istnieje niebezpieczeństwo odstrojenia się algorytmu od śledzenia przydźwięku w wypadku silnie interferujących komponentów widmowych w pobliżu tonu sieciowego.**

Gdy ścieżka przydźwięku zostanie już znaleziona, stosunek wykrytej częstotliwości do częstotliwości nominalnej przekłada się na wartości globalnej krzywej PVC. Filtracja medianowa pozwala na usunięcie nieprawidłowych biegunów interferujących składowych

3. Zniekształcenie Wow

widmowych, a wygładzenie za pomocą krzywej ruchomej poprawia kształt ~~w celu lepszego wykorzystania krzywej przy rekonstrukcji sygnału.~~

3.2.4.2 Wysokoczęstotliwościowy prąd podkładu

Niemalże wzorcowym tonem do śledzenia jest **prąd podkładu**. Jego niepowtarzalnymi zaletami są ~~występowanie na częstotliwości~~ wysoko powyżej pasma akustycznego, ~~wysoki poziom w stosunku do szumów oraz obecność we wszystkich analogowych nagraniach~~ (których to właśnie dotyczy problem zniekształceń typu wow&flutter). Jednak pierwsza z wymienionych zalet jest zarazem jego wadą. **Głowice odtwarzające magnetofonów nie przenoszą tak wysokich częstotliwości, żeby zgranie nagrania z prądem podkładu było bezpośrednio możliwe.** Katedra Systemów Multimedialnych na wydziale ETI Politechniki Gdańskiej zaproponowała 3 sposoby zgrywania taśm wraz z prądem podkładu – są one przedstawione w kilku publikacjach [4][5][7].

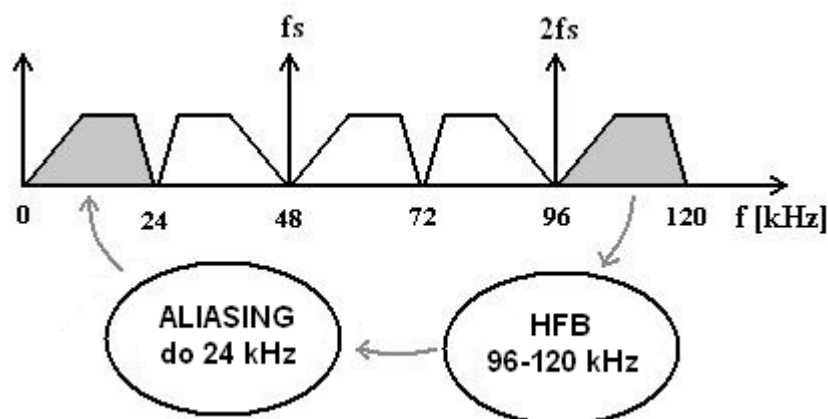
Pierwszy ze sposobów polega na wykorzystaniu zjawiska aliasingu do „sprowadzenia”

wysokoczęstotliwościowych komponentów widmowych do podstawowego pasma $\left(0; \frac{f_s}{2}\right)$

, ~~gdzie f_s to częstotliwość próbkowania.~~ Stosuje się w tym celu podpróbkowanie sygnału. Zjawisko demonstruje Rys. 51.

Metoda ta wiąże się jednak z dwiema trudnościami. Pierwszą z nich jest wbudowany filtr antyaliasingowy, który ~~od razu~~ odfiltrowuje wszystkie składowe sygnału powyżej częstotliwości próbkowania. Drugą natomiast jest konieczność analogowego odfiltrowania wszelkich sygnałów poniżej interesującego nas pasma prądu podkładu, żeby podczas podpróbkowania się na siebie nie nałożyły.

3. Zniekształcenie Wow



Rysunek 51: Wykorzystanie aliasingu przy podpróbkowaniu prądu podkładu

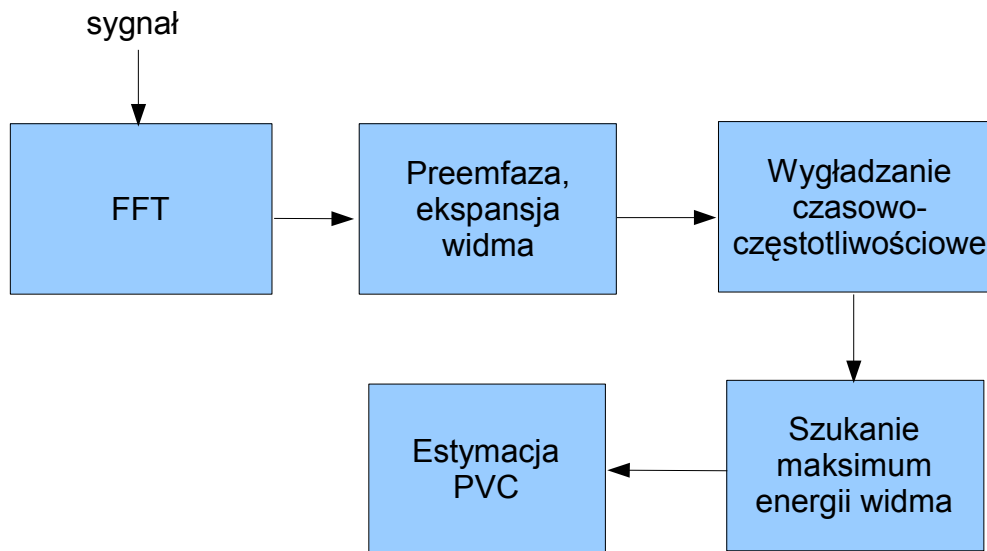
Drugi ze sposobów opiera się na zjawisku heterodynowania, czyli sumowania pewnego sygnału o danej częstotliwości f_s z lokalnie wytwarzanym sygnałem o częstotliwości f_h zbliżonej do f_s , w wyniku czego otrzymujemy sygnały o częstotliwościach będącej różnicą i sumą f_s i f_h [41]. Jeśli weźmiemy sygnał o częstotliwości $f = f_s - f_h$, będzie się on mieścić w pasmie akustycznym, a więc możliwe będzie normalne zgranie i śledzenie HFB. Konieczne będzie jednak użycie specjalnego interfejsu składającego się z miksera sygnałów i przetwornika analogowo-cyfrowego [10]. Problem może stanowić konieczność wiedzy o częstotliwości na jakiej znajduje się prąd podkładu.

Najprostsze podejście reprezentuje ostatni ze sposobów na zgranie tonu prądu podkładu. Polega on na spowolnieniu tempa odtwarzania nagrania, a więc bez korzystania z żadnych dodatkowych urządzeń. Przy zwolnionym tempie **zwiększona jest niejako częstotliwość próbkowania**, co pozwala na zarejestrowanie sygnału w pasmie wystarczającym do przechwycenia HFB. Jednocześnie zgrywany jest sam sygnał użyteczny, co stanowi dodatkową zaletę metody. Problemem jest jednak zapewnienie stałej prędkości odtwarzania, gdyż modulacje prędkości przesuwu taśmy **(które przy mechanicznym spowalnianiu mają duże szanse wystąpić)** spowodują dodatkowe zniekształcenia w sygnale.

Dysponując gotowym do analizy sygnałem, poddajemy go standardowym operacjom FFT, a następnie stosujemy preemfazę i ekspansję widma. Po wygładzeniu

3. Zniekształcenie Wow

częstotliwościowym i czasowym (~~celem zwiększenia dokładności określania częstotliwości HFB~~), maksimum energii widma powyżej pasma akustycznego określa nam szukaną częstotliwość prądu podkładu. Schemat blokowy algorytm w uproszczonej postaci znajduje się na Rys. 52. W kolejnym rozdziale pracy zostanie on omówiony szerzej.



Rysunek 52: Schemat blokowy algorytmu metody śledzenia prądu podkładu

4. Realizacja algorytmów do śledzenia podnośnych

Celem niniejszej pracy było, by w oparciu o dotychczasową wiedzę zaproponować algorytm do śledzenia pilota MTS w celu ~~późniejszej rekonstrukcji nagrań zniekształconych przez wów~~. Zadanie było ~~częścią~~ międzynarodowego projektu PrestoSpace [46] mającego na celu ochronę archiwów audiowizualnego ~~dziedzictwa europejskiego~~. Zdecydowano by sprawdzić możliwość zastosowania opracowanego wcześniej ~~w ramach projektu~~ algorytmu śledzącego wysokoczęstotliwościowy prąd podkładu do śledzenia pilota MTS.

4.1 Wysokoczęstotliwościowy prąd podkładu

Algorytm działa w sposób iteracyjny dla całego wczytanego do pamięci fragmentu nagrania. Ilość kroków zależy od jego długości w próbkach (która zależy od czasu trwania nagrania i częstotliwości próbkowania), parametru zakładkowania i w pewnym stopniu od rozmiaru okna, a wyraża się wzorem:

```
liczba_kroków = zaokrąglij_w_dół((długość_nagrania - długość_okna) /  
długość_zakładki)
```

Nominalne wartości parametrów wynoszą:

```
długość_okna = 1024  
długość_zakładki = 768
```

Algorytm składa się z 3 oddzielnych pętli w których realizowane są kolejno:

1. ~~preprocessing~~, tworzenie spektrogramu nagrania, wygładzanie częstotliwościowe,
2. wygładzanie czasowe,
3. poszukiwanie prądu podkładu.

W pętli pierwszej w każdym kroku wykonywane są następujące operacje:

- ~~wycinanie z nagrania fragmentu stosując okno Hanna~~,
- liczenie FFT,
- obliczenie modułu współczynników FFT → przejście na widmo amplitudowe,

4. Realizacja algorytmów do śledzenia podnośnych

- zastosowanie filtru górnoprzepustowego,
- zastosowanie preemfazy,
- zastosowanie ekspansji widma,
- wygładzenie częstotliwościowe widma za pomocą średniej ruchomej.

Filtr górnoprzepustowy wycina całą zawartość nagrania w pasmie akustycznym. Częstotliwość odcięcia przyjęta jest na $f_c = 25 \text{ kHz}$. Filtracja realizowana jest jako zerowanie współczynników widmowych poniżej f_c . Zastosowana funkcja preemfazy zdefiniowana jest wzorem:

$$p[n] = \left(\frac{n}{N} \right)^k, \quad n = 1, 2, \dots, N \quad (18)$$

gdzie k to współczynnik preemfazy (nominalnie 2)

Kształt tak zdefiniowanej funkcji oraz jego zależność od współczynnika k ilustruje wykres na Rys. 53. Operacja stosowania preemfazy polega na wymnożeniu próbek widma z odpowiadającymi im wartościami funkcji $p[n]$. Daje to efekt uwypuklenia wyższych częstotliwości, jako że ton pilota znajduje się w zakresie górnej części pasma. Jest to o tyle przydatne, że często szum obecny w nagraniu ma charakterystykę szumu różowego [44], czyli o energii malejącej z częstotliwością, a więc zastosowanie preemfazy pozwala na jego wyrównanie. W ten sposób tło szumowe, na którym znajduje się pilot jest energetycznie równomiernie rozłożone w widmie, co chroni przed błędami detekcji HFB.

Operacja ekspansji polega na podniesieniu widma do 4 potęgi, żeby zwiększyć odstęp sygnału od szumu.

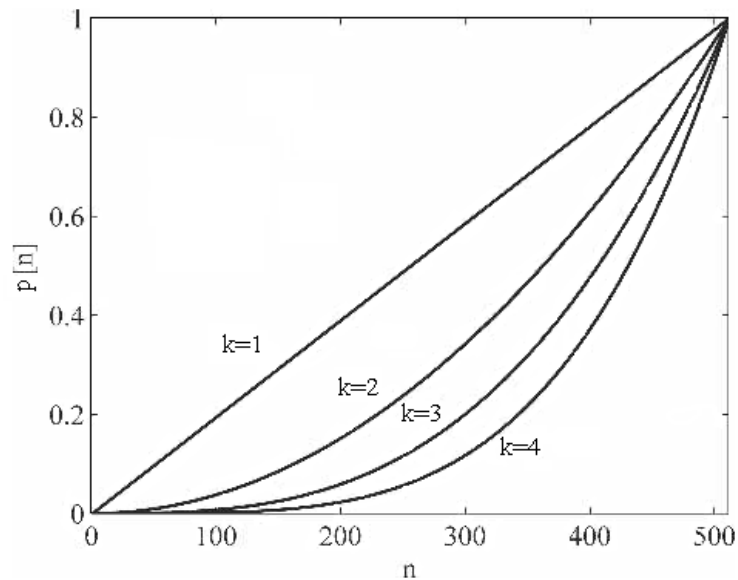
Wygładzanie widma polega na przepuszczeniu go przez filtr średniej ruchomej o parametrze 3, tj. każdą kolejną próbkę widma interpolujący jako średnią z siebie samej i dwóch sąsiednich próbek. W przypadku braku próbki sąsiadującej (przy próbce o indeksie 1), za sąsiadującą przyjmuje się próbkę będącą średnią całego widma.

Kolejna pętla jest potrzebna do wygładzenia czasowego widma. Powodem, dla którego nie zostało ono zaimplementowane w pętli poprzedniej jest konieczność dysponowania całą długością czasową spektrogramu. Wygładzanie realizowane jest jako ta

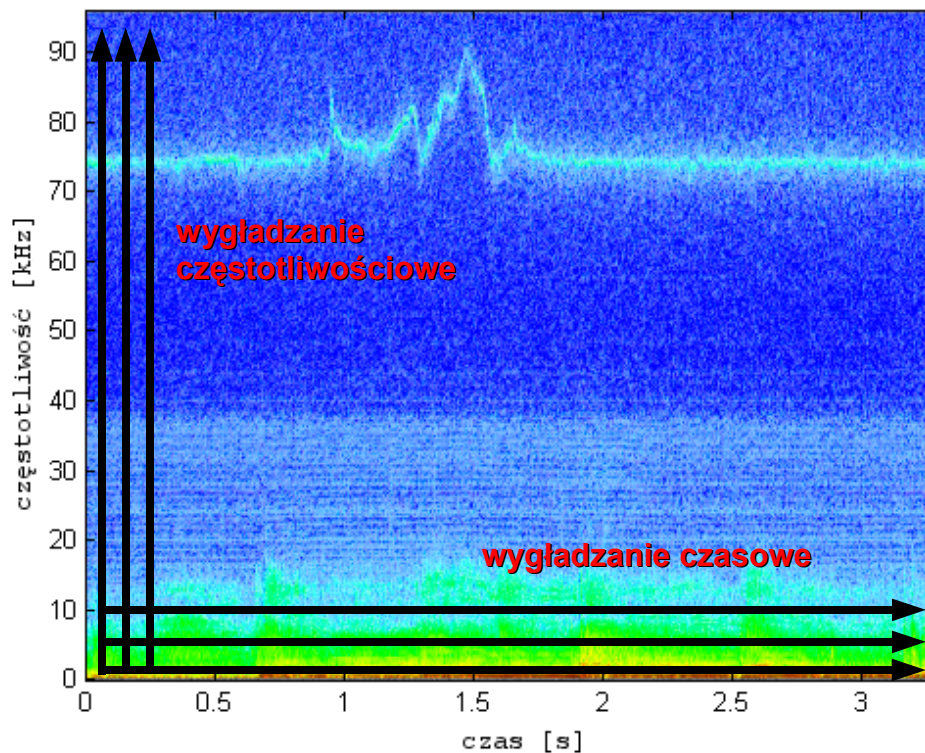
4. Realizacja algorytmów do śledzenia podnośnych

sama funkcja średniej ruchomej.

Kierunki obu typów wygładzania zostały zaznaczone na przykładowym spektrogramie (Rys. 54).



Rysunek 53: Wpływ wartości współczynnika k na funkcję preemfazy o długości 512 próbek



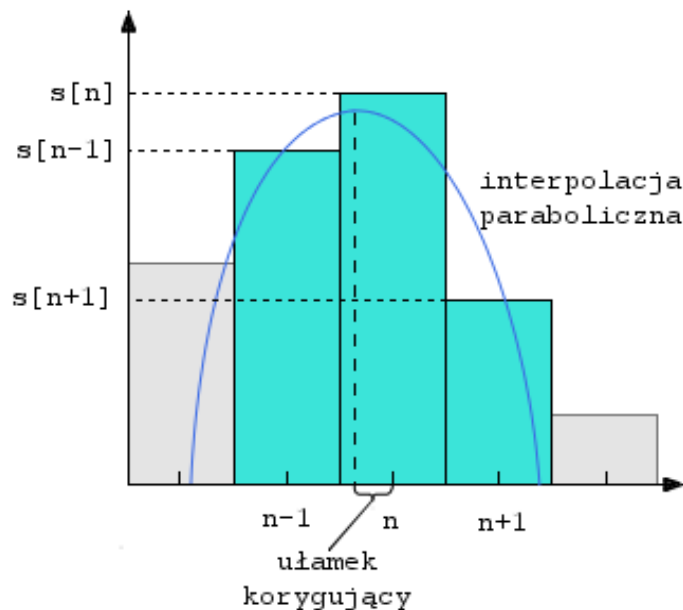
Rysunek 54: Kierunki wygładzania zaznaczone na przykładowym spektrogramie

4. Realizacja algorytmów do śledzenia podnośnych

W tak przygotowanym spektrum częstotliwość prądu podkładu znajdowana jest jako indeks maksimum ~~energetycznego~~. Aby estymacja była dokładniejsza, stosuje się dodatkowo interpolację paraboliczną wspomnianą w podrozdziale 3.2.1. Zakładając, że maksimum energii zostało odnalezione w indeksie częstotliwości o numerze n , któremu odpowiada współczynnik spektralny $s[n]$, ułamek korygujący znajdujemy na podstawie równania:

$$i_k = \begin{cases} \frac{s[n-1] - s[n+1]}{2(s[n-1] + s[n+1] - 2s[n])} & \text{dla } s[n] \neq \frac{s[n-1] + s[n+1]}{2} \\ 0 & \text{dla } s[n] = \frac{s[n-1] + s[n+1]}{2} \end{cases} \quad (19)$$

Interpretacja graficzna ułamka korygującego została przedstawiona na Rys. 55.



Rysunek 55: Zwiększenie precyzji określania częstotliwości wykorzystując interpolację paraboliczną

Ekstrakcja PVC na podstawie znalezionej śladu prądu podkładu nie dotyczy już samego algorytmu, dlatego nie zostanie w niniejszej pracy opisana.

4. Realizacja algorytmów do śledzenia podnośnych

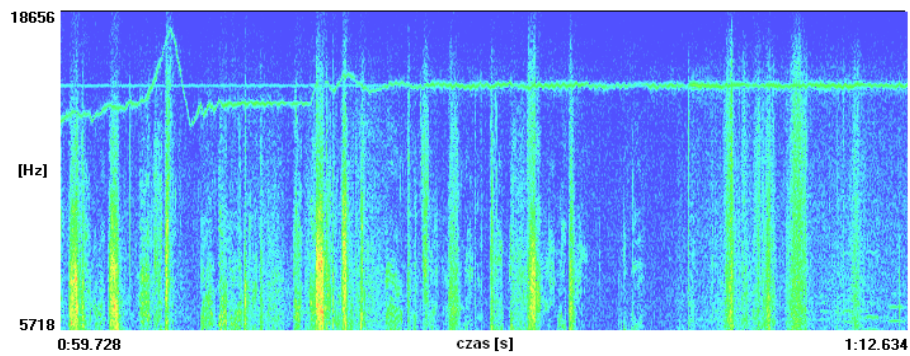
4.2. Pilot MTS

Punktem wyjścia dla stworzenia algorytmu do śledzenia pilota MTS był algorytm opisany w poprzednim podrozdziale. Założono, że dokonując pewnych modyfikacji uda się dostroić algorytm do śledzenia nie prądu podkładu, a tonu pilota MTS. Podstawowymi różnicami między oboma sygnałami są:

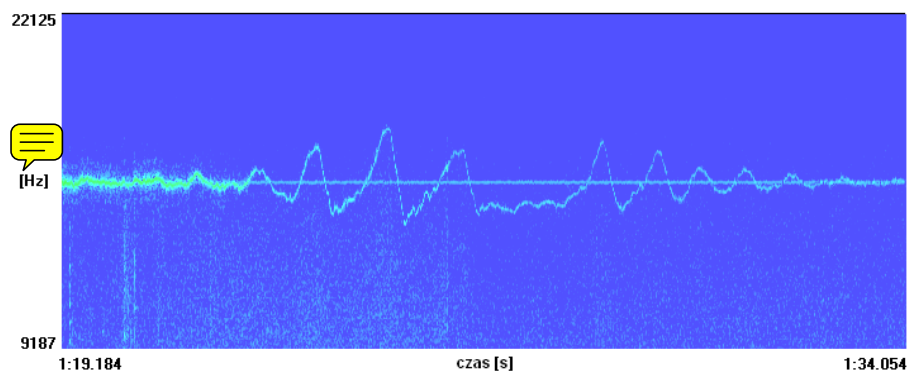
- ton pilota MTS znajduje się wewnątrz pasma akustycznego, a co za tym idzie może być maskowany przez komponenty szumowe lub impulsowe sygnału użytecznego (Rys. 56),
- podczas kopiowania nagrań NTSC, dodawany jest nowy pilot MTS (poprzedni traktowany jest jak sygnał akustyczny), a więc powstaje zjawisko zwane dalej pilotem tła (Rys. 56-58),
- ton pilota ma niewielki i zmienny w czasie poziom (mogą występować samoistne zaniki) (Rys. 57,58)

Konsekwencją powyższych różnic jest konieczność wykluczenia metody poszukiwania maksimum widma powyżej pewnej częstotliwości odcięcia filtru górnoprzepustowego jako określającej częstotliwość tonu pilota z co najmniej trzech powodów. Pierwszym z nich jest możliwość wystąpienia sygnałów o większej energii niż pilot w jego pobliżu (Rys. 56), co w przypadku kryterium opartym na szukaniu maksimum energii spowoduje powstanie licznych błędów grubych. Drugi powód jest zbliżony – istnieje bowiem możliwość występowania w pobliżu szukanego tonu pilota tła o porównywalnej energii (Rys. 57). Trzecim powodem jest zaś brak wiedzy o maksymalnej możliwej głębokości pasożytniczej modulacji FM (Rys. 58), a więc o minimalnej częstotliwości odcięcia filtru GP, jaki możemy zastosować, żeby nie wyciąć z widma tonu pilota, a pozbyć się wysokoenergetycznych sygnałów akustycznych. W efekcie konieczne jest wprowadzenie działania adaptacyjnego oraz szukanie tonu pilota w precyzyjnie określonym wąskim przedziale widma.

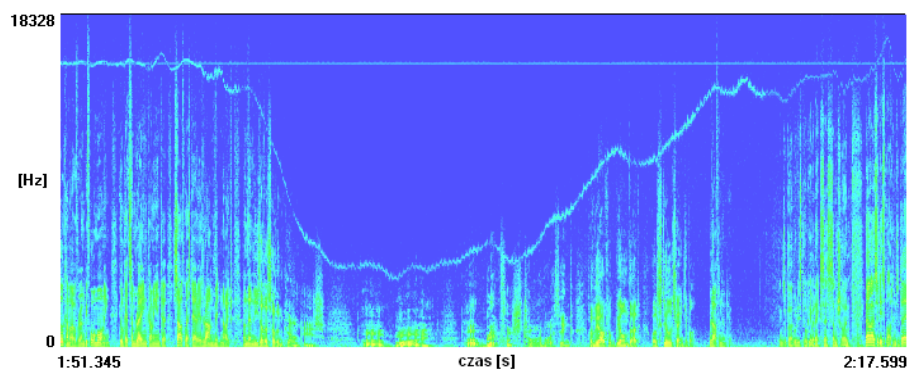
4. Realizacja algorytmów do śledzenia podnośnych



Rysunek 56: Spektrogram fragmentu nagrania z maskowaniem pilota MTS przez sygnały akustyczne



Rysunek 57: Spektrogram fragmentu nagrania z pilotem MTS o niewielkim poziomie w obecności pilota tła



Rysunek 58: Spektrogram fragmentu nagrania z obszarem o głębokiej aperyodycznej modulacji częstotliwości

4. Realizacja algorytmów do śledzenia podnośnych

4.2.1 Wprowadzone modyfikacje

~~Pierwszym podjętym krokiem~~ było zostawienie jedynie wąskiego marginesu widma w okolicach poszukiwanego tonu pilota, by sygnały zakłócające miały jak najmniejszy wpływ na dalszą analizę. Szybko okazało się, że na sztywno przyjęte wartości szerokości pasma takiego filtru środkowoprzepustowego nie sprawdzały się dla głębokich modulacji, gdzie całe widmo ulega kompresji, w wyniku czego wysokoenergetyczne komponenty mogą wejść do zakresu analizy. Przyjęto zatem, że rozmiar marginesu zależeć będzie od częstotliwości w stosunku wprost proporcjonalnym. Ponadto okazało się, że właściwe będzie zróżnicowanie rozmiarów górnego i dolnego marginesu ze względu na logarytmiczną skalę zmian częstotliwości i większą energię widma poniżej niż powyżej tonu pilota. Przy rozmiarach jednakowych, zbyt mała wartość marginesów powodowała trudności z nadążeniem za szybkimi zmianami głębokości modulacji (szczególnie gdy modulacja zmniejszała się gwałtownie). Jednocześnie przy marginesach zbyt dużych następowało większe zakłócanie przez sygnały akustyczne. Kompromisem pomiędzy tymi dwoma faktami było **wprowadzone zróżnicowanie**.

Kolejną modyfikacją było wprowadzenie filtru wagowego wewnątrz zawężonego już przedziału częstotliwości. Redukowało to w dalszym stopniu sygnały akustyczne wchodzące w jego granice, a ponadto w znacznym stopniu zmniejszało efekt maskowania pilota przez sygnały szumowe i impulsowe. Sygnały te mają charakterystykę zbliżoną do szumu różowego, który jest wybielany poprzez zastosowanie preemfazy, a więc sygnały te wchodząc do analizowanego przedziału pasma mają w nim stały poziom energii. Zastosowanie filtru spowodowało, że miejsce o największym poziomie energii pozostaje na środku analizowanego przedziału i nie występują już gwałtowne skoki spowodowane wykryciem odległego maksimum energetycznego nie związanego z poszukiwanym pilotem. Ewentualność ich wystąpienia została także zminimalizowana poprzez wprowadzenie ograniczenia, które określa maksymalną możliwą zmianę głębokości modulacji w pojedynczym kroku.

Kolejnym problemem był pilot tła. Aby przeciwdziałać sytuacjom, w którym szybkie zmiany ~~periodycznie~~ oscylującego pilota w obecności pilota tła powodowały ~~zmylenie~~ algorytmu i błąd podążania za niewłaściwym, wprowadzono predykcję liniową oczekiwanej częstotliwości tonu szukanego pilota w kolejnej ramce czasu. Zastosowanie predykcji

4. Realizacja algorytmów do śledzenia podnośnych

liniowej 1-go rzędu dało dobre rezultaty, jednak przy gwałtownych zmianach monotoniczności, algorytm mógł zboczyć ze ścieżki, szczególnie jeśli w pobliżu danego ekstremum pilota był obecny silny interferujący sygnał. Dopiero zastosowanie predykcji rzędu 2-go (kwadratowej) pozwoliło uwzględnić zmiany monotoniczności przy przewidywaniu wystąpienia kolejnych częstotliwości tonu.

Zdecydowano się dodatkowo na zastąpienie kryterium wykrywania częstotliwości na podstawie prążka o największej energii ~~środkiem~~ ciężkości widma. Zapewniło to mniejszą podatność na przypadkowe odchylenia powodujące zboczenie algorytmu z prawidłowej ścieżki pilota będące konsekwencją zastosowania predykcji liniowej. Wadą tego kroku jest jednak to, że ścieżka pilota ma charakter ~~pewnej~~ estymacji. Jest ona bardzo dokładna w przypadku braku innych niż ton pilota MTS komponentów widmowych w analizowanym pasmie. W przypadku pojawienia się jakichkolwiek sygnałów zakłócających, środek ciężkości widma przesuwają się, jednak zachowana jest stabilność, co stanowi niepodważalną zaletę takiego podejścia.

Kolejnym krokiem mającym na celu zredukowanie wpływu obecności pilota tła było wprowadzenie jego filtrowania. Bardzo wąski filtr ustawiony precyzyjnie na częstotliwość tonu MTS zaczął działać tylko wtedy, gdy ujawniał się wpływ pilota tła, tj. wtedy, gdy właściwy pilot oscylował wokół nominalnej częstotliwości (wykazywał cechy niestalości). Dokładną częstotliwość pilota tła określa się na samym początku działania algorytmu (na początku nagrania – konieczne jest zatem założenie, że ~~on~~ nie jest ~~on~~ zniekształcony) poprzez znalezienie środka ciężkości widma w bardzo wąskim zakresie w okolicach nominalnej częstotliwości 15,734 kHz.

4.2.2. Opis algorytmu

Algorytm do śledzenia pilota MTS, w odróżnieniu od poprzednika, działa w jednej pętli. Stało się to możliwe za sprawą zmienionej zasady wygładzania czasowego spektrogramu, tj. z usunięcia nieprzyczynowości **funkcji**. Algorytm korzysta z **k** ostatnich próbek (k zwane jest parametrem funkcji wygładzającej), zamiast z $\frac{k-1}{2}$ próbek przed i po bieżącej.

Liczba kroków iteracji pozostała ~~w tej samej zależności~~, co w algorytmie bazowym.

4. Realizacja algorytmów do śledzenia podnośnych

Jako optymalne zostały również wybrane te same wartości parametrów związanych z tworzeniem spektrogramu. Dla przypomnienia, wspomniana zależność określona jest jako:

```
liczba_kroków = zaokrąglij_w_dół((długość_nagrania - długość_okna) /  
długość_zakładki)
```

Podczas gdy nominalne wartości występujących we wzorze parametrów, to:

```
długość_okna = 1024  
długość_zakładki = 768
```

Przed wejściem w główną pętlę algorytmu, dokonywana jest wspomniana w poprzednim podrozdziale estymacja częstotliwości pilota, używana później do filtrowania pilota tła. Estymacja ta polega na liczeniu środka ciężkości widma w wąskim pasmie wokół nominalnej częstotliwości pilota MTS. Widmo liczone jest dla pierwszej ramki sygnału po uprzednim zastosowaniu funkcji preemfazy, która dla całego algorytmu ma wartość współczynnika $k=3$. Zakres, dla którego liczony jest środek ciężkości widma to:

```
zakres_dolny = nominalna_częstotliwość_pilota - odległość_filtrowania_pilota / 4  
zakres_górny = nominalna_częstotliwość_pilota + odległość_filtrowania_pilota / 4
```

Parametry (wielkości wyrażone w hercach) określone są bazowo jako:

```
nominalna_częstotliwość_pilota = 15734  
odległość_filtrowania_pilota = 1000
```

Pierwszy z parametrów określa nominalną częstotliwość pilota MTS (teoretyczną) wokół której należy szukać rzeczywistego tonu pilota.

Drugi z parametrów dotyczy maksymalnej odległości (na skali częstotliwości) śledzonego pilota od pilota tła, dla której należy przeprowadzić operację filtrowania tego drugiego. Ma to zapobiec niepotrzebnemu przeprowadzaniu operacji filtracji części pasma, która i tak znajduje się poza zakresem, dla którego liczymy środek ciężkości, by określić częstotliwość tonu śledzonego pilota.

Następnie na podstawie znalezionej częstotliwości określa się raz charakterystykę filtru środkowozaporowego mającego w razie potrzeby wycinać ton pilota tła. Filtr ten ma charakterystykę opisaną wzorem:

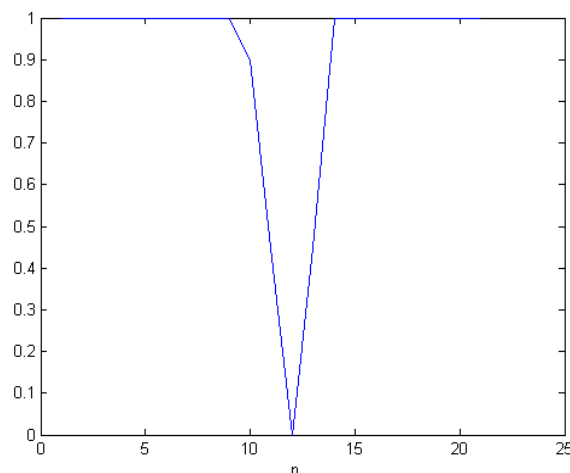
4. Realizacja algorytmów do śledzenia podnośnych

$$filtr[n] = \begin{cases} 0,9 \frac{|n-k_p|}{m_f} & \text{dla } n \in \langle k_p - m_f; k_p + m_f \rangle \\ 1 & \text{dla } n \notin \langle k_p - m_f; k_p + m_f \rangle \end{cases} \quad (20)$$

gdzie k_p to określona rzeczywista częstotliwość pilota (tła) MTS,

m_f to połowa pasma filtru wycinającego pilota tła (nominalnie 100 Hz)

Przykładową charakterystykę filtru dla parametrów wynoszących $k_p=12$ i $m_f=2$ przedstawia wykres na Rys. 59. Wartości te wyrażone są w próbkach FFT, a więc żeby przenieść je w dziedzinę częstotliwości, należy znać długość okna stosowanego do analizy i częstotliwość próbkowania sygnału dyskretnego. Przykładowo dla długości okna 1024 i częstotliwości próbkowania wynoszącej 48 kHz, częstotliwość środkowa przedstawionego filtru wynosiłaby 562,5 Hz, a jego szerokość – 2*93,75 Hz.



Rysunek 59: Filtr środkowozaporowy dla $k_p=12$ i $m_f=2$

Po określeniu występującej dla danego nagrania częstotliwości pilota MTS oraz charakterystyki filtru wycinającego, rozpoczyna się główna pętla programu.

Po dokonaniu standardowych operacji okienkowania (pozostaje okno Hanna), zastosowaniu preemfazy (o wyższym współczynniku wynoszącym 3), następuje obliczenie zerowego momentu widmowego (patrz rozdział 2.1.3.3 - Momenty widmowe), który wyraża energię sygnału. Aby uniknąć dzielenia przez zero w dalszej części programu, sprawdza się, czy nie mamy do czynienia z sygnałem o zbyt niskim poziomie. Jeśli tak będzie, wartość

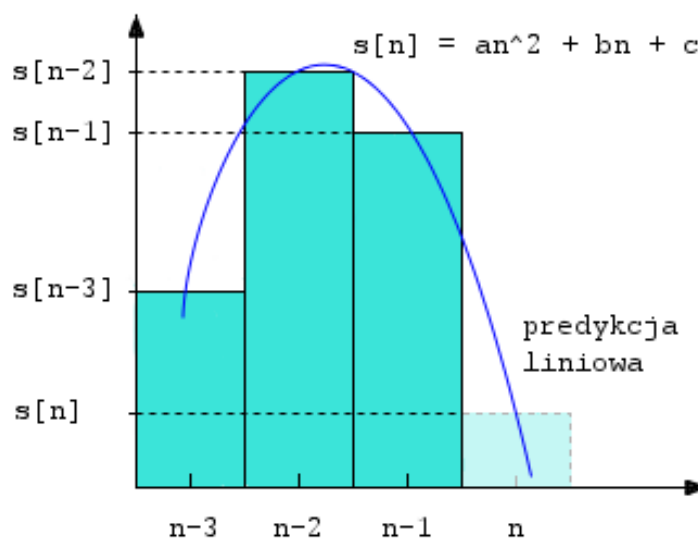
4. Realizacja algorytmów do śledzenia podnośnych

częstotliwości tonu pilota MTS przypisywana jest jako niekreślona (NaN – Not a Number) i program przechodzi do kolejnego kroku pętli. W przeciwnym razie zaczyna się właściwe działanie algorytmu.

Algorytm jest podzielony na dwa etapy. Przez pierwsze 10 kroków działa podobnie do algorytmu śledzenia prądu podkładu. Wycinane jest widmo sygnału akustycznego z częstotliwością odcięcia obniżoną do wartości $f_c = 14\text{kHz}$. Kolejno następują: ekspansja widma (podniesienie wartości współczynników spektralnych do 4 potęgi), wygładzanie częstotliwościowe (o parametrze 3), wygładzanie czasowe (tylko jeśli wykonuje się już co najmniej 3 krok pętli) oraz znajdowanie maksimum energii widma. W celu zwiększenia dokładności stosuje się interpolację paraboliczną (przedstawioną na Rys. 3). Jeśli zmiana częstotliwości w porównaniu do częstotliwości z poprzedniego kroku przekracza wartość parametru maksymalnego nachylenia śladu pilota, wtedy za aktualną wartość częstotliwości przyjmuje się częstotliwość pilota z poprzedniego kroku. Nominalna wartość parametru (w ~~hereach~~) wynosi:

```
maksymalne_nachylenie = 800
```

Po 10 krokach algorytm zaczyna działać adaptacyjnie. Pierwszą operacją po sprawdzeniu energii sygnału jest liniowa predykcja kolejnej częstotliwości pilota MTS. Opiera się ona o 3 ostatnie próbki i to m.in. dlatego nie mogła być stosowana od samego początku sygnału. Zasadę jej działania obrazuje Rys. 60.



Rysunek 60: Predykcja liniowa 3 rzędu

4. Realizacja algorytmów do śledzenia podnośnych

Predykcja liniowa polega na interpolacji k poprzednich próbek funkcją stopnia k (stąd parametr k wyraża rząd predykcji), aby po uzyskaniu odpowiednich parametrów móc wyliczyć bezpośrednio wartość próbki bieżącej. Wartości tych parametrów po odpowiednich modyfikacjach wprowadzonych przez autora wyrażają się wzorami:

$$a = \frac{s[n-3] - s[n-1]}{2} - s[n-2] \quad (21)$$

$$b = s[n-2] - s[n-3] - a \quad (22)$$

$$c = s[n-3] \quad (23)$$

Na ich podstawie wartość próbki w chwili czasu n oblicza się ze wzoru:

$$s[n] = 9 \cdot a + 3 \cdot b + c \quad (24)$$

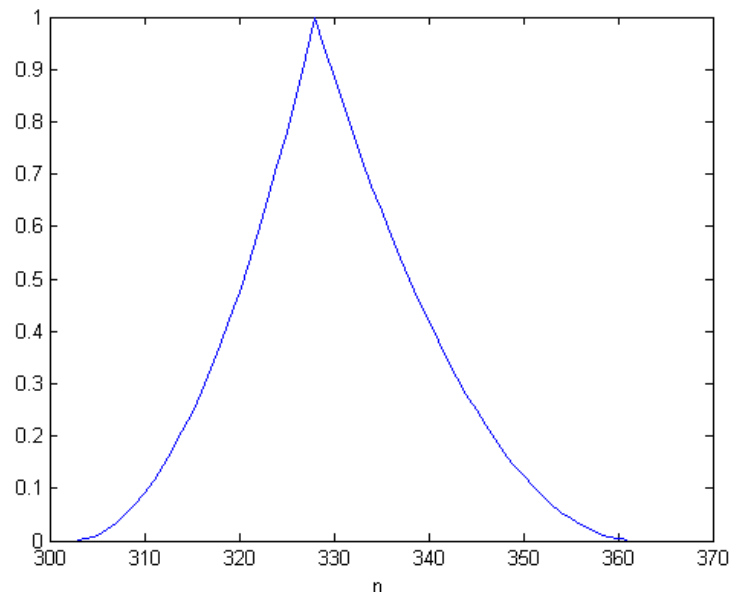
Obliczona częstotliwość wyraża nam częstotliwość środkową filtru ŚP, którego pasmo ma szerokość równą 17,5% tej częstotliwości (z czego 7,5% przypada w kierunku niższych częstotliwości, a 10% w kierunku wyższych). Charakterystyka filtru ma dla podanych wyżej wartości marginesów ogólną postać:

$$filtr[n] = \begin{cases} \left(\frac{1}{0,075 f_p} n - \frac{0,925}{0,075} \right)^2 & \text{dla } n \in \langle 0,925 f_p; f_p \rangle \\ \left(-\frac{1}{0,1 f_p} n + \frac{1,1}{0,1} \right)^2 & \text{dla } n \in \langle f_p; 1,1 f_p \rangle \\ 0 & \text{dla } n \notin \langle 0,925 f_p; 1,1 f_p \rangle \end{cases} \quad (25)$$

Wykres funkcji przykładowego filtru dla indeksu częstotliwości $f_p = 328$ któremu dla przykładowych parametrów długości okna równej 1024 i próbkowaniu z częstotliwością 48 kHz odpowiada częstotliwość 15375 Hz znajduje się na Rys. 61. Szerokość takiego filtru wyznaczają częstotliwości $f_d = 14203 \text{ Hz}$ i $f_g = 16922 \text{ Hz}$. Nie odpowiada to dokładnie wartościom $0,925 \cdot 15375 \text{ Hz} = 14222 \text{ Hz}$ oraz $1,1 \cdot 15375 \text{ Hz} = 16913 \text{ Hz}$ ze względu na zaokrąglenia indeksów częstotliwości używane w algorytmie (muszą mieć one wartość całkowitą). Błąd jest jednak względnie niewielki (odpowiednio 0,13% oraz 0,05%), a dodatkowo nie jest wymagana bezwzględna precyzja w określaniu częstotliwości filtru ŚP, tak jak to ma miejsce w wypadku znajdowania częstotliwości pilota

4. Realizacja algorytmów do śledzenia podnośnych

lub jej predykcji.



Rysunek 61: Filtr środkowoprzepustowy dla $f_p=328$

Przed zastosowaniem filtracji ŚP sprawdzane jest, czy odfiltrowany powinien być ~~jeszcze~~ pilot tła. W tym celu na skali częstotliwości jest obliczana odległość od **niego** (na podstawie określonej przed główną pętlą faktycznej częstotliwości pilota dla danego nagrania i przewidzianej częstotliwości dla bieżącego kroku). Jeśli jest mniejsza od parametru marginesu stabilności, a ponadto była mniejsza w każdym z ostatnich 10 kroków pętli, filtrowanie pilota nie następuje (zakłada się wtedy, że sygnał jest stabilny na właściwej częstotliwości). Parametr marginesu ma nominalną wartość wyrażoną w ~~hercach~~:

```
margines_stabilności = 200
```

Jeśli jednak sygnał pilota wykazywał zmiany, a dodatkowo odległość od właściwej dla danego **grania** rzeczywistej częstotliwości pilota MTS jest mniejsza niż parametr maksymalnej odległości filtrowania pilota, stosowany jest wąski filtr środkowozaporowy o charakterystyce wyliczonej na początku działania algorytmu.

Kolejną operacją jest usuwanie niepotrzebnych składników widma, tj. stosowanie filtracji ~~środkowoprzepustowej o omówionej w poprzednich akapitach charakterystyce.~~

W dalszej kolejności stosowana jest ekspansja widma (podnoszenie do 4 potęgi) i

4. Realizacja algorytmów do śledzenia podnośnych

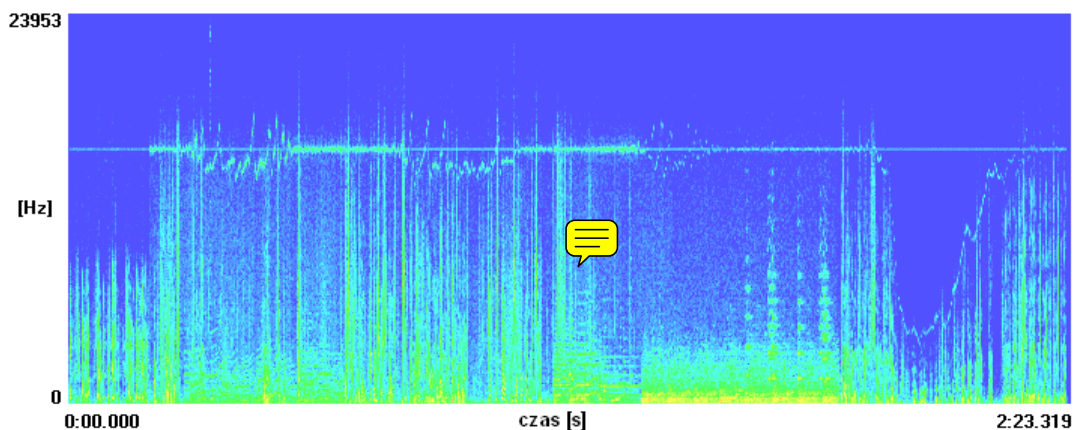
wygładzanie czasowe – jest to krok wspólny dla obu faz działania algorytmu.

Kończącym krokiem jest znalezienie częstotliwości pilota MTS na podstawie badania środka ciężkości widma przetworzonego za pomocą wszystkich opisanych wcześniej operacji.

Po wykonaniu wszystkich iteracji dysponujemy ścieżką częstotliwości pilota MTS w nagraniu, która może posłużyć do estymacji PVC dowolnymi metodami, które – podobnie jak w przypadku algorytmu bazowego – nie zostaną tutaj opisane.

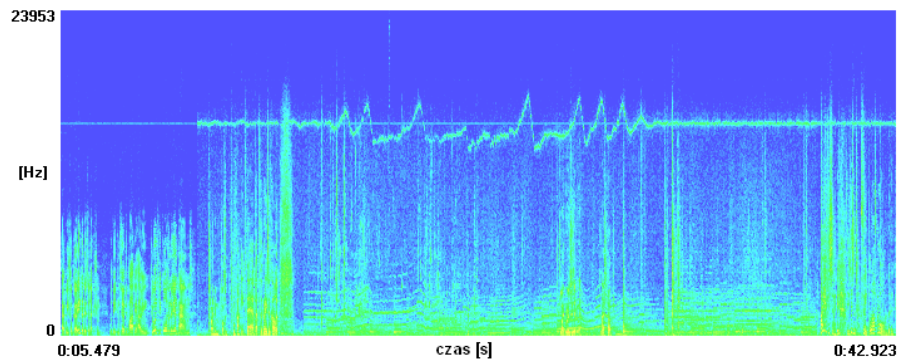
4.2.3 Eksperyment i wyniki

Algorytm był tworzony na podstawie 143-sekundowego nagrania (Rys. 62) z wyraźnymi rejonami objętymi zniekształceniem wow. Trzy z nich (między 0:20 a 0:30 – Rys. 63, 0:45 a 1:05 – Rys. 64 i 1:20 a 1:30 – Rys. 57,65) mają charakter oscylacyjny, czwarty zaś to obszar pojedynczej aperiodycznej modulacji z niewielkimi oscylacjami pod koniec (zaczyna się w okolicach 1:55 – Rys. 58,66).

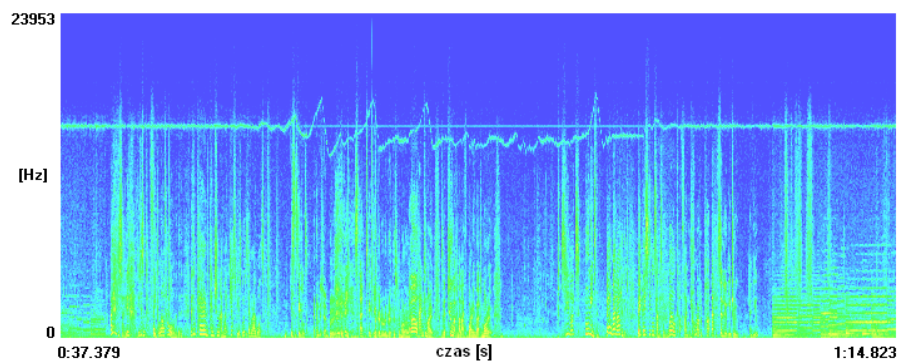


Rysunek 62: Spektrogram nagrania testowego

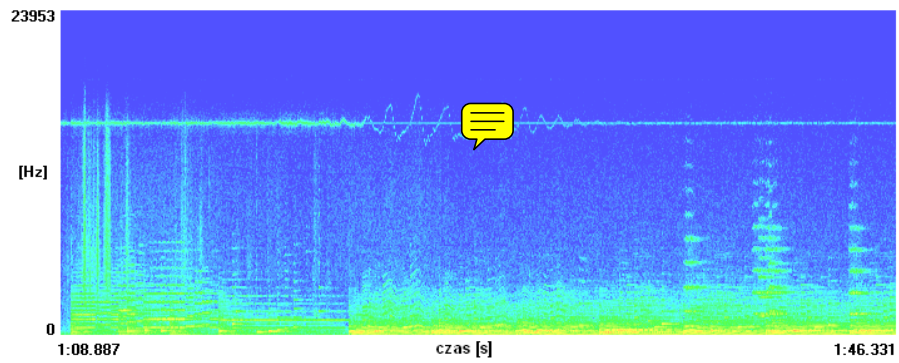
4. Realizacja algorytmów do śledzenia podnośnych



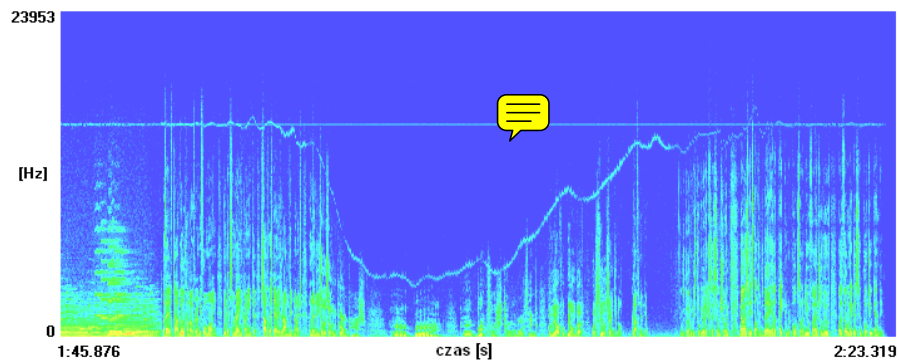
Rysunek 63: Spektrogram pierwszego zniekształconego fragmentu nagrania



Rysunek 64: Spektrogram drugiego zniekształconego fragmentu nagrania



Rysunek 65: Spektrogram trzeciego zniekształconego fragmentu nagrania

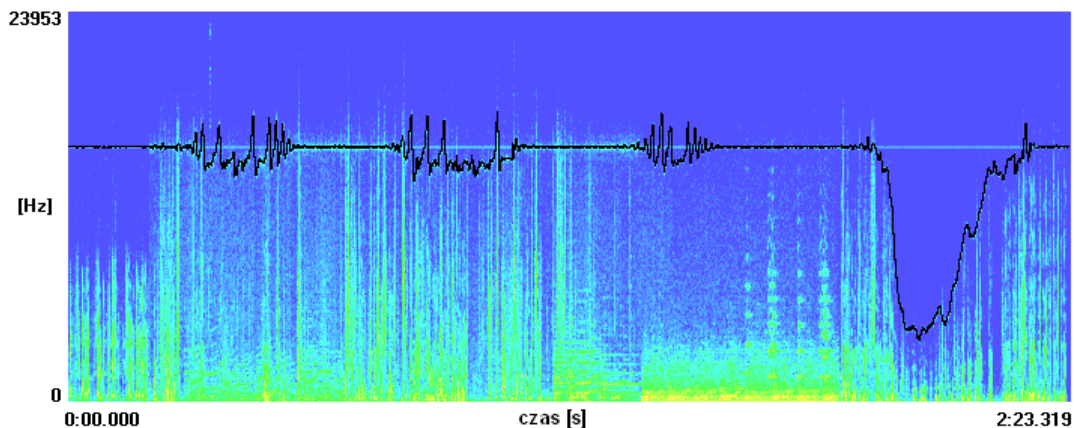


Rysunek 66: Spektrogram czwartego zniekształconego fragmentu nagrania

4. Realizacja algorytmów do śledzenia podnośnych

Pilot tła jest zauważalny w całej długości nagrania. Poziom właściwego pilota MTS jest zmienny – zwiększa się w okolicach 11 sekundy (w rejonie pierwszego zniekształcenia), spada w okolicach 1:22 (wraz z pojawieniem się zniekształcenia w rejonie trzecim), po czym ponownie po 2:13 (końcowa część regionu czwartego zniekształcenia). Region pierwszy nie sprawił kłopotów przy śledzeniu, głównie ze względu na wysoki poziom pilota. W regionie drugim występują silne „zakłócenia” ze strony sygnału użytecznego, co utrudniało podążanie za właściwym tonem pilota. Region trzeci charakteryzował się porównywalnym poziomem pilota w stosunku do pilota tła, w wyniku czego algorytm wymagał modyfikacji dotyczących filtracji pilota tła, by poprawnie śledzić ton MTS w tym fragmencie. Fragment czwarty był najbardziej podatny na zakłócenia ze strony wysokoenergetycznych komponentów widmowych sygnału użytecznego, szczególnie w okolicach 2:05. Oprócz tego dalszy spadek poziomu sygnału pilota utrudnił właściwe śledzenie pod koniec fragmentu, w sąsiedztwie pilota tła i szumowych komponentów.

Po odpowiednim dobraniu metod i parametrów je charakteryzujących, jak to zostało opisane wyżej, uzyskano satysfakcjonujące wyniki śledzenia. Spektrogram nagrania z zaznaczoną ścieżką pilota MTS znajduje się na Rys. 67.



Rysunek 67: Spektrogram nagrania testowego z zaznaczoną ścieżką pilota

5. Wnioski i podsumowanie

Zaprezentowany algorytm z całą pewnością nie jest idealnym ani jedynym możliwym rozwiązaniem. Nie można go też traktować jako rozwiązanie ostateczne, gdyż każdą z zastosowanych koncepcji można rozwinąć bądź ulepszyć. Jego forma utrwaliła się jednak wraz z osiągnięciem pozytywnych wyników dla opisanego w podrozdziale 4.2.3 – Eksperymenty i wyniki zniekształconego materiału, jaki był do dyspozycji autora. Konieczność pracy na rzeczywistych zniekształceniach, jak i rzeczywistych sygnałach NTSC w połączeniu z utrudnionym dostępem do tychże (archiwa europejskie nie dysponują szeroko dostępnymi nagraniami w tym amerykańskim standardzie) spowodowała, że nie przeprowadzone zostały testy na szerszej liczbie próbek. Jest to z całą pewnością wada przedstawionego algorytmu, jednak pewną tego rekompensatą jest wprowadzenie wielu łatwo dostępnych parametrów, które mogą w znaczący sposób wpłynąć na dostrojenie algorytmu do konkretnych (typów) nagrań.

Kolejną wadą algorytmu jest wspólna wada wszystkich algorytmów wykorzystujących adaptacyjność – możliwość destabilizacji i generowania znacznych błędów. Dzieje się tak w przypadku wyjścia poza ścieżkę tonu MTS i przejście w obszar szumów tudzież samej zawartości sygnału audio. Sposobem walki z tą przypadłością w przypadku automatycznej pracy algorytmu dla masowej ilości nagrań byłoby wprowadzenie współczynnika pewności, który umożliwiałby rozpoznanie sytuacji zgubienia śladu pilota i jego ponowne znalezienie w tym samym kroku, bez narażania się na propagację błędu.

Należy też pamiętać, że przy wykorzystaniu zaprezentowanej metody, mamy do czynienia z pewną estymacją ścieżki pilota (kompromis między stabilnością a dokładnością). Dysponując jednak taką estymatą, można – za pomocą innych metod [1][9] – określić częstotliwość tonu pilota bardziej precyzyjnie. W takim wypadku nawet wystąpienie błędów grubych nie spowoduje destabilizacji śledzenia, gdyż nie będzie już stosowana predykcja liniowa kolejnych częstotliwości (jako że będą one znane i ustalone).

6. Bibliografia

- [1] Auger F., Flandrin P., Improving the readability of time-frequency and time-scale representations by the reassignment method, *IEEE Transactions on Signal Processing*, vol. 43, pp. 1068-1089, May 1995.
- [2] Broersen P.M.T., Wensink H.E., On finite sample theory for autoregressive model order selection, *IEEE Transactions on Signal Processing*, vol. 41, pp. 194–204, 1993.
- [3] Czyżewski A.: Dźwięk cyfrowy, pp. 390-392, AOW EXIT Warszawa 2001.
- [4] Czyżewski A., Ciarkowski A., Kaczmarek A., Kotus J., Kulesza M., Maziewski P., DSP Techniques for Determining “Wow” Distortion, *Journal of Audio Engineering Society*, vol. 55, No. 4, pp. 266-284, 2007.
- [5] Czyżewski A., Dziubiński M., Ciarkowski A., Kulesza M., Maziewski P., Kotus J., New Algorithms for Wow and Flutter Detection and Compensation in Audio, 118th Audio Engineering Society Convention, Barcelona, Spain (28-31 May 2005).
- [6] Czyżewski A., Maziewski P., Dziubiński M., Kaczmarek A., Kostek B., Wow detection and compensation employing spectral processing of audio, 117th Audio Engineering Society Convention, San Francisco, CA, USA (28–31 October 2004).
- [7] Czyżewski A., Maziewski P., Dziubiński M., Kaczmarek A., Kulesza M., Ciarkowski A., Kostek B., Methods for Detection and Removal of Parasitic Frequency Modulation in Audio Recordings, AES 26th International Conference, Denver, Colorado, USA (7-9 July 2005).
- [8] de Hoon M.J.L., van der Hagen T.H.J.J., Schoonewelle H., van Dam H., Why Yule-Walker should not be used for autoregressive modelling, Interfaculty Reactor Institute, Delft University of Technology, 1996.
- [9] Fitz K., Haken L., On the Use of Time–Frequency Reassignment in Additive Sound Modeling, *Journal of Audio Engineering Society*, vol. 50, No. 11, pp. 879-893, November 2002.
- [10] Howarth J., Wolfe P., Correction of Wow and Flutter Effects in Analog Tape Transfers, 117th Audio Engineering Society Convention, San Francisco, California, USA (28-31 October 2004).
- [11] Kaczmarek A., Materiały do przedmiotu Akustyka Mowy, Katedra Systemów Multimedialnych, WETI, Politechnika Gdańska, 2007.
- [12] Lagrange M., Marchand S., Rault J.B., Sinusoidal parameter extraction and component selection in a nonstationary model, *Proc. of the 5th Int. Conference on Digital Audio Effects*, Hamburg, September 2002.
- [13] Litwic Ł., Maziewski P., Evaluation of wow defects based on tonal components detection and tracking, *Proc. Int. Symposium of Sound Engineering and Tonmeistering*, pp. 145-150, Kraków, Poland (June 2005).
- [14] Lyons R., Windowing Functions Improve FFT Results, *Test & Measurement World*, str. 37, wydanie 6: 1 czerwca, 1998.
- [15] Masri P., Computer Modeling of Sound for Transformation and Synthesis of Musical Signals, praca doktorska, University of Bristol, 1996.
- [16] McAulay R.J., Quatieri T.F., Speech analysis/synthesis based on a sinusoidalrepresentation, *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 34(4), pp. 744-754, 1986.
- [17] Priestley M.B.: Spectral Analysis and Time Series, Academic Press London 1994.
- [18] Rodet X., Musical Sound Signal Analysis/Synthesis: Sinusoidal + Residual and Elementary Waveform Models, *Proc. IEEE Symp. Time-Frequency and Time-Scale Analysis*,

6. Bibliografia

1997.

- [19] Szwoch G., Materiały do przedmiotu Pomiary elektroakustyczne – "Pomiary urządzeń rejestrujących i odtwarzających", Katedra Systemów Multimedialnych, WETI, Politechnika Gdańska, 2007.
- [20] Urbański B.: Technika Zapisywania i Odczytywania Dźwięków, WKiŁ Warszawa 1978.
- [21] Young, Hugh D.: University Physics, 8th Ed., Addison-Wesley 1992.
- [22] Zieliński T.: Cyfrowe Przetwarzanie Sygnałów: od teorii do zastosowań, WKŁ Warszawa 2007.
- [23] High Frequency Bias Requirements for Magnetic Tape Recording, Sound Talk, vol. 1, nr 2, 1968
- [24] AES Standard AES6-1982 (r2003), Method for Measurement of Weighted Peak Flutter of Sound Recording and Reproducing Equipment (Reaffirmed 2003).
- [25] Understanding Multichannel Television Sound (MTS), Sencore Tech Tips #213, <http://www.sencore.com/custsup/techtips.htm>.
- [26] [http://en.wikipedia.org/wiki/Dbx_\(noise_reduction\)](http://en.wikipedia.org/wiki/Dbx_(noise_reduction)), dostęp z dnia 12.06.2007.
- [27] http://en.wikipedia.org/wiki/Multichannel_television_sound, dostęp z dnia 10.05.2007.
- [28] <http://en.wikipedia.org/wiki/Ntsc>, dostęp z dnia 11.06.2007.
- [29] <http://en.wikipedia.org/wiki/Phonograph>, dostęp z dnia 11.05.2007.
- [30] http://en.wikipedia.org/wiki/Sound_reproduction, dostęp z dnia 12.10.2006.
- [31] <http://en.wikipedia.org/wiki/STFT>, dostęp z dnia 02.06.2007.
- [32] http://en.wikipedia.org/wiki/Wow_%28recording%29, dostęp z dnia 14.05.2007.
- [33] http://en.wikipedia.org/wiki/Wow_and_flutter_measurement, dostęp z dnia 20.02.2007.
- [34] <http://history.sandiego.edu/gen/recording/jacquard1.html>, dostęp z dnia 10.09.2006.
- [35] <http://history.sandiego.edu/gen/recording/precursors.html>, dostęp z dnia 10.09.2006.
- [36] <http://history.sandiego.edu/gen/recording/scott.html>, dostęp z dnia 10.09.2006.
- [37] <http://hyperphysics.phy-astr.gsu.edu/hbase/solids/hyst.html>, dostęp z dnia .
- [38] http://nl.wikipedia.org/wiki/Wow_en_flutter, dostęp z dnia 20.02.2007.
- [39] http://pl.wikipedia.org/wiki/Dyskretna_transformata_Fouriera, dostęp z dnia 12.06.2007.
- [40] <http://pl.wikipedia.org/wiki/FFTW>, dostęp z dnia 12.04.2007.
- [41] <http://pl.wikipedia.org/wiki/Heterodyna>, dostęp z dnia 19.05.2007.
- [42] http://pl.wikipedia.org/wiki/Prawo_Webera-Fechnera, dostęp z dnia 05.06.2007.
- [43] <http://pl.wikipedia.org/wiki/PrzydźBAwi%C4%99k>, dostęp z dnia 05.06.2007.
- [44] http://pl.wikipedia.org/wiki/Szum_r%C3%B3w%20C5%BCowy, dostęp z dnia 04.06.2007.
- [45] http://pl.wikipedia.org/wiki/Transformacja_Fouriera, dostęp z dnia 12.06.2007.
- [46] <http://prestospace.org/>, dostęp z dnia 12.06.2007.
- [47] <http://www.funky-lady.com/pl/artykuly/index.php?id=ewolucja>, dostęp z dnia 06.04.07.

7. Załączniki

Załączniki obejmują kody źródłowe głównej funkcji `ntsc(s, fs)` realizującej śledzenie pilota MTS w sygnale `s` o częstotliwości próbkowania `fs`. Druga funkcja to funkcja wykorzystywana wewnątrz pierwszej, `estFreqFromMax(max, lval, rval)`, realizująca interpolację paraboliczną (a konkretnie zwracająca ułamek korygujący) dla 3 sąsiadujących punktów o wartościach odpowiednio `lval`, `max` i `rval`.

```
function [b, mxi] = ntsc(s, fs)

% in:
% s - signal
% fs - sampling frequency
%
% out:
% b - calculated spectrogram of the signal s
% mxi - MTS pilot tone track (vector of frequency indices)

winSize=1024;
distSize=512+256;

nominal_pilot = 15734 / fs * winSize;
pilot_filtering_distance = 1000 / fs * winSize;

steadiness_margin = 200 / fs * winSize;

max_slope = 800 / fs * winSize;

[a,b]=size(s);
if(b==1)
    s=s';          % has to be line, not column vector
end

%----- HIGH PASS FILTER CUT-OFF -----
cutoff=14000;
ci=round((winSize/2+1) * cutoff/(fs/2));

%----- FREQ DOMAIN SMOOTHING ORDER ----
%N=winSize/32;
```

7. Załączniki

```
N=3;
k=ones(1, N)/N;
%----- TIME DOMAIN SMOOTHING ORDER -----
N2=3;
k2=ones(1, N2)/N2;
%-----

steps=floor((length(s)-winSize)/distSize);

calm(1:steps) = 1; % an initialization to prevent filtering pilot at the beginning

win=hanning(winSize)';

b=zeros(winSize/2+1, steps);

%----- PREEMPHASIS -----
nrm=1:winSize/2+1;
nrm=nrm/(winSize/2+1);
nrm=nrm.^3;

%----- FINDING ACTUAL PILOT TONE -----
sp=abs(fft(s(1:winSize).*win));
sp=sp(1:winSize/2+1).*nrm;

actual_pilot = sum(sp(round(nominal_pilot-
pilot_filtering_distance/4):round(nominal_pilot+pilot_filtering_distance/4)).*[roun
d(nominal_pilot-
pilot_filtering_distance/4):round(nominal_pilot+pilot_filtering_distance/4)])/sum(s
p(round(nominal_pilot-
pilot_filtering_distance/4):round(nominal_pilot+pilot_filtering_distance/4)))-1;

actual_pilot_filtering_margin = 100 / fs * winSize; % filtering +-100 Hz around
actual_pilot frequency

actual_pilot_filter(1:length(round(actual_pilot-
actual_pilot_filtering_margin):round(actual_pilot+actual_pilot_filtering_margin)))
= abs((round(actual_pilot-
actual_pilot_filtering_margin):round(actual_pilot+actual_pilot_filtering_margin)) -
actual_pilot)/max(abs((round(actual_pilot-
actual_pilot_filtering_margin):round(actual_pilot+actual_pilot_filtering_margin)) -
actual_pilot))*0.9;

%actual_pilot_filter=[0.75 0.5 0.1 0.2 0.85];

%figure; % uncomment this when using partial graphs

for i=1:steps
    ['Progress: ' num2str(100*(i-1)/(steps-1))]
```


7. Załączniki

```

start=(i-1)*distSize+1;
stop=start+winSize-1;

sig=s(start:stop);

sp=sig.*win;          % apply window function
sp=abs(fft(sp));      % fft
sp=sp(1:winSize/2+1);
original_spectrum(:,i)=sp';
sp=sp./nrm; % apply preemphasis
energy = sum(sp); % zero spectral moment - energy of signal
if energy < 0.01 % checking to avoid division by zero
    mxi(i) = NaN;
else
    if i<10
        sp(1:ci)=0; % HP filtering
    else
        %----- 2nd ORDER LINEAR PREDICION OF PILOT TONE FREQUENCY -----
        a = ( mxi(i-3) + mxi(i-1) ) / 2 - mxi(i-2);
        be = mxi(i-2) - mxi(i-3) - a;
        ce = mxi(i-3);
        predicted_fq = round( a * 9 + be * 3 + ce);

        %----- 1st ORDER LINEAR PREDICION OF PILOT TONE FREQUENCY -----
        % predicted_fq = round( 2 * (mxi(i-1) - 1) - (mxi(i-2) - 1));

        %----- 0th ORDER LINEAR PREDICION OF PILOT TONE FREQUENCY -----
        % predicted_fq = round(mxi(i-1) - 1);

        %----- LOWER MARGIN AS CERTAIN FRACTION OF PREDICTED FQ -----
        margin_dol = 0.075 * predicted_fq;
        margines_dol = round(margin_dol);
        margines = round(margines_dol * 4 / 3);
        % lower margin - 7,5% of predicted fq, upper - 10%

        %----- WEIGHT FUNCTION INSIDE MARGIN -----
        mask = 1:margines/margines_dol:margines+1;
        mask(margines_dol+2:margines+margines_dol+1) = margines:-1:1;
        % margines+1 - central element (predicted fq)
        mask = mask / (margines + 1); % normalization
    
```

7. Załączniki

```
%----- ALTERNATIVE WEIGHT FUNCTION INSIDE MARGIN -----
%
%   sinmask = 1:margines*2+1;
%   sinmask = sinmask / (2*margines + 1);
%   sinmask = sinmask * pi;
%   mask = sin(sinmask);

mask = mask.^2;

%----- CHECKING WHETHER SIGNAL IS CALM -----
dist_to_pilot = abs(predicted_fq - actual_pilot);
if (dist_to_pilot < steadiness_margin)
    calm(i) = 1;      % incrementing steadiness variable
else
    calm = calm * 0;  % resetting steadiness variable
end;

%----- CHECKING WHETHER PILOT NEEDS TO BE FILTERED -----
if ( dist_to_pilot < pilot_filtering_distance && sum(calm) < 10)
    sp(round(actual_pilot-
actual_pilot_filtering_margin):round(actual_pilot+actual_pilot_filtering_margin))=s
p(round(actual_pilot-
actual_pilot_filtering_margin):round(actual_pilot+actual_pilot_filtering_margin)).*
actual_pilot_filter;
end;

%----- CLEARING UNNECESARRY PARTS OF SPECTRUM -----
sp(1:predicted_fq - margines_dol) = 0;
sp(predicted_fq - margines_dol:predicted_fq + margines) =
sp(predicted_fq - margines_dol:predicted_fq + margines).*mask;
sp(predicted_fq + margines:winSize/2+1) = 0;
end;

sp=sp.*sp;    % apply spectrum expansion
sp=sp.*sp;    % apply spectrum expansion

%----- FREQ-DOMAIN SMOOTHING -----
%sp=movingAvg(sp, N); %function rewritten in order to simplify C++
implementation
sp=filtfilt(k,1,sp); %Matlab function
b(:,i)=sp';

%----- PEAKs DETECTION -----
if (i>2)
    sp=(b(:,i-2)+b(:,i-1)+b(:,i))/3;    % time-domain smoothing
```

7. Załączniki

```

else
    sp=b(:,i);
end;
if i<10
    [vv,ii]=max(sp);
    if i==1
        if (ii<winSize/2+1)
            ii = ii + estFreqFromMax(sp(ii), sp(ii-1), sp(ii+1));
            % find fractional index (parabolic interpolation)
        end;
    elseif (abs(ii-mxi(i-1)) < max_slope) && (ii<winSize/2+1) && (ii>1)
        ii = ii + estFreqFromMax(sp(ii), sp(ii-1), sp(ii+1));
        % find fractional index (parabolic interpolation)
    else
        ii = mxi(i-1);
    end;
else
    ii = sum(sp(predicted_fq - margines:predicted_fq +
margines)'.*[predicted_fq - margines:predicted_fq + margines])/sum(sp(predicted_fq
- margines:predicted_fq + margines));
    end;
    mxi(i)=ii;
end;

%----- PARTIAL GRAPHS -----
%
%   if (mod(i,1000) == 0)
%       biasFq=(fs/2)*(mxi-1)/(winSize/2);
%       F=0:winSize/2; F=F/max(F); F=F*fs/2;
%       T=0:i-1; T=(T*distSize+winSize/2)/fs;
%
%       imagesc(T,F/1000, log(abs(original_spectrum)));
%       hold on
%       plot(T-winSize/fs, biasFq/1000,'k');
%       xlabel('Time (s)')
%       ylabel('Frequency (kHz)')
%       title('Track of the NTSC MTS pilot tone')
%       axis xy;
%       hold off
%       ['Press any key']
%       pause
%   end;

```

7. Załączniki

```
end

%----- SPECTROGRAM -----

biasFq=(fs/2)*(mxi-1)/(winSize/2);
F=0:winSize/2; F=F/max(F); F=F*fs/2;
T=0:steps-1; T=(T*distSize+winSize/2)/fs;

figure;imagesc(T,F/1000, log(abs(original_spectrum)));
hold on
xlabel('Time (s)')
ylabel('Frequency (kHz)')
title('Track of the NTSC MTS pilot tone')
axis xy;
hold off
['Press any key']
pause
figure;imagesc(T,F/1000, log(abs(original_spectrum)));
hold on
plot(T-winSize/fs, biasFq/1000,'k');
xlabel('Time (s)')
ylabel('Frequency (kHz)')
title('Track of the NTSC MTS pilot tone')
axis xy;
hold off
['Press any key']
pause
figure;imagesc(T,F/1000, log(abs(b)));
hold on
plot(T-winSize/fs, biasFq/1000,'w');
xlabel('Time (s)')
ylabel('Frequency (kHz)')
title('Track of the NTSC MTS pilot tone')
axis xy;
hold off
```

7. Załączniki

```
function [indMax]=estFreqFromMax(max,lval,rval);  
c=max;  
b=(rval-lval)/2;  
a=lval + (rval-lval)/2 - max;  
if a==0  
    indMax=0;  
else  
    indMax = -b/(2*a);  
end
```