

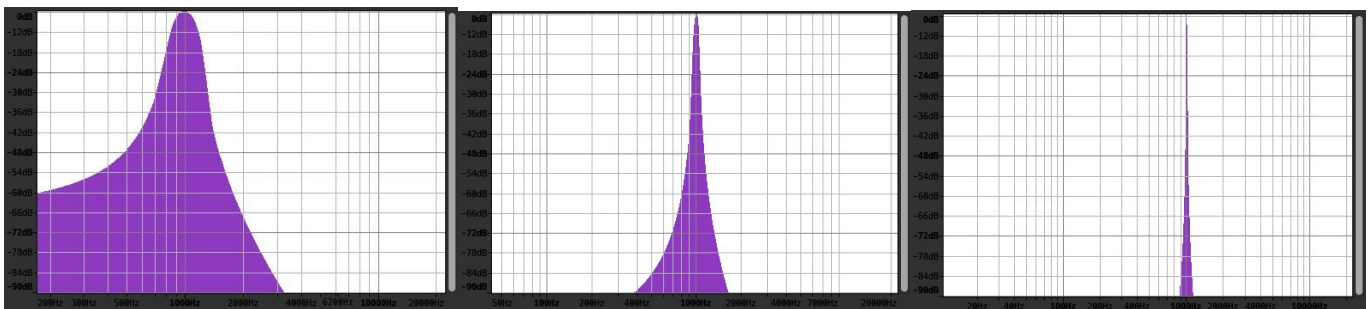
WMM — Dźwięk — Laboratorium 1:

Marfenko Mykhailo 323558

1 Analiza widma i spektrogramu dźwięków

1.1 Pojedyncze tony

- Wygeneruj ton o częstotliwości 1000 Hz i amplitudzie 0,5.
- Wyświetl widmo sygnału (Analizuj - Narysuj widmo) i sprawdź, jaki wpływ na wygląd widma ma zmiana rozmiaru okna analizy widmowej (Rozmiar: ...)



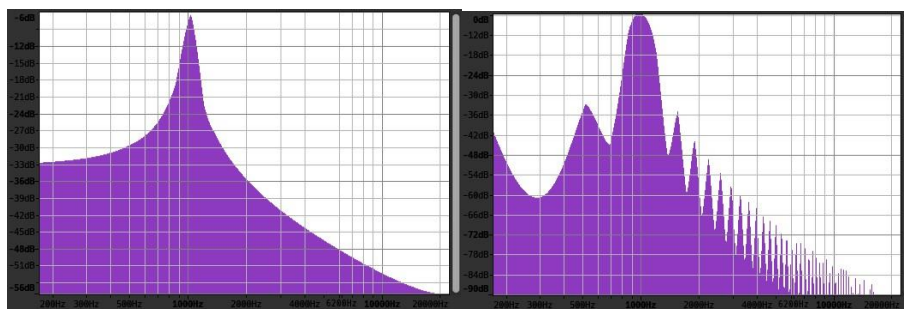
(a) Widmo dla okna = 256

(b) Widmo dla okna = 1024

(c) Widmo dla okna = 4096

Dla małych rozmiarów okna, widmo jest rozlane na inne wartości. W miarę zwiększania tego rozmiaru, widmo wydaje się być bardziej skoncentrowane wokół zamierzonej wartości.

- Sprawdź, jaki wpływ na widmo sygnału ma rodzaj zastosowanego okna (Funkcja: ...)



(d) Widmo dla okna prostokątnego

(e) Widmo dla okna Barletta

Wszystkie widma powyżej mają rozmiar okna 256 — najłatwiej widać przy nich różnice.

Okno prostokątne najbardziej rozlewa widmo.

Okno Barletta umieszczam natomiast dlatego, że podoba mi się to, jak śmiesznie ono fałuje.

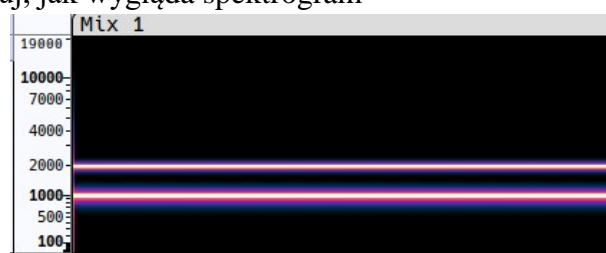
- Klikając czarny trójkąt przy nazwie ścieżki zmień widok na „Spektro-gram”
- Zaobserwuj, jak wygląda spektrogram dla wygenerowanego tonu (pojedynczej częstotliwości)



(f) Spektrogram dla wygenerowanego sygnału

Spektrogram ma jedną linię, która jest skupiona wokół zadanej częstotliwości.

- Wygeneruj kolejny ton (będąc „odklikniętym” z poprzedniej ścieżki), o częstotliwości 2000 Hz i amplitudzie 0,3
- Zaznacz obie ścieżki i w menu wybierz Ścieżki – Miksuj – Miksuj i renderuj
- Zaobserwuj, jak wygląda spektrogram



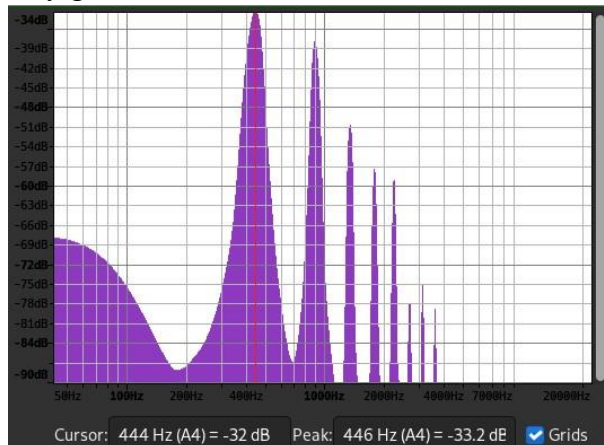
(g) Spektrogram dla połączonych tonów

Spektrogram jest połączeniem dwóch poprzednich, bez żadnego nakładania się na siebie tonów. Obie linie są stałe, choć ta na $f = 2000$ Hz jest cieńsza (widoczna mniejsza amplituda).

1.2 Dźwięki muzyczne

- Zaimportuj ścieżkę „flet.wav”

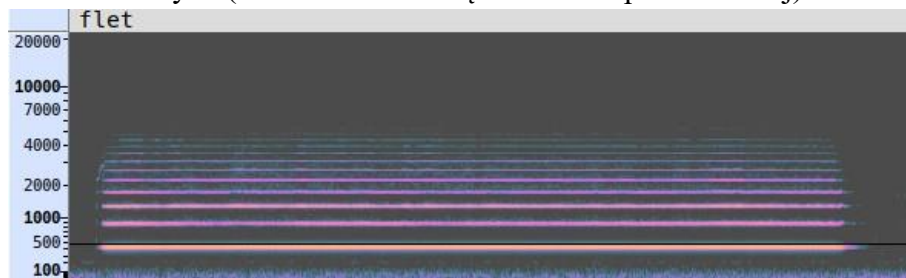
- Posłuchaj i wyświetlając widmo sygnału, sprawdź, jaka jest częstotliwość dźwięku, który gra flecistka



(h) Widmo dźwięku granego przez flecistkę

Częstotliwość to 446 Hz (peak na widmie).

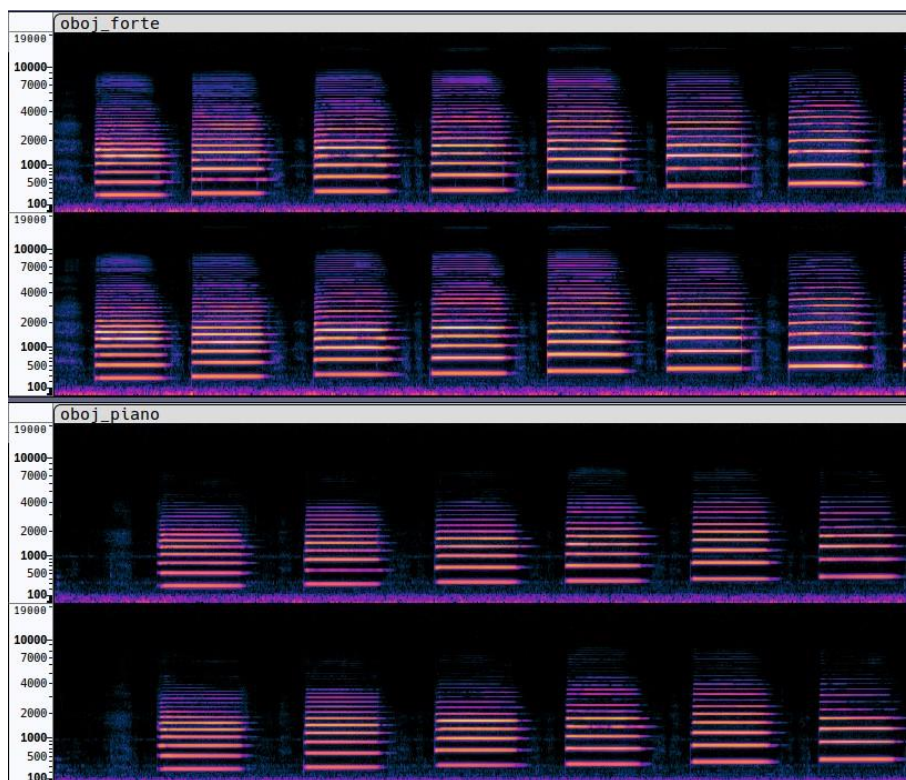
- Zaobserwuj, jak wygląda na spektrogramie dźwięk instrumentu muzycznego – zwróć uwagę na występowanie wielu częstotliwości harmoniczných (wielokrotności częstotliwości podstawowej)



(i) Spektrogram dźwięku granego przez flecistkę

Rzeczywiście widoczne są częstotliwości harmoniczne.

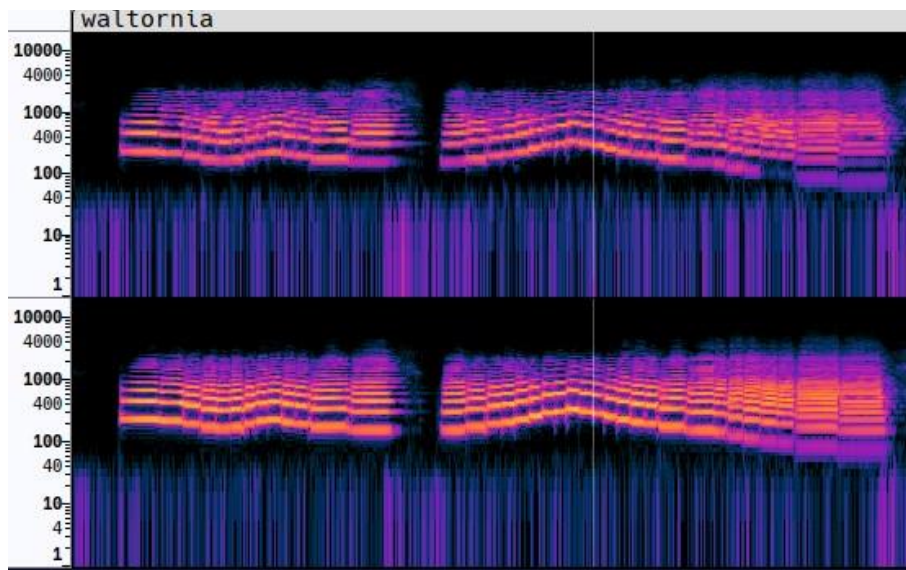
- Zaimportuj ścieżkę „oboj_piano.wav” oraz osobno „oboj_forte.wav”
- Posłuchaj kolejno każdej ze ścieżek i zobacz na spektrogramie, w jakisposób różnica w barwie dźwięku pomiędzy instrumentem grającym cicho (piano) i głośno (forte) widoczna jest na spektrogramie



(j) Spektrogramy powyższych ścieżek

Dźwięk grany głośniej jest jaśniejszy (ponownie, większa amplituda) oraz widzimy na nim więcej częstotliwości harmoniczných.

- Zimportuj ścieżkę „waltornia.wav”
- Zmień pionową skalę spektrogramu z liniowej na logarytmiczną (kliknięcie prawym przyciskiem myszy na skali po lewej stronie wykresu) i słuchając nagrania zobacz, jak na spektrogramie widoczna jest melodia, którą gra waltornista.

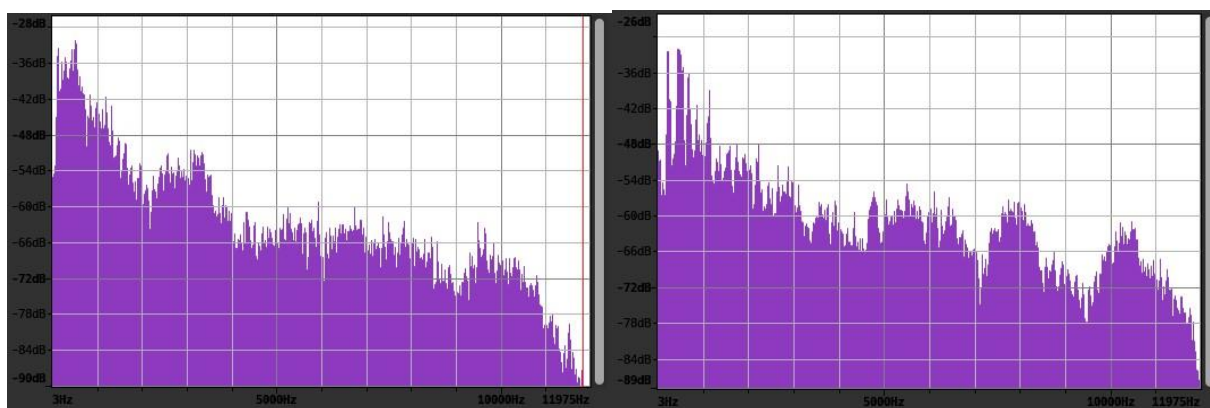


(k) Spektrogram nagrania gry waltornisty

Względem fletu — gdzie gra była ciągła przy każdym dźwięku — widzimy różnice w melodii (falowanie linii na spektrogramie).

1.3 Mowa

- Zaimportuj ścieżkę „mowa_mezczyzna.wav” oraz osobno „mowa_kobieta.wav”
- Na podstawie widma obu sygnałów przeanalizuj, w jakich zakresach częstotliwości więcej energii ma dźwięk mowy męskiej, a w jakich mowy kobiecej (lepiej będzie to widać ustawiając liniową skalę poziomą wykresu)



(l) Głos mężczyzny

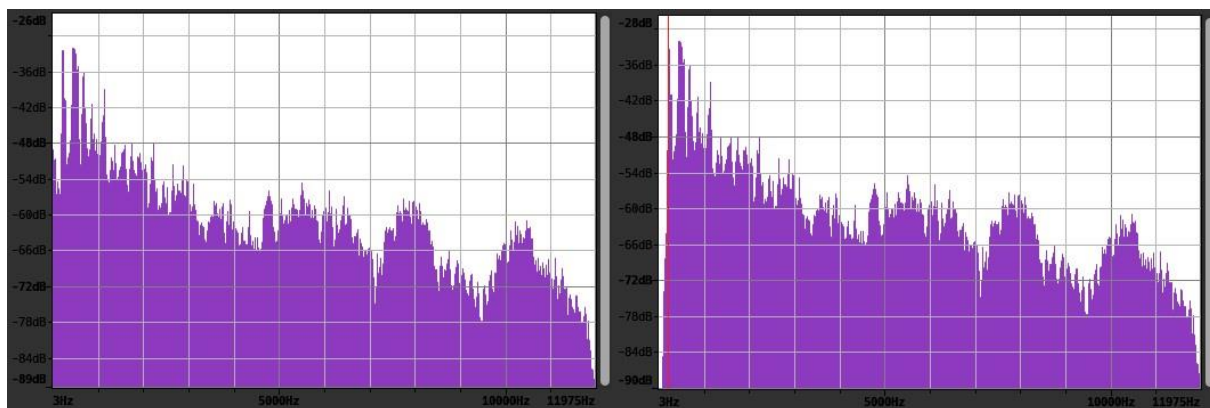
(m) Głos kobiety

Mowa męska ma więcej wartości po lewej stronie wykresu (dźwięki niższe), a kobieca po prawej (dźwięki wyższe).

2 Edycja dźwięków

2.1 Mowa

- Otwórz ścieżkę „mowa_kobieta.wav” oraz osobno „mowa_mezczyzna.wav”
- Zastosuj filtrację górnoprzepustową (zaznacz ścieżkę, Efekt – Filtr górnoprze-pustowy) do ścieżki z mową kobiecą, tj. odfiltruj dolne częstotliwości dźwięku, stosując ustawienia: częstotliwość graniczna 200 Hz, rolloff 24 dB/oktawę
- Zobacz różnicę w widmie przed i po filtracji. Posłuchaj ścieżki mowykobiecej po filtracji – czy utracona została jakaś istotna część sygnału?



(n) Głos kobiety przed filtracją

(o) Głos kobiety po filtracji

Wartości poniżej 200 Hz zostały w większości “wycięte”, choć nie w całości, bo filtry nie działają tak precyzyjnie. Zastosowany filtr nie ma jednak wpływu na wyrazistość mowy i nie słyszę żadnej różnicy w jej brzmieniu.

- Zastosuj taką samą filtrację górnoprzepustową (te same parametry), do mowy męskiej. Zaobserwuj widmo przed i po filtracji. Posłuchaj ścieżki

mowy męskiej – czy przy filtracji utracona została istotna część sygnału w porównaniu do mowy kobiecej?

Tak, mowa męska odfiltrowana w ten sposób brzmi bardzo “sztucznie” i traci swoją głębię. Jest to spowodowane stratą charakterystycznych dla niej niższych częstotliwości, których utrata w mowie kobiecej nie była tak dotkliwa.

- Do mowy kobiecej zastosuj dodatkowo filtrację dolnoprzepustową, tj. odfiltruj wysokie częstotliwości sygnału, stosując ustawienia np: 8 000 Hz, 24 db/okt

Tym razem, mowa kobieca brzmi sztucznie — analogicznie do poprzedniego punktu z mową męską — co jest spowodowane odfiltrowaniem charakterystycznych dla niej wyższych częstotliwości.

- Stosując różne wartości częstotliwości granicznej filtru, sprawdź, przy jakiej częstotliwości utracona zostaje wyrazistość mowy, a przy jakiej zrozumiałość.

Dla filtru górnoprzepustowego (reszta ustawień jak wyżej):

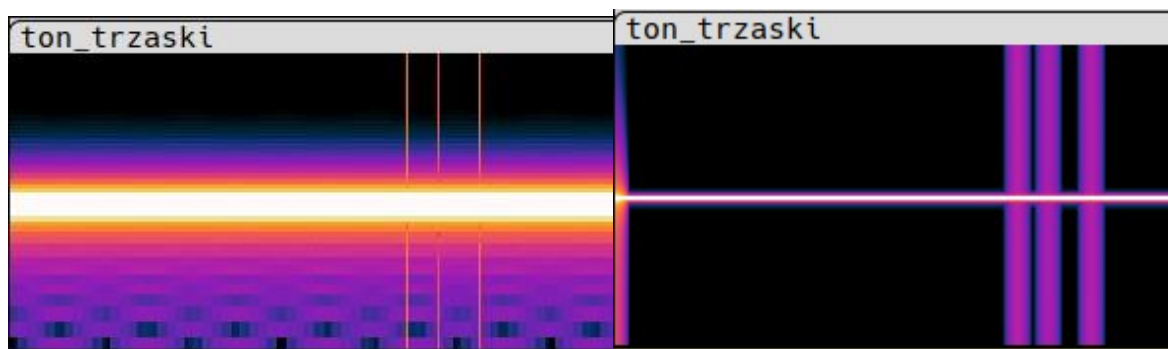
mowa	zanik wyrazistości	zanik zrozumiałości
męska	500 Hz	5 kHz
kobieca	1 kHz	7 kHz

Dla filtru dolnoprzepustowego (reszta ustawień jak wyżej):

mowa	zanik wyrazistości	zanik zrozumiałości
męska	1,5 kHz	500 Hz
kobieca	4 kHz	500 Hz

2.2 Eliminacja zakłóceń

- Zaimportuj ścieżkę „ton_trzaski.wav” i posłuchaj – w nagraniu znajduje się ton i zakłócenia (trzaski)
- Przełącz widok ścieżki na spektrogram, zaobserwuj występowanie trzasków, które nie były widoczne na przebiegu czasowym sygnału, a następnie ponownie rozwiń menu przy nazwie ścieżki i włącz „Ustawienia spektrogramu”
- Sprawdź, jak rozmiar okna analizy fft wpływa na wygląd spektrogramu. Zmieniaj po kolei rozmiar okna z 1024 na coraz mniejsze, a później coraz większe i zobacz, który rozmiar jest najlepszy dla uzyskania największej rozdzielczości w dziedzinie czasu (pozioma oś), a który zapewnia najlepszą rozdzielczość w dziedzinie częstotliwości (pionowa oś).



(a) Spektrogram dla $n=128$

(b) Spektrogram dla $n=4096$

Najlepiej dla analizy w dziedzinie czasu nadaje się $n=128$ ze względu na idealne wyłapanie pionowych prążków. Natomiast do analizy w dziedzinie częstotliwości, dobrze spisuje się $n=4096$, które dobrze pokazuje częstotliwości, jednocześnie nie zlewając wszystkich okolicznych pionowych prążków w jeden.

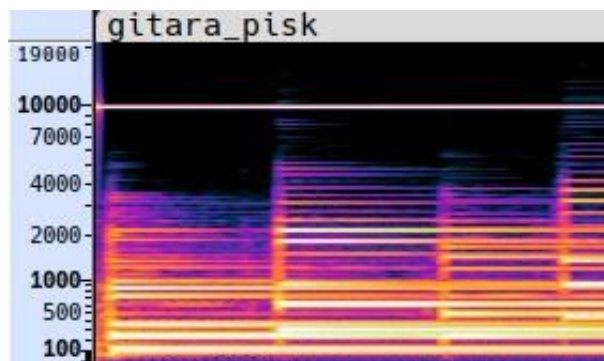
Bierze się to stąd, że rozdzielczość w d. częstotliwości jest wprost proporcjonalna do n , a rozdzielczość w d. czasu odwrotnie proporcjonalna.

- Przy optymalnych ustawieniach okna analizy dla widoczności trzasków (których czas trwania jest bardzo krótki) spróbuj usunąć zakłócenia w następujący sposób:
 1. W widoku spektrogramu zaznacz fragment, w który występuje trzask
 2. Przełącz widok na przebieg czasowy
 3. W menu wybierz Zaznacz – Na miejscach przejść przez zero (tę opcję pozwoli uniknąć nieciągłości sygnału na krańcach zaznaczenia)
 4. Usuń zaznaczony fragment (pojedynczy trzask) naciskając delete
 5. Posłuchaj nowej wersji ścieżki

W wyniku powtarzania tej operacji, udało się precyzyjnie wyciąć trzaski.

- Zaimportuj ścieżkę „gitara_pisk.wav”
- Zaobserwuj na spektrogramie, jakiego rodzaju zakłócenie występuje w tym nagraniu i spróbuj je usunąć w następujący sposób:
 1. Zaznacz odpowiedni fragment spektrogramu (dany zakres częstotliwości w całym czasie trwania zakłócenia)

2. Spróbuj usunąć zakłócenie używając filtra notch (Efekt – FiltrNotch) o odpowiednich parametrach
3. Posłuchaj nowej wersji ścieżki



(c) Spektrogram nagrania gitary

Na spektrogramie widzimy, że jest to pisk o stałej częstotliwości 10 kHz, rozciągnięty po całej długości nagrania.

Filtr Notch z parametrami 10 kHz i $Q=1$ pozbył się szumu. Oczywiście, większe Q ograniczyłoby wycięte częstotliwości i zawężyło działanie filtra, jednak moje drewniane ucho nie widzi specjalnej różnicy, więc nie będę na siłę tego opisywał.

3 Lateralizacja źródła dźwięku

- Do tego zadania niezbędne są słuchawki!
- Zaimportuj ścieżkę „lektor.wav”
- Zaznacz zaimportowaną ścieżkę i powiel ją (Edycja – Powiel)
- Jedną ze ścieżek ustaw w panoramie (suwak L – P pod nazwą ścieżki) na prawo, a drugą na lewo
- Kliknij kursorem w jakieś miejsce ścieżki, gdzie amplituda sygnału jest duża
- Klikając „lupkę” przybliż przebieg czasowy tak, aby na skali były tysięczne części sekundy
- Przesuń jedną ze ścieżek w prawo lub w lewo o tysięczne części sekundy

- Posłuchaj, jak po przesunięciu zmienia się położenie pozornego źródła dźwięku
- Jeśli przesuniesz ścieżkę o zbyt dużą wartość, zniknie wrażenie przesuwania się źródła



(d)

Opóźnienie lewej ścieżki o 0.002 ms, prowadzi do wrażenia, że głos dochodzi z prawej strony. Jest to związane z tym, jak ludzkie uszy rejestrują dźwięk (głos z prawej zostanie najpierw wyłapany przez prawe ucho, dopiero potem przez lewe).

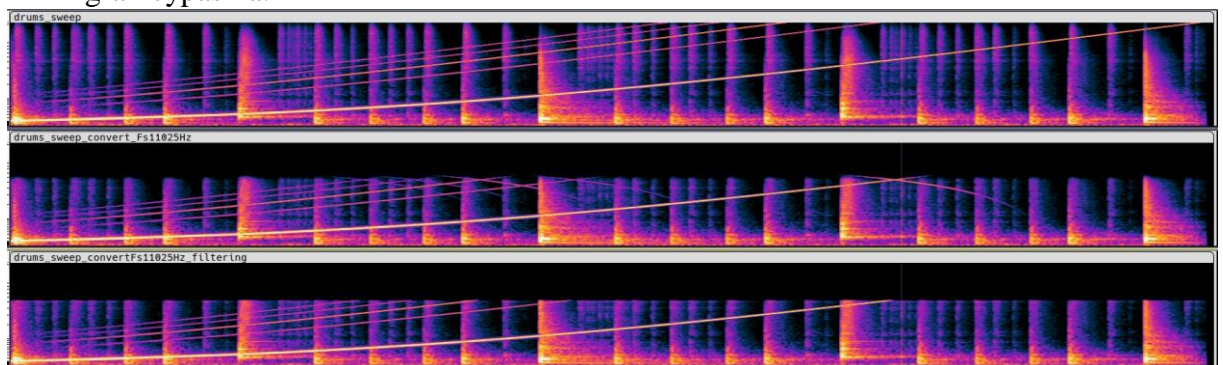
Większe opóźnienia rzeczywiście dają już efekt echa (choć te tylko trochę większe dają dziwny efekt ni to echa, ni to zmiany położenia).

4 Próbkowanie i kwantyzacja

- zaimportuj pliki “drums_sweep.wav”, “drums_sweep_convert_Fs11025Hz.wav” i “drums_sweep_convertFs11025Hz_filtering.wav”
- posłuchaj i oceń brzmienie każdego pliku (drugi z nich to konwersja $f_s = 44\,100\text{ Hz}$ na $11\,025\text{ Hz}$ bez odpowiedniej filtracji sygnału, a trzeci to konwersja z $44\,100\text{ Hz}$ na $11\,025\text{ Hz}$, ale z włączoną filtracją anty-aliasingową),

Zgodnie z warunkiem Nyquista, w plikach po konwersji, nie może się znaleźć częstotliwość wyższa niż $11025/2\text{Hz} \approx 5,5\text{kHz}$, więc naturalnie obcięte zostały wyższe od tej częstotliwości, przez co nagrania po konwersji nie kłują tak w uszy.

- posłuchaj, czy słyszysz składowe, które nie występowały w oryginalnym sygnale? W pliku bez należytej filtracji pojawia się opadający dźwięk pod koniec nagrania.
- zobaczyć na spektrogramie w jaki sposób wygląda aliasing na granicy pasma.



(e)

Na drugim i trzecim nagraniu widzimy ucięcie częstotliwości wyższych (wcześniej opisany warunek) oraz opadający dźwięk z nagrania drugiego, a także pozbycie się go z nagrania trzeciego.

- zaimportuj pliki “drums.wav”, “drums_8bits.wav” i “drums_8bits_dith1.0.wav”

- posłuchaj i oceń wybrzmiewanie uderzenia stopy i werbla oraz brzmieniatalerzy po kwantyzacji do 8 bitów

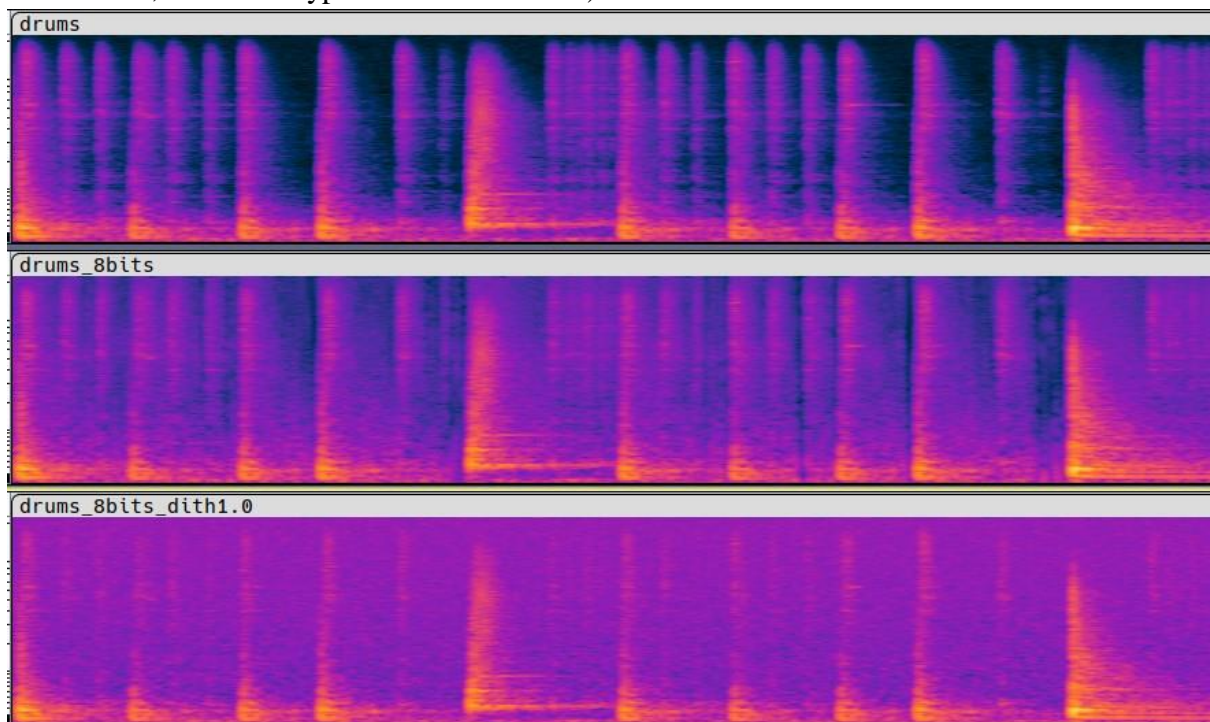
Werbel i talerze nie są zbyt dotkliwe, natomiast uderzenia stopą zostawiają po sobie szumy.

- posłuchaj i oceń wybrzmiewanie uderzenia stopy i werbla oraz brzmieniatalerzy po kwantyzacji do 8 bitów, ale z dodaniem sygnału dither’a.

Czy odzyskało wybrzmienie, ale kosztem mniejszego SNR?

Tak, teraz szum jest bardziej stały i nie “podbija” się tak przy uderzeniach stopą.

- obejrzyj i porównaj spektrogramy każdego sygnału (najlepiej ustawienia: Gain-0dB, Range-120dB, Max Freq-22000Hz, Windows size - 4096, Windows type: Blackman Harris)



(f)

Tak jak wcześniej dało się usłyszeć, a teraz widzieć, na pierwszym nagraniu nie ma szumów, na drugim są one “skokowo”, a na trzecim są one równomiernie rozłożone.