

LAB 02: LOKALIZACJA ŹRÓDŁA DŹWIĘKU, CZ. 1

Karol Działowski

nr albumu: 39259
przedmiot: Sygnały akustyczne

Szczecin, 23 marca 2021

Spis treści

1	Cel laboratorium	2
2	Wygenerowanie sygnału z opóźnionym kanałem	2
3	Wyznaczanie przesunięcia na podstawie sygnału	5
3.1	Korelacja wzajemna	5
3.2	GCC-PHAT	6
4	Kąt źródła dźwięku	7
5	Międzyuszna różnica poziomów	7
6	Badania wygenerowanych sygnałów	7
6.1	Wnioski	9
7	Badanie dla indywidualnego pliku	9
	Bibliografia	11

1 Cel laboratorium

Celem laboratorium było zapoznanie się z podstawowymi pojęciami wykorzystywanymi w lokalizacji źródła dźwięku. W tym celu należało wygenerować sygnał stereo z opóźnionym jednym kanałem. Przy pomocy korelacji wzajemnej i GCC-PHAT wyznaczyć przesunięcie kanałów w czasie, aby obliczyć kąt źródła dźwięku. Dodatkowo należało zaimplementować wzór na ILD (ang. Interaural level difference).

2 Wygenerowanie sygnału z opóźnionym kanałem

Pierwszym zadaniem było wygenerowanie sygnału sinusoidalnego. W tym celu wybrano niską częstotliwość, tj. 440Hz.

Częstotliwość tą wybrano z dwóch powodów. Po pierwsze dla sygnałów o wyższej częstotliwości po znacznym przesunięciu w fazie nie będzie możliwe wyznaczenie tego przesunięcia na podstawie samego sygnału.

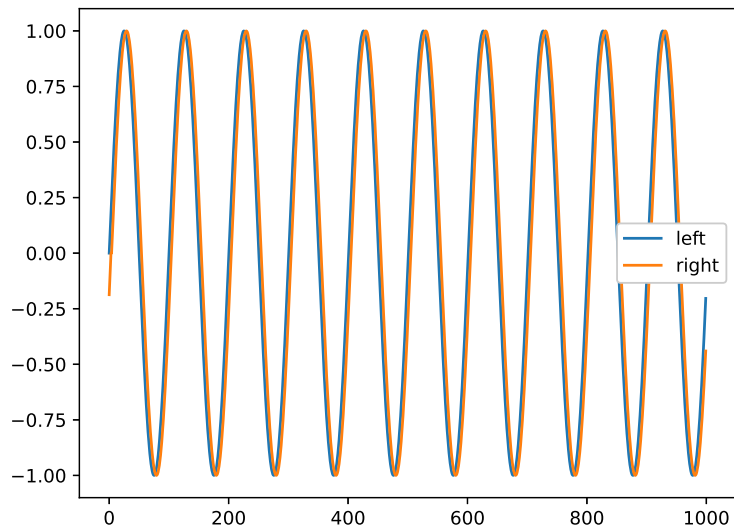
Drugi argument wybrania takiej częstotliwości wynika ze słów Moore'a z książki Wprowadzenie do psychologii słyszenia. Według Moore'a niejednoznaczności na podstawie międzyusznych różnic czasu zaczynają pojawiać się wtedy, gdy pół długości fali ma ok. 23 cm, tj. wtedy, gdy częstotliwość fali wynosi około 750 Hz. Wtedy do obu uszu dochodzą fale, które mają przeciwne fazy. Nie jest możliwa jednoznaczna ocena położenia źródła dźwięku, ponieważ fala dochodząca do jednego ucha może o pół cyklu wyprzedzać lub być o pół cyklu opóźniona względem ucha drugiego. Różnice fazowe stają się bardzo mylące dla częstotliwości większych od 1500 Hz [1].

Kod źródłowy 1: Stworzenie sygnału 440Hz o długości 5s z częstotliwością próbkowania 44100Hz
Źródło: Opracowanie własne

```
1 def generate_sin(sr=44100, frequency=440, length=5):
2     t = np.linspace(0, length, sr * length)
3     y = np.sin(frequency * 2 * np.pi * t)
4
5     return y, sr
6
7 shift_seconds = 0.002
8 shift = int(shift_seconds * sr)
9 y_shifted = np.roll(y, shift=shift)
10 y_stereo_shifted = np.column_stack((y, y_shifted))
```

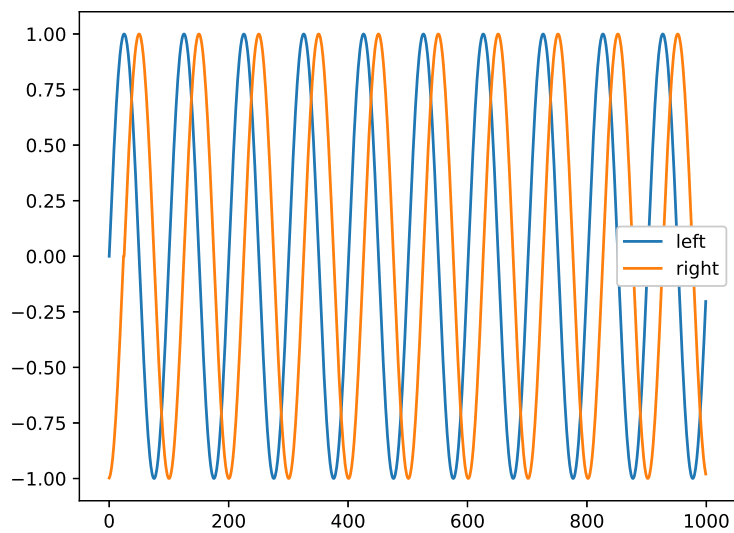
Wygenerowano sygnały z jednym kanałem przesuniętym o {0.0001, 0.00105, 0.001525, 0.002} sekundy, co było równoważne przesunięciu o {4, 25, 46, 67, 88} próbek.

Wykresy sygnałów przedstawiono na poniższych rysunkach.



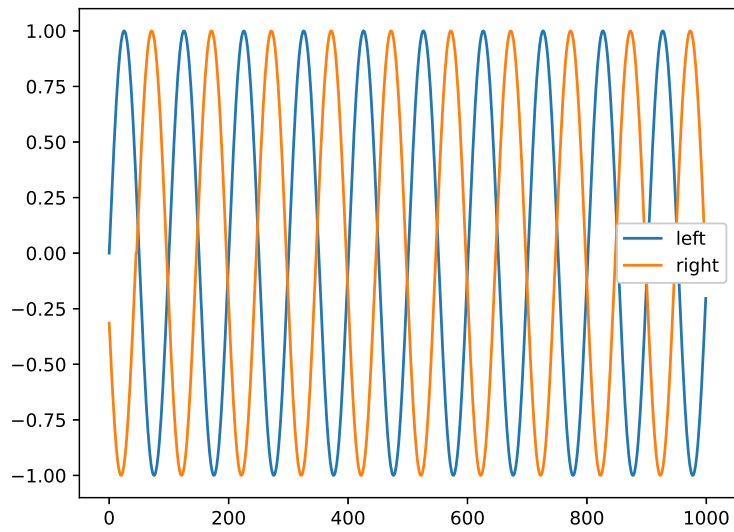
Rysunek 1: Sygnał z prawym kanałem przesunięty o 0.0001 s

Źródło: Opracowanie własne



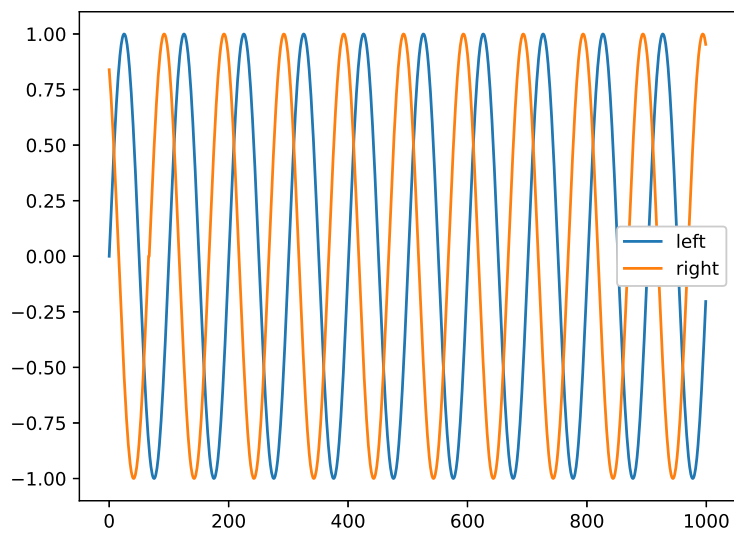
Rysunek 2: Sygnał z prawym kanałem przesunięty o 0.000575 s

Źródło: Opracowanie własne



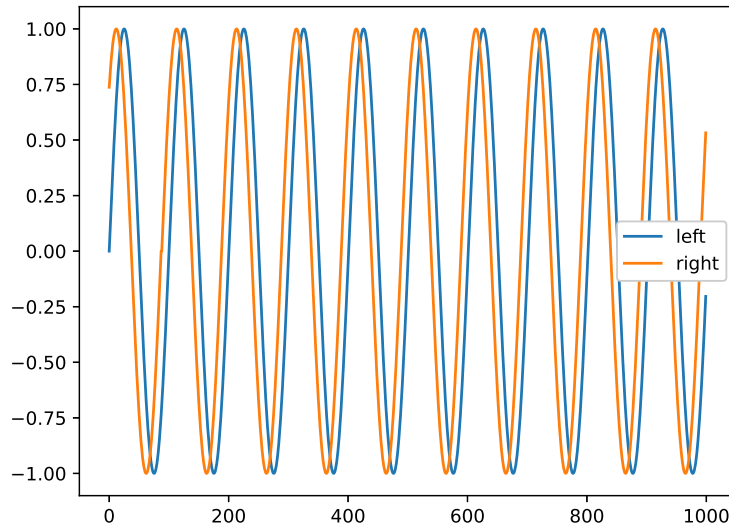
Rysunek 3: Sygnał z prawym kanałem przesuniętym o 0.00105 s

Źródło: Opracowanie własne



Rysunek 4: Sygnał z prawym kanałem przesuniętym o 0.001525 s

Źródło: Opracowanie własne



Rysunek 5: Sygnał z prawym kanałem przesuniętym o 0.002 s

Źródło: Opracowanie własne

3 Wyznaczanie przesunięcia na podstawie sygnału

Międzysłyszna różnica czasów - ITD (ang. Interaural time difference). Atrybut ten wyznacza się na podstawie porównania sygnałów docierających do uszu. Do porównania sygnałów i określenia opóźnienia między nimi stosuje się funkcję autokorelacji wzajemnej (lub jej wariant liczony w dziedzinie częstotliwości nazywany GCC-PHAT) [2].

3.1 Korelacja wzajemna

$$g(k) = \sum_{n=0}^{N-1-k} x_L(n) \cdot x_R(n-k)$$

W celu określenia opóźnienia poszukuje się największego pików w sygnale wynikowych, którego pozycja określa opóźnienie:

$$\Delta t = \underset{k}{\operatorname{argmax}}[g(k)]$$

Zaimplementowano dwie metody znajdowania przesunięcia za pomocą korelacji wzajemnej. Pierwsza z nich to implementacja naiwna w języku Python. Druga implementacja wykorzystuje bibliotekę *scipy* w celu szybszego liczenia korelacji wzajemnej [2].

Kod źródłowy 2: Implementacja korelacji wzajemnej

Źródło: Opracowanie własne

```
| 1 def crosscorrelation_mine(y):
```

```

2     length = y.shape[0]
3     g = np.zeros(length)
4     for k in range(length):
5         for n in range(length - 1 - k):
6             g[k] += y[n, 1] * y[n - k, 0]
7     return g
8
9 g = crosscorrelation_mine(y_stereo_shifted)
10 recovered_shift = np.argmax(g)
11
12 def crosscorrelation_scipy(y1, y2, sr):
13     length = y1.size
14     xcorr = correlate(y1, y2)
15     dt = np.arange(1 - length, length)
16     shift = dt[xcorr.argmax()]
17     return shift

```

3.2 GCC-PHAT

Generalized Cross Correlation with Phase Transform (GCC-PHAT)

$$g(k) = \mathcal{F}^{-1} \left(\frac{X_L(f) \cdot X_R^*(f)}{|X_L(f) \cdot X_R^*(f)|} \right)$$

gdzie:

$\mathcal{F}^{-1}(\cdot)$ - odwrotna transformata Fouriera, X_L, X_R - transformaty Fouriera sygnałów x_L oraz x_R , X_R^* - wektor wartości sprzężonych do X_R [2].

Kod źródłowy 3: Implementacja GCC-PHAT

Źródło: Opracowanie własne

```

1 def gcc_phat(y1, y2):
2     n = y1.size + y2.size
3
4     x_l = np.fft.rfft(y1, n=n)
5     x_r = np.fft.rfft(y2, n=n)
6     x_conj = np.conj(x_r)
7
8     g = np.fft.irfft((x_l * x_conj) / (np.abs(x_l * x_conj)))
9     dt = np.arange(-y1.size, y1.size)
10    shift = dt[g.argmax()]
11    return shift

```

4 Kąt źródła dźwięku

Znając przesunięcie czasowe sygnałów można określić kąt, pod jakim dociera czoło fali płaskiej emitowanej przez źródło dźwięku:

$$\Delta t \cdot c = d \cdot \sin(\alpha)$$

gdzie: Δt - opóźnienie dźwięku docierające do uszu, c - prędkość dźwięku, d - odległość między uszami [2].

Kąt źródła dźwięku w płaszczyźnie poziomej określa zależność:

$$\alpha = \arcsin\left(\frac{\Delta t \cdot c}{d}\right)$$

Kod źródłowy 4: Wyznaczanie kąta padania

Źródło: Opracowanie własne

```
1 def find_angle(shift, speed=343, ear_distance=20):  
2     return np.arcsin((shift * speed) / ear_distance)
```

5 Międzyuszna różnica poziomów

Międzyuszna różnica poziomów - ILD (ang. Interaural level difference) - atrybut ten jest obliczany na podstawie średniej mocy sygnałów docierających do uszu.

$$ILD = 10 \cdot \log_{10} \left(\frac{\sum_{k=0}^{K-1} x_L^2(k)}{\sum_{k=0}^{K-1} x_R^2(k)} \right)$$

Czasami ILD jest obliczane w dziedzinie częstotliwości [2].

Kod źródłowy 5: Wyznaczanie ILD

Źródło: Opracowanie własne

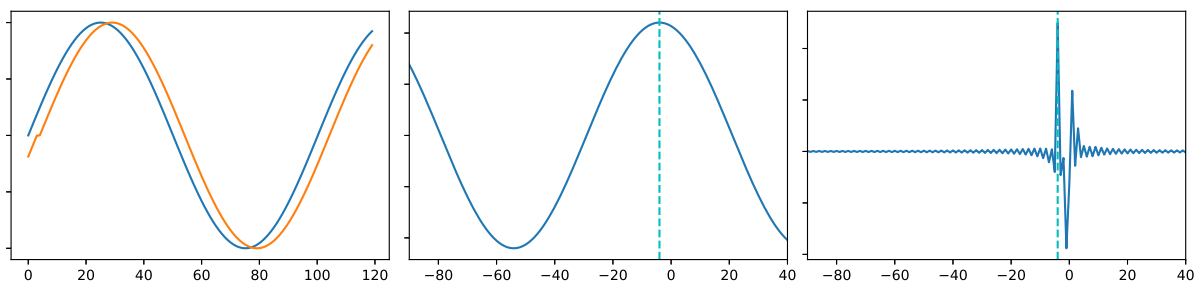
```
1 def ild(y1, y2):  
2     return 10 * np.log10(np.sum(y1 ** 2) / np.sum(y2 ** 2))
```

6 Badania wygenerowanych sygnałów

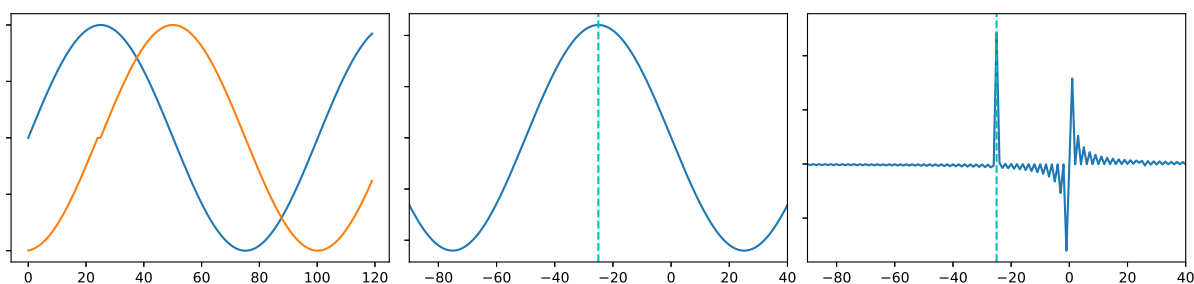
Zbadano wygenerowane sygnały z sekcji 2 za pomocą wcześniej zaimplementowanych metryk. Wyniki przedstawiono w tabeli.

Tabela 1: Pomiary dla wygenerowanych sygnałów

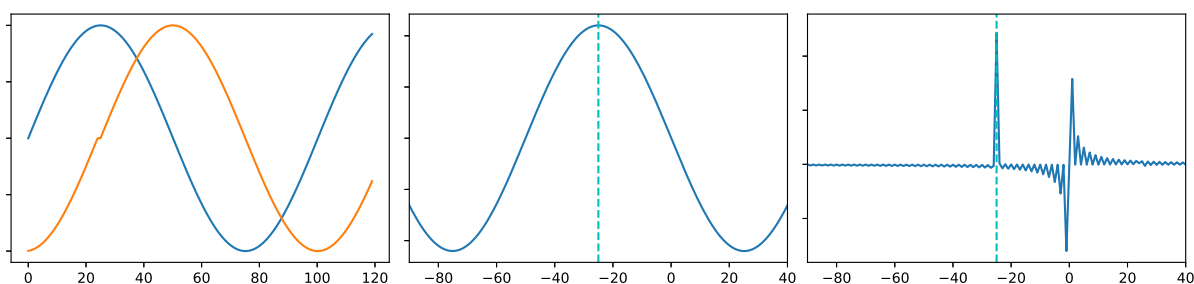
przesunięcie [próbki]	XCORR	GCC-PHAT	Kąt	ILD
4	-4	-4	0.098262	-4.821637e-16
25	-25	-25	0.565017	-4.821637e-16
46	-46	-46	1.031810	0.000000e+00
67	33	-67	1.498670	-4.821637e-16
88	12	-88	1.965631	0.000000e+00



Sygnał przesunięty o 4 próbki



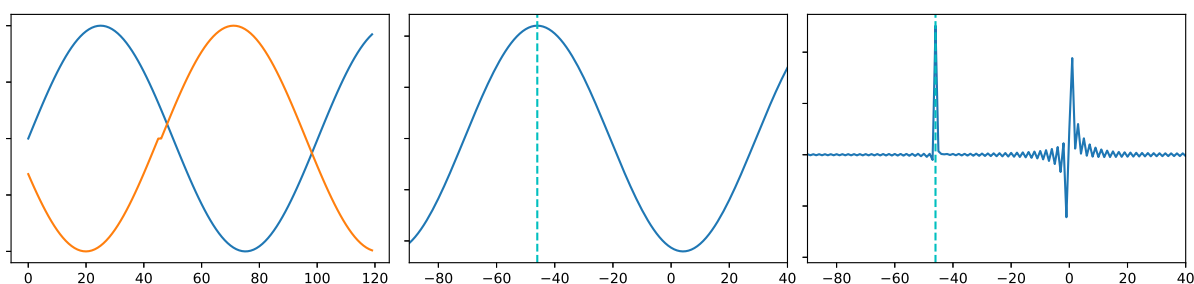
Sygnał przesunięty o 25 próbek



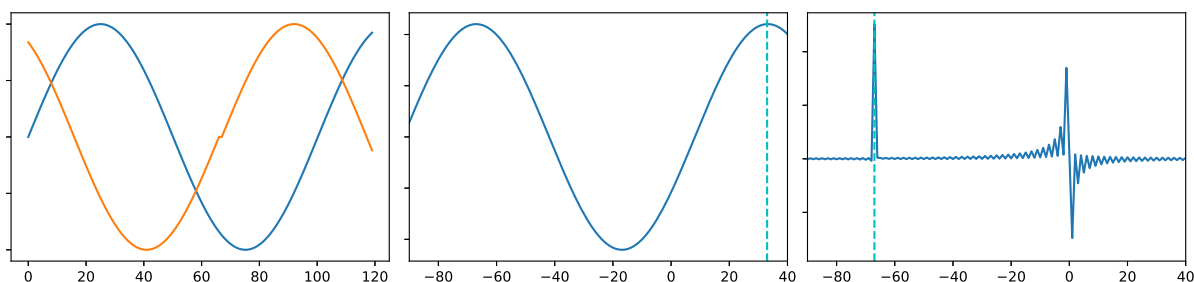
Sygnał przesunięty o 25 próbek

Rysunek 6: Porównanie wyznaczanych metryk dla różnych przesunięć sygnału. Wykresy są przycięte do interesujących przedziałów.

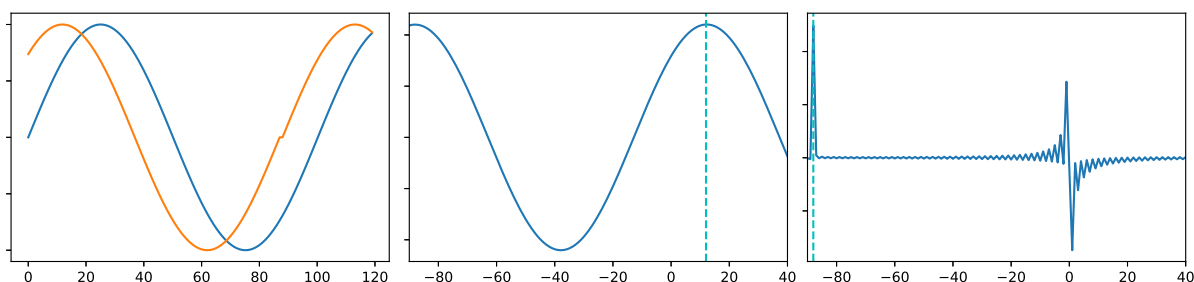
Źródło: Opracowanie własne



Sygnal przesunięty o 46 próbek



Sygnal przesunięty o 67 próbek



Sygnal przesunięty o 88 próbek

Rysunek 7: Porównanie wyznaczanych metryk dla różnych przesunięć sygnału. Dla przesunięcia o 67 i 88 próbek metoda korelacji nie daje poprawnych wyników.

Źródło: Opracowanie własne

6.1 Wnioski

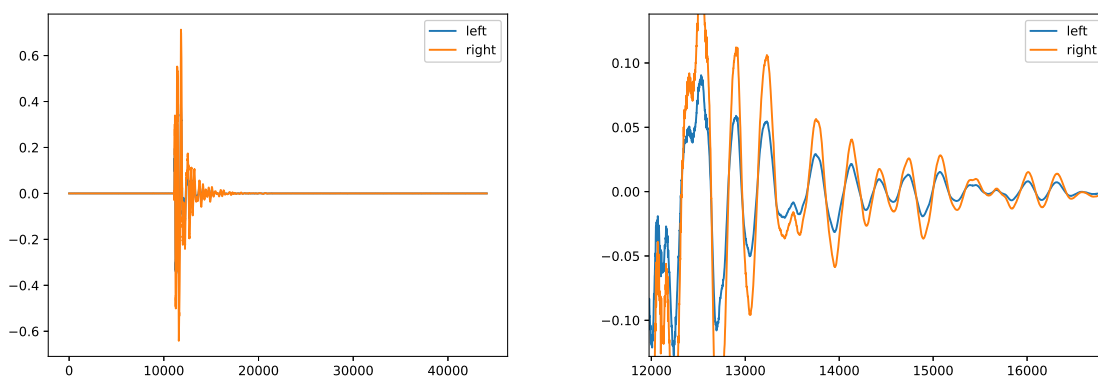
Dla wygenerowanego sygnału harmonicznego 440Hz przy przesunięciach od 0.0001 do 0.002 sekundy obie metody wyznaczania przesunięcia dają dokładne wyniki. Testując sygnały o wyższych częstotliwościach zauważono występujące błędy w metodzie korelacji wzajemnej przy dużych przesunięciach sygnału. Powód takich błędów wyjaśniono w sekcji 2.

7 Badanie dla indywidualnego pliku

W zadaniu należało też wyznaczyć zaimplementowane metryki dla dostarczonego sygnału.

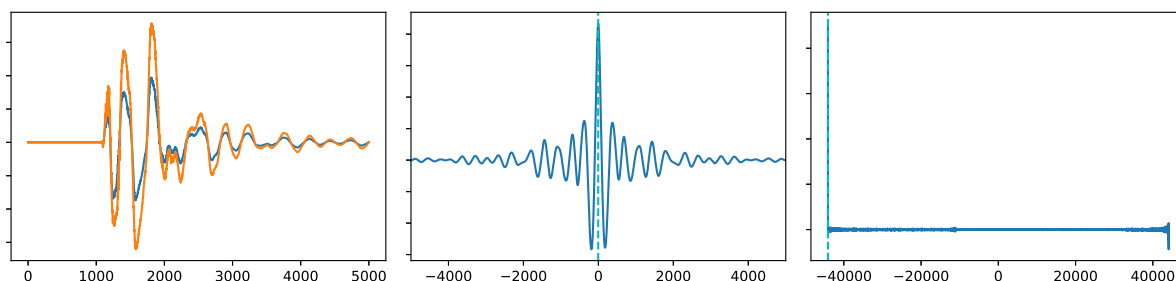
Tabela 2: Uzyskane wyniki dla sygnału *lab_02-03-DK.wav*

metryka	wynik
xcorr	-5 próbek = 0.00022 s
GCC-PHAT	-44100
kąt	-0.222
ILD	-5.1796



Rysunek 8: Sygnał *lab_02-03-DK.wav*

Źródło: Opracowanie własne



Rysunek 9: Analiza dźwięku uderzenia perkusji. Metoda GCC-PHAT nie dała poprawnych wyników.

Źródło: Opracowanie własne

Analizowany sygnał trwa 2 sekundy. Jego częstotliwość próbkowania to 22050 Hz. Na podstawie subiektywnej analizy wykresy, można stwierdzić, że kanał prawy jest przesunięty o około 5 próbek sygnału.

Uzyskane wyniki pokrywają się z obserwacjami wykresu. Za pomocą korelacji wzajemnej poprawnie wyznaczono przesunięcie sygnału. Metoda GCC-PHAT nie dała poprawnych rezultatów.

Bibliografia

- [1] Moore B. C.: *Wprowadzenie do psychologii słyszenia*, Wydawnictwo Naukowe PWN, 1999.
- [2] Mąka T.: *Lokalizacja i śledzenie źródeł dźwięku*, Sygnały akustyczne, 2021.