

Spis treści:

1. Wstęp
2. Cel i zakres pracy
3. Stosowane sposoby realizacji pogłosu
 - 3.1. Główne idee
 - 3.2. Opis stosowanych sposobów implementacji poszczególnych elementów pogłosu
4. Zjawiska psychoakustyczne przy percypowaniu pogłosu
5. Wybrane sposoby porównywania sygnałów pod kątem zawartości widmowej i czasowej sygnału pogłosowego
6. Testy stosowanych rozwiązań implementacji pogłosu
 - 6.1. Wzmocnienia linii opóźniających i poszczególnych odbić
 - 6.2. Filtracja linii opóźniających
 - 6.3. Wykorzystanie zjawisk psychoakustycznych
 - 6.4. Synteza późnego ogona pogłosowego
7. Implementacja docelowego pogłosu (~~przy implementowaniu pogłosu nastąpią modyfikacje metod opisanych w rozdziale 5. W tym rozdziale, przy opisie implementacji, zostaną opisane także te modyfikacje~~)
 - 7.1. Wprowadzenie (~~w tym dobór narzędzia~~)
 - 7.3. Architektura aplikacji i opis klas stworzonych w języku C++
 - 7.5. Połączenie wybranych rozwiązań (i ich modyfikacje) w celu uzyskania docelowego pogłosu
 - 7.5.1. Linie opóźniające i sposób obliczania aktualnego bufora audio
 - 7.5.2. Implementacja pierwszych odbić
 - 7.5.3. Tłumienie odbić
 - 7.5.4. Filtracja linii opóźniających
 - 7.5.5. Zastosowanie międzyuszných zjawisk psychoakustycznych
 - 7.5.6. Synteza późnego ogona pogłosowego
7. Analiza działania wynikowego pogłosu
8. Podsumowanie i wnioski
9. Bibliografia

1. Wstęp

W odróżnieniu do badań przeprowadzanych pod koniec XX wieku, które wymagały skomplikowanego toru audio oraz zwykle potrzeby wykorzystania pola swobodnego oraz dyfuzyjnego do wykonywania testów, teraźniejsze możliwości znacznie to ułatwiają. Przede wszystkim możliwość odwzorowania większości zjawisk w domenie cyfrowej, jak i znacznie ułatwiające badania nieporównywalnie większe możliwości obliczeniowe. W literaturze z tego okresu można zaobserwować obszerne opisy dotyczące przygotowania warunków, które było niezbędne to wykonania badań. W odróżnieniu do tej sytuacji, w przypadku potrzeby dokonania niewielkiej zmiany w parametrach, jedyne niezbędne czynności do wykonania to niewielka modyfikacja kodu oraz odczekanie niewielkiej ilości czasu w celu otrzymania nowo skompilowanej aplikacji. O wiele mniejsze możliwości obliczeniowe w tamtych czasach skutkowały także bardzo częstym i popularnym podejściem subiektywnym do problemu klasyfikacji i oceny zjawisk akustycznych. W dzisiejszych czasach wciąż istnieje bardzo duża liczba zjawisk trudnych do obiektywnego opisu, jednak szybkość obliczeń pozwala na większą ilość testów w tym samym czasie, a co za tym idzie łatwość opisu - teraźniejsze warunki zdecydowanie to ułatwiają.

W pionierskiej pracy Schroedera z 1962 roku [4] zostają poruszone fundamentalne dziś zasady, na który opiera się większość systemów i architektur implementacyjnych sztucznych pogłosów. Problemy rozwiązywane wtedy wymagały o wiele więcej pracy (dostosowanie czasu

2. Cel i zakres pracy

Celem pracy jest projekt i implementacja pogłosu w formacie VST. Celem jest również próba obiektywizacji parametrów pojawiających się w literaturze za pomocą opisu subiektywnego. Praca obejmie przegląd głównych idei realizacji sztucznego pogłosu,

3. Stosowane sposoby realizacji pogłosu

Od wczesnych lat rozwoju prac nad implementacją sztucznego pogłosu, najbardziej popularne jest rozgraniczenie go na dwie części: pierwsze odbicia (pierwsze z nich to około 80 ms od dźwięku bezpośredniego, zależy to jednak od wielkości pomieszczenia) oraz późny pogłos (pozostała część pogłosu) [1], [2]. Pierwsze odbicia to część pogłosu złożona z dyskretnych odbić imitujących wczesne odbicia od ścian lub innych elementów w pomieszczeniu. O rozpoznaniu kierunku, z którego dochodzi dźwięk decyduje pierwsza fala dźwiękowa. Odbicia dźwięku przychodzące w ciągu pierwszych 30 ms od pierwszej fali wpływają słabo na odczucie kierunku [11]. Czas opóźnienia w stosunku do dźwięku bezpośredniego oraz stosunki amplitud wczesnych odbić zależą w głównej mierze od kształtu pomieszczenia i pozycji źródła względem odbiornika. Odbicia te odgrywają główną rolę w subiektywnym odczuciu przestrzenności dźwięku, człowiek kojarzy percepcyjnie właściwości sygnału z warunki akustycznymi, jakie panowałyby w rzeczywistych warunkach. Pomagają

więc w lokalizacji źródła dźwięku przez odbiorcę. Późny pogłos, w odróżnieniu od pierwszych odbić jest niezależny od położenia źródła dźwięku i odbiornika - jest silnie zależny od geometrii pomieszczenia oraz właściwości pochłaniających i rozpraszających materiałów elementów w nim się znajdujących. W celu zaprojektowania algorytmu sztucznego pogłosu Jot [1] proponuje procedurę przyjęcia odpowiednich parametrów, aby uzyskać końcowy efekt w postaci sygnału 'mono', co w praktyce odpowiada sygnałowi o identycznej zawartości dla wszystkich kanałów. Następnie, należy rozważyć zjawiska binauralne oraz inne rozwiązania w celu nadania pogłosowi realizmu.

Pomimo obecności w literaturze w większości opisu subiektywnego, W celu ograniczenia wyboru metod na podstawie subiektywnych wrażeń słuchowych, przeprowadzono analizę właściwości sygnałów za pomocą zaproponowanych obiektywnych parametrów.

Na podstawie badań wykazano, że pogłos oparty na sieci linii opóźniających nie jest subiektywnie rozróżnialny z pogłosem z rzeczywistych pomieszczeń. Badania wskazały, iż zjawisko zachodzi dla nie tylko sygnałów takich jak mowa czy muzyka, ale także dla bardzo krótkich odcinków szerokopasmowego szumu Gaussa. [?]

3.1. Główne idee

3.2. Opis stosowanych sposobów implementacji poszczególnych elementów pogłosu

Wzmocnienia linii opóźniających i poszczególnych odbić

Barron [2] zwrócił uwagę, że czas pogłosu nie jest jedynym wyznacznikiem jakości dźwięku w pomieszczeniu. Przeprowadził badania mające na celu zrozumienie jak ważny przy modelowaniu sztucznego pogłosu jest wpływ pierwszych odbić. Badania te polegały na analizie subiektywnych wrażeń osób badanych, którym przedstawiane były próbki dźwiękowe zawierające pogłos o różnych parametrach. Został zbadany między innymi wpływ pierwszych odbić bocznych na subiektywne zjawisko 'wrażenia przestrzenności' pogłosu. To zaobserwowane zjawisko miało wynikać z właściwości wczesnych odbić bocznych w pomieszczeniu (10 – 80ms). Na podstawie jego badań został także wysunięty wniosek, iż wrażenie przestrzenności jest tym większe, im stosunek docierających do odbiorcy bocznych do pozostałych. Ze względu na najwyższy poziom dźwięku, odbicia te są najbardziej znaczące w percypowaniu pogłosu. Zjawisko 'wrażenia przestrzenności' odczuwane przez osoby badane pojawiało się dla symulacji odbić bocznych w zakresie >10ms. Powodowało to wrażenie 'poszerzenia się' źródła dźwięku. Przy zwiększeniu poziomu dźwięku odbicia bocznego wrażenie przestrzenności zwiększało się. W badaniach został przebadany także wpływ odbicia od sufitu na wrażenie przestrzenności - wpływ odbicia of sufitu ma negatywny (niewielki) wpływ na 'wrażenie przestrzenności'.

Według Jota [1], jeśli poszczególne linie opóźniające mają niejednakowy czas zaniku, skutkuje to wyraźnie słyszalnymi składowymi i ujawnia obecność tych linii. W celu upewnienia

się, że sytuacja taka nie będzie miała miejsca, należy upewnić się, że wszystkie linie mają jednakowy czas zaniknięcia dźwięku. Również [3] zwraca uwagę, iż jedną z fundamentalnych zasad przy tworzeniu sztucznego pogłosu jest zadbanie o to, aby amplituda wszystkich składowych zanikała jednakowo prawie jednakowo szybko. Ma to na celu zapewnienie jednakowego opadania różnych składowych częstotliwościowych zanikającego dźwięku.

Czasy opóźnienia linii opóźniających

Dobór czasów opóźnienia linii opóźniających nie jest kluczowy, pod warunkiem, że liczby je reprezentujące nie są przez siebie podzielne [1]. Według [5] dobór nieskorelowanych wartości jest niezbędny, aby uniknąć zjawiska 'flutter echa' i uzyskać płaską charakterystykę częstotliwościową. Aby uniknąć zmniejszenia gęstości odbić oraz superpozycji, pożądane jest użycie liczb nieproporcjonalnych do siebie [4].

Powyższe fakty prowadzą do wniosku, iż w celu uniknięcia negatywnych, wymienionych wyżej zjawisk, niezbędne jest ograniczenie zbioru liczb reprezentujących czasy opóźnienia do liczb wzajemnie pierwszych.

Rys .. przedstawia obwiednię amplitudy obu przypadków. Rys .. przedstawia widmo obu przypadków.

Według [5] krótkie czasy opóźnienia skutkują większym ubarwieniem dźwięku, jakie ma miejsce w małych pomieszczeniach. Zjawisko takie byłoby nienaturalne w pomieszczeniach o wysokim czasie pogłosu. Według autorów zawartość ubarwień w złożonej sieci nie jest łatwa do oceny, kontrola zawartości tych ubarwień wymaga używania metody prób i błędów. Prawdopodobnie zjawisko to występuje w mniejszym stopniu dla sieci zawierających choć niewielką ilość długich linii opóźniających, jednak wtedy trudniej kontrolować gęstość pogłosu. Wy tłumaczenie tego zjawiska proponuje Barron [2] jako interferencję między dźwiękiem bezpośrednim, a opóźnionym o niewielkie wartości, co skutkuje pojawieniem się filtru grzebieniowego. Wskazuje na fakt, iż zjawisko ma miejsce dla opóźnień około 10-50ms, a szczególnie w okolicach 20ms.

Prawdopodobnie zjawisko nazwane 'zakolorowaniem' przez autora ma swoje odzwierciedlenie w widmie sygnału. W celu potencjalnej formalizacji tego terminu na rys .. przedstawia się widmo sygnałów z niskimi oraz wysokimi czasami opóźnienia.

Schroeder [4] - czasy opóźnienia w stosunku 1:1.5. Rys .. Porównanie odpowiedzi impulsowej z czasami opóźnienia w tym zakresie i w o wiele szerszym zakresie.

Jednym z głównych celów testowania różnych sposobów filtracji jest ocena gęstości widma późnego pogłosu. Według Jota [1] istnieją następujące subiektywne sposoby na identyfikację niewystarczającej gęstości widmowej późnego pogłosu:

- Odpowiedź na sygnał impulsowy sygnał zawierać będzie "dzwonienie" poszczególnych składowych

- Odpowiedź na quasi-stacjonarny sygnał będzie zawierała nadmierny poziom niektórych częstotliwości (na przykład na niektórych nutach przy grze na instrumencie).

Filtracja linii opóźniających

Schroeder zaproponował powiązanie wartości tłumienia linii opóźniających z ich zawartością częstotliwościową. Miało to na celu odwzorowanie faktu dłuższego czasu pogłosu dla niskich częstotliwości. Dodatkowo, liczba powtórzeń dźwięku przy powstaniu filtru grzebieniowego powstałego przy zsumowaniu dźwięku bezpośredniego z jego kopią przesuniętą o 40ms to 25 [4]. Sytuacja ta powoduje znaczne zagęszczenie ilości odbić w jednostce czasu.

[5] zaproponowali, aby stosować filtr dolnoprzepustowy na wyjściu każdej linii opóźniającej w celu imitacji tłumienia dźwięku przez powietrze. W ich podejściu częstotliwości odcięcia filtrów zależą od czasu opóźnienia konkretnej linii.

Aby pogłos brzmiał naturalnie, pożądane jest, aby charakterystyka częstotliwościowa była płaska, lecz nie zbyt płaska [8].

Późny ogon pogłosowy

4. Zjawiska psychoakustyczne przy percypowaniu pogłosu

*Dodam podrozdział o ocenie subiektywnej i obiektywnej (artykuł w folderze z literaturą).
Albo wpleść te tematy w któryś rozdział.*

Teoria dupleksowa

Mechanizmy słyszenia przestrzennego pozwalają na odczucie kierunku, z którego dobiega dźwięk, a także w pewnej mierze odległości od źródła dźwięku [11]. ~~W naturalnych warunkach (pole rozproszone)~~

Teoria wygłoszona na podstawie badań audiometrycznych głosi, iż w sytuacji, w której do uszu słuchacza dochodzi ten sam sygnał, słuchacz lokalizuje sygnał jako będący z przodu. W przypadku słyszenia sygnału przez słuchawki słuchacz odczuwa dźwięk w środku głowy [11].

Dwa najważniejsze zjawiska opisujące słyszenie przestrzenne oraz możliwość lokalizacji źródła dźwięku przez człowieka to [11]:

- ITD. (*Interaural Time Difference*) oraz ILD (*Interaural Level Difference*). Zjawisko ILD polega na różnicy w natężeniach dźwięku docierającego do obu uszu słuchacza w momencie, gdy źródło dźwięku nie znajduje się dokładnie naprzeciw słuchacza. Zjawisko jest związane z faktem istnienia cienia akustycznego generowanego przez głowę odbiorcy (cień akustyczny zaczyna występować dla częstotliwości, przy których połowa długości fali staje się mniejsza od rozmiarów głowy). Dzięki temu, na

podstawie przeszłych doświadczeń słuchacz jest w stanie z dużą rozdzielczością stwierdzić, z jakiego kierunku w płaszczyźnie horyzontalnej dochodzi dźwięk.

- ITD to międzyuszną różnicę fazy biorącą się z różnicy czasów dotarcia sygnału do obu uszu. Na tym zjawisku oparta jest kierunkowość słyszenia niższych częstotliwości.

Zjawiska te i fakt działania ich obu dla różnych zakresów częstotliwości opisuje teoria *Duplex Theory* i została zaproponowana przez Reyleigha w 1907 roku [17].

W warunkach rzeczywistych, odbiorca dźwięku ma także do dyspozycji możliwość ruchów głową, dzięki którym w przypadku nieoczywistego kierunku dochodzenia dźwięku pozwalają na wykrycie różnic fazowych i poprawną lokalizację. W przypadku odbierania sygnału sztucznego zjawisko występuje jedynie słuchając sygnału korzystając z głośników (nie występuje korzystając ze słuchawek) [11].

Większość badań naukowych dotyczących pogłosu, zarówno teraźniejszych jak i przeszłych, jest badaniami subiektywnymi – są przeprowadzane na grupie ludzi, którzy zostają poddani testom słuchowym. Ich zadaniem jest podanie badającemu informacji np. o tym, w którym momencie dźwięk zaczął być słyszalny, lub czy słyszą różnicę między dwoma sygnałami. W niniejszej pracy zostaje podjęta próba zniwelowania czynnika subiektywnego w celu powtarzalności wyników badań oraz powołania się fakty, a nie opinie.

/////

Według Beranek [6] najważniejszym wyznacznikiem jakości pogłosu jest czas pierwszego odbicia w stosunku do dźwięku bezpośredniego i jest to nawet ważniejsze, czas pogłosu w pomieszczeniu, a optymalna wartość tego parametru to czas poniżej 20ms.

Zostało także stwierdzone [7 – Keet], że stopień wrażenia przestrzenności jest związany ze stopniem niekoherencji sygnałów z obu kanałów.

Efekt związany z lokalizacją dźwięku przez czas pierwszych odbić nie występuje, jeżeli odstęp czasowy między pierwszą falą, a pierwszym odbiciem jest krótszy niż 1 ms, w takim przypadku źródło dźwięku jest lokalizowane w położeniu pomiędzy kierunkami obu tych dźwięków.

W zależności od rodzaju sygnału mogą zachodzić różne zjawiska percepcyjne takie jak: dwuuszne odmaskowanie, dwuuszne dudnienia, efekt pierwszeństwa i inne [11]. *Tutaj mogę opisać te zjawiska i napisać dla jakich sygnałów działają. I że zjawiska mogą chodzić w różnym stopniu w zależności od tego, który sygnał badany sygnał najbardziej przypomina.*

///// *Jeżeli ktoś twierdzi że słychać dzwonięcie, to faktycznie - większość ludzi może słyszeć dzwonięcie. Jednak nie będzie to 100% ludzi. Natomiast jeżeli powołujemy się na liczby, to wszyscy muszą się z tym zgodzić. Poza tym można znaleźć nieskończoną liczbę sygnałów, które można opisać tym określeniem.*

5. Wybrane sposoby porównywania sygnałów pod kątem zawartości widmowej i czasowej sygnału pogłosowego

Opisane w podrozdziale 3.2 sposoby implementacji poszczególnych elementów pogłosu zawierają wskazówki dotyczące pożądanej zawartości widmowej i czasowej pogłosu oraz cech sygnału wpływających na wrażenie przestrzenności. W celu późniejszej oceny zastosowanych metod przedstawia się parametry statystyczne mogące być */adekwatne do ich oceny./* w tym pomocne.

A). Widmowa gęstość mocy - określa, jak moc sygnału jest rozprowadzona w dziedzinie częstotliwości. W przypadku sygnałów cyfrowych moc definiuje się abstrakcyjnie jako kwadrat wartości sygnału. Ogólnie, średnia wartość mocy zdefiniowana jako [19]:

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |x(t)|^2 dt.$$

Jeśli sygnał może być potraktowany jako sygnał stacjonarny, to widmowa gęstość mocy jest transformatą Fouriera

$$\hat{x}(\omega) = \frac{1}{\sqrt{T}} \int_0^T x(t) e^{-i\omega t} dt.$$

funkcji autokorelacji tego sygnału, co w rezultacie daje [20]:

$$S(f) = \int_{-\infty}^{\infty} R(\tau) e^{-2\pi i f \tau} d\tau = \mathcal{F}(R(\tau))$$

Moc sygnału dla danego pasma częstotliwości można wyliczyć wykonując całkowanie widmowej gęstości mocy po dodatnich i ujemnych częstotliwościach [20]:

$$P = \int_{F_1}^{F_2} S(f) df + \int_{-F_2}^{-F_1} S(f) df$$

Definicja widmowej gęstości mocy może być przełożona do dziedziny dyskretnej, wtedy wzór przyjmuje formę [19]:

$$\tilde{S}_{xx}(\omega) = \frac{(\Delta t)^2}{T} \left| \sum_{n=1}^N x_n e^{-i\omega n \Delta t} \right|^2$$

B). Płaskość widmowa – parametr określający, jak bardzo sygnał zbliżony jest do sygnału szumowego, w opozycji do sygnału tonalnego. Wysoka wartość płaskości widmowej (bliska 1.0 dla szumu białego) oznacza, że widmo ma podobną moc we wszystkich pasmach częstotliwości. Płaskość widmowa jest zdefiniowana jako stosunek średniej geometrycznej widma mocy sygnału do średniej arytmetycznej widma mocy sygnału [21]:

$$\text{Flatness} = \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\frac{\sum_{n=0}^{N-1} x(n)}{N}} = \frac{\exp\left(\frac{1}{N} \sum_{n=0}^{N-1} \ln x(n)\right)}{\frac{1}{N} \sum_{n=0}^{N-1} x(n)}$$

Wielkość ta często jest przedstawiana w decybelach, gdzie wartość szczytową parametru przyjmuje się jako 0 dB [21].

C). Koherencja sygnałów:

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)},$$

gdzie $G_{xy}(f)$ jest gęstością spektralną krzyżową [22] między sygnałami x i y , a $G_{xx}(f)$ i $G_{yy}(f)$ są gęstościami spektralnymi sygnałów samych z sobą. Wielkość ta opisuje relację pomiędzy dwoma sygnałami ... [24].

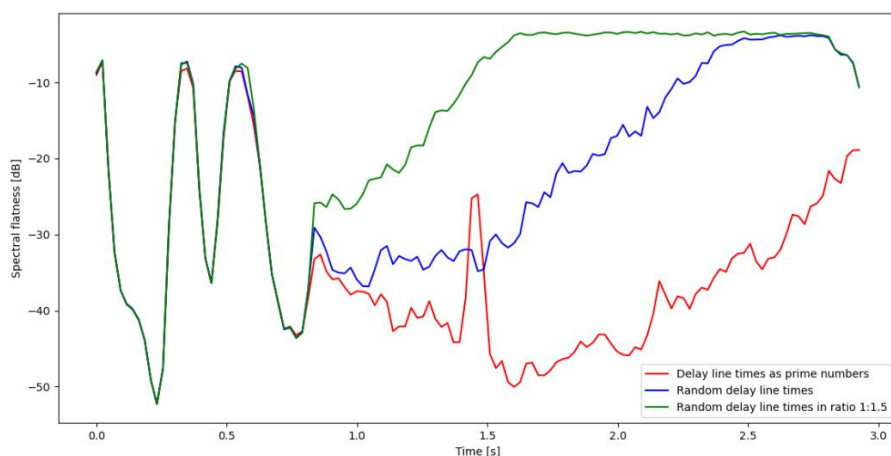
6. Parametryzacja sygnałów pogłosowych oraz testy stosowanych rozwiązań implementacji pogłosu

Obiektywna ocena przydatności poszczególnych metod wymaga odpowiedniego doboru parametrów statystycznych i ich zastosowania w celu uzyskania wartości liczbowych będących podstawą do uzasadnienia oceny danej metody na tle innej.

Czasy opóźnienia:

W celu zbadania wpływu doboru czasów opóźnienia na brzmienie pogłosu, dokonano symulacji zjawisk następujących zjawisk.

- Czasy opóźnienia jako liczby losowe,
- Czasy opóźnienia jako liczby losowe wzajemnie pierwsze,
- Czasy opóźnienia zawierające się w zakresie 1:1.5 [4],



Rys (..) Płaskość widmowa sygnału mowy dla różnego rodzaju generowania linii opóźniających.

W celu obiektywnego porównania sygnałów pogłosowych przedstawia się parametr będący odzwierciedleniem następujących założeń - sygnał najbardziej przydatny to ten, którego:

- Płaskość widma największą w największej ilości punktów w stosunku do pozostałych
- Widmowa gęstość mocy jest największa

Powyższe założenia można przedstawić jako zależność:

$$//[(G_{xy} + G_{xz})/2 * E(S_{xx}(x(n)))] / [\sigma(Flatness(x(n))) * \sigma(S_{xx}(x(n)))]$$

$$B := \frac{\frac{(G_{xy} + G_{xz})}{2} E(S_{xx}(x(n)))}{\sigma(Flatness(x(n))) * \sigma(S_{xx}(x(n)))}$$

Analogicznie, dla porównania sygnału y(n):

$$//[(G_{xy} + G_{yz})/2 * E(S_{yy}(y(n)))] / [\sigma(Flatness(y(n))) * \sigma(S_{yy}(y(n)))]$$

$$B := \frac{\frac{(G_{xy} + G_{yz})}{2} E(S_{yy}(y(n)))}{\sigma(Flatness(y(n))) * \sigma(S_{yy}(y(n)))}$$

Oraz z(n):

$$//[(G_{xz} + G_{yz})/2 * E(S_{zz}(z(n))) / [\sigma(Flatness(z(n))) * \sigma(S_{zz}(z(n)))]$$

$$B := \frac{\frac{(G_{yz} + G_{xz})}{2} E(S_{zz}(z(n)))}{\sigma(Flatness(z(n))) * \sigma(S_{zz}(z(n)))}$$

gdzie:

//B - zdefiniowany współczynnik służący do porównania sygnałów

G_{xy} - współczynnik wyrażony liczbą z zakresu [0,1] będący stosunkiem liczby punktów, w których płaskość widma pierwszego sygnału $x(n)$ jest większa, niż płaskość widma sygnału $y(n)$ do liczby wszystkich punktów wektora płaskości widma.

$x(n)$ - sygnał pogłosowy zawierający linie opóźniające o czasach opóźnienia będącymi liczbami wzajemnie pierwszymi

$y(n)$ - sygnał pogłosowy zawierający linie opóźniające o czasach opóźnienia będącymi liczbami losowymi

$z(n)$ - sygnał pogłosowy zawierający linie opóźniające o czasach opóźnienia będącymi liczbami losowymi z zakresu liczb o stosunku 1:1.5

σ - średnie odchylenie standardowe

E - średnia arytmetyczna

(reszta symboli opisana w rozdziale 5)

	$x(n)$	$y(n)$	$z(n)$
B			

Filtracja

Do testów związanych z filtracją zastosowano wersję pogłosu o czasach opóźnienia linii opóźniających będącymi liczbami wzajemnie pierwszymi

W tym celu jako jeden ze sposobów implementacji filtracji zastosowano następującą zależność: im wyższy czas opóźnienia linii opóźniającej, tym wyższa częstotliwości odcięcia filtru dolnoprzepustowego.

W celu zbadania rodzaju filtracji na właściwości sygnału zaimplementowano następujące rozwiązania:

- Filtracja dolnoprzepustowa każdej linii opóźniającej w jednakowy sposób- ~~symulacja tłumienia przez powietrze~~
- Filtracja dolnoprzepustowa o częstotliwości odcięcia tym wyższej, im wyższa wartość opóźnienia linii opóźniającej (a nie niższej? - bo na końcu zwykle zostają niskie f)

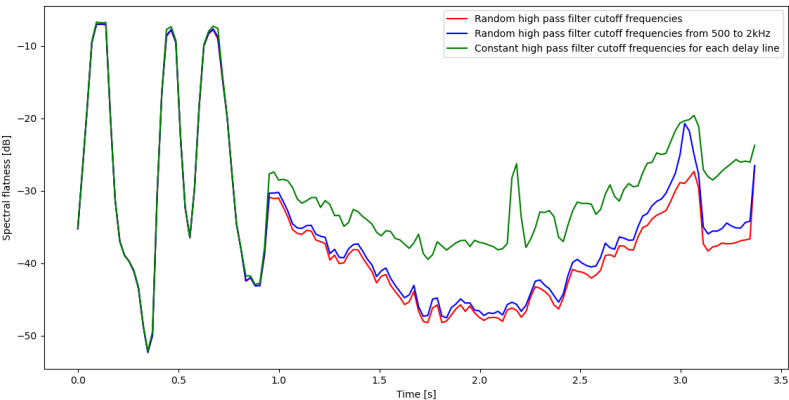
- Filtracja w zakresie pochłaniania materiałów (Większość materiałów, z których składają się ściany mają największy stopień pochłaniania energii akustycznej w zakresie 500 Hz – 2 kHz [5]).

Na trudność obiektywnej oceny jakości pogłosu wskazują [5]. Autorzy zauważyli, że dobór parametrów filtrów stosowanych na pierwszych odbiciach mocno wpływa na jakość pogłosu, natomiast trudno znaleźć ilościowy parametr pozwalający na obiektywną, liczbową ich ocenę.

// "Perhaps a statistical method of choosing the early reflection properties, combined with some perceptually meaningful constraints, can be found." - [5].

Ze względu na fakt niezalezienia obiektywnego parametru przez w literaturze, w celu oceny wpływu filtracji na jakość pogłosu dokonano próby obiektywnej oceny tego zjawiska. Wysłunięto także przypuszczenie, iż stwierdzenie [5] o trudności obiektywnej oceny dotyczy także pogłosu jako całość, a nie odnosi się to wyłącznie do pierwszy odbić. Wykorzystując więc brak aktualnego obiektywnego parametru oraz wskazówkę w postaci cytatu, zaproponowano porównanie właściwości *całego* pogłosu parametrem ... Rys .. przedstawia porównanie ..

~~Zgodnie z tym stwierdzeniem...~~



Rys (..) Płaskość widma symulacji różnych wariantów filtracji

Porównanie sygnałów pogłosowych z różnymi wariantami filtracji

//pogłosz z różnymi filtracjami również porównywane również współczynnikiem B

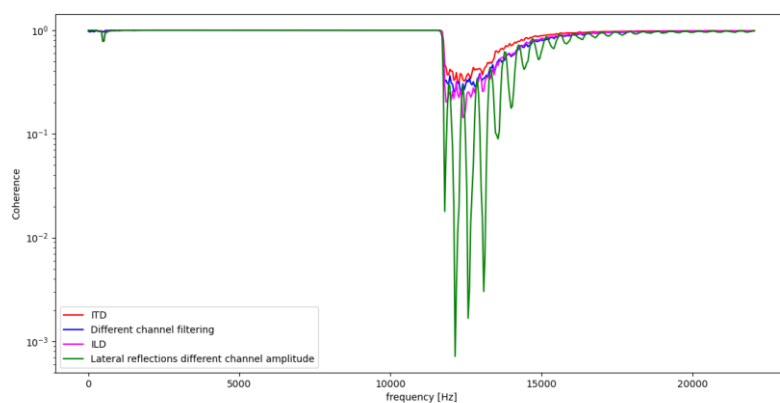
	$x_f(n)$	$y_f(n)$	$z_f(n)$
--	----------	----------	----------

B			
---	--	--	--

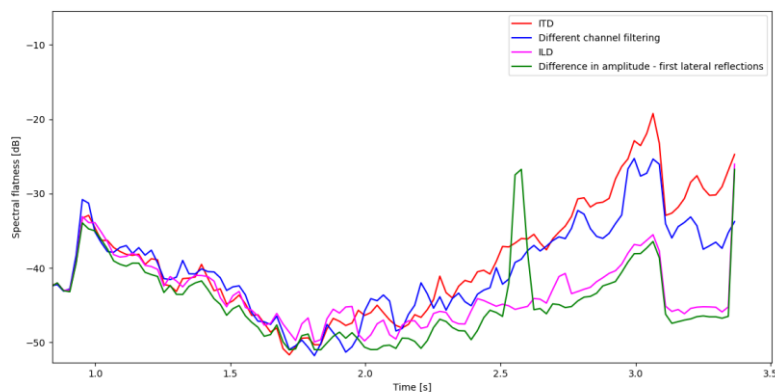
Wykorzystanie zjawisk psychoakustycznych/przestrzenność pogłosu

W celu zbadania wpływu zjawisk psychoakustycznych na właściwości sygnału zaimplementowano następujące rozwiązania:

- Różnica w poziomie pierwszych odbić bocznych w celu potwierdzenia zaobserwowanej w [2] tezy o większej przestrzenności przy wyższym poziomie tych odbić,
- Zastosowanie różnicy czasu opóźnienia między kanałami dla linii opóźniających,
- Zastosowanie różnicy amplitudy pomiędzy kanałami dla linii opóźniających,
- Ze względu na fakt, iż jednym ze sposobów zwiększenia przestrzenności dźwięku jest implementacja niewielkich różnic pomiędzy parametrami pogłosu [7], zdecydowano się na implementację niewielkiej różnicy pomiędzy filtracją (różnica w częstotliwości odcięcia filtru dolnoprzepustowego w granicach 0-5%) odbić dla obu kanałów.



Rys (..) Koherencja między kanałami dla różnych zjawisk przestrzennych



Rys (..) Płaskość widmowa dla różnych zjawisk przestrzennych

Jak opisano w podrozdziale 3.2, stopień wrażenia przestrzenności związany jest ze stopniem niekoherencji sygnałów z obu kanałów. Dlatego też pożądane jest, aby koherencja sygnałów z obu kanałów była jak najmniejsza. Równie pożądana jest jednak wysoka widmowa gęstość mocy. Przyjęto założenie, że im bardziej wartości koherencji skupione wokół średniej tym lepiej. Podobnie, jak odchylenie od wartości średniej widmowej gęstości mocy, która pożądana jest na stałym poziomie. Założenia te można opisać poniższym parametrem:

$$//\text{przestrzenność} = (E(S_{xx}(x(n))) / (C_{xy}(x(n)) * \sigma(S_{xx}(x(n))) * \sigma(C_{xy}(x(n))))$$

$$\text{przestrzenność} = \frac{E(S_{xx}(x(n)))}{C_{xy}(x(n)) * \sigma(S_{xx}(x(n))) * \sigma(C_{xy}(x(n)))}$$

gdzie:

$$\sigma = \sqrt{E((X - E(X))^2)}$$

$$\bar{S}_{xx}(\omega) = \frac{(\Delta t)^2}{T} \left| \sum_{n=1}^N x_n e^{-i\omega n \Delta t} \right|^2$$

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)},$$

E(x) - średnia arytmetyczna

(wzory opisane w rozdziale 5)

Parametr ten będzie wyznacznikiem przestrzenności pogłosu.

Tab . Wartości parametrów statystycznych

	ITD	Różna filtracja obu kanałów	ILD	Różnica w amplitudzie pierwszych odbić bocznych
Wartość średnia gęstości widma				
Średnie odchylenie standardowe gęstości widma				
Wartość średnia koherencji kanałów				
Średnie odchylenie standardowe koherencji kanałów				

Korzystając ze wzoru .. otrzymujemy wartości parametru przestrzenność:

Tab . Wartości parametru przestrzenność dla symulacji różnych zjawisk przestrzennych

	ITD	Różna filtracja obu kanałów	ILD	Różnica w amplitudzie pierwszych odbić bocznych
przestrzenność (0, 1)				

Dodatkowo, przy zastosowaniu różnicy czasu opóźnienia między kanałami, zyskujemy większą gęstość odbić.

Ze względu na fakt powstawania filtra grzebieniowego przy opóźnieniu do 40ms, zastosowano właśnie taką wartość jako maksymalną różnicę czasu opóźnienia między

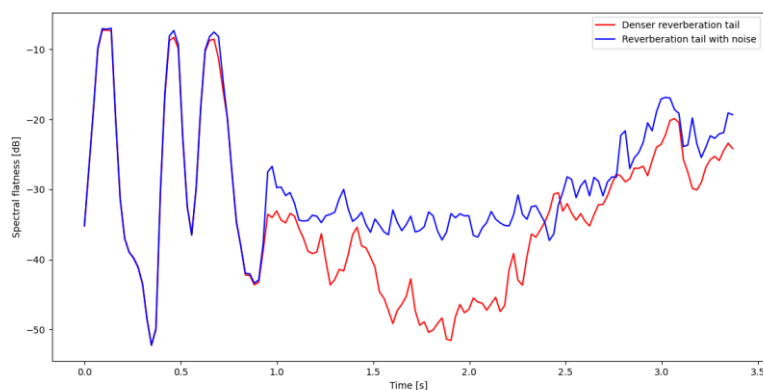
kanalami. Różnice te są różne dla każdego odbicia – skutkuje to uzyskaniem największej gęstości widmowej, dostajemy bowiem różny filtr grzebieniowy dla każdej linii opóźniającej.

W celu kompensacji wrażenia znajdowania się źródła dźwięku po bardziej po lewej lub prawej stronie, wartości różnicy parametrów (dotyczy to pierwszego odbicia): czasów opóźnienia między kanałami, poziomu oraz filtracji dostosowano tak, aby odnosiło się wrażenie umiejscowienia źródła naprzeciwko słuchacza.

Późny ogon pogłosowy

//“The remainder, called the *late reverberation*, is so densely populated with echoes that it is best to characterize the response *statistically* in some way” - [8]

- Symulacja późnego pogłosu jako sygnał będący wynikiem mnożenia późnego ogona pogłosowego złożonego z linii opóźniających ze sprzężeniem - z filtrowanym szumem białym - cross fade z opadającym ogonem pogłosowym
- Zagęszczenie odbić w późnej fazie pogłosu przez dodanie tam większej liczby linii **to**



Rys (..) Płaskość widma dla obu wersji implementacji późnego pogłosu

//Mamy dwie płaskości widma opisane wektorami – sprawdzam w ilu punktach płaskość pierwsza jest większa niż płaskość druga. Przedstawiam w procentach (0,1). Mnożę przez średnią gęstość widma i dzielę przez średnie odchylenie standardowe gęstości widma.

$$C := \frac{G_{xy} * E(S_{xx}(x(n)))}{\sigma(S_{xx}(x(n)))}$$

$$C := \frac{G_{xy} * E(S_{yy}(y(n)))}{\sigma(S_{yy}(y(n)))}$$

$x(n)$ - sygnał z pogłosem późnym pogłosem zagęszczonym odbiciami

$y(n)$ - sygnał z późnym pogłosem z szumem

(reszta symboli opisana w rozdziale 5)

	Późny pogłos zagęszczony odbiciami	Późny pogłos z szumem
C		

Do wyznaczenia parametrów: koherencja, widmowa gęstość mocy oraz płaskość widma zostały wykorzystane funkcje z biblioteki *scipy* w języku Python, odpowiednio: *coherence* [25], *spectral_flatness* [26], *welch* [27].

7. Implementacja docelowego pogłosu

W tym rozdziale będę też pisał o rzeczach, które wprawdzie są oczywiste, ale w sumie tylko dla mnie i promotora – reszta tego nie wie. Np. o tym, że losowanie wartości za każdym razem innych skutkuje klikami, i - że w celu uniknięcia sytuacji, kiedy pojawiają się kliki przy modyfikacji wartości pokrętkła zastosowano jakieś tam rozwiązanie. - czyli w jednym z problemów w moim przypadku to 'budzenie' dalszych linii przy zwiększaniu rozmiaru pogłosu. Linie są posortowane, jednak przy wrażeniach słuchowych nie ma to znaczenia, bufor jest bowiem wypuszczany na wyjście dopiero po zsumowaniu wszystkiego. Filtry są sortowane odwrotnie, w celu większej filtracji linii opóźniających z większym opóźnieniem. Różnica czasów opóźnienia jest z jakiegoś +- zakresu, dla linii nie będących w stanie tego obsłużyć (czas opóźnienia mniejszy niż potencjalna wartość opóźnienia) różnica czasu jest mniejsza.

Będę też opisywał próby, które podejmowałem - jak mówił prof. Kleczkowski – wszystko co robimy opisywać, nawet jak nie wychodzi. Czyli np. próba zagęszczenia późnego pogłosu, np. też to, że przy tej wersji z feedbackiem nie słyszalne są już dalsze odbicia ze względu na konieczność ograniczenia amplitudy poszczególnych linii w celu zniwelowania negatywnych

efektów sprzężenia. I mogę napisać, że w celu osiągnięcia efektu przybliżonego do sytuacji z feedbackiem zastosowano odbicia będące dwukrotnością odbić, cichszych.

6.1. Wprowadzenie

Technologia VST popularna i pozwala na...

Jak wspomniano we Wstępie (1) możliwości obliczeniowe i wygoda vs przygotowywanie stanowiska do przeprowadzania testów. Realizacja zjawisk akustycznych w domenie cyfrowej może być jednak nietrywialna ze względu na konieczność zaawansowanej znajomości technik przetwarzania sygnałów cyfrowych (*Digital Signal Processing*) oraz języków programowania, a także sposobów implementacyjnych zapewniających wydajność obliczeniową, szczególnie w przypadku konieczności działania algorytmu w czasie rzeczywistym. Istnieje jednak coraz więcej gotowych narzędzi umożliwiających zaawansowane operacje związane DSP.

Zarówno testowe implementacje elementów pogłosu (rozdział 5), jak i docelowej aplikacji zostały zrealizowane z wykorzystaniem framework'u JUCE. JUCE to wieloplatformowy otwarto-źródłowy framework pozwalający na realizację aplikacji m. in. w technologii VST w języku C++ [16]. Narzędzie to pozwala także na realizację interfejsu użytkownika. Połączenie tych funkcji daje wygodne narzędzie do tworzenia aplikacji związanych z dźwiękiem.

6.3. Architektura aplikacji i opis klas stworzonych w języku C++

W celu obsługi wszystkich elementów składających się na wynikowy pogłos, przy jednoczesnym zachowaniu czytelności kodu, stworzono szereg klas i struktur języku C++, z których każda odpowiada za inną część obliczeń prowadzących do przetworzenia sygnału w celu uzyskania wynikowego sygnału.

Rys .. Schemat przedstawiający hierarchię oraz opis funkcji klas w programie

6.5. Połączenie wybranych rozwiązań (i ich modyfikacje) w celu uzyskania docelowego pogłosu

6.5.1. Linie opóźniające i sposób obliczania aktualnego bufora audio

Docelowym założeniem było stworzenie aplikacji w formacie VST ...

Potrzeba zastosowania bufora kołowego wynika z obecności problemu producenta i konsumenta [18]. Wybór docelowej struktury danych i narzędzi może mieć kluczowy wpływ na

Architektura aplikacji opiera się na

Read pointer to zmienna typu wskaźnik - wskazuje na pierwszy element tablicy, w tym wypadku na pierwszą wartość zapisaną aktualnym buforze audio. Mając informację zawierającą miejsce pierwszego elementu bufora oraz jego wartość jesteśmy w stanie ...

Przy projektowaniu należy zwrócić szczególną uwagę na poprawne *sklejenie* buforów

//Napiszę tutaj o sumowaniu buforów w przypadku, gdy wychodzi poza zakres bufora opóźniającego - że przesuwam wskaźnik z początku bufora na trochę dalej i w tamto miejsce wklejam resztę bufora, w przypadku niezrobienia tego idealnie mogą wystąpić niepożądane artefakty. np. w przypadku takim jak ja miałem, czyli należy zwrócić uwagę na poprawne ustawianie wskaźników i branie odpowiednich fragmentów buforów.

//Mogę dać podrozdział wskazówki implementacyjne – widziałem w jakiejś publikacji taki podrozdział. Mogę też nazwać go jakoś ogólniej, bo to słowo wskazówki w sumie sugeruje, że to będzie coś małego. Mogę też zrobić osobno taki podrozdział i dopisać tam to czego nie napiszę we właściwym opisie implementacji. Albo po prostu pisać w opisie implementacji, że aby zrobić coś tam należy zastosować/zrobić coś tam. Mogę też pisać co zrobiono, a potem, nawet jeśli będę powtarzał niektóre rzeczy to wypisać od kropek to co jest ważne, jeśli chodzi o implementację, np.: "długość bufora opóźniającego powinna być odpowiednio długa, określa jaki czas 'widzimy' w przeszłość, tym samym jak daleko z przeszłości możemy brać sygnał".

W celu stworzenia linii opóźniającej zawierającej kopię sygnału, zastosowano nowo stworzony bufor, o długości ...

6.5.2. Implementacja pierwszych odbić

Zgodnie z [2] podjęto działania mające na celu symulację pierwszych odbić. Symulacja polegała na dodaniu kopii sygnału o parametrach zgodnie z wartościami zaproponowanymi w podrozdziale 5.1 (oraz 5.2?). Odbicie boczne pod kątem 90 stopni (z lewej strony słuchacza) jest kopią sygnału bezpośredniego o zmniejszonej amplitudzie, opóźnieniu odpowiadającemu geometrii teoretycznie istniejącego pomieszczenia (80 ms) oraz filtracji dolnoprzepustowej imitującej tłumienie przez powietrze oraz pochłanianie przez materiał znajdujący się na teoretycznej ścianie.

6.5.3. Tłumienie odbić

6.5.4. Filtracja linii opóźniających

6.5.5. Zastosowanie międzyuszných zjawisk psychoakustycznych

W celu nadania przestrzenności sygnałowi zaimplementowano międzyuszne zjawiska psychoakustyczne zgodnie z opisem w rozdziale 3.

6.5.6. Synteza późnego ogona pogłosowego

W tym podrozdziale

8. Podsumowanie i wnioski

Celem niniejszej pracy było stworzenie cyfrowej symulacji zjawiska pogłosu w oparciu o analizę istniejących rozwiązań implementacyjnych oraz świadomości istnienia percepcyjnych zjawisk akustycznych. Aplikacja docelowo została zrealizowana w formacie VST. W niniejszej pracy dokonano także parametryzacji poszczególnych elementów pogłosu – na tej podstawie dokonany został wybór konkretnych parametrów/sposobów implementacyjnych w celu realizacji finalnego pogłosu. Przedstawione zostały najbardziej popularne podejścia do realizacji elementów sztucznego pogłosu oraz ich analiza. Przedstawiono także zaprojektowaną procedurę zrealizowania sztucznego pogłosu z zastosowaniem wybranych rozwiązań i ich modyfikacjami.

Historia realizacji tego typu projektów pokazuje, iż czynnik oceny subiektywnej jest nieodłącznym jej elementem – w przeszłości ze względu fakt niewielkich możliwości obliczeniowych w czasach publikacji kluczowych dla tematyki prac. Fakt uznawania oceny subiektywnej jako wiarygodnej wynikał także z powtarzalności oceny przy testach odsłuchowych. ...

//Być może istnieją bardziej efektywne i mniej skomplikowane sposoby rozwiązania powstałych problemów przy jednoczesnym uzyskaniu lepszego efektu końcowego. Z pewnością dalsze badania i próby są konieczne w celu rozszerzania wiedzy w tym zakresie.

Faktyczna sytuacja, jaka na miejsce w pomieszczeniu jest znacznie bardziej skomplikowana niż zaproponowane rozwiązania i implementacja auralizacji konkretnego pomieszczenia wymaga innego podejścia oraz bardziej szczegółowej analizy. Istnieje także konieczność proponowania innych parametrów pozwalających na obiektywne porównywanie sygnałów pod kątem zjawisk towarzyszących propagacji dźwięku w pomieszczeniu pod kątem pogłosu. Metodą na ocenę zamodelowanych zjawisk akustycznych mogą być także badania na grupie słuchaczy, którzy będą oceniać sygnał pod względem parametrów zaproponowanych parametrów i kryteriów ocen – tak jak miało to miejsce w przeszłości.

Należy pamiętać, iż zjawiska zastosowane w implementacji dotyczące wrażeń przestrzennych związane są jedynie z lokalizacją źródła dźwięku w płaszczyźnie poziomej. Istnieje więcej czynników wpływających na wrażenie przestrzenności, takie jak np. wpływ małżowin usznych lub geometrii głowy i ciała słuchacza ze szczególnym wskazaniem na Funkcję Transmitycji Głowy (*Head Related Transfer Function*). A także innych zjawisk, dzięki którym możliwe jest lokalizowanie źródła poza płaszczyzną poziomą. Dokładna identyfikacja i parametryzacja tych zjawisk jest wciąż obiektem badań, dalszy ich rozwój pozwoli na dokładniejsze odwzorowanie wrażenia przestrzenności w przypadku chęci ich stosowania w podobnych projektach. Ponadto, ze względu na fakt działania zjawisk psychoakustycznych w różnych zakresach częstotliwości, wrażenie przestrzenności silnie zależy od rodzaju sygnału i zawartości jego widma.

9. Bibliografia

[1] Jot J. – M., Chaigne A., Digital Delay Networks for Designing Artificial Reverberators, 90th Convention of Audio Engineering Society, luty 1991

[2] Barron

[3] Schroeder - Colorless Sounding Artificial Reverberation

[4] Schroeder - Natural Sounding Artificial Reverberation

[5] J. Stautner and M. Puckette - Designing Multi-Channel Reverberators

[6] Beranek

[7] <http://musicweb.ucsd.edu/~trsmth/space175/space175.pdf>

[8] https://www.dsprelated.com/freebooks/pasp/Artificial_Reverberation.html

[9] Moorer – About this reverberation business

Commented [AK1]: Follow Up: [9] Moorer – About this reverberation business

[10] V. Välimäki, B. Holm-Rasmussen, B. Alary and H. M. Lehtonen – Late Reverberation Synthesis Using Filtered Velvet Noise

[11] P. Kleczkowski – Percepcja Dźwięku

[] https://librosa.github.io/librosa/generated/librosa.feature.spectral_flatness.html

[] <https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.coherence.html>

[16] <https://github.com/WeAreROLI/JUCE>

... strony zawierające dokumentację konkretnych klas z JUCE

... git

[17] <http://diamonddissertation.blogspot.com/2010/05/rayleigh-duplex-theory.html>

[18] <https://embedjournal.com/implementing-circular-buffer-embedded-c/>

[19]

https://en.wikipedia.org/wiki/Spectral_density#:~:text=Energy%20spectral%20density%20describes%20how%20the%20energy%20of,transients%E2%80%94that%20is%2C%20pulse-like%20signals%E2%80%94having%20a%20finite%20total%20energy.

[20] https://pl.wikipedia.org/wiki/Widmowa_g%C4%99sto%C5%9B%C4%87_mocy

[21] https://en.wikipedia.org/wiki/Spectral_flatness

[22] https://en.wikipedia.org/wiki/Spectral_density#Cross-spectral_density

[24] [https://en.wikipedia.org/wiki/Coherence_\(signal_processing\)](https://en.wikipedia.org/wiki/Coherence_(signal_processing))

[25] Dokumentacja funkcji *coherence* z biblioteki *scipy* w języku *Python*:

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.coherence.html>

[26] Dokumentacja funkcji `spectral_flatness` z biblioteki *scipy* w języku *Python*:

https://librosa.org/librosa/generated/librosa.feature.spectral_flatness.html

[27] Dokumentacja funkcji `welch` z biblioteki *scipy* w języku *Python*:

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.welch.html#scipy.signal.welch>