

Spis treści:

1. Wstęp

1.1. Cel i zakres pracy

2. Symulacja i percepcja zjawiska pogłosu

2.1. Główne metody realizacji

2.1.1. Pogłos splotowy

2.1.2. Rozwiązania elektromechaniczne

2.1.3. Cyfrowe symulacje pogłosu

2.2. Sposoby implementacji elementów pogłosu

2.3. Zjawiska psychoakustyczne przy percypowaniu pogłosu

3. Implementacja i właściwości istniejących rozwiązań

3.1. Opis wybranych parametrów statystycznych do analizy sygnałów

3.2. Testy wybranych sposobów realizacji elementów pogłosu

4. Propozycja i implementacja metody hybrydowej

4.1. Koncepcja realizacji

4.2. Architektura aplikacji

4.3. Połączenie testowanych rozwiązań w docelowy pogłos

4.3.1. Linie opóźniające i sposób obliczania aktualnego bufora audio

4.3.2. Implementacja pierwszych odbić

4.3.3. Tłumienie odbić

4.3.4. Filtracja linii opóźniających

4.3.5. Uprzestrzennienie pogłosu

4.3.6. Późny ogon pogłosowy

4.4. Część funkcjonalna aplikacji

5. Podsumowanie i wnioski

Bibliografia

Commented [1]: tytuł nie za bardzo odpowiada zawartości; może lepiej: symulacja i percepcja zjawiska pogłosu?

Commented [2]: czy nie lepszy byłby np. tytuł: implementacja i właściwości istniejących rozwiązań? albo podobny – lepiej oddający zawartość

Commented [3]: tu i w 3.1 słowo „opis” nie jest potrzebne

1. Wstęp

Zjawisko pogłosu akustycznego towarzyszy człowiekowi w niemal każdym momencie życia. Pełni kluczową rolę w odczuwaniu przestrzeni akustycznej oraz lokalizacji źródła dźwięku w pomieszczeniu. Bardzo często pogłos towarzyszący dźwiękowi bezpośredniemu wiąże się z jego atrakcyjnością dla człowieka - zarówno naturalny pogłos pomieszczenia (np. w sali koncertowej), jak i przy odsłuchu we własnych warunkach odsłuchowych (nagranie z naturalnym, zarejestrowanym pogłosem lub pogłosem sztucznym dodanym na etapie miksowania utworu).

Badania dotyczące pogłosu prowadzone pod koniec XX wieku wymagały skomplikowanego toru sygnałowego oraz zwykle potrzeby wykorzystania pola swobodnego i dyfuzyjnego do wykonywania testów lub badań [2]. Współczesne możliwości znacznie to ułatwiają. Przede wszystkim możliwość odwzorowania większości zjawisk w domenie cyfrowej, jak również znacznie ułatwiające badania o wiele większe możliwości obliczeniowe. W literaturze z tego okresu można zaobserwować obszerne opisy [2][12]. W odróżnieniu od tej sytuacji, korzystając z narzędzi cyfrowych, w przypadku potrzeby dokonania niewielkiej zmiany w parametrach pogłosu, jedyne niezbędne czynności do wykonania to niewielka modyfikacja kodu oraz odczekanie niewielkiej ilości czasu w celu otrzymania nowej symulacji pogłosu. Co więcej, o wiele mniejsze możliwości obliczeniowe skutkowałyby pozwoleniem na większą ilość testów w tym samym czasie, a co za tym idzie łatwość opisu.

W związku z powyższym otwierają się szersze możliwości eksperymentowania - w niniejszej pracy możliwość ta zostaje wykorzystana do stworzenia cyfrowej symulacji pogłosu. Implementacja jest poprzedzona obiektywną oceną jakości brzmienia poszczególnych jego elementów.

1.1. Cel i zakres pracy

Celem pracy są projekt i implementacja cyfrowej symulacji zjawiska pogłosu w formacie VST poprzedzone oceną elementów składowych pogłosu przy użyciu zaproponowanych w ramach pracy parametrów.

Praca obejmuje przegląd głównych idei realizacji sztucznego pogłosu oraz percepcyjnych zjawisk przestrzennych, testy wybranych rozwiązań realizacji poszczególnych elementów pogłosu i ich ocena na podstawie zaproponowanych parametrów w sposób obiektywny, jak również implementacja docelowej aplikacji w formacie VST. Docelowa aplikacja będzie pogłosem w postaci efektu, w odróżnieniu do uzyskania jak najbardziej realistycznego pogłosu. Implementacja będzie hybrydą łączącą elementy istniejących rozwiązań na podstawie ich oceny zaproponowanymi przez autora parametrami. Praca omawia również szczegóły koncepcji realizacji docelowego pogłosu oraz algorytm prowadzący do uzyskania finalnego efektu.

Commented [4]: gramatyka, sens

Commented [5]: lepiej: dzisiejsze, współczesne

Commented [Na6]: gramatyka

Commented [Na7]: zawartość tego rozdziału jest dobra; na koniec warto dopisać, że w związku z tym „nasuwa się pomysł (czy potrzeba, albo konieczność) aby...” albo „możliwość eksperymentowania” - i tu w 2-3 zdaniach wprowadzić w cel i zakres, które pojawiają się za chwilę; chodzi o „domknięcie” wstępu

Commented [8]: gramatyka

Commented [Na9]: zawartość tego rozdziału jest dobra; na koniec warto dopisać, że w związku z tym „nasuwa się pomysł (czy potrzeba, albo konieczność) aby...” albo „możliwość eksperymentowania” - i tu w 2-3 zdaniach wprowadzić w cel i zakres, które pojawiają się za chwilę; chodzi o „domknięcie” wstępu

Commented [Na10]: wiele cytowań: [2][12] – bez przecinka pomiędzy; te odnośniki akurat lepiej dać na końcu zdania

Commented [Na11]: zawartość tego rozdziału jest dobra; na koniec warto dopisać, że w związku z tym „nasuwa się pomysł (czy potrzeba, albo konieczność) aby...” albo „możliwość eksperymentowania” - i tu w 2-3 zdaniach wprowadzić w cel i zakres, które pojawiają się za chwilę; chodzi o „domknięcie” wstępu

Commented [12]: to trzeba lepiej sformułować

Commented [13]: gramatyka, sens

Commented [14]: zawartość tego rozdziału jest dobra; na koniec warto dopisać, że w związku z tym „nasuwa się pomysł (czy potrzeba, albo konieczność) aby...” albo „możliwość eksperymentowania” - i tu w 2-3 zdaniach wprowadzić w cel i zakres, które pojawiają się za chwilę; chodzi o „domknięcie” wstępu

Commented [15]: otwierają się szersze możliwości

Commented [16]: żeby to pasowało do pierwszej części zdania, tą pierwszą część trzeba inaczej sformułować (eksperymentowania z czym?)

Commented [17]: są, albo (jeżeli „jest”) to „poprzedzona”

Commented [18]: proszę to nieco lepiej wyjaśnić (ja rozumiem różnicę, ale mimo wszystko to nie jest zbyt klarowne)

Commented [19]: hybrydą łączącą elementy istniejących rozwiązań

Commented [20]: Praca omawia również szczegóły koncepcji

Commented [21]: uwaga: nie komentuję błędów gramatyki i stylu w dalszych rozdziałach – proszę je znaleźć i poprawić zgodnie z sugestiami umieszczonymi do tego miejsca

2. Symulacja i percepcja zjawiska pogłosu

Pogłos akustyczny składa się z opóźnionych i osłabionych kopii dźwięku bezpośredniego. Zawartość częstotliwościowa każdego z odbić zależy od kierunkowości źródła oraz pochłaniania odbijających powierzchni [dafx]. Zjawisko to występuje naturalnie w większości środowisk i ma mocniejszy efekt przy powierzchniach odbijających.

Pogłos to czas, po którym poziom natężenia dźwięku spadnie o 60 dB od wyłączenia źródła dźwięku [Drobner nowszy]. Fale akustyczne odbijają się od elementów wnętrza pomieszczenia z coraz to większą stratą energii [Drobner]. Fale emitowane ze źródła dźwięku zwane są dźwiękiem bezpośrednim, odbite od elementów pomieszczenia tworzą fale wtórne, które nakładają się na fale dźwięku bezpośredniego. Po każdym odbiciu dźwięku bezpośredniego następują odbicia wtórne - dzięki temu podczas emisji wzrasta energia akustyczna dźwięku odbitego. Po wyłączeniu źródła następuje faza zaniku dźwięku odbitego, wskutek zmniejszania się ciśnienia akustycznego [Oleszkowicz]. Dwa najważniejsze czynniki wpływające na pogłos to całkowita objętość wnętrza oraz łączne pochłanianie przez wszystkie elementy wnętrza. Pomiary praktyczne Sabine'a doprowadziły do powstania wzoru opisującego czas pogłosu w pomieszczeniu [Drobner]:

$$T = 0,164 * V/A$$

V - objętość pomieszczenia,

A - średni współczynnik pochłaniania wszystkich materiałów w pomieszczeniu,

T – czas pogłosu

Objętość pomieszczenia i średni współczynnik pochłaniania nie są od siebie zależne. Zwiększenie objętości, podobnie jak zmniejszenie średniego współczynnika pochłaniania zwykle skutkuje zwiększeniem czasu pogłosu [Drobner]. Skorygowany wzór nazywany jest wzorem Eyringa i przyjmuje następującą postać [Drobner nowszy]:

$$T = -0,0713 * V / (P * \log(1 - \alpha))$$

alfa - współczynnik pochłaniania

Pogłos może być także wytworzony w sposób sztuczny na wiele sposobów [28].

2.1. Główne metody realizacji

2.1.1. Rozwiązania elektromechaniczne

W klasycznym podejściu do realizacji pogłosu wykorzystywano komorę pogłosową - pomieszczenie zaprojektowane specjalnie po to, aby uzyskać w nim jak najdłuższy czas pogłosu. W pomieszczeniu tym, długi czas pogłosu uzyskuje się dzięki odpowiednio dużej objętości, silnie odbijającym, nierównoległym ścianach oraz zastosowaniu dodatkowych, silnie odbijających powierzchni. Realizując nagranie w takim pomieszczeniu uzyskuje się

Commented [Na22]: tytuł nie za bardzo odpowiada zawartości; może lepiej: symulacja i percepcja zjawiska pogłosu?

sygnał wzbogacony o jego naturalny pogłos. Można uzyskać wiele wersji sygnału w zależności od pozycji mikrofonu w pomieszczeniu oraz pozycji źródła.

Jednym z najwcześniejszych rozwiązań pozwalających na zastosowanie pogłosu w nagraniach był pogłos płytowy. Jego konstrukcja opierała się na prostokątnej, stalowej, obramowanej płycie z przymocowanym przetwornikiem elektromechanicznym. Przetwornik rejestrował sygnał dźwiękowy, którego drgania były przenoszone na płytę. Następnie drgania płyty były odbierane przez przetwornik elektroakustyczny, tym sposobem był uzyskiwany sygnał z pogłosem powstałym dzięki drganiu płyty. Odpowiedź impulsowa takiego układu różni się od odpowiedzi impulsowej pomieszczenia - występuje brak wyraźnych odbić, a jako całość jest bardziej gładka i bardziej przypominająca szum [dafx].

Podobnym rozwiązaniem jest pogłos sprężynowy, w którym przetwornik elektromechaniczny wprawia w drgania sprężynę. Na drugim końcu sprężyny znajduje się przetwornik elektroakustyczny odbierający sygnał z dodanym pogłosem powstałym dzięki drganiu sprężyny. W odróżnieniu do pogłosu płytowego, sprężyna jest w stanie wykonywać drgania zarówno podłużne, jak i poprzeczne. Złożenie tych dwóch rodzajów drgań prowadzi do charakterystycznego brzmienia łączącego w sobie cechy brzmienia rzeczywistego pomieszczenia z pogłosem płytowym [dafx].

2.1.2. Pogłos splotowy

Najważniejszym pojęciem w akustyce pomieszczeń jest odpowiedź impulsowa pomieszczenia. Opisuje ona akustykę pomieszczenia i jest realizowana jako rejestracja nieruchomego źródła dźwięku w jednym punkcie. Jeśli odpowiedź impulsowa pomieszczenia jest znana, najbardziej wierny pogłos sygnał pogłosowy można uzyskać wykonując operację splotu sygnału z odpowiedzią impulsową. Operację taką można wykonać traktując każdą próbkę odpowiedzi impulsowej jako współczynnik filtru FIR, a następnie dokonując nim filtracji sygnału wejściowego [dafx]. Definicja splotu przedstawia się jako zależność [29]:

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m],$$

gdzie:

f, g - sygnały poddawane operacji splotu

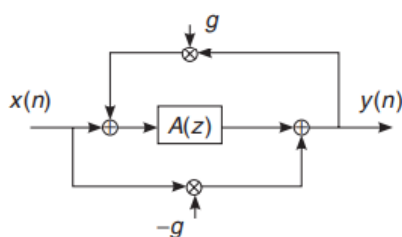
Ze względu na małą wydajność obliczeniową operacji splotu w dziedzinie czasu, stosuje się podejście wykorzystujące tożsamość operacji splotu w dziedzinie czasu z operacją mnożenia w dziedzinie częstotliwości. Wykonywanie obliczeń w dziedzinie częstotliwości jest znacznie bardziej wydajne obliczeniowo. Często jest to jedyne możliwe rozwiązanie w przypadku konieczności działania algorytmu w czasie rzeczywistym [dafx].

2.1.2. Rozwiązanie oparte o filtry wszechprzepustowe i grzebieniowe

Jednymi z najważniejszych, pionierskich prac dotyczących realizacji sztucznego pogłosu są prace M. Schroeder'a z lat 60 XX wieku [3][4]. Zaproponował on podejście wykorzystujące filtry wszechprzepustowe i grzebieniowe jako narzędzie umożliwiające symulację pogłosu. Filtr wszechprzepustowy przedstawia się jako zależność [dafx]:

$$y(n) = -g \cdot x(n) + x(n - m) + g \cdot y(n - m)$$

gdzie m oznacza długość opóźnienia w próbkach.



Rys. 2.2. Struktura filtru wszechprzepustowego [dafx].

W schemacie filtru wszechprzepustowego przedstawionego na Rys. 2.2., współczynnik g oznacza wzmacnienie, a $A(z)$ linię opóźniającą. Zastosowanie szeregu takich filtrów pozwala na uzyskanie odpowiedzi impulsowej o płaskiej charakterystyce [30]. Podejście oparte o filtry wszechprzepustowe krótkim czasie stało się standardem w niemal wszystkich stosowanych algorytmach sztucznego pogłosu [dafx].

2.1.2. Feedback Delay Network

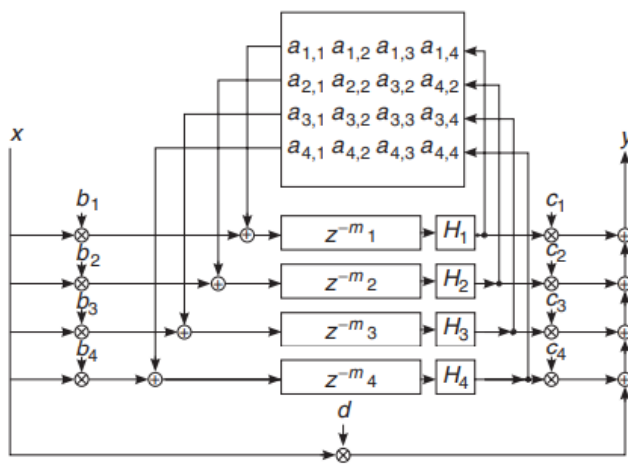
W 1982 roku J. Stautner i M. Puckette jako strukturę realizującą pogłos zaproponowali sieć linii opóźniających połączonych w pętli sprzężenia zwrotnego [5]. W późniejszym czasie struktura taka została nazwana siecią Feedback Delay Network. Pierwotnie zaproponowana sieć miała postać [dafx]:

$$y(n) = x(n - m) + g \cdot y(n - m)$$

Linia opóźniająca m w filtrze wszechprzepustowym została zastąpiona szeregiem linii o różnych długościach, a parametr g oznaczający wzmacnienie – macierzą sprzężenia zwrotnego daną jako:

$$\mathbf{G} = g \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} / \sqrt{2}$$

Dzięki zastosowaniu sieci pogłosowej FDN uzyskać można wysoką gęstość późnego ogona pogłosowego [dafx]. Zależy ona jednak od liczby linii opóźniających i wzajemnych relacji długości linii opóźniających. Według teorii, największą gęstość późnego ogona pogłosowego można uzyskać, kiedy stosunki długości linii opóźniających są liczbami pierwszymi [1][4][5].



Rys. 2.3. Schemat podstawowej sieci pogłosowej FDN.

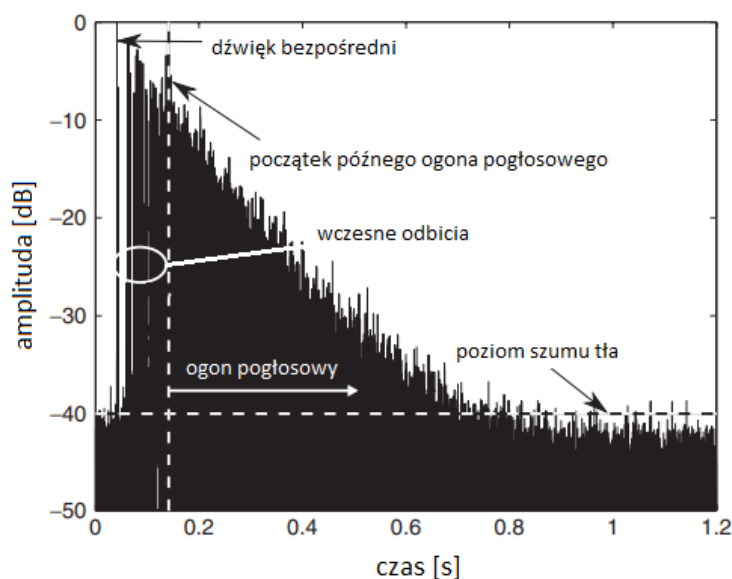
Badania nad rozwojem sieci FDN zostały rozwinięte w pracy Jota [1].

2.1.3. Cyfrowe symulacje pogłosu

Od wczesnych lat rozwoju prac nad realizacją sztucznego pogłosu, najbardziej popularne jest rozgraniczenie go na dwie części: wczesne 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][dafx] (Rys 2.1.). Wczesne 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 wczesnych 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 wynikowy sygnał, a następnie rozważyć zjawiska binauralne oraz inne rozwiązania w celu nadania pogłosowi realizmu.

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 [?].



Rys. 2.1. Schematyczne przedstawienie odpowiedzi impulsowej z wyszczególnieniem dźwięku bezpośredniego, wczesnych odbić oraz późnego ogona pogłosowego [dafx].

2.2. Sposoby implementacji elementów pogłosu

A). 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 – 80 ms). 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 powyżej „poszerzenia się” źródła dźwięku. Przy zwiększeniu poziomu dźwięku odbicia boczne wrażenie przestrzenności zwiększało się. W badaniach został przebadany także wpływ odbicia od sufitu na wrażenie przestrzenności - wpływ odbicia od 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.

B). Czasy opóźnienia linii opóźniających

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

Krótkie czasy opóźnienia skutkują większym ubarwieniem dźwięku - sytuacja taka ma miejsce w małych pomieszczeniach. Zjawisko takie byłoby nienaturalne w pomieszczeniach o wysokim czasie pogłosu. Według autorów [5] 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 [5]. Zjawisko to ma wytłumaczenie w interferencji 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 [2].

Commented [AK23]: chodzi o publikację: M. Barron - The subjective effects of first reflections in concert halls – the need for lateral reflections, 1969

Commented [24]: między wartość a jednostkę (tu i wcześniej/dalej) lepiej wstawić twardą spację (w LibreOffice: ctrl+shift+spacja, nie wiem jak MS)

C). Filtracja linii opóźniających

Schroeder zaproponował [4] 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. Sytuacja ta powoduje znaczne zagęszczenie ilości odbić w jednostce czasu.

J. Stautner i M. Puckette 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 [5].

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

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].

Istnieją następujące subiektywne sposoby na identyfikację niewystarczającej gęstości widmowej późnego pogłosu [1]:

- 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).

W celu wyłączenia czynnika subiektywnego z oceny pogłosu, w rozdziale 3 znajduje się zaproponowany w ramach pracy sposób oceny jakości w zależności od rodzaju filtracji.

D). Późny ogon pogłosowy

2.4. Zjawiska psychoakustyczne przy percypowaniu pogłosu

Mechanizmy słyszenia przestrzennego pozwalają na odczucie kierunku, z którego dobiega dźwięk, a także w pewnej mierze odległości od źródła dźwięku [11].

Teoria oparta na podstawie wyników subiektywnych badań słuchowych 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].

Commented [25]: przy wypunktowaniu: punkty od małych liter, po każdym przecinek, po ostatnim kropka

Dwie najważniejsze wielkości 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.

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, że stopień wrażenia przestrzenności jest związany ze stopniem niekoherencji sygnałów z obu kanałów [7].

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].

3. Ocena jakości symulacji pogłosu

Obiektywna ocena jakości pogłosów wymaga odpowiedniego doboru parametrów statystycznych i ich zastosowania w celu uzyskania wartości liczbowych będących podstawą do uzasadnienia oceny danego wariantu na tle innego. Posłużą do tego zdefiniowane na potrzeby pracy parametry (w podrozdziale 3.2) – będą one wyznacznikiem jakości pogłosów. Do zdefiniowania tych parametrów zostaną wykorzystane istniejące, znane parametry statystyczne stosowane do analizy widmowej i czasowej sygnału oraz do oceny podobieństwa sygnałów. Parametry te opisane są w podrozdziale 3.1. W podrozdziale 3.2 opisane są także warianty pogłosów zaimplementowanych przez autora. Przedstawione są również wartości niektórych parametrów statystycznych – w formie graficznej w dziedzinie częstotliwości lub czasu, jak również wartości liczbowe stworzonych na potrzeby pracy parametrów (tabele). Testy te przeprowadzone będą dla trzech sygnałów: mowa ludzka, sygnał MLS oraz nagranie instrumentu – gitary klasycznej w warunkach bezechowych. Następnie dokonana zostanie interpretacja wyników oraz wysunięte wnioski mające wpływ na opisaną w rozdziale 4 implementację docelowego pogłosu.

Commented [Na26]: o to chodziło - dobrze

Commented [27]: dobrze by było pokazać dobrej jakości spektrogramy – wówczas dalsze rys. w tym rozdziale będą miały więcej sensu (bo będzie baza, z którą będzie je można porównać)

Commented [28]: o to chodziło - dobrze

3.1. Opis wybranych parametrów statystycznych do analizy sygnałów

Opisane w rozdziale 2 sposoby implementacji poszczególnych elementów pogłosu zawierają wskazówki dotyczące pożądanej zawartości widmowej i czasowej pogłosu oraz cech sygnału wpływających na wrażenie przestrzenności. W celu oceny jakości pogłosu przedstawia się stosowane parametry statystyczne mogące być 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 jest 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]:

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

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

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

gdzie:

P – moc sygnału,

[F₁,F₂] - przedział częstotliwości

S(f) - widmowa gęstość mocy sygnału

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

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

gdzie:

S_{xx}(ω) – widmowa gęstość mocy,

T = NΔt,

n – kolejna próbka sygnału,

N – liczba próbek sygnału,

t – czas.

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 wzajemną gęstością widmową [22] między sygnałami x i y , a $G_{xx}(f)$ i $G_{yy}(f)$ są gęstościami widmowymi sygnałów samych z sobą. Wielkość ta opisuje relację pomiędzy dwoma sygnałami ... [24].

3.2. Testy stosowanych rozwiązań implementacji pogłosu

Implementacje poniższych symulacji pogłosu zostały zrealizowane korzystając z frameworku JUCE w języku C++ - dokładny opis implementacji oraz uzasadnienie wyboru narzędzia znajdują się w rozdziale 4.

Warianty pogłosów zaimplementowane przez autora:

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

- pogłos złożony z linii opóźniających o czasach opóźnienia jako liczby losowe,
- pogłos złożony z linii opóźniających o czasach opóźnienia jako liczby wzajemnie pierwsze (w nawiązaniu do [1][4][5]),
- pogłos złożony z linii opóźniających o czasach opóźnienia zawierających się w zakresie 1:1.5 (w nawiązaniu do [4]).

B). Filtracja linii opóźniających:

- filtracja dolnoprzepustowa każdej linii opóźniającej w jednakowy sposób,
- filtracja dolnoprzepustowa o częstotliwości odcięcia tym niższej, im wyższa wartość opóźnienia linii opóźniającej (w nawiązaniu do [4]),
- filtracja w zakresie pochłaniania materiałów (Większość materiałów, z których składają się ściany mają największy stopień pochłaniania energii akustycznej w zakresie 500 Hz – 2 kHz) (w nawiązaniu do [5]).

C). Późny ogon pogłosowy:

- pogłos złożony z linii opóźniających ze sprzężeniem mnożony przez filtrowany szum biały,
- pogłos złożony z linii opóźniających ze sprzężeniem zagęszczony odbiciami w późnej fazie pogłosu przez dodanie tam większej liczby linii opóźniających.

W celu obiektywnego porównania zaimplementowanych wersji, na potrzeby niniejszej pracy zdefiniowano parametr będący podstawą do oceny wpływu doboru wartości czasów opóźnienia linii opóźniających, późnego pogłosu oraz sposobu filtracji na jakość pogłosu. Zdefiniowany parametr jest odpowiedzią na powtarzające się w literaturze [1][3][4] podkreślenia wpływu wysokiej gęstości oraz płaskości widma na jakość pogłosu. Zdefiniowany przez autora parametr przedstawia się jako zależność:

Commented [Na29]: jeżeli to Pana wzór, to proszę to wyraźnie zaznaczyć (a rozumiem, że tak jest); w przeciwnym wypadku podać źródło

$$B_x := \frac{\frac{(H_{xy} + H_{xz})}{2} E(S_{xx}(x(n)))}{\sigma(Flatness(x(n))) * \sigma(S_{xx}(x(n)))} \quad (5)$$

gdzie:

H_{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 - sygnał pogłosowy zawierający linie opóźniające o czasach opóźnienia będącymi liczbami wzajemnie pierwszymi

y - sygnał pogłosowy zawierający linie opóźniające o czasach opóźnienia będącymi liczbami losowymi

z - 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

Wzór przyjmuje formę analogiczną dla sygnałów $y(n)$ oraz $z(n)$.

D). Psychoakustyczne zjawiska przestrzenne

- pogłos z różnicą w poziomie pierwszych odbić bocznych – w celu potwierdzenia zaobserwowanej w tezy o większej przestrzenności przy wyższym poziomie tych odbić (w nawiązaniu do [2]),
- pogłos z zastosowaniem różnicy czasu opóźnienia między kanałami dla linii opóźniających (w nawiązaniu do [11]),
- pogłos z zastosowaniem różnicy amplitudy pomiędzy kanałami dla linii opóźniających (w nawiązaniu do [11]),
- pogłos z zastosowaniem różnicy pomiędzy filtracją linii opóźniających obu kanałów (w nawiązaniu do [7]).

Jak opisano w podrozdziale 2.2, 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, iż 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 opisano zdefiniowanym na potrzeby pracy parametrem *przestrzenność*. Zdefiniowanie parametru jest odpowiedzią na powtarzające się w literaturze podkreślanie istotności cech sygnału wpływające na jakość pogłosu [1][3][4] oraz jego przestrzenność [2][7]. Cechy te to: wysoka gęstość widma oraz stopień niekoherencji

Commented [30]: mamy trochę kolizji – x i y oznaczają jednocześnie różne rzeczy w G_{xy} i w B_x – to trzeba uzgodnić

Commented [Na31]: a może inne oznaczenia zamiast xyz - bardziej czytelne?

Commented [Na32]: niejasne

sygnałów z obu kanałów. Zdefiniowany przez autora parametr *przestrzenność* przedstawia się jako:

Commented [33]: może jednak jakiś krótszy symbol? będzie wygodniejszy na wykresach i w tabelach

$$\textit{przestrzenność} = \frac{E(S_{xx}(x(n)))}{E(C_x(x_L(n), x_R(n))) * \sigma(S_{xx}(x(n))) * \sigma(C_x(x_L(n), x_R(n)))}$$

gdzie:

σ - średnie odchylenie standardowe

S_{xx} - widmowa gęstość mocy

C_{xy} - koherencja lewego i prawego kanału sygnału

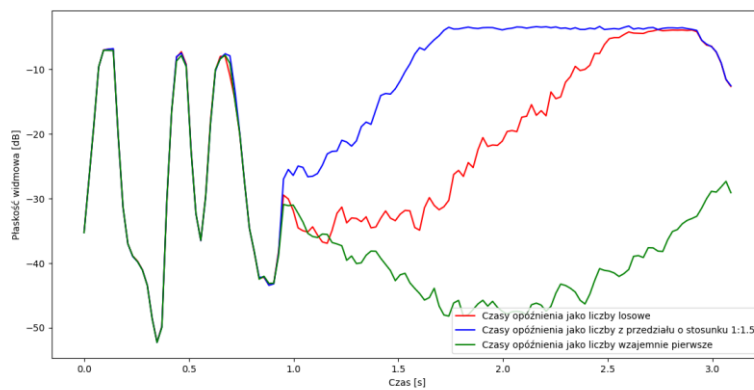
E - średnia arytmetyczna

$x_L(n)$ - lewy kanał sygnału $x(n)$

$x_R(n)$ - prawy kanał sygnału $x(n)$

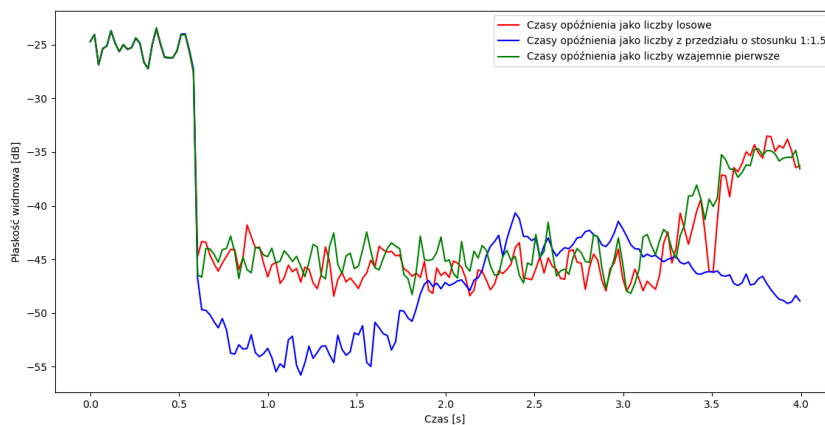
Do wyznaczenia parametrów: koherencja, widmowa gęstość mocy oraz płaskość widma zostały wykorzystane funkcje z biblioteki *scipy* w języku Python, odpowiednio: *coherence* [25], *spectral_flatness* [26], *welch* [27]. //tutaj będzie opisana szczegółowo metodologia wyznaczania wartości parametrów (parametr B oraz Pr)

Na rys. 3.1. widać, iż dla sygnału mowy z liniami opóźniającymi o czasach opóźnienia będącymi liczbami zawierającymi się z zakresie 1:1.5, płaskość widma przez większość czasu trwania sygnału jest największa. Parametr ten ma bezpośredni wpływ na wartość parametru B będący wyznacznikiem jakości pogłosu (wzór 5). Wartości parametru B przedstawiono w tab. 3.1. //ten komentarz będzie bardziej rozbudowany i odniesie się też do innych sygnałów i innych wykresów oraz wartości w tabelach.

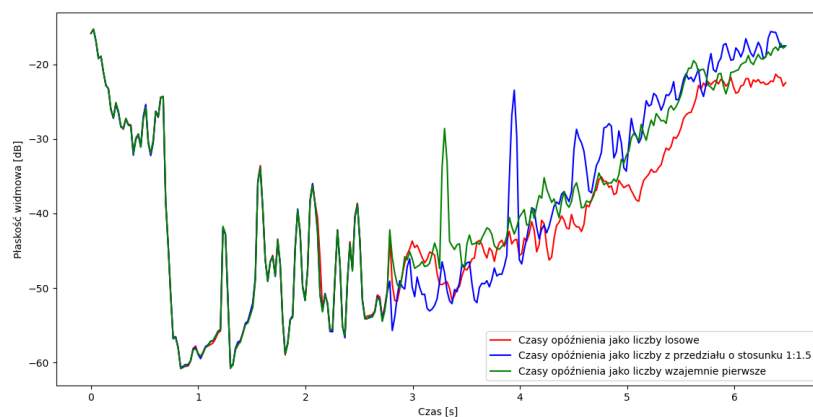


Rys. 3.1. Płaskość widmowa dla różnego rodzaju generowania linii opóźniających - sygnał mowy.

Commented [34]: w miarę możliwości opis rysunków i wykresów też powinien być w języku polskim



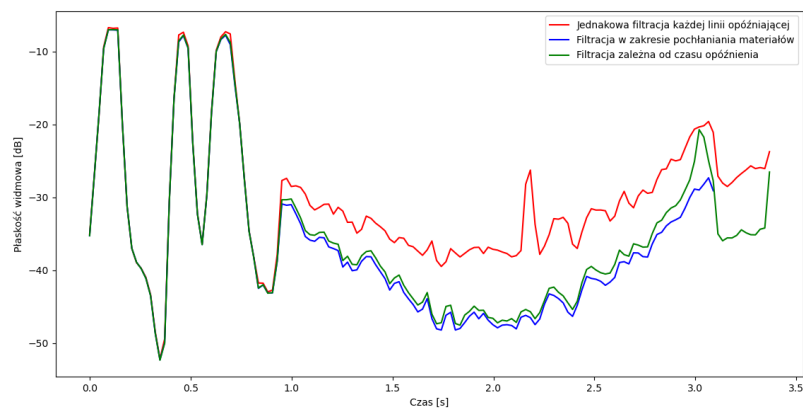
Rys. 3.2. Płaskość widmowa dla różnego rodzaju generowania linii opóźniających - sygnału MLS.



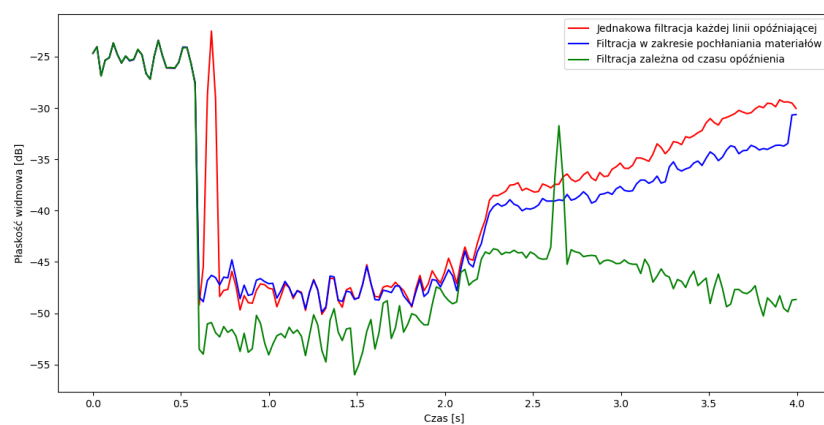
Rys. 3.3. Płaskość widmowa dla różnego rodzaju generowania linii opóźniających - gitara klasyczna.

Tab. 3.1. Wartości parametru B dla różnego rodzaju generowania czasów linii opóźniających

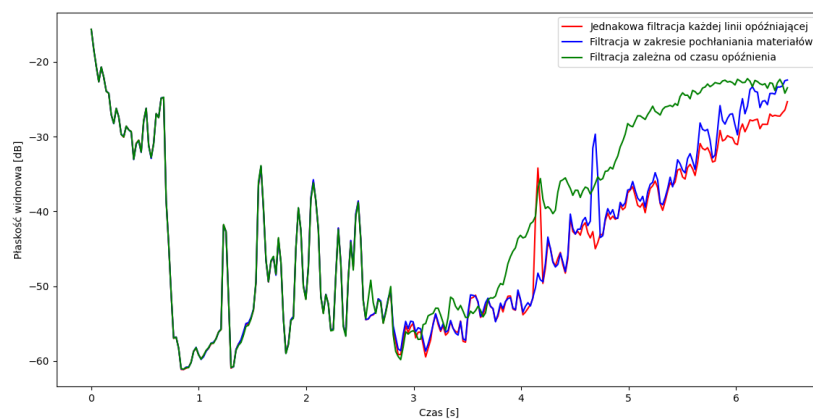
B	Losowe czasy opóźnienia	Czasy opóźnienia jako liczby pierwsze	Czasy opóźnienia jako liczby zawierające się w zakresie liczb w stosunku 1:1.5
Sygnał mowy	0,57	0,22	1,0
MLS	0,64	1,0	0,39
Gitara klasyczna	0,80	1,0	0,98



Rys. 3.4. Płaskość widma symulacji różnych wariantów filtracji - sygnał mowy



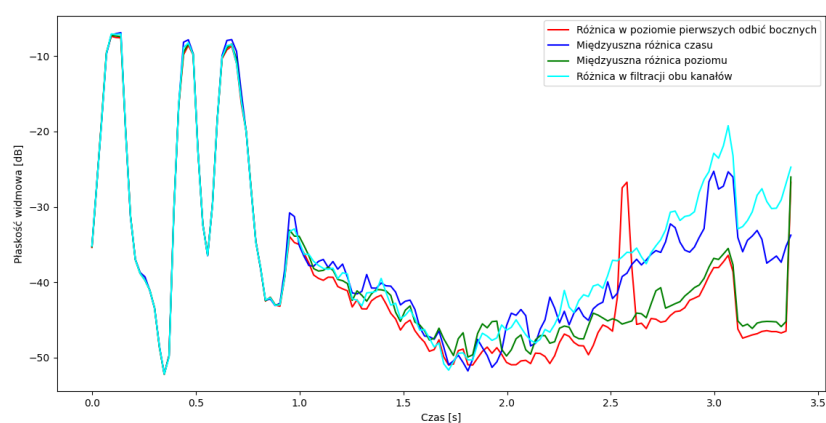
Rys. 3.4. Płaskość widma symulacji różnych wariantów filtracji - sygnał MLS



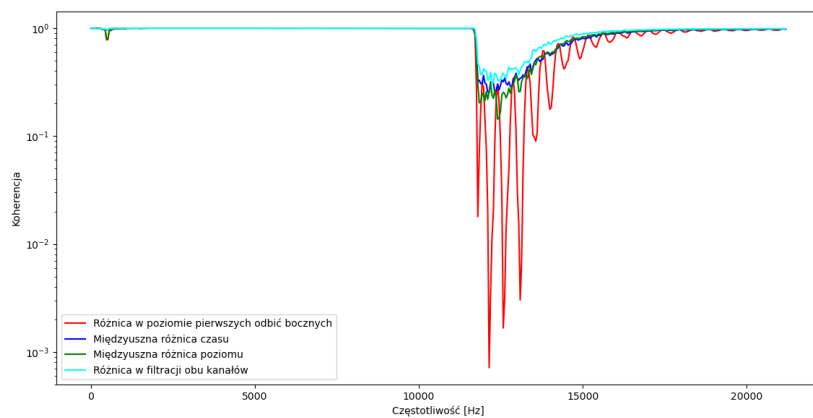
Rys. 3.4. Płaskość widma symulacji różnych wariantów filtracji – gitara klasyczna

Tab. 3.2. Wartości parametru B dla pogłosów z różnymi wariantami filtracji

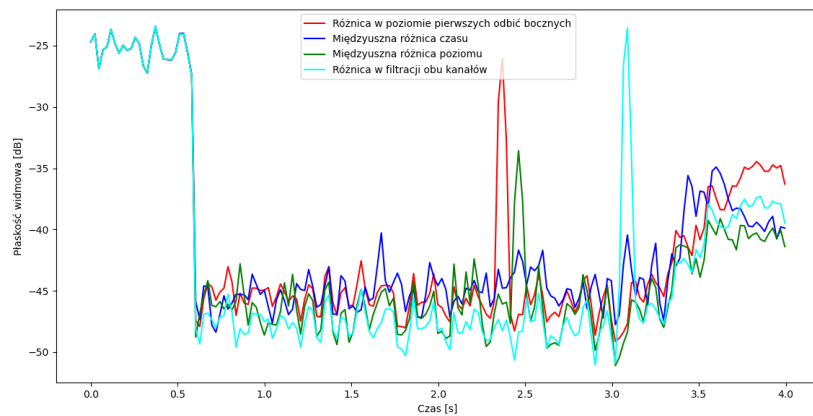
B	Filtracja każdej linii w jednakowy sposób	Filtracja w zakresie pochłaniania materiałów	Tym niższa częstotliwość odcięcia filtru, im wyższy czas opóźnienia
Sygnał mowy	1,0	0,41	0,07
MLS	1,0	0,79	0,11
Gitara klasyczna	0,47	0,92	1,0



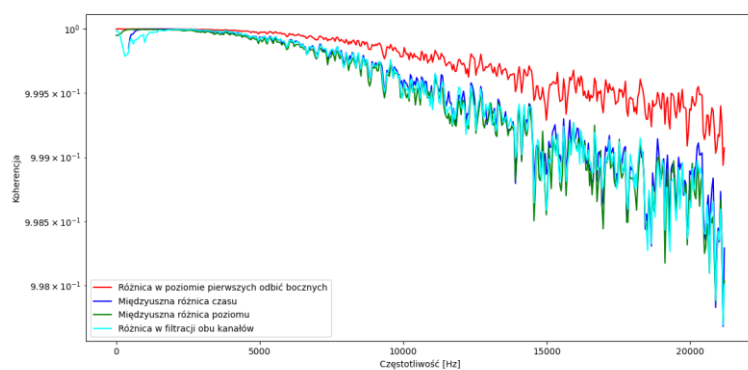
Rys. 3.7 Płaskość widmowa dla różnych zjawisk przestrzennych - sygnał mowy.



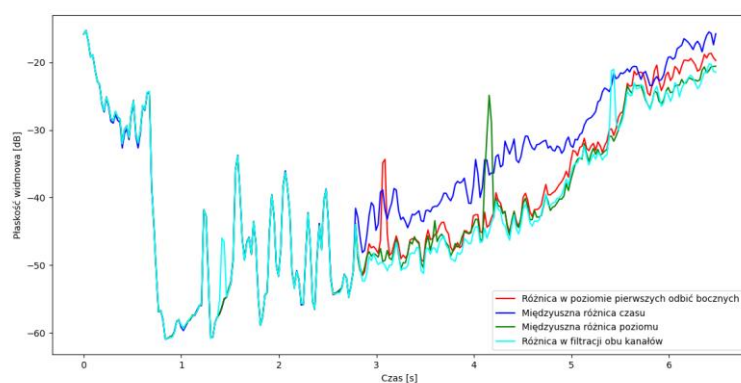
Rys. 3.10 Koherencja między kanałami dla różnych zjawisk przestrzennych - sygnał mowy



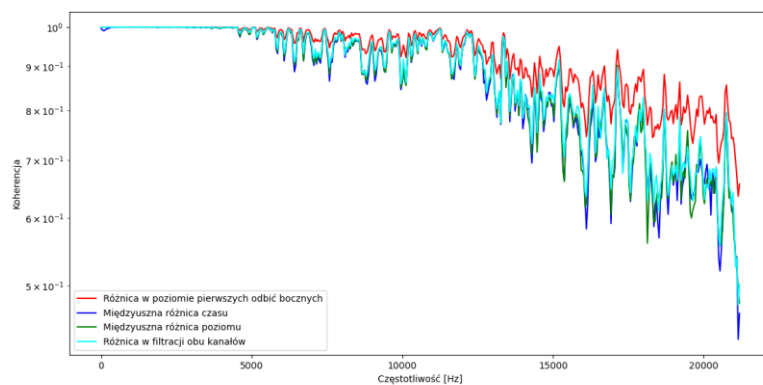
Rys. 3.7 Płaskość widmowa dla różnych zjawisk przestrzennych - sygnał MLS.



Rys. 3.10 Koherencja między kanałami dla różnych zjawisk przestrzennych - sygnał MLS.



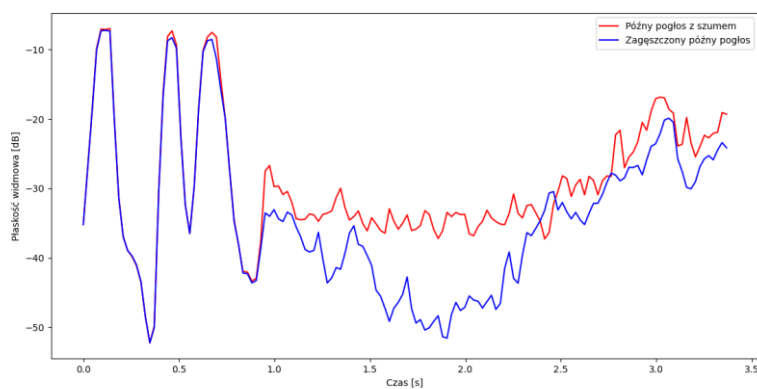
Rys. 3.7 Płaskość widmowa dla różnych zjawisk przestrzennych – gitara klasyczna.



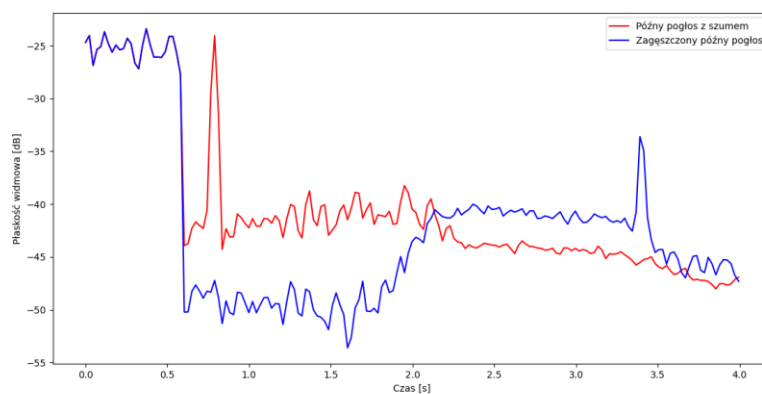
Rys. 3.10 Koherencja między kanałami dla różnych zjawisk przestrzennych - gitara klasyczna.

Tab. 3.3 Wartości parametru Pr dla symulacji różnych zjawisk przestrzennych

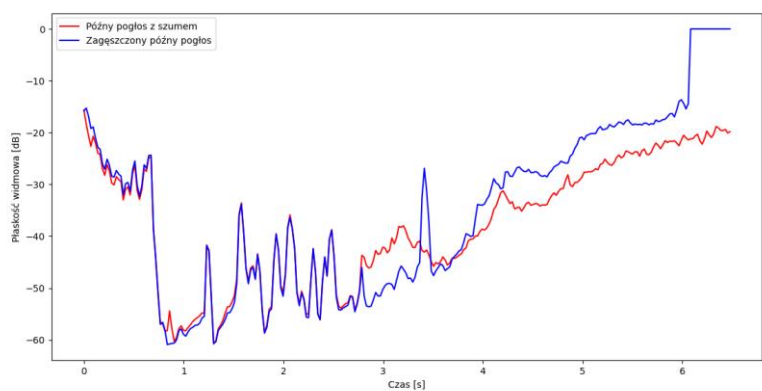
Pr	ITD	Różna filtracja obu kanałów	ILD	Różnica w amplitudzie pierwszych odbić bocznych
Sygnał mowy	0,88	1,0	0,83	0,68
MLS	0,51	0,48	0,47	1,0
Gitara klasyczna	0,67	0,69	0,66	1,0



Rys 3.7 Płaskość widmowa dla obu wersji implementacji późnego pogłosu - sygnał mowy



Rys 3.7 Płaskość widmowa dla obu wersji implementacji późnego pogłosu - sygnał MLS.



Rys 3.7 Płaskość widmowa dla obu wersji implementacji późnego pogłosu - gitara klasyczna

Tab. 3.4. Wartości parametru B dla pogłosów z różnymi wariantami generowania ogona pogłosowego

B	Późny pogłos z szumem	Późny pogłos zagęszczony odbiciami
Sygnał mowy	1,0	0,086
MLS	0,98	1,0
Gitara klasyczna	0,84	1,0

4. Implementacja docelowego pogłosu

Jak wspomniano we Wstępie (1) możliwości obliczeniowe w obecnych czasach znacznie ułatwiają badania nad realizacją sztucznego pogłosu w stosunku do warunków w XX wieku. Realizacja zjawisk akustycznych w domenie cyfrowej wymaga jednak konieczności 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 z DSP.

//tutaj będzie uzasadnienie wyboru narzędzia

Zarówno testowe implementacje elementów pogłosu (rozdział 3), jak i docelowej aplikacji zostały zrealizowane z wykorzystaniem `frameworku` 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 efektywne narzędzie do tworzenia aplikacji związanych z dźwiękiem.

//tutaj ciąg dalszy wstępu

Commented [AK35]: niestety nie znalazłem na razie adekwatnego tłumaczenia

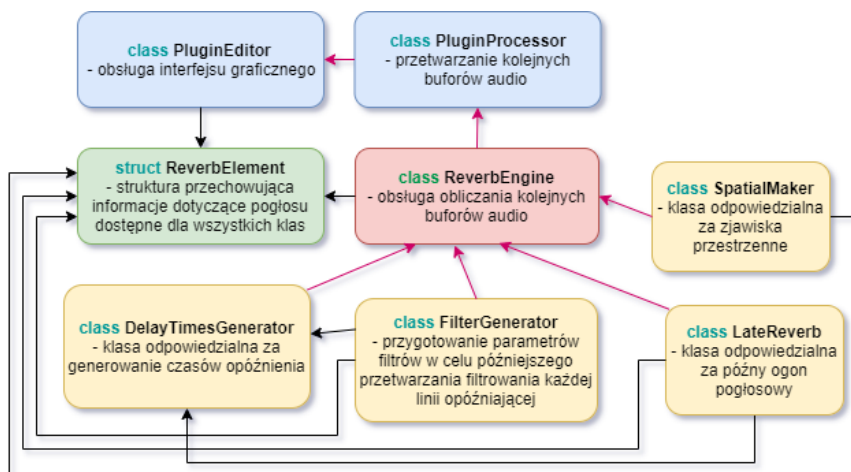
Wszystkie klasy oraz funkcje występujące w opisie implementacji w tym rozdziale są częścią frameworku JUCE.

4.1. Koncepcja realizacji

//tutaj zostanie opisana koncepcja realizacji

4.2. Opis architektury aplikacji

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



Rys. 4.1 Schemat przedstawiający hierarchię oraz opis klas w programie

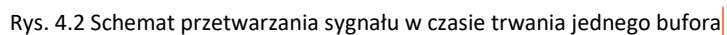
A więc obliczenia w klasach odpowiedzialnych za przygotowanie tablicy czasów opóźnienia linii opóźniających, działania filtrów i zjawisk przestrzennych oraz za realizację późnego pogłosu wykonywane są przed rozpoczęciem wykonywania obliczeń związanych z przetwarzaniem aktualnego bufora audio.

W przypadku zmiany przez użytkownika wielkości pogłosu (podrozdział 4.4.) nie jest konieczne ponowne losowanie wartości. Dzięki temu zmiana ta nie skutkuje również niepożądanymi artefaktami. W przypadku

//tutaj będzie szczegółowy opis klas

4.3. Połączenie wybranych rozwiązań w docelowy pogłos

//tutaj będzie dokładny opis obliczania aktualnego bufora audio oraz opis rysunku 4.2



A). Utworzenie sygnału sprzężonego z jego opóźnioną kopią.

W celu stworzenia bufora opóźniającego skorzystano z klasy *AudioBuffer*, realizującej wielokanałowy bufor audio złożony z liczb zmiennoprzecinkowych reprezentujących próbki sygnału [31]. W obrębie klasy istnieją metody pozwalające na dodawanie i kopiowanie

buforów, ustawianie wzmocnień oraz ustalanie wartości konkretnych próbek bufora, metody te okażą się przydatne w realizacji dalszych częściach algorytmu. Zastosowanie tej klasy pozwala również na wykorzystanie próbek sygnału w tej postaci przez inne klasy wbudowane frameworku.

Stworzony bufor opóźniający zapewnia dodatkowy nośnik informacji, niezbędny do wykonywania obliczeń wykorzystujących kopie pierwotnego sygnału - kopie te są głównym składnikiem pogłosu.

Aby uzyskać sygnał będący sumą sygnału pierwotnego z jego opóźnioną kopią, w pierwszej kolejności stworzono kopię sygnału pierwotnego o długości odpowiadającej długości bufora opóźnionego. Długość bufora opóźniającego została ustalona na ilość próbek odpowiadającą 4 sekundom (jest zależna od częstotliwości próbkowania ustalonej przez użytkownika) – maksymalne opóźnienie linii opóźniającej jest nie większe niż 4 sekundy. Tak więc jest to ilość próbek wystarczająca do tego, aby w każdym momencie w buforze opóźniającym znajdował się cały sygnał pogłosowy. Następnie za pomocą funkcji bibliotecznej *AudioBuffer::copyFrom()* wykonane zostaje wypełnienie bufora opóźniającego następującymi po sobie buforami przygotowanymi pierwotnie do wysłania na wyjście (bufory sygnału oryginalnego). Funkcja ta powoduje stworzenie kopii określonego fragmentu bufora. Długość fragmentu bufora, dokładne miejsce wyklejenia, numer kanału oraz źródło kopiowania określone są w argumentach ww. funkcji.

//bufor kołowy

Następnie z wypełnionego bufora opóźnionego, kopiowany jest bufor z pewnego miejsca wcześniej (miejsce to zależy od aktualnego czasu opóźnienia - podrozdział 4.3.4) i podmieniany z aktualnym buforem.

Aby uzyskać sumę bufora aktualnego z jego opóźnioną kopią, w następnym kroku do bufora opóźnionego dodawany jest aktualny bufor - dzięki wcześniejszym krokom zawiera on już swoją opóźnioną wersję. Tym sposobem podczas przetwarzania kolejnego bufora, przy kopiowaniu aktualnego bufora do bufora opóźnionego aktualny bufor zawiera już sumę sygnału i jego kopii. Mając w buforze wejściowym aktualny fragment sygnału i jego opóźnioną kopię, przetwarzanie kolejny bufor zapewnia istnienie sprzężenia zwrotnego.

B).

Pomiędzy operacjami

Stworzone linie opóźniające istnieją przez czas trwania (krótki bardzo).

4.3.2. Implementacja pierwszych odbić

4.3.3. Tłumienie odbić

4.3.4. Filtracja linii opóźniających

Zastosowano różną filtrację dla każdego odbicia. Sytuacja ta różni się od rzeczywistej....

4.3.5. Uprzestrzennienie pogłosu

4.3.6. Późny ogon pogłosowy

4.4. Część funkcjonalna aplikacji

5. Podsumowanie i wnioski

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

//podsumowanie będzie poszerzone o komentarz wyników, opis procedury itp.

Stworzone na potrzeby niniejszej pracy parametry i sposób ich zastosowania należy traktować jako próbę wskazania drogi, która docelowo ma prowadzić do pokrywania się wartości parametrów oceniających pogłos z subiektywnymi wrażeniami słuchaczy. Nasuwa się więc pomysł zweryfikowania, czy jakość pogłosu wyrażona otrzymanymi wartościami pokryje się z jakością wskazaną przez słuchaczy, lub zaproponowania nowych parametrów i ich weryfikacji również porównując otrzymane wartości ze wskazaniami subiektywnymi.

Należy pamiętać również o tym, iż duża ilość zmiennych przy badaniach (m. In. minimalna i maksymalna wartość czasu opóźnienia linii opóźniającej, zakresy częstotliwości odcięcia filtrów, różnica w amplitudzie odbić bocznych) powoduje, iż w wszystkie możliwe przypadki są praktycznie niemożliwe to zbadania (ze względu na ich liczbę), w zależności od kombinacji wyniki mogą się różnić. Nawet w przypadku wykonania badań przy parametrach zalecanych przez literaturę, w przypadku manipulacji którejs z zmiennych przez użytkownika podczas działania programu może wpłynąć na zdefiniowane parametry.

W przypadku znalezienia parametrów zbieżnych co do oceny ze wskazaniami subiektywnymi,

Commented [37]: to wymaga poszerzenia w odpowiednich miejscach – zarówno od strony opisu procedury, jak i komentarzy do wyników

... , a co za tym idzie będzie istniała możliwość oceny jakości pogłosu w sposób obiektywny z dużą szansą, iż wskazania parametrów pokrywają się z wrażeniami słuchowymi, bez konieczności wykorzystywania dużej grupy ludzi do badań.

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. Ocena subiektywna wynikała także z powtarzalności oceny przy testach odsłuchowych - przez powtarzalność opinii osób badanych na temat sygnałów uznawano metodę badawczą za wiarygodną.

Faktyczna sytuacja związana z propagacją fali w pomieszczeniu jest znacznie bardziej skomplikowana niż symulacja będąca wynikiem zaimplementowanego rozwiązania. Należy jednak pamiętać, iż przedstawione rozwiązanie miało na celu realizację pogłosu w postaci formie efektu, w odróżnieniu do uzyskania jak najbardziej realistycznego pogłosu. Istnieje także konieczność proponowania innych parametrów pozwalających na obiektywne porównywanie sygnałów pod kątem zjawisk towarzyszących propagacji dźwięku w pomieszczeniu. Metodą na ocenę zamodelowanych zjawisk akustycznych mogą być także stosowane dotąd badania na grupie słuchaczy, którzy będą oceniać sygnał pod względem zaproponowanych parametrów i kryteriów ocen.

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 Transmittancji Głowy (*Head Related Transfer Function*). A także innych zjawisk, dzięki którym możliwe jest lokalizowanie źródła poza płaszczyzną poziomą. Dokładna identyfikacja i parametryzacja tych zjawisk wymaga dalszych badań, pozwoli to na dokładniejsze odwzorowanie wrażenia przestrzenności w przypadku chęci ich stosowania w podobnych projektach. Ponadto, ze względu na fakt działania różnych zjawisk psychoakustycznych w różnych zakresach częstotliwości, wrażenie przestrzenności silnie zależy od rodzaju sygnału i zawartości jego widma.

Bibliografia

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

[2] M. Barron - The subjective effects of first reflections in concert halls – the need for lateral reflections, 1969

[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

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

[11] P. Kleczkowski – Percepcja Dźwięku

[12] A. H. Marshall – A note on the importance of room cross-section in concert halls

[] 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>
- [28] Davis, Gary (1987). *The sound reinforcement handbook* (2nd ed.). Milwaukee, WI: Hal Leonard. p. 259. [ISBN 9780881889000](#). Retrieved February 12, 2016.
- [29] <https://pages.iuh.edu/~signals/discreteconv2/index.html>
- [30] https://ccrma.stanford.edu/~jos/pasp/Schroeder_Allpass_Sections.html
- [31] Dokumentacja klasy *AudioBuffer* frameworku *Juce*:
<https://docs.juce.com/master/classAudioBuffer.html>