Spis treści:

1. Wstęp

1.1. Cel i zakres pracy

2. Teoretyczne podstawy 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. Ocena jakości symulacji pogłosu

3.1. Opis wybranych parametrów statystycznych do analizy sygnałów

3.2. Testy wybranych sposobów realizacji elementów pogłosu

4. Implementacja docelowego pogłosu

4.1. Koncepcja realizacji

4.2. Opis architektury 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. Synteza późnego ogona pogłosowego

4.4. Część funkcjonalna aplikacji

5. Podsumowanie i wnioski

Bibliografia

**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 miksu 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]. Teraźniejsze możliwości znacznie to ułatwiają. Przede wszystkim możliwość odwzorowania większości zjawisk w domenie cyfrowej, jak również znacznie ułatwiające badania większe o wiele możliwości obliczeniowe. W literaturze z tego okresu można zaobserwować obszerne opisy [2],[12] dotyczące przygotowania warunków, które były niezbędne to wykonania badań nad pogłosem. 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 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.

W związku z powyższym istnieje większa możliwość 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 jest projekt i implementacja cyfrowej symulacji zjawiska pogłosu w formacie VST poprzedzone oceną elementów składowych pogłosu zaproponowanymi w ramach pracy parametrami.

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 wynikiem połączenia poszczególnych elementów pogłosu na podstawie ich oceny zaproponowanymi parametrami. Przedstawiona zostanie również koncepcja realizacji docelowego pogłosu oraz algorytm prowadzący do uzyskania finalnego efektu.

**2.** **Teoretyczne podstawy zjawiska pogłosu**

*// Opisanie czym jest pogłos*

**2.1. Główne metody realizacji**

//tutaj będzie opis poniższych metod

**2.1.1. Pogłos splotowy**

**2.1.2. Rozwiązania elektromechaniczne**

**2.1.3. Cyfrowe symulacje pogłosu**

Od wczesnych lat rozwoju prac nad realizacją 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 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 [?].

*~~W pionierskiej pracy S chroedera z 1962 roku [4] zostają poruszone fundamentalne dziś zasady, na który opiera się większość systemów i architektur implementacyjnych sztucznych pogłosów.~~*

// tutaj o Feedback Delay Network

**2.2. Sposoby implementacji 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]. 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].

*~~Schroeder [4] - czasy opóźnienia w stosunku 1:1.5.~~*

~~//~~*~~Jeżeli ktoś twierdzi że słychać dzwonienie, to faktycznie - większość ludzi może słyszeć dzwonienie. 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 określeniem dzwonienia.~~*

**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].

Jednym z głównych celów testowania w ramach pracy różnych sposobów filtracji jest ocena gęstości widma późnego pogłosu. 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 brzmienia w zależności od rodzaju filtracji.

**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].

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 międzyuszna różnica fazy biorąca 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 niektóre parametry statystyczne – 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, MLS oraz nagranie instrumentu muzycznego w warunkach bezechowych. Następnie dokonana zostanie interpretacja wyników oraz wysunięte wnioski mające wpływ na opisaną w rozdziale 4 implementację docelowego pogłosu.

**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]:



gdzie:  
P - średnia wartość mocy sygnału,  
x(t) - przebieg czasowy sygnału,

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



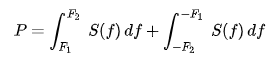
gdzie:  
ω - częstość kołowa,  
x(t) - przebieg czasowy sygnału,

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



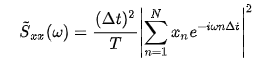
gdzie:  
f - częstotliwość,  
R(t) – przebieg czasowy autokorelacji sygnału.

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



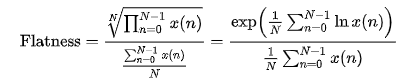
gdzie:  
P – moc sygnału,  
[F1,F2] - przedział częstotliwości   
S(f) - widmowa gęstość mocy sygnału

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



gdzie:  
Sxx(ω) **-** widmowa gęstość mocy,  
T = NΔt,  
n - 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]:



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

C). Koherencja sygnałów:

,

gdzie Gxy(f) jest wzajemną gęstością widmową [22] między sygnałami x i y, a Gxx(f) i Gyy(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:

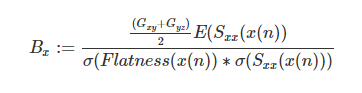
A). W celu zbadania wpływu doboru czasów opóźnienia na brzmienie pogłosu, dokonano symulacji zjawisk następujących pogłosów:

* Pogłos złożony z linii opóźniających o czasach opóźnienia jako liczby losowe,
* Pogłos złożony z linii opóźniających o czasach opóźnienia jako liczby losowe wzajemnie pierwsze (w nawiązaniu do [1], [4], [5]),
* Pogłos złożony z linii opóźniających o czasach opóźnienia zawierające 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]).

W celu obiektywnego porównania zaimplementowanych wersji, na potrzeby niniejszej pracy zdefiniowano parametr będący podstawą do oceny wpływu doboru wartości czasów opóźnienia linii opóźniających, jak również 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 parametr przedstawia się jako zależność:



gdzie:

*Gxy*- 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

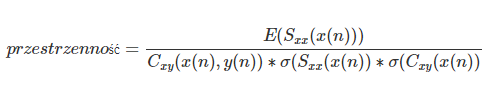
*σ-* średnie odchylenie standardowe

*E* - średnia arytmetyczna

C). Psychoakustyczne zjawiska przestrzenne

* Różnica w poziomie pierwszych odbić bocznych w celu potwierdzenia zaobserwowanej w tezy o większej przestrzenności przy wyższym poziomie tych odbić (w nawiązaniu do[2]),
* Zastosowanie różnicy czasu opóźnienia między kanałami dla linii opóźniających (w nawiązaniu do [11]),
* Zastosowanie różnicy amplitudy pomiędzy kanałami dla linii opóźniających (w nawiązaniu do [11]),
* Zastosowanie różnicy pomiędzy filtracją odbić dla 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 jakoś pogłosu. W przypadku zjawisk przestrzennych cechy te to: wysoka gęstość widma oraz stopień niekoherencji sygnałów z obu kanałów. Dlatego też, parametr ten będzie wyznacznikiem przestrzenności pogłosu. Parametr *przestrzenność* zdefiniowano jako:



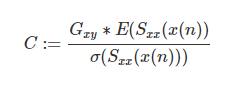
gdzie:

*σ* - średnie odchylenie standardowe  
Sxx - widmowa gęstość mocy   
Cxy - koherencja sygnałów   
E- średnia arytmetyczna

D). Późny ogon pogłosowy:

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

Założenia opisano zdefiniowanym na potrzeby pracy parametrem *<nazwa parametru>*. Zdefiniowanie parametru jest odpowiedzią na powtarzające się w literaturze podkreślanie istotności posiadania wysokiej gęstości widma późnego ogona pogłosowego. Dlatego też parametr ten będzie wyznacznikiem jakości późnego pogłosu. Parametr <nazwa parametru> zdefiniowano jako:



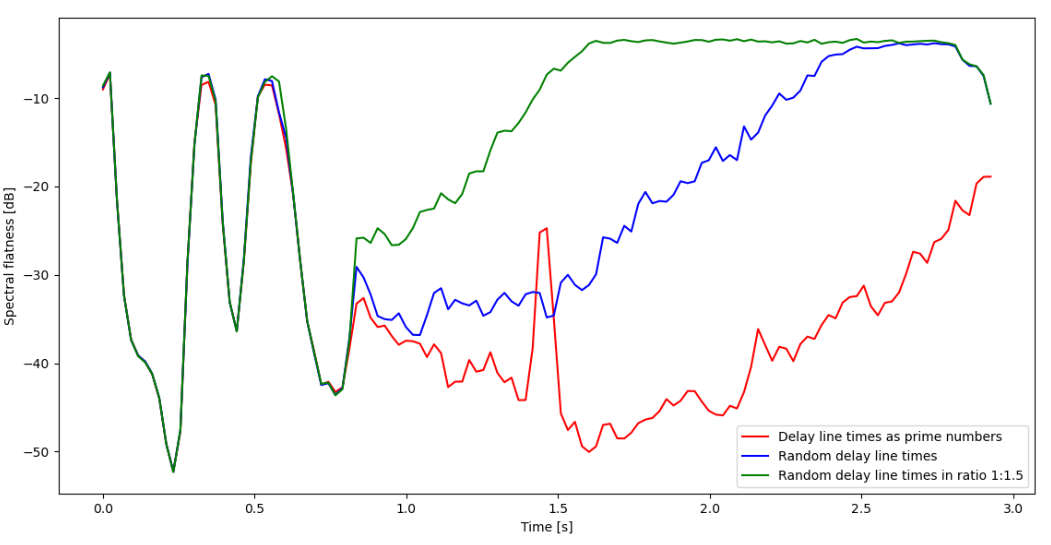
Gdzie:

Gxy - 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 z pogłosem późnym pogłosem zagęszczonym odbiciami  
y(n) - sygnał z późnym pogłosem z szumem   
E - średnia arytmetyczna,  
*σ* - średnie odchylenie standardowe  
Sxx - widmowa gęstość mocy

~~Fakty opisane w podrozdziale 2.2 dotyczące doboru czasów opóźnienia linii opóźniających prowadzą do wniosku, iż w celu uniknięcia negatywnych, wymienionych tam zjawisk, niezbędne jest ograniczenie zbioru liczb reprezentujących czasy opóźnienia do liczb wzajemnie pierwszych.~~

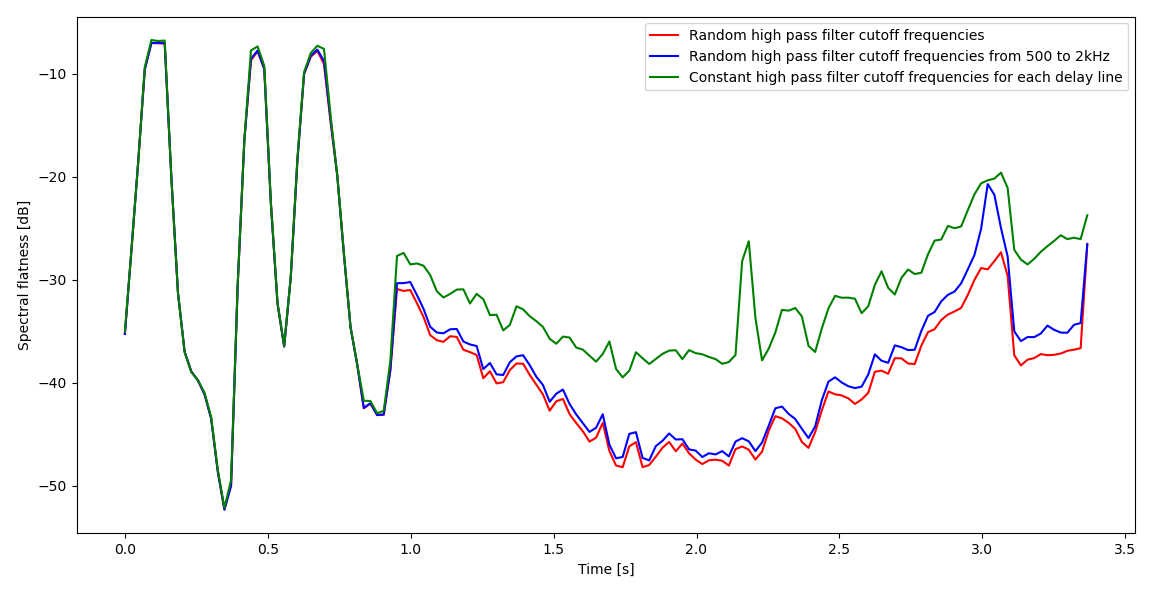
Tab. 3.1 Tabela zawierająca wartości parametru B dla wszystkich wersji

|  |  |  |  |
| --- | --- | --- | --- |
|  | Bx | By | Bz |
|  |  |  |  |



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

*~~Odwołując się do trudności w obiektywnej ocenie jakości pogłosu opisanej w podrozdziale 3.2, w celu oceny wpływu filtracji na jakość pogłosu dokonano próby obiektywnej oceny tego zjawiska.~~*

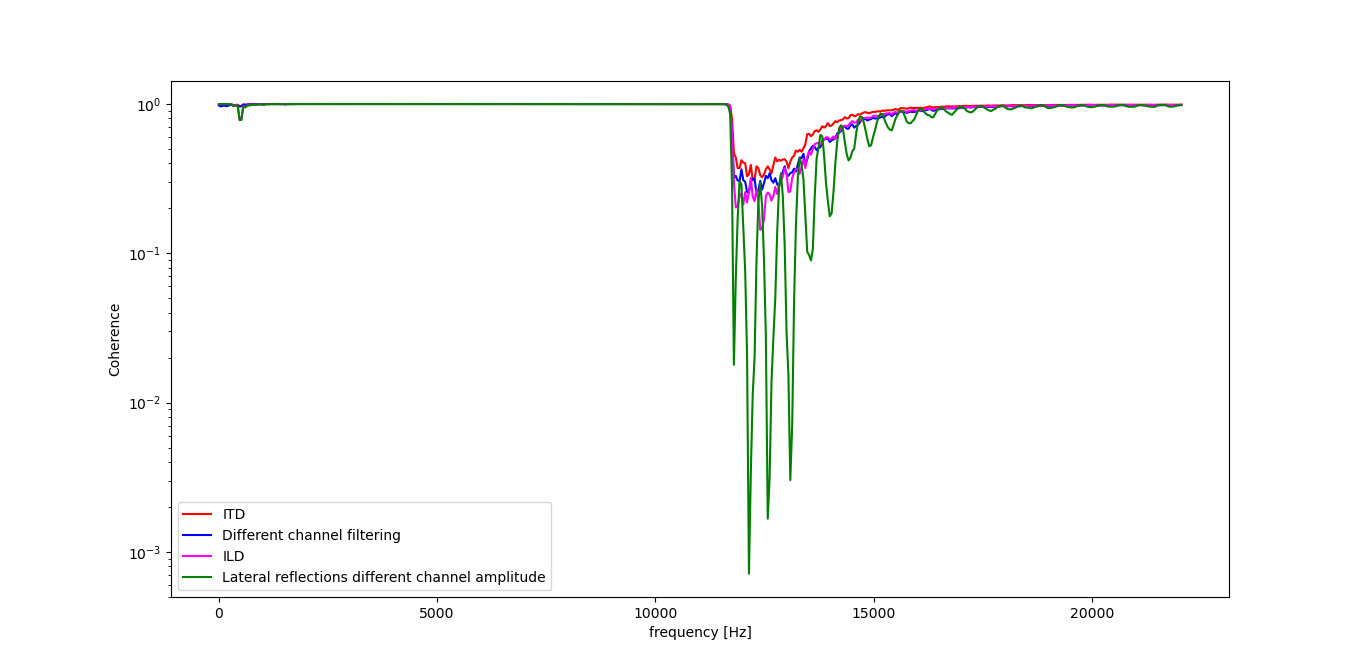


Rys (..) Płaskość widma symulacji różnych wariantów filtracji

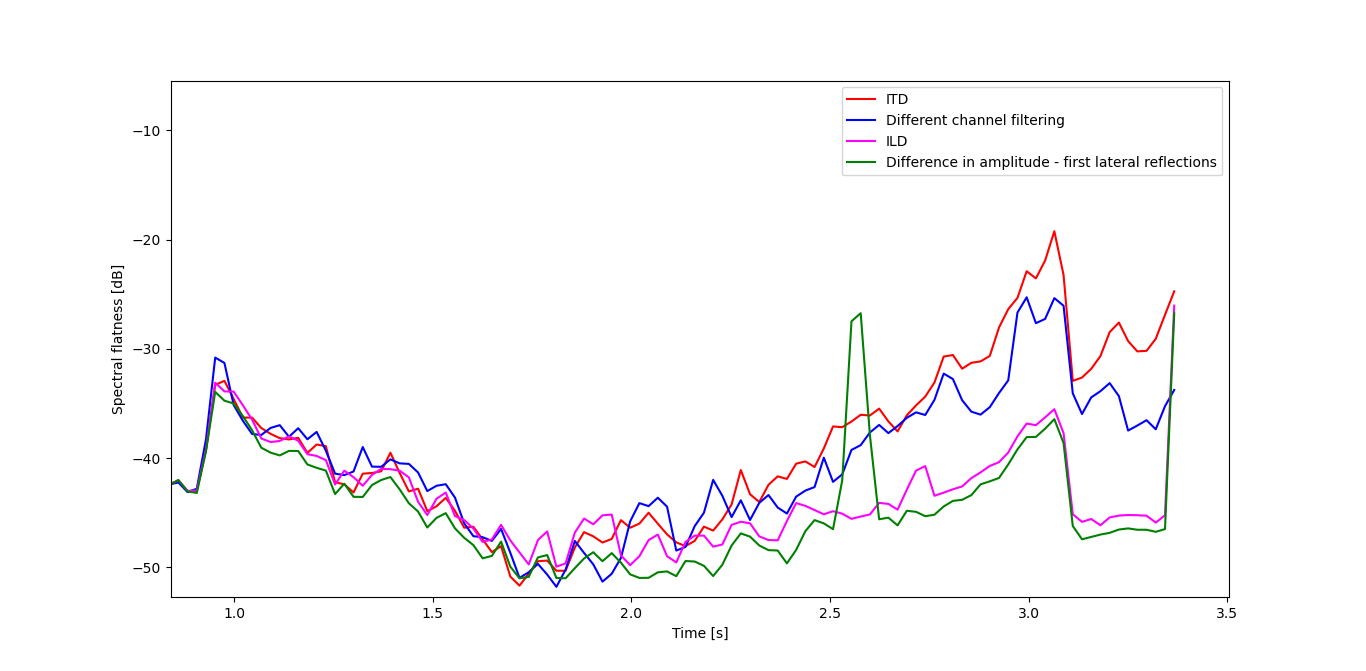
Porównanie sygnałów pogłosowych z różnymi wariantami filtracji

|  |  |  |  |
| --- | --- | --- | --- |
|  | xf(n) | yf(n) | zf(n) |
| B |  |  |  |

**Wykorzystanie zjawisk psychoakustycznych/przestrzenność pogłosu**



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

Rys (..) Płaskość widmowa dla różnych zjawisk przestrzennych

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 (n) 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) |  |  |  |  |

Ze względu na fakt powstawania filtra grzebieniowego przy opóźnieniu do 40ms, zastosowano właśnie taką wartość jako maksymalną różnicę czasu opóźnienia między kanałami. Różnice te są różne dla każdego odbicia – skutkuje to uzyskaniem największej gęstości widmowej, dostajemy bowiem różny filtr grzebieniowy dla każdej linii opóźniającej, dzięki czemu jest mała szansa interferencje tych filtrów. Ma to potwierdzenie w parametrze *przestrzenność - (jeśli się potwierdzi). Aczkolwiek na pewno na wykresie widać, że dla dużego fragmentu ITD jest najwyżej.*

Parametr *przestrzenność*

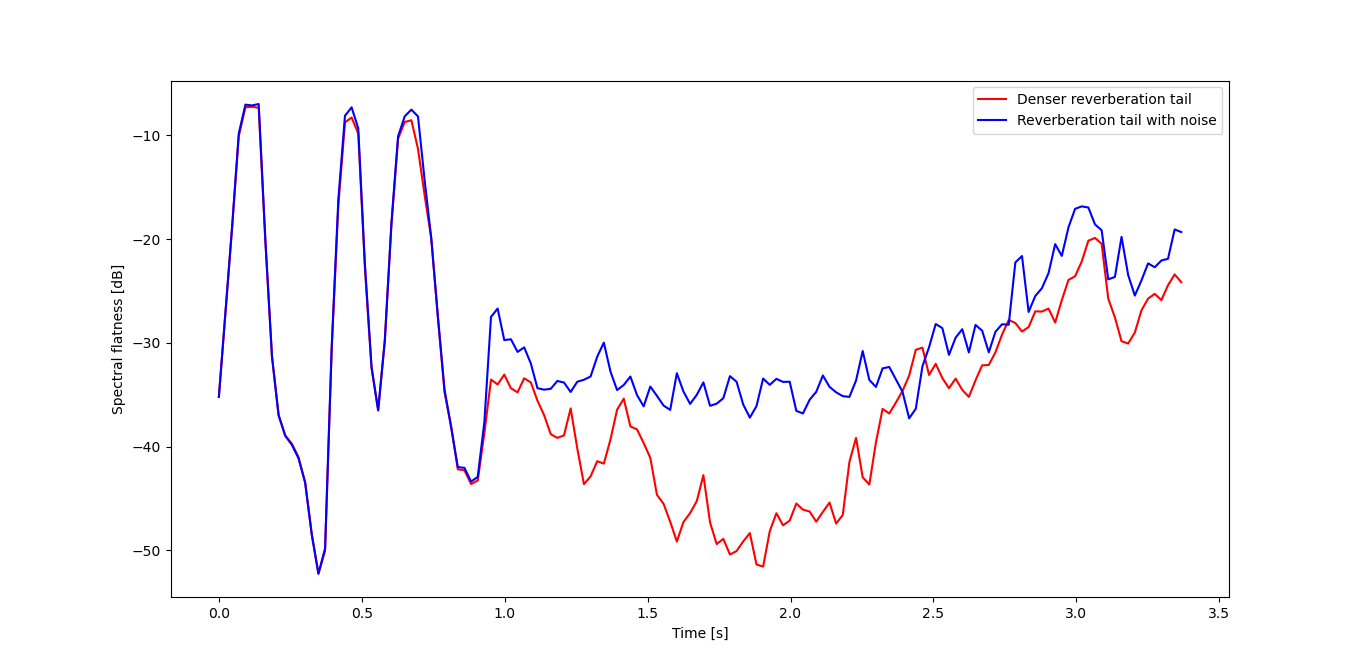
~~W celu kompensacji wrażenia znajdowania się źródła dźwięku po bardziej po lewej lub prawej stronie, wartości różnicy parametrów (dotyczy to pierwszego odbicia): czasów opóźnienia między kanałami, poziomu oraz filtracji dostosowano tak, aby odnosiło się wrażenie umiejscowienia źródła naprzeciwko słuchacza.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

**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]

Z celu oceny jakości późnego pogłosu zaimplementowano poniższe rozwiązania:

* 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 opóźniających



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

|  |  |  |
| --- | --- | --- |
|  | Późny pogłos z szumem | Późny pogłos zagęszczony odbiciami |
| 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].

**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 DSP.

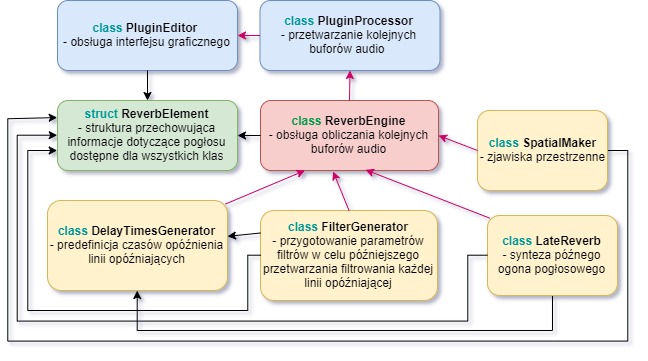
Zarówno testowe implementacje elementów pogłosu (rozdział 3), jak i docelowej aplikacji zostały zrealizowane z wykorzystaniem frameworku JUCE.   
JUCE to wieloplatformowy otwarto-źródłowy framework pozwalający na realizację aplikacji m. in. w technologii VST w języku C++ [16]. Narzędzie to pozwala także na realizację interfejsu użytkownika. Połączenie tych funkcji daje wygodne narzędzie do tworzenia aplikacji związanych z dźwiękiem.

**4.1. Koncepcja realizacji**

//Tutaj zostanie opisane 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 w języku C++, z których każda odpowiada za inną część obliczeń realizujących przetworzenie sygnału w celu uzyskania sygnału wynikowego.

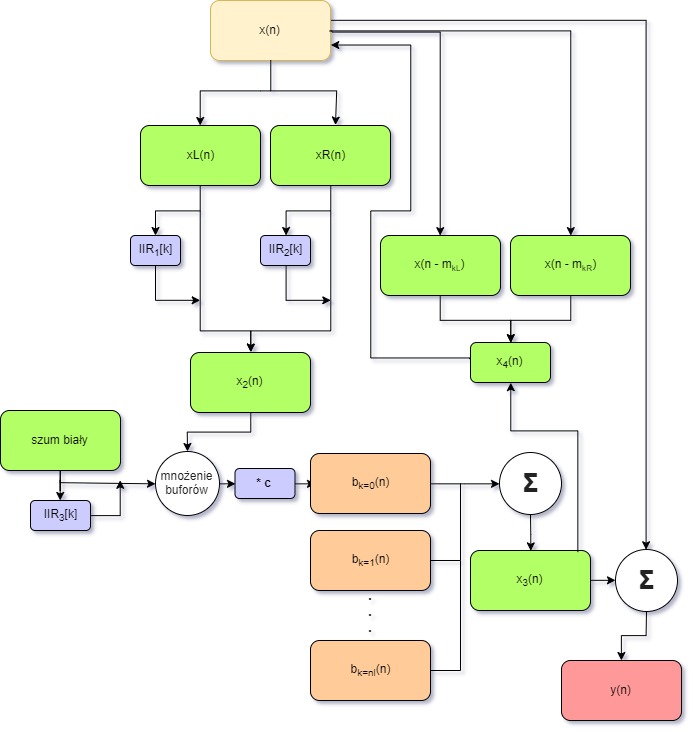


Rys. 4.2 Schemat przedstawiający hierarchię oraz opis klas w programie

**4.3. Połączenie wybranych rozwiązań w docelowy pogłos**

**4.3.1. Linie opóźniające i sposób obliczania aktualnego bufora audio**

~~Właściwie mamy do czynienia z~~



Rys. 4.1 Schemat przetwarzania bufora audio w aplikacji

gdzie:

x(n) - aktualny bufor audio  
xL(n) - lewy kanał aktualnego bufora audio  
xR(n) - prawy kanał aktualnego bufora audio  
k – iteracja   
IIR1 - tablica filtrów dla lewego kanału  
IIR2 - tablica filtrów dla prawego kanału  
IIR3 - tablica filtrów dla buforów szumowych  
mkL -   
mkR -   
x2(n) - dwukanałowy bufor audio po filtracji  
x3(n) -   
x4(n) -   
y(n) - bufor wyjściowy  
c - współczynnik tłumienia

**4.3.2. Implementacja pierwszych odbić**

Zgodnie z [2] podjęto działania mające na celu symulację pierwszych odbić. Symulacja polegała na dodaniu kopii sygnału o parametrach o zgodnie z wartościami zaproponowanymi w podrozdziale ~~3.2~~. 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), filtracji dolnoprzepustowej imitującej tłumienie przez powietrze oraz pochłanianie przez materiał znajdujący się na teoretycznej ścianie.

**4.3.3. Tłumienie odbić**

**4.3.4. Filtracja linii opóźniających**

**4.3.5. Uprzestrzennienie pogłosu**

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

**4.3.6. Implementacja ogona pogłosowego**

**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.

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 Transmitancji 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 Convetion 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/~trsmyth/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.](https://en.wikipedia.org/wiki/Spectral_density" \l ":~:text=Energy spectral density describes how the energy of,transients—that is%2C pulse-like signals—having a finite total energy.)

[20] [https://pl.wikipedia.org/wiki/Widmowa\_g%C4%99sto%C5%9B%C4%87\_mocy](https://pl.wikipedia.org/wiki/Widmowa_gęstość_mocy)

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

[22] [https://en.wikipedia.org/wiki/Spectral\_density#Cross-spectral\_density](https://en.wikipedia.org/wiki/Spectral_density" \l "Cross-spectral_density)

[24] <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](https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.welch.html" \l "scipy.signal.welch)