Polish
Audio Engineering Society

XIX Międzynarodowe Sympozjum Nowości w Technice Audio i Wideo NTAV2022

Wrocław, 13-15 października 2022

PTA

Materiały konferencyjne

ORGANIZATORZY ▪ ORGANIZERS

Polska Sekcja Audio Engineering Society
Polskie Towarzystwo Akustyczne oddział we Wrocławiu
Katedra Akustyki, Multimediów i Przetwarzania Sygnałów
Politechnika Wrocławska

KOMITET NAUKOWY ▪ SCIENTIFIC COMMITTEE

Przewodniczący: Krzysztof Opieliński
Andrzej Brzoska
Andrzej Czyżewski
Andrzej Dobrucki
Tadeusz Kamisiński
Piotr Kleczkowski
Bożena Kostek
Mirosław Meissner
Andrzej Miśkiewicz
Janusz Piechowicz
Anna Preis
Tomira Rogala
Aleksander Sęk
Ewa Skrodzka
Paweł Strumiłło
Jerzy Wiciak
Wiesław Woszczyk
Jan Żera

KOMITET ORGANIZACYJNY ▪ ORGANIZING COMMITTEE

Przewodniczący:	Przemysław Plaskota *(edycja)*
Sekretarz:	Michał Łuczyński *(redakcja)*
Skarbnik:	Paweł Dziechciński
Członkowie:	Bartłomiej Kruk
	Maurycy J. Kin
	Piotr Kozłowski
	Jędrzej Borowski
	Agnieszka Wielgus
	Romuald Bolejko
	Maciej Walczyński
	Zbigniew Świetach
	Agnieszka Paula Pietrzak

Wydawnictwo InfoArt
ISBN 978-83-966051-0-8

WSPIERAJĄ NAS

Dolby

softserve

ZYLIA

NFM | NARODOWE FORUM MUZYKI

SPIS TREŚCI

PROJEKTOWANIE ORAZ TWORZENIE SYSTEMÓW ZARZĄDZAJĄCYCH W NIEKONWENCJONALNYCH INSTALACJACH DŹWIĘKU PRZESTRZENNEGO

Jan SKORUPA[1]
Maciej GŁOWIAK[2]
[1]Instytut Chemii Bioorganicznej PAN – Poznańskie Centrum Superkomputerowo-Sieciowe
jskorupa@man.poznan.pl
[2]Instytut Chemii Bioorganicznej PAN – Poznańskie Centrum Superkomputerowo-Sieciowe
mac@man.poznan.pl

1. WPROWADZENIE

Dźwięk przestrzenny to zjawisko obecne w przemyśle rozrywkowym od początku lat 40 XX w. Oczywistym jest, że pierwsze skojarzenia z tą dziedziną wiążą się z kinematografią jako narzędziem pogłębiającym doznania wizyjne. Oprócz środowiska filmowego wykorzystanie systemów dźwięku wielokanałowego było silnie eksplorowane przez kompozytorów i inżynierów związanych z ówczesną muzyką elektroniczną. W swoich poszukiwaniach szukali nowych narzędzi, adekwatnych do sztuki dźwiękowej, w której źródłem dźwięku przestaje być instrument a głośnik odtwarzający nagrania, przygotowywane pierwotnie na płytach szelakowych, gramofonowych czy taśmach magnetycznych. Już we wczesnych latach 50 XX w. Pierre Shaeffer wraz ze swoim asystentem Pierre Henrym we współpracy z inżynierami z Radiodiffusion-Télévision Française opracowali system wielokanałowego odsłuchu czterech niezależnych kanałów audio do 5 niezależnych głośników ustawionych wokół publiczności. Praktyki tego typu były chętnie podejmowane w kolejnych latach przez takich twórców jak IIanis Xenakis czy Karlheinz Stockhausen. Xenakis kompozytor oraz architekt w roku 1958 wraz z Corbusierem zaprojektowali pawilon multimedialny Philipsa zaprezentowany podczas Expo w Brukseli. W pawilonie zbudowana została instalacja dźwięku przestrzennego składająca się z 350 głośników oraz oddzielnej instalacji świetlnej zsynchronizowanej z warstwą dźwiękową. Podczas demonstracji w pawilonie odtwarzana była kompozycja Edgara Varèse Poème Électronique oraz Concert Ph Xenakisa. Stockhausen natomiast w roku 1970 był odpowiedzialny za stworzenie kompozycji dźwiękowych w pawilonie zaprezentowanym podczas Expo w Osace. Na potrzeby utworów stworzono sferyczną instalację składającą się z 50 głośników. Innym przykładem wartym wymienienia jest działalność Franccois Byle' a oraz francuskiego GRM (fr. *Groupe de Recherches Musicales*), który w latach 70 XX w. pracował nad *Acousmonium.* Idea tej instalacji opierała się o stworzenie orkiestry głośnikowej, były to wielogłośnikowe systemy składające się z kilku do kilkudziesięciu głośników, gdzie każdy był inny – charakteryzował się różnym pasmem przeniesienia, głośnością czy kierunkowością. Kompozytor w tym układzie przygotowywał utwór elektroniczny, który podczas koncertu był uprzestrzenniany „na żywo" za pośrednictwem miksera audio poprzez sterowanie wzmocnieniem poszczególnych głośników. Wskazani eksperymentatorzy w trakcie swoich poszukiwań opracowywali rozwiązania znacząco wykraczające poza ówcześnie znane systemy dźwięku przestrzennego, proponowali rozwiązania nowatorskie budzące zainteresowanie zarówno z punktu widzenia kompozytora muzyki elektronicznej jak i inżyniera dźwięku. Praktyki tego typu obecne są do dnia

dzisiejszego a z racji na znacznie większe możliwości techniczne możliwe jest tworzenie systemów dających pełną swobodę dystrybucji kanałów oraz miksowania ich w dowolnych konstelacjach głośnikowych. W ramach referatu omówione zostaną dwie autorskie instalacje wielogłośnikowe dedykowane prezentacji kompozycji przestrzennych: Akuzmonium prezentowanie po raz pierwszy podczas festiwalu Fazma oraz system wielogłośnikowy w „Pokoju do Słuchania" stworzony do autorskiej instalacji kompozytorki Hani Rani oraz studia Architektonicznego Zmir. Omówione zostaną szczegółowo zagadnienia techniczne związane z konstruowaniem instalacji oraz dedykowane systemy dystrybucji sygnałów elektroakustycznych.

2. AKUZMONIUM

Rys. 1. zdjęcie przedstawiające instalację Akuzmonium podczas koncertu

Instalacja (Rys. 1.) w swojej formie odnosi się do praktyk francuskiego kompozytora franccois Byle' a. W ramach festiwalu muzyki elektronicznej pt: Fazma skonstruowana została współczesna wersja historycznego systemu pozwalającego na dystrybucję nieograniczonej ilości kanałów audio do 42 niezależnych głośników. Wielokanałowy system opierał się o 7 rzędów, gdzie w każdym ustawiono po 6 głośników. Dla instalacji za pośrednictwem MaxMSP zaprojektowany został specjalny mikser dźwięku pozwalający na kontrolę systemu wielokanałowego. Na Akuzmonium powstały cztery dedykowane kompozycje elektroniczne od początku stworzone przy wykorzystaniu pełnej instalacji.

3. „POKÓJ DO SŁUCHANIA”

Z inicjatywy kompozytorki Hani Rani oraz Studia Architektonicznego Zmir w Warszawskim Pawilonie Architektonicznym Zodiak powstała instalacja artystyczna podejmująca tematykę relacji architektury oraz muzyki. W ramach przedsięwzięcia zaprojektowana została specjalna drewniana konstrukcja o powierzchni ponad 60 m^2 oraz dedykowana tej przestrzeni ścieżka dźwiękowa. Stworzona forma składała się z trzech oddzielnych pomieszczeń, gdzie w każdym wybrzmiewały różne partie kompozycji. Częścią stworzonej konstrukcji była instalacja dźwiękowa specjalnie zaprojektowana na potrzeby wystawy (Rys. 2.). Projekt systemu dźwiękowego składał się z trzech różnych konstelacji głośnikowych dedykowanych poszczególnym pomieszczeniom. W sumie przestrzeń została nagłośniona 25 monitorami bliskiego pola, umieszczonymi w ścianach oraz suficie całej konstrukcji. Oprócz projektu, instalacji elektroakustycznej w celu uprzestrzennienia stworzonej kompozycji, stworzono specjalną aplikację pozwalającą na dowolne umieszczanie oraz poruszanie źródłami dźwiękowymi w dedykowanej przestrzeni jak i kontrolę rezonansów czy głośności całej kompozycji. Podczas tworzenia systemu przetestowano wiele różnych algorytmów panoramowania jak i zewnętrznych bibliotek dedykowanych konstruowaniu wielokanałowych systemów dźwiękowych. Stworzony system pozwalał na kompleksowe uprzestrzennie oraz miksowanie kompozycji muzycznej w powstałej instalacji.

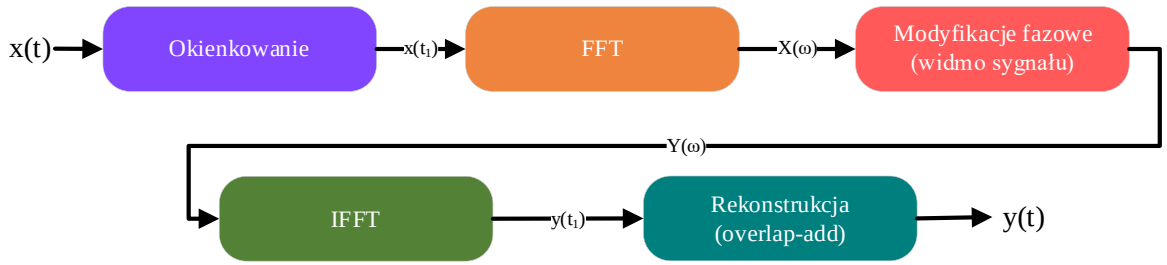
Rys. 2. zdjęcie przedstawiające instalację dźwiękową w jednym z pomieszczeń instalacji „Pokój do Słuchania”

Za projekt obydwu instalacji odpowiedzialne było laboratorium Art&Science Poznańskiego Centrum Superkomputerowo Sieciowego.

ALGORYTM UPRZESTRZENNIAJĄCY SYGNAŁY DŹWIĘKOWE BAZUJĄCY NA PRZESUNIĘCIACH FAZOWYCH

Kamil ZIMNY[1]
Teresa MAKUCH[2]
[1]AGH Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie,
al. A. Mickiewicza 30,30-059 Kraków
kzimny@agh.edu.pl
[2]AGH Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie,
al. A. Mickiewicza 30, 30-059 Kraków
tmakuch@agh.edu.pl

W obecnych czasach istnieje wiele metod wpływania na przestrzenność sygnałów dźwiękowych. Zazwyczaj opierają się one na modelowaniu zjawisk fizycznych, które naturalnie wywołują to wrażenie, tj. nakładanie pogłosu, generowanie powtórek sygnału, czy wprowadzanie opóźnień. Praktyczne implementacje (algorytmy) tego typu oferują wiele parametrów umożliwiających precyzyjną kontrolę symulowanych procesów. Jednakże, problematyczne okazuje się ich przełożenie na subiektywne wrażenie przestrzenności sygnału dźwiękowego.
Dotychczas przeprowadzone badania pokazują, że percepcja przestrzenności sygnału może być znacznie bardziej skomplikowana i realizowana w dużej mierze przez wyższe piętra układu słuchowego. Tłumaczy to trudność w przełożeniu parametrów odpowiadających za fizyczne właściwości modelowanego procesu na subiektywne wrażenia. Prace badawcze, w tym dr. D. Griesingera, wykazują istotny wpływ koherencji (spójności) fazowej składowych harmonicznych na subiektywne wrażenie przestrzenności [2]. Jednak faza składowych sygnału może też być odpowiedzialna za wrażenia związane z barwą dźwięku, zakolorowaniem i klarownością wrażenia wysokości [3].
W związku z powyższym autorzy podjęli próbę stworzenia algorytmu uprzestrzenniającego sygnał dźwiękowy bazującego wyłącznie na modyfikacjach (przesunięciach) fazy składowych sygnału. Schemat blokowy algorytmu przedstawiono na rys. 1.



Rys. 1. Schemat blokowy wraz z przepływem sygnału w algorytmie uprzestrzenniającym

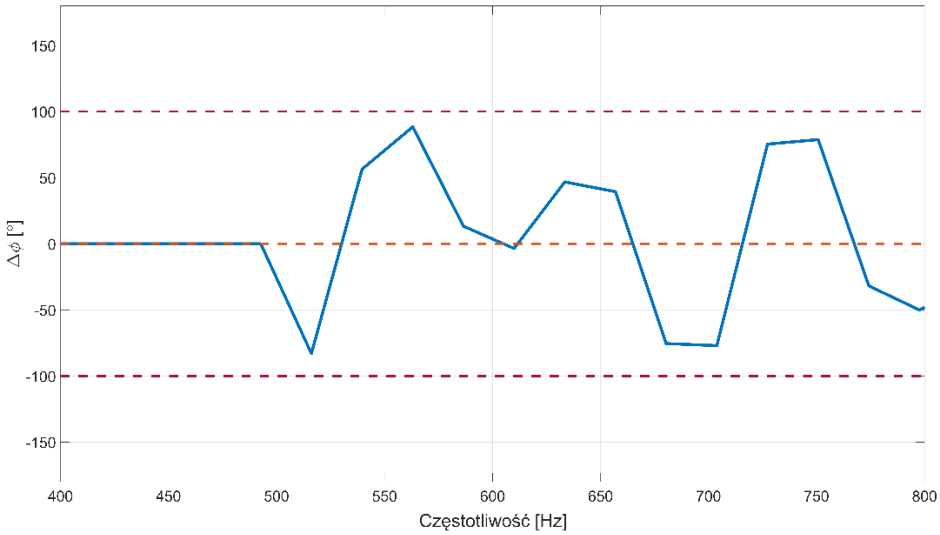
Zgodnie z powyższym schematem, sygnał w pierwszej kolejności poddawany jest procesowi okienkowania, z parametrami (okno Hamminga, odpowiednia długość i parametry funkcji okna) dobranymi tak, aby umożliwić późniejszą rekonstrukcję za pomocą metody *overlap-add* [1]. Następnie dla każdego okna sygnału wykonywana jest Szybka Transformata Fouriera (FFT – ang. *Fast Fourier Transform*) o liczbie punktów równej lub wyższej od rozmiaru okna. Dalej, operując w dziedzinie widma sygnału, wykonywane są modyfikacje fazowe dla harmonicznych o rozpoznanych uprzednio częstotliwościach. Następnie sygnał zostaje przeniesiony ponownie

do dziedziny czasu z wykorzystaniem Odwrotnej Szybkiej Transformaty Fouriera (IFFT – ang. *Inverse Fast Fourier Transform*), po czym finalnie zostaje zrekonstruowany za pomocą metody *overlap-add*, zachowując swoją pierwotną długość. Algorytm został zaimplementowany w środowisku MATLAB.

Zasadnicze działanie algorytmu polega na wprowadzaniu modyfikacji fazowych w przetwarzanym sygnale. W celu uniknięcia przy tym słyszalnych zniekształceń, niezbędnie było wprowadzenie dodatkowych założeń, m.in określenie maksymalnego przesunięcia fazowego, dobór odpowiedniej długości okna, czy zapobieganie zbyt dużym skokom fazy między kolejnym oknami. W opracowanym algorytmie przesunięcia fazowe są realizowane na kilka sposobów:

- losowo, z uwzględnieniem innych rozkładów niż normalny,
- zgodnie z przebiegami funkcji matematycznych,
- w całym sygnale lub tylko dla poszczególnych harmonicznych lub pasm częstotliwości.

Przykład działania algorytmu z zastosowaniem losowych przesunięć fazy przedstawiono na rys. 2.



Rys. 2. Przykład wprowadzonych losowo przesunięć fazowych do sygnału przetwarzanego przez algorytm

Bazując na modyfikacjach fazowych, opisywany algorytm prezentuje nowe podejście do kontroli i modyfikacji przestrzenności sygnałów dźwiękowych. Zaproponowany sposób przetwarzania w niewielkim stopniu ingeruje w przetwarzany sygnał, w odróżnieniu od np. dodawania pogłosu. Dzięki wykorzystaniu środowiska MATLAB, docelowo planowane jest przygotowanie algorytmu w postaci wtyczki programowej VST, a więc wersji algorytmu działającej w czasie rzeczywistym. W ten sposób działanie algorytmu będzie mogło być testowane przez szersze grono odbiorców, w tym realizatorów dźwięku i producentów materiałów muzycznych.

LITERATURA

[1] ALLEN J.B., RABINER L.R., *A Unified Approach to Short-Time Fourier Analysis and Synthesis*, Proceedings of the IEEE vol. 65, no. 11, 1977

[2] GRIESINGER D., *Phase Coherence as a Measure of Acoustic Quality. Part One: the Neural Mechanism*, Proceedings of 20th International Congress of Acoustics (ICA), 2010

[3] LAITINEN M-V., DISCH S., PULKKI V., *Sensitivity of Human Hearing to Changes in Phase Spectrum*, AES paper, 2013

PERCEPTUAL EVALUATION OF FIRST AND THIRD ORDER AMBISONIC RECORDINGS

Agnieszka Paula PIETRZAK, Amaia SAGASTI[1]
Ricardo SAN MARTÍN, Rubén EGUINOA [2]
[1]Politechnika Warszawska,, ul. Nowowiejska 15/19, 00-665 Warszawa
A.Pietrzak@ire.pw.edu.pl
[2] Public University of Navarre

First-order or higher-order ambisonic microphones are used for ambisonic recordings. The higher the order of ambisonics, the more accurately the acoustic field can be recorded, giving more information about the location of the sound sources. However, in the case of subsequent binaural reproduction of the ambisonic sound, errors and distortions may occur, which make the listener's ability to locate the sound source in the recordings made with a higher-order microphone not significantly higher than in the case of a first-order microphone. Higher resolution of spatial audio information can also be achieved by using upmixing techniques. In this work, auditory tests were carried out on the ability to locate the sound source in recordings made using first and third order ambisonic microphones. Perceptual evaluation of the quality of recordings upmixed from first to third order ambisonic recordings was also carried out.

AUTOMATYCZNA KLASYFIKACJA MOWY PATOLOGICZNEJ

Martyna WŁOSZCZYŃSKA[1], Bożena KOSTEK[2]

[1]Politechnika Gdańska, Wydział ETI, G. Narutowicza 11/12, 80-233 Gdańsk
mkwloszczynska@gmail.com
[2]Politechnika Gdańska, Wydział ETI, Laboratorium Akustyki Fonicznej, G. Narutowicza 11/12, 80-233 Gdańsk
bozena.kostek@pg.edu.pl

Zaburzenia mowy mają istotny wpływ na komunikację osób z otoczeniem. Trudności w wypowiadaniu pojedynczych głosek, pełnych słów i zdań zaburzają proces efektywnego porozumiewania się, co znacznie utrudnia adaptację w środowisku społecznym oraz może mieć wpływ na funkcje poznawcze zarówno w przypadku dzieci, jak i starszych osób. Głównym powodem powstawania patologii głosu/wymowy jest nieprawidłowe funkcjonowanie aparatu mowy. Wśród przyczyn wymienia się m.in. budowę anatomiczną aparatu głosowego, problemy ze słuchem, a także zaburzenia rozwoju czy schorzenia neurologiczne.

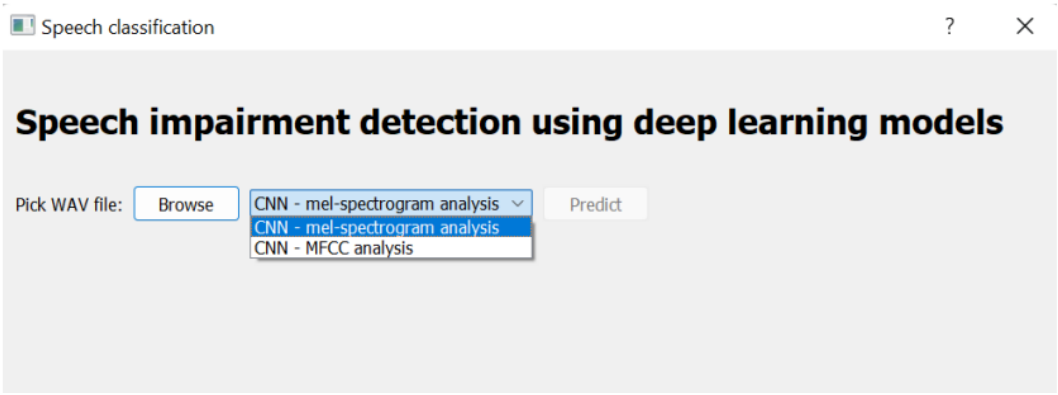
Metody automatycznej detekcji lub identyfikacji patologii w mowie są obecne w literaturze od kilku dekad. Jednak dopiero pojawienie się algorytmów uczenia głębokiego spowodowało realną możliwość zastosowania tego typu metod w diagnostyce medycznej w celu wsparcia foniatry w praktyce.

Celem pracy jest przedstawienie aplikacji, która służy do automatycznego wykrywania mowy patologicznej na podstawie bazy nagrań. W pierwszej kolejności przedstawiono założenia leżące u podstaw przeprowadzonych eksperymentów wraz z wyborem bazy mowy patologicznej. Zaprezentowano również zastosowane algorytmy oraz cechy sygnału mowy, które pozwalają odróżnić mowę niezaburzoną od mowy patologicznej. Wytrenowane sieci neuronowe zostały następnie wykorzystane w aplikacji, która umożliwia przeprowadzenie klasyfikacji binarnej na sygnale mowy. Uzyskane wyniki klasyfikacji mowy niezaburzonej i patologicznej zostały porównane ze źródłami literaturowymi. W podsumowaniu eksperymentów zamieszczono również wnioski oraz propozycje rozwoju prowadzonych badań.

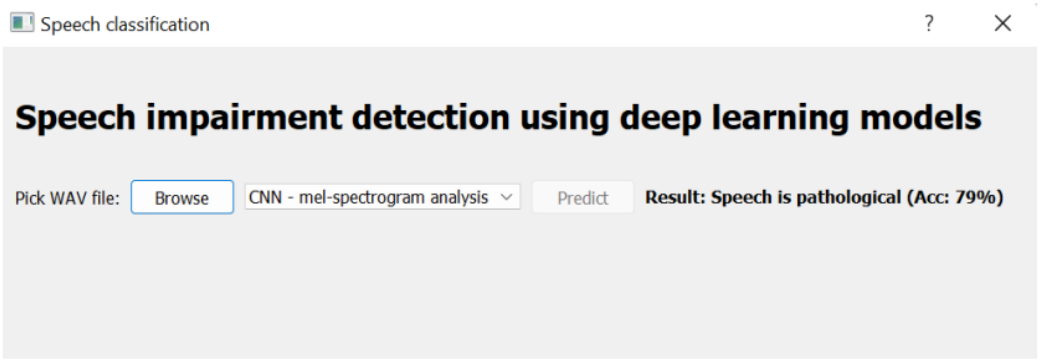
Projekt aplikacji demonstracyjnej ma na celu umożliwienie użytkownikowi przeprowadzenie detekcji zaburzenia na sygnale fonicznym i otrzymania wyniku klasyfikacji binarnej (mowa zaburzona/niezaburzona). W tle przygotowanej aplikacji pracują dwa algorytmy do klasyfikacji sygnału mowy, z możliwością wyboru modelu przez użytkownika. Na etapie przygotowania aplikacji została wyznaczona dokładność (i inne miary ilościowe, takie jak precyzja, czułość, swoistość i F1-score) różnych konfiguracji parametryzacji sygnału mowy oraz architektur i parametrów modeli (wektor cech, jak również reprezentacje 2D, tj. spektrogramy oraz spektrogramy w skali melowej). W wyniku przeprowadzonych licznych eksperymentów zauważono, że modele sieci splotowej z wyznaczonymi spektrogramami w skali melowej oraz z podanym na wejściu wektorem cech zawierającym współczynniki mel-cepstralne (MFCCs, Mel Frequency Cepstrum Coefficients) uzyskały najwyższą skuteczność. W treningu modeli wykorzystane zostały trzy bazy danych: SVD (Saarbrücken Voice Database; zawierająca 71 klas zaburzeń mowy), PC-GITA (nagrania mowy w języku hiszpańskim, pochodzące od osób z chorobą Parkinsona) oraz ADReSSo (zestaw próbek mowy pacjentów cierpiących na demencję wywołaną chorobą Alzheimera).

Aplikacja została przygotowana przy użyciu biblioteki PyQt5 [53], która pozwala na tworzenie prostych aplikacji okienkowych, zawierających określone komponenty. Jak wspomniano wcześniej – w przypadku modeli sieci splotowej obliczane są współczynniki mel-cepstralne (MFCCs) lub podawana

jest reprezentacja 2D w postaci spektrogramów w skali melowej. Wybór algorytmu − sieć splotowa: trening z wykorzystaniem spektrogramów w skali melowej/wektora współczynników MFCCs − warunkuje postać danych, które są obliczane i pobierane na wejście modelu. Na rys. 1 przedstawiono widok aplikacji z wybranym modelem klasyfikacji. Z kolei na rys. 2 zaprezentowano widok aplikacji po przeprowadzeniu predykcji na załadowanym pliku fonicznym zawierającym próbkę mowy zaburzonej.



Rys. 1. Widok aplikacji z wybranym algorytmem do binarnej klasyfikacji zaburzeń mowy



Rys. 2. Widok aplikacji po przeprowadzeniu predykcji na załadowanym pliku fonicznym zawierającym próbkę mowy zaburzonej

W podsumowaniu można stwierdzić, że automatyczna detekcja/identyfikacja zaburzeń mowy/wymowy jest możliwa do przeprowadzenia z wystarczającą skutecznością. Porównanie reprezentacji danych, na których przeprowadzono trening modeli, pozwoliło na stwierdzenie, że analizowanie spektrogramów w skali melowej jest dobrym podejściem w przypadku wykrywania zmian patologicznych w głosie i pozwala na uzyskanie większej dokładności niż w przypadku zastosowania współczynników mel-cepstralnych lub wykorzystaniu spektrogramów w konfiguracji sieci splotowej. Uzyskane wyniki są porównywalne z wartościami dokładności prezentowanymi w literaturze, tj. zawierają się (średnio) w przedziale (61%-71%) w przypadku badań własnych oraz (64%-98%) w przypadku danych literaturowych w zależności od zastosowanych algorytmów i baz danych.

Głównym problemem, który się pojawia w realizacji tego typu badań jest brak baz danych zawierających zbalansowane klasy zaburzeń (podobna ilość danych przypisana do danej klasy), jak również poprawna adnotacja poszczególnych problemów wymowy. Dodatkowym problemem jest stosowanie różnych oznaczeń klas zaburzeń przez różnych specjalistów w procesie tworzenia tego typu zbiorów danych. Dlatego główne działania pozwalające na rozwój tego typu badań należałoby skierować na stworzenie bazy wieloklasowej z danymi równolicznymi i z ujednoliconą adnotacją.

WYKORZYSTANIE TESTU MUSHRA W BADANIU KORZYŚCI UŻYTKOWANIA PROTEZ SŁUCHOWYCH

Piotr SZYMAŃSKI[1], Tomasz POREMSKI[2], Bożena KOSTEK[3]

[1]Sonova Audiological Care Sp. z o.o, Gabriela Narutowicza 130, 90-146 Łódź
Piotr.Szymanski@geers.pl
[2]Advanced Bionics Polska, Sonova Audiological Care Polska Sp. z o.o., Gabriela Narutowicza 130, 90-146 Łódź
Tomasz.Poremski@advancedbionics.com
[3]Politechnika Gdańska, Wydział ETI. Laboratorium Akustyki Fonicznej, Narutowicza 11/12, 80-233 Gdańsk
bozena.kostek@pg.edu.pl

Ocena jakości dopasowania aparatów słuchowych w kontekście korzyści, jakie może przynieść proteza jest złożonym zagadnieniem. W łatwy sposób można wyznaczyć obiektywne parametry aparatów takie, jak wzmocnienie, zniekształcenia harmoniczne, pasmo przenoszenia, itd.

Parametry te jednak nie zawsze mają bezpośredni i decydujący wpływ na subiektywną ocenę jakości dopasowania protezy słuchowej przez pacjenta. Obecne aparaty słuchowe posiadają szereg zaawansowanych rozwiązań, które ułatwiają i poprawiają zwłaszcza rozumienie mowy w różnych, trudnych sytuacjach akustycznych, ale ich porównanie lub pomiar nie jest w pełni możliwy.

Pomiary efektywności aparatu słuchowego mogą dotyczyć wielu aspektów, między innymi kompensacji niedosłuchu, akceptacji, zysku czy też satysfakcji z protezowania.

Autorzy przedstawiają modyfikację powszechnie stosowanego kwestionariusza oceny korzyści użytkowania aparatów słuchowych APHAB (Abbreviated Profile of Hearing Aid Benefit) polegającą na połączeniu go z testem MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor).

APHAB jest to kwestionariusz zamknięty, wypełniany samodzielnie przez użytkownika aparatów słuchowych. Użytkownik ocenia słyszenie zarówno bez aparatów, jak i w aparatach. Kwestionariusz składa się z 24 elementów (twierdzeń dotyczących sytuacji dźwiękowych) w czterech podkategoriach (po 6 elementów na kategorię). Są to:

- EC (*Ease of Communication*) – zdolność komunikacji w ciszy, wysiłek związany z komunikacją w relatywnie łatwych warunkach odsłuchu;
- RV (*Reverberation*) – zdolność komunikowania się w obecności echa, opisuje rozumienie mowy w warunkach umiarkowanego pogłosu;
- BN (*Background Noise*) – zdolność komunikowania się w obecności szumu otoczenia opisuje rozumienie mowy w obecności wielu rozmówców lub innych konkurencyjnych warunkach akustycznych (hałas środowiskowy);
- AV (*Aversiveness of Sounds*) – stopień akceptacji nieprzyjemnych dźwięków, opisuje negatywne reakcje na dźwięki środowiskowe.

Ocenie podlegają określone sytuacje dźwiękowe. Oceniane są w skali 7-stopniowej. Każdy stopień skali od A do G zawiera opis i związaną z nim wartość procentową (minimalna wartość 1%, maksymalna 99%). Korzyść wynikającą z zastosowania aparatu słuchowego można ocenić, analizując średnie wartości procentowe dla poszczególnych kategorii (EC, RV, BN, AV), jak również obliczając wartość średnią dla trzech (EC, RV, BN) lub czterech kategorii.

Formularza APHAB jest elementem zaproponowanej przez autorów metody oceny skuteczności

użytkowania aparatu słuchowego (krótkotrwałego, jak i długotrwałego) u pacjentów z różnym stopniem ubytku słuchu. Porównując wyniki uzyskane dla dwóch grup badanych, stwierdzono, że członkowie jednej z grup uzyskali w ogólności lepsze wyniki. Może to wynikać ze skali oceny słyszenia stosowanej w formularzu APHAB. Połączenie skali literowej, procentowej i opisowej może stanowić trudność w interpretacji dla osób badanych, którymi w przeważającej części są osobami w podeszłym wieku.

W związku z powyższym została opracowana koncepcja modyfikacji kwestionariusza, polegająca tj. przemapowaniu skali APHAB na skalę zgodną ze skalą testu MUSHRA przy jednoczesnym wykorzystaniu logiki rozmytej.

Test MUSHRA znajduje zastosowanie w ocenie jakości dźwięku aparatów słuchowych, zarówno przez osoby z ubytkiem słuchu, jak i słuchem normalnym.

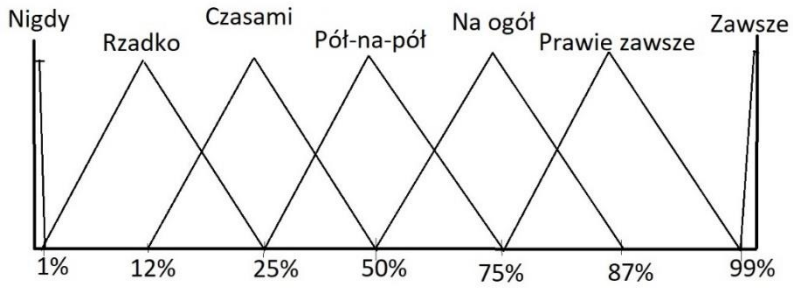
Zadaniem użytkownika aparatów słuchowych, jest ocena słyszenia (jakości dźwięku) bez aparatów i w aparatach według skali 100-punktowej, według mapowania przedstawionego rysunku 1, dla przykładowej sytuacji:

Kategoria BN: „Gdy siedzę przy stole, podczas obiadu, w towarzystwie kilku rozmawiających osób, i próbuję z jedną z nich rozmawiać, mam trudności ze zrozumieniem mowy.”

Skala MUSHRA	0% →						100%
Etykieta MUSHRA	Nigdy →						Zawsze
Skala APHAB	1%	12%	25%	50%	75%	87%	99%
Etykieta APHAB	Nigdy	Rzadko	Czasami	Pół-na-pół	Na ogół	Prawie Zawsze	Zawsze

Rys. 1. Mapowanie skali APHAB na skalę MUSHRA

Odpowiedzi użytkownika zostaną przypisane do odpowiedniej funkcji przynależności dla danej kategorii. Przykładową funkcję przynależności przedstawiono na rysunku 2.



Rys. 2. Przykładowa funkcja przynależności dla kategorii BN

Na bazie oceny użytkownika, przy zastosowaniu reguł logiki rozmytej, zostanie wyznaczony zysk dla określonej kategorii i/lub kilku kategorii.

ANALIZA ZAUTOMATYZOWANEJ METODY POMIARU SYNCHRONIZACJI POMIĘDZY URZĄDZENIAMI GŁOŚNIKOWYMI

Jędrzej BOROWSKI
Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
jedrzej.borowski@pwr.edu.pl

1. WSTĘP

Celem niniejszej pracy jest przedstawienie zautomatyzowanej metody pomiaru synchronizacji pomiędzy urządzeniami głośnikowymi. Prezentowana metoda ma zastosowanie przy pomiarach jakości systemów dźwiękowych składających się z wielu niezależnych bezprzewodowych urządzeń głośnikowych. W przypadku takich systemów poprawna syncronizacja jest niezmiernie ważna, gdyż może zapobiegać niepożądanym odsłuchowo problemom fazowym a także zapewnia poprawne działanie dodatkowego przetwarzania sygnałów, jak np. wirtualizacja dźwięku przestrzennego. Na problemy z synchronizacją mogą wpływać np. niska jakość połączenia sieciowego albo utrata pakietów.

2. METODOLOGIA

Pomiaru synchronizacji dokonuje się poprzez odtworzenie na systemie specjalnie przygotowanego sygnału, zawierające krótkie (ok. 0.1 s) skorelowane impulsy przedzielone 3 sekundami ciszy. Za pomocą pary (lub więcej) mikrofonów pomiarowych rejestrowany jest sygnał odtwarzany przez każde z urządzeń głośnikowych. System oraz mikrofony wysterować należy tak, aby zapobiec przesłuchom pomiędzy kanałami a także otrzymać w zarejestrowanym sygnale odpowiedni odstęp sygnału od szumu. Zarejestrowany przebieg czasowy dzielony jest na równe 3-sekundowe próbki. Dla każdej z tych próbek wyliczana jest wzajemna korelacja sygnałów w dwóch kanałach. Ewaluację metody wykonano dla systemu dwukanałowego, lecz jest możliwe rozwinięcie metody do systemów wielokanałowych. Na podstawie wzajemnej korelacji możemy określić przesunięcie sygnałów względem siebie w próbkach, co znając częstotliwość próbkowania pozwala na wyliczenie przesunięcia czasowego.

3. EWALUACJA METODY

W celu sprawdzenia poprawności działania opracowanej metody przeprowadzono dwa eksperymenty w kontrolowanych warunkach oraz jeden pomiar z użyciem faktycznego systemu bezprzewodowych urządzeń głośnikowych.

3.1. TEST W WARUNKACH KONTROLOWANYCH

Pierwsza weryfikacja polegała na zastąpieniu w systemie pomiarowym sygnału z mikrofonów przez syntetyczny sygnał, w którym jeden z kanałów opóźniono ręcznie o znaną liczbę próbek. Przygotowano sygnały o częstotliwości próbkowania 48 kHz, zawierające impulsy sinusoidalne oraz szum wąskopasmowy. Zastosowano opóźnienia 10, 20, 30 ms. Otrzymane wyniki mieściły się w granicach błędu pomiarowego co oznacza, że sposób wyliczenia wzajemnej korelacji był poprawny.

Druga weryfikacja polegała na sprawdzeniu metody w systemie przewodowych urządzeń głośnikowych. Wykonano pomiar z użyciem dwóch urządzeń głośnikowych i dwóch mikrofonów pomiarowych. Przygotowano sygnały o częstotliwości próbkowania 48 kHz, zawierające impulsy sinusoidalne oraz szum wąskopasmowy. Zastosowano mniejsze opóźnienia celem sprawdzenia dokładności pomiaru z użyciem mikrofonów – 100 μs, 150 μs. Otrzymane wyniki ponownie mieściły się w granicach błędu. Nie zauważono różnicy w wynikach pomiędzy sygnałem szumowym a tonalnym.

3.1. TEST NA URZĄDZENIACH BEZPRZEWODOWYCH

Finalną weryfikacją metody był test na systemie zawierającym urządznenie Amazon Fire Stick, podłączone przez złącze HDMI eARC do telewizora. Przez łącze HDMI odegrano sygnał zapisany na nośniku USB. Do urządzenia Fire Stick podłączono przez sieć WiFi dwa urządzenia bezprzewodowe Amazon Echo, które posłużyły do odegrania sygnału. Za pomocą dwóch mikrofonów pomiarowych dokonano pomiaru synchronizacji tego systemu. W celu zwiększenia dokładności pomiaru zastosowano przy akwizycji danych częstotliwość próbkowania 96 kHz.

Uzyskane wyniki pozwoliły ocenić błąd synchronizacji w badanym systemie. W wyniku pomiarów uzyskano wartości w przedziale 100 – 150 μs.

4. PODSUMOWANIE

Przeprowadzone eksperymenty potwierdziły skuteczność badanej metody pomiarów synchronizacji bezprzewodowych urządzeń głośnikowych. Zarówno test na syntetycznych danych jak i test w kontrolowanych warunkach dały wyniki w granicach błędów pomiarowych. Test na prawdziwym bezprzewodowym systemie audio pokazał że metoda może mieć zastosowanie do sprawdzania jakości połączenia w takich systemach. Rodzaj sygnału testowego (szum lub sinus) nie miał wpływu na wynik pomiaru. Największą trudnością podczas pomiaru było odpowiednie wysterowanie systemu tak aby osiągnać odpowiedni odstęp sygnału od szumu.

KONCEPCJA MATRYCY FALOWODOWEJ DO KSZTAŁTOWANIA FRONTU FALI AKUSTYCZNEJ

Tomasz NOWAK[1]
Andrzej DOBRUCKI[2]
[1]Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
T.Nowak@pwr.edu.pl
[2]Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
Andrzej.Dobrucki@pwr.edu.pl

1. WPROWADZENIE

Nauka oraz inżynieria stojąca za projektowaniem przetworników elektroakustycznych oraz urządzeń głośnikowych spotyka się ze skomplikowanymi ograniczeniami wynikającymi z niedopasowania kątowego zakresu promieniowania źródeł oraz kształtów ich frontów falowych. Fizyczne wymiary i rozmieszczenie głośników, także urządzeń głośnikowych, w relacji do generowanych długości fali związane są z występowaniem interferencji na powierzchni wspólnego frontu falowego. W wyniku niedopasowania fazowego oraz amplitudowego powstaje źródło dźwięku zniekształcające sygnał akustyczny w zmienny sposób w całej objętości pola akustycznego. Dla rozwiązania wyżej wymienionych problemów autor proponuje zastosowanie soczewki w formie matrycy falowodów, które dzielą czoło fali na skończoną ilość fragmentów oraz wprowadzają kontrolowane opóźnienie każdego z nich. W efekcie powstaje narzędzie wpływające na rozkład fazy na powierzchni źródła.

Wyżej wspomniana koncepcja była już przedstawiona w publikacjach naukowych. [1] [2] Dotychczasowe rozwiązania opierały się na obrocie macierzy wyjściowej względem wejściowej. Obrotowe soczewki pozwalają na regulację krzywizny czoła fali tylko w jednym kierunku. Dla soczewek, o kształcie wielokąta foremnego lub koła, za pomocą kąta obrotu oraz grubości struktury kształtowaniu podlega promień krzywizny fali kulistej, nie ma możliwości osiągnięcia kształtu fali elipsoidalnej lub cylindrycznej.

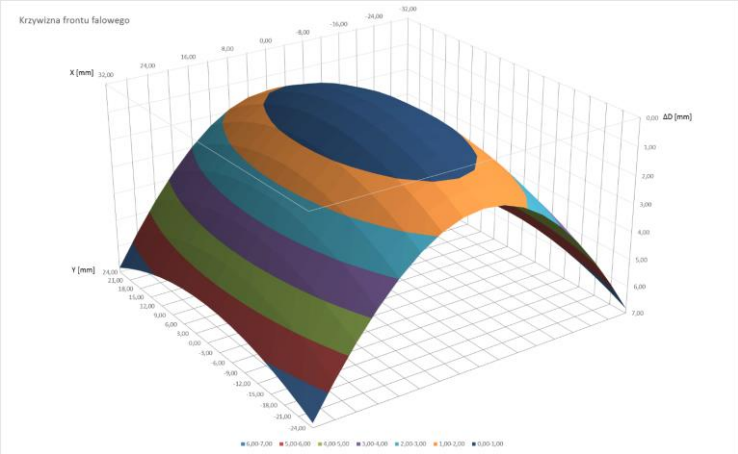
2. PROPONOWANE ROZWIĄZANIE

Propozycją struktury dyskretyzującej, która pozwala na regulację krzywizny generowanej fali w dwóch płaszczyznach jest macierz wyjściowa o rozmiarach i pozycji innej niż wejściowa. Skalowanie rozmiaru oraz pozycji wyjściowej macierzy w obu kierunkach umożliwia kształtowanie czoła fali, jako płaskie, cylindryczne, elipsoidalne, oraz inne krzywizny przestrzenne.

Zakładając wymiary i rozkład komórek macierzy, ze wzoru (1) można obliczyć drogę, czyli długość każdego tunelu łączącego komórkę wejściową z wyjściową. Dla każdej komórki macierzy można obliczyć drogę. Gdy macierz wyjściowa znajdująca się w pozycji t = 20 mm (grubość soczewki) zostanie powiększona poprzez iloczyn współrzędnych X_i ze współczynnikiem skalującym A_x oraz analogicznie współrzędnych Y_i ze współczynnikiem skalującym A_y, otrzymana zostaje macierz wyjściowa o symetrycznym rozkładzie opóźnień, których wartość rośnie dla każdej komórki wraz z odległością od środka symetrii. Dystrybucję opóźnień $\Delta t_{m,n}$ dla każdej z komórek wyjściowych obliczono ze wzoru (2).

$$D_{m,n} = \sqrt{(A_x X_m - X_m)^2 + \left(A_y Y_n - Y_n\right)^2 + t^2} \tag{1}$$

$$\Delta t_{m,n} = \frac{D_{m,n} - h}{c} \tag{2}$$



Rys. 1. Wizualna reprezentacja frontu falowego za pomocą przestrzennej dystrybucji przyrostu drogi ΔD, dla soczewki o grubości $t = 20mm$ oraz współczynnikach skalowania: $A_x = 2, A_y = 1{,}5$

Gdzie $D_{m,n}$ - droga między odpowiadającymi sobie komórkami macierzy wejściowej i wyjściowej, $\Delta t_{m,n}$ - opóźnienie w poszczególnych komórkach soczewki, A_x, A_y - współczynniki skalowania macierzy wyjściowej soczewki, kolejno dla osi X oraz Y. Wykorzystując obliczony przyrost drogi, dla każdej z komórek wyjściowych soczewki, można wizualnie przedstawić front falowy jako powierzchnię łączącą punkty tego przyrostu na wykresie przestrzennym. (Rys. 1)

W pracy obliczono oraz przedstawiono inne przykłady frontów falowych, łącznie z parametrycznymi, gdzie występuje zmiana kształtu w funkcji jednego lub obu wymiarów macierzy wyjściowej. Wykonano prototyp soczewki dla przetwornika AMT, który znacząco zwiększył jego kątowy zakres promieniowania w płaszczyźnie pionowej. Działanie urządzenia zostało potwierdzone pomiarami.

LITERATURA

[1] BERSTIS V., *3D PRINTED ACOUSTIC LENS FOR DISPERSING SOUND*, J. Audio Eng. Soc., vol. 66, no. 12, pp. 1082–1093, 2018

[2] DAHL T., EALO J.L., PAPAKONSTANTINOU K., PAZOS J.F., *Design of Acoustic Lenses for Ultrasonic Human-Computer Interaction*, IEEE International Ultrasonics Symposium, 2011

[3] HEIL C., *United States Patent Number 5,163,167 (US005163167A)*, 1992

ANALIZA PARAMETRÓW AKUSTYCZNYCH DŁUGIEGO DOMU W LOFOTR

Piotr KSIĄŻEK[1]
Adam PILCH[2]
[1]AGH Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie,
al. Mickiewicza 30 30-059 Kraków.
pioksiazek@student.agh.edu.pl
[2]AGH Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie,
al. Mickiewicza 30 30-059 Kraków.
apilch@agh.edu.pl

1. WPROWADZENIE

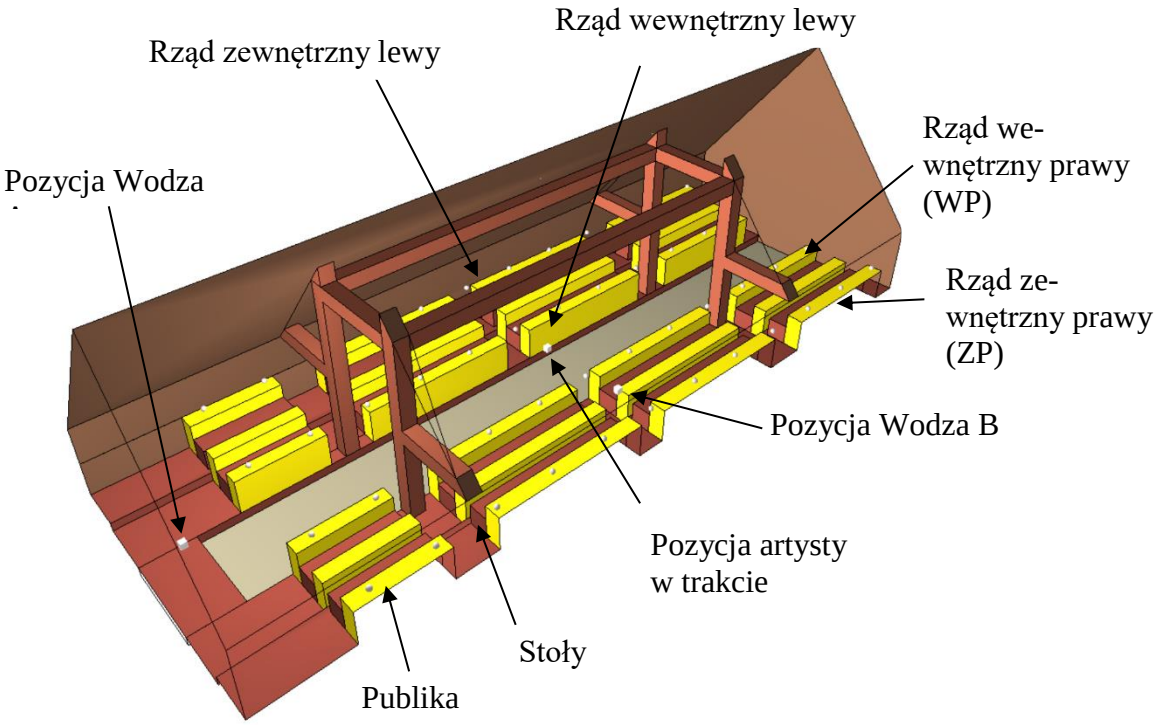
Gdy skandynawski najeźdźca ruszał do walki na europejskich ziemiach wiedział, że jeśli w boju przyjdzie mu polec, spocznie na wieczność w Wallhalli [5], wielkim, długim domu, w którym nigdy nie zabraknie jedzenia ani rozrywki [5]. Wierzenia te najpewniej mają swoją genezę – długi dom jest historycznie bardzo popularną konstrukcją [4]. W szczególności rozpoznawalne są długie domy budowane przez ludy Skandynawii zamieszkujące tereny sięgające od współczesnej Danii aż po Grenlandię. Przypadek rozpatrywany w niniejszym referacie jest przypadkiem szczególnym zarówno z punktu widzenia archeologicznego jak i kulturowego. Długi dom w Lofotr który został odkryty przez archeologów na terenie wioski Borg w archipelagu Lofoten na północy Norwegii jest największym odkrytym dotąd długim domem [1, 3]. Przestrzeń tegoż domu nie była poświęcona jedynie sprawom gospodarczym – znajdowała się tam bowiem wielka hala jadalna w której odbywały się Tingi, czyli walne spotkania w trakcie których obradowano nad sprawami wielkiej wagi oraz oczywiście wystawne uczty w czasie których Skaldowie zabawiali ucztujących swoimi występami scenicznymi zarówno poprzez poezję jak i muzykę [6]. Rozpatrywanym do analizy budynkiem jest długi dom wodza w Lofotr, którego fundament został odkryty przez archeologów w trakcie wykopalisk w latach 80 XX wieku. Wykopaliska odkryły pozostałości drewnianego budynku o długości 83 metrów oraz szerokości 9,5 metra. Obecnie nieopodal miejsca wykopalisk znajduje się rekonstrukcja tego budynku wznosząca się na wysokość 9 metrów zadaszona dachem z drewnianego gontu [1]. Budynek był wewnętrznie podzielony na sekcje. Sekcja zawierająca salę biesiadną miała długość 27 metrów [3] i prawdopodobnie zawierała cztery rzędy ław biesiadnych ustawionych na podniesionych drewnianych podestach. Założenie na temat umeblowania sali tworzone jest na podstawie rekonstrukcji domu w Lofotr i jest w pewnym stopniu kwestią dyskusyjną [3]. Tron wodza w obecnym stanie rekonstrukcji budynku znajduje się w szczycie sali przy krótszej jej ścianie. Takie umiejscowienie tronu jednak jest podawane w wątpliwość na bazie faktu, iż kolumny podtrzymujące dach zasłaniają wodzowi siedzącemu na tronie część sali biesiadnej. Modelowanie akustyczne zrozumiałości mowy przemawiającego wodza może pomóc w określeniu jakie umiejscowienie jego tronu byłoby odpowiednie.

2. MODEL AKUSTYCZNY

Na podstawie literatury, fotografii wnętrz rekonstruowanego budynku dostępnych w internecie oraz ortofotomapy okolicy, w której rekonstrukcja owa się znajduje został stworzony przybliżony model akustyczny wnętrza. Model został stworzony za pomocą oprogramowania SKETCHUP MAKE 2017, następnie zaimportowany do programu Blender oraz wyeksportowany do formatu GEO akceptowanego przez oprogramowanie CATT-Acoustic. Przypisanie materiałów modelowanym powierzchniom stanowiło pewien problem z powodu braku dużej bazy danych pogłosowych współczynników pochłaniania naturalnych materiałów. Problem ten przezwyciężono poprzez przybliżenie parametrów akustycznych materiałów tworzących wnętrze do zbadanych materiałów o znanych pogłosowych współczynnikach pochłaniania dźwięku.

W modelu zdefiniowano 48 odbiorników umieszczonych 50cm ponad powierzchnią widowni w czterech rzędach. Zdefiniowano także powierzchnie generowania graficznego rozkładu parametrów akustycznych w pomieszczeniu. Do modelu wprowadzono trzy źródła mające na celu symulowanie trzech rozpatrywanych scenariuszy: Wódz przemawiający w pozycji A, Wódz przemawiający w pozycji B oraz występ muzyczny odbywający się na środku sali. Źródła zostały wprowadzone jako źródła wszechkierunkowe o mocy akustycznej zdefiniowanej przez oprogramowanie CATT-Acoustic we wzorcu „Lp_voice_raised".

Najistotniejszym badanym w modelu parametrem jest parametr STI (Speech Transfer Index) oraz parametr C-80 określający klarowność odbioru muzyki. Dla obliczeń parametru STI wykorzystano poziom szumu tła określony na poziomie 38dBA. Wykorzystano typ STI „IEC ED4 male".



Rys. 1. Wizualizacja modelu akustycznego z oznaczonymi istotnymi miejscami

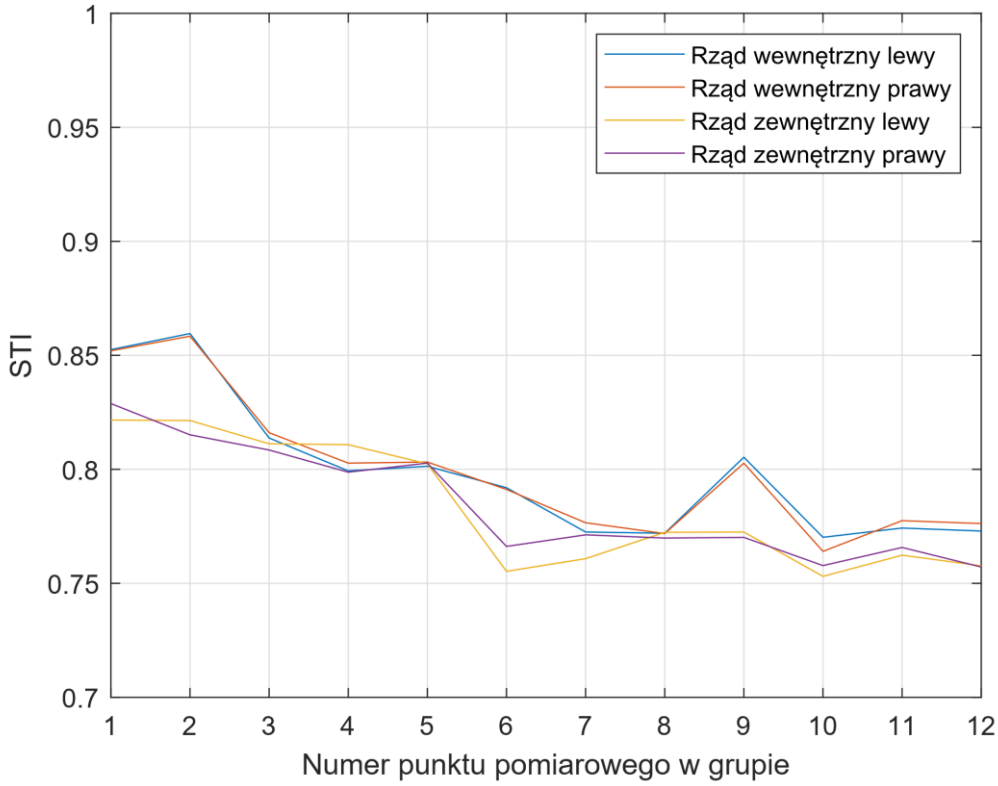
Tab. 1. Tabela pogłosowych współczynników pochłaniania dźwięku α materiałów wykorzystanych w modelu

Typ materiału	125Hz	250Hz	500Hz	1000Hz	2000Hz	4000Hz
Klepisko [7]	0,15	0,25	0,40	0,55	0,60	0,60
Powierzchnie drewniane lite [8]	0,11	0,11	0,12	0,11	0,10	0,08
Powierzchnie drewniane z pustką wewnętrzną [9]	0,40	0,30	0,20	0,17	0,15	0,10
Publika [2]	0,24	0,40	0,78	0,98	0,96	0,87

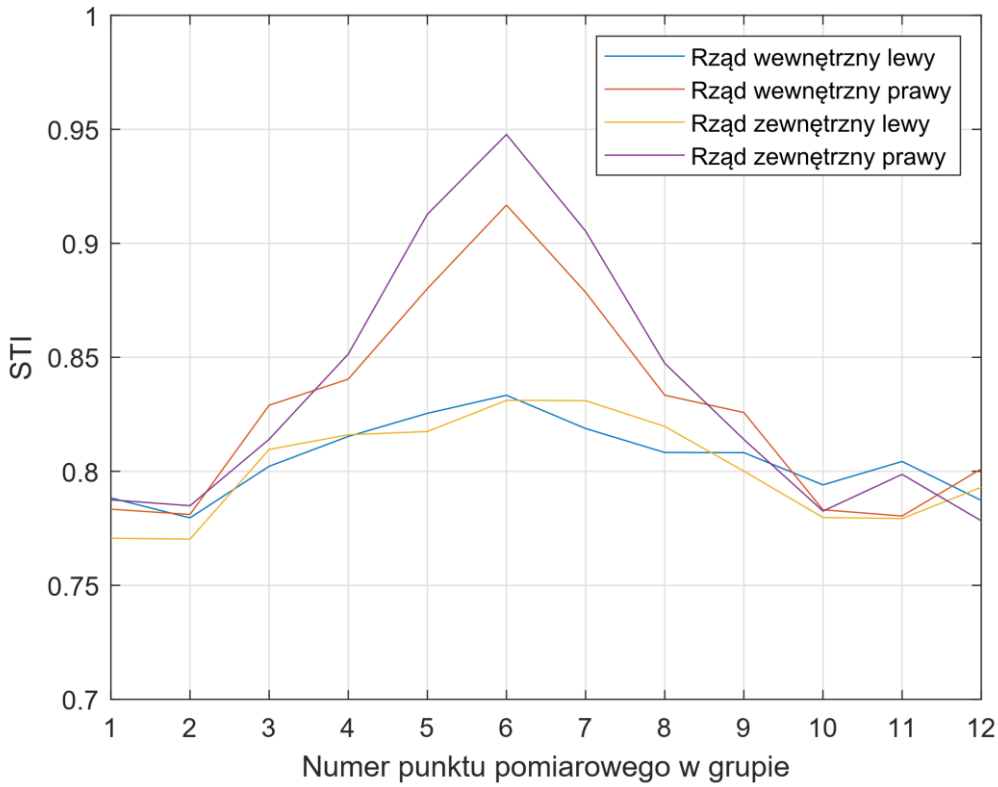
3. WYNIKI I WNIOSKI

Tab. 2. Tabela wyników poziomu STI w modelu

Sytuacja	LT	LS	PT	PS	Ca-łość
Wódz w pozycji A	0,80	0,80	0,78	0,78	0,79
Wódz w pozycji B	0,81	0,83	0,80	0,84	0,82



Rys. 2. Wykres parametru STI w grupach odbiorników wodza przemawiającego w pozycji A



Rys. 3. Wykres parametru STI w grupach odbiorników wodza przemawiającego w pozycji B

Wódz przemawiając zarówno w pozycji A jak i pozycji B zostałby zrozumiany na poziomie znakomitym w większości przestrzeni przy założeniu poziomu hałasu tła zbliżonego do 35dBA. Widoczne jest, iż mowa wodza przemawiającego w pozycji B jest zrozumiała w sposób wyśmienity dla rzędów najbardziej zbliżonych do niego, natomiast wskaźnik ten opada nieznacznie dla rzędów po przeciwległej stronie sali. Wskaźnik STI mowy wodza w przypadku pozycji A opada wraz z kolejnymi punktami pomiarowymi określonymi w modelu co jest związane ze zwiększającą się odległością między mówcą a odbiorcą dla każdego kolejnego punktu pomiarowego w grupie (rzędzie). Wyniki uśrednione są wysoce zbliżone do siebie, nie rzutują więc jednoznacznie na prawdopodobieństwo umiejscowienia tronu wodza w pozycji A bądź też B mimo faktu, iż wynik uśrednionego wskaźnika STI dla wodza przemawiającego w pozycji B jest wyższy o 0,03. Średnia wartość wskaźnika klarowności muzyki C-80 osiągnęła w obliczeniach modelu poziom 5,4dB, co wskazuje, iż nie jest to wnętrze spełniające współczesne oczekiwania wobec przestrzeni koncertowych.

4. DALSZE MOŻLIWOŚCI ROZWOJU

Dalsze możliwości rozwoju tego typu działań są szerokie. Przede wszystkim model akustyczny stworzony na potrzeby niniejszego eksperymentu może zostać poddany kalibracji po wykonaniu pomiaru kontrolnego w miejscu rekonstrukcji długiego domu z Lofotr. Dalsze zgłębianie tematyki modelowania akustycznego struktur rekonstruowanych przez nauki historyczne jest tematyką, którą warto zbadać. Możliwe są także dalsze eksploracje modelu akustycznego stworzonego na potrzeby niniejszego eksperymentu – przykładowe adaptacje akustyczne mające na celu obniżenie wskaźnika C-80 korzystając jedynie z materiałów dostępnych w epoce średniowiecza oraz rejonie Skandynawii.

LITERATURA

[1] „Lofotr Wikingmuseum - The Chieftain's house," 09 26 2022. [Online]. Available: https://www.lofotr.no/en/the-history/the-longhouse.

[2] KUTTRUFF. H., Room acoustics, CRC Press, 2019.

[3] NARMO L. E, „The Unexpected," Acta Archaeologica Lundensia Series in 8°, 2011.

[4] ZORI, BYOCK, ERLENDSSON, MARTIN, WAKE, EDWARDS, „Feasting in Viking Age Iceland: sustaining a chiefly political economy in a marginal environment," Antiquity, 2013.

[5] SŁUPECKI L.P, Mitologia Skandynawska w epoce Wikingów, Kraków: Zakład wydawniczy Nomos, 2014.

[6] GARDEŁA L., „What the Vikings did for fun? Sports and pastimes in medieval northern Europe," World Archaeology, 2012.

[7] EGAN, Architectural Acoustics, J. Ross Publishing, 2007.

[8] LAWRENCE A., Architectural Acoustics, 1970.

[9] Baza danych oprogramowania CATT-Acoustic

FLEXIBLE ROOM ACOUSTICS SIMULATION PLATFORM FOR IMPULSE RESPONSE ESTIMATION

Aleksander KACZMAREK[1,2], Piotr CENDA[1], Jakub WASILEWSKI[1], Tomasz WOŹNIAK[1], Maria PEŃSKO[1] and Łukasz JANUSZKIEWICZ[1]

[1]SoftServe

[2]Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław

A.Kowalski@pwr.wroc.pl

Measuring room impulse responses of complex acoustic environments is a crucial starting point of many audio engineering scenarios. Usually performing a series of necessary experiments in all required acoustic arrangements is demanding in terms of time and effort. Using synthetic or simulated RIRs is one of the data augmentation techniques used for developing spatial audio applications based on machine learning algorithms. In this paper we present a customized simulation platform that allows for flexible estimation of impulse responses in various acoustic scenarios. Simulations are based on k-Wave – an open-source acoustics toolbox for MATLAB and C++ to compute a time-domain acoustic wave propagation based on the k-space pseudo-spectral method. We demonstrate the capabilities of the platform by simulating selected physical phenomena such as inverse square law and early reflections. We also show the results of applying our platform to a specific use case: spherical microphone array impulse response estimation. We compare our results with the analytical calculations of impulse responses based on Bessel-family functions and Legendre functions.

1. INTRODUCTION

Room impulse responses (RIR) are key characteristics of complex acoustic environments for various audio engineering scenarios. RIR allows to describe acoustic properties of real places such as e.g., concert halls, but due to the recent development of virtual and augmented reality, RIRs of simulated virtual rooms are also needed to properly mimic real world acoustic experience. Usually, to obtain an impulse response (IR) a series of experiments in many acoustic arrangements is needed, for example recording an impulse signal emitted from every possible position of a speaker inside a room [1, 2, 3]. This can be demanding in terms of time and effort. Therefore, having a reliable simulation environment for generating synthetic RIRs of arbitrary rooms would benefit in alleviating the need for timely real experiments and allow to develop algorithms for virtual rooms.

There are several software tools allowing for this kind of simulations e.g., Acoustics Module for COMSOL Multiphysics [4]. For simulations run in the frequency domain it employs the Helmholtz equation, whereas in the time domain, the classical scalar wave equation is used. In the frequency domain, both finite element method (FEM) [5] and boundary element method (BEM) [6] are available, as well as hybrid FEM–BEM. The main motivation for our work was to develop a simulation environment capable of generating impulse responses of spaces with a given set of parameters. We used the existing open-source simulation environment, that could be further adjusted to more specialized use cases. It also lets us optimize computational time and run simulations on high-power machines like Google Cloud Platform.

In this paper we present a customized simulation platform that allows for flexible estimating RIRs of various acoustic scenarios based on k-Wave – an open-source acoustics toolbox for MATLAB and C++. We demonstrate the capabilities of the platform by simulating selected physical phenomena such as inverse square law and early reflections. We also show the results of applying our platform to a specific use case: spherical microphone array (SMA) IR estimation. We compare our results with the analytical calculations of IRs based on Bessel-family functions and Legendre functions [7, 8].

2. ENVIRONMENT DESCRIPTION

2.1. K-WAVE

K-Wave is an open-source acoustics toolbox for MATLAB and C++ [9]. The software is designed for time-domain simulation of propagating acoustic and ultrasound waves in 1D, 2D, or 3D in complex and tissue-realistic media. The simulation functions are based on the k-space pseudo-spectral method. The toolbox includes an advanced numerical model that can account for both linear and nonlinear wave propagation, an arbitrary distribution of heterogeneous material parameters and power law acoustic absorption [9, 10, 11]. The toolbox has also the ability to model pressure and velocity sources, including photo-acoustic sources, and diagnostic and therapeutic ultrasound transducers, and to specify arbitrary detection surfaces with directional elements, with options to record acoustic pressure, particle velocity, and acoustic intensity. An optimized C++ version of the code allows to maximize computational performance for more complex simulations.
Preparation for the simulation begins with determining the physical parameters of the simulation space and the medium in which the signal propagates, such as the size of the room, variations in density and speed of sound propagation, spatial and temporal resolution, etc. Next, it is necessary to arrange the sensors and sound sources in the simulation room and the signals to be generated. Once the data is prepared in this way, a simulation is run calculating pressure value at each voxel of the room every time step.
Fig. 1 presents a snapshot from the simulation performed using k-Wave toolbox in Octave. Acoustic waves are propagating in a small box with the SMA in the center. Sound source is placed near the boundary of the box and the excitation signal is the Dirac impulse played at time zero. The yellow and blue shades indicate positive and negative relative pressure, respectively, and the black circle denotes the ABS plastic microphone casing sphere.

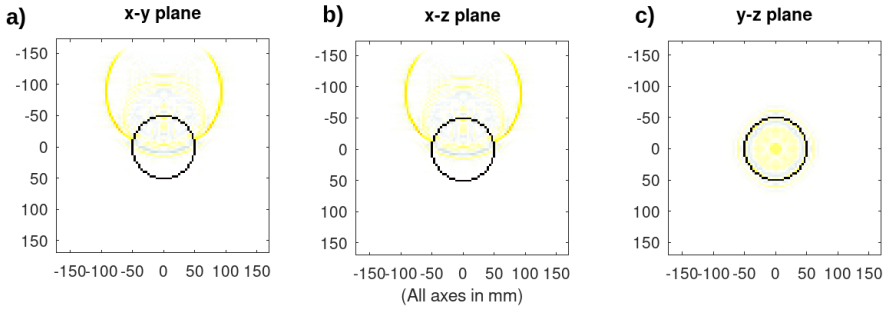


Fig.1 Example of the simulation run using k-Wave toolbox.

2.2. PIPELINE

In order to tackle the problems with computational complexity and memory requirements associated with simulating realistic size rooms, we developed a pipeline that uses a simple configuration file with the simulation parameters to run simulation on the computing cluster. The pipeline consists of two main steps. In the first one Matlab scripts use the k-Wave toolbox to generate simulation parameters that describe the properties of the signal sources, acoustic sensors, and simulation room. The latter one is based on Python scripts that use binary files derived from the k-Wave toolbox to run the simulation and generate plots. Due to the large size of the simulated environments, as well as the required spatial and time resolution, Google Cloud Platform was used, where calculations were carried out using a virtual machine with 40 Intel Xeon processors and 1TB of RAM.

2.3 INPUT DATA

Simulation generated using the k-Wave toolbox can be specified and described using a variety of available inputs. It is possible to specify both high-level and low-level properties of the simulation. The high-level parameters are the size of the simulation room, the number and locations of sound sources and sensors and the parameters of the sound source signal. The low-level ones describe the physical properties of the simulation environment, such as the propagation speed of the wave in the given medium and its density, as also the spatial and time resolution of simulation.

2.3.1 PHYSICAL INPUT PARAMETERS

The physical parameters of the simulation environment in which the simulations are carried out determine how it interacts with the given signal. In each simulation, parameters such as the properties of the medium, the spatial and time resolution and the properties of the simulation room boundaries that absorb the propagating waves are set top-down. The medium is defined by specifying its density as 1.2 [kg/m3], its absorption coefficient and exponent as 1.1795 and 1.4484, respectively, and its speed of sound propagation as 343 [m/s] (all parameters correspond to real air properties [12]). The spatial resolution, which defines the dimensions of the voxels from which the simulation is constructed, is set to 0.0035729 [m] and is derived from the quotient of the speed of sound propagation in the medium and the target sampling frequency. It defines the highest possible frequency that can propagate in the simulation environment. The time resolution is set by calling the appropriate k-Wave toolbox function, which calculates the required time unit for which computational stability will be maintained. The boundaries of the simulation room are encircled by an absorbing PML - Perfectly Matched Layer [9]. It ensures that any signal going outside the simulation room is attenuated and does not appear on the opposite side of the room, what happens when the layer is absent. The PML is 50 grids thick and has an absorption coefficient of 2, minimizing the layer penetration and reflection of signals.

2.4 OUTPUT SIMULATION DATA

The toolbox allows to record the change of pressure over time at each point in the simulated space with the accuracy of the dimensions of a single voxel. This allows for accurate and detailed analysis of the signal propagating in the medium of different acoustic properties, containing objects and obstacles. Using a single Dirac pulse as the excitation signal, the platform can be used to simulate IRs of a variety of sound source-sensor configurations.

3 EXAMPLES OF SIMULATED PHENOMENA

As we mentioned in Section 1, k-Wave is a package developed mostly for ultrasound simulations in complex and tissue-realistic media that we adjusted for our purposes. For that reason, we first checked if this software correctly accounts for various physical phenomena that occur during audio engineering scenarios but may be not so important in modeling of tissue-realistic media.

3.1 INVERSE SQUARE LAW

The sound intensity is the product of the in-phase component of the Root Mean Square (RMS) particle velocity and the RMS sound pressure. Both quantities are inversely proportional to the distance from the source. Therefore, the intensity of the sound should follow an inverse-square behavior. To demonstrate that this law is satisfied in our simulation environment, a simulation was run in a room of size 0.8x0.4x0.4 [m], with a source generating a Dirac pulse of 20 [Pa]. A series of sensors was set every voxel from the source across a distance of 0.6 [m]. For each sensor, the intensity of the measured signal

was calculated and the dependence on the distance from the source was plotted, which can be seen in Fig. 2, thus confirming the law.

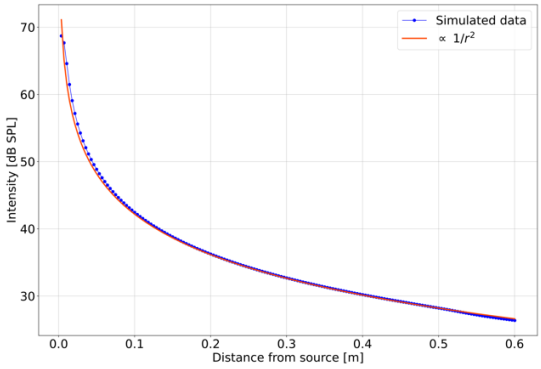


Fig.2 Intensity of the signal recorded by the microphones distributed along the radial direction from the source.

3.2 EARLY REFLECTIONS

Another important feature of k-Wave package is that it considers the reflections of waves that propagate in the computational grid. This was already probably seen in section 2.1, but we wanted to demonstrate it on the simulation of the realistic scenario. Therefore, instead of an anechoic chamber, we simulate a room with one reflecting wall, which resembles ABS material. That is the absorption coefficient and density of simulated material are 0.3 and 1000 [kg/m3] respectively. We place a speaker and a microphone in a 3x2x1 [m] room, 2 [m] apart and 1 [m] from the wall. We expect to observe at least two picks in the time domain of the recorded signal. First due to direct signal and later the second one due to the signal reflected by the wall. The results of the simulation are presented in fig. 3. As expected, we obtained two picks. Moreover, the time which separates the picks agrees with what one would expect given the fact that the path the reflected sound takes is 0.828 [m] longer and the sound speed we set in the simulation is 343 [m/s]. Which agrees with observations - the sound travels 2 [m] in 0.005831 [s], and 2.828 [m] in 0.008246 [s].

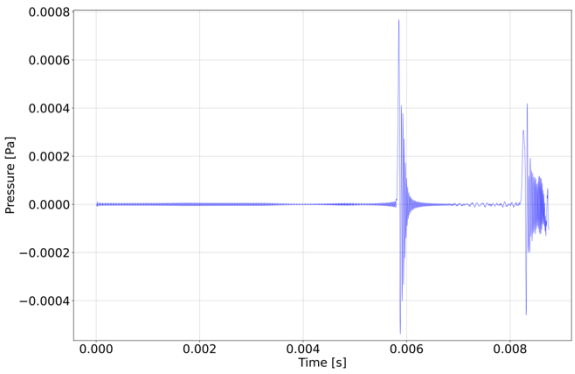


Fig.3 Pressure signal detected by a microphone in a 3x2x1 [m] room with reflecting wall.

Research is also underway to further explore the properties and capabilities of the developed simulation platform. Properties such as diffraction and Doppler effect will be tested, describing the behavior of the simulation during signal scattering and interaction with encountered obstacles and during sound source movement in the simulated room.

4 SPHERICAL MICROPHONE ARRAY SIMULATIONS

4.1 SPHERICAL MICROPHONE ARRAY MODELING

We consider a simulation of a realistic scenario in which a set of IRs of SMA is measured in an anechoic chamber. The SMA's radius is 5 [cm] and its casing is modeled to have a sound speed and the density of ABS plastic. The number of microphones is 20 and they are placed in the vertices of dodecahedron. We placed the SMA in the center of the simulated room of size 1.64x1.64x0.5 [m]. The speakers that will generate a Dirac impulse are distributed on a Lebedev grid of order 974 (so there will be 974 speakers) of radius 0.8 [m]. The center of the Lebedev grid coincides with the center of SMA. For demonstration purposes we choose a microphone 5 which coordinates are phi -45.00 [◦], theta -35.26 [◦] and radius 0.8 [m]. In Fig. 4 we plot a set of pressure signals registered by the microphone from the circle of speakers lying in the equatorial plane of the sphere. The plots show the differences in time of arrival of the signal from different directions – very similar for both compared simulation methods.

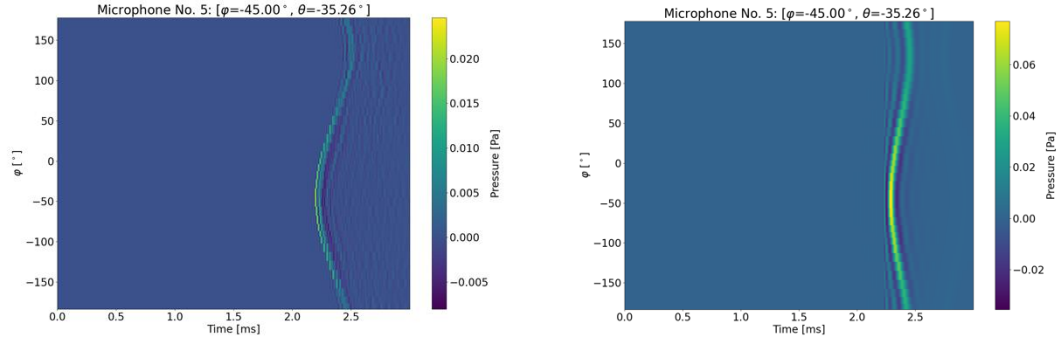


Fig.4 IRs of the simulated spherical microphone array obtained with k-Wave simulation (left) and with analytical calculations (right).

4.2 COMPARISON WITH ANALITYCAL CALCULATIONS

To validate our simulation environment, we compare the results from section 4.1 with analytical calculations (Fig. 5). In this method the computation of the frequency and time domain response is based on the theoretical expansion of a scalar incident plane wave field to a series of wavenumber-dependent Bessel-family functions and direction-dependent Fourier or Legendre functions. The achieved results are in line with the expected IR derived from the analytical calculations. The effect of a wave passing through a medium of different density than air is visible, which is reflected on the graph with lower signal magnitudes received by microphones placed on the side of the sphere opposite to the signal source. Moreover, since k-Wave simulation takes more effects into account we see more distortions in the IR for frequencies above 10 kHz.

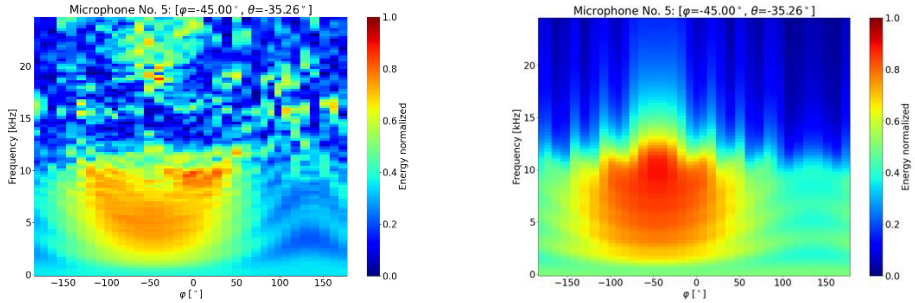


Fig.5 Frequency response of the simulated spherical microphone array obtained with k-Wave simulation (left) and with analytical calculations (right).

5 CONCLUSIONS

We proposed the room acoustics simulation platform that was verified with several simulation examples. The results demonstrate a decent compliance with the physical properties of wave propagation, and large flexibility in terms of both, creating various acoustic scenarios and capabilities to be adjusted to more specific use cases. We also used the platform to estimate a set of IRs of an SMA in an anechoic environment and confirmed that it can be used for specific acoustic RIR simulations. Our further work will concentrate on optimization of computational time on cloud platforms to prepare the platform for efficient generation of large amount of data for machine learning applications.

REFERENCES

[1] Ivanic, J., Ruedenberg, K. *Rotation Matrices for Real Spherical Harmonics. Direct Determination by Recursion Page: Additions and Corrections.* Journal of Physical Chemistry A, 102(45), 9099?9100, 1998.
[2] Politis, A. *Microphone array processing for parametric spatial audio techniques*, Doctoral Dissertation, Department of Signal Processing and Acoustics, Aalto University, Finland, 2016.
[3] Teutsch, H. *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*, Springer, 2007.
[4] *www.comsol.com/acoustics-module*
[5] Bathe, Klaus-Jürgen. *Finite element method*. Wiley encyclopedia of computer science and engineering, 2007, 1-12.
[6] Kirkup, Stephen Martin. *The boundary element method in acoustics*. Integrated sound software, 2007.
[7] Williams, Earl G. *Chapter 6 - Spherical Waves*, Fourier Acoustics, Academic Press, 1999, pp. 183-234, ISBN 9780127539607, doi.org/10.1016/B978-012753960-7/50006-1.
[8] Williams, Earl G. *Chapter 4 - Cylindrical Waves*, Fourier Acoustics, Academic Press, 1999, pp. 115-148, ISBN 9780127539607, doi.org/10.1016/B978-012753960-7/50004-8.
[9] Treeby, B. E., Cox, T. B. *k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields*, Journal of Biomedical Optics, vol. 15, no. 2, p. 021314, 2010.
[10] Treeby, B. E., Jaros, J., Rendell, P. A., Cox, T. B. *Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method*, Journal of the Acoustical Society of America, vol. 131, no. 6, pp. 4324–4336, 2012.
[11] Treeby, B. E., Cox, T. B. *Modeling power law absorption and dispersion for acoustic propagation using the fractional Laplacian*, The Journal of the Acoustical Society of America, vol. 127, pp. 2741–48, 05 2010.
[12] *www.engineeringtoolbox.com*

POLSKA SEKCJA AES − Z PERSPEKTYWY 30 LAT DZIAŁALNOŚCI

Bożena KOSTEK

Politechnika Gdańska, Wydział ETI, Laboratorium Akustyki Fonicznej, G. Narutowicza 11/12, 80-233 Gdańsk
bozena.kostek@pg.edu.pl

W prezentacji przywołano pokrótce najważniejsze działania, które towarzyszyły powstaniu i funkcjonowaniu Polskiej Sekcji towarzystwa naukowego *Audio Engineering Society* (PS AES). Wskazano także jej twórców, prof. Mariannę Sankiewicz i prof. Gustawa Budzyńskiego. Zaprezentowano skład Zarządu PS AES w kolejnych kadencjach. Dla przypomnienia − PS AES została powołana w celu rozwijania inżynierii dźwięku, popularyzowania inżynierii dźwięku jako dziedziny naukowej i dydaktycznej ze szczególnym uwzględnieniem obszarów ważnych dla nauki, edukacji, kultury, społeczeństwa, gospodarki narodowej i ochrony środowiska oraz w celu nawiązania łączności naukowej z osobami pracującymi twórczo w inżynierii dźwięku i pokrewnych dziedzinach. Dlatego warto zauważyć rolę PS AES w promowaniu osiągnięć naukowców i młodych adeptów nauki, prezentujących swoje prace w ramach udziału w sympozjach ISSET (Międzynarodowe Sympozjum Inżynierii i Reżyserii Dźwięku) oraz NTAV (Nowości w Nowości w Technice Audio i Wideo) − odbywających się pod auspicjami PS AES, jak też prezentowaniu polskiej inżynierii dźwięku na forum europejskim i światowym. Podkreślono także udział PS AES w corocznych konkursach na najlepsze nagranie (*recording competition*) oraz najlepszą konstrukcję (*design competition*) na Konwencjach AES oraz Sympozjach poprzez prace zgłaszane przez studentów, młodych naukowców i inżynierów dźwięku, co zaowocowało licznymi nagrodami otrzymanymi na forach krajowym i międzynarodowym. Warto też wspomnieć o organizacji 138 Konwencji AES w Warszawie, która zgromadziła ok. 1500 osób ze świata naukowego, ale również przedstawicieli przemysłu i wystawców technologii. W podsumowaniu nie zabraknie odniesień do licznych nagród uzyskanych przez członków PS AES, jak również zajmowanych funkcji w zarządzie towarzystwa naukowego AES. Ten krótki przegląd zakończy się wnioskami, które nasuwają się z perspektywy ponad 30 lat działalności Polskiej Sekcji, a także zasygnalizowanie działań na przyszłość.

ZASTOSOWANIE ŚRODOWISKA OBLICZENIOWEGO MATLAB DO SYNTEZY DŹWIĘKU PRZESTRZENNEGO Z WYKORZYSTANIEM ZINDYWIDUALIZOWANYCH POMIARÓW HRTF

Zbigniew ŚWIĘTACH [1]
Przemysław PLASKOTA [1]

[1] Politechnika Wrocławska,Wyspiańskiego 27, 50-370 Wrocław
przemyslaw.plaskota@pwr.wroc.pl
zbigniew.swietach@pwr.wroc.pl

WPROWADZENIE

W niniejszej pracy przedstawiono wyniki syntezy dźwięku przestrzennego na podstawie zadanego monofonicznego sygnału dźwiękowego. W tym celu wykorzystano zbiór przestrzennie skorelowanych odpowiedzi impulsowych lub równoważnie transmitancji przestrzennych mierzonych dookólnie względem środka głowy potencjalnego słuchacza (HRTF – head related transfer functions) [1], [2]. We wzmiankowanej metodzie zakłada się liniowy model propagacji fali akustycznej w przestrzeni otaczającej słuchacza, przy czym fizyczne nieograniczone czasowo odpowiedzi impulsowe przybliża się odpowiednio dobranymi skończonymi ciągami próbek (dyskretyzacja odpowiedzi impulsowych).

Do celów syntezy dźwięku przestrzennego wykorzystano środowisko obliczeniowe Matlab. Wybór Matlaba podyktowany jest prostotą i przejrzystością kodu źródłowego, który jest plikiem tekstowym. Ponadto sposób zapisu problemu technicznego czy matematycznego w Matlabie jest wyjątkowo intuicyjny i niewiele odbiega od zapisu takiego problemu przy użyciu standardowej notacji matematycznej.

SYNTEZA DŹWIĘKU PRZESTRZENNEGO

Przestrzenne odpowiedzi impulsowe HRTF zostały uprzednio zmierzone i były dostępne w formacie xml [4]. Ze względu na dużą ilość danych niezbędnych do przeprowadzenia syntezy dźwięku wybrano dwie metody analizy. Pierwsza metoda polega na wykorzystaniu pomiarów HRTF pochodzących od jednej osoby. W drugiej metodzie uśredniono pomiary HRTF pochodzące od kilku osób. Rozważa się syntezę dźwięku przestrzennego w przypadkach, gdy źródło dźwięku przemieszcza się wirtualnie, jak i gdy źródło dźwięku znajduje się w określonym położeniu względem słuchacza.

Jeżeli źródło dźwięku przemieszcza się wirtualnie względem słuchacza, wówczas w obydwu metodach istotna jest długość ramki czasowej pomiędzy dwoma kolejnymi zmianami położenia źródła dźwięku. Wirtualny ruch źródła dźwięku realizowany jest przez splatanie sygnału monofonicznego źródła dźwięku z kolejno wybraną odpowiedzią impulsową ze zbioru dostępnych HRFT. W pracy [3] rozważa się możliwość predykcji długości wzmiankowanej ramki czasowej, jednak w chwili obecnej nie udało się skutecznie wykorzystać tej możliwości. Będzie to przedmiotem dalszych prac. Aktualnie ustala się arbitralnie długość ramki czasowej dla całego procesu syntezy dźwięku przestrzennego. Podczas konferencji dostępne będą przykłady omawianej syntezy dźwięku przestrzennego w formie plików *.wav.

LITERATURA

[1] ANDRZEJ DOBRUCKI, PRZEMYSLAW PLASKOTA, PIOTR PRUCHNICKI, MICHAL PEC, MICHAL BUJACZ, PAWEL STRUMILLO. *Measurement System for Personalized Head-Related Transfer Functions and Its Verificationby Virtual Source Localization Trials with Visually Impaired and Sighted Individuals*. J. Audio Eng. Soc., Vol. 58, No. 9, 2010 September.
[2] COREY I. CHENG, AND GREGORY H. WAKEFIELD. *Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space*. J Audio Eng Soc, Vol 49, No 4, 2001 April.
[3] FLORIAN PAUSCH, SHAIMA'A DOMA, JANINA FELS. *Hybrid multi-harmonic model for the prediction of interaural time differences in individual behind-the-ear hearing-aid-related transfer functions.* Acta Acustica 2022, 6, 34.
[4] PRZEMYSŁAW PLASKOTA, JAKUB STASIAK. *Baza danych zawierająca wyniki pomiarów hrtf w formacie xml*. Konferencja Postępy Akustyki, Gliwice 2017.

WYBRANE ASPEKTY CHARAKTERYSTYK KIERUNKOWOŚCI GŁOŚNIKÓW MODÓW ROZPROSZONYCH

Karol CZESAK[1]
Piotr KLECZKOWSKI[1]
[1]Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, 30-059 Kraków,
al. Mickiewicza 30, czesak@agh.edu.pl

1. WPROWADZENIE

Głośniki modów rozproszonych (ang. Distributed Mode Loudspeakers – DML) charakteryzują się znacząco odmiennymi właściwościami od tych, opisujących głośniki tłokowe. Dzieje się tak, gdyż założenia konstrukcyjne tych przetworników w istotny sposób odbiegają od założeń przyświecających konstrukcji głośników tłokowych. Także sposób oscylacji DML jest zupełnie inny niż w przypadku konwencjonalnego przetwornika. Głośniki DML generują ciśnienie akustyczne poprzez wprawienie powierzchni głośnika (w większości przypadków – prostokątnej płyty) w drgania giętne [1]. Ten sposób pracy przekłada się na dość oryginalne charakterystyki amplitudowo – częstotliwościowe uzyskiwane w czasie pomiarów charakterystyk amplitudowo – częstotliwościowych DML na półsferze. Ich wyróżnikiem są znaczące minima lokalne, jakich nie zwykliśmy obserwować przy pomiarach charakterystyk amplitudowo – częstotliwościowych przetworników tłokowych. Zależnie od wyboru punktu pomiarowego na półsferze, te minima pojawiają się dla różnych częstotliwości.

Wbrew przewidywaniom, wiązka promieniowania DML nie ulega zwężeniu wraz ze wzrostem częstotliwości pobudzenia. Zauważalne jest jednak zjawisko pojawiania się listków bocznych, co zaczyna być widoczne już od częstotliwości 500 Hz. Poniżej tej częstotliwości, można uznać charakterystykę kierunkowości DML za dookólną. Wraz ze wzrostem częstotliwości pobudzenia, ilość zauważalnych listków bocznych w wiązce promieniowania zwiększa się. Również największa skuteczność DML nie jest notowana na osi głośnika, dlatego też wynik pomiaru poziomu ciśnienia akustycznego w osi przetwornika nie powinien być uznawany za referencyjny przy normalizacji pomiarów wykonanych z innych kierunków [2].

2. PROCEDURA POMIARU

Referat przedstawia skrócony wyciąg z badań charakterystyk amplitudowo-częstotliwościowych głośników modów rozproszonych w funkcji kąta azymutalnego oraz elewacji, w warunkach bezechowych. Przebadanych zostało sześć egzemplarzy przetworników DML. Trzy z jednym wzbudnikiem oraz trzy z dwoma. Każdy z nich został poddany wymuszeniu sinusem przestrajanym liniowo, z mocą równą 1 W. Seria pomiarowa dla każdego z egzemplarzy obejmowała 325 punktów rozlokowanych na półsferze ze stałą rozdzielczością kątową, równą 10 stopni. Precyzyjne zadawanie kąta azymutalnego umożliwił stolik obrotowy sterowany silnikiem krokowym. Kąt elewacji zadawany był poprzez ruch zrobotyzowanego ramienia - na którym zamocowany był mikrofon pomiarowy - wykonywany po łuku. Aparatura pracowała z częstotliwością próbkowania 96 kHz oraz rozdzielczością bitową równą 24 bitom. Dzięki takiej konfiguracji możliwe stało się prowadzenie pomiarów w zakresie częstotliwości sięgającym 24 kHz oraz wyeliminowanie artefaktów wynikających z działania filtracji antyaliasingowej.

Dzięki liniowemu sposobowi przestrajania sinusoidy wymuszającej drgania DML, była możliwa dokładna rejestracja uzyskiwanych przy ich pomocy charakterystyk amplitudowo – częstotliwościowych,

wiarygodnie oddająca lokalizacje minimów w widmie. Tak zarejestrowane odpowiedzi głośników zostały poddane okienkowaniu oraz analizie fourierowskiej. Zastosowano również poprawki korekcyjne, wynikające z analizy charakterystyk amplitudowo – częstotliwościowych urządzeń składających się na tor pomiarowy, wykorzystany przy wyznaczaniu odpowiedzi częstotliwościowej DML w komorze bezechowej (wzmacniacz mocy, przetworniki A/C i C/A, urządzenia zapewniające separację galwaniczną elementów toru).

Na podstawie zbioru charakterystyk amplitudowo – częstotliwościowych przetworników, sporządzonych w powtarzalny sposób, możliwe stało się wykreślenie ich charakterystyk kierunkowości dla konkretnych częstotliwości pobudzenia, dla trzech egzemplarzy DML. Arbitralnie zostały wybrane częstotliwości, zwyczajowo podawane jako częstotliwości środkowe pasm tercjowych, co dało po 25 „balonów kierunkowości" na jeden egzemplarz głośnika. Przy prezentacji danych pominięto najniższe częstotliwości, tj. poniżej 80Hz, co ma związek z gwałtownie spadającą skutecznością badanych DML poniżej tej częstotliwości oraz ryzyko uszkodzenia przetwornika przy wzbudzeniu sygnałem o częstotliwości niższej niż 70 Hz.

3. OBSERWACJE I WNIOSKI

„Balony kierunkowości" wsparte zostały również wykresami dwuwymiarowymi, prezentującymi skuteczność głośnika dla konkretnej częstotliwości w funkcji kąta elewacji, gdzie dla każdego z kątów elewacji zostały uśrednione w sposób średniokwadratowy wyniki pomiarów z wszystkich punktów o zadanej elewacji. Inne wykresy dwuwymiarowe obrazują skuteczność DML dla zadanej częstotliwości i elewacji, jednak w funkcji kąta azymutalnego.

Na uwagę zasługuje pewna asymetryczność uzyskiwanych charakterystyk kierunkowości oraz niewielkie różnice pomiędzy wynikami uzyskanymi dla poszczególnych egzemplarzy badanych urządzeń.

Znajomość charakterystyk kierunkowości DML w funkcji częstotliwości umożliwia wyliczenie dla nich indeksu kierunkowości oraz określenie a priori subiektywnej głośności uzyskiwanej z ich pomocą w realnym pomieszczeniu [3]. W obliczu zaobserwowanych cech charakterystycznych DML pojawia się jednak pytanie, na ile dotychczasowo znane i stosowane narzędzie do opisu przetworników elektroakustycznych są adekwatne do przetworników typu DML.

LITERATURA

[1] HARRIS N., Spatial Bandwidth of Diffuse Radiation in Distributed-Mode Loudspeakers , w: 111th AES Convention, 2001 September 21–24 New York, NY, USA, AES, USA, 2001

[2] GONTCHAROV V. P., HILL N. P. R., TAYLOR V. J., Measurement Aspects of Distributed Mode Loudspeakers, w: 106th AES Convention, 1999 May 8–11 Munich, Germany, AES, USA, 1999

[3] HARRIS N., FLANAGAN S., Loudness – a study of the subjective Loudness Difference between DML and Conventional Loudspeakers, w: 106th AES Convention, 1999 May 8–11 Munich, Germany, AES, USA, 1999

MICROPHONE VERSUS LASER DISPLACEMENT SENSOR IN MEASURING THE DIAPHRAGM MOVEMENT OF DYNAMIC LOUDSPEAKERS

Michał KMIECIK[1]
[1]AGH Akademia Górniczo-Hutnicza im. S. Staszica w Krakowie,
al. A. Mickiewicza 30,30-059 Kraków
miszkoo@gmail.com

The distortions generated by the loudspeaker are largely due to the processing of electrical signal into mechanical movement. However, displacement of the loudspeaker coil does not exactly match the changes in acoustic pressure and velocity. Therefore, recording movement of the loudspeaker coil based on the generated acoustic signal is subject to significant errors, particularly in the low-frequency range. Obtaining information about constant displacement related to e.g. reluctance force is impossible. Contactless measurement of speaker diaphragm movement is problematic and requires an optical method. A high-speed laser displacement sensor is needed to provide a wide band of the measurement, which is then limited, however, by the sensor's amplitude range and accuracy.

The paper presents a comparison of the displacement of the speaker diaphragm measured with the laser sensor and the displacement calculated on the basis of the acoustic signal recorded with a microphone. It discusses the problem of the measurement range and the error related to nonlinearity (sensitivity) of the sensor as a function of frequency.

ANALIZA WPŁYWU WYBRANYCH TECHNIK KODOWANIA NA OCENĘ JAKOŚCI RÓŻNYCH GATUNKÓW MUZYKI

Stefan BRACHMAŃSKI[1]
Maurycy KIN[2]
Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
[1]stefan.brachmanski@pwr.edu.pl, [2]maurycy.kin@pwr.edu.pl

1. WPROWADZENIE

Rozwój technologii cyfrowych dostarczył nowych możliwości dla przetwarzania sygnałów fonicznych, a najważniejszą zaletą sygnałów cyfrowych jest łatwość ich przechowywania, przesyłania i przetwarzania. Wciąż jednak istnieją ograniczenia związane z magazynowaniem oraz przepustowością kanałów transmisyjnych. W wielu przypadkach wskazane jest, aby sygnały były zakodowane jak najwydajniej. Do popularnych metod kodowania można zaliczyć kodowanie mp3, stosowane powszechnie w Internecie, oraz AAC oraz HE-AAC, które znalazły zastosowanie w radiofonii cyfrowej i telewizji satelitarnej. Wymienione metody należą do grupy metod wykorzystujących kodowanie stratne, związane z utratą części informacji [1,2], wykorzystujące w algorytmie właściwości psychoakustyczne.

Podstawowymi parametrami, mającymi wpływ na jakość sygnału podlegającemu procesowi kompresji stratnej są rodzaj kodowania oraz szybkość bitowa, od której zależy objętość przesyłanej informacji [3]. Głównym celem pracy jest zbadanie jaki wpływ na subiektywną ocenę jakości różnych gatunków muzyki ma rodzaj i szybkość bitowa kodowania. Przeprowadzono eksperyment, w ramach którego słuchacze oceniali ogólną jakość dźwięku oraz dwa atrybuty wrażenia słuchowego: barwę dźwięku i przestrzenność nagrań z zakresu sześciu różnych gatunków muzyki.

2. PRZEPROWADZONE BADANIA

W eksperymencie wzięło udział 21 osób. Znaczącą większość, (15 osób), czyli 71,4% to osoby z przedziału wiekowego 18 - 26 lat, 4 osoby z przedziału wiekowego 27 - 35 lat. Najmniej liczną grupą byli niepełnoletni oraz osoby powyżej 36 roku życia; obie grupy stanowiły 4,8% wszystkich uczestników. Na pytanie związane z grą na instrumentach muzycznych 15 osób odpowiedziało, że nie gra na żadnym instrumencie natomiast 6 osób ma doświadczenie muzyczne w postaci gry na instrumentach.

Ponieważ badania dotyczyły oceny jakości różnych gatunków muzycznych, postanowiono sprawdzić, jakie są ulubione gatunki u słuchaczy. Okazało się, że słuchacze preferują niemal w równym stopniu muzykę pop, bluesa, rock czy hip-hop. Zdecydowanie najmniejszą popularnością cieszyła się muzyka klasyczna oraz jazz.

Badano dwie wartości szybkości bitowych: 64 oraz 128 kb/s. Oprogramowaniem, jakie użyto do przygotowywania próbek testowych był GX-Transcoder, natomiast edycji długości materiału dokonywano za pomocą edytora Samplitude v.8. Odsłuchu badanych próbek dźwiękowych słuchacze dokonywali w warunkach domowych, przy pomocy formularza Google. Ocenę ogólnej jakości dźwięku oraz badanych atrybutów dokonano przy zastosowaniu metody ACR (Absolute Category Rating) z pięciostopniową skalą ocen [4].

3. WYNIKI BADAŃ

Na podstawie otrzymanych wyników można stwierdzić, że dla przepływności 64 kb/s żaden z badanych gatunków nie został oceniony jako dobry (MOS ≥4) tak pod względem ogólnej oceny jakości, jak i pod kątem badanych atrybutów. Należy zaznaczyć, że dla tej przepływności najwyższe oceny uzyskano dla gatunku Hip hop, co może świadczyć o sposobie produkcji oraz zgrania materiału, gdzie stosunkowo najmniej jest przestrzeni, a wszystkie elementy dźwiękobrazu usytuowane są blisko środka panoramy oraz poddane mocnej kompresji, co skutkuje małym wrażeniem głębi obrazu.

Najniższe oceny otrzymano dla próbek muzyki klasycznej: przy przepływności 64 kb/s wartości MOS zawierają się w przedziale 2÷2,5, natomiast dla przepływności 128 kb/s – 3,1÷3,7. Jest to niewątpliwie związane z większym znaczeniem i udziałem przestrzenności oraz większą wrażliwością słuchaczy na zmiany barwy dźwięku instrumentów naturalnych.

Uzyskane wyniki potwierdzają wpływ technik kodowania na ocenę jakości brzmienia utworu muzycznego dla każdego badanego gatunku [5]. Zastosowanie największej możliwej przepływności w przypadku badanych gatunków muzycznych nie zawsze wpływa na uzyskanie najlepszej barwy dźwięku, czy ogólnej jakości dźwięku. Optymalną przepływnością dla każdego badanego gatunku jest 128 kb/s, ponieważ w zdecydowanej większości przypadków daje ona zadowalające rezultaty.

LITERATURA

[1] SAYOOD K., Introduction to Data Compression, 4th Edition, Elsevier, Waltham, MA,2021.
[2] BOSI M., GOLDBERG R., Introduction to Digital Audio Coding and Standards, SPRINGER, 2002.
[3] BRANDENBURG K., Mp3 and AAC explained, Fraunhofer Institute for Integrated Circuits FhG-IIS A, Erlangen, Germany.
[4] ITU-T P.800, Methods for objective and subjective assessment of quality.
[5] RONAN M., WARD N., SAZDOV R., LEE H., The Perception of Hyper-compression by Mastering Engineers, J. Audio Eng. Soc., vol.65, No.7/8, 2017, s. 613 – 621.

SUBIEKTYWNA OCENA JAKOŚCI SYGNAŁU VIDEO KODOWANEGO W STANDARDACH H.264 I H265

Stefan BRACHMAŃSKI[1]
Michał ŁUCZYŃSKI[2]
Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
[1]stefan.brachmanski@pwr.edu.pl, [2]michal.luczynski@pwr.edu.pl

1. WPROWADZENIE

W ostatnich latach można zauważyć wśród młodego pokolenia rosnące zainteresowanie przekazami m.in. filmów za pośrednictwem Internetu. Przewaga Internetu nad telewizją wynika przede wszystkim z tego, że mając na przykład smartfona wszystko jest dostępne z dowolnego miejsca, a oglądanie programów nie zależy od ramówki danego kanału.

Przekaz video realizowany jest z wykorzystaniem różnych standardów kodowania, International Telecommunication Union (ITU) zaleca korzystanie ze standardów H.264 [3] i nowszego H.265 [4]. Na jakość transmisji video wpływ mają: szybkość transmisji, a także rozdzielczość obrazu.

Głównym celem pracy było zbadanie wpływu: kodowania H.264, H.265, szybkości bitowej (od 300 kb/s do 6000kb/s) oraz rozdzielczości (640 x 360, 1280 x 720 i 1920 x 1080) na ocenę jakości sygnału video przez młodego użytkownika końcowego.

2. EKSPERYMENT

Spośród różnych metod oceny jakości video [1], [40], [5], [6], [7] w badaniach zastosowano metodę Double Stimulus Impairment Scale Method (DSISM) [5]. Pomiar polega na porównaniu wzorcowego sygnału video o wysokiej jakości (prezentowany jako pierwszy) z sygnałem ocenianym (sygnał drugi) i określeniu stopnia pogorszenia jakości sygnału badanego (drugiego) w odniesieniu do sygnału wzorcowego (pierwszego) w pięciostopniowej skali ocen. Według tej skali ocena 5 -oznacza pogorszenie jakość niedostrzegalne, a 1 – bardzo dokuczliwe [5].

Wzorcowym materiałem testowym była 20 sekundowa sekwencja video (bez dźwięku) o rozdzielczości 1920 x 1080 w formacie avi zawierająca sceny dynamiczne – fragment startu wyścigów konnych. Ta sekwencja video została poddana kodowaniu H.264 i H265 z różnymi szybkościami bitowymi i różną rozdzielczością. Zastosowano trzy rozdzielczości: 640 x 360, 1280 x 720 i 1920 x 1080. Dla danej techniki kodowania i określonej rozdzielczości różne warunki transmisji symulowane były osiemnastoma szybkościami bitowymi: 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000 kb/s. Tak przygotowany materiał testowy prezentowany był widzom w laboratorium zaadoptowanym do oceny sygnałów video na ekranie telewizora o przekątnej 60 cali. Prezentacja sygnałów testowych o różnych warunkach transmisji (szybkość bitowa) dokonywana była w sposób losowy.

Zespół obserwatorów tworzyli studenci Politechniki Wrocławskiej w wieku 20 – 21 lat o prawidłowej ostrości widzenia i poprawnym rozróżnianiu kolorów. Zgodnie z zaleceniem International Telecommunication Union BT. 500-14 [1] minimalna liczba obserwatorów powinna wynosić 15. W badaniach utworzono 2 grupy. Każda grupa oceniała inny typ kodowania. Liczebność poszczególnych grup była różna i wynosiła dla kodowania H.264 – 45 osób, a dla

H.265 – 35 osób. Przed rozpoczęciem pomiarów uczestnicy zapoznali się z metodą oceny i odbyli jedną sesję treningową. Podczas treningu obserwatorzy zapoznali się ze sposobem prezentacji materiału testowego i oceny zmian jakości video

Subiektywna ocena jakości video dostarcza dużej ilości danych (ocen), które muszą być poddane obróbce statystycznej, pozwalającej wyeliminować oceny wykraczające poza przyjęty przedział ufności. Wyniki poddano analizie statystycznej zgodnie z procedurą z zalecenia ITU-R BT.500 [1]

3.PODSUMOWANIE

Ocena jakości sygnału video dokonana przez młodego widza potwierdza teoretyczne sugestie o większych możliwościach standardu H.265 w stosunku do H.264 zarówno w odniesieniu do parametrów kompresji jak i jakości kompresowanego sygnału video. Standard H265 dla sygnału video Full HD pozwala otrzymać bardzo dobrą jakość (MOS = 4,5) począwszy od szybkości bitowej wynoszącej 3000 kb/s. Dla standardu H.264, ocenę MOS = 4,5 uzyskuje się dla szybkości powyżej 3500 kb/s.

Przyjmując za dopuszczalną ocenę MOS = 4,0 oznaczającą zauważalne, lecz niedokuczliwe pogorszenie jakości można stwierdzić, że standard H.265 zapewnia taką ocenę już dla szybkości 1500 kb/s, natomiast H.264 dla szybkości powyżej 2000 kb/s.

LITERATURA

[1] ITU-R Recommendation BT. 500-14, *Methodologies for the subjective assessment of the quality of television images* International Telecommunication Union, 2019

[2] ITU-R Recommendation BT. 1129-2, *Subjective assessment of standard definition digital television (SDTV) systems*, International Telecommunication Union, 2019

[3] ITU-T Recommendation H.264, *Advanced video coding for generic audiovisual services,* International Telecommunication Union, 2021.

[4] ITU-T Recommendation H.265, *High efficiency video coding,* International Telecommunication Union, 2021

[5] ITU-T Recommendation P.910, *Subjective video quality assessment methods for multimedia applications,* International Telecommunication Union, 1998

[6] PINSON M., WOLF S., *Comparing subjective video quality testing methodologies*, Visual Communications and Image Processing 2003. SPIE, 2003. p. 573-582.

[7] WINKLER S., *Video quality measurement standards – current status and trends*, 2009 7th International Conference on Information, Communications and Signal Processing (ICICS). IEEE, 2009. p. 1-5.

UPROSZCZONA METODA POMIARU PRZESTRZENNYCH ODPOWIEDZI IMPULSOWYCH I JEJ WYKORZYSTANIE DO OCENY JAKOŚCI AKUSTYCZNEJ POMIESZCZEŃ I GENEROWANIA POGŁOSU SURROUND

Witold MICKIEWICZ, Grzegorz PAWEŁKIEWICZ, Kaja KOSMENDA

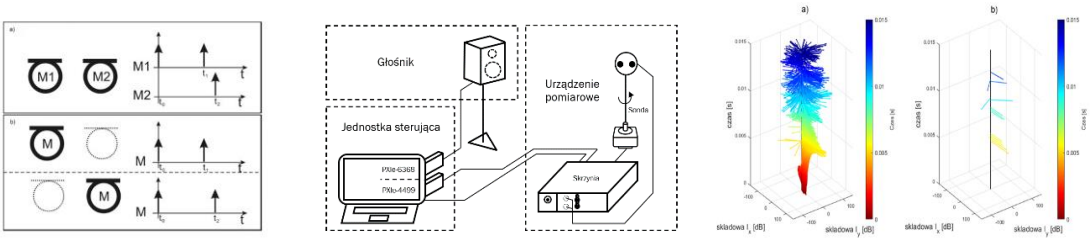
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Al. Piastów 17, 70-310 Szczecin

1. WPROWADZENIE

Zapewnienie odpowiednich właściwości akustycznych pomieszczeń to zagadnienie ciągle aktualne i dotyczy nie tylko pomieszczeń o tzw. akustyce kwalifikowanej. Ciągle istnieją braki w poprawnym modelowaniu zjawisk wpływających na percepcję dźwięku w tego typu obiektach i mówi się, że projektowanie sal koncertowych itp. to swego rodzaju sztuka. Badanie akustyki wnętrz w pomieszczeniach istniejących w celu eliminacji wad akustycznych oraz modelowanie propagacji dźwięku w pomieszczeniach projektowanych, gdzie na propagację dźwięku wpływa wiele detali architektonicznych jest zagadnieniem ciągle aktualnym i ważnym.
Ciśnienie akustyczne jest łatwe w pomiarze i dlatego we współczesnej akustyce technicznej większość parametrów bazuje na pomiarze tej wielkości skalarnej i nie zawiera informacji o kierunku przepływu energii akustycznej. Mimo to, ze względu na wspomnianą łatwość pomiaru ciśnienia akustycznego za pomocą mikrofonu, na przestrzeni lat zdefiniowano wiele obiektywnych parametrów akustyki pomieszczeń na bazie tzw. odpowiedzi impulsowej pomieszczenia, czyli czasowego przebiegu zmian ciśnienia akustycznego mierzonego w określonym punkcie pomieszczenia pobudzonego sygnałem impulsowym. Mimo ich użyteczności posiadają jednak pewną wadę. Na wyznaczane wielkości nie ma wpływu kierunek, z którego rejestrowana fala akustyczna przybywa w kolejnych chwilach czasu, gdyż odpowiedzi impulsowe mierzone są za pomocą dookólnego mikrofonu ciśnieniowego.
Akustyka pomieszczeń jest ściśle związaną z percepcją dźwięku przez człowieka, który posiada dwoje uszu i słyszy przestrzennie. Ważnym więc aspektem, który należy wziąć pod uwagę przy modelowaniu i pomiarach jest kierunek propagacji dźwięku. Wielkością fizyczną, która dostarcza informacji o kierunku przepływu dźwięku jest natężenie dźwięku. Dzięki zarejestrowaniu odpowiedzi impulsowej pomieszczenia, która ukazuje przebieg wartości chwilowych natężenia dźwięku w czasie, można w znacznie szerszym stopniu analizować badane pomieszczenie. Pomiar natężenia dźwięku jest bardziej złożony i wymaga zastosowania specjalistycznych czujników – tzw. sond natężeniowych. W zależności od konstrukcji składają się one z kilku (co najmniej dwóch) precyzyjnie dopasowanych mikrofonów ciśnieniowych, lub z jednego mikrofonu i czujników prędkości akustycznej. Ze względu na detale konstrukcyjne takie rozwiązania są drogie, co bezpośrednio wpływa na to, że są mało rozpowszechnione. Poniższa praca aspiruje do bycia przyczynkiem w rozpowszechnianiu się stosowania metod natężeniowych w akustyce pomieszczeń. Autorzy chcą przybliżyć uproszczoną tanią metodę mierzenia przestrzennych natężeniowych odpowiedzi impulsowych, zaproponować pewne sposoby wizualizacji wyników oraz na bazie przeprowadzonych eksperymentów wskazać obszary potencjalnego zastosowania metody w akustyce architektonicznej i technologii nagrań dźwiękowych.

2. UPROSZCZONA METODA POMIARU NATĘŻENIA DŹWIĘKU

Wykorzystanie sondy dwumikrofonowej umożliwia rejestrację przebiegu pojedynczej składowej natężenia dźwięku dla dowolnego źródła dźwięku. W przypadku akustyki architektonicznej pomieszczenie pobudza się znanym sygnałem i rejestruje odpowiedź układu na to pobudzenie. Dzięki współczesnej technologii można powtarzać taki eksperyment pomiarowy wielokrotnie przy zachowaniu pełnej powtarzalności wymuszenia i synchronizacji akwizycji. W tej specyficznej sytuacji przy założeniu stabilności parametrów termodynamicznych obiektu (problemy dla wyższych częstotliwości) oraz liniowości badanego pola akustycznego (poziomy ciśnień nieprzekraczające 120 dB SPL) pomiar natężenia dźwięku może zostać uproszczony i wykonany przy użyciu jednego mikrofonu, który w kolejnych pomiarach będzie zmieniał swoje położenie (zgodnie z geometrią sondy wielomikrofonowej), mierząc tym samym przebieg ciśnienia akustycznego w kilku pozycjach. W dalszej części artykułu przedstawiono przykłady zautomatyzowanych pomiarów 2D, a więc wymagających 4 pozycji pomiarowych do otrzymania rozkładu wektora natężenia dźwięku na płaszczyźnie.



Rys.1. Idea metody, kompletny system pomiarowy, wizualizacja odpowiedzi 2D + czas

2. PODSUMOWANIE

Zaproponowana w artykule metoda pomiarowa oraz przedstawione przykłady jej wykorzystania mają na celu wzbudzić na nowo poszukiwania w znajdowaniu korelacji między obiektywnymi wskaźnikami bazującymi na akustyce wektorowej i subiektywnymi ocenami jakości akustycznej pomieszczeń. Zastąpienie analizy czasowej ciśnieniowych odpowiedzi impulsowych analizą czasowo-przestrzenną odpowiedzi natężeniowych wymaga opracowania nowych metod wizualizacji. Jedną z takich metod zaproponowano w pracy. Na przykładzie współczynnika odbić bocznych pokazano, że rejestracja odpowiedzi kierunkowej daje większe możliwości analizy przestrzennych właściwości pola akustycznego w fazie opracowywania danych. Jest to również podstawa to kreowania wielokanałowych odpowiedzi impulsowych wykorzystywanych w efektach pogłosowych surround. Podsumowując powinniśmy chętniej śledzić i brać pod uwagę w pomiarach akustycznych metody czasowo-przestrzenne. Naszym zdaniem drzemie w nich nieodkryty jeszcze potencjał, dzięki któremu kształtowanie akustyczne pomieszczeń oraz akustyka wirtualna może być w przyszłości jeszcze bardziej efektywna.

ZADANIA SŁUCHOWE W GRZE KOMPUTEROWEJ WSPOMAGAJĄCEJ ROZWIJANIE UMIEJĘTNOŚCI W ZAKRESIE SOLFEŻU BARWY

Paulina BIELESZ[1]
[1]Uniwersytet Śląski w Katowicach
paulina.bielesz@gmail.com

Gra komputerowa o roboczej nazwie Sound Jobs jest grą edukacyjną posiadającą wszelkie cechy gry czyli grafikę, fabułę, rywalizację. W pracy nad projektem podjęłam próbę uprzystępnienia i zaadoptowania ćwiczeń z zakresu solfeżu barwy do świata wirtualnego. Celem badania stanowi prototyp takiej gry.

Od pewnego czasu interesuję się zagadnieniami związanymi z solfeżem barwy. Jako że zajmuję się również grami komputerowymi, szczególnie tematem warstwy dźwiękowej do gier postanowiłam zrealizować program, grę komputerową opartą na ćwiczeniach zawierających się w temacie solfeżu barwy. Wraz z programistą i grafikiem projektujemy innowacyjne narzędzie, środowisko komputerowe oparte na silniku gier Unity, współpracującym z dźwiękowym silnikiem FMOD, który umożliwia kształtowanie barwy dźwięku. Edukacyjna gra, operująca spójnym, zakomponowanym światem dźwięków ma być również narzędziem dydaktycznym, wspomagającym rozwijanie umiejętności w zakresie solfeżu barwy posiadającym wszelkie cechy gry komputerowej, czyli grafikę, fabułę i rywalizację. W przebiegu projektu został stworzony Game Design Document, czyli dokument opisujący fabułę gry, opis poziomów, mechanikę, przedstawiający grafikę i opis warstw dźwiękowych projektu, na podstawie, którego została zbudowana gra. Przeprowadziłam również analizę dostępnych silników gier pod kątem kształtowania barwy dźwięku.

Ćwiczenia z zakresu solfeżu barwy mają na celu doskonalenie umiejętności identyfikacji i dyskryminacji zmian barwy dźwięku. Zadania utrwalają charakterystyczne brzmienia dźwięku w pamięci długotrwałej. Barwę dźwięku można kształtować poprzez użycie poszczególnych urządzeń/ pluginów takich jak korektor, kompresja dynamiki, użycie różnych efektów, efekty pogłosowe takie jak reverb oraz delay a także zmiany w obrazie stereo. Aby prawidłowo używać tych narzędzi należy poznać efekty jakie możemy uzyskać dzięki nim. Umiejętność usłyszenia zmian w bazie stereo, wszelkich nieprawidłowości, jest istotną umiejętnością potrzebną realizatorowi dźwięku czy też kompozytorowi.

Fabuła gry jest oparta na prawdziwej historii. Czas akcji gry umiejscowiony został w latach 70, kiedy na świecie nie było telefonów komórkowych. 60 % społeczeństwa posiadało telefony stacjonarne - choć rozmowy telefoniczne były dostępne, nie były takie tanie! Gracz wciela się w postać młodego konstruktora, który gdzieś w dolinie Krzemowej w zaciszu swojego pokoju, tworzy urządzenie które zrewolucjonizuje połączenia telefoniczne BlueBox - dzięki któremu można zrealizować połączenie telefoniczne do każdego abonenta, nawet do prezydenta, za darmo – musi tylko rozpoznać odpowiedni rodzaj sygnału, rozpoznać konkretną zmianę barwy dźwięku.

W edukacji warto skorzystać z gamifikacji, inaczej zwaną gryfikacją. To świadome i celowe wykorzystanie mechanizmów stosowanych przy projektowaniu gier. Ma na celu aktywizowanie oraz motywowanie grupy uczestników/ studentów. Gamifikacja w edukacji zakłada użycie jednego bądź kilku elementów gry w procesie nauki. Postrzegana jest jako zbiór strategii, taktyk i produktów, których zastosowanie pozwala na osiągnięcie celów edukacyjnych. Gra jest zamkniętą całością, tzn. ma początek, przebieg i zakończenie. Głównym zadaniem wykorzystania

gry w procesie uczenia jest osiągnięcie wzrostu motywacji i zaangażowania. Dodatkowo mechanizmy, które są wykorzystywane w grach, mogą wpłynąć na stymulację kreatywności, samozaparcia czy też systematyczności u wielu osób biorących czynny udział w procesie nauki. Wyższa skuteczność uczenia się może być osiągana dzięki zwiększeniu atrakcyjności procesu dydaktycznego.

Prototyp gry składa się z okna menu, intro wprowadzającego do historii oraz ekranu gry wraz z tutorialem. Ekran gry na ten moment zawiera 5 etapów. Pierwsze 4 etapy dotyczą rozpoznawania formantów, częstotliwości wybierane są losowo z następujących wartości: 63, 125, 250, 500, 1k, 2k, 4k, 8k, 16 kHz. Piąty etap poświęcony jest wybranym efektom takim jak kompresja, zniekształcenia, pogłos, delay, zmiany bazy stereo.

Opis poszczególnych etapów (pięciu domków):

1.Filtr pasmowo przepustowy (szum różowy i przykłady muzyczne)

2.Filtr dolnoprzepustowy (szum różowy i przykłady muzyczne)

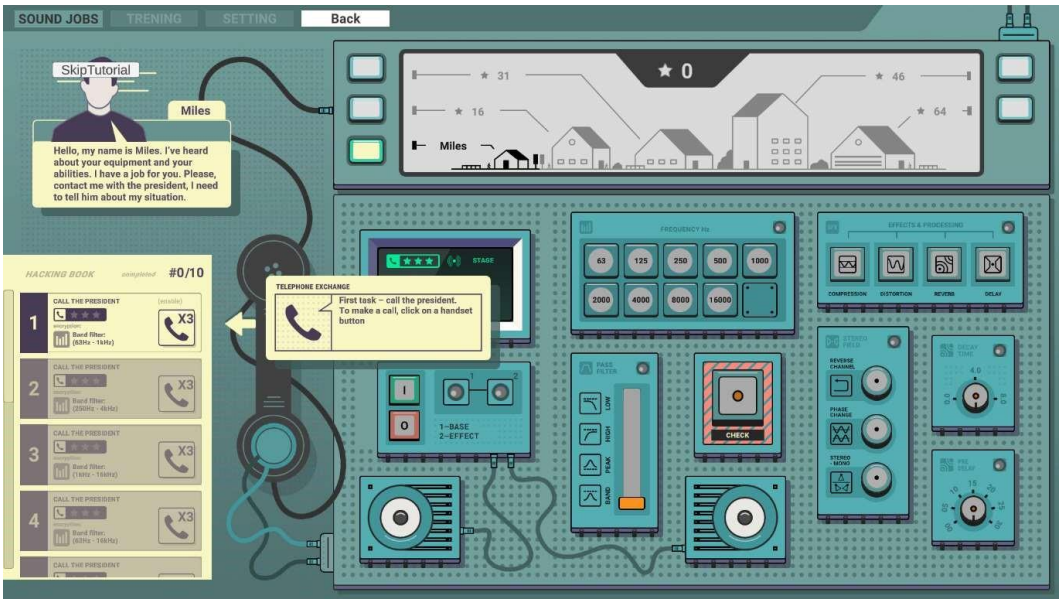
3.Filtr górnoprzepustowy (szum różowy i przykłady muzyczne)

4.Filtr pasmowy (szum różowy i przykłady muzyczne)

5.Efekty (kompresja, distortion, reverb, delay, zmiany stereo (zamiana kanałów lewego i prawego, zmiana fazy i stereo do mono))

Projekt w dalszym ciągu jest kontynuowany i rozwijany.

Okno menu gry.



Okno programu

ADAPTACYJNY SYSTEM KSZTAŁTOWANIA WIĄZKI W POLU BLISKIM W OPARCIU O UCZENIE MASZYNOWE

Agnieszka WIELGUS[1]
[1]Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
agnieszka.wielgus@pwr.edu.pl

1. WPROWADZENIE

Zagadnienie kształtowania wiązki (ang. *beamforming*) pojawia się w wielu rzeczywistych systemach: bezprzewodowej komunikacji, radarach, sonarach, macierzach mikrofonów i anten [2,3], zarówno po stronie nadawczej, jak i odbiorczej. W systemach, w których należy odebrać sygnał pożądany bardzo często mamy do czynienia z niepożądanym sygnałem lub sygnałami zakłócającymi. Jeżeli pasmo częstotliwość sygnału pożądanego pokrywa się częściowo lub w pełni z pasmem częstotliwości sygnału zakłócającego, wówczas zastosowanie klasycznych metod filtracji częstotliwości (ang. *classical temporal filtering*) bazujących na wzmacnianiu bądź tłumieniu konkretnych składowych nie jest możliwe. W takich przypadkach wykorzystuje się tak zwaną filtrację przestrzenną (ang. *spatial filtering*), która w oparciu o zjawisko interferencji sygnałów umożliwia rozdzielenie sygnałów pożądanego i zakłócającego.

Zastosowanie metod filtracji przestrzennej wymaga wykorzystania macierzy sensorów. W przypadku problemu odbioru sygnału z zakłóceniami, których źródło znajduje się poza obszarem, z którego pochodzi sygnał pożądany wykorzystywana jest macierz mikrofonów. Każdy z mikrofonów traktowany jest wówczas jako filtr (o skończonej bądź nieskończonej odpowiedzi impulsowej), którego współczynniki powinny zostać wyznaczone tak, aby spełnić zadane kryterium, które opisuje jakość systemu (np. maksymalizacja poziomu sygnału pożądanego, minimalizacja poziomu zakłóceń, bądź ich eliminacja).

Wspomniane podejście jest metodą dobrze znaną i doczekało się wielu modyfikacji [4, 5]. Jednakże, jeżeli położenie mówcy ulega zmianie, wówczas odpowiedź systemu (wartość funkcji celu) również może ulec znaczącej zmianie. Zatem na jakość systemu ma wpływ nie tylko konfiguracja macierzy i jej parametry, ale też położenie mówcy.

Dotychczasowe prace z obszaru optymalizacji macierzy mikrofonów koncentrowały się głównie na doborze odpowiedniej konfiguracji macierzy, jej rozmiaru oraz współczynników filtrów, wag poszczególnych sygnałów. W pracy [1] wykazano, że poszczególne elementy macierzy nie muszą być równomiernie rozmieszczone. Podejście takie umożliwia uzyskanie znacznie lepszych wyników niż klasyczne konfiguracje macierzy. Kolejno wykazano, że zwiększanie rozmiaru macierzy nie zawsze jest optymalnym podejściem [7]. Zastosowanie technik przerzedzania (ang. *thinning technique*) może znacząco poprawić jakość systemu.

W oparciu o powyższe rezultaty, w niniejszej pracy zaproponowano metodą, która wykorzystuje technikę przerzedzania dużej macierzy mikrofonów i doboru współczynników poszczególnych filtrów, tak, aby macierz ta adaptowała się do poruszającego się źródła sygnału pożądanego (mówcy). Podejście to wykorzystuje metodę uczenia maszynowego zwaną uczeniem ze wzmocnieniem (ang. *reinforcement learning*) [6] i nie było do tej pory rozpatrywane w literaturze. Rozważany problem polega na zaprojektowaniu adaptacyjnego system macierzy mikrofonów, która będzie dostosowywała zbiór aktywnych w danej chwili mikrofonów, tak aby wyjście systemu (jego odpowiedź) była jak najbardziej zbliżona do pożądanej w sensie normy $l2$.

LITERATURA

[1] Z. G. Feng, K. F. C. Yiu, and S. E. Nordholm, Placement design of microphone arrays in near-field broadband beamformers. IEEE Transactions on Signal Processing, 60(3): 1195–1204, 2012.
[2] D. H. Johnson, D. E. Dudgeon, *Array Signal Processing: Concepts and Techniqu*es. New York, NY, USA: Simon & Schuster, Inc., 1992.
[3] W. Liu, S. Weiss, *Design of frequency invariant beamformers for broadband arrays,* IEEE Transactions on Signal Processing, 56(2): 855–860, 2008.
[4]. R. A. Kennedy, D. B. Ward, and T. D. Abhayapala, *Nearfield beamforming using radial reciprocity*, IEEE Transactions on Signal Processing, 47(1): 33–40, 1999.
[5]. S. Nordebo, I. Claesson, S. Nordholm*, Weighted chebyshev approximation for the design of broadband beamformers using quadratic programming*. IEEE Signal Processing Letters, 1(7): 103–105 1994.
[6] Watkins, Christopher J.C.H. *Learning from Delayed Rewards* (PhD thesis). King's College, Cambridge, UK, 1989.
[7] A. Wielgus, B. Szlachetko, *A Simulation of Thinning of Microphone Array in Near-field Broadband Beamformers*, 32(2), 2021.

KOREKCJA PARAMETRÓW SYSTEMU DŹWIĘKOWEGO DOSTOSOWANA DO WARUNKÓW ATMOSFERYCZNYCH

Tomasz DUDA[1]
Marcin LEWANDOWSKI[2]
[1] tomasz.duda@post.pl
[2] Politechnika Warszawska, Wydział Elektroniki i Technik Informacyjnych,
ul. Nowowiejska 15/19, 00-665 Warszawa
marcin.lewandowski@pw.edu.pl

1. WPROWADZENIE

Warunki atmosferyczne mają znaczący wpływ na propagowanie się fali dźwiękowej w przestrzeni. Zmiana czynników atmosferycznych, takich jak temperatura, wilgotność oraz ciśnienie atmosferyczne powoduje, że fala dźwiękowa przemieszcza się z różną prędkością i zmiennym tłumieniem. Urządzenia, które automatycznie korygują parametry systemu dźwiękowego i dostosowują je do warunków atmosferycznych praktycznie nie istnieją.

2. PROJEKT URZĄDZENIA

Jednym z zadań prezentowanego urządzenia jest precyzyjne wyznaczenie prędkości dźwięku oparte na trzech parametrach: temperaturze, wilgotności oraz ciśnienia atmosferycznego. Urządzenie posiada trzy czujniki zintegrowane w jednym module do pomiaru ww. czynników atmosferycznych. Wyznaczana prędkość dźwięku służy do obliczania czasów opóźnień dla linii opóźniających znajdujących się w systemie nagłośnieniowym. Kolejnym zadaniem procesora jest kompensacja strat charakterystyki częstotliwościowej. W tym wypadku urządzenie niweluje skutki tłumienia fali w powietrzu i przywraca zamierzoną charakterystykę częstotliwościową. Do wyliczeń strat wynikających ze zmiany warunków atmosferycznych wykorzystane zostały dwie metody obliczeniowe. Norma ANSI S1.26-1995 służy do obliczeń absorpcji energii fali dźwiękowej w powietrzu, natomiast do precyzyjnego wyznaczania prędkości dźwięku oraz opóźnień dla systemu nagłaśniającego służy metoda opracowana przez Owena Cramera w 1993. W obu wymienionych metodach obliczeń wykorzystywane są bieżące wartości czynników atmosferycznych tj. temperatury, wilgotności oraz ciśnienia atmosferycznego, a w normie ANSI S1.26-1995 obliczany jest współczynnik tłumienia dla konkretnej częstotliwości środkowej pasma tercjowego, który przemnożony przez odległość w metrach daje wartość tłumienia w dB dla podanej odległości.

Urządzenie jest przystosowane do współpracy z systemem dźwiękowym poprzez interfejs MIDI. Ma możliwość obsługi do trzech systemów strefowych jednocześnie, czyli trzech niezależnych systemów nagłośnieniowych umiejscowionych w różnych odległościach od sceny. Do każdej ze stref nagłośnienia jest wysyłana informacja dla korektora tercjowego lub innego korektora o zmianach charakterystyki częstotliwościowej oraz informacja o wartości opóźnienia w danej strefie dla procesora linii opóźniającej. Głównym zadaniem urządzenia jest poinformowanie użytkownika, że w wybranym paśmie częstotliwości może zmienić się tłumienie lub wzmocnienie energii fali dźwiękowej zależnych od aktualnych warunków atmosferycznych. Próg decyzyjny zadziałania urządzenia podany w skali decybelowej ustala użytkownik systemu. Z głównego menu można wybrać ustawienia urządzenia i tam ustalana jest częstotliwość środkowa interesującego nas pasma, odległość od źródła oraz wartość progowa w dB, po której przekroczeniu lub spadku poniżej procesor poinformuje użytkownika o konieczności podjęcia decyzji

o zmianach w systemie nagłośnienia. Urządzenie może działać w dwóch trybach: automatycznym oraz ręcznym. W trybie ręcznym to realizator dźwięku decyduje, czy wysłać informację do systemu nagłośnieniowego i zmienić parametry systemu lub odczekać do kolejnego utworu muzycznego i w przerwie pomiędzy utworami dokonać korekcji systemu. Wybór trybu automatycznego pozwala na zmiany w systemie nagłośnieniowym samoczynnie co 3 minuty.

Przy założeniu, że w systemie nagłośnieniowym jest tylko jeden korektor charakterystyki częstotliwościowej i to on służy również do optymalizacji czy nastrojenia całego systemu nagłośnieniowego taką optymalizację można przeprowadzić bezpośrednio z urządzenia wybierając kolejne pasma oraz wartość tłumienia w dB. Urządzenia zapamiętuje nastawy i po włączeniu zasilania przywraca wszystkie ustawienia. Dodatkowo urządzenie posiada prosty kalkulator do obliczeń absorpcji w danych warunkach atmosferycznych na podstawie informacji o odległości w metrach od źródła i częstotliwości środkowej pasma. Obudowa zaprojektowanego urządzenia ma standardowe rozmiary przystosowane do szaf typu RACK 19” o wysokości 1U. Posiada wyświetlacz OLED oraz enkoder do poruszania się po menu.

WYKORZYSTANIE METOD ZDALNYCH W TRENINGU SŁUCHU MUZYCZNEGO

Sara SAJA
Bartłomiej KRUK
Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
saraoliwiasaja@gmail.com
Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
bartlomiej.kruk@pwr.edu.pl

WPROWADZENIE

W dzisiejszych czasach dostęp do technologii jest powszechny, a popularnymi stały się metody nauki w formie zdanej. Dzięki takim rozwiązaniom można zdobywać doświadczenie oraz wiedzę we własnym zakresie i wykorzystywać w produktywny sposób czas wolny. Klasyczne testy odsłuchowe wymagają specjalistycznych pomieszczeń, a także poświęcenia czasu ze strony zarówno uczniów, jak i nauczyciela. Pandemia związana z Covid-19, przyczyniła się do oswojenia społeczeństwa z różnego rodzaju narzędziami do pracy zdalnej, a także do rozszerzenia ich dostępności. Na rynku istnieje dużo rozwiązań dotyczących ćwiczeń słuchowych, jednakże skierowane są one głównie do muzyków. Rozwiązania dla inżynierów dźwięku, w szczególności tych rozpoczynających swoją przygodę z nagraniami oraz miksem utworów muzycznych są mało powszechne.

FORMA TESTÓW ODSŁUCHOWYCH

Ćwiczenia wykonano w oparciu o rekomendację ITU-R BS.1116-1. Nagrania dźwiękowe przygotowane były w taki sposób, aby wyraźnie była oddzielona od siebie faza nauki oraz faza, w trakcie której były wykonywane testy. Faza nauki polegała na prezentacji próbek dźwiękowych, które następnie były modyfikowane (w sposób zależny od tematu danego ćwiczenia). Wynikiem tego był zestaw składający się z dwóch różnych próbek – próbka referencyjna (oryginalna) oraz ta pokazująca różnicę.
Liczba zestawów była zależna od ilości prezentowanych zmian, które powiązane były z konkretnym ćwiczeniem. Czas nauki trwał tyle, ile użytkownik potrzebował do zapamiętania zmiany, ponieważ materiałem uczącym było nagranie przekazywanie uczniom. Istniała zatem możliwość swobodnego przewijania lekcji do momentu gotowości ucznia. Kolejnym etapem była faza testowa, w trakcie której sprawdzano pamięć związaną z wcześniej zaprezentowanymi zmianami. Skorzystano z metody odsłuchowej określonej we wspomnianej wcześniej rekomendacji jako podwójnie ślepa potrójnie bodźcowa z ukrytą referencją (ang. “double-blind triple-stimulus with hidden reference”). Zajęcia odbywały się przy użyciu platformy Google Classroom, gdzie w łatwy sposób można było zgromadzić grupę docelowa, informować ich o kolejnych zajęciach oraz zbierać informację zwrotną dotyczącą ich wyników.

PROBLEMATYKA

Powiązania formy zajęć wraz z niewiadomymi czynnikami takimi jak akustyka pomieszczenia, brak kontroli nad wyborem najlepszego możliwego sprzętu odsłuchowego spowodowało, iż

zalecono wykonywanie ćwiczeń przy użyciu słuchawek. Ograniczono w ten sposób wpływ pomieszczenia na percepcję próbek dźwiękowych. Jednakże nie jest pewnym, iż osoby badane zastosowały się do prośby autora. Badanych w żaden sposób nie ograniczał termin oddawania testów – brak zalecenia. Lekcje rozsyłane były do grupy kilka razy w tygodniu. Większość osób biorących udział w zajęciach była studentami w wieku do dwudziestu pięciu lat oraz uczniowie w wieku od czternastego do osiemnastego roku życia. Zadbano, aby prowadzone badania nie kolidowały z życiem szkolnym i nie wymuszano odsyłania wyników w konkretnych ramach czasowych. Elastyczny termin odpowiedzi sprawił, że osoby mogły ćwiczyć w czasie wolnym, co mogło wyeliminować czynnik zmęczenia.

PROGRAM ĆWICZENIOWY

Badania miały charakter empiryczny. Oznacza to, że analizowano poszczególne wyniki (grupy systematycznej) z testów na bieżąco oraz dostosowywano poziom trudności zgodnie z aktualnie osiągniętymi wynikami grupy w kolejnych ćwiczeniach. Podczas ćwiczeń wykorzystano częstotliwości oktawowe zgodne z ISO lub pasma oktawowe, których środkowe częstotliwości zawierały się w zakresie od 250 Hz do 4k Hz. Ograniczenie dolnego oraz górnego zakresu częstotliwości wynikało z niepewności dotyczącej typu urządzenia, na którym były wykonywane odsłuchy.

Większość testów opierała się o wzmacnianie lub tłumienie wspomnianych pasm, lub rozpoznawanie częstotliwości sygnału sinusoidalnego (także w połączeniu go z muzyką). Pierwsze z nich prezentowane były na próbkach nagrań prostych - tzn. samego wokalu (męskiego oraz żeńskiego), pianina, gitary oraz źródeł złożonych - muzyki popularnej czy rockowej. Wyboru źródeł prostych oraz muzyki popularnej dokonano ze względu na ich wszechobecność w różnego rodzaju mediach lub życiu codziennym, przez co większość uczniów mogła być bardziej zaciekawiona ćwiczeniami oraz mogło skutkować ich lepszym rozpoznawaniem ze względu na „osłuchanie". Muzykę rockową wybrano natomiast ze względu na jej „jaśniejszy" charakter barwy, tak aby zróżnicować ćwiczenia.

Podczas kursu dodano możliwość oceny słownej pokazywanych zmian w trakcie fazy nauki. Ważnym był nacisk na analizę techniczną utworu i odwrócenie uwagi słuchaczy od samego wykonania muzycznego. Ponadto w sytuacjach profesjonalnych częstym jest używanie specyficznej nomenklatury, gdzie zadaniem inżyniera dźwięku jest przetłumaczenie pewnych subiektywnych wrażeń określanych przy pomocy słów (skojarzeniowych) na odpowiednie parametry.

WNIOSKI

Przy ćwiczeniach, które mają na celu wytworzenie nowej umiejętności, ważnym jest systematyczność, aby dobrze i efektywnie kształtować techniczne słuchanie. Wyniki grupy systematycznej były znacznie lepsze od pozostałych osób. Na podstawie pracy rozpatrzono także wskazówki dotyczące wykonywania ćwiczeń w przyszłości. Wskazane jest również zachęcanie uczniów do nazywania percypowanych zmian barwy –umiejętność opisu. Dzięki temu mogą w łatwy sposób w późniejszej pracy precyzować i kojarzyć za co odpowiedzialne są konkretne pasma częstotliwościowe oraz rozumieć ich korelację. Pierwsze efekty zauważono już po pierwszym teście, jednakże wyniki zadowalające osiągnięto już po pięciu tygodniach nauki. Najlepsze efekty osiągały osoby, które wykonywały testy regularnie. Metody całkowicie zdalne

sprawdzają się, pomimo braku kontroli nad uczestnikami kursu. Słuchacze, którzy nie wykazali samodyscypliny, rezygnowali po pewnym czasie z kursu. Grupa osób, która przeszła kurs wykazywała chęci do rozwoju słuchu muzycznego. Poprawa wyników indywidualnych zauważona została dla wszystkich uczestników kursu.

ANALIZA METODOLOGII ALGORYTMÓW STOSOWANYCH W STROJENIU SYSTEMÓW NAGŁAŚNIANIA

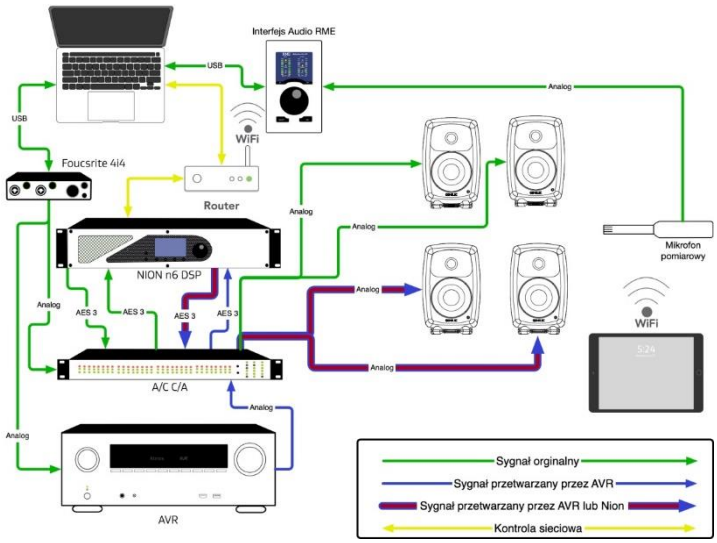
Łukasz BUREK
Bartłomiej KRUK
[1]Politechnika Wrocławska, Wyspiańskiego 27, 50-370 Wrocław
luaksz.k.burek@gmial.com, bartlomiej.kruk@pwr.edu.pl

1. WSTĘP

Celem pracy jest analiza metod strojenia, które wykorzystują algorytmy automatycznej korekcji sygnału oraz zadane manualne algorytmy. Do realizacji powyższego założenia konieczne jest wykonanie następujących zadań: analiza dostępnych metod strojenia systemów nagłaśniania,• zaprojektowanie stanowiska pomiarowego, wykonanie pomiarów z wykorzystaniem metod subiektywnych oraz obiektywnych,

2. STANOWISKO BADAWCZE

Badanie trzech różnych strojeń wymagało przygotowania odpowiednio zaawansowanego stanowiska badawczo – pomiarowego. Głównym założeniem podczas projektowania stanowiska było logiczne połączenie wszystkich urządzeń w jeden system elektroakustyczny. Istotnym było zachowanie wszystkich standardów związanych z transmisją sygnałów elektroakustycznych, a także zapewnienie jak najlepszych warunków odsłuchowych w pomieszczeniu odbiorczym. Dodatkowym celem było opracowanie możliwie wygodnego i przejrzystego sposobu symultanicznego przełączania się pomiędzy strojeniami. Założenie to wynikało przede wszystkim z konieczności przeprowadzenia subiektywnych testów odsłuchowych, które wymagały płynnej zmiany strojenia. Zapewnienie takiej funkcjonalności pozwoliło także na łatwiejsze przełączanie podczas badań obiektywnych.



Rys. 1 Schemat połączeń

2.1 METODYKA BADAŃ

Z uwagi na brak gotowej metody badań, została wprowadzona nowo opracowana metoda przez autora pracy, która bazowała na metodach opisanych w normach ITU-R -BS.11163_2015_02 oraz EBU - tech.3286. Główną przeszkodą w istniejących zaleceniach była obecność materiału referencyjnego. Ze względu na charakter badań, nie było to możliwe. Porównanymi obiektami nie były próbki dźwiękowe, a różnego rodzaju wrażenia związane z poprawnym zestrojeniem systemu odsłuchowego.

Każdy z słuchaczy miał za zadanie odsłuchać dwa utwory wybrane przez autora oraz dwa znane mu utwory, które zostały wybrane przez nich samych. W trakcie odsłuchu każda z osób mogła dowolnie przewijać fragmenty utworów, a także zmieniać je odpowiednio w stosunku do preferencji. Do oceny zostały podane cztery możliwości strojeń, które nie były znane słuchaczowi.

3. PODSUMOWANIE

Proces strojenia systemu nagłaśniania jest pracą wymagającą dużej znajomości systemów elektroakustycznych oraz akustyki pomieszczenia. Wiedza ta jest bardzo przydatna podczas analizy jakości systemu oraz przy rozwiazywaniu problemów które mogą wystąpić w trakcie tego procesu. Automatyczna kalibracji nawet w przypadku starszych algorytmów przynosi lepsze wrażenia słuchowe niż brak strojenia systemu. Dlatego może być ona dobrym punktem wyjścia dla inżyniera wykonującego kalibrację w sposób zaawansowany. Profesjonalne rozwiązania takie jak algorytm GLM przynoszą rezultaty porównywalne ze strojeniem manualnym, przy wymaganej mniejszej ilości sprzętu. Dużą zaletą takiego rozwiązania jest fakt, że urządzenia te są wstanie jeszcze lepiej zarządzać ustawieniami wzmacniaczy mocy oraz częstotliwości podziału. Gdzie w przypadku strojenia aktywnych zestawów głośnikowych w innych rozwiązaniach jest to nie możliwe.

Każdy człowiek ma własne upodobnia co do brzmienia systemu z tego powodu niemożliwym do osiągniecia jest idealne strojenie, które będzie zadowalało wszystkich. Ta zależność została dobrze przedstawiona na badanej grupie osób których oceny nie były całkowicie zbieżne. Bardzo subiektywne okazało się także percypowanie przestrzeni stereo. W przypadku strojenia manualnego okazała się ona bardzo szeroka w porównaniu do strojenia automatycznego wykonanego przez urządzenia Genelec. Było to spowodowane prawdopodobnie użyciem zbyt różnych korekcji na lewym i prawym kanale. Część słuchaczy w swobodnej rozmowie wskazywało to jako wadę cześć jako zaletę.

RESTAURACJA STUDIA DO NAGRAŃ SŁOWNYCH Z LAT 50-TYCH XX WIEKU W POLSKIM RADIO W WARSZAWIE

Radosław SMOLIŃSKI
Polskie Radio S.A., Al. Niepodległości 77/85, 00-977 Warszawa
Radoslaw.Smolinski@polskieradio.pl

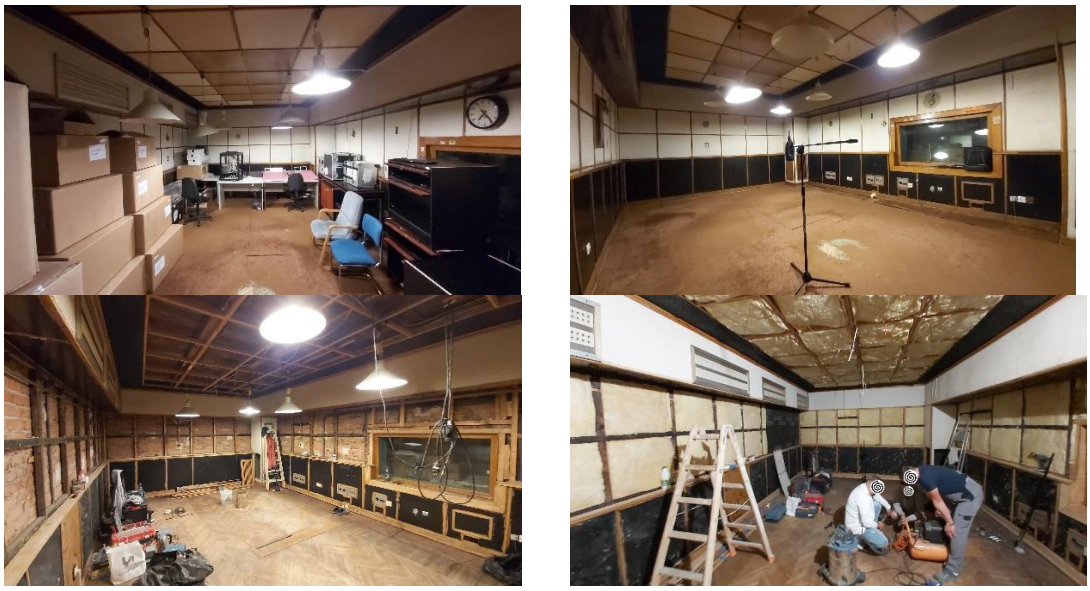
1. WPROWADZENIE

Polskie Radio S.A. ma główną siedzibę w Warszawie w Al. Niepodległości 77/85. Budynek powstał w latach 50-tych XX wieku. Obecnie w Polskim Radiu funkcjonują nowoczesne studia emisyjne, produkcyjne i muzyczne ale oprócz tego kilka fragmentów budynku zachowało wystrój architektoniczny z początków istnienia rozgłośni. Należy do nich studio do nagrań słownych S6, które przetrwało w niezmienionej formie architektonicznej z lat 50-tych XX w. Ze względu na wyeksploatowanie techniczne zostało wyłączone z użytkowania. Obecnie podjęto prace mające na celu restaurację studia - odrestaurowano zabytkowe elementy wystroju oraz uzupełniono wyposażenie w nowoczesny sprzęt nagraniowy i nowe meble technologiczne. W studiu dokonano korekty w zakresie akustyki wnętrz uwzględniając współczesne wymagania techniczne. Uzupełniono zabytkowe elementy akustyczne wkomponowując w nie nowoczesne rozwiązania dźwiękochłonne i rozpraszające fale dźwiękowe. W studiu zachowano oryginalny, unikalny historyczny wystrój, dostosowując równocześnie studio do współczesnych realiów technologicznych, akustycznych i wizerunkowych. Obecnie trwają w studiu prace instalacyjne, po ich zakończeniu studio zostanie przekazane do eksploatacji.

1. RESTAURACJA REŻYSERNI R6

Zespół studyjny składa się z dwóch pomieszczeń - studia i reżyserni. Największy zakres prac restauracyjnych dotyczył pomieszczenia reżyserni. W stanie przed restauracją reżyserni pełniła funkcje pomocnicze w zakresie testowania sprzętu. Usunięto z reżyserni wszystkie urządzenia i meble. Przystąpiono do oceny wyeksploatowania materiałów i ustrojów akustycznych ściennych, sufitowych oraz podłogi. Ze względu na wyeksploatowanie warstw wykończeniowych ustrojów akustycznych i degradację materiałów porowatych wypełniających akustyczne elementy perforowane, zdecydowano o demontażu większości elementów akustycznych ze ścian i sufitu oraz demontażu wykładziny podłogowej. Usunięto wyeksploatowane materiały porowate wypełniające ustroje akustyczne aż do odsłonięcia ścian murowanych i drewnianej konstrukcji ciesielskiej sufitu. Następnym etapem prac było odtworzenie wypełnienia ustrojów akustycznych wykonane z wełny szklanej i zabezpieczenie tej warstwy przed pyleniem (Fot. 1-4.).

Elementy akustyczne zdemontowane ze ścian i sufitu reżyserni poddano starannej renowacji. Usunięto z elementów stare powłoki malarskie, uzupełniono ubytki, wykonano zabezpieczenia przeciwpożarowe, zainstalowano ponownie na konstrukcjach drewnianych, odtworzono powłoki malarskie. Odsłonięto oryginalny, dębowy parkiet, który został wycyklinowany a następnie olejowany.



Fot. 1-4. Reżysernia R6, etapy renowacji – stan użytkowania jako pomieszczenie pomocniecze, usunięcie wyposażenia, demontaż elementów akustycznych ze ścian i sufitu, do odsłonięcia ścian murowanych i drewnianej konstrukcji ciesielskiej sufitu, odtworzenie wypełnienia ustrojów akustycznych wykonane z wełny szklanej.

Równolegle opracowano szczegółowy, trójwymiarowy projekt wnętrza w środowisku BIM. Na tej podstawie wykonano meble technologiczne i aranżację funkcjonalną wnętrza (Rys 1. i Fot. 5.) Wykonano również analizę akustyczną reżyserni. Pomiary wykonane przed renowacją wykazały wartości czasu pogłosu za duże w porównaniu do obecnych wymagań w zakresie realizacji dźwięku $T_{500\text{-}1000\ Hz} = 0{,}35$ s. Zdecydowano się zmniejszyć czas pogłosu reżyserni do wartości $T_{500\text{-}1000\ Hz} = 0{,}25$ s stosując wolnostojące elementy dźwiękochłonno – rozpraszające za stanowiskiem realizatora dźwięku oraz wyspowe dźwiękochłonne elementy sufitowe.

Rys. 1. i Fot. 5. Reżysernia R6, etapy renowacji – projekt BIM i zdjęcie reżyserni po renowacji.

Pomiary akustyczne i oceny słuchowe reżyserni po renowacji potwierdziły trafność przyjętych założeń.