



OSKA

Ogólnopolska Studencka
Konferencja Akustyków

27-30.03 2025

Warszawa

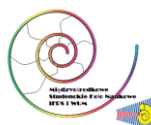
Wspierają nas



Patroni



Organizatorzy



Partnerzy Wydarzenia



www.oska-konferencja.pl

Skład:

Marianna Jarzombek, Maciej Marcinkiewicz
Łukasz Pokorzyński, Katarzyna Stec

Stowarzyszenie Producentów Silikatów „Białe murowanie”

Stowarzyszenie Producentów Silikatów „Białe murowanie” powstało w 2012 roku i od tego czasu działa na rzecz upowszechniania wiedzy dotyczącej produkcji i zastosowania wyrobów silikatowych w budownictwie oraz ich wpływu na środowisko naturalne.



We współpracy z najbardziej uznanymi ekspertami Stowarzyszenie wydaje publikacje poświęcone najważniejszym wymaganiom, jakie powinny spełniać budynki. Dotychczas ukazały się: „Izolacyjność od dźwięków powietrznych i uderzeniowych. Regulacje prawne, obliczenia i rozwiązania konstrukcyjne na przykładzie ścian z silikatów”, autorstwa dr. Leszka Dulaka, „Oszczędność energii i ochrona ciepła. Regulacje prawne, obliczenia i rozwiązania konstrukcyjne na przykładzie ścian z silikatów”, autorstwa prof. Dariusza Bajno oraz „Mury skrupowane z elementów silikatowych”, autorstwa prof. Łukasza Drobca.



Stowarzyszenie współpracuje z kołami naukowymi studentów i doktorantów, regularnie wspierając wydarzenia przez nie organizowane. Ponadto, eksperci Stowarzyszenia co roku biorą udział w konferencjach studenckich.

SAMSUNG

Samsung R&D Institute Poland jest jednym z największych centrów badawczo-rozwojowych w Polsce. W ramach prac prowadzonych w ośrodku powstaje wysokiej jakości oprogramowanie dla

produktów Samsung Electronics, które znajduje się w topowych produktach marki, zwłaszcza w smartfonach i telewizorach, które trafiają do użytkowników na całym świecie. W centrum projektowane i wdrażane są rozwiązania z takich obszarów, jak AI, multimedia, serwisy w chmurze, analityka danych, bezpieczeństwo IT czy technologie internetu rzeczy. Centrum specjalizuje się w zagadnieniach związanych z algorytmami sztucznej inteligencji, uczeniem maszynowym, analizą dźwięku i mowy, przetwarzaniem obrazu i języka naturalnego oraz analityką danych i bezpieczeństwem informacji. Współtworzona jest tutaj usługa głosowa Bixby, platforma Tizen, usługi płatności mobilnych, Samsung Wallet, usługi mobile-health, system optymalizacji gier mobilnych oraz Samsung Cloud. W obszarze badań nad technologiami sztucznej inteligencji, jako jeden z największych zespołów AI w Europie centrum zajmuje się technologiami i algorytmami wykorzystywanymi m.in. przy tworzeniu systemów rekomendacyjnych czy nawigacyjnych, jest również jednym z wiodących ośrodków specjalizujących się w przetwarzaniu języka naturalnego, w szczególności w zakresie tłumaczenia maszynowego, agentów dialogowych oraz szeroko pojętej analityki tekstu. Tworzone jest tutaj także oprogramowanie na urządzenia do odbioru telewizji cyfrowej najnowszej generacji oraz panele wielkoformatowe. Ważnym elementem i obszarem prowadzonych badań są prace związane z projektowaniem, tworzeniem i testowaniem rozwiązań dla sieci dostępowej i szkieletowej w technologii 5G w oparciu o najnowsze rozwiązania chmurowe.



Wstęp



Szanowny Oskowiczu!

Z OSKA swój pierwszy kontakt miałem nieco ponad rok temu, kiedy okazało się, że zostaje ona reaktywowana po długim okresie zastoju i niepewności. Po doświadczeniu tego wydarzenia i zawarciu wielu wspólnych znajomości w Krakowie, mogę powiedzieć, że jest jedyne w swoim rodzaju.

Wydarzenie, w którym macie okazję uczestniczyć, jest efektem ogromnego zaangażowania i ciężkiej pracy wielu osób, które studiuje i pracują nie tylko na Politechnice Warszawskiej, ale także na wielu innych uczelniach technicznych i uniwersytetach w całej Polsce. Ta książeczka i fakt, że czytacie właśnie te słowa, są tego największym dowodem.

Chciałbym z całego serca podziękować wszystkim członkom Komitetu Organizacyjnego, Naukowego oraz Jury konkursu OSKArec edycji 2025. Bez Waszego wkładu i determinacji to wydarzenie nie mogłoby się odbyć. Dziękuję również firmom, które postanowiły wesprzeć naszą inicjatywę w tym roku, bowiem ich wkład także jest nieoceniony.

Konferencja naukowa to nie tylko sesje referatowe i odkrywanie nowych tematów, ale również wspólna zabawa i produktywnie spędzony czas z osobami o zbliżonych pasjach. Mam nadzieję, że podczas tego wydarzenia nie tylko wzbogacie swoją wiedzę i poszerzycie horyzonty, ale także zawrzecie wiele nowych znajomości oraz wymienicie się doświadczeniami

inż. Łukasz Pokorzyński

Przewodniczący Komitetu Organizacyjnego
OSKA 2025



Komitet naukowy



Przewodniczący: dr inż. Agnieszka Paula Pietrzak

dr hab. inż. Artur Janicki, prof. PW

dr hab. Inż. Tomasz Zieliński

dr Rafał Bielas

dr inż. Aleksandra Chojak

dr inż. Bartłomiej Chojnacki

dr Przemysław Danowski

dr inż. Paweł Dziechciński

dr inż. Maciej Jasiński

dr inż. Piotr Kozłowski

dr inż. Marcin Lewandowski

dr inż. Michał Łuczyński

dr inż. Karolina Marciniuk

dr inż. Dorota Młynarczyk

dr inż. Mateusz Modrzejewski

dr inż. Bartłomiej Mróz

dr inż. Anna Ostaszewska-Lizewska

dr inż. Przemysław Plaskota

dr Bartosz Przysucha

dr inż. Grzegorz Szczepański

dr inż. Marcin Zastawnik



Jury OSKArec



Przewodniczący: dr inż. Marcin Lewandowski

dr Jan Felcyn

dr inż. Bartłomiej Kruk

dr Przemysław Danowski

dr inż. Bartłomiej Mróz



Komitet organizacyjny



Przewodniczący: inż. Łukasz Pokorzyński

mgr inż. Tomasz Murawski

mgr inż. Karolina Pondel-Sycz

mgr inż. Andrzej Szczepańczyk

inż. Andrzej Budny

lic. Marianna Jarzombek

inż. Daria Kałowska

lic. Nina Lignar

inż. Maciej Marcinkiewicz

inż. Aleksandra Sawczuk

inż. Katarzyna Stec

inż. Mateusz Zych

Paweł Pławczyk

Oprawa graficzna: inż. Aleksandra Gątarczyk



Abstrakty referatów



Sesja anglojęzyczna – 27 marca

Referat A.1

Soundscape approach in environmental noise assessment

Jakub Dumanowski

Uniwersytet im. Adama Mickiewicza w Poznaniu

The traditional approach to environmental noise assessment, based on equivalent sound level measurements (L_{den} and L_n) and on annoyance assessments of individual sources, often neglects context and treats environmental sounds as undesirable, ignoring desirable sounds. The soundscape approach represents a paradigm shift in noise management, considering environmental sounds as a “resource” rather than “waste.” It focuses on the perception of the sound environment and its subjective evaluations, and on distinguishing between desirable and undesirable sounds. This perspective opens up new possibilities for improving the quality of the sound environment beyond noise level reduction. One of the approaches is to mask undesirable sounds by emphasizing desirable ones. A standardized soundscape assessment includes eight attributes (pleasant, vibrant, eventful, chaotic, annoying, monotonous, uneventful, calm), focusing on both positive and negative aspects of the sound environment. Our research group's current work focuses on translating and validating soundscape attributes from ISO/TS 12913-2:2018 as part of the international Sound Attributes Translation Project to ensure cross-language compatibility and comparability of soundscape research results. The presentation will compare the soundscape approach with traditional noise management and highlight the significance of the translation project for global standardization of soundscape assessment.

Referat A.2

ACOUSTIC SPECTRUM OF MECHANICAL IMPACT COMPARED TO THE RESONANT FREQUENCIES OF THE SRS IMPACT RESPONSE SPECTRUM

Jarosław Jóźwik

CTO S.A./Politechnika Gdańska

Mechanical shock measurements are often performed using shock machines and shock transducers mounted directly on the tested object. During the analysis, the shock response spectrum (SRS) is most often used, which describes the properties of the resonance frequency spectrum of oscillators with one degree of freedom (SDOF). These spectra look similar in the form of graphs, but the analysis of the sound level in bands with specific filter center frequencies, which describes the frequency content of the signal itself and shows what frequencies and amplitudes are present in the analyzed signal. The SRS spectrum, on the other hand, illustrates the response of the mechanical system by a set of SDOF oscillators to the shock. It describes what vibrations (usually expressed by vibration acceleration) will have individual systems with different natural frequencies, subjected to the same shock. The article presents the results of shock measurements performed using a shock machine using a shock transducer (accelerometer) along with simultaneous measurement of the acoustic spectrum of the shock using a microphone. The aim of this work is to attempt to determine the correlation between the noise spectrum generated by the impact and the SRS impact response spectrum measured directly on the tested object.

Referat A.3

A comparative study of deep End-To-End automatic speech recognition models for doctor-patient conversations in Polish in a realistic acoustic environment

Karolina Pondel-Sycz

Politechnika Warszawska

The goal of the study was to find the most effective Automatic Speech Recognition (ASR) methods that can be used to build a transcription system for a doctor-patient conversation in Polish. In such a scenario, ASR models face challenges such as ambient sounds, noise, reverberation and overlapping conversations, and the medical terminology. The study evaluated four ASR models based on deep neural networks: XLSR-53 large Polish, STT PL Quartznet15x5, STT PL FastConformer Hybrid Transducer-CTC Large P&C and Whisper-large. The study used the author's dataset of real-world recordings and transcriptions of doctor-patient conversations, which is known not to have been used to train any of the models tested. The recordings were made using 5 microphones, of varying quality. The length of the recordings varies from 1.36 to 6.34 minutes (average 3.06 minutes). Originally made as 2-channel recordings, due to the requirements of the FastConformer and QuartzNet models, the recordings were converted to 1-channel, with a sampling rate of 16 kHz. Two classic ASR metrics, Word Error Rate (WER) and Character Error Rate (CER), were used to evaluate the models under study, along with additional metrics including: Match Error Rate (MER), Word Accuracy (WAcc), Word Information Preserved (WIP), Word Information Lost (WIL), Levenshtein distance (Lev. Dist.), Levenshtein ratio (LR), Jaro - Winkler similarity (JW sim.) and Jaccard index. The Whisper-large model performed best for all metrics, except for Jaro - Winkler similarity, for which the result is similar to Wav2Vec. Whisper's results are less dispersed, so the model makes errors on a narrower scale.

Referat A.4

This bee does not exist: Honeybee sound synthesis

Piotr Książek

Politechnika Wrocławska

Honeybee monitoring and sound analysis is a somewhat explored and established field in bioacoustics. Many, at times successful, attempts have been made recently to analyse and interpret sounds produced by honeybees in the context of creating apiary support systems and interpreting various bee colony state parameters. Another, connected avenue in research about honeybee sound is the synthesis of sounds meaningful to bees to aid in manipulating the bee colony state. This work demonstrates a method of synthesizing honeybee sounds using wavelet decomposition and machine learning techniques to achieve an algorithm capable of generating an infinite honeybee buzz sound texture signal from given input parameters, which may be utilized in in-vivo studies of non-trivial signal impact on honeybee colony behaviour. Impact of utilization of different wavelets on the signal characteristics is discussed, as well as the potential use of generated signals in different fields such as virtual soundscape synthesis and music production.

Referat A.5

Sonic Hues: Exploring Analogies Between Sound and Color Through Wave Analysis

Tomasz Murawski

Politechnika Warszawska

This thesis undertakes a comparative analysis of sound and colour - two fundamental perceptual phenomena with wave analysis as their common denominator. Although acoustic and electromagnetic waves, their physical carrier, differ in nature and properties, this study focuses on identifying analogies between these sensory modalities. A key element of the work is the exploration of the possibilities of colour editing using digital sound processing techniques and tools (sound plug-ins). Visualising the results of these experiments is an innovative conceptual exercise, shedding new light on the relationship between acoustics and colour theory. This work takes a novel approach, demonstrating how electroacoustic operations can be adapted to manipulate timbres, and examines the interrelationship between timbre and sound. The potential of transferring image editing techniques to the modification of audio signals is indicated here, opening up new perspectives in both fields. The historical context outlined in the work further highlights the historical links and inspirations between the two fields.

Sesja konkursowa I – 28 marca

Referat 1.1

Ocena wierności auralizacji nagrań sygnałów instrumentów muzycznych

Paweł Kabata

Politechnika Warszawska

W referacie dokonano analizy wierności procesu auralizacji nagrań sygnałów instrumentów muzycznych. Proces auralizacji polega na splocie nagrań bezechowych z odpowiedzią impulsową danego pomieszczenia, co umożliwia symulację odsłuchu dźwięku w wybranej przestrzeni. Do rejestracji nagrań referencyjnych wykorzystano manekin akustyczny. Zbadano różne przestrzenie pomiarowe, takie jak komora bezechowa i studio nagrań. Opracowano oprogramowanie w języku Python z graficznym interfejsem użytkownika (GUI), umożliwiające przeprowadzanie procesu auralizacji oraz analizę wyników za pomocą współczynnika korelacji. Przeprowadzono testy na różnych instrumentach muzycznych i oceniono skuteczność odwzorowania przestrzenności dźwięku.

Referat 1.2

Pomiar indywidualnych funkcji HRTF w warunkach bezpogłosowych

Jakub Tokarski, Paweł Skowron, Paweł Baran, Kamil Zych
Politechnika Warszawska

Rosnąca popularność multimediiów sprawia, że coraz ważniejsze staje się zwiększenie immersji słuchacza w wirtualnych scenach dźwiękowych. Aby uzyskać wrażenie przestrzenności i naturalności dźwięku w odsłuchu słuchawkowym, wykorzystuje się funkcję HRTF (Head-Related Transfer Function). Wzrost immersji można osiągnąć dzięki personalizacji tej funkcji, czyli indywidualnemu dopasowaniu do cech anatomicznych użytkownika. Celem niniejszej pracy było przeprowadzenie serii pomiarów indywidualnych funkcji HRTF dla kilku osób, weryfikacja dokładności zebranych danych oraz analiza możliwości usprawnienia i optymalizacji samego procesu pomiarowego. Badania zrealizowano w komorze bezchowej, używając odpowiednio przystosowanego sprzętu. Dźwięk emitowało dziesięć głośników rozmieszczonych co 15 stopni kąta elewacji (w zakresie od -45° do 90°). Osoba badana siedziała na obrotowym stoliku, w kanałach słuchowych mając umieszczone mikrofony. Sygnały dźwiękowe były generowane oraz rejestrowane poprzez system pomiarowy Pulse. W efekcie uzyskano sześć zestawów funkcji HRTF.

Referat 1.3

Nagrania binauralne jako pomoc w rozwoju orientacji przestrzennej

Klaudia Piątek

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Celem projektu jest stworzenie narzędzia wspierającego rozwój orientacji przestrzennej u dzieci z niepełnosprawnościami wzrokowymi poprzez bodźce słuchowe oraz zmniejszenie lęku przed wychodzeniem z domu.

Metoda treningu opiera się na słuchowiskach, które uczą zaimków przestrzennych i orientacji w przestrzeni. Zaprojektowana forma i struktura słuchowisk umożliwia naukę zarówno z mentorem, jak i samodzielnie. W ramach ćwiczeń wykorzystano nagrania binauralne, które wiernie odzwierciedlają przestrzeń dźwiękową. Stworzone słuchowiska oraz ich analizy są udostępniane na platformie YouTube w postaci filmów oraz formularzy umożliwiających przesyłanie odpowiedzi przez słuchaczy.

Analiza tych odpowiedzi pozwala ocenić odbiór słuchowisk i percepcję dźwięków w nich zawartych. Wstępne badania wykazują, że nagrania binauralne mogą wspierać słuchacza w rozwijaniu orientacji przestrzennej. Dodatkowo potwierdzają informację z literatury o rehabilitacji nagraniami binauralnymi.

Projekt jest rozwijany o mapę dźwiękową miejsc w Polsce, która będzie zawierać nagrania statyczne konkretnych miejsc oraz nagrania spacerów poszczególnymi ulicami, z dokładnym opisem co zawiera się w nagraniu. Pozwoli to na cyfrowe poznanie świata zewnętrznego z perspektywy słuchu.

Referat 1.4

Badanie skuteczności działania indywidualnych ochronników słuchu

Anna Frączkowska

Politechnika Wrocławska

Celem pracy jest zbadanie skuteczności działania indywidualnych ochronników słuchu, a dokładniej pasywnych wkładek przeciwhałasowych. Opracowano dwie metody badawcze, które miały na celu ocenę tłumienia ochronników w sposób obiektywny i subiektywny. Podstawowym parametrem, który zbadano jest tłumienie wkładki obliczane jako algebraiczna różnica wartości zmierzonej bez ochronnika i z ochronnikiem. W celu realizacji pomiarów obiektywnych zastosowano dwa sygnały pomiarowe: szum różowy filtrowany pasmami tercjowymi oraz sygnał sinusoidalny. Pomiary obiektywne wykonano przy użyciu fantomu sztucznej głowy, który umożliwił właściwe umieszczenie wkładki w kanale. Pomiary subiektywne zrealizowano jako badania audiometryczne wykorzystując standardowe sygnały sinusoidalne, zgodnie z normą dotyczącą tego typu pomiarów. Uzyskane wyniki wykazały duże rozbieżności dla metody obiektywnej i obu rodzajów sygnałów pomiarowych między sobą, natomiast wyniki badań subiektywnych wykazały rezultaty najbardziej zbliżone do wartości podawanych przez producentów ochronników.

Sesja konkursowa II – 28 marca

Referat 2.1

Badanie możliwości zastosowania wielokomórkowych metamateriałów membranowych do zwiększania izolacyjności akustycznej

Sara Kopeć, Klara Chojnacka

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Membranowe struktury rezonansowe, znane również jako metamateriały membranowe, stanowią innowacyjne rozwiązanie w dziedzinie materiałów dźwiękoizolacyjnych, szczególnie w zakresie niskich częstotliwości, przekraczając ograniczenia tradycyjnych rozwiązań. Konwencjonalne materiały izolacyjne działają zgodnie z zasadą prawa masy, izolacyjność akustyczna wzrasta proporcjonalnie do ich masy.

Metamateriały membranowe umożliwiają jednak uzyskanie wysokiej izolacyjności akustycznej w niskich pasmach częstotliwości bez konieczności znaczącego zwiększania ciężaru całego układu.

Referat skupia się na opracowaniu i walidacji modelu numerycznego opisującego zachowanie metamateriałów membranowych, a także na analizie metod manipulacji charakterystyką izolacyjności akustycznej. Ponadto, przeanalizowano możliwości integracji pojedynczych membran w wielokomórkowe struktury rezonansowe w celu uzyskania szerszego zakresu częstotliwości działania ustroju.

Weryfikację teoretycznych założeń przeprowadzono na podstawie badań eksperymentalnych, w tym pomiarów akustycznych wykonanych w rurze impedancyjnej zgodnie z normą ASTM E2611-09. Przedstawiono wyniki porównania przewidywań modelu numerycznego z wynikami pomiarów eksperymentalnych. Wyniki te otwierają nowe perspektywy projektowania wysoko wydajnych materiałów dźwiękoizolacyjnych opartych na strukturach rezonansowych.

Referat 2.2

Projektowanie i analiza dźwiękoizolacyjnych metamateriałów akustycznych typu soft solid z wtrąceniami periodycznymi

Agata Maciuszek, Klara Chojnacka

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Metamateriały wibroakustyczne przyjmują różne formy i mają różnorodne zastosowania. Jednym z nich jest zwiększanie izolacyjności akustycznej w zakresach, w których tradycyjne materiały jednorodne dają niewystarczające efekty. Aby osiągnąć taki efekt, należy wpłynąć na fale rozchodzące się w elemencie podstawowym poprzez rozmieszczenie w odpowiednich odstępach na dostępnej powierzchni elementy rezonansowe. W pracy analizie poddano metamateriały wibroakustyczne oparte na strukturze soft solid wraz z periodycznymi wtrąceniami, składające się z materiału porowatego oraz cyklicznych elementów wykonanych z materiału jednorodnego. Dodane wtrącenia do materiału porowatego umożliwiają poszerzenie możliwości struktury w kontekście izolacyjności akustycznej. Efektywność struktury zależy zarówno od parametrów materiałowych jak i geometrycznych wykorzystanego materiału porowatego i ciała stałego. W ramach prac wykonany został model w programie COMSOL, który uwzględnia interakcję domeny akustycznej z ciałem stałym. Analizie poddano wpływ zmian parametrów materiałowych oraz geometrycznych, takich jak grubość materiału porowatego, kształt wtrąceń oraz ich wypełnienie, na efektywność struktury. Bazując na wynikach symulacji numerycznych dobrano materiały o odpowiednich parametrach i wykonano prototyp opisanej wyżej struktury. Dla prototypu przeprowadzono pomiary izolacyjności akustycznej w polu rozproszonym w celu weryfikacji efektywności struktury. Kolejnym etapem było wykonanie optymalizacji parametrów geometrycznych struktury za pomocą metody PSO, a następnie wzorując się na otrzymanych wartościach, wykonanie kolejnej struktury.

Referat 2.3**Projekt adaptacji akustycznej sali seminaryjnej****Anna Celińska**

Politechnika Warszawska

W referacie przedstawiono projekt adaptacji akustycznej sali seminaryjnej Instytutu Radioelektroniki i Technik Informacyjnych na Wydziale Elektroniki i Technik Informacyjnych Politechniki Warszawskiej. Celem opisanych prac było dostosowanie parametrów akustycznych pomieszczenia do wymagań normy PN-B-02151-4, zapewniając tym samym odpowiednie warunki do prowadzenia zajęć, wykładów i seminariów. W pierwszym etapie projektu przeprowadzono rejestrację odpowiedzi impulsowych wnętrza, co umożliwiło ocenę zastanych warunków akustycznych. Na podstawie wyników czasu pogłosu (T_{30}) skalibrowano model akustyczny pomieszczenia opracowany w oprogramowaniu symulacyjnym Odeon. Adaptacja akustyczna opierała się na rozwiązaniach, które nie ingerowały w konstrukcję budowlaną obiektu, takich jak instalacja paneli sufitowych, pokrycie podłogi wykładziną oraz zastosowanie absorberów ściennych. Na podstawie obliczeń modelu wyznaczono i przeanalizowano cztery kluczowe parametry akustyczne: czas pogłosu (T_{30}), wskaźnik transmisji mowy (STI), czas wczesnego zaniku (EDT) oraz czytelność dźwięku (C_{50}). Ocena tych parametrów przed i po adaptacji potwierdziły skuteczność zaproponowanych rozwiązań, które spełniły zarówno wymagania normatywne, jak i zalecenia literaturowe w zakresie akustyki wnętrz.

Sesja III – 28 marca

Referat 3.1

Analiza parametrów akustycznych wybranych hal sportowych w Krakowie

Jakub Woźniak, Adam Pilch

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Projekt dotyczy analizy parametrów akustycznych wybranych hal sportowych w Krakowie oraz zaproponowanie ewentualnych adaptacji akustycznych. W ramach pracy przeprowadzono badania na trzech halach sportowych o różnych objętościach oraz geometrii. W trakcie badań zastosowano jedną z metod geometrycznych, metodę promieniową, która została szczegółowo opisana w pracy. Parametrami, na których skupiono główną uwagę, był czas pogłosu, który wymieniany jest jako jeden z największych problemów hal sportowych oraz wskaźnik transmisji mowy STI. Analizę otrzymanych wyników przeprowadzono w programie CATT-Acoustic. Głównym celem pracy była analiza przeprowadzonych badań oraz sprawdzenie czy wymienione hale sportowe spełniają wymagane normy akustyczne. W przypadku gdy dany obiekt nie spełnia wymogów pozwalających na komfortowe użytkowanie pod względem akustyki, przeprowadzano symulacje z dodatkowymi adaptacjami akustycznymi, mającymi na celu poprawę wyżej wymienionych parametrów akustycznych.

Referat 3.2

Wpływ hałasu szpitalnego na rekonwalescencję pacjentów

Jadwiga Hyla, Wojciech Kotala, Aleksandra Sawczuk, Tadeusz Wszótek
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

W salach chorych na różnych oddziałach szpitalnych można zaobserwować wiele źródeł dźwięku, takich jak aparatura medyczna, rozmowy pacjentów, lekarzy i pielęgniarek, a także hałas technologiczny generowany przez działające urządzenia. Wszystkie te zdarzenia dźwiękowe mogą być źródłem dokuczliwości i negatywnie wpływać na pacjentów, prowadząc do zaburzeń snu i utrudniając regenerację po przebytych zabiegach.

W ramach projektu przeprowadzono pomiary monitoringowe hałasu występującego na obszarze jednego z krakowskich szpitali, które następnie poddano analizie i wyznaczono wartości podstawowych parametrów akustycznych użytecznych w ocenie dokuczliwości. W szczególności skoncentrowano się na wyznaczaniu poziomów statystycznych, poziomów równoważnych oraz wartości maksymalnych, oddzielnie dla każdej godziny w ciągu doby. Zapoznano się również z doniesieniami literaturowymi na temat czynników wpływających na dyskomfort pacjentów, sposobami prowadzonych ankietyzacji badanych oraz sposobami pomiarów hałasu w placówkach medycznych, m.in. z wykorzystaniem telefonów komórkowych.

Na dalszym etapie rozwoju projektu pozwoli to na pogłębioną analizę i przypisanie cyklicznych zdarzeń akustycznych konkretnym czynnościom, aparaturze, czy zabiegom. Dodatkowo przewiduje się, w ramach współpracy z personelem medycznym, przeprowadzenie badań ankietowych pod kątem identyfikacji dokuczliwych źródeł hałasu wskazanych przez pacjentów. Korelacja wyników badań ankietowych z wynikami pomiarowymi hałasu powinna umożliwić bardziej precyzyjne wskazanie dokuczliwych źródeł hałasu w porze dziennej i wybudzeń w porze nocnej.

Sesja IV – 28 marca

Referat 4.1

Trenowanie modelu detekcji mowy syntezowanej metodą voice cloning

Mateusz Zych

Politechnika Warszawska

Wraz z rozwojem technologii generowania mowy syntezowanej, w tym metod voice-cloning, coraz częściej występują zagrożenia związane z jej wykorzystaniem w atakach socjotechnicznych. W niniejszej pracy przeprowadzono analizę skuteczności metod automatycznej oceny naturalności mowy oraz detekcji mowy syntezowanej. Przeprowadzono badania dwóch modeli głębokich: Non-Intrusive Speech Quality Assessment (NISQA) oraz Audio-deepfake-detection. Model NISQA wykazał wysoką skuteczność (90%) w klasyfikacji próbek mowy testowej, natomiast model Audio-deepfake-detection początkowo cechował się niską skutecznością (50%) oraz głęboką generalizacją. W celu poprawy jego działania zastosowano dotrenowanie modelu na bazie danych LJSpeech oraz na próbkach syntezowanych metodą voice-cloning. Po przeprowadzeniu dodatkowego treningu skuteczność modelu wzrosła do 95%. Wyniki wskazują, że odpowiednie przygotowanie bazy treningowej znacząco wpływa na zdolność modeli do rozpoznawania mowy syntezowanej. Praca kontynuuje badania nad skutecznością generalizacji dotrenowanego modelu, w tym jego testowanie na nowych, różnorodnych bazach danych oraz wdrażanie technik poprawy generalizacji. Badania te mają istotne znaczenie dla bezpieczeństwa informacji i mogą znaleźć zastosowanie w systemach biometrycznych oraz w wykrywaniu deepfake audio.

Referat 4.2**Porównanie modeli sieci neuronowych do automatycznego rozpoznawania emocji w głosie w języku polskim****Katarzyna Stec**

Politechnika Warszawska

Zagadnienie automatycznego rozpoznawania emocji w głosie (z ang. Automatic Speech Emotion Recognition - ASER) jest przedmiotem intensywnych badań od wielu lat, lecz wciąż brakuje kompleksowych rozwiązań dla języka polskiego. W niniejszej pracy zaprezentowano wyniki testów czterech modeli głębokich sieci neuronowych, które wykazują potencjał w rozpoznawaniu emocji spontanicznie wyrażanych w języku polskim. Dwa z nich to wielojęzyczne modele ASER, a pozostałe dwa zostały stworzone z myślą o języku rosyjskim (zaliczanym do tej samej grupy językowej co polski). Analiza przeprowadzona na korpusie BES – bazie polskich, naturalnych (spontanicznych) wypowiedzi emocjonalnych – pokazała, że najlepszy model osiągnął dokładność na poziomie około 59%. Uzyskane wyniki podkreślają trudności związane z rozpoznawaniem bardziej subtelnych i niejednoznacznych ekspresji emocjonalnych w rzeczywistych warunkach oraz wskazują na potrzebę dalszych badań nad rozwojem modeli optymalnie dostosowanych do języka polskiego.

Sesja V – 29 marca

Referat 5.1

Analiza wpływu rozmieszczenia ustrojów rozpraszających dźwięk w akustyce małych pomieszczeń z wykorzystaniem algorytmów obliczania stopnia rozproszenia pola akustycznego

Maria Brzóška, Bartłomiej Chojnacki

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Ustroje rozpraszające dźwięk są powszechnie wykorzystywane do poprawy akustyki analizowanego wnętrza. Ich zastosowanie umożliwia niwelację wad akustycznych przy jednoczesnym zachowaniu większości energii akustycznej w pomieszczeniu. Dotychczas zaproponowano w literaturze szereg parametrów do oceny stopnia rozproszenia pola, jednak ograniczają się one do pomieszczeń o dużych rozmiarach. We wnętrzach o małych kubaturach panują odmienne warunki akustyczne, w szerszym zakresie częstotliwości występują zjawiska falowe, przez co wymagane jest odrębne podejście. W pracy zbadano zaproponowane w literaturze parametry do oceny stopnia rozproszenia pola, takie jak krzywa echo – density oraz krzywa znormalizowanej kurtozy. Do algorytmów zaimplementowano poprawki dla małych pomieszczeń oraz zbadano skuteczność ich zastosowania. W związku z zaleceniem stosowania źródła kierunkowego do pomiarów akustyki małych pomieszczeń oraz wykorzystaniem przez autorów algorytmów źródła wszechkierunkowego, w niniejszej pracy przeanalizowano wpływ zmiany charakterystyki kierunkowości zastosowanego źródła na uzyskane wyniki. Ostatnim etapem pracy było przeanalizowanie skuteczności działania rozpraszacza dźwięku w zależności od jego lokalizacji z wykorzystaniem opracowanych parametrów.

Referat 5.2

Elastyczne metamateriały akustyczne do zwiększania izolacyjności akustycznej

Tytus Nowak, Klara Chojnacka, Aleksandra Sawczuk

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Metamateriały akustyczne to sztucznie wytworzone materiały, których właściwości akustyczne wynikają z ich parametrów geometrycznych jak i materiałowych. Charakteryzują się one zdolnością do manipulowania falami dźwiękowymi poprzez lokalne rezonanse, umożliwiając osiągnięcie właściwości pochłaniających, rozpraszających dźwięk, czy zwiększania izolacyjności akustycznej. Metamateriały stosowane do zwiększania izolacyjności akustycznej zazwyczaj są budowane ze sztywnych elementów, co ogranicza ich zastosowanie w sytuacjach wymagających większej elastyczności. W odpowiedzi na te ograniczenia zaczęto rozwijać metamateriały elastyczne, które wykorzystują miękkie materiały jako elementy konstrukcyjne, a poprzez zjawisko lokalnych rezonansów wewnątrz struktury, oferują szerokopasmową redukcję hałasu, także w zakresie niskich częstotliwości. W ramach projektu wykonano model numeryczny metamateriału elastycznego w środowisku COMSOL Multiphysics z uwzględnieniem domeny akustycznej i strukturalnej. Model pozwolił na określenie izolacyjności akustycznej projektowanej struktury i umożliwił analizę wpływu zmian parametrów materiałowych i geometrycznych na skuteczność metamateriału w tłumieniu dźwięku

Referat 5.3

Projektowanie metamateriałowych soczewek akustycznych o labiryntowej geometrii

Jakub Słoboda, Klara Chojnacka, Miłosz Szarek

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Soczewki akustyczne to przykład metamateriałów akustycznych, ponieważ ich właściwości wynikają z precyzyjnie zaprojektowanej geometrii komórek podstawowych, a nie z właściwości materiału, z którego są wykonane. Umożliwiają one kontrolowanie i kształtowanie propagacji fal dźwiękowych poprzez odpowiednie opóźnianie ich fazy, co pozwala na manipulowanie kierunkowością, skupieniem i rozkładem fal w przestrzeni. Zastosowanie soczewek akustycznych obejmuje szeroki zakres technologii, od poprawy jakości odsłuchu w systemach audio, przez zaawansowane techniki obrazowania ultradźwiękowego, aż po nowoczesne rozwiązania haptyki w powietrzu i lewitacji akustycznej. Wykorzystanie metamateriałów umożliwia osiągnięcie precyzyjnej kontroli nad fazą i kierunkiem propagacji fal dźwiękowych, co otwiera nowe możliwości w zakresie kształtowania wiązki akustycznej. W projekcie przedstawiono proces projektowania, symulacji oraz weryfikacji właściwości soczewki akustycznej, której budowa oparta jest na komórkach z kanałami opóźniającymi. W środowisku COMSOL Multiphysics wykonano model numeryczny struktury. Wykorzystując moduł akustyczny przeprowadzono symulacje metodą elementów skończonych. Na podstawie symulacji określono geometrię kolejnych komórek podstawowych soczewki, które pozwalają na uzyskanie pożądanych parametrów skupienia oraz charakterystyki kierunkowości wypromieniowanej fali akustycznej.

Sesja VI – 29 marca

Referat 6.1

Wykorzystanie nowoczesnych technologii w badaniach zmysłów – Kapsuła Badań Zmysłów

Arkadiusz Moskwa⁴, Skarżyński P.H.^{1,2,3}, Czajka N.¹, Talarek M.¹,
Będziński W.¹, Piłka A.¹, Lignar N.^{1,4}, Skarżyński H.¹

1. Światowe Centrum Słuchu, Instytut Fizjologii i Patologii Słuchu,
Warszawa/Kajetany

2. Instytut Narządów Zmysłów, Kajetany

3. Zakład Niewydolności Serca i Rehabilitacji Kardiologicznej,
Warszawski Uniwersytet Medyczny, Warszawa

4. Międzyośrodkowe Studenckie Koło Naukowe
przy Światowym Centrum Słuchu Instytutu Fizjologii i Patologii Słuchu
oraz Warszawskim Uniwersytecie Medycznym, Warszawa

Nowoczesne technologie znacząco wpłynęły na badania zmysłów poprzez rozwój narzędzi wspomagających diagnostykę i terapię. Kapsuła Badań Zmysłów (KBZ) to zintegrowany system urządzeń do diagnostyki i telekonsultacji schorzeń narządów zmysłów, z których korzystać mogą osoby od 7-go roku życia. Urządzenie zawiera wystandaryzowane testy do prowadzenia badań przesiewowych 5 zmysłów: słuchu, wzroku, węchu, smaku i równowagi.

Kapsuła pozwala na wykonywanie badań przesiewowo-diagnostycznych u dzieci w wieku szkolnym, osób dorosłych, grup zawodowych narażonych na różne zaburzenia zmysłów oraz osób w wieku senioralnym. Urządzenie umożliwia szybką identyfikację problemu i ukierunkowuje na dalsze postępowanie i leczenie. Dodatkowo opracowano specjalny portal dedykowany dla pacjentów i specjalistów, na który wysyłane są wyniki poszczególnych badań. Lekarze mają możliwość szybkiego kontaktu z pacjentem i wskazania mu ewentualnych zaleceń.

Referat 6.2

Diagnostyka szumów usznych oraz ich wpływ na funkcje układu słuchowego

Marianna Jarzombek

Uniwersytet im. Adama Mickiewicza w Poznaniu

Szum uszny (tinnitus) to wrażenie słuchowe powstające bez fizycznej obecności dźwięku stymulującego, wpływające negatywnie na codzienne funkcjonowanie osoby dotkniętej tym zjawiskiem. Mimo wieloletnich badań prowadzonych w zakresie diagnozowania i leczenia subiektywnych szumów usznych, zjawisko tinnitusa w dalszym ciągu jest dużym wyzwaniem dla lekarzy i protetyków słuchu. Głównym celem aktualnie prowadzonych badań jest analiza, a następnie wdrożenie kompleksowych metod badawczych, pozwalających na jak najszerszą diagnostykę tinnitusa. Badania skupiają się zarówno na wpływie szumów usznych na codzienne funkcjonowanie, jak i na określeniu pełnego spektrum parametrów szumów usznych, takich jak: określenie częstotliwości i proggu maskowania, wpływu na otoemsje akustyczne i zrozumiałość mowy na tle szumu. Tak szczegółowa diagnostyka będzie bazą do tworzenia modelu powstawania tinnitusa na bazie najbardziej uznanych w międzynarodowym środowisku naukowym modeli słyszenia. To z kolei będzie podstawą do tworzenia skuteczniejszych metod terapii szumów usznych.

Referat 6.3

Wpływ kształtu przebiegu czasowego hałasu na oceny jego dokuczliwości

Julia Wesołowska

Uniwersytet im. Adama Mickiewicza w Poznaniu

Według WHO hałas jest jednym z największych zagrożeń dla zdrowia fizycznego i psychicznego oraz dobrego samopoczucia. Hałas generuje skutki zdrowotne, do których przede wszystkim należą niedokrwienność serca, przeszkadzanie we śnie oraz dokuczliwość. Dokuczliwość jest tym skutkiem działania hałasu, który dotyczy praktycznie wszystkich źródeł hałasu, a w szczególności dotyczy hałasu komunikacyjnego. Wyniki badań dokuczliwości hałasu komunikacyjnego przedstawia się zwykle w postaci funkcji przedstawiającej zależność ekspozycji hałasu od reakcji tzw. Exposure-response function (ERF), gdzie ekspozycja wyrażona jest poprzez wartość poziomu dźwięku LDWN a reakcja przez ocenę skrajnej dokuczliwości (%HA) jaką dany hałas wywołuje. Funkcje te mają różny kształt dla hałasu samochodowego, szynowego i lotniczego ze względu na różne wartości poziomu dźwięku generujące tę samą dokuczliwość. Problem w tym, że nawet dla jednego typu hałasu np. samochodowego ten sam poziom dźwięku może dotyczyć różnego przebiegu czasowego danego hałasu. W tej prezentacji zaprezentowane zostaną oceny dokuczliwości hałasów o tym samym poziomie dźwięku, ale o różnych kształtach przebiegu czasowego. Analiza wyników pokaże, czy dla pojedynczych przejazdów samochodu tramwaju i samolotu istnieje różna ocena dokuczliwości hałasu w zależności od kształtu obiedni czasowych tych hałasów.

Sesja VII – 29 marca

Referat 7.1

Program edukacyjny „SonicHarmonic” – eksplorowanie barw dźwięku poprzez parametry częstotliwości harmoniczych, napisany w języku MATLAB.

Hubert Kuś

Uniwersytet im. Adama Mickiewicza w Poznaniu

Niniejszy referat opisuje proces tworzenia, zasadę działania, główne funkcje oraz zastosowanie programu „SonicHarmonic”. Program został stworzony w języku MATLAB na potrzeby zrealizowania warsztatów, które miały na celu przybliżenie uczestnikom wybranych parametrów związanych z generowaniem dźwięków muzycznych oraz modyfikacjami częstotliwości i amplitudy sygnałów składowych. Program pozwala na manipulowanie parametrami związanymi z generacją sygnału oraz na odtwarzanie jego przebiegu, obliczonego zgodnie podanymi parametrami. W celu prezentacji wizualnej wyświetlane są również zawarty jest przebieg sygnału oraz widmo częstotliwościowe. Program może być wykorzystany jako narzędzie dydaktyczne lub w celach muzycznych jako podstawowy syntezytor ponieważ pozwala na wizualizowanie zmian w czasie rzeczywistym w formie wykresów oraz generowanie zmiennych sygnałów dźwiękowych.

Referat 7.2**Redukcja wpływu rezonansów wkładki i talerza z wykorzystaniem wibroizolacji w gramofonach typu lekkiego****Aleksandra Sawczuk, Bartłomiej Chojnacki**

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Wibroizolacja pełni istotną rolę w ograniczaniu przenoszenia drgań między źródłem a otoczeniem, szczególnie w przypadku delikatnych urządzeń, w których wszelkie dodatkowe wstrząsy są niepożądane. Tradycyjne metody redukcji drgań, takie jak zwiększanie masy urządzenia, nie są możliwe do zaimplementowania w przypadku lekkich gramofonów o masie poniżej 10 kg. W takich urządzeniach szczególnie problematyczne są drgania o niskiej częstotliwości, które wywołują rezonans wkładki i ramienia, wpływając negatywnie na jakość odtwarzanego dźwięku oraz mogąc prowadzić do uszkodzenia elementów mechanicznych. Aby je zminimalizować należy odizolować gramofon od podłoża, na przykład przy pomocy platformy antywibracyjnej. W pracy przedstawiono wyniki badań, które obejmowały analizę różnych metod pobudzania drgań w gramofonie oraz ocenę skuteczności materiałów wibroizolacyjnych. W pierwszym etapie sprawdzono techniki pobudzenia drgań, takie jak szum różowy, stukacz młotkowy i praca własna urządzenia. W drugim etapie badano, jak różne materiały wibroizolacyjne, takie jak sprężyny, elastomery i wibroizolatory liniowe, wpływają na ograniczenie amplitudy przyspieszeń drgań. Do oceny wyników zastosowano analizę widmową oraz funkcję przejścia, co pozwoliło na wybór najbardziej efektywnej metody generowania drgań oraz odpowiedniego materiału wibroizolacyjnego.

Referat 7.3

Analiza porównawcza parametrów akustycznych emulacji mikrofonów z ich pierwowzorami na bazie wybranych mikrofonów modelujących

Paweł Kortas

Uniwersytet im. Adama Mickiewicza w Poznaniu

Na rynku pro audio znajduje się kilka specjalnie zaprojektowanych mikrofonów z zaawansowanym procesorem dsp, który ma umożliwić symulowanie brzmienia określonych topowych, mikrofonowych "klasyków". Procedura przebiega w ten sposób, że dokonuje się normalnego nagrania owym specjalnym mikrofonem a następnie, korzystając z odpowiedniego procesora, wybiera się mikrofon, którego brzmienie chce się zasymulować. Praca polega na praktycznym zbadaniu takiego systemu i porównaniu jego efektów brzmieniowych z rzeczywistymi mikrofonami. W tym celu przeprowadzono analizę obiektywną, czyli zbadano oraz porównano odpowiedzi częstotliwościowe dwóch mikrofonów modelujących oraz pięciu mikrofonów "zwykłych". Następnie dokonano oceny subiektywnej za pomocą testu ABX. Wykorzystano w tym celu próbki zarejestrowane w studiu, które były nagrywane przez wcześniej wspomniane siedem mikrofonów.

Sesja VIII – 30 marca

Referat 8.1

Mobilna wersja gry fabularnej przeznaczonej dla osób niedowidzących i niewidzących

Daria Kałowska

Politechnika Warszawska

Projekt jest realizowany w ramach pracy magisterskiej. Skupia się na wykonaniu mobilnej gry fabularnej inspirowanej papierowymi grami typu RPG (role-playing game), która wykorzystuje dźwięk przestrzenny i jest przeznaczona głównie dla osób niewidzących i niedowidzących. Plan pracy zakłada rozplanowanie mechanik, na których będzie opierać się rozgrywka tak, aby gra była spójna i ciekawa dla gracza. Następnie wybrany system zostanie zaimplementowany w silniku do gier komputerowych - Godot. Po utworzeniu działającego wzorca możliwe będzie dodanie elementu fabularnego oraz dopasowanie optymalnego wpływu poszczególnych wyborów gracza na etapy rozgrywki. Posiadając gotowy program rozpocznie się nagrywanie części dźwiękowej, czyli dialogów, efektów dźwiękowych, komunikatów, dźwięków środowiskowych, muzyki oraz dodanie ich do projektu za pomocą dodatkowego silnika Fmod. Większość nagrań zostanie zrealizowana samodzielnie, a pozostałe dźwięki zostaną pozyskane z plików open source. Część nagrań zostanie wykonana za pomocą mikrofonów ambisonicznych, a reszta plików zostanie zakodowana w przestrzeni 3D w programie do edycji dźwięku. Po dodaniu dźwięków do silnika możliwe będzie zaimportowanie gry na wersję mobilną i ostateczne dostosowanie interfejsu do osób niewidzących. Efekty pracy zostaną poddane ocenie na grupie osób z wadami wzroku i na grupie bez takich wad. Z badania uzyskane zostaną informacje o dokładności lokalizacji dźwięku, odbioru fabuły, zrozumiałości rozgrywki i skuteczności dostosowania do osób niewidomych.

Referat 8.2

Lokalizacja źródeł dźwięku w polu rozproszonym wspomagana metodami uczenia maszynowego

Maciej Marcinkiewicz

Politechnika Warszawska

Lokalizacja źródeł dźwięku w przestrzeni przy pomocy matryc mikrofonowych jest zadaniem, do którego od lat 70. XX wieku są wykorzystywane metody cyfrowego przetwarzania sygnałów. Metody te są skuteczne w polu swobodnym, ale charakteryzują się mniejszą skutecznością w polu rozproszonym (pogłosowym). W celu poprawy jakości detekcji lokalizacji możliwe jest zastosowanie technik uczenia maszynowego i sieci neuronowych. Praca porównuje liniowe oraz okrągłe matryce mikrofonowe w kontekście detekcji lokalizacji źródeł dźwięków, analizuje dotychczasowo używane w tym celu metody cyfrowego przetwarzania sygnałów. Omawiana jest również propozycja systemu hybrydowego łączącego wspomniane metody wraz z rekurencyjną siecią neuronową służącą do detekcji liczby źródeł. Praca również prezentuje proces przygotowania danych treningowych do uczenia zaproponowanego systemu. W kolejnych etapach projektu będzie przeprowadzony trening i test systemu zarówno w polu swobodnym, jak i rozproszonym, a wyniki będą porównywane z istniejącymi rozwiązaniami pod kątem dokładności estymowanej liczby źródeł i ich lokalizacji w przestrzeni.

Referat 8.3**Analiza wpływu konfiguracji głośników ambisonicznych na percepcję sceny dźwiękowej****Andrzej Budny**

Politechnika Warszawska

Celem niniejszego projektu badawczego jest analiza wpływu konfiguracji układu głośników ambisonicznych na percepcję oraz immersję użytkowników w scenie dźwiękowej. Dźwięk przestrzenny stanowi kluczowy element tworzenia immersyjnych doświadczeń VR, a odpowiednie rozmieszczenie głośników może znacząco poprawić jakość oraz realizm dźwięku przestrzennego. Badania obejmują testowanie różnych układów geometrycznych głośników dla dźwięku w formacie ambisonicznym pierwszego i drugiego rzędu. Pomiary zostaną przeprowadzone w specjalistycznej komorze bezchowej, eliminującej pogłos i odbicia, wykorzystując zarówno obiektywne parametry akustyczne (poziom dźwięku, widmo, zniekształcenia), jak i subiektywne testy słuchowe z udziałem grupy słuchaczy. Uczestnicy będą oceniać dokładność lokalizacji dźwięku, jakość brzmienia oraz poziom immersji w skali MOS (Mean Opinion Score). Wyniki badań pozwolą zidentyfikować optymalne konfiguracje układów głośników, które poprawią realistyczne postrzeganie użytkownika w wirtualnym środowisku dźwiękowym.

Sesja IX – 30 marca

Referat 9.1

Rozszerzanie modalności warunkowania w modelach text-to-music

Mikołaj Szawerda

Politechnika Warszawska

Niniejsza praca przedstawia nowatorskie podejście do warunkowania muzycznych modeli generatywnych, koncentrując się na adaptacji inwersji tekstowej do modelu MusicGen – transformera syntezy audio warunkowanego tekstem. Zastosowana metoda umożliwia wzbogacenie wiedzy modelu bez typowych ograniczeń związanych z fine-tuningiem, wymagając przy tym mniejszych zasobów obliczeniowych. Technika umożliwia rozszerzenie sygnału warunkującego model na inne modalności. Dodatkowo pozwala na efektywną wymianę wyuczonych zanurzeń pomiędzy użytkownikami. W ramach badań przeprowadzono eksperymenty mające na celu ocenę skuteczności inwersji tekstowej dla różnych aspektów muzycznych, takich jak wspólny gatunek, brzmienie, tempo, emocje oraz dane spoza rozkładu treningowego. Dodatkowo przeanalizowano wpływ ilości danych oraz rozmiaru reprezentacji na jakość wyników, uwzględniając złożoność muzyczną analizowanych utworów. W celu optymalizacji procesu uczenia zaproponowano modyfikacje funkcji straty, które przyspieszają jednocześnie uczenie wielu konceptów oraz skracają czas treningu pojedynczego konceptu. Wyniki eksperymentów porównano z bazowym warunkowaniem tekstowym, modelem wzbogaconym o ekstraktor stylu oraz modelem warunkowanym audio – VampNet. Przeprowadzono również wstępne analizy adaptacji podejścia do innych modalności, takich jak obrazy determinujące generowany gatunek muzyczny. Uzyskane rezultaty wskazują na potencjał inwersji tekstowej jako elastycznego narzędzia do precyzyjnego sterowania muzycznymi modelami generatywnymi.

Referat 9.2

MidiTok Visualizer - Rozwój narzędzia do wizualizacji i analizy muzyki symbolicznej MIDI

Łukasz Pokorzyński

Politechnika Warszawska

Badania nad muzyką symboliczną odgrywają kluczową rolę w uczeniu maszynowym związanym z muzyką. W badaniach tych bardzo istotne są dane w formacie MIDI, które mogą być zbyt złożone dla osób bez wcześniejszego kontaktu z tym formatem danych lub bez doświadczenia muzycznego. Aby rozwiązać te problemy stworzyliśmy MidiTok Visualizer, aplikację internetową zaprojektowaną w celu ułatwienia eksploracji i wizualizacji różnych metod tokenizacji MIDI z pakietu MidiTok Python. MidiTok Visualizer oferuje wiele konfigurowalnych parametrów, umożliwiając użytkownikom przesyłanie plików MIDI w celu wizualizacji tokenizowanych danych wraz z interaktywną rolką fortepianu. Program jednocześnie oferuje przystępny interfejs użytkownika, który znacznie upraszcza i przyspiesza proces zapoznawania się ze strukturą plików MIDI. W referacie przybliżamy kolejne etapy rozwoju aplikacji, zastosowaną technologię, biblioteki, a także przedstawiamy dalsze pomysły na rozwój narzędzia.

Sesja plakatowa – 30 marca

Plakat P.1

Wykorzystanie archiwum dźwięków bbc w ocenie progu dyskomfortu słyszenia

Ewa Czepiczek, Bieńkowski O., Soszyńska N., Panek M.,
Hojan-Jezińska D., Stieler O.
Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu

Ocena progu dyskomfortu słuchowego jest kluczowym elementem w procesie dopasowania aparatów słuchowych i terapii dźwiękowych. Celem pracy było opracowanie materiału dźwiękowego, który mógłby zostać wykorzystany w gabinetach protetyki słuchu do subiektywnej oceny progu dyskomfortu dla dźwięków otoczenia. Wykorzystano materiał z archiwum dźwięków BBC Sound Effects, oferujące szeroką gamę dźwięków naturalnych o zróżnicowanej charakterystyce, które zostały poddane analizie spektrograficznej i widmowej. Wybrano sygnały akustyczne, które mogą posłużyć jako narzędzie wspomagające pracę specjalistów w ocenie poziomu akceptacji dźwięków w warunkach gabinetu. Ocena subiektywna umożliwia precyzyjne ustawienie maksymalnego poziomu wyjściowego i układów kompresji w aparatach słuchowych.

Plakat P.2**Powierzchnia soniczna do manipulacji małymi obiektami****Paweł Nikitin**

Politechnika warszawska

Metoda lewitacji sonicznej polega na unoszeniu i manipulacji obiektem z użyciem fal ultradźwiękowych za pośrednictwem powietrza. Umożliwia ona bezkontaktowe w kontekście ciał stałych przemieszczanie i obracanie obiektu w ograniczonej przestrzeni oddziaływania urządzenia do lewitacji. Artykuł ma charakter weryfikacyjny, autorzy podjęli się reprodukcji urządzenia do lewitacji sonicznej zaprojektowanego i opisanego w pracy Rafael Morales et al.[1]. Zweryfikowany przez nas w działaniu kompletny projekt powierzchni sonicznej składa się z zaprogramowanych algorytmów symulujących propagację i interferencję ultradźwięków w powietrzu, projektu elektronicznego i mechanicznego sprzętowej platformy sonicznej oraz oprogramowania do jej kontroli. Urządzenie odtworzone zgodnie z instrukcjami z powodzeniem unosi niewielkie obiekty i umożliwia przemieszczanie ich przez tworzenie krajobrazu dolin i pagórków sonicznych w kontrolowanej przestrzeni.

1. Morales, R.; Ezcurdia, I.; Irisarri, J.; Andrade, M.A.B.; Marzo, A. Generating Airborne Ultrasonic Amplitude Patterns Using an Open Hardware Phased Array. Appl. Sci. 2021, 11, 2981. <https://doi.org/10.3390/app11072981>

Plakat P.3

Ekstrakcja warstwy wokalnejs utworu muzycznego przy pomocy sieci neuronowych

Wojciech Polak

Politechnika Warszawska

Tematem tego artykułu był projekt aplikacji realizującej separację warstwy wokalnejs z pliku utworu muzycznego, wybieranego przez użytkownika. Utworzenie tego narzędzia było zmotywowane jego wykorzystaniem w procesie produkcji utworów muzycznych, przede wszystkim zrozumienie struktury akompaniamentu poprzez analizę oddzielnych części utworu. Omówiono w nim budowę i działanie splotowych sieci neuronowych, ze szczególnym uwzględnieniem architektury HDemucs, wykorzystywanejs w rozwiązaniu. Przedstawiono procedurę treningową sieci i jej implementację w tej pracy. Zaproponowano wykorzystanie wytrenowanego modelu w strukturze strony internetowej, umożliwiającej użytkownikowi interakcję z nim. W celu utworzenia interfejsu aplikacji i kodu serwera wykorzystano języki Python, HTML, CSS i JavaScript. Kod treningowy został napisany z użyciem biblioteki języka Python - PyTorch i skompilowany w środowisku Google Colab. Przeprowadzona procedura treningowa pozwoliła na stworzenie modelu separującego warstwę wokalną na poziomie SDR równym 3,5 dB. Zauważono, że operacja decymacji danych treningowych ma duży wpływ na jakość otrzymanego narzędzia. Na podstawie tych doświadczeń zaproponowano powtórzenie procesu treningowego, przy lepszych warunkach eksperymentalnych, co pozwoliłoby na wykorzystanie danych w oryginalnej częstotliwości próbkowania.



Kategoria I – Nagranie akustyczne

Nagranie K1.1

Adrian Szaposznikow - Sonata na flet i harfę

Jan Olejniczak

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Nagranie zrealizowano w Sali Koncertowej Akademii Muzycznej w Łodzi przy ul. Żubardzkiej 2A w czerwcu 2021 roku. Podczas sesji nagraniowej trwającej około 3 godziny zarejestrowano Sonatę na flet i harfę Adriana Szaposznikowa (XIX w.). Do konkursu zgłoszone zostało nagranie tylko I części sonaty.

W nagraniu uczestniczyli muzycy:

- Przemysław Szczypiński – flet
- Roksana Puchniarz – harfa

Ja byłem odpowiedzialny za reżyserię nagrania, montaż oraz miks.

Nagranie K1.2

Losa

Marcin Studniarz

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Eksplikacja nagrania utworu *Losa* Emmanuela Sejourne

Wykonawcy:

- Marimba – Nadia Mikołajczyk
- Wibrafon – Aniela Remiszewska

Nagranie zrealizowałem w maju 2024 roku w Sali Koncertowej Uniwersytetu Muzycznego Fryderyka Chopina w Warszawie.

Nagranie K1.3

Marta Ptaszyńska - Concerto Grosso cz. III

Wiktor Szymański

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Nagranie zostało zrealizowane na płytę (ukaze się w 2025 roku) dla wydawnictwa Naxos z utworami p. Marty Ptaszyńskiej. Jest to trzecia część koncertu na dwoje skrzypiec i orkiestrę kameralną. Jest to wymagający skład pod względem fonograficznym. Podczas nagrania zależało mi na zbudowaniu przestrzeni w obrazie stereofonicznym adekwatnej do materiału muzycznego. Ważnym aspektem było budowanie planów i balansu pomiędzy instrumentami. Soliści skrzypkowie zostali ustawieni przed kwintetem smyczkowym, a sekcja instrumentów dętych za nim. W najdalszym planie rozstawione były instrumenty perkusyjne. Nagranie było realizowane w korzystnej akustycznie sali koncertowej, jednak podczas postprodukcji materiału zdecydowałem się na powiększenie tej przestrzeni z użyciem sztucznych pogłosów.

Artyści:

- I skrzypce – Wojciech Koprowski
- II skrzypce – Roksana Kwaśnikowska
- Chopin University Chamber Orchestra
- Dyrygent – Mariusz Smolij

Kompozytorka: Marta Ptaszyńska

Kategoria II – Nagranie studyjne

Nagranie K2.1

HOPE- HANIA DEREJ

Antoni Cieplik

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Nagrania dokonano 18 października 2024 roku w Studio S1 na Uniwersytecie Muzycznym Fryderyka Chopina.

Podczas sesji nagraniowej zarejestrowania 4 utwory autorstwa Hani Derej z jej zespołem „Hania Derej trio”, które nie były nigdzie wcześniej publikowane.

Kompozytor: Hanna Derej

Zespół „Hania Derej Trio” w składzie:

Hanna Derej – Fortepian,

Kosma Góra-Gitara basowa,

Jakub Długoborski – Zestaw Perkusyjny.

Realizacja nagrania i Postprodukcja: Antoni Cieplik

Nagranie K2.2

Nardis - Miles Davis

Bartłomiej Dudziński

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Utwór „Nardis” Milesa Davisa to kompozycja o mrocznym i tajemniczym charakterze. Cały proces nagrania zaplanowałem tak, aby podkreślić istotę materiału muzycznego, i ukształtować przestrzenno-brzmieniowe właściwości dzieła. Zgodnie z praktyką nagraniową lat 50. XX wieku umieściłem muzyków blisko siebie, bez paneli akustycznych. Sesja nagraniowa została przeprowadzona w formie live session – muzycy grali akustycznie, bez odsłuchów.

Wykonawcy:

- Fortepian – Bartłomiej Mazurek
- Gitara elektryczna – Runc Jonatans Marks
- Kontrabas - Tymoteusz Wójtewicz
- Zestaw perkusyjny - Szymon Froń

Nagranie K2.3

Konstelacje

Piotr Kluzowicz

Politechnika Wrocławska

Utwór “Konstelacje” wykonany przez zespół Gosza w składzie: perkusja, gitara basowa, gitara elektryczna, instrumenty klawiszowe oraz wokół został nagrany na setkę. Dodatkowe wokale, chórki oraz organy hammonda zostały dograne po głównym nagraniu.

Kategoria III – Utwór immersyjny

Nagranie K3.1

Isthar

Jan Olejniczak

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Nagrania zrealizowano w Sali Koncertowej oraz studiu S1 UMFC w Warszawie w 2024 roku. Z racji na aranżację oraz akustykę i dostępność pomieszczeń nagraniowych nagranie zostało podzielone na dwie sesje. Pierwsza była realizowana w sali koncertowej bez udziału perkusjonaliów, które następnie były dogrywane do wcześniej nagranych materiałów w studiu S1. W nagraniach wzięli udział członkowie zespołu Heroes Orchestra. Podczas realizacji tego projektu byłem odpowiedzialny za reżyserię nagrań, montaż oraz miks i zgranie.

Nagranie K3.2

Tujon - „Pan w kapeluszu”

Szczepan Buśko

Akademia Muzyczna im. Feliksa Nowowiejskiego w Bydgoszczy

Współczesna technologia dźwięku przestrzennego otwiera nowe możliwości w dziedzinie produkcji muzycznej, pozwalając na osiągnięcie niezwykle realistycznego i angażującego efektu odsłuchowego. Jednym z najciekawszych rozwiązań w tej dziedzinie jest system Dolby Atmos 7.1.4, który umożliwia precyzyjne umiejscowienie dźwięków w trójwymiarowej przestrzeni. Niniejszy referat przedstawia proces nagrania i postprodukcji utworu „Pan w kapeluszu” zespołu Tujon (muz. Małgorzata Gołąb-Buśko, tekst Karolina Dudek), ze szczególnym uwzględnieniem technicznych aspektów pracy nad wersją immersyjną. Wykorzystanie zaawansowanych technik miksu, takich jak precyzyjna automatyzacja panoramy, separacja elementów w pionie i poziomie oraz kreatywne wykorzystanie efektów przestrzennych, pozwoliło na stworzenie immersyjnego środowiska dźwiękowego, które otacza słuchacza z każdej strony.

Nagranie K3.3

La Muerte del Ángel

Wiktor Szymański

Uniwersytet Muzyczny Fryderyka Chopina w Warszawie

Nagranie zrealizowałem w systemie 7.1.4. Do zarejestrowania warstwy 7.0 użyłem Neumannów U87 o charakterystyce dookólnej, natomiast do kanałów warstwy sufitowej użyłem Shcoepsów MK41 (charakterystyka superkardoidalna). Dla porównania użyłem też Schoepsów MK2S o charakterystyce dookólnej, jednak nie wykorzystałem ich w ostatecznym zgraniu. Każdy instrument został oddzielnie omikrofonowany. Korzystałem także z sygnałów ze wzmacniaczy gitarowych i basowych.

Artyści:

- Skrzypce – Olivia Bujnowicz
- Bandoneon – Kacper Kaczmarek
- Kontrabas – Jakub Kaczmarczyk
- Fortepian – Szymon Drabiński
- Gitara elektryczna – Tomasz Dziubiński

Kompozytor: Astor Piazzolla



Koła naukowe



Koło Naukowe Elektroakustyki Politechniki Warszawskiej

Koło Naukowe Elektroakustyki Politechniki Warszawskiej działa przy Zakładzie Elektroakustyki Instytutu Radioelektroniki i Technik Multimedialnych na Wydziale Elektroniki i Technik Informacyjnych Politechniki Warszawskiej. Opiekunem naukowym koła jest dr inż. Agnieszka Paula Pietrzak.

Głównym celem Koła Naukowego Elektroakustyki jest rozwijanie wiedzy i umiejętności członków poprzez organizację wykładów, warsztatów oraz praktycznych projektów. W obszarze działań KN jest nie tylko elektroakustyka, ale również sesje nagraniowe, realizacja dźwięku przestrzennego, zastosowanie AI do sygnałów dźwiękowych, akustyka mowy i wiele innych.

Członkowie KN Elektroakustyki prowadzą swoje projekty w studio nagraniowym oraz komorze bezdechowej, wspierają również organizację różnych konferencji i targów. Współpracuje również z innymi kołami naukowymi i organizacjami studenckimi, organizując wspólne inicjatywy takie jak realizacja koncertów studyjnych RA Studyjnie w studio nagrań Zakładu Elektroakustyki we współpracy z rozgłośnią radiową PW – Radio Aktywne. Ostatni projekt, jakim może się pochwalić, to realizacja albumu "M. kontra S. Musical (Original Cast Recording)" wraz z grupą musicalową Samodzielnego Zespołu Teatralnego Organizowanego w Liceum im. Stanisława Staszica w Warszawie!

Zapraszamy do śledzenia strony koła na Facebook
<https://www.facebook.com/knelektroakustyki>

Koło Naukowe Akustyki Architektonicznej AGH

Koło Naukowe Akustyki Architektonicznej jest obecnie największym kołem naukowym na Inżynierii Akustycznej na AGH. Zajmujemy się wieloma aspektami akustyki, nie tylko architektoniczną. Nie są nam straszne takie dziedziny, jak elektroakustyka, przetwarzanie sygnałów, tworzenie systemów pomiarowych, czy też projektowanie i modelowanie numeryczne metamateriałów akustycznych. Praca w Kole umożliwia nam również poszerzenie swoich umiejętności z zakresu przeprowadzania pomiarów i symulacji, niezbędnych do zbadania parametrów akustycznych wnętrz. Pomiary wykonujemy zarówno w obrębie uczelni w Laboratorium Akustyki Technicznej: komorze pogłosowej i bezechowej, jak i w terenie - na obszarach różnych inwestycji. W ramach działalności Koła organizujemy również różnego rodzaju pokazy, m.in. podczas Małopolskiej Nocy Naukowców, a także warsztaty np. Hackathon MetaSound, w trakcie którego projektowaliśmy oraz modelowaliśmy metamateriały akustyczne. Co roku bierzemy również udział w interdyscyplinarnych warsztatach "Nowa Przestrzeń" we współpracy z innymi uczelniami krakowskimi, na których projektujemy przestrzenie publiczne na terenie miasta Krakowa.

Międzyośrodkowe Studenckie Koło Naukowe IFPS i WUM

Międzyośrodkowe Studenckie Koło Naukowe przy Światowym Centrum Słuchu Instytutu Fizjologii i Patologii Słuchu oraz SKN Sekcja Otolaryngologiczna przy Zakładzie Niewydolności Serca i Rehabilitacji Kardiologicznej Warszawskiego Uniwersytetu Medycznego jest platformą do wymiany doświadczeń i zdobywania nowej wiedzy oraz dzielenia się pasją studentów medycyny, audiofonologii, informatyki i akustyki, psychologii, pedagogiki, ochrony środowiska, zarządzania, elektroniki, polonistyki i lingwistyki. Wielospecjalistyczny charakter Międzyośrodkowego Studenckiego Koła Naukowego otwiera możliwość współpracy i wymiany naukowej między różnymi ośrodkami akademickimi i badawczo-rozwojowymi, zarówno w kraju, jak i zagranicą. Działając przy Instytucie Fizjologii i Patologii Słuchu w Warszawie członkowie koła mają możliwość czynnego udziału w wielu ciekawych projektach badawczych i klinicznych na najwyższym poziomie, a dzięki wykorzystaniu zaawansowanych technik informatycznych i telemedycznych mogą uczestniczyć w codziennej pracy klinicznej Instytutu.

Studenckie Koło Naukowe Akustyków SONIC

Studenckie Koło Naukowe Akustyków SONIC to prawdopodobnie najgłośniejsze Koło na całym Uniwersytecie im. Adama Mickiewicza. Sonic skupia się na wydarzeniach naukowych oraz popularnonaukowych. Chętnie dzielimy się tajnikami świata dźwięku z uczniami poznańskich szkół średnich. Najbardziej kojarzyć nas można z recepcji i badań psychoakustycznych na festiwalu Soundedit w Łodzi. Sonic łączy ludzi z pasją do akustyki w każdym wydaniu – od zapalonych reżyserów dźwięku i samych muzyków, przez ludzi interesujących się nowinkami technologicznymi w branży audiologicznej, aż do osób poszukujących jeszcze swojej drogi w tej dziedzinie. Z legendarnym jeżem też mamy trochę wspólnego – niebieski kolor w logo i bycie najszybszą drużyną w sztafecie na Dniu Sportu UAM.

Koło Naukowe Inżynierii Dźwięku i Obrazu Politechniki Gdańskiej

Koło Naukowe Inżynierii Dźwięku i Obrazu (KNIDiO) od wielu lat aktywnie rozwija i promuje wiedzę z zakresu technologii audiowizualnych, multimedialnych oraz cyfrowego przetwarzania sygnałów. Członkowie naszego Koła realizują projekty badawcze, które z powodzeniem prezentują zarówno na konferencjach krajowych, jak i międzynarodowych.

Ponadto angażujemy się w lokalne inicjatywy multimedialne, takie jak organizacja warsztatów z zakresu realizacji nagrań audio i wideo oraz transmisji na żywo z ważnych wydarzeń na Politechnice Gdańskiej. Naszym celem jest nie tylko zdobywanie wiedzy, ale również dzielenie się nią oraz inspirowanie innych do rozwoju w dziedzinie nowoczesnych technologii multimedialnych.

Strona KN: <https://www.facebook.com/KNIDiO>