

POLITECHNIKA WARSZAWSKA

DYSCYPLINA NAUKOWA AUTOMATYKA, ELEKTRONIKA,
ELEKTROTECHNIKA I TECHNOLOGIE KOSMICZNE
DZIEDZINA NAUK INŻYNIERYJNO-TECHNICZNYCH

ROZPRAWA DOKTORSKA

mgr inż. Maciej Jasiński

**Wyznaczanie parametrów akustycznych pomieszczeń
z zastosowaniem przestrzennych odpowiedzi impulsowych**

Promotor
prof. dr hab. inż. Jan Żera

WARSZAWA, 2023

*Pragnę serdecznie podziękować promotorowi,
prof. Janowi Żerze za pomoc w przygotowaniu niniejszej
rozprawy, wszelkie rady i wskazówki, a przede wszystkim –
co najcenniejsze – poświęcony czas.*

*Koleżankom i Kolegom z Zakładu Elektroakustyki PW,
a także wszystkim współpracownikom za nieustanną
motywację do pisania niniejszej pracy.*

*Mojej narzeczonej Dorocie dziękuję z całego serca
za wsparcie, wyrozumiałość i cierpliwość w każdym
momencie.*

Streszczenie

Przedmiotem niniejszej rozprawy są badania możliwości wykorzystania współczesnych mikrofonów ambisonicznych pierwszego i trzeciego rzędu do pomiarów parametrów stosowanych w ocenie akustyki pomieszczeń. Mikrofony ambisoniczne pozwalają na wyznaczenie tzw. przestrzennej odpowiedzi impulsowej pomieszczenia, SRIR (ang. *Spatial Room Impulse Response*), co następnie, w połączeniu z funkcją przenoszenia głowy, HRTF (ang. *Head Related Transfer Function*), charakteryzującą wpływ obecności osoby w polu akustycznym, pozwala na wyznaczenie binauralnej przestrzennej odpowiedzi impulsowej pomieszczenia, BRIR (ang. *Binaural Room Impulse Response*). Użycie odpowiedzi impulsowych SRIR i BRIR umożliwia wyznaczenie odpowiednio wszechkierunkowych parametrów akustycznych pomieszczenia (czasu pogłosu, wczesnego czasu zaniku, czytelności i przejrzystości dźwięku, czasu środkowego echogramu) oraz, co najważniejsze, międzyusznej korelacji skrośnej sygnałów docierających do prawego i lewego ucha, która jest wskaźnikiem informującym o przestrzenności obrazu dźwiękowego. Badania obejmowały ocenę różnic wartości parametrów akustycznych sal uzyskanych za pomocą mikrofonów ambisonicznych oraz z użyciem mikrofonu pomiarowego, dwukanałowego manekina akustycznego i tzw. sztucznej głowy. Odstępstwa wartości parametrów wyznaczonych za pomocą mikrofonów ambisonicznych, w stosunku do pomiarów referencyjnych, nie przekraczają w większości przypadków jednego do dwóch progów (JND, ang. *Just Noticeable Difference*) postrzegania zmian parametrów akustycznych pomieszczenia. Z obserwacji tej wynika wniosek, że możliwe jest wyznaczanie parametrów akustycznych pomieszczenia na podstawie pojedynczego pomiaru przeprowadzonego mikrofonem ambisonicznym. Metoda taka, oparta na najnowszych technikach rejestracji dźwięku przestrzennego mikrofonami ambisonicznymi, może w istotny sposób uprościć procedurę pomiaru akustyki pomieszczeń.

Słowa kluczowe: ambisonia, dźwięk przestrzenny, parametry akustyczne pomieszczeń, przestrzenna odpowiedź impulsowa pomieszczenia SRIR, binauralna odpowiedź impulsowa pomieszczenia BRIR, funkcja przenoszenia głowy HRTF, HRIR

Abstract

The aim of the dissertation was to examine the application of first- and third-order ambisonic microphones in the measurement of room acoustics parameters. Ambisonic microphones enable to determine the Spatial Room Impulse Response (SRIR), which, when combined with the HRTF (Head Related Transfer Function), representing the filtering process that occurs when a person is present in the acoustic field, leads to the Binaural Room Impulse Response. The use of SRIR and BRIR impulse responses makes it possible to determine the omnidirectional room acoustics parameters (reverberation time, early decay time, sound clarity and intelligibility) and, more importantly, to determine the interaural cross-correlation coefficient of signals arriving at the left and right ear, which is an index of sound spaciousness in the room. The study comprised an evaluation of room acoustic parameters determined with the use of ambisonic microphones by a comparison with the parameters measured using a current standard method, employing a measurement microphone, a two-channel acoustic manikin and an artificial head. The results have shown that the values measured with the use of ambisonic microphones did not differ, in most cases, by more than one or two just-noticeable differences (JND) from the reference value of the given acoustic parameter. This finding indicates that room acoustic parameters can be determined with sufficient accuracy from a single measurement made with an ambisonic microphone which significantly simplifies the measurement procedure and is compatible with the latest trends in sound recording technology.

Keywords: ambisonics, spatial sound, room acoustic parameters, spatial room impulse response SRIR, binaural room impulse response BRIR, head related transfer function HRTF, HRIR

Spis treści

1. Wstęp.....	9
2. Parametry akustyczne pomieszczeń: definicje, kryteria, sposoby pomiaru	12
2.1. Parametry akustyczne pomieszczeń.....	12
2.2. Progowe różnice parametrów JND.....	14
2.3. Pomiary parametrów akustycznych z użyciem odpowiedzi impulsowych pomieszczenia.....	14
3. Techniki przetwarzania sygnałów ambisonicznych	16
3.1. Przestrzenna odpowiedź impulsowa pomieszczenia SRIR.....	16
3.2. Binauralna odpowiedź impulsowa pomieszczenia BRIR	17
3.3. Matryce mikrofonowe a mikrofony ambisoniczne.....	17
3.4. Rozkład sygnału na harmoniki sferyczne	18
3.5. Kodowanie sygnałów ambisonicznych.....	21
3.6. Mikrofony ambisoniczne analizowane w pracy	22
3.6.1. Budowa mikrofonu pierwszego rzędu.....	22
3.6.2. Budowa mikrofonu trzeciego rzędu	23
3.7. Aliasing przestrzenny.....	24
4. Funkcje przenoszenia głowy HRTF.....	25
4.1. Definicja HRTF	25
4.2. Metody binauralizacji sygnałów ambisonicznych oceniane pod kątem wyznaczania parametrów akustycznych pomieszczenia.....	27
4.2.1. Metoda Spatial Resampling	27
4.2.2. Metoda Least Squares.....	28
4.2.3. Metoda Time Alignment	28
4.2.4. Metoda Magnitude Least Squares.....	28
5. Przegląd literatury.....	30
5.1. Zastosowanie przestrzennych odpowiedzi impulsowych do wyznaczania parametrów akustycznych wewnątrz.....	30
5.2. Wyznaczanie kierunku dochodzenia dźwięku z użyciem sygnałów ambisonicznych	32
5.3. Metody binauralizacji	34
6. Cel, teza i zakres pracy.....	38
7. Analiza przestrzennej odpowiedzi impulsowej pomieszczenia	40

7.1.	Metodyka badań	40
7.2.	Wyznaczenie binauralnych odpowiedzi pomieszczeń BRIR.....	40
7.3.	Wyznaczenie funkcji HRTF manekina B&K4100-D	41
7.4.	Badane pomieszczenia	42
7.5.	Aparatura pomiarowa.....	43
8.	Wyniki badań.....	44
8.1.	Obserwacja echogramów	44
8.2.	Parametry akustyczne wyznaczone z dookólnej odpowiedzi impulsowej	49
8.2.1.	Czas pogłosu T30.....	49
8.2.2.	Czas wczesnego zaniku EDT.....	51
8.2.3.	Czytelność dźwięku C50	53
8.2.4.	Przejrzystość dźwięku C80.....	54
8.2.5.	Czas Środkowy Ts.....	56
8.3.	Międzyszyna korelacja skrośna IACC wyznaczana z binauralnej odpowiedzi impulsowej BRIR	57
8.3.1.	Porównanie pomiaru IACC przeprowadzonego przy pomocy manekina B&K 4100-D i sztucznej głowy Neumann KU-100.....	57
8.3.2.	Metoda Least Squares (LS).....	60
8.3.3.	Metoda Magnitude Least Squares (MagLS).....	62
8.3.4.	Metoda Time Alignment (TA).....	63
8.3.5.	Metoda Spatial Resampling (SPR).....	65
8.4.	Propozycja rachunku błędu pomiaru mikrofonem ambisonicznym	67
8.4.1.	Przykład oszacowania błędu pomiarowego dla mikrofonu Sennheiser Ambeo i Zylia ZM1-3E	70
9.	Podsumowanie i dyskusja.....	72
10.	Wnioski	76
	Bibliografia	79
	Załącznik nr 1: Definicje parametrów akustycznych pomieszczeń.....	86
	Załącznik nr 2: Schematy punktów pomiarowych w pomieszczeniach.....	88
	Załącznik nr 3: Analiza wariancji danych.....	91

1. Wstęp

Opis i ocena jakości akustyki pomieszczeń wymaga zdefiniowania zestawu wielkości i pojęć fizycznych oraz zastosowania właściwych metod, które jednoznacznie i powtarzalnie porównują wnętrza pod kątem ich akustyki. Zarówno definicja pojęć, jak i metody opisu właściwości akustycznych pomieszczeń były zawsze związane ze zjawiskami fizycznymi występującymi w pomieszczeniu oraz odpowiednimi cechami wrażeniowymi, a także były nierozwalnie związane z rozwojem techniki, w szczególności elektroniki oraz technik przetwarzania sygnałów.

Intensywny rozwój technik ambisonicznych jest ściśle powiązany z postępowaniem w dziedzinie multimedialnej nakierowanej na wykorzystanie dźwięku przestrzennego w systemach rozszerzonej i wirtualnej rzeczywistości. Dźwięk przestrzenny w takich systemach ma w jak najwiedzijszym stopniu zapewnić symulację lub odtworzenie zarejestrowanej trójwymiarowej przestrzeni dźwiękowej. Prace nad możliwie dokładną rejestracją przestrzennych obrazów dźwiękowych spowodowały postęp w konstrukcji mikrofonów ambisonicznych, a także w algorytmach przetwarzania sygnałów.

Celem niniejszej pracy jest sprawdzenie możliwości wykorzystania współczesnych mikrofonów ambisonicznych do pomiarów kompletnego zestawu parametrów służących ocenie akustycznej pomieszczeń, w szczególności wymienianych w normie PN-EN ISO 3382-1 [73] oraz parametrów najnowszych, związanych z binauralnością słyszenia. W tym celu wymagane jest wyznaczenie binauralnych odpowiedzi impulsowych BRIR (ang. *Binaural Room Impulse Response*) [98, 106] wykorzystujących transmitancję odniesioną do głowy HRTF (ang. *Head Related Transfer Function*) [106] do wyznaczania wskaźnika międzyuszy korelacji skrośnej IACC (ang. *Interaural Cross Correlation*) [7, 32] na podstawie rejestracji przeprowadzonej mikrofonami ambisonicznymi.

Pierwsze opisy akustyki pomieszczeń sięgają czasów rzymskich i polegały na słownym opisie pogłosowości oraz rozmieszczenia echa w teatrach [82]. Dopiero na przełomie wieków XIX i XX dzięki pracom Sabine'a [83, 84] zdefiniowano pojęcie czasu pogłosu oraz określono pierwszą metodę jego wyznaczania. Sabine do określenia czasu pogłosu w funkcji częstotliwości wykorzystywał piszczałki organowe o różnych wysokościach dźwięku oraz grupę przeszkolonych słuchaczy dla oceny momentu zaniku dźwięku (czas mierzony był stoperem). Miał to być czas mierzony od momentu wyłączenia źródła dźwięku w pomieszczeniu do chwili, w której dźwięk stawał się całkowicie niesłyszalny. Odpowiednikiem fizycznym jest określenie go jako spadek ciśnienia akustycznego milion razy, czyli o 60 dB. W późniejszych czasach,

wykorzystanie głośnika jako źródła dźwięku pozwoliło na zastąpienie nim piszczałek organowych zaproponowanych przez Sabine'a, a rozwój aparatury pozwolił na zastąpienie osoby mierzącej czas zaniku dźwięku przy użyciu stopera mikrofonowymi technikami zapisu i rejestracją sygnału przeprowadzaną analogowymi przyrządami pomiarowymi. Czas pogłosu był wyznaczany na podstawie analizy zarejestrowanej krzywej zaniku dźwięku w pomieszczeniu, przy pobudzeniu szumem pasmowym w kolejnych oktawach. Niezależną techniką była bezpośrednia rejestracja odpowiedzi pomieszczenia na sygnał impulsowy np. w postaci wystrzału pistoletowego. W takim przypadku, krzywa zaniku wyznaczana była na podstawie wstecznego całkowania zarejestrowanego impulsu. Dzięki jej wykreśleniu, możliwe było wprowadzenie innych, również energetycznych, parametrów pozwalających na ocenę jakości wnętrza pod kątem akustycznym (czas wczesnego zaniku EDT [37], wskaźniki przejrzystości C50, C80 [80], definicja D [57] i siła dźwięku G [41]). W przypadku techniki pobudzenia impulsowego, generowany impuls nie był powtarzalny i mógł powodować niepożądane zniekształcenia w torze pomiarowym. Obecnie stosowane metody, dzięki zastosowaniu cyfrowych technik przetwarzania sygnałów wprowadzają pośredni pomiar odpowiedzi impulsowej, przy pobudzeniu sygnałem o znanej sekwencji pseudolosowej MLS (ang. *Maximum Length Sequence*) [81, 99] lub tonem o przestrajanej częstotliwości [24, 25, 64]. Metody te pozwalają na zwiększenie powtarzalności uzyskiwanych wyników, a także dają zwiększenie ich dokładności poprzez redukcję wpływu zakłóceń. Dzięki swoim zaletom, metody te są ustandaryzowane [72, 73] i powszechnie stosowane.

Rozwój technik pomiarowych łączy się także z możliwościami dostępnych przetworników pomiarowych. Powszechnie stosowane mikrofony o charakterystyce dookólnej pozwalają na rejestrację podstawowych, uśrednionych przestrzennie parametrów akustycznych [4, 39, 40], czyli czasu pogłosu (T60, T30, T20), czasu wczesnego zaniku (EDT), czytelności dźwięku (C50), przejrzystości dźwięku (C80, czasu środkowego echogramu (Ts), tzw. siły dźwięku (G) oraz parametru Stage Support (ST) charakteryzującego warunki akustyczne na estradzie. Wielkości opisujące przestrzenność dźwięku we wnętrzu, tj. współczynnik międzyusznej korelacji skrośnej IACC oraz stosunek energii bocznej LF (ang. *Lateral Fraction*) wymagają odpowiednio zastosowania mikrofonów umieszczonych u wejścia kanałów słuchowych akustycznych symulatorów głowy (manekinów akustycznych), bądź mikrofonów o charakterystyce ósemkowej [73].

Dzięki możliwości arbitralnego kształtowania charakterystyki sygnałów zarejestrowanych techniką ambisoniczną [77, 111], wspomniane wyżej parametry pomieszczeń można wyznaczyć z przestrzennej odpowiedzi impulsowej SRIR (ang. *Spatial Room Impulse Response*)

[111]. Jej rejestrację przeprowadza się analogicznie do pomiarów z wykorzystaniem pojedynczego mikrofonu, poprzez rejestrację tonu przestrajanego, bądź sygnału MLS przez wszystkie kanały mikrofonu ambisonicznego. Zarejestrowane sygnały należy przetransformować do odpowiedzi impulsowych, a następnie przetworzyć z użyciem technik ambisonicznych, aby móc kształtować ich kierunkowość. Co więcej, filtracja ambisonicznych odpowiedzi impulsowych przez funkcje przenoszenia głowy HRTF osoby lub manekina akustycznego pozwala na uzyskanie binauralnej odpowiedzi impulsowej BRIR [98, 106] i wyznaczenie parametrów charakteryzujących przestrzeń dźwięku w pomieszczeniu w miejsce pomiaru bezpośredniego z użyciem manekina.

W niniejszej pracy opisano proces wyznaczania binauralnej przestrzennej odpowiedzi impulsowej SRIR i BRIR i obliczania na ich podstawie zestawu parametrów charakteryzujących własności akustyczne pomieszczenia. W drugim rozdziale opisano podstawowe parametry akustyczne pomieszczeń oraz współcześnie stosowane techniki pomiarowe. Trzeci rozdział zawiera informacje o technikach ambisonicznych. W czwartym rozdziale opisano funkcje HRTF wykorzystywane do binauralizacji sygnałów ambisonicznych. Piąty rozdział stanowi podsumowanie literatury. Cel, tezę i zakres niniejszej rozprawy opisano w rozdziale szóstym. Rozdział siódmy i ósmy opisują wyniki uzyskanych pomiarów. W ostatnim, dziewiątym rozdziale podsumowano wyniki przeprowadzonych badań.

Celem badań przeprowadzonych w rozprawie było uzyskanie zestawu parametrów akustycznych wewnątrz z użyciem mikrofonów ambisonicznych pozwalających na otrzymanie przestrzennej odpowiedzi impulsowej SRIR, a następnie – przy użyciu HRTF – binauralnej odpowiedzi impulsowej BRIR, i porównanie skuteczności tej metody z metodami klasycznymi, które wykorzystują mikrofon wszechkierunkowy i manekin akustyczny lub sztuczną głowę. Badania przeprowadzono w odniesieniu do sal o różnych warunkach pogłosowych: od pola swobodnego komory bezchowej, do pola dyfuzyjnego w komorze pogłosowej, uwzględniając warunki akustyczne studio średniej wielkości, sali koncertowej i korytarza jako pomieszczenia użytkowego o względnie dużym pogłosie. Porównano cztery znane z literatury przedmiotu metody uproszczenia transmitancji HRTF stosowane przy konstrukcji binauralnej odpowiedzi impulsowej pomieszczenia BRIR z ambisonicznej odpowiedzi impulsowej, LS (ang. *Least Squares*) [23, 87], MagLS (ang. *Magnitude Least Squares*) [87], TA (ang. *Time Alignment*) [109] oraz SPR (ang. *Spatial Resampling*). Kryterium oceny było porównanie wartości parametrów charakteryzujących sale otrzymanych na podstawie przestrzennych odpowiedzi impulsowych SRIR i BRIR z wartościami pochodzącymi z pomiarów mikrofonem pomiarowym i manekinem akustycznym.

2. Parametry akustyczne pomieszczeń: definicje, kryteria, sposoby pomiaru

2.1. Parametry akustyczne pomieszczeń

W niniejszym rozdziale przedstawiono parametry pomieszczeń [4, 39, 40], które w dalszej części pracy są wykorzystywane przy ocenie metod ambisonicznych ich wyznaczania. Historycznie najwcześniejszym parametrem jest czas pogłosu wprowadzony przez Sabine'a [83, 84] jak to przedstawiono w pracach opisanych wcześniej we wstępie niniejszej rozprawy. Kolejne parametry akustyczne pomieszczeń: definicja D [57], czas wczesnego zaniku EDT [37], wskaźniki przejrzystości C50, C80 [80], i siła dźwięku G [41] pojawiały się poczynając od lat pięćdziesiątych XX wieku w celu uzyskania zgodności pomiarów fizycznych z ocenami jakości dźwięku w salach. Zbiór parametrów akustycznych pozwala na obiektywną ocenę jakości wnętrza, a także porównywanie między sobą pomieszczeń o podobnym przeznaczeniu. Parametry są także wykorzystywane na etapie projektowania – do ustalenia wytycznych, które w końcowym etapie realizacji umożliwiają weryfikację spełnienia założeń projektowych. W sensie fizycznym, opisują one wpływ pomieszczenia na propagację dźwięku pomiędzy parą źródło – odbiornik umieszczoną w określonych punktach pomiarowych w pomieszczeniu. Definicje parametrów określają właściwości pola akustycznego w określonych punktach położenia słuchacza, a możliwości ich wyznaczenia – jak wspomniano we wstępie – są ściśle związane z możliwościami i rozwojem technik pomiarowych.

Wielkości opisujące cechy pola akustycznego we wnętrzach można podzielić na dwie podstawowe grupy, z których pierwsza, z uwagi na pomiar przy pomocy pojedynczego mikrofonu o charakterystyce dookólnej, pozwala na określenie wielkości, zasadniczo niezależnych od kierunku dochodzenia fali akustycznej. Do tych parametrów zaliczają się wspomniane: czas pogłosu (T60, T30, T20), czas wczesnego zaniku EDT, czytelność dźwięku C50, przejrzystość dźwięku C80, czas środkowy echogramu Ts oraz siła dźwięku G.

Druga grupa parametrów ma na celu określenie przestrzennych cech pola akustycznego o charakterze kierunkowym. Wiąże się to ze stosowaniem mikrofonów o charakterystyce ósemkowej lub układu dwóch mikrofonów, np. takich, jakie występują w przypadku manekina akustycznego, z mikrofonami zamontowanymi u wejścia przewodu słuchowego po obu stronach głowy. Na podstawie pomiaru przy użyciu mikrofonu ósemkowego wyznacza się stosunek energii bocznej LF, będący stosunkiem energii bocznej odbić docierających do słuchacza.

Natomiast przy użyciu manekina akustycznego wyznacza się współczynnik międzyusznej korelacji skrośnej IACC (ang. *Interaural Cross-Corelation Coefficient*).

Współczynnik międzyusznej korelacji skrośnej IACC sygnałów docierających do lewego i prawego ucha słuchacza [32, 73] ma podstawowe znaczenie, ponieważ za jego pomocą można wyznaczyć wszystkie pozostałe parametry przestrzenne, takie jak wskaźnik jakości odwzorowania binauralnego BQI [7] (ang. *Binaural Quality Index*), czy też postrzegane otoczenie słuchacza dźwiękiem LEV [8] (ang. *Listener Envelopment*). Badania psychoakustyczne polegające na ocenach słuchowych dźwięku w pomieszczeniu pokazują, że słabe skorelowanie sygnałów lewego i prawego ucha jest związane ze wzrostem wrażenia przestrzenności dźwięku w audytorium lub sali koncertowej, i odwrotnie, silna korelacja jest świadectwem niekorzystnej, znikomej przestrzenności dźwięku, ponieważ pełna korelacja ($IACC \approx 1$) odpowiada sygnałowi monofonicznemu podanemu dwuusznie [32, 67]. Parametr IACC jest wyznaczany na podstawie jednoczesnej rejestracji odpowiedzi impulsowych pomieszczenia w punktach u wejścia do lewego i prawego kanału słuchowego. Ten rodzaj odpowiedzi impulsowej określany jest w literaturze jako BRIR (ang. *Binaural Room Impulse Response*). Pomiar BRIR i dalej wyznaczenie IACC przeprowadza się dwuusznie typowo z wykorzystaniem manekina akustycznego o cechach antropometrycznych, wyposażonego w dwa mikrofony położone w miejscach odpowiadających położeniem uszom słuchacza, którego przykładem są manekin Brüel & Kjaer (B&K) typ 4101-D [14], oraz tzw. sztuczna głowa Neuman KU-100 [66] zasadniczo przeznaczona do studyjnych nagrań dźwiękowych.

Wymienione parametry akustyczne pomieszczeń odpowiadają następującym parametrom wrażeniowym sali [4, 39, 40]. Czas pogłosu i EDT służą określeniu pogłosowości sali, przy czym obecnie parametr EDT uważa się za ważniejszy, ponieważ obejmuje wczesną część przedziału zaniku dźwięku, silnie związaną z postrzeganym wrażeniem pogłosowości. Wskaźnik przejrzystości dźwięku C50 charakteryzuje czytelność mowy, a wskaźnik przejrzystości C80 związany jest z czytelnością muzyki. Wskaźniki te są silnie skorelowane ze sobą, ponieważ różnią się jedynie przedziałem uśredniania początkowego fragmentu echogramu. Czas uśredniania energii w przedziale 50 ms w przypadku wskaźnika C50 jest właściwszy dla oceny czytelności mowy, natomiast uśrednienie energii w pierwszych 80 ms odpowiedzi impulsowej wskaźnika C80 uznawany jest za właściwy do oceny właściwości sali na rzecz określonego zespołu i gatunku muzycznego. Wskaźniki binauralnej, takie jak BQI, LEV i ASW są pochodną międzyusznej korelacji skrośnej IACC [67], która przede wszystkim określa przestrzenność zależną od odbić bocznych w pomieszczeniu.

2.2. Progowe różnice parametrów JND

Dla oceny wartości parametrów sal i występujących różnic, co ma również znaczenie przy ocenie metod opisanych w dalszej części niniejszej pracy, istotne są wypracowane w badaniach progowe wartości zmian tych parametrów JND (ang. *Just Noticeable Difference*), których przekroczenie prowadzi do zmian wrażeniowych w sali. Wartości przyjętych progów JND dla poszczególnych parametrów przedstawiono w tabeli 2.1, a wynikają one z porównań pomiarów fizycznych z ocenami słuchowymi. Wartości JND są także przytoczone w normie PN-EN ISO 3382-1 [73] i stanowią istotne kryterium służące do porównywania sal, a także metod pomiarowych, wykorzystywane w wielu publikacjach [76, 105].

Tabela 2.1 Progowe wartości (JND) postrzegania różnic parametrów akustycznych według [73, 105]

Parametr	JND
Czas pogłosu, T30	5%
Czas wczesnego zaniku, EDT	5%
Czytelność dźwięku, C50	1 dB
Przejrzystość dźwięku, C80	1 dB
Czas środkowy, Ts	10 ms
Współczynnik międzyusznej korelacji skrośnej IACC	0,075

Ogólnie, w stosunku do parametrów czasowych, takich jak EDT (a także czasu pogłosu), próg postrzegania JND jest wyrażony jako 5% zmiany. Natomiast w przypadku pomiarów energetycznych, typową wartością JND jest 1 dB, co jest związane z tym, że parametry te są wynikiem znacznego uśredniania warunków w salach z uwagi na liczbę punktów pomiarowych.

2.3. Pomiary parametrów akustycznych z użyciem odpowiedzi impulsowych pomieszczenia

Z punktu widzenia teorii sygnałów, pomieszczenie stanowi system liniowy, niezmienny w czasie [68, 99]. Jak już to przedstawiono, pierwsze klasyczne techniki pomiarowe w akustyce pomieszczeń opierały się na pomiarach energetycznych i ocenie spadku energii w pomieszczeniu po wyłączeniu źródła. Te techniki nie pozwalały na wyznaczenie

transmitancji. techniki cyfrowe Natomiast pozwalają dla danej kombinacji źródło-odbiornik wyznaczyć transmitancję o charakterze zespolonym, opisaną za pomocą odpowiedzi impulsowej.

Od strony pomiarowej, powszechnie wykorzystywaną metodą wyznaczenia odpowiedzi impulsowej jest pobudzenie pomieszczenia sygnałem pseudolosowym MLS (ang. *Maxium Length Sequence*) [81, 99] lub przestrajającym tonem (ang. *swepped sine*) [24, 25], a następnie wyznaczenie odpowiedzi impulsowej odpowiednio poprzez transformatę Hadamarda (MLS) lub Fouriera (*swepped sine*). Techniki te są dotychczas najpopularniejszymi metodami rejestracji odpowiedzi impulsowych pomieszczenia i jako takie są przywołane w normie PN-EN ISO 3382 [73] oraz PN-EN ISO 18233 [72] standaryzujące metody pomiaru parametrów charakteryzujących akustykę pomieszczeń. Ugruntowana pozycja tych metod pozwala na łatwe uogólnienie w kierunku przestrzennych odpowiedzi impulsowych sali.

3. Techniki przetwarzania sygnałów ambisonicznych

Terminem ambisonii określa się odwzorowanie dźwięku z zachowaniem jego cech przestrzennych w przestrzeni trójwymiarowej. Pojęcie to, a także pierwsze podstawy teoretyczne zostały wprowadzone przez Gerzon'a w 1973 roku [28]. Cechą szczególną ambisonii jest wykorzystanie do rejestracji dźwięku zespołu mikrofonów, często kierunkowych kardioidalnych, które pozwalają na zachowanie informacji o rozkładzie geometrycznym pola akustycznego w punkcie położenia mikrofonu, zarówno w płaszczyźnie poziomej, jak i pionowej. Odwzorowanie pola akustycznego przy odtwarzaniu wymaga dokonania konwersji umożliwiającej właściwą reprodukcję dźwięku za pomocą głośników.

Z określeniem kierunkowości dźwięku przestrzennego w pomieszczeniu związane są takie pojęcia, jak przestrzenna odpowiedź impulsowa (SRIR – ang. *Spatial Room Impulse Response*) [111] oraz binauralna odpowiedź impulsowa pomieszczenia (BRIR – ang. *Binaural Room Impulse Response*) [106] opisane w punktach 3.1 i 3.2. Obliczana odpowiedź binauralna BRIR łączy cechy przestrzennej odpowiedzi impulsowej SRIR oraz właściwości głowy i małżowin usznych umieszczonych w polu akustycznym, wyrażonych transmitancją HRTF (ang. *Head Related Transfer Function*) [106] lub jej odpowiednikiem w dziedzinie czasu – odpowiedzią impulsową HRIR (ang. *Head Related Impulse Response*), a jest obliczeniowym odpowiednikiem pomiaru z użyciem manekina akustycznego. W dalszej części rozdziału, omówiono konstrukcje mikrofonów ambisonicznych oraz sposoby kodowania i dekodowania sygnałów ambisonicznych niezbędne do wykorzystania w analizie przestrzennej odpowiedzi impulsowej pomieszczeń.

3.1. Przestrzenna odpowiedź impulsowa pomieszczenia SRIR

Przestrzenna odpowiedź impulsowa pomieszczenia SRIR, oprócz informacji o amplitudzie i czasie docierania poszczególnych odbić, pozwala także na poznanie kierunku dochodzenia fali akustycznej. Genezą analizy przestrzennej odbić są metody geometryczne, stosowane w modelowaniu akustyki pomieszczeń. Przestrzenną odpowiedź impulsową wyznacza się na podstawie śledzenia promieni akustycznych docierających od zamodelowanego źródła dźwięku do odbiornika. Na tej podstawie, w punkcie, w którym umieszczony jest odbiornik można wyznaczyć kierunek odbić, a także ich amplitudę i moment czasowy, w którym są rejestrowane. Odpowiednikiem pomiarowym odpowiedzi impulsowej wynikającej z symulacji

numerycznych pomieszczenia jest jej uzyskanie z użyciem kierunkowych matryc mikrofonowych typu ambisonicznego.

3.2. Binauralna odpowiedź impulsowa pomieszczenia BRIR

Jak wcześniej wspomniano, binauralna odpowiedź impulsowa pomieszczenia BRIR uwzględnia wpływ filtracji wprowadzanej przez głowę, małżowinę uszną i tors na dźwięk docierający do wejścia przewodów słuchowych. W związku z tym, odbiornikiem nie jest pojedynczy mikrofon pomiarowy, lecz para mikrofonów umieszczonych u wejścia do przewodów słuchowych. Z punktu widzenia teorii sygnałów, w najbardziej ogólnym podejściu, tak zdefiniowana BRIR stanowi splot odpowiedzi impulsowej pomieszczenia z binauralną odpowiedzią impulsową odniesioną do głowy w postaci HRIR (ang. *Head Related Impulse Response*), opisaną w rozdziale 4.

3.3. Matryce mikrofonowe a mikrofony ambisoniczne

Pod pojęciem matrycy mikrofonowej opisuje się dowolną liczbę pojedynczych kapsuł mikrofonowych działających współbieżnie. W zależności od układu geometrycznego mikrofonów, wyróżnia się matryce: liniową, w której ułożenie mikrofonów pokrywa się z linią prostą; dwumikrofonową, jako minimalną liczbę mikrofonów matrycy liniowej, ale często z ich rozdzieleniem przeszkodą akustyczną, co ma miejsce w przypadku mikrofonów umieszczonych u wejścia przewodów słuchowych manekina usznego; kołową, w której mikrofony umieszczone są na obwodzie koła; powierzchniową, której mikrofony rozmieszczone są w określony sposób na zdefiniowanej płaszczyźnie; matryce mikrofonowe sferyczne – z mikrofonami umieszczonymi na powierzchni sfery, często stanowiąc wierzchołki wielościanów foremnych. Matryce mikrofonowe różnych kształtów, w tym w szczególności matryce powierzchniowe są od dawna stosowane w pomiarach akustycznych [11, 19]. Natomiast w wersji jako mikrofony ambisoniczne pojawiają się dopiero przy rozwoju tej techniki do zagadnień studyjnych.

W użyciu równoległe funkcjonują określenia sferycznych matryc mikrofonowych oraz mikrofonów ambisonicznych. W rozdziale 3.4 opisano harmoniki sferyczne, które reprezentują kierunkowość sygnałów ambisonicznych poszczególnych rzędów. Fizycznie możliwe jest bezpośrednio uzyskanie sygnału ambisonicznego maksymalnie rzędu pierwszego poprzez zastosowanie mikrofonu dookólnego oraz ortogonalnie ustawionych mikrofonów o charakterystyce dwukierunkowej – ósemkowej. Przykładem takiej realizacji mikrofonu ambisonicznego, którego wyjście odpowiada bezpośrednio kodowaniu ambisonicznemu jest mikrofon Nimbus-

Halliday [18]. Idealny mikrofon ambisoniczny powinien mieć kapsuły mikrofonowe umieszczone w jednym punkcie. Ponieważ nie jest to fizycznie możliwe, został przez Gerzona [28] opracowany mikrofon, który składa się z czterech kapsuł o charakterystyce kardoidalnej umieszczonych na planie czworościanu foremego. Mikrofon ten często nazywany jest Soundfield, z uwagi na nazwę firmy, która go pierwotnie wytwarzała.

Odrębną konstrukcję stanowią sferyczne matryce mikrofonowe. Miniaturowe mikrofony są w nich umieszczone na powierzchni sztywnej sfery. Komercyjnym przykładem takich rozwiązań są mikrofony Zylia ZM1-3E (19 kapsuł) [112] oraz Eigenmike (32 lub 64 kapsuły) [59, 60].

W niniejszej pracy termin mikrofonu ambisonicznego odnoszony jest zarówno do klasycznej konstrukcji typu Soundfield, jak i sferycznych matryc mikrofonowych pozwalających na konwersję sygnału do formatu ambisonicznego.

3.4. Rozkład sygnału na harmoniki sferyczne

Rozkład sygnału na harmoniki sferyczne jest rozkładem charakterystyki kierunkowości przestrzennej na ortogonalną bazę funkcji, w pewnym sensie realizowaną w sposób analogiczny do ortogonalnego rozkładu sygnału czasowego na funkcje trygonometryczne w analizie Fouriera [77, 98, 104, 107, 111].

Aby poprawnie zarejestrować dźwięk dochodzący z dowolnego kierunku w przestrzeni, należałoby wykorzystać matrycę o nieskończonej liczbie punktowych mikrofonów umieszczonych w jednym punkcie i skierowanych promieniście we wszystkich kierunkach. Każdy z tych mikrofonów musiałby mieć nieskończenie wąską charakterystykę kierunkowości. Wszystkie te założenia czynią to podejście fizycznie nierealizowalnym. Praktycznym ograniczeniem jest liczba kanałów możliwa do wykorzystania oraz kierunkowość mikrofonów.

Koncepcja przestrzennej analizy Fourierowskiej kierunków na sferze odnosi się do ortogonalnego zbioru tzw. harmonik sferycznych [77, 104] i sięga wstecz koncepcji pochodzących od Legendre'a i Laplace'a. Kierunkowość sygnału w polu akustycznym może być opisana nieskończoną sumą ortogonalnych harmonik sferycznych, których zwiększający się rząd opisuje z coraz większą szczegółowością rozdzielczość kierunkowości. Kierunkowość sygnału opisywana we współrzędnych przestrzennych może być wyrażona współczynnikami wagowymi odnoszonymi się do poszczególnych harmonik sferycznych, co jest odpowiednikiem reprezentacji Fourierowskiej w dziedzinie częstotliwości.

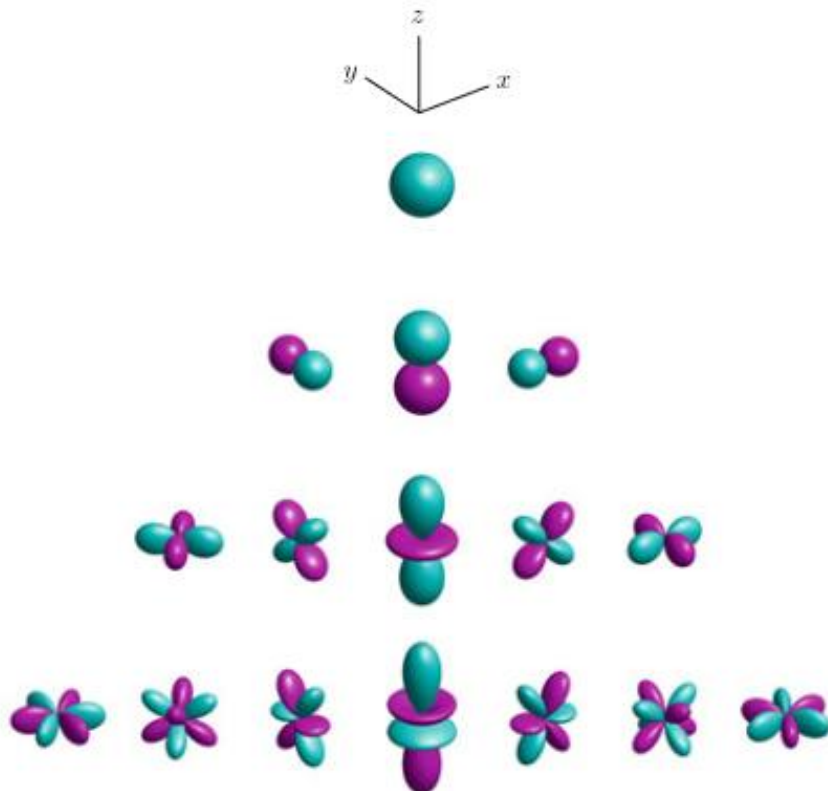
Ponieważ analizowane w tej pracy mikrofony ambisoniczne są mikrofonami pierwszego i trzeciego rzędu, na rysunku 3.1 i w tabeli 3.1 przedstawiono kształty (tzw. biegunowe wykresy balonowe) harmonik od zerowego do trzeciego rzędu dla uwidocznienia szczególności z jaką analizowana jest charakterystyka kierunkowości. Należy zauważyć, że inne analizy, np. analiza kierunkowości funkcji HRTF z użyciem harmonik przestrzennych mogą wykorzystywać harmoniki sferyczne nawet trzydziestego lub wyższych rzędów [23, 87, 109]. Harmoniki tego rzędu byłyby trudne do wizualizacji w postaci przedstawionych na rys. 3.1.

Każda z harmonik sferycznych [77, 104] zdefiniowana jest na jednostkowej sferze, jako (3.1):

$$Y_n^m(\theta, \phi) \equiv \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \quad (3.1)$$

gdzie: n – określa rząd funkcji, m – stopień funkcji, P_n^m – odpowiadającą postać funkcji Legendre’a.

W tabeli 3.1 przedstawiono postaci harmonik sferycznych do rzędu 3, czyli dla wartości $n = 0, 1, 2$ lub 3 , a wówczas m przyjmuje kolejne wartości całkowite z przedziału od $-n$ do n dla utworzenia zbioru harmonik sferycznych związanych z określonym rzędem n .



Rys. 3.1. Postacie harmonik sferycznych rzędów $n = 0$ do $n = 3$.

Tabela 3.1. Funkcje harmonik sferycznych rzędów $n = 1, 2, 3$

$n = 0$	$Y_0^0(\theta, \phi) = \sqrt{\frac{1}{4\pi}}$
$n = 1$	$Y_1^{-1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi}$
	$Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos \theta$
	$Y_1^1(\theta, \phi) = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi}$
$n = 2$	$Y_2^{-2}(\theta, \phi) = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{-2i\phi}$
	$Y_2^{-1}(\theta, \phi) = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{-i\phi}$
	$Y_2^0(\theta, \phi) = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$
	$Y_2^1(\theta, \phi) = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}$
	$Y_2^2(\theta, \phi) = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{2i\phi}$
$n = 3$	$Y_3^{-3}(\theta, \phi) = \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{-3i\phi}$
	$Y_3^{-2}(\theta, \phi) = \sqrt{\frac{105}{32\pi}} \cos \theta \sin^2 \theta e^{-2i\phi}$
	$Y_3^{-1}(\theta, \phi) = \sqrt{\frac{21}{64\pi}} (5 \cos^2 \theta - 1) \sin \theta e^{-i\phi}$
	$Y_3^0(\theta, \phi) = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$
	$Y_3^1(\theta, \phi) = -\sqrt{\frac{21}{64\pi}} (5 \cos^2 \theta - 1) \sin \theta e^{i\phi}$
	$Y_3^2(\theta, \phi) = \sqrt{\frac{105}{32\pi}} \cos \theta \sin^2 \theta e^{2i\phi}$
	$Y_3^3(\theta, \phi) = -\sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{3i\phi}$

W przypadku rzędu $n = 0$ występuje jedynie dookólna charakterystyka kierunkowości. W przypadku rzędu $n > 0$ charakterystyka kierunkowości wynika z sumowania harmonik sferycznych rzędów od 0 do n , przy zastosowaniu odpowiednich współczynników wagowych. Liczba harmonik sferycznych do opisanie n -tego rzędu ambisonii wynosi $(n+1)^2$.

Bezpośrednia realizacja harmonik sferycznych zerowego i pierwszego rzędu jest możliwa za pomocą mikrofonu wszechkierunkowego i zestawu mikrofonów ósemkowych. Przykładem takiego mikrofonu jest wspomniany wcześniej mikrofon Nimbus-Halliday [18], składający się z mikrofonu wszechkierunkowego oraz dwóch mikrofonów ósemkowych. Innego typu mikrofony ambisoniczne, takie jak mikrofon Gerzona [28], dysponujący czterema kapsułami, muszą mieć przekodowanie sygnałów otrzymywanych bezpośrednio z mikrofonu, czyli tzw. formatu A, do formatu B (patrz punkt 3.6.1). Mikrofony wyższych rzędów wymagają analizy sygnału dającej rozkład kierunkowości na harmoniki sferyczne.

3.5. Kodowanie sygnałów ambisonicznych

Sygnały z mikrofonów ambisonicznych w większości nie odpowiadają fizycznie sygnałom o charakterystykach opisanych harmonikami sferycznymi. Stąd przyjęto konwencję kodowania sygnałów w formatach ambisonicznych. Każdy z sygnałów w formacie A odpowiada sygnałom bezpośrednio zarejestrowanym przez każdy z mikrofonów. Aby uzyskać sygnał w formacie B, odpowiadającym poszczególnym harmonikom sferycznym – należy zastosować odpowiednią kombinację liniową sygnałów formatu A. Istnieją dalsze formaty C i D, które stosuje się przy reprodukcji dźwięku z użyciem wielokanałowych systemów odsłuchowych [77, 98, 107, 111].

Zapis sygnału ambisonicznego w formacie B wymaga ujednoczenia konwencji przypisania kanałów odpowiadającym poszczególnym harmonikom sferycznym. Odpowiednie przypisanie kanałów jest niezbędne w celu zapewnienia właściwego działania stosowanych algorytmów przetwarzania sygnałów. Obecnie najpopularniejszą konwencją jest system AmbiX [111] stosujący porządkowanie kanałów według konwencji ACN (ang. *Ambisonic Channel Number*). Konwencja ta umożliwiła uszeregowanie poszczególnych postaci harmonik sferycznych pod względem polaryzacji, a przypisanie odpowiednich kanałów odbywa się zgodnie z zależnością $ACN = n^2 + n + m$, gdzie n i m to odpowiednio rząd i stopień harmoniki, analogicznie do wzoru (3.1), punkt 3.4.

Poprawne działanie algorytmów przetwarzania sygnałów ambisonicznych, niezbędnych przy dekodowaniu sygnałów wymaga zastosowania normalizacji amplitud odpowiadających poszczególnym harmonikom sferycznym. W formacie AmbiX zastosowano normalizację z wykorzystaniem współczynników SN3D (*Schmidt quasi-normalization*) określaną według zależności (3.2):

$$N_{n,m}^{SN3D} = \sqrt{(2 - \delta_m) \frac{(n - |m|)!}{(n + |m|)!}}, \quad \delta_m \begin{cases} 1 \text{ dla } m = 0 \\ 0 \text{ dla } m \neq 0 \end{cases} \quad (3.2)$$

gdzie: n - rząd harmoniki sferycznej, m - jej postać. Tak zdefiniowana normalizacja kanałów pozwala na zapewnienie, że poziomy sygnałów pochodzących z harmonik wyższych rzędów nie przekroczą poziomu sygnału harmoniki rzędu zerowego.

Historycznie istniały także inne konwencje porządkowania sygnałów, których przykładem jest system FuMa opracowany przez Furse'a i Malhama [45, 111]. Jest to obecnie mniej popularny system, którego literowa notacja pozwalała na wykorzystanie w ambisonii maksymalnie czwartego rzędu [65]. W tej konwencji stosowano normalizację z wykorzystaniem

współczynników maxN [45], która zakładała normalizację każdego z kanałów w taki sposób, aby żaden nie przekraczał wzmocnienia równego jeden. W przeciwieństwie do normalizacji SN3D, współczynniki maxN w przypadku wyższych rzędów ambisonii, są wyliczane numerycznie na podstawie znajomości funkcji opisującej daną postać harmoniki i nie mogą być wyrażone za pomocą prostej zależności [45], której przykładem jest zależność (3.2).

3.6. Mikrofony ambisoniczne analizowane w pracy

W badaniach przeprowadzonych w pracy użyto mikrofon ambisoniczny pierwszego rzędu Sennheiser Ambeo [88] oraz trzeciego rzędu Zylia ZM1-3E [112]. Podstawowe informacje odnoszące się do mikrofonów przedstawiono w punktach 3.6.1 i 3.6.2.

3.6.1. Budowa mikrofonu pierwszego rzędu

Zwyczajowo poszczególne kanały mikrofonu Sennheiser Ambeo pierwszego rzędu w formacie A oznaczone są jako LF, LB, RF, RB (litery oznaczające odpowiednio lewą i prawą stronę oraz tył i przód: ang. *left, right, front, back*). Oznaczenie związane z położeniem mikrofonów w wierzchołkach czworościanu foremego odpowiednio ukierunkowanych w sposób łącznie uwzględniający płaszczyznę pionową i poziomą.

Dodatkowo, ze względu na ukośne ustawienie mikrofonów względem układu XYZ (XY – płaszczyzna pozioma, Z – kierunek pionowy), możliwe są dwie kombinacje rozmieszczenia kapsuł mikrofonu ambisonicznego U (ang. *up*) i D (ang. *down*), co daje dwie różne kombinacje ustawienia mikrofonów: FLU-FRD-BLD-BRU (Stosowana w mikrofonie Soundfield [28], i Sennheiser Ambeo [88]), oraz FLD-FRU-BLU-BRD (np. historycznie stosowane w mikrofonie DPA-4).

Dla mikrofonu pierwszego rzędu Sennheiser Ambeo, wykorzystywanego w niniejszej pracy, związek pomiędzy poszczególnymi sygnałami, a harmonicznymi sferycznymi jest następujący:

$$X = LF - LB + RF - RB$$

$$Y = LF + LB - RF - RB$$

$$Z = LF - LB - RF + RB$$

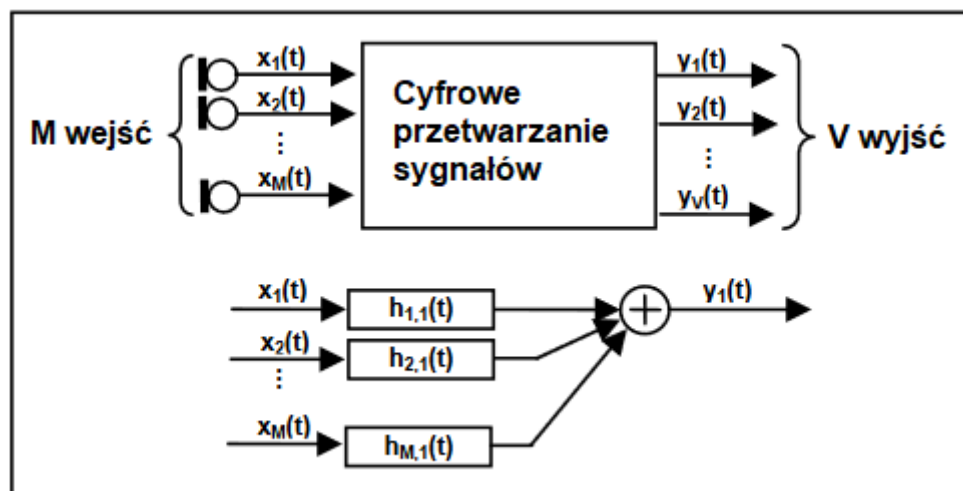
$$W = LF + LB + RF + RB$$

W przypadku większej niż 4 liczby kapsuł mikrofonowych nie stosuje się symbolicznego opisu kanałów, jak to pokazano powyżej na przykładzie niniejszego mikrofonu pierwszego rzędu.

3.6.2. Budowa mikrofonu trzeciego rzędu

Mikrofonem ambisonicznym wyższego rzędu jest matryca mikrofonowa Zylia ZM1-3E [112] składająca się z 19 kapsuł mikrofonowych typu MEMS pozwalających na uzyskanie sygnału ambisonicznego trzeciego rzędu.

Konwersja sygnałów 19 przetworników mikrofonu Zylia z formatu A do formatu B trzeciego rzędu wymaga uzyskania 16 kanałów odpowiadających liczbie harmonik sferycznych mikrofonu trzeciego rzędu [1]. Ogólny schemat przetwarzania M sygnałów z mikrofonów na V sygnałów B-formatu przedstawiono na rysunku 3.2 [27].



Rys. 3.2. Schemat przetwarzania sygnałów z formatu A do B [27]

Przetwarzanie wymaga zastosowania zestawu $M \times V$ filtrów. Zakładając, że sygnał x_m jest sygnałem dowolnego z M mikrofonów, y_v wyjściem sygnału B-formatu z każdej z harmonik sferycznych, a $h_{m,v}$ jest macierzą filtrów – proces przetwarzania można wyrazić za pomocą wzoru (3.3):

$$y_v(t) = \sum_{m=1}^M x_m(t) * h_{m,v}(t) \quad (3.3)$$

Istnieją dwie techniki uzyskiwania zestawu filtrów tworzących sygnał z wirtualnych mikrofonów o kierunkowości opisanej postaciami harmonik sferycznych [27]. W pierwszej

z nich – obliczeniowej – M sygnałów jest przetwarzana w celu uzyskania odpowiedniej liczby harmonik sferycznych. Zbiór filtrów h jest obliczany poprzez rozwiązania równań bazujących na równaniu falowym, zakładając, że mikrofony nie różnią się między sobą. Drugie podejście zakłada uzyskanie filtrów h bezpośrednio w pomiarach odpowiedzi impulsowych mikrofonu ambisonicznego w komorze bezchowej. Definiuje się problem odwrotny, w którym znana jest postać sygnału – określona harmonikami sferycznymi. Na tej podstawie, znając odpowiedź impulsową z każdej z kapsuł mikrofonowych, można obliczyć postać filtru h generującego harmonikę sferyczną.

3.7. Aliasing przestrzenny

Jak wiadomo, w zjawiskach związanych z propagacją fali występuje aliasing, nazywany w tym przypadku aliasingiem przestrzennym [10, 79], który wynika z niedostatecznego próbkowania pola akustycznego za pomocą skończonej liczby harmonik sferycznych w sensie relacji do długości fali. W technikach ambisonicznych zjawisko aliasingu przestrzennego spowodowane jest ograniczonym rzędem wykorzystywanych mikrofonów. Częstotliwość graniczną [10], od której występuje zjawisko aliasingu przestrzennego określa się zależnością (3.4):

$$f_a = \frac{c}{2R\gamma}, \quad (3.4)$$

gdzie:

c – prędkość dźwięku [m/s]

R – promień sfery na której umieszczone są przetworniki [m]

γ – maksymalny kąt pomiędzy dwoma przetwornikami, który jest zdefiniowany przez rząd mikrofonu.

Dla mikrofonu Ambeo pierwszego rzędu, stosowanego w niniejszej pracy – górna częstotliwość graniczna f_a , powyżej której występuje zjawisko aliasingu przestrzennego, wynosi ok. 6 kHz [29, 71]. Dla mikrofonu Zylia ta częstotliwość graniczna to ok. 3,3 kHz [53].

4. Funkcje przenoszenia głowy HRTF

Jak wspomniano w rozdziale 2, wyznaczenie BRIR, ale także odtworzenie sygnału bin-auralnego z sygnału ambisonicznego w formacie B wymaga uwzględnienia wpływu obecności głowy słuchacza, bądź też manekina pomiarowego na dźwięk docierający do wejść przewodów słuchowych. Wpływ ten opisuje para funkcji HRTF (ang. *Head Related Transfer Functions*) opisanej w punkcie 4.1.

4.1. Definicja HRTF

Charakter filtracji dźwięku docierającego z dowolnego punktu w przestrzeni do wejścia zewnętrznego przewodu słuchowego silnie zależy od rozmiarów i kształtu struktur anatomicznych. Opisanie funkcji transmitancji układu słuchowego jest możliwe za pomocą pary filtrów o transmitancji HRTF, czyli funkcji przejścia z punktu położenia źródła dźwięku w przestrzeni do wejścia przewodu słuchowego, z uwzględnieniem obecności głowy i torsu. Definicję pary HRTF H_L oraz H_P [90, 106] odpowiednio dla ucha lewego i prawego przedstawiono równaniami (4.1):

$$\begin{aligned} H_L &= H_L(r, \theta, \varphi, f, a) = \frac{P_L(r, \theta, \varphi, f, a)}{P_0(r, f)}, \\ H_P &= H_P(r, \theta, \varphi, f, a) = \frac{P_P(r, \theta, \varphi, f, a)}{P_0(r, f)}, \end{aligned} \quad (4.1)$$

gdzie H_L i H_P to HRTF odpowiednio dla ucha lewego i prawego, P_L i P_P to widmo ciśnienia akustycznego rejestrowane u wejścia kanału słuchowego obojga uszu. W celu normalizacji funkcji HRTF konieczna jest rejestracja P_0 , ciśnienia akustycznego (w postaci widma) zarejestrowanego w miejscu środka głowy badanej osoby, lecz bez jej obecności [106].

Równanie (4.1) pokazuje zależność HRTF nie tylko od położenia źródła dźwięku w przestrzeni, które jest określone współrzędnymi biegunowymi (r, θ, φ) , lecz także jako głównego czynnika od częstotliwości f źródła dźwięku oraz dodatkowo indywidualnych cech anatomicznych reprezentowanych we wzorze (4.1) parametrem a . Na indywidualne cechy anatomiczne składają się między innymi kształt i rozmiar głowy, małżowiny usznej oraz tułowia. Różnice cech anatomicznych skutkują silną personalizacją HRTF. Parametr a we wzorze (4.1) reprezentujący cechy anatomiczne ma znaczenie formalne, nie ma miejsca jego wyznaczenie w drodze pomiarów. Przy odległości źródła $r < 1$ m, wynik pomiaru HRTF zależy od odległości r . Funkcje HRTF rejestrowane są jednak zazwyczaj w polu dalekim, przy $r > 1$ m ($f > 5$ kHz),

w którym funkcja HRTF jest niezależna od odległości r . Funkcjom HRTF odpowiada para odpowiedzi impulsowych HRIR (ang. *Head Related Impulse Response*), będąca ich transformata Fouriera. W kategoriach pomiarowych, przy udziale osoby lub manekina akustycznego, przeprowadza się rejestrację odpowiedzi impulsowych HRIR w funkcji kąta (θ, φ) stosując techniki pomiarowe wykorzystujące MLS [2, 22] i ton przestrajany [44, 48].

Koncepcja funkcji HRTF wynika z wczesnych badań E. Shaw'a [89, 90], w których w postaci charakterystyk amplitudowo-częstotliwościowych ucha wyznaczono wpływ odbić i zacieniania akustycznego wynikającego z obecności małżowiny usznej (ze szczególnym znaczeniem muszli ucha), głowy i torsu na dźwięk docierający do zewnętrznego przewodu słuchowego. Zmiany wzmocnienia w funkcji kąta przy częstotliwości powyżej 5 kHz sięgają 15-20 dB, co stanowi o jednousznych możliwościach lokalizacyjnych słuchu. Pierwsze prace o charakterze badawczym, które dotyczą funkcji HRTF jako transmitancji zespolonej Wightmana i Kistler [102, 103] stanowią punkt wyjścia do dzisiejszych szerokich zastosowań funkcji HRTF. Istotne prace w zakresie pomiarowym to prace [31, 63, 90], w których wykazano, że transmitancje od punktu w polu akustycznym do bębienka można rozdzielić na osobne transmitancje od punktu w polu akustycznym do zablokowanego wejścia przewodu słuchowego oraz transmitancję samego przewodu słuchowego. Prace te uzasadniają umieszczanie mikrofonu rejestrującego sygnał zaraz u wejścia do przewodu słuchowego.

Niezależnie od bezpośrednich pomiarów z udziałem osoby lub manekina akustycznego [22, 35, 42], najczęściej przeprowadzanych w warunkach pola swobodnego komory bezechowej, możliwe jest dopasowanie funkcji HRTF z istniejących licznych baz danych (np. CIPIC [2], ARI [44], SADIE-II [3]) będących efektem pracy szeregu laboratoriów w świecie. Obecnie rozwinięte są metody doboru funkcji HRTF z baz danych wykorzystujące algorytmy polegające na testach odsłuchowych lokalizacji dźwięku [92, 97], a także na doborze HRTF pod względem podobieństwa cech antropometrycznych uwzględnionych w bazach danych [86, 91]. W ostatnich latach powstały metody wyznaczania funkcji HRTF na podstawie skanu geometrii głowy [12, 110]. W niniejszej pracy metoda ta została wykorzystana w odniesieniu do wyznaczenia funkcji HRTF manekina B&K 4100-D, uzyskując zgodność z wcześniej przeprowadzonymi pomiarami HRTF z użyciem tego manekina [35, 36]. W przypadku głowy studyjnej KU100, HRTF został przyjęty z bazy danych SADIE-II, w której wyznaczony był na podstawie pomiarów [3].

4.2. Metody binauralizacji sygnałów ambisonicznych oceniane pod kątem wyznaczania parametrów akustycznych pomieszczenia

Uzyskanie sygnału binauralnego z sygnału ambisonicznego polega w ogólności na właściwym dobraniu filtrów HRTF z bazy danych. Zasadniczym czynnikiem jest właściwe dopasowanie zbioru punktów HRTF do kierunków nadejścia dźwięku zarejestrowanych w sygnale ambisonicznym po jego przekształceniu do formatu B. Istota problemu polega na minimalizacji błędu wynikającego z różnicy siatki pomiarów HRTF i rozdzielczości przestrzennej sygnału ambisonicznego związanej z dekompozycją na harmoniki sferyczne przy ograniczeniu wynikającego z rzędu ambisonii. Dopasowanie zbioru funkcji HRTF polega na ich dekompozycji za pomocą harmonik sferycznych, które muszą odpowiadać rzędowi przetwarzanego sygnału ambisonicznego [23]. Funkcje HRTF zawierają istotną część energii aż do rzędu $n = 30$ (co odpowiada 961 kanałom ambisonicznym) [109]. W związku z tym, redukcja rzędu HRTF do niskich rzędów ambisonii wprowadza znaczne zniekształcenia sygnału związane z aliasingiem przestrzennym [23, 108, 109, 111].

W punktach od 4.2.1 do 4.2.4 przedstawiono cztery podejścia pozwalające na minimalizację błędu związanego z ograniczaniem rzędu harmonik sferycznych funkcji przenoszenia głowy HRTF wybrane dla wyznaczania parametru IACC pomieszczenia. Obejmują one metody Least Squares (LS), Magnitude Least Squares (MagLS), Time Alignment (TA) i Spatial Resampling (SPR).

4.2.1. Metoda Spatial Resampling

Binauralizacja pośrednia polega na zdekodowaniu zarejestrowanego sygnału ambisonicznego do sygnału możliwego do odtworzenia wirtualnym, trójwymiarowym systemem odsłuchowym. Podejście to jest określane z języka angielskiego jako Spatial Resampling lub Spatial Subsampling [9, 23]. W tym przypadku, możliwe jest ograniczenie rzędu funkcji przenoszenia głowy HRTF poprzez ograniczenie liczby wykorzystywanych transmitancji do kierunków zgodnych z rozmieszczeniem głośników w wirtualnej przestrzeni. Liczba niezależnych kanałów, dla których wybiera się HRTF powinna wynosić $(n + 1)^2$, gdzie n jest rzędem ambisonii. Każdy z kanałów jest następnie filtrowany przez HRTF (lub splatany z odpowiadającą odpowiedzią impulsową HRIR). W tym podejściu zniekształcenia spowodowane aliasingiem są obecne, jednak wprowadzają mniejszy błąd niż w przypadku bezpośredniej redukcji rzędu HRTF [9].

4.2.2. Metoda Least Squares

W tej kategorii algorytmów przetwarzanie polega na bezpośrednim wykorzystaniu sygnału ambisonicznego opisanego z użyciem harmonik sferycznych n -tego rzędu, czyli w formacie B. Dopasowanie zbioru HRTF polega na jego dekompozycji do odpowiadających harmonik sferycznych n -tego rzędu. Jak wcześniej wspomniano, redukcja rzędu zbioru funkcji HRTF skutkuje wprowadzeniem zniekształceń widmowych wynikających z aliasingu przestrzennego. Można je ograniczać, stosując minimalizację błędu z wykorzystaniem metody najmniejszych kwadratów, a także stosując wstępne przetwarzanie transmitancji HRTF według przytoczonych algorytmów.

W najprostszym podejściu, nazywanym metodą najmniejszych kwadratów – LS (ang. *Least Squares*) [87, 109] minimalizowany jest wprost błąd ograniczenia HRTF do określonego rzędu ambisonii. W tym podejściu, z powodu wymienionych ograniczeń, obserwowane są zniekształcenia widma w zakresie wyższych częstotliwości [87]. Jedną z metod poprawy algorytmu LS jest wyrównanie widma HRTF zastosowane po ograniczeniu ich rzędu [6]. Pomimo poprawy tłumienia widma w zakresie wysokich częstotliwości, testy odsłuchowe [87, 109] wykazały, że dla niższych rzędów ambisonii w sygnale pozostają pewne zniekształcenia.

4.2.3. Metoda Time Alignment

Kolejnym wariantem metody najmniejszych kwadratów jest zastosowanie algorytmu przesunięcia czasowego TA (ang. *Time Alignment*) [109] przed ograniczeniem rzędu zbioru transmitancji HRTF. Operacja ta polega na usunięciu informacji o międzyusznej różnicy czasu ITD (ang. *Interaural Time Difference*), co pozwala na znaczne ograniczenie rzędu funkcji HRTF. Przykładowe ograniczenie do rzędu $n = 5$, pozwala na zachowanie nawet 90% energii przy częstotliwości 10 kHz. Ocena tego algorytmu przez Brinkmanna i Weinzierla [13] wykazała, że przy ograniczeniu rzędu do $n = 3$ (stosowanego również w niniejszej pracy) zdolności lokalizacyjne, jak i wprowadzane zniekształcenia widmowe były podobne, jak w przypadku ograniczenia HRTF do rzędu $n = 19$ i nie stosowania przesunięcia czasowego TA.

4.2.4. Metoda Magnitude Least Squares

Trzecią opisywaną metodą jest MagLS (ang. *Magnitude Least Squares*) [87], która polega na pominięciu różnic fazowych w transmitancjach powyżej pewnej częstotliwości odcięcia f_c , przyjmowanej często jako równą częstotliwości granicznej aliasingu f_a dla danego rzędu

ambisonii. W przeciwieństwie do metody Time Alignment, metoda MagLS pozwala na redukcję rzędu zbioru HRTF przy zachowaniu informacji o międzyusznej różnicy czasu. Badania Schörkhubera i in. [87] wykazały, że błędy fazy przy ograniczeniu rzędu HRTF do $n = 3$ nie są zauważalne powyżej częstotliwości 2 kHz w przypadku sygnałów ciągłych (np. mowy) lub powyżej 4 kHz w przypadku sygnałów przerywanych.

5. Przegląd literatury

Jak wcześniej wspomniano, wykorzystanie techniki ambisonicznej w pomiarach akustycznych sal zostało zaproponowane już w 1975 roku przez Gerzon'a [30]. Wybiegając w przyszłość, autor przedstawił koncepcję utworzenia przestrzennych odpowiedzi impulsowych pomieszczenia uzyskanych przy użyciu mikrofonu ambisonicznego. Koncepcja ta polegała na rejestracji odpowiedzi impulsowej, np. strzału pistoletowego (Gerzon [30] wspomina już o innych technikach wyznaczania odpowiedzi impulsowej gwarantujących lepszy stosunek sygnału do szumu), z użyciem czterech kanałów mikrofonu ambisonicznego do dalszego wyznaczania przestrzennej odpowiedzi impulsowej. Opisane techniki stały się możliwe w realizacji dopiero później, wraz z możliwościami techniki cyfrowego zapisu i przetwarzania danych. Za punkt wyjścia można przyjąć pracę Fariny i Ayalon'a z 2003 roku [26], w której autorzy bezpośrednio powołują się na koncepcję Gerzona [30] i wprowadzają realizację tej idei.

Prac poświęconych wyznaczaniu parametrów akustycznych sal w sposób ściśle odpowiadający tradycyjnie ugruntowanym parametrom akustycznym jest niewiele i zaczynają się od wspomnianej pracy Fariny i Ayalona [26]. Przegląd tych prac przedstawiony jest w punkcie 5.1. Ponieważ możliwości wyznaczenia przestrzennych odpowiedzi impulsowych oraz parametrów binauralnych sal są ściśle związane z technikami wyznaczania kierunku dochodzenia dźwięku od poszczególnych źródeł, niektóre prace z tego zakresu, te które wykorzystują w tym celu mikrofony ambisoniczne, ale nie zajmują się klasycznymi parametrami akustycznymi sal, zostały przedstawione w punkcie 5.2.

Poważnym problemem obliczeniowym i pomiarowym jest binauralizacja przestrzennej odpowiedzi impulsowej w drodze jej połączenia z HRTF i ocena dokładności tego procesu. Wiele prac z tego zakresu pojawiło się w ostatnich dziesięciu latach. Ich krótki przegląd jest przedstawiony w punkcie 5.3.

5.1. Zastosowanie przestrzennych odpowiedzi impulsowych do wyznaczania parametrów akustycznych wnętrz

Farina i Ayalon [26] wykorzystali technikę ambisoniczną z użyciem mikrofonu ambisonicznego Soundfield ST-250 (B-format pierwszego rzędu) do rejestracji przestrzennych odpowiedzi impulsowych dziewięciu sal teatralnych i koncertowych (Włochy, Grecja, Japonia) i wyznaczenia parametrów akustycznych T30, C50, C80, D, Ts i G z użyciem kanału wszechkierunkowego W, a także parametrów przestrzennych LF/LFC z użyciem kanału Y

pierwszego rzędu. Natomiast parametr międzyusznej korelacji skrośnej IACC nie był wyznaczany mikrofonem ambisonicznym, ale został zmierzony z użyciem sztucznej głowy Neumann KU-100. Para mikrofonów kardioidalnych w układzie techniki ORTF (charakterystyki kardioidalne pod kątem 110° i rozsuniecie mikrofonów) służyła dodatkowo do weryfikacji prawidłowości rejestracji przestrzenności pola akustycznego rejestrowanego mikrofonem ambisonicznym. Praca ta, pomimo uwzględnienia parametrów akustycznych sal, w większym stopniu koncentruje się na rejestracji dźwięku mikrofonem ambisonicznym do późniejszego przetwarzania dla uzyskania przestrzennego obrazu dźwiękowego. Aspekt pomiaru parametrów akustycznych ogranicza się do przedstawienia nowej metodologii pomiarowej, bez analizy uzyskanych parametrów.

Pollack i in. [74] przeprowadzili pomiary parametrów akustycznych 14 paryskich sal teatralnych, operowych i koncertowych z użyciem mikrofonu ambisonicznego 1-go rzędu Soundfield ST-250. Do obliczenia parametrów T30, EDT, TS, G i C50 wykorzystano ambisoniczny kanał dookólny W, a do wyznaczenia parametru przestrzennego LF dodatkowo kanał Y. Celem pracy było porównanie sal, a nie zagadnienia metodologiczne. Jak widać, z punktu widzenia metodologii pomiarowej, praca ta jest podobna do wcześniej przedstawionej pracy [26].

Clapp i in. [16, 17] wykorzystywali w pomiarach sferyczną matrycę mikrofonową o ambisonii drugiego rzędu zaprojektowaną i wykonaną we własnym laboratorium. Pomimo zaawansowanej, 16-mikrofonowej matrycy, obie prace pokazują wyniki w sposób ograniczony. Pracy [16] na podstawie harmonik sferycznych wyznaczono rejestrację kanałem W parametrów T30 i C80, natomiast międzyuszną korelację skrośną IACC, podobnie jak w pracy [26], wyznaczono na podstawie odrębnych pomiarów binauralnej odpowiedzi impulsowej z użyciem manekina akustycznego. W pracy [17] przedstawiono rozkłady pola akustycznego w określonych miejscach na widowni.

W pracy Panton i in. [70] wykorzystano w pomiarach ambisoniczny 4-go rzędu mikrofon Eigenmike. Przedmiotem badań była analiza parametrów akustycznych estrady. Zarejestrowane parametry to czas pogłosu T30, siła dźwięku G, a także parametr Stage Support [7, 39, 40] określający wzajemną słyszalność muzyków na estradzie. Istotnymi elementami tej pracy było porównanie wyników z uzyskanymi za pomocą referencyjnego wszechkierunkowego mikrofonu pomiarowego (B&K typ 4189) oraz przeprowadzenie pomiaru parametru G i Stage Support w sposób kierunkowy. W tym celu wykorzystano kierunkowość mikrofonu ambisonicznego z podziałem na 6 kierunków dochodzenia odbić w układzie sześciangu, w poziomie i pionie. Porównanie z referencyjnym mikrofonem pomiarowym pokazało, że niewielki jest wpływ charakterystyki mikrofonu ambisonicznego na uzyskiwane wyniki parametrów.

W mniej typowych pomieszczeniach niż sale audytoryjne i teatralne były prowadzone prace [20, 46, 47] zespołu Akademii Górniczo Hutniczej w Krakowie. W pracy [47] poddano badaniom akustykę ośmiu cerkwi prawosławnych w południowej Polsce z użyciem mikrofonu ambisonicznego 1-go rzędu Soundfield ST-350, przy czym główne parametry akustyczne wyznaczono kanałem W, a wskaźnik energii bocznej LF z użyciem kanału Y. Prace [20, 46] odnosiły się do akustyki obszarów lodowych Spitsbergenu i Svalbardu, a w pomiarach stosowano mikrofon ambisoniczny pierwszego rzędu Soundfield SPS-200. W tym przypadku do wyznaczania parametrów akustycznych używano kanału W, a zachowanie informacji o kierunkowej odpowiedzi impulsowej służyło do możliwości późniejszej auralizacji przestrzennej dźwięku tych obszarów.

Na szczególną uwagę zasługuje praca Rafaely i Avni [78], w której przeprowadzono oszacowanie dokładności wyznaczenia międzyuszej korelacji skrośnej IACC przeprowadzanej na szereg sposobów, w tym czysto obliczeniowych, przy wykorzystaniu oprogramowania do modelowania akustyki wewnątrz Odeon i funkcji HRTF z bazy CIPIC oraz mikrofonu sferycznego Eigenmike, w połączeniu z funkcjami HRTF manekina KEMAR. Odniesieniem były bezpośrednie pomiary parametru IACC z użyciem tego manekina. Praca ta wykazała dużą zgodność pomiarów przeprowadzonych dwiema metodami, a jej znaczenie wynika ze szczególnego podjęcia tematu wykorzystania wyższych rzędów ambisonii.

Zgodnie z najlepszą wiedzą autora, rozwijanie metod pomiaru z użyciem mikrofonów ambisonicznych w minionych piętnastu latach obejmuje niewiele prac, ponieważ nie udało się wyodrębnić innych prac poza wymienionymi, co świadczy o nieugruntowaniu tych metod w praktyce pomiarowej. Celem większości prac było zastosowanie ambisonii do tworzenia wirtualnej przestrzeni dźwiękowej 3D w drodze tzw. „renderingu” 3D. Wskazuje to wobec tego na słuszność podjęcia tematyki wyznaczania mikrofonami ambisonicznymi parametrów akustycznych pomieszczeń oraz celu przyjętego w tej pracy porównania wyników uzyskanych z użyciem techniki mikrofonu ambisonicznego z klasycznymi technikami pomiaru pojedynczym mikrofonem wszechkierunkowym oraz w układzie binauralnym manekina akustycznego.

5.2. Wyznaczanie kierunku dochodzenia dźwięku z użyciem sygnałów ambisonicznych

W niniejszym punkcie omówiono wybrane prace związane z lokalizacją źródeł dźwięku z użyciem mikrofonów ambisonicznych z uwagi na bliski związek tej tematyki z wyznaczaniem przestrzennej odpowiedzi impulsowej i jej binauralizacji z użyciem funkcji HRTF. Zagadnienie

to występuje przy określaniu kierunku dojścia fali bezpośredniej i wczesnych odbić pomieszczenia.

Koncepcyjnie wyznaczanie kierunku dojścia fali DOA (ang. *Direction of Arrival*) sprawdzane jest do wyznaczenia wektora natężenia dźwięku na podstawie danych z mikrofonu ambisonicznego pierwszego rzędu, co oznacza odpowiednie użycie kanału W oraz kierunkowych X, Y, Z. W literaturze mówi się niekiedy o wektorze pseudonatężenia dźwięku dla odróżnienia od technik ściśle pomiarowych, np. takich, jakie omówione są w pracach [62, 100].

Artykuł Merimaa i Pulkki [56] wyznacza wektory natężenia dźwięku w przestrzeni 3D na podstawie kanałów W, X, Y, Z, przy czym dla późniejszego umiejscowienia źródła dźwięku (rendering) w obrazie kreowanym przez system odsłuchowy wystarczająca jest znajomość kąta położenia wektora natężenia w przestrzeni. Dla techniki tej przyjęto nazwę Spatial Impulse Response Rendering (SIRR).

Menzer i in. [55] wyznacжали składowe I_x , I_y i I_z wektora natężenia dźwięku na podstawie sumy iloczynów próbek pochodzących odpowiednio z kanałów ambisonicznych X, Y i Z i kanału W. Ze względu na ósemkową charakterystykę kierunkowości kanałów X, Y, Z oraz dookólną kanału W, odpowiada iloczynowi składowych prędkości akustycznej v_x , v_y , v_z i ciśnienia akustycznego p . A zatem kierunek, z którego dochodzi dźwięk obliczano na podstawie orientacji wypadkowego wektora pseudonatężenia dźwięku w układzie współrzędnych sferycznych. Wektor pseudonatężenia wyznaczano wyłącznie dla dźwięku bezpośredniego, stanowiącego wycinek z pierwszych 10 ms przestrzennej odpowiedzi impulsowej pomieszczenia. Określenie kierunku wydzielonej składowej bezpośredniej było punktem wyjścia zaproponowanej metody binauralizacji odpowiedzi impulsowej pomieszczenia wspomnianej dalej w punkcie 5.3.

Wierzbicki, Małecki i Wiciak [101] porównali obie wymienione wyżej metody i zaproponowali własną, która zwiera pewne modyfikacje obliczeniowe. Porównawczo metoda Merimaa i Pulkki [56] opiera się na oszacowaniu wartości wektora X, Y lub Z, metoda Menzer i Faller'a, na obliczeniu sumy iloczynów próbek kanałów X, Y lub Z i kanału W, natomiast metoda autorów, sumy iloczynów kanałów X, Y lub Z i znaku kanału W. Artykuł wskazuje na poprawę dokładności lokalizacji uzyskaną ostatnią metodą.

Odrębne podejście wnosi praca Jarret i in. [34], w której do wyznaczenia wektora natężenia dźwięku używane są wiązki własne (eigenbeams) harmonik sferycznych pierwszego rzędu. Autorzy wykazują wyższą skuteczność tej metody w stosunku do metody SRP (Steered Response Power) [58, 59] polegającej na beamformingowym skanowaniu mapy energetycznej w sferycznym układzie współrzędnych kąta azymutu i elewacji.

W pracy Tervo i in. [96], wyznaczenie przestrzennej odpowiedzi impulsowej jest związane z dekompozycją harmonik sferycznych w zbiór źródeł pozornych, które później mogą być również powiązane z lokalizacją dźwięku w systemach odsłuchowych. W kontekście pomiarów, tego typu analiza może być przydatna przy lokalizacji wczesnych odbić w kontekście doboru funkcji HRTF.

Natomiast w pracy McCormack i in. [51] ma miejsce rozwinięcie wcześniej omówionej metody SIRR [41] na wyznaczenie kierunku nadejścia dźwięku przy zastosowaniu mikrofonów ambisonicznych wyższych rzędów (większych od 1). Artykuł ten pokazuje, że wyniki obliczeń są dokładniejsze w stosunku do oryginalnej metody SIRR [56], jak i również metody lokalizacji proponowanej w pracy Tervo i in. [96].

Trudność w zagadnieniach związanych z wyznaczeniem kierunku położenia źródła dźwięku leży również w zależnościach częstotliwościowych związanych z długością fali zgodnie ze zjawiskiem aliasingu przestrzennego. W związku z tym, w często importowanych, rozwiązaniach matematycznych dotyczących zagadnień antenowych takich jak np. algorytm MUSIC [85] lub MVDR [15] pochodzących z okresów, w których zagadnienia związane z akustyką nie były jeszcze analizowane, konieczne jest dopasowanie analizy przez dekompozycję harmonik sferycznych do częstotliwości sygnału. Praca Khaykin i Rafaely [38] jest przykładem skutecznej realizacji wygładzonej analizy częstotliwościowej sygnału do wykorzystania przez wspomniane algorytmy opierające się na relacyjnej macierzy sygnałów przetworników pochodzących z macierzy mikrofonowej.

5.3. Metody binauralizacji

W niniejszym punkcie przedstawiono prace opisujące metody binauralizacji, czyli połączenia przestrzennej odpowiedzi impulsowej pomieszczenia SRIR zarejestrowanej w określonym punkcie pomiarowym z funkcją przenoszenia głowy HRTF (często jako HRIR). Należy podkreślić, że rozwijane rozwiązania nie miały na celu realizacji jakichkolwiek pomiarów akustycznych, lecz tworzenie (rendering) przestrzennego obrazu dźwiękowego dla wielokanałowych głośnikowych systemów odsłuchowych lub odsłuchu z użyciem słuchawek. Niejednokrotnie więc użycie odpowiedzi impulsowej pomieszczenia i funkcji HRTF miało charakter uproszczony, nie do końca odwzorujący fizyczne właściwości danych, lecz posługujący się pewnymi zasadami związanymi z percepcją fali bezpośredniej, wczesnych odbić i późnego pogłosu.

Praca Pörschmanna i in. [75] jest przykładem takiego projektu, którego celem jest jedynie binauralizacja wszechkierunkowej (nie przestrzennej) odpowiedzi impulsowej pomieszczenia. Autorzy wykorzystali funkcje HRTF dla utworzenia obrazu przestrzennego, a synteza kierunkowości fali bezpośredniej i wczesnych odbić polegała na przyjęciu kątów nadejścia fali wyznaczonych na podstawie arbitralnie wybranego modelu sali prostopadłościennej (typu „shoebox”). Kąty określone na podstawie modelu sali przyjmowano dla wyboru właściwej funkcji HRTF. Późna pogłosowość sali została zsyntezowana na podstawie splotu późnych odbić pogłosowych z szumem dwuusznym. Metoda ta pozwalała na uzyskanie obrazu przestrzennego sali w kontekście określonego słuchacza (HRTF) i odpowiedniego punktu sali (określona odpowiedź impulsowa sali) ze znaczną oszczędnością obliczeniową, jednak nie stanowi ona właściwego wyznaczenia funkcji BRIR. Niemniej, metoda Pörschmanna i in. [75] była wystarczająca percepcyjnie z uwagi na ograniczenia naszej zdolności słuchu do szczegółowej analizy pogłosu sali.

W pracy Menzera i in. [55] występuje podobne uproszczenie w konstrukcji BRIR z uwagi na ograniczenie HRTF. W tej propozycji jedynie dźwięk bezpośredni jest wprost filtrowany przez HRTF właściwy dla danego kierunku (por. str 34). W zakresie czasowym późnego dźwięku dyfuzyjnego (odbicia wielokrotne) korzysta się z kombinacji liniowej kanałów B-formatu osobno dla każdej częstotliwości tak by uzyskać widmowe i międzyuszne właściwości (koherencja) takie, jakie występują w rzeczywistej BRIR, w tym przypadku BRIR zmierzzonego z użyciem manekina KEMAR. Ta praca jest ścisłym koncepcyjnie nawiązaniem do pracy Merimaa i Pulkki [56] omówionej w punkcie 5.2.

W odróżnieniu od w znacznym stopniu sztucznej syntezy opisanej w pracy [75], realne zastosowanie mikrofonu ambisonicznego i binauralizacji funkcją HRTF, do której jest zbliżona realizacja przedstawiona w obecnej pracy, opisane jest w artykułach Zaunschrim'a i in. [108, 109]. W pracy [108] motywację stanowi działanie dynamiczne, przy którym zakłada się ruchy głową słuchacza. Zatem chodzi o dynamiczne tworzenie BRIR zgodnie z kolejnymi kątami ustawienia głowy. Przestrzenna odpowiedź impulsowa pomieszczenia rejestrowana była mikrofonem ambisonicznym Soundfield ST450 pierwszego rzędu. Wyznaczenie binauralnej odpowiedzi impulsowej BRIR polegało na splotcie z właściwym HRTF wyznaczanym jako kombinacja linowa najbliższych HRTF ze zbioru danych. Przeprowadzono dwa eksperymenty oceny słuchowej uzyskanego obrazu przestrzennego. W pierwszym dokonano oceny skutków interpolacji funkcji HRTF, w drugim syntezę BRIR z przestrzennej odpowiedzi impulsowej i HRTF. Odniesieniem były obrazy dźwiękowe uzyskane na podstawie pomiarów BRIR przeprowadzonych z użyciem sztucznej głowy Neumann KU-100. W pracy wykazano skuteczność

kombinacji ambisonicznej odpowiedzi impulsowej z HRTF sztucznej głowy dla utworzenia przestrzeni dźwiękowej w porównaniu do danych z pomiarów BRIR sztuczną głową. Generalnie w pracy [108] porównano obrazy wynikające z pomiarów głową KU100 i użyciem statycznie położonego mikrofonu ambisonicznego pierwszego rzędu z rozszerzeniem rzędu do 3, 5 i 7 przy użyciu metody dekompozycji przestrzennej sygnału ambisonicznego ASDM (ang. *Ambisonic Spatial Decomposition Method*). Metoda ta polega na dodaniu harmoniki wyższego rzędu z odpowiednim współczynnikiem ważącym w kierunku, z którego dochodzi fala akustyczna rozumiana jako fala lokalnie płaska. Autorzy stwierdzają, że przy rozszerzeniu rzędu ambisonii do $n = 5$, obrazy dźwiękowe przy rejestracji przestrzennej odpowiedzi impulsowej są lepsze od wynikających z nagrań KU100.

Istotą pracy [109] jest spostrzeżenie, że sieć punktów HRTF jest zbyt gęsta w stosunku do rozdzielczości harmonik sferycznych przy ambisonii 1 lub 3 rzędu, czy przy zastosowaniu rozszerzenia rzędu do $n = 5$. Autorzy postulują redukcję rozdzielczości HRIR/HRTF z zastosowaniem wyrównania czasowego TA (ang. *time alignment*), które polega na usunięciu różnic czasowych lewego i prawego ucha w zakresie wysokich częstotliwości powyżej przyjętej częstotliwości granicznej. Jest to więc zmniejszenie rzędu harmonik sferycznych wymaganych do reprezentacji HRTF w zakresie wysokich częstotliwości. Autorzy badają wpływ operacji na dokładność odwzorowania ITD, ILD i korelacji międzyusznej IACC w zakresie wysokich częstotliwości. Metoda stanowi wskazówkę, jak należy upraszczać HRTF do uzyskania porównywalnego rzędu rozdzielczości przestrzennej z rozdzielczością, jaką jest przedstawiana przestrzenna odpowiedź impulsowa pomieszczenia.

Nadmiar danych koniecznych dla opisu przestrzennego dźwięku w sali i przestrzennych właściwości słuchu funkcjami HRTF zawsze stanowił problem przy realizacjach wirtualnego dźwięku przestrzennego. Dlatego już we wcześniejszych pracach takich jak [49] podejmowano wysiłki ograniczania liczby danych przez faktoryzację funkcji HRTF dla odrzucenia informacji niezależnie od kierunku dojścia fali, oraz wyodrębnienie z odpowiedzi impulsowej fali bezpośredniej i wczesnych odbić. Syntezę przeprowadzano na podstawie części pogłosowej uzyskiwanej metodami stosującymi analizę pasmową i dekorelację sygnałów międzyusznych.

Podsumowując, stosowane metody dążą do rozdzielenia pomiarów przestrzennej odpowiedzi impulsowej pomieszczenia SRIR i funkcji HRTF dla uzyskania korzystnego ograniczenia zbioru danych. Niemożliwe jest bowiem stosowanie maczyzy danych dla wszystkich kombinacji punktów pomiarowych, kątów położenia głowy i wzajemnego położenia źródła i obserwatora. Co więcej, dane musiałyby być dopasowane do indywidualnych funkcji HRTF.

Bardzo istotnym dla przetwarzania HRTF w kontekście stanowiącym podsumowanie szeregu wcześniejszych prac jest artykuł Engela i in. [23] z 2022 roku. Autorzy analizują sześć technik redukcji wymiarów HRTF [5, 6, 9, 33, 54, 111]. Do omawianych metod ograniczających rząd HRTF należą: (1) proste obcięcie harmonik sferycznych powyżej założonego rzędu N , (2) dodatkowa filtracja [6] wyrównująca składowe w zakresie wysokich częstotliwości w celu zrównoważenia z zakresem niskich częstotliwości dominujących po ograniczeniu rzędu harmonik sferycznych do N , (3) stosowanie przestrzennego okna Hanna [33] nałożonego na składowe harmonik sferycznych dla ograniczenia aliasingu przestrzennego po ograniczeniu do rzędu N , (4) wyrównanie czasowe [5] związane z usunięciem międzyuszej różnicy czasowej ITD. [5], (5) stosowanie MagLS [87, 111], czyli ograniczania różnic fazowych w HRIR powyżej pewnej częstotliwości granicznej, często przyjmowanej jako równej częstotliwości aliasingu, (6) spatial resampling [54, 109] polegająca na zdekodowaniu sygnału do zbioru wirtualnych głośników umieszczonych w przestrzeni, a następnie przypisaniu ich kierunkom odpowiednich funkcji HRTF. W artykule jako siódma praca jest przedstawiona metoda autorów określona skrótem BiMagLS, będąca połączeniem metod wyrównania czasowego dla częstotliwości niższych od częstotliwości aliasingu, a dla wyższych MagLS.

W pracy Engel i in. [23] przeprowadzono analizę błędów porównując oryginalne HRTF z HRTF utworzonego z użyciem harmonik sferycznych z ograniczeniem rozdzielczości odpowiednimi wymienionymi metodami i resyntezie. Autorzy wykazali, że błędy amplitudowy i fazowy narastają znacząco po przekroczeniu częstotliwości aliasingu związanej z rzędem N harmonik sferycznych. W pracy [23] szczególną analizę przeprowadzono dla ograniczenia harmonik sferycznych reprezentujących HRTF do rzędu $N = 3$, co jest o tyle istotne, że w obecnej pracy było przeprowadzone podobne uproszczenie HRTF w odniesieniu do danych uzyskanych mikrofonem trzeciego rzędu Zylia ZM1-3E. Krótkie porównanie uzyskanych wyników z wnioskami wyprowadzonymi w pracy Engela i in. [23] przeprowadzono w rozdziale 9, stanowiącym dyskusję wyników niniejszej pracy.

6. Cel, teza i zakres pracy

Szczegółowym celem pracy jest wyznaczenie parametrów akustycznych pomieszczenia z użyciem rejestracji odpowiedzi impulsowej pomieszczenia mikrofonami ambisonicznymi pierwszego i trzeciego rzędu i określenie różnic, jakie mogą występować względem pomiarów przeprowadzonych mikrofonem pomiarowym o charakterystyce dookólnej i manekinem akustycznym. Badania obejmują wyznaczenie parametrów dookólnych: T30, EDT, C50, C80 i Ts i wyznaczenie międzyusznej korelacji skrośnej IACC. W pierwszym przypadku wykorzystywany jest kanał W mikrofonów ambisonicznych, w drugim przestrzenna odpowiedź impulsowa sali SRIR dająca w połączeniu z HRTF stosowanych manekinów akustycznych binauralną odpowiedź impulsową BRIR. W analizie pominięto obliczanie parametrów siły dźwięku G i wsparcia estrady Stage Support.

W ramach pracy wyznaczana jest przestrzenna odpowiedź impulsowa pomieszczenia SRIR, a następnie z użyciem funkcji HRTF binauralna odpowiedź impulsowa pomieszczenia BRIR. Funkcje SRIR i BRIR są odpowiednio wykorzystane do wyznaczenia parametru międzyusznej korelacji skrośnej IACC. Jak wspomniano wcześniej, wyznaczanie parametrów wszechkierunkowych kanałem W mikrofonu ambisonicznego nie jest nowością, jednak w niniejszej pracy istotne jest porównanie tych wyników ze standardem pomiaru mikrofonem pomiarowym w różnych warunkach akustycznych. W odniesieniu do współczynnika IACC jest to pokazanie możliwości zastosowania odpowiedzi impulsowej BRIR do celów pomiarowych. Obecnie wyznaczanie BRIR wykorzystywane jest przede wszystkim w systemach rozszerzonej rzeczywistości i dźwięku przestrzennego, stąd obserwowana niemalejąca popularność badań nad modyfikacją i próbkowaniem przestrzennym zarówno BRIR jak i HRIR [21, 49, 52, 55, 94, 98] wskazuje na rozpatrzenie możliwości zastosowania w pomiarach akustycznych.

Cel pracy wiąże się z udowodnieniem tezy pracy, którą można sformułować następująco: **możliwe jest uzyskanie zestawu parametrów akustycznych pomieszczenia, ze szczególnym uwzględnieniem binauralnych parametrów przestrzennych, za pomocą jednego przetwornika pomiarowego w postaci sferycznej matrycy mikrofonowej o kontrolowanej kierunkowości przy uwzględnieniu niezależnie mierzonych lub modelowanych funkcji HRTF.**

Biorąc pod uwagę rozwój konstrukcji matryc mikrofonowych niniejsza praca wpisuje się w kierunek badań akustycznych związany z rozwijaniem metod prowadzących do jak najlepszej reprezentacji w pomiarach subiektywnie odbieranej przestrzenności dźwięku i jakości pola akustycznego w pomieszczeniach takich jak sale koncertowych i audytoria.

Temat powyższy był podejmowany w literaturze, niemniej jest reprezentowany bardzo skromnie, co wykazała analiza w rozdziale 5. Większość bowiem prac jest związana z dźwiękiem przestrzennym w kontekście rejestracji nagrań ambisonicznych i filtrowanie przez funkcje przeniesienia głowy HRTF dla odwzorowania w układzie wielokanałowym lub słuchawkowym dźwięku 3D w określonych salach lub symulacji i syntezy warunków przestrzennych 3D. Dodatkowym uzasadnieniem podjęcia zagadnienia związanego z pomiarami sal, a zwłaszcza parametru IACC, jest nie tylko dążenie do wykorzystania równolegle istniejących najnowszych osiągnięć technologicznych, ale także korzyść, jak się wydaje, wynikająca z uproszczenia procedury pomiarowej przez ograniczenie liczby i rodzaju pomiarów, jakie należy przeprowadzać. Pomiary (podwójne) mikrofonem wszechkierunkowym i niezależnie z użyciem manekina akustycznego można zastąpić pomiarem wyłącznie mikrofonem ambisonicznym. Jest to ważne ze względu na znaczną liczbę punktów pomiarowych na widowni i estradzie, typową dla pomiarów dużych sal koncertowych i audytoryjnych.

W ramach pracy podjęto również próbę określenia dokładności pomiarów z użyciem mikrofonów ambisonicznych jako konstrukcji wieloprzetwornikowych o przetwornikach nieodpowiadających standardom dokładności mikrofonów pomiarowych.

7. Analiza przestrzennej odpowiedzi impulsowej pomieszczenia

7.1. Metodyka badań

Badania ukierunkowano na dwa mikrofony ambisoniczne: pierwszego rzędu Sennheiser Ambeo i trzeciego rzędu Zylia ZM1-3E, z których użyciem przeprowadzono rejestrację przestrzennych odpowiedzi impulsowych pomieszczenia SRIR. W celu uzyskania binauralnych odpowiedzi impulsowych pomieszczenia BRIR, zastosowano funkcje HRTF manekina pomiarowego B&K4100-D oraz nagraniowej sztucznej głowy Neumann KU-100. Dodatkowy wybór sztucznej głowy Neumann KU-100 wynikał z powszechnego wykorzystania tego manekina do rejestracji dźwięku w pracach związanych z ambisonią [23, 108, 109] .

Wszechkierunkowe parametry akustyczne sal, czas pogłosu T30, czas wczesnego zaniku EDT, czytelność dźwięku C50, przejrzystość dźwięku C80 i czas środkowy Ts, zostały wyznaczone na podstawie sygnału wszechkierunkowego W mikrofonów ambisonicznych. Parametr przestrzenny międzyusznej korelacji skróśnej IACC został wyznaczony na podstawie uzyskanych binauralnych odpowiedzi impulsowych pomieszczenia BRIR.

Wartości parametrów opisujących akustykę pomieszczeń uzyskanych z rejestracji mikrofonami ambisonicznymi i binauralizacji zostały zweryfikowane przez pomiary przeprowadzone z użyciem wszechkierunkowego mikrofonu pomiarowego B&K typ 4189 oraz pomiary z użyciem manekina akustycznego B&K typ 4100-D oraz sztucznej głowy Neumann KU-100, których HRTF był używany do tworzenia funkcji BRIR.

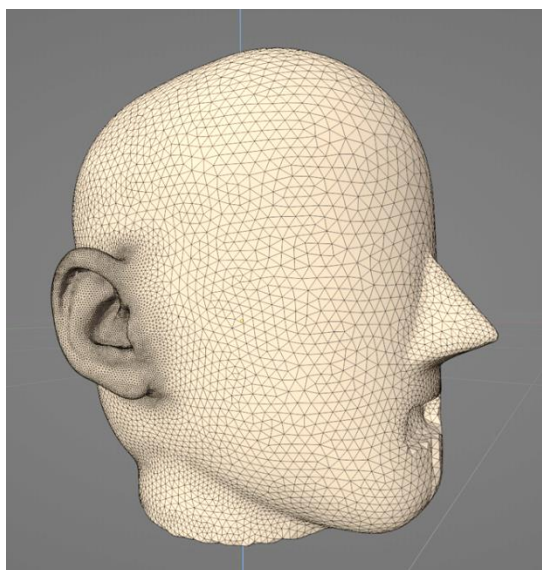
Wyniki pomiarów parametrów wszechkierunkowych T30, EDT, C50, C80 i Ts zostały przedstawione w punkcie 8.2, a wyniki parametrów binauralnych w punkcie 8.3.

7.2. Wyznaczenie binauralnych odpowiedzi pomieszczeń BRIR

Binauralne odpowiedzi impulsowe pomieszczeń BRIR otrzymane na podstawie przestrzennej odpowiedzi impulsowej SRIR zarejestrowanej przy użyciu każdego z mikrofonów ambisonicznych (Sennheiser Ambeo i Zylia ZM1-3E) zostały wyznaczone przy użyciu czterech opisanych wcześniej algorytmów binauralizacji sygnału ambisonicznego: Least Squares (LS), Magnitude Least Squares (MagLS), Time Alignment (TA) i Spatial Resampling (SPR) zgodnie z procedurami określonymi w pracy Schonhuber [87] z wykorzystaniem opracowanego przez autorów toolboxu SPARTA [50].

7.3. Wyznaczenie funkcji HRTF manekina B&K4100-D

Zbiór transmitancji HRTF manekina pomiarowego B&K 4100-D uzyskano na podstawie modelowania metodą elementów brzegowych z wykorzystaniem oprogramowania Mesh2HRTF [12, 95, 110] na podstawie cyfrowego modelu 3D. Model 3D został uzyskany w procesie skanowania z użyciem skanera SMARTTECH3D med [93] i jego firmowego oprogramowania¹. Dokładność skanowania obejmowała siatkę punktów z rozdzielczością 0,23 mm i dokładnością pomiaru +/- 0,06 mm na podstawie której wyznaczono siatkę trójkątów do obliczeń (ang. *mesh*). W celu optymalizacji modelu manekina pod kątem obliczeń, siatka trójkątów stanowiąca model manekina miała większą rozdzielczość w obrębie małżowiny usznej (bok trójkąta: 0,5 mm) oraz mniejszą rozdzielczość w pozostałej części głowy (bok trójkąta do 10,0 mm), zgodnie z zaleceniami autorów oprogramowania [69]. Rysunek 7.1 przedstawia cyfrowy, trójwymiarowy model manekina B&K4100-D wykorzystany do obliczeń zbioru funkcji HRTF.

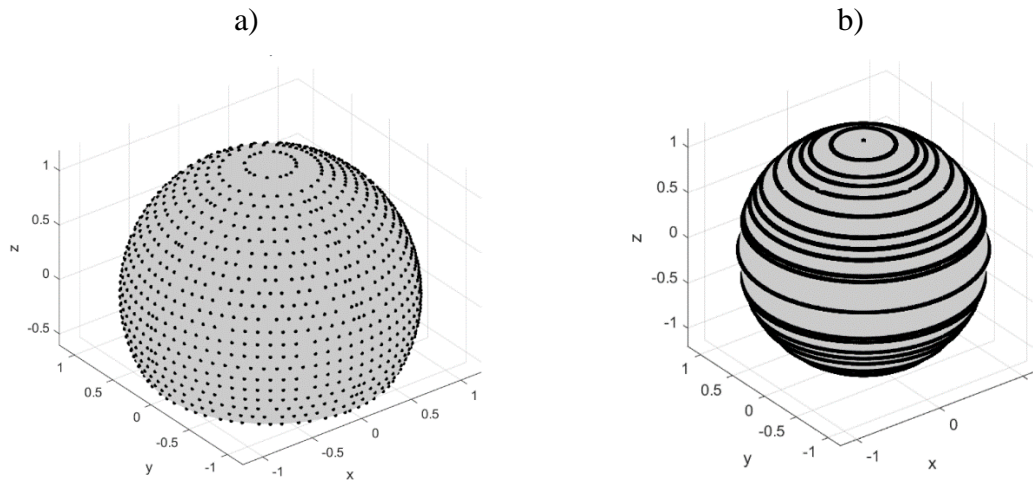


Rys. 7.1. Model manekina B&K4100-D wykorzystany do obliczeń z użyciem Mesh2HRTF

Funkcje HRTF sztucznej głowy Neumann KU-100 zaczerpnięte z bazy SADIE-II [3] występowały w dwóch rozdzielczościach: gęstej siatki 8802 punktów, która obejmowała z rozdzielczością pomiaru 1° płaszczyznę azymutu oraz 24 kąty elewacji, Rzadszej siatki, 1550 punktów z rozdzielczością 5° w płaszczyźnie azymutu oraz 10° w płaszczyźnie elewacji.

¹ Skan 3D manekina uzyskano dzięki uprzejmości firmy SMARTTECH sp. z o.o., <https://www.smarttech3d.com/>

W odniesieniu do HRTF obu manekinów przyjęto wymienione dwie rozdzielczości siatki punktów 1550 punktów oraz 8802 punkty. Obie siatki wykorzystano do utworzenia odpowiedzi impulsowej BRIR tak, że rozdzielczość siatki HRTF stanowi jeden z badanych czynników.



Rys. 7.2. Siatka kątów funkcji HRTF: a) 1550 punktów; b) 8802 punkty

7.4. Badane pomieszczenia

Możliwości zastosowania techniki ambisonicznej, HRTF i BRIR do wyznaczania parametrów akustycznych pomieszczeń zweryfikowano na podstawie pomiarów przeprowadzonych w pięciu pomieszczeniach, zróżnicowanych pod względem warunków pogłosowych. Koncepcja doboru pomieszczeń polegała na wyborze: od pomieszczenia, w którym w całości występuje akustyczne pole fali swobodnej, do pomieszczenia o polu dyfuzyjnym, poprzez kilka wewnątrz o wzrastającej pogłosowości. Skrajnymi pod względem właściwości akustycznych były komora bezechowa w Gmachu Elektroniki Politechniki Warszawskiej oraz komora pogłosowa Zakładu Fizyki Ciepłej, Akustyki i Środowiska Instytutu Techniki Budowlanej w Warszawie. Pomieszczeniami pośrednimi były Studio Dźwiękowe Zakładu Elektroakustyki Politechniki Warszawskiej, korytarz w Gmachu Elektroniki Politechniki Warszawskiej oraz sala koncertowa Filharmonii Pomorskiej im. Ignacego Paderewskiego w Bydgoszczy. Szczegółowa informacja odnosząca się do wymienionych pomieszczeń jest następująca:

1. Komora bezechowa Zakładu Elektroakustyki PW: kubatura 250 m^3 , częstotliwość odcięcia 110 Hz, czas pogłosu 70 ms (średnia 500 i 1000 Hz), poziom szumów tła poniżej 10 dB(A).

2. Korytarz Gmachu Elektroniki w Zakładzie Elektroakustyki PW: kubatura 120 m³, wymiary: 17 x 2 x 3 m (dł. x szer. x wys.) czas pogłosu 1,3 s, ściany odbijające, sufit o typowym wyciszeniu biurowym.
3. Studio dźwiękowe Zakładu Elektroakustyki PW: kubatura 300 m³, czas pogłosu 0,8 s, adaptacja akustyczna i wyposażenie w szereg ustrojów akustycznych o charakterze zarówno rezonansowym, jak i rozpraszającym.
4. Sala koncertowa Filharmonii Pomorskiej w Bydgoszczy im. Ignacego Paderewskiego: kubatura 6300 m³ i 886 miejsc siedzących, czas pogłosu: 2,1 s.
5. Komora pogłosowa Instytutu Techniki Budowlanej: kubatura 200 m³, czas pogłosu 6,1 s

7.5. Aparatura pomiarowa

Rejestrację odpowiedzi impulsowej przeprowadzono zgodnie z metodyką pomiarową opisaną w normie PN-EN ISO 3382-1 [73]. Do pobudzenia pomieszczeń, w każdym z przypadków, wybrano sygnał tonu przestrajanego logarytmicznie (ang. *log sine sweep*).

W celu uzyskania zestawu do porównania wyników różnych typów odpowiedzi impulsowych, zarejestrowano je z wykorzystaniem 6 typów przetworników pomiarowych:

1. Mikrofony wszechkierunkowe B&K 4189 (pomiaru dookólne), Microtech Gefell typ MK225 (pomiaru dookólne w Filharmonii w Bydgoszczy).
2. Manekin akustyczny B&K 4101-D (pomiaru BRIR).
3. Sztuczna głowa Neumann KU100 (pomiaru BRIR).
4. Mikrofon ambisoniczny 1-go rzędu Sennheiser Ambeo (pomiaru SRIR 1-rzędowy).
5. Sferyczna matryca mikrofonowa 3-go rzędu Zylia ZM1-3E (pomiaru SRIR 3-go rzędowy).

Pomiary przeprowadzono z użyciem wszechkierunkowego źródła dźwięku B&K 4292, wzmacniacza B&K 2734, interfejsów RME-UC i Soundcraft Ui-24. Do rejestracji i przetwarzania odpowiedzi impulsowych wykorzystano oprogramowanie Dirac v6.0 i Easera v1.2.18. Metody upraszczania HRTF zastosowano na podstawie opracowanego przez ich autorów oprogramowania w postaci pluginów SPARTA [50].

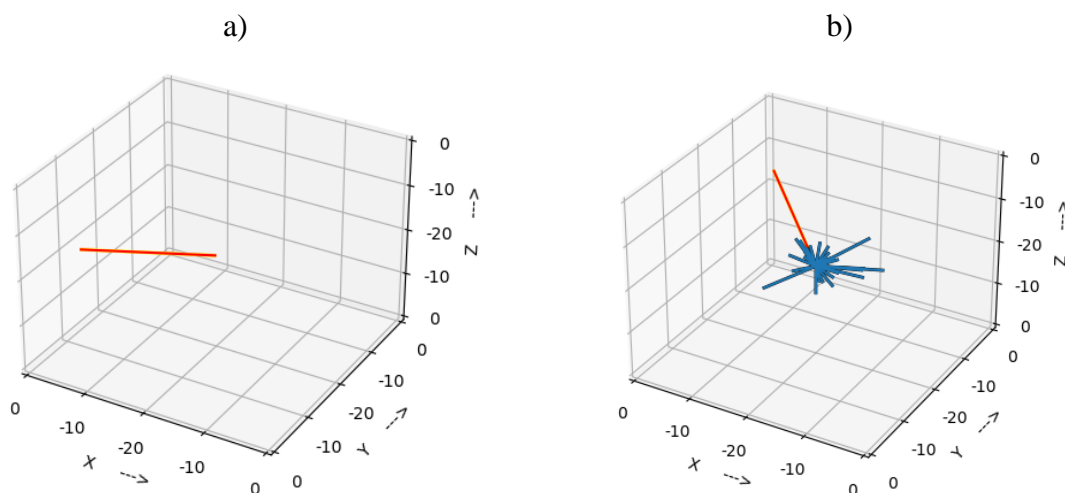
Szczegóły obejmujące rozmieszczenie punktów pomiarowych w pomieszczeniach podane są w załączniku 2.

8. Wyniki badań

W niniejszym rozdziale przedstawiono wyniki pomiarów, które pokazują skuteczność rejestracji odpowiedzi impulsowych pomieszczeń z wykorzystaniem technik ambisonicznych przy określaniu parametrów akustycznych pomieszczeń na podstawie pomiarów przeprowadzonych w pięciu opisanych wcześniej pomieszczeniach o różnych właściwościach akustycznych. W punkcie 8.1 przedstawiono różnice między echogramami pomieszczeń uzyskanymi z użyciem mikrofonów Sennheiser Ambeo, Zylia ZM1-3E oraz mikrofonem pomiarowym. W punkcie 8.2 przedstawiono wyznaczone na podstawie dookólnej odpowiedzi impulsowej (kanał W) czas pogłosu T30, czas wczesnego zaniku EDT, czytelność dźwięku C50, przejrzystość dźwięku C80 i czas środkowy echogramu Ts. Analogicznie, w punkcie 8.3 przedstawiono parametr IACC wyznaczony na podstawie binauralnych odpowiedzi impulsowych uwzględniających funkcje HRTF upraszczane z użyciem algorytmów opisanych w punkcie 4.2. Wyniki pomiaru parametrów akustycznych pomieszczeń zostały odniesione do pomiarów z użyciem manekina B&K 4100-D oraz sztucznej głowy Neumann KU-100 oraz dookólnego mikrofonu pomiarowego.

8.1. Obserwacja echogramów

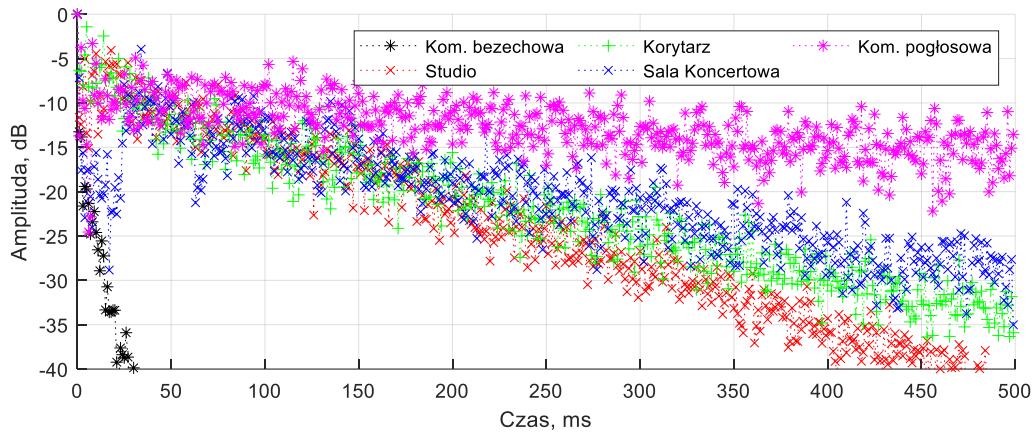
Najbardziej ogólną i jednocześnie najprostszą metodą analizy odpowiedzi impulsowej pomieszczenia jest obserwacja echogramów przedstawiających amplitudę w funkcji czasu opóźnienia poszczególnych odbić dźwięku docierających od źródła do punktu rejestracji. Zastosowanie techniki ambisonicznej pozwala na rozszerzenie tej analizy dodatkowo o wymiar przestrzeni i obserwację z jakiego kierunku docierają poszczególne odbicia. Przykład kierunkowego echogramu otrzymanego w komorze bezechowej i studio dźwiękowym pokazuje rysunek 8.1. W naturalny sposób w echogramie w komorze bezechowej widoczny jest wyłącznie kierunek dojścia dźwięku bezpośredniego. Analiza tego typu przestrzennych echogramów jest skomplikowana i nie jest kontynuowana w dalszej części pracy. Przedstawione echogramy są ograniczone do rejestracji z pomocą kanału W mikrofonów ambisonicznych oraz wszechkierunkowego mikrofonu pomiarowego.



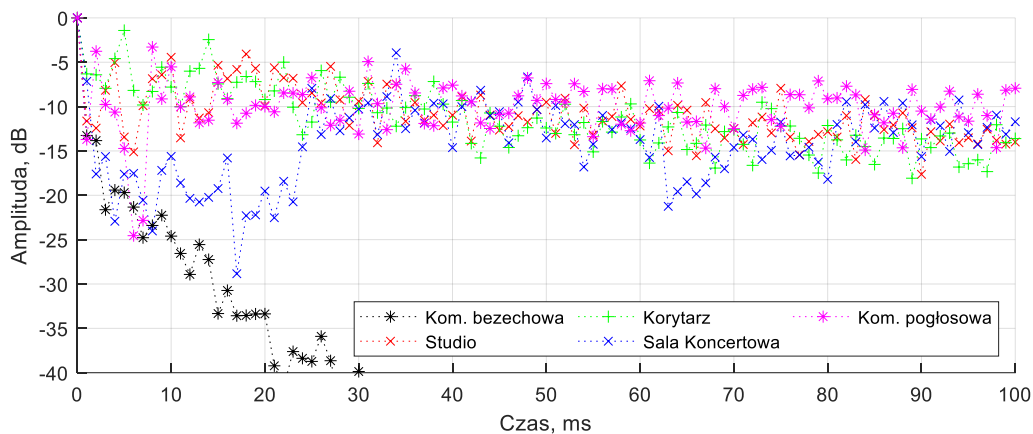
Rys. 8.1. Echogramy kierunkowe otrzymane z użyciem algorytmu *DOA-Estimation* obliczania wektora pseudonatężenia dźwięku opisanego w pracy [55]; a) komora bezechowa; b) studio dźwiękowe

Jako wstępny przegląd właściwości pomieszczeń, w których przeprowadzono pomiary, na rys. 8.2 i 8.3 przedstawione są echogramy wszystkich badanych pomieszczeń. Rysunek 8.2 przedstawia echogramy w zakresie 0 - 500 ms, natomiast dla uwypuklenia wczesnych odbić rys. 8.3 w zakresie do 100 ms. Rozdzielczość echogramów wynosi 1 ms, zrealizowaną w taki sposób, że spośród próbek sygnału zarejestrowanych w każdym przedziale 1 ms, pokazano największą wartość.

Na 8.2 i 8.3 wyraźnie wyróżnia się pole swobodne obserwowane w komorze bezechowej, jako że zanik echogramu ma miejsce już po 30 ms. Pomieszczenie o przeciwstawnej charakterystyce – komora pogłosowa – charakteryzuje się zbliżonym poziomem obić przez znaczną część odpowiedzi impulsowej. Dominację pola pogłosowego można zaobserwować na rys. 8.2. Dokładniej widoczne szczegóły echogramu na rys. 8.3 pokazują, że w echogramie sali koncertowej, zaobserwować można wyraźny wpływ dźwięku bezpośredniego oraz wczesnych odbić docierających po około 25 ms. W przypadku studio dźwiękowego ten odstęp wynosi około 5 ms. W echogramie korytarza można zauważyć powtarzające się, dominujące odbicia świadczące o występującym zjawisku echa trzepoczącego (ang. *flutter echo*).



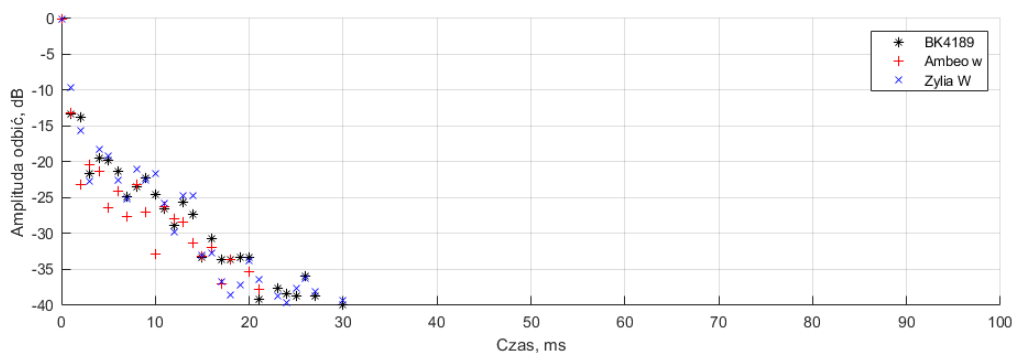
Rys. 8.2. Echogramy późnych odbić dźwięku zarejestrowane w badanych pomieszczeniach przy użyciu referencyjnego mikrofonu pomiarowego. Przedział 0 - 500 ms



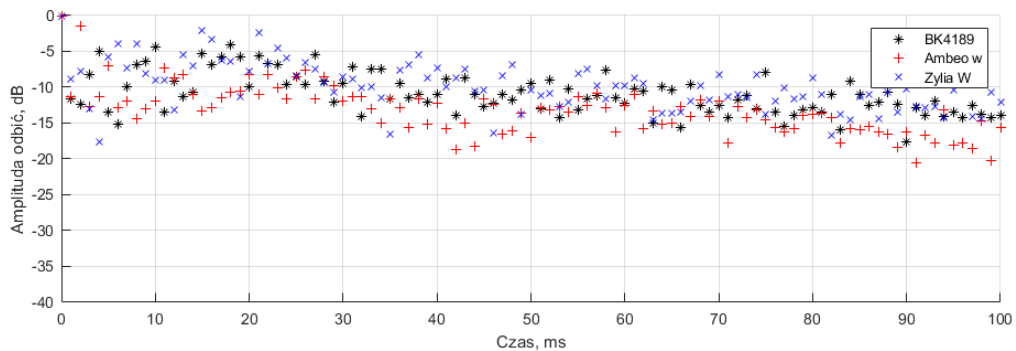
Rys. 8.3 Echogramy wczesnych odbić dźwięku zarejestrowane w badanych pomieszczeniach przy użyciu referencyjnego mikrofonu pomiarowego. Przedział 0 - 100 ms

Rysunki 8.4 - 8.7 przedstawiają echogramy zarejestrowane z wykorzystaniem wszechkierunkowego kanału W obu mikrofonów ambisonicznych i pomiarowego mikrofonu wszechkierunkowego w poszczególnych pomieszczeniach, w których przeprowadzane były badania. W komorze bezechowej (rys. 8.4) przy zaniku dźwięku po 30 ms różnice pomiędzy pomiarami z użyciem różnych mikrofonów wynoszą około ± 5 dB, bez systematycznych różnic pomiędzy mikrofonami. W studio dźwiękowym (rys. 8.5) w pojedynczych przypadkach amplitudy odbić zarejestrowane mikrofonami ambisonicznymi różnią się od zarejestrowanych mikrofonem pomiarowym o 10 dB. Charakterystyczne jest, że mikrofon Ambeo przedstawia wartości poniżej otrzymanych z użyciem mikrofonu pomiarowego w 70% przypadków, a mikrofon Zylia powyżej (60-70% przypadków). W przypadku korytarza (rys. 8.6), oba mikrofony ambisoniczne przedstawiają wyniki pomiarów w 80% przypadków mniejsze niż zmierzone

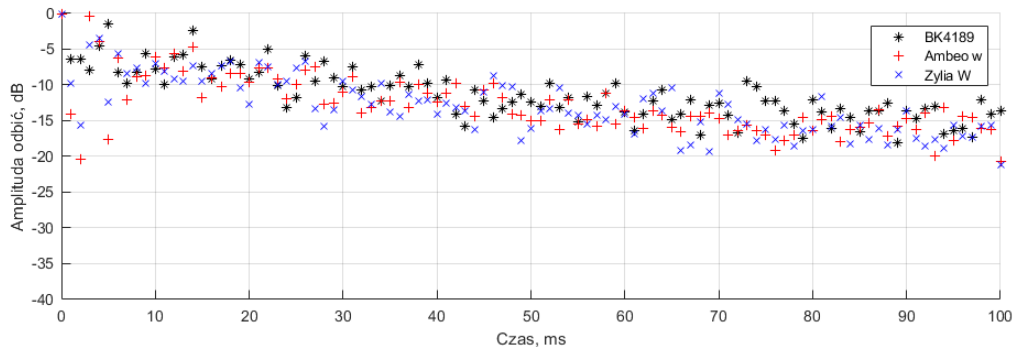
mikrofonem pomiarowym, w granicach 5 dB. W początkowej części wszystkich echogramów zauważalne są szczyty przebiegu, które wskazują na trzepoczące echo. W przypadku sali koncertowej (rys. 8.7) wspomniane wczesne odbicia pojawiające się po pierwszych 25 ms są widoczne zarówno w rejestracji mikrofonem Ambeo, jak i pomiarowym mikrofonem referencyjnym. Amplitudy zmierzone mikrofonem Ambeo są w 70% przypadków mniejsze, jednakże w granicach pojedynczych dB. W zapisie echogramu komory pogłosowej (rys. 8.7) widoczne są wczesne odbicia o opóźnieniu poniżej 10 ms i wyrównany późny pogłos, w odniesieniu do którego jednak rejestracja mikrofonem Zylia dała odczyt amplitud prawie zawsze mniejszy o 5-10 dB niż rejestracja mikrofonem pomiarowym.



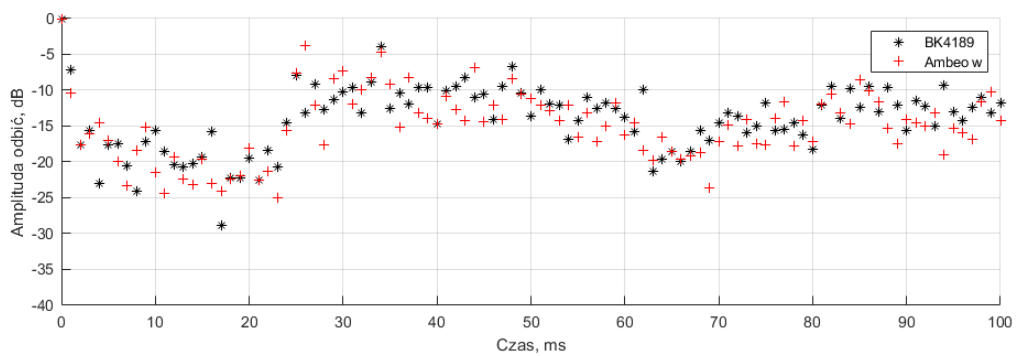
Rys. 8.4 Echogramy odpowiedzi impulsowych komory bezechowej. Rejestracja kanałem W mikrofonów ambisonicznych Sennheiser Ambeo 1-go rzędu i Zylia ZM1-3E 3-go rzędu oraz mikrofonem pomiarowym B&K 4189



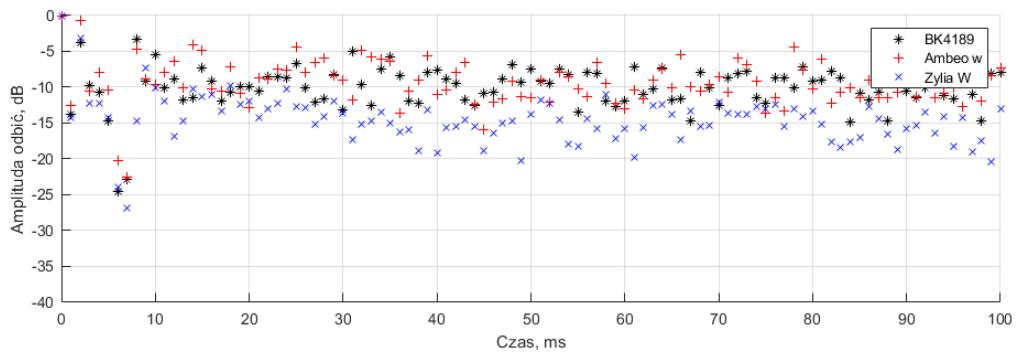
Rys. 8.5 Echogramy odpowiedzi impulsowych studia dźwiękowego. Rejestracja kanałem W mikrofonów ambisonicznych Sennheiser Ambeo 1 go rzędu i Zylia ZM1-3E 3-go rzędu oraz mikrofonem pomiarowym B&K 4189



Rys. 8.6 Echogramy odpowiedzi impulsowych korytarza. Rejestracja kanałem W mikrofonów ambisonicznych Sennheiser Ambeo 1 go rzędu i Zylia ZM1-3E 3-go rzędu oraz mikrofonem pomiarowym B&K 4189



Rys. 8.7 Echogramy odpowiedzi impulsowych sali koncertowej. Rejestracja kanałem W mikrofonów ambisonicznych Sennheiser Ambeo 1 go rzędu oraz mikrofonem pomiarowym MK 225



Rys. 8.8 Echogramy odpowiedzi impulsowych komory pogłosowej. Rejestracja kanałem W mikrofonów ambisonicznych Sennheiser Ambeo 1 go rzędu i Zylia ZM1-3E 3-go rzędu oraz mikrofonem pomiarowym B&K 4189

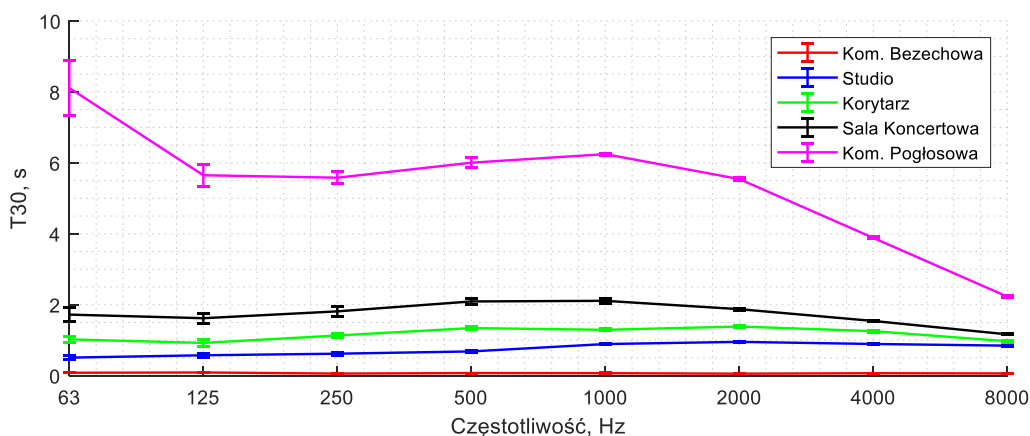
W podsumowaniu można stwierdzić, że użycie kanału W mikrofonu ambisonicznego uwidacznia cechy echogramów rejestrowane w klasyczny sposób mikrofonem pomiarowym.

8.2. Parametry akustyczne wyznaczone z dookólnej odpowiedzi impulsowej

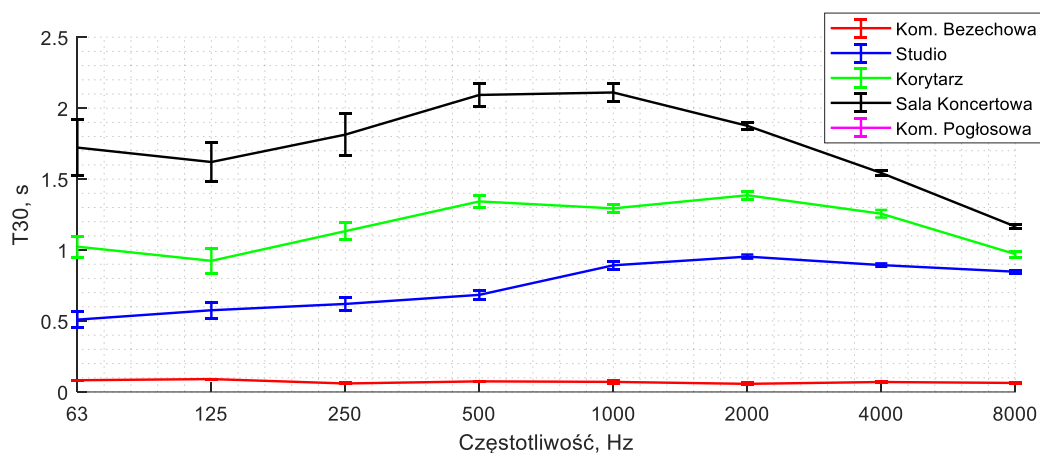
W niniejszym rozdziale przedstawiono wyniki pomiarów czasu pogłosu T_{30} , czasu wczesnego zaniku EDT, czytelności dźwięku C50, przejrzystości dźwięku C80 i czasu środkowego echogramu T_s z użyciem kanału W mikrofonów Ambeo i Zylia, przeprowadzając porównanie z analogicznymi wynikami uzyskiwanymi dookólnym mikrofonem pomiarowym. Wyniki uzyskane dla poszczególnych parametrów akustycznych przedstawione są w punktach 8.2.1 do 8.2.5.

Struktura prezentacji danych jest taka, że w pierwszej kolejności przedstawiane są parametry pomieszczeń zmierzone mikrofonem wszechkierunkowym B&K 4189 lub MK225, a następnie przedstawiane są różnice, jakie występują w określaniu tych wartości przy pomiarach z użyciem mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E. Różnica zdefiniowana jest przez odjęcie wartości uzyskanej mikrofonem ambisonicznym od wartości uzyskanej mikrofonem pomiarowym. W przypadku parametrów czasowych T_{30} i EDT jest ona wyrażona w procentach. Wszystkie wyniki stanowią uśrednienie punktów pomiarowych przedstawionych w załączniku 2.

8.2.1. Czas pogłosu T_{30}

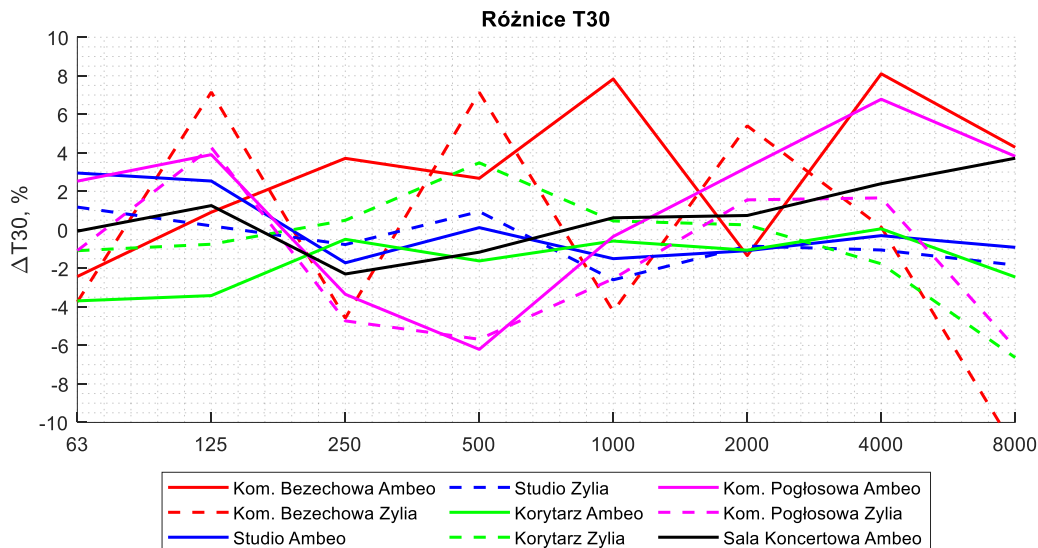


Rys. 8.9. Czasu pogłosu T_{30} zmierzony mikrofonem pomiarowym B&K 4189 w pięciu badanych pomieszczeniach. Średni czas pogłosu (pasma 500 i 1000 Hz): komora bezechowa – 70 ms, studio – 0,8 s, korytarz – 1,3 s, sala koncertowa – 2,1 s, komora pogłosowa – 6,1 s.



Rys. 8.10. Czasu pogłosu T_{30} przy rozszerzonej skali czasu pogłosu. Szczegóły jak na rys. 8.9.

Charakterystyki częstotliwościowe czasu pogłosu T_{30} (rys. 8.9 i 8.10) zarejestrowane przy użyciu pomiarowego mikrofonu wszechkierunkowego wskazują na znaczące i spodziewane różnice pomiędzy pomieszczeniami. Czas pogłosu T_{30} w komorze bezechowej zawiera się w przedziale od 60 ms w paśmie 2 kHz do 90 ms w paśmie 125 Hz. W studio dźwiękowym, adaptowanym akustycznie zmienia się w przedziale od 0,5 s w paśmie 63 Hz do 0,9 s w paśmie 2 kHz. W przypadku korytarza, stanowiącego przykład pomieszczenia niezaadaptowanego akustycznie, czas pogłosu zmienia się od 0,7 s (125 Hz) do 1,4 s (2 kHz). W sali koncertowej o znacznie większej kubaturze czas pogłosu zmienia się od 2,1 s w paśmie 1 kHz do 1,2 s w paśmie 8 kHz. Jak należy przewidywać, najwyższe wartości czasu pogłosu zarejestrowano w komorze pogłosowej, w której zmieniał się w przedziale od 2,2 s w paśmie 8 kHz do nawet 8,1 s w paśmie 63 Hz. Na rysunku 8.10 przedstawiono czas pogłosu T_{30} w większej skali, co pozwala na zauważenie, że charakterystyka częstotliwościowa T_{30} w studio dźwiękowym i sali koncertowej jest typowa dla pomieszczeń tego rodzaju.



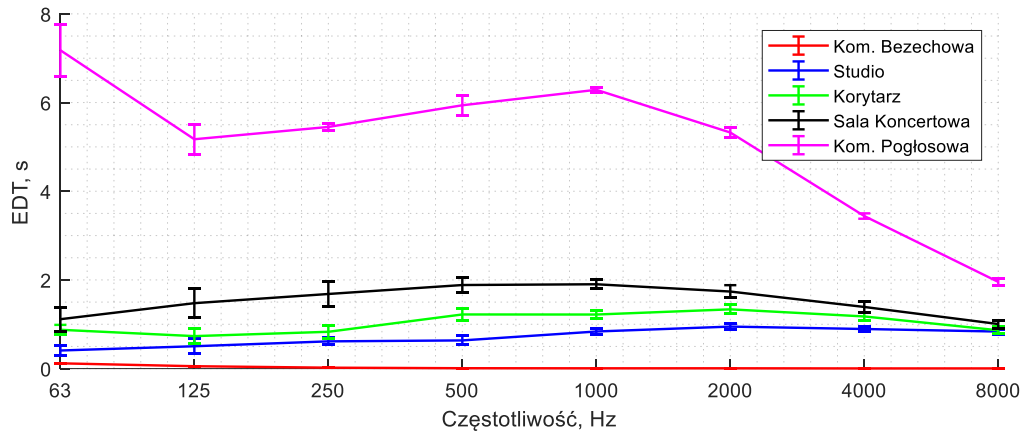
Rys. 8.11. Czas pogłosu T30. Różnice zmierzone mikrofonem pomiarowym B&K 4189 i przy użyciu kanału W mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E

Różnice czasu pogłosu T30 uzyskane z pomiaru z użyciem mikrofonu pomiarowego i użyciem mikrofonów Ambeo i Zylia przedstawiono na rys. 8.11. We wszystkich pomieszczeniach użytkowych (studio, korytarz, sala koncertowa), największe odstępstwa przy pomiarze mikrofonem ambisonicznym w stosunku do mikrofonu pomiarowego nie przekraczają $\pm 4\%$, co jest wartością mniejszą od progowej różnicy zauważania zmiany czasu pogłosu JND (por. tab. 2.1). Jedynie w pomieszczeniach laboratoryjnych, tj. komorze bezechowej i komorze pogłosowej te różnice są większe od JND, ale nie przekraczają 8-9% (ogółem 9 przypadków). Największe odstępstwo sięgające 11,3% występuje w przypadku mikrofonu Zylia, w komorze bezechowej przy 8 kHz i należy tę wartość łączyć z bardzo małym czasem pogłosu rejestrowanej w tym pomieszczeniu. Analiza wariancji, której szczegółowe wyniki podane są w załączniku 3 wykazała, że obserwowane na rys. 8.11 większe różnice są istotne statystycznie na poziomie $p = 0,05$. Tabela Z.3.1 w załączniku 3 podaje szczegółową informację, przy których częstotliwościach zaobserwowane różnice były istotne.

8.2.2. Czas wczesnego zaniku EDT

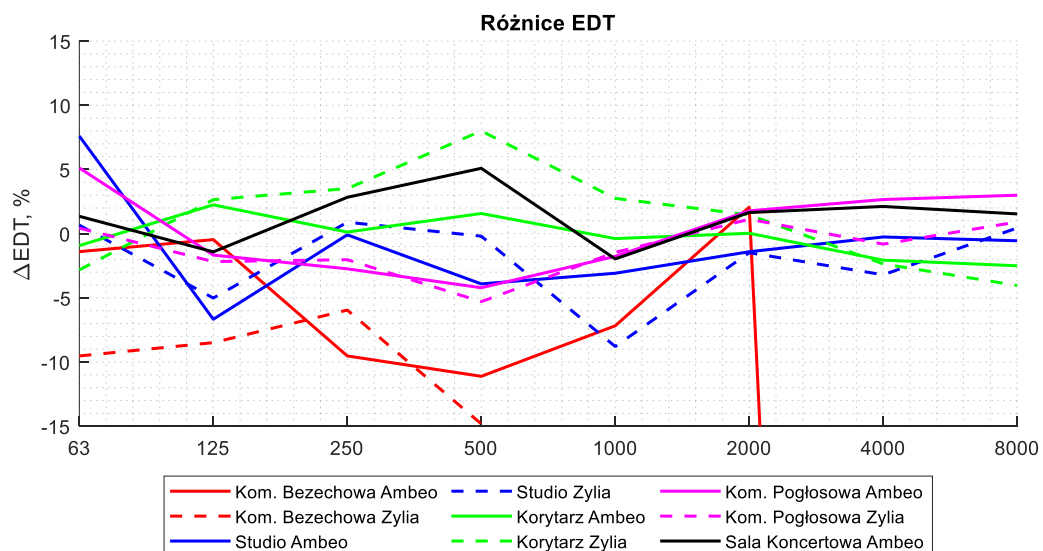
Uśrednione we wszystkich punktach pomiarowych wartości czasu wczesnego zaniku EDT przedstawiono na rysunku 8.12. W przypadku komory bezechowej, zarejestrowane EDT było najniższe i zmieniało się w przedziale od 4 ms w paśmie 4 kHz do 12 ms w paśmie 63 Hz. W studio dźwiękowym wartości EDT wynoszą od 0,4 s w paśmie 63 Hz do 0,9 s dla 2 kHz, w korytarzu mieszczą się w przedziale od 0,7 s dla 125 Hz do 1,3 s dla 2 kHz. W sali

koncertowej, EDT zmienia się od 1,0 s w paśmie 8 kHz do 1,9 s w paśmie 1 kHz, a w komorze pogłosowej pomiędzy 2,0 s w paśmie 8 kHz do 7,2 s w najniższym paśmie częstotliwości 63 Hz.



Rys. 8.12. Czas wczesnego zaniku EDT zmierzony mikrofonem pomiarowym B&K 4189 w pięciu badanych pomieszczeniach.

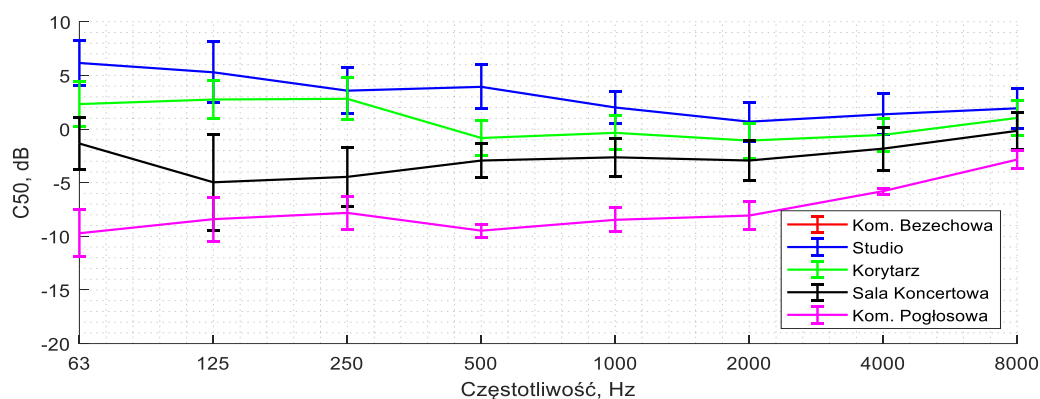
Podobnie jak w przypadku czasu pogłosu T30, określono względne różnice w czasie wczesnego zaniku przy pomiarach z użyciem mikrofonu pomiarowego i z użyciem mikrofonów ambisonicznych. Charakterystykę częstotliwościową tych różnic przedstawiono na rys. 8.13.



Rys. 8.13. Czas wczesnego zaniku EDT. Różnice zmierzone mikrofonem pomiarowym B&K 4189 i przy użyciu kanału W mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E

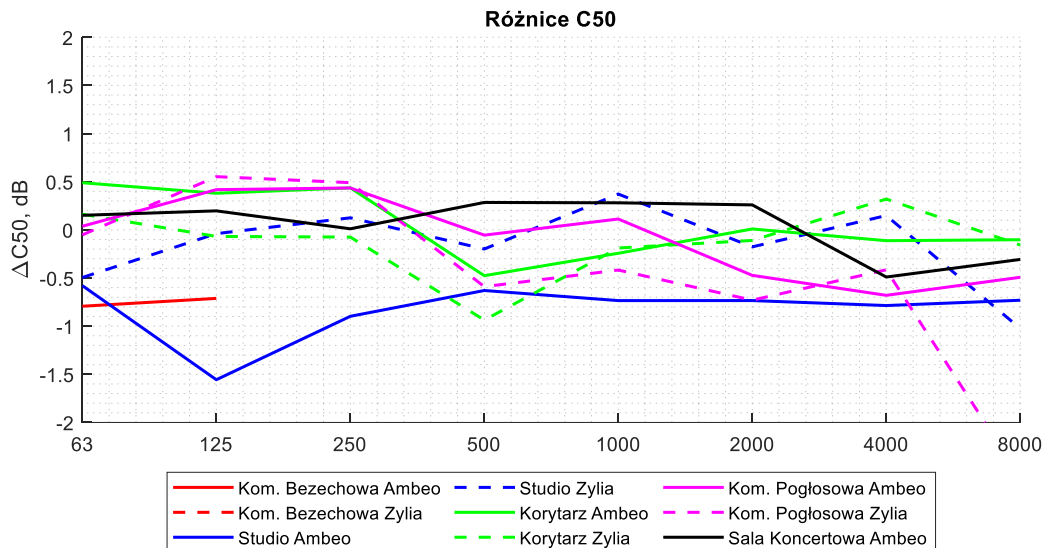
Zarówno w komorze pogłosowej, jak i pozostałych pomieszczeniach częściowo pogłosowych (studio, korytarz, sala koncertowa), odstępstwa przy pomiarze EDT nie przekraczają 10%, a w większości przypadków mieszczą się w granicach progu JND wynoszącego podobnie jak w przypadku czasu pogłosu 5%. Jest to wynik spodziewany, ponieważ wartości EDT są w jakimś stopniu skorelowane z wartościami czasu pogłosu. Wyjaśnienia wymagają natomiast znaczne różnice uzyskane w przypadku komory bezechowej. Nie są one reprezentatywne, ponieważ wynikają z niezwykle małych zmierzonych wartości EDT, które dla 8 kHz wynoszą 0,3 ms. Analiza wariancji przedstawiona w załączniku 3 wskazuje, że prawie we wszystkich przypadkach, z wyjątkiem wyników uzyskanych w komorze bezechowej, różnice przedstawione na rys. 8.13 nie są istotne statystycznie. Podobna sytuacja występuje dla parametrów C50, C80 i Ts przedstawionych w rozdziałach 8.2.3 do 8.2.5.

8.2.3. Czytelność dźwięku C50



Rys. 8.14. Wskaźnik czytelności dźwięku C50 zmierzony mikrofonem pomiarowym B&K 4189 w pięciu badanych pomieszczeniach.

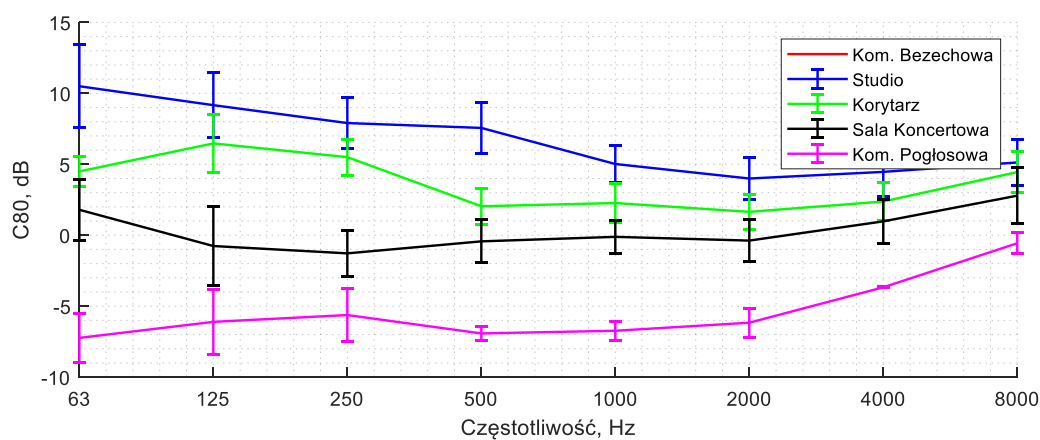
Przebieg wskaźnika czytelności dźwięku C50 w funkcji częstotliwości został przedstawiony na rys. 8.14 i wskazuje na prawidłową zależność wartości C50 od charakteru pomieszczenia. Największe wartości (0,4 – 6,2 dB) zaobserwowano w studio, mniejsze w korytarzu (-1,0 do 2,8 dB), sali koncertowej (-5 do -0,2), aż wreszcie najmniejsze dla komory pogłosowej (-9,7 do -2,8 dB). A zatem, zgodnie z oczekiwaniami, przy tym uszeregowaniu pomieszczeń maleje czytelność mowy zgodnie ze wzrostem pogłosowości. W danych pominięte są wyniki dla komory bezechowej, w której silne wyłumienie uniemożliwiało uzyskanie prawidłowych wartości C50, poza najniższymi częstotliwościami 63 Hz i 125 Hz, przy których C50 przekraczało 25 dB.



Rys. 8.15. Czytelność dźwięku C50. Różnice zmierzone mikrofonem pomiarowym B&K 4189 i przy użyciu kanału W mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E

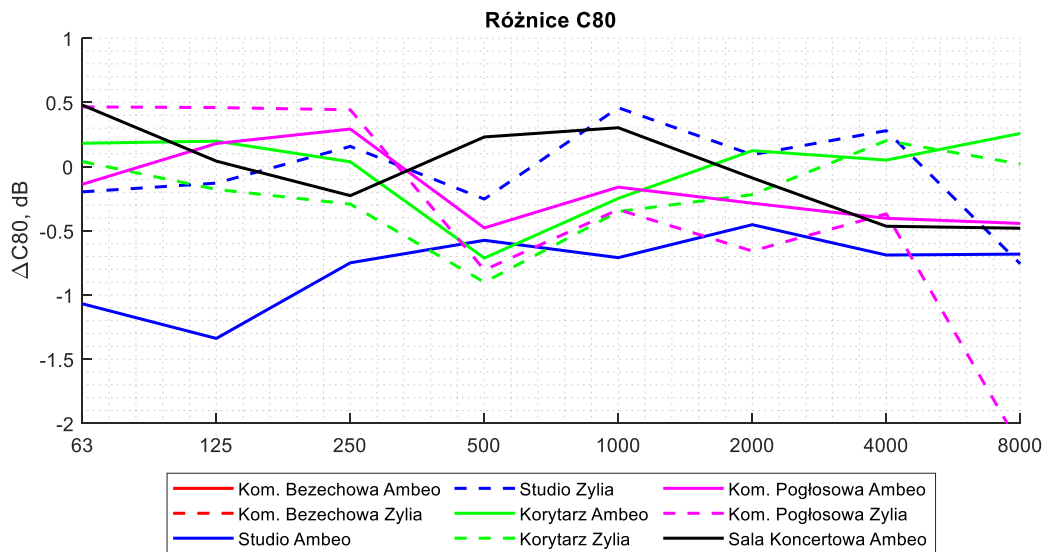
Na rys. 8.15 przedstawiono różnicę $\Delta C50$ wartości uzyskanej z użyciem mikrofonu pomiarowego i z użyciem mikrofonu ambisonicznego. Dla wskaźnika C50, progowa zmiana JND określana jest na 1 dB (por. tab 2.1). W odniesieniu do $\Delta C50$, wartość ta, jest przekraczana w nielicznych przypadkach: przy 8 kHz (mikrofon Zylia, studio dźwiękowe i komora pogłosowa) oraz przy 125 Hz (mikrofon Ambeo, studio dźwiękowe). Natomiast charakterystyczne jest, że w większości przypadków wartości pochodzące z pomiaru mikrofonami ambisonicznymi są minimalnie większe od wartości pochodzących z pomiaru mikrofonem pomiarowym.

8.2.4. Przejrzystość dźwięku C80



Rys. 8.16. Wskaźnik przejrzystości dźwięku C80 zmierzony mikrofonem pomiarowym B&K 4189 w pięciu badanych pomieszczeniach.

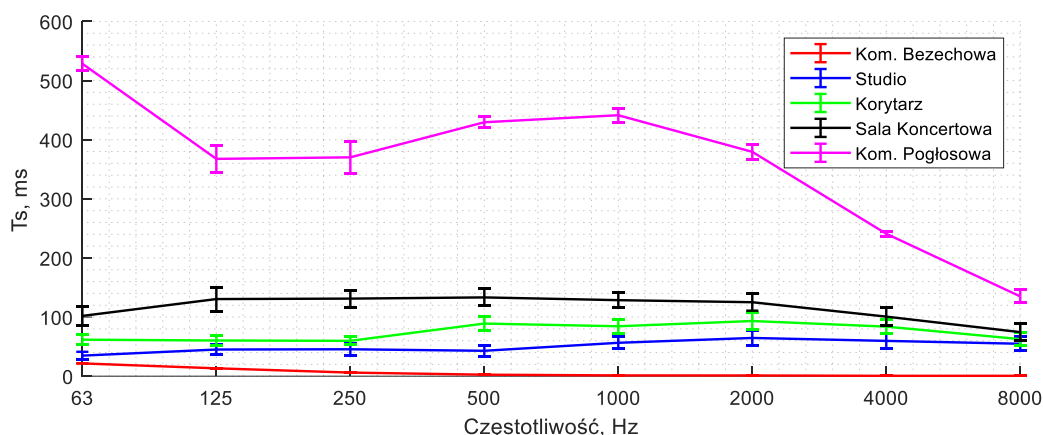
Przebieg czytelności dźwięku C80 w funkcji częstotliwości został przedstawiony na rys. 8.16. Podobnie, jak w przypadku C50, w danych pominięte są wyniki dla komory bezechowej, której wytlumienie całkowicie uniemożliwiło uzyskanie prawidłowych wartości wskaźnika C80. Występuje podobne uporządkowanie wartości C80 do uszeregowania C50, z uwagi na ich bliski związek. W zależności od charakteru pomieszczenia, C80 przyjmuje wartości od (4,0 – 10,5 dB) w studio, (-1,6 do 6,5 dB) w korytarzu, (-1,3 do -2,8 dB) w sali koncertowej, do (-7,0 do -0,6 dB) w komorze pogłosowej. A zatem, zgodnie z oczekiwaniami, przy tym uszeregowaniu pomieszczeń, tak samo jak w przypadku C50, przejrzystość dźwięku C80 maleje zgodnie ze wzrostem pogłosowości.



Rys. 8.17. Czytelność dźwięku C80. Różnice zmierzone mikrofonem pomiarowym B&K 4189 i przy użyciu kanału W mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E

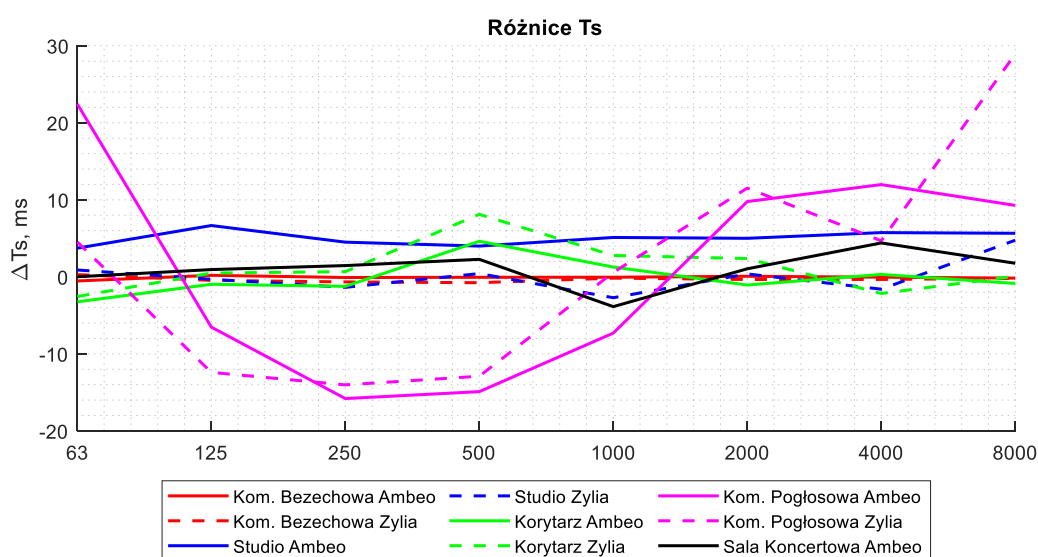
Różnice wartości C80 uzyskanych z użyciem pomiarowego mikrofonu referencyjnego i mikrofonu ambisonicznego $\Delta C80$ przedstawiono na rysunku 8.17. Dla wskaźnika C80, progowa zmiana JND wynosi 1 dB, tak jak dla wskaźnika C50 (por tab 2.1). W przypadku $\Delta C80$, minimalne przekroczenie ($\Delta C80 > 1 \text{ dB}$) występuje wyłącznie w trzech przypadkach (mikrofon Ambeo w Studio, -1,1 dB przy 63 Hz i -1,3 dB przy 125 Hz oraz w komorze pogłosowej dla mikrofonu Zylia -2,2 dB w paśmie 8 kHz). Jak można było przypuszczać, różnice w przypadku wskaźnika przejrzystości dźwięku C80 są podobne do tych, które występowały w przypadku czytelności dźwięku C50.

8.2.5. Czas Środkowy Ts



Rys. 8.18. Wskaźnik czasu środkowego Ts zmierzony mikrofonem pomiarowym B&K 4189 w pięciu badanych pomieszczeniach.

Na rys 8.18 przedstawiono ostatni z rozpatrywanych parametrów wszechkierunkowych – czas środkowy Ts. Jest to parametr bezpośrednio zależny od czasu pogłosu i dlatego jego wartości w poszczególnych salach układały się zgodnie z charakterystykami czasu pogłosu. Najmniejsze wartości w przedziale od 0,6 ms (8 kHz) do 22 ms (63 Hz) występowały dla komory bezechowej, następnie w sali studio dźwiękowego od 35 ms (63 Hz) do 65 ms (2 kHz), dla korytarza od 60 ms (250 Hz) do 93 ms (2 kHz), a dla sali koncertowej 75 ms (8 kHz) do 133 ms (500 Hz). Jak należy oczekiwać, skrajne wartości występowały dla komory pogłosowej od 135 ms (8 kHz) do ponad 0,5 s (63 Hz).



Rys. 8.19. Czas środkowy Ts. Różnice zmierzone mikrofonem pomiarowym B&K 4189 i przy użyciu kanału W mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E

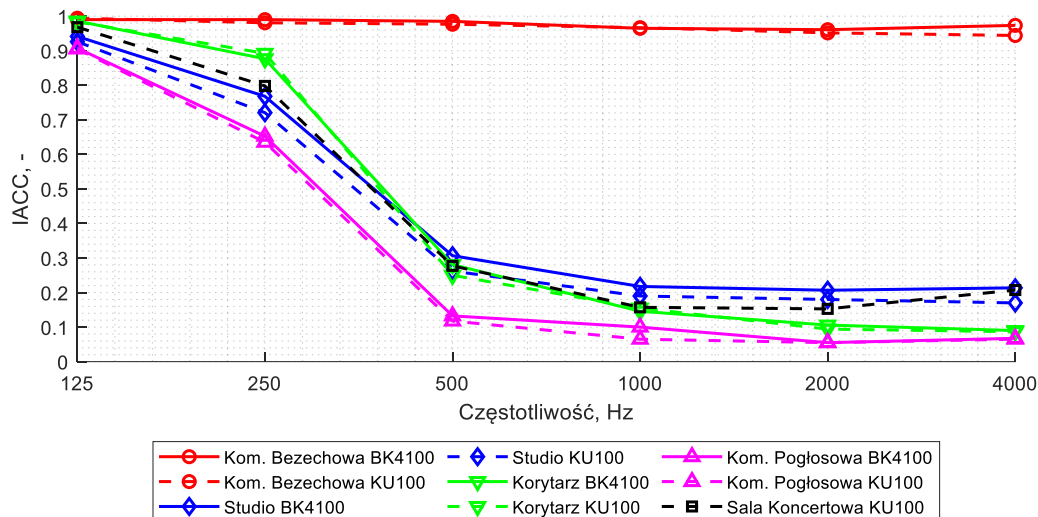
Różnice T_s uzyskane dla mikrofonów ambisonicznych Sennheiser Ambeo i Zylia Z1-3E przedstawione na rys. 8.19 praktycznie nie odbiegały od tych wartości, ponieważ pomijając pomiary w komorze pogłosowej, nie różniły się o więcej niż 10 ms, co stanowi JND w przypadku czasu T_s . Nieco większe różnice w przypadku komory pogłosowej są nieistotne w przypadku dużych wartości T_s , jakie występowały w pomiarach w tych warunkach pogłosowych. Należy podkreślić, że wyniki uzyskane w komorze bezehowej są nieistotne, ponieważ czas T_s jest praktycznie bliski zeru.

8.3. Międzyuszna korelacja skrośna IACC wyznaczana z binauralnej odpowiedzi impulsowej BRIR

Jak szczegółowo opisano w punkcie 7.2, uzyskanie binauralnej odpowiedzi impulsowej z sygnału ambisonicznego wymaga przeprowadzenia binauralizacji tego sygnału, w skrócie polegającej na określeniu kierunku nadejścia dźwięku oraz jego filtracji przez właściwie dobrany, odpowiadający manekinowi pomiarowemu lub sztucznej głowie zestaw filtrów HRTF. Wyniki obliczeń parametru IACC przeprowadzonych na podstawie binauralnej odpowiedzi impulsowej z wykorzystaniem czterech opisanych w punkcie 4.2 metod upraszczania HRTF przedstawiono w punktach 8.3.2 do 8.3.5. W każdym przypadku, analizowano wyniki pomiarów przy użyciu tak 1550 punktowego HRTF, jak i 8802 punktowego HRTF. Natomiast w punkcie 8.3.1 porównano pomiary IACC przeprowadzone z użyciem manekina B&K 4100-D i sztucznej głowy Neumann KU-100.

8.3.1. Porównanie pomiaru IACC przeprowadzonego przy pomocy manekina B&K 4100-D i sztucznej głowy Neumann KU-100

Wykresy na rys 8.20 przedstawiają przebieg parametru IACC zmierzonego z użyciem manekina B&K 4100-D i sztucznej głowy Neumann KU-100 w komorze bezehowej, studio dźwiękowym, na korytarzu i w komorze pogłosowej. W sali koncertowej przeprowadzony był pomiar wyłącznie z użyciem sztucznej głowy Neumann KU-100.



Rys. 8.20. Współczynnik międzyuszej korelacji skrośnej IACC zmierzony z użyciem manekina pomiarowego B&K4100-D i sztucznej głowy KU-100 w pięciu badanych pomieszczeniach

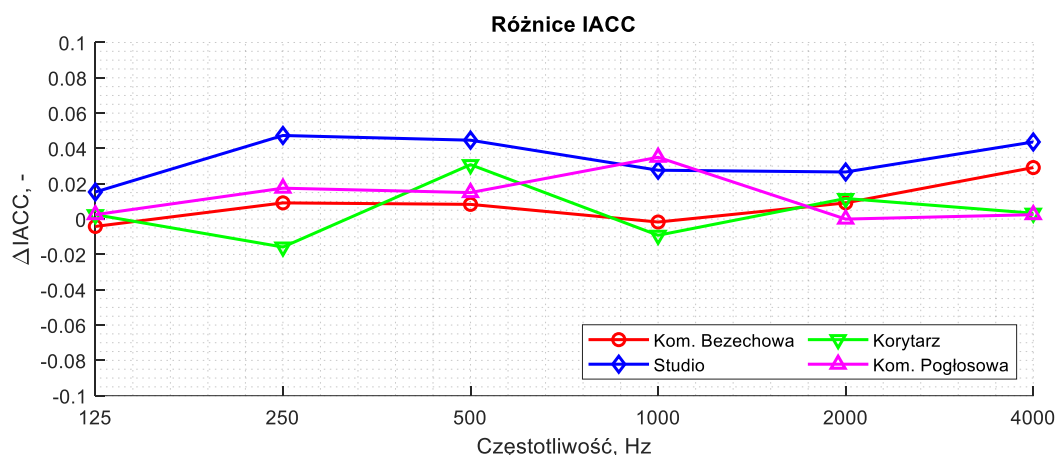
Przebieg IACC w pasmach oktaowych w odniesieniu do studio dźwiękowego, korytarza, sali koncertowej i komory pogłosowej jest zbliżony i charakterystyczny tym, że poniżej częstotliwości 500 Hz IACC przyjmuje duże wartości, co świadczy o podobieństwie sygnałów w prawym i lewym uchu wynikającym z jednorodności pola akustycznego i braku zacienienia akustycznego przez głowę. Dla częstotliwości 500 Hz i wyższych, IACC spada do wartości między 0,02 i 0,20, co jest związane z rozproszeniem pola akustycznego. Tego rodzaju przebieg jest typowo obserwowany w pomieszczeniach.

Różnice pomiędzy pomieszczeniami są znaczące, przyjmując jako kryterium JND dla IACC, które wynosi 0,075. W komorze pogłosowej, w całym zakresie wartość IACC jest mniejsza niż w przypadku pozostałych pomieszczeń, co jest zgodne z oczekiwaniami biorąc pod uwagę dyfuzyjność pola akustycznego. W zakresie niskich częstotliwości (250 Hz), IACC zwiększa się kolejno od komory pogłosowej, do studio dźwiękowego, sali koncertowej i korytarza w przedziale od ok. 0,6 do ponad 0,9. Większe znaczenie ma zmienność IACC w zakresie częstotliwości od 500 Hz, która wynosi od 0,05 do ponad 0,2 (wartości przy 2 kHz), zwiększając się kolejno od komory pogłosowej, do korytarza, sali koncertowej i studio dźwiękowego. Jest to wzrost związany z malejącym czasem pogłosu, za wyjątkiem studio prawdopodobnie z uwagi na zastosowane ustroje rozpraszające w tym pomieszczeniu.

Charakterystyczna jest wartość $IACC > 0,99$ (z danych) prawie niezależna od częstotliwości w przypadku rejestracji w warunkach bezechowych. Jest to zgodne z oczekiwaniami, ponieważ w tych warunkach całkowitego braku odbić, sygnał rejestrowany w prawym i lewym

uchu jest identyczny przy braku opóźnienia niwelowanego nie tylko procedurą obliczania współczynnika międzyusznej korelacji skrośnej, ale przy ustawieniu głowy skierowanej na źródło dźwięku.

Na rysunku 8.21 przedstawiono różnice $\Delta IACC$ w pomiarach manekinem B&K 4100-D i sztuczną głową KU-100. Różnice te wskazują, że stosowanie obu manekinów jest równoważne, przyjmując wspomniany wcześniej JND dla IACC równy 0,075 (zgodnie z kryterium w normie PN-EN ISO 3382-1 [73]).



Rys. 8.21. Różnice współczynnika międzyusznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i przy użyciu sztucznej głowy Neumann KU-100.

Różnice uzyskane w pomiarach przy użyciu obu manekinów nie przekraczają $\pm 0,03$ w komorze bezechowej i korytarzu oraz $+0,05$ w studio dźwiękowym i komorze pogłosowej. Wartości te są wyraźnie mniejsze od uznanego kryterium, zatem można przyjąć, że różnice pomiędzy pomiarami przy użyciu obydwu manekinów nie są istotne.

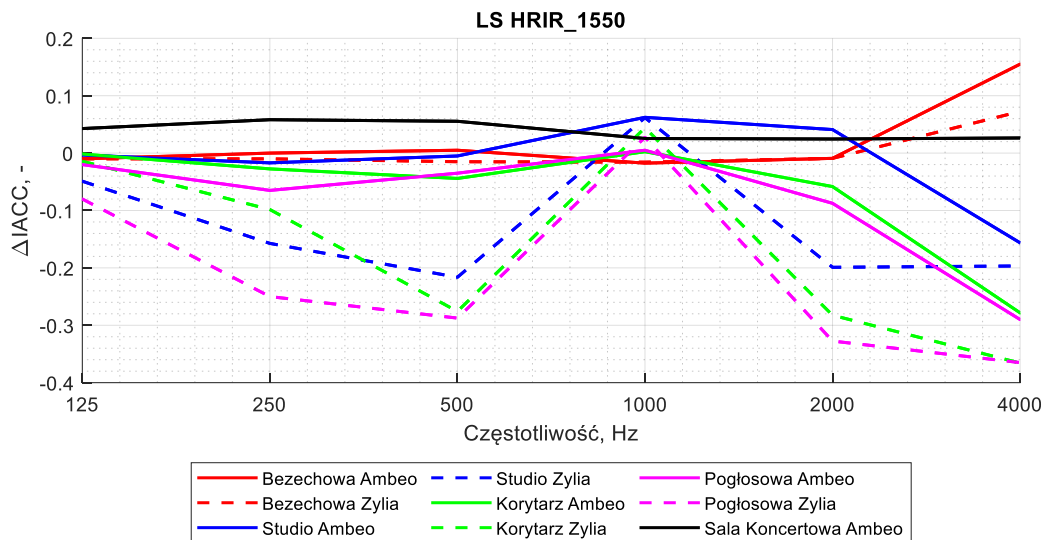
W dalszej analizie, przedstawionej w punktach 8.3.2 do 8.3.5, ocenę wyników obliczeń IACC na podstawie BRIR uzyskanego przy użyciu czterech procedur binauralizacji przeprowadzono przez porównanie z pomiarami z użyciem manekina B&K 4100-D w komorach bezechowej i pogłosowej oraz sali studio i korytarzu. Wybór ten podyktowany był tym, że manekin B&K4100-D jest wyposażony w mikrofony pomiarowe. W przypadku danych dotyczących sali koncertowej w Bydgoszczy, porównanie przeprowadzono w odniesieniu do pomiarów sztucznej głowy Neumann KU-100, ponieważ w tej sali pomiary z użyciem manekina B&K 4100-D nie były przeprowadzane.

W załączniku 3 tab. Z3.2 i Z3.3 zamieszczono wyniki analizy wariancji istotności różnic ($p = 0,05$) w wyznaczaniu parametru IACC na podstawie przestrzennej odpowiedzi SRIR i binauralnej odpowiedzi impulsowej pomieszczenia BRIR, a zmierzonymi wartościami manekina akustycznego B&K4100-D i sztucznej głowy Neumann KU-100 przy uwzględnieniu

metod upraszczania i rozdzielczości HRTF. Analiza wariancji potwierdziła istotność różnic w odniesieniu do największych różnic wykazanych w punktach 8.3.2 do 8.3.6.

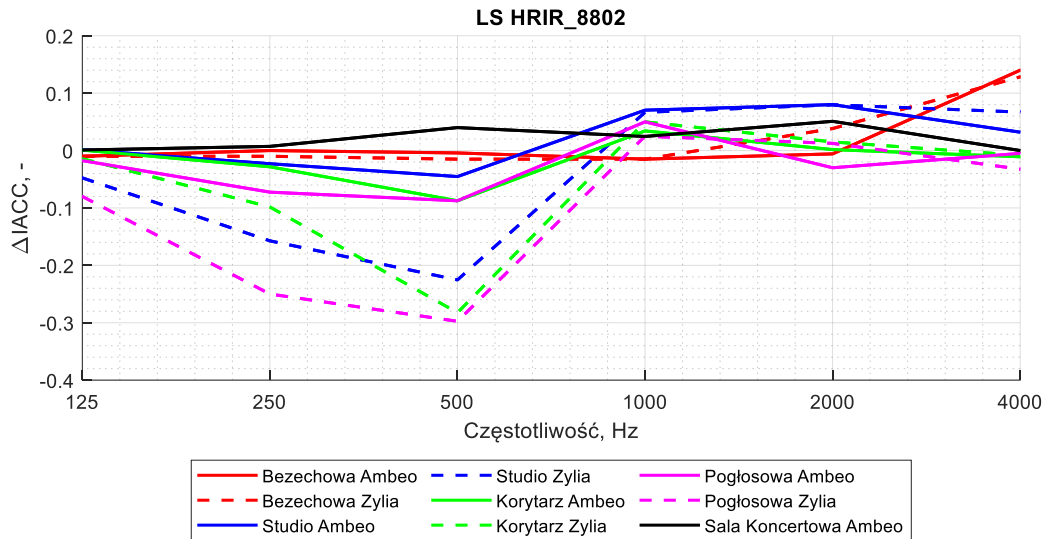
8.3.2. Metoda Least Squares (LS)

Różnice ($\Delta IACC$) pomiędzy wartościami międzyusznej korelacji skrośnej uzyskanej na podstawie BRIR i metody LS do utworzenia uproszczonego HRTF oraz otrzymanych w pomiarze z użyciem manekina akustycznego przedstawiono na rys. 8.22 i 8.23 odpowiednio dla oryginalnej rozdzielczości HRTF równej 1550 lub 8802 punkty w przestrzeni.



Rys. 8.22. Metoda LS upraszczania HRTF, rozdzielczość 1550. Różnice współczynnika międzyusznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

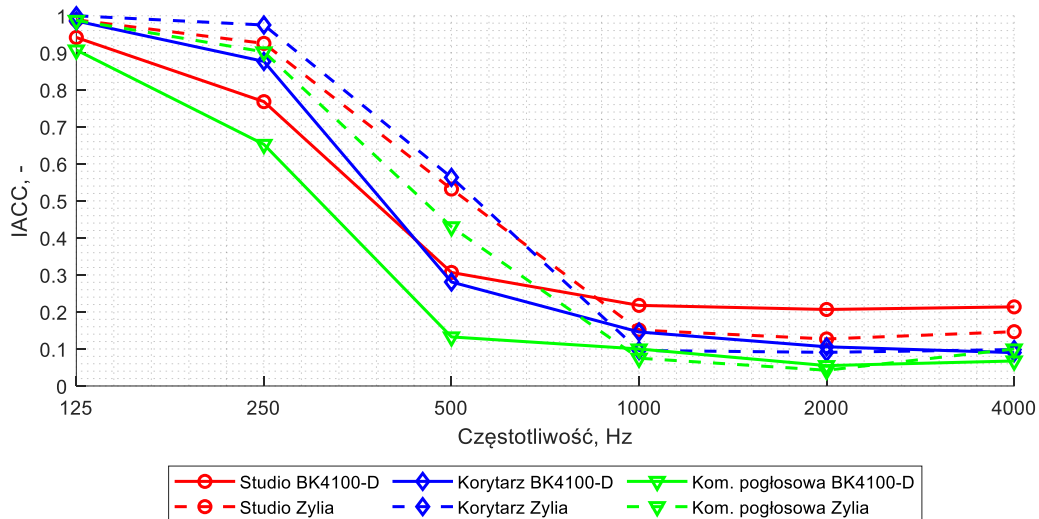
W przypadku niższej rozdzielczości punktowej HRTF (rys. 8.22) zasadnicze odstępstwa sięgające 4 JND (0,300) zachodzą w przypadku mikrofonu Zylia w zakresie częstotliwości poniżej 1 kHz oraz dla najwyższych częstotliwości 2 i 4 kHz w odniesieniu do komory pogłosowej, korytarza i studio dźwiękowego, a więc przy większej pogłosowości pomieszczenia. Dla mikrofonu Ambeo, różnice zasadniczo nie przekraczają 1 JND, z wyjątkiem częstotliwości 4 kHz.



Rys. 8.23. Metoda LS upraszczania HRTF, rozdzielczość 8802. Różnice współczynnika między-usznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

W przypadku większej (8802) rozdzielczości punktowej HRTF (rys. 8.23) następuje znaczne zmniejszenie różnic w stosunku do pomiaru manekinami akustycznymi w zakresie wysokich częstotliwości, szczególnie istotne w przypadku mikrofonu Zylia. W zakresie poniżej 1 kHz odstępstwa sięgające 4 JND (0,300) zachodzące dla mikrofonu Zylia w zakresie częstotliwości poniżej 1 kHz pozostają takie, jakie były w przypadku HRTF o rozdzielczości 1550 punktów. Dla mikrofonu Ambeo, podobnie jak dla HRTF o niższej rozdzielczości, różnice zasadniczo nie przekraczają 1 JND, z wyjątkiem częstotliwości 4 kHz.

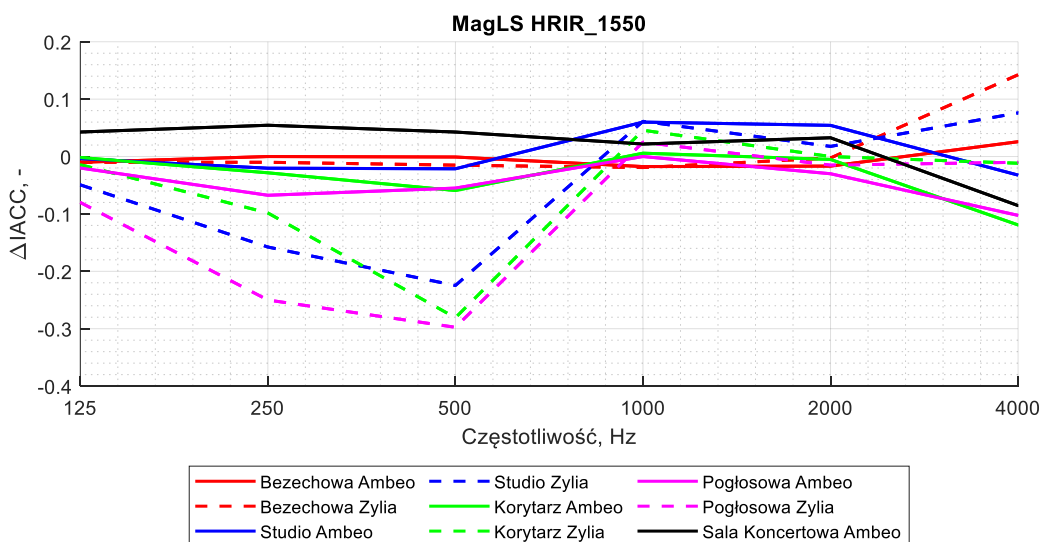
Wyżej przedstawione zdefiniowanie błędu $\Delta IACC$ oznacza, że wartości IACC uzyskane za pomocą mikrofonu Zylia przewyższają wartości zmierzone za pomocą manekina B&K4100-D w taki sposób, że dla najniższych częstotliwości współczynnik IACC praktycznie równał się jedności. Dla ilustracji tego zjawiska na rys. 8.24 przedstawiono wartości IACC zmierzone z użyciem mikrofonu Zylia w różnych pomieszczeniach i przetworzone za pomocą HRTF odpowiednio o rozdzielczości 1550 i 8802 punkty w przestrzeni.



Rys. 8.24. Metoda LS upraszczania HRTF, rozdzielczość 8802. Współczynnika międzyuszej korelacji skrośnej IACC zmierzony manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

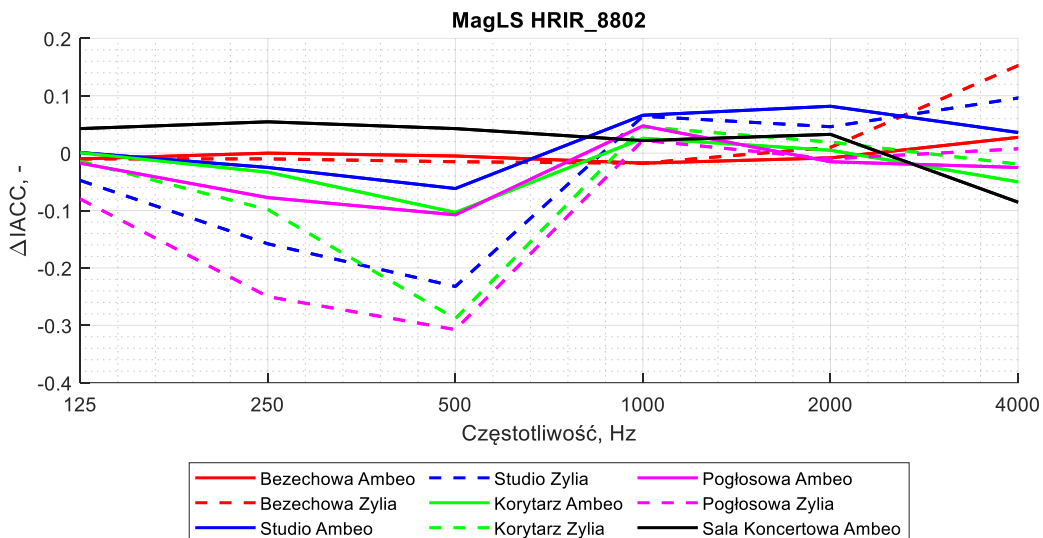
8.3.3. Metoda Magnitude Least Squares (MagLS)

Wynik zastosowania metody MagLS do utworzenia uproszczonego HRTF używanego do uzyskania odpowiedzi impulsowej BRIR, przedstawiono na rys. 8.25 i 8.26 odpowiednio dla rozdzielczości HRTF równej 1550 i 8802 punkty w przestrzeni.



Rys. 8.25 Metoda MagLS upraszczania HRTF, rozdzielczość 1550. Różnice współczynnika międzyuszej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

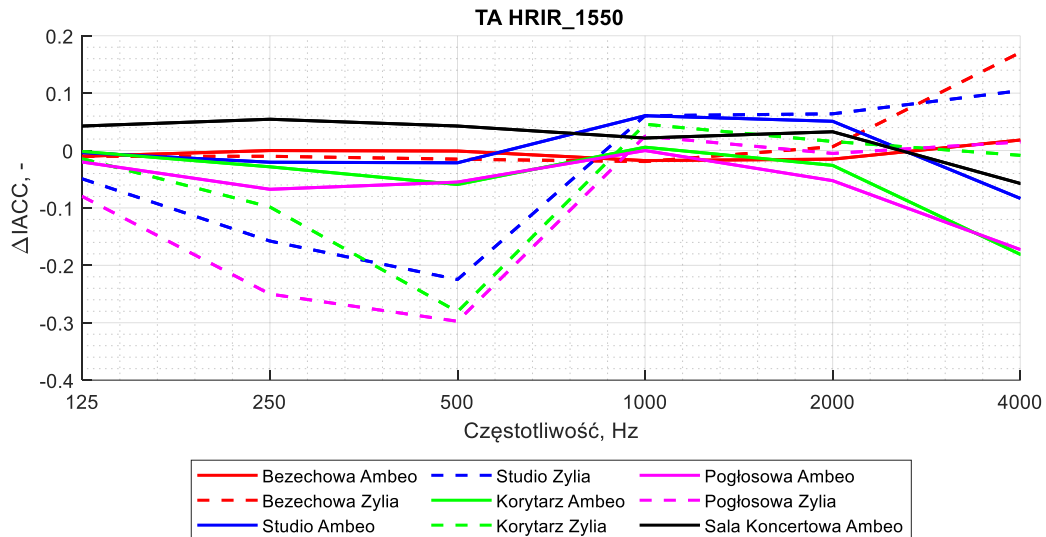
Zauważyć można, że w przypadku metody MagLS wpływ oryginalnej rozdzielczości HRTF jest nieduży, ponieważ wykresy przedstawione na rys. 8.25 i 8.26 są praktycznie identyczne. Główna różnica w stosunku do poprzednio rozważanej metody LS odnosi się do zakresu wysokich częstotliwości, co jest szczególnie widoczne w przypadku mikrofonu Zylia przy częstotliwościach 2 i 4 kHz, bowiem odstępstwa od wartości IACC uzyskanych za pomocą manekina akustycznego są znacznie mniejsze, niż były z użyciem metody LS. W zakresie częstotliwości poniżej 1 kHz, wyniki są zbliżone do przypadku zastosowania metody LS. Ma tu miejsce względnie małe odstępstwo od pomiaru manekinem akustycznym w przypadku mikrofonu Ambeo i sięgające 4 JND przewyższenie wartości zmierzonej z użyciem manekina akustycznego przez obliczenia wykorzystujące pomiar mikrofonem Zylia.



Rys. 8.26. Metoda MagLS upraszczania HRTF, rozdzielczość 8802. Różnice współczynnika międzyuszy korelacji skróśnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

8.3.4. Metoda Time Alignment (TA)

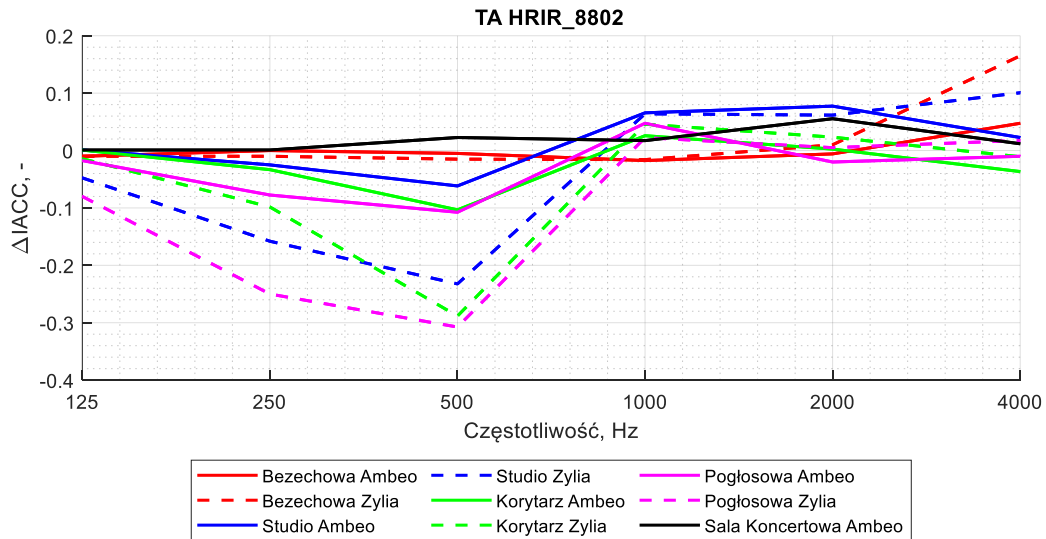
Wynik zastosowania metody TA do utworzenia uproszczonego HRTF używanego do uzyskania odpowiedzi impulsowej BRIR, przedstawiono na rys. 8.27 i 8.28 ponownie odpowiednio dla rozdzielczości HRTF równej 1550 i 8802 punkty w przestrzeni.



Rys. 8.27. Metoda TA upraszczania HRTF, rozdzielczość 1550. Różnice współczynnika międzyuszy korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

W przypadku niższej rozdzielczości punktowej HRTF (rys. 8.28) zaobserwowane odstępstwa wartości IACC wynikających z obliczeń z danymi pochodzącymi z mikrofonu Zylia od zmierzonych manekinem akustycznym są bardzo podobne do wcześniej obserwowanych przy metodach LS i MagLS. A więc, poniżej 1 kHz i pomieszczeń – studio dźwiękowe, korytarz i sala koncertowa – odstępstwa od IACC zmierzonego manekinem sięgają 4 JND. Funkcja IACC wyznaczona na podstawie BRIR ma charakter zbliżony do pokazanej wcześniej na rys. 8.24. Natomiast w przypadku mikrofonu Ambeo występuje głębsze, sięgające 2 JND odstępstwo od pomiarów z użyciem manekina B&K 4100-D.

W przypadku mikrofonu Ambeo, znacząca poprawa występuje, jeśli HRTF ma rozdzielczość 8802 punktu. Praktycznie przy wszystkich częstotliwościach odstępstwo od pomiarów manekinem jest najmniejsze. W przypadku mikrofonu Zylia, wartości $\Delta IACC$ praktycznie nie zależą od rozdzielczości HRTF.

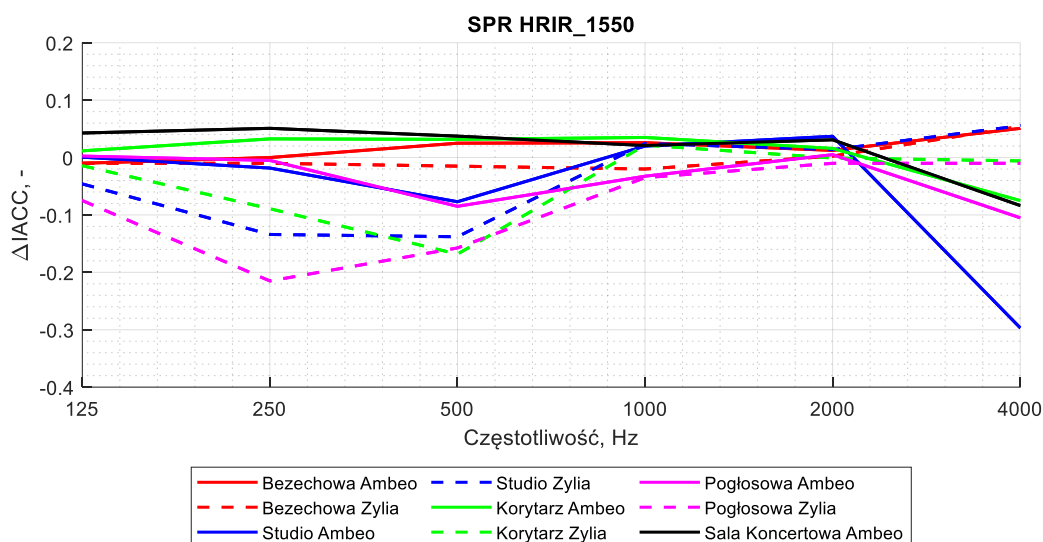


Rys. 8.28. Metoda TA upraszczania HRTF, rozdzielczość 8802. Różnice współczynnika międzyusznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

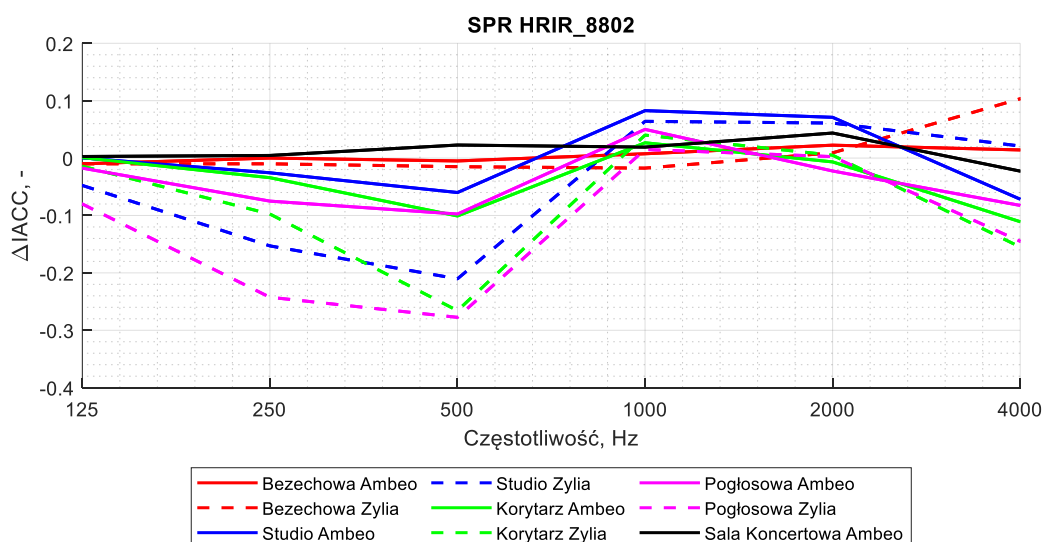
8.3.5. Metoda Spatial Resampling (SPR)

Wynik zastosowania metody SPR do utworzenia uproszczonego HRTF używanego do uzyskania odpowiedzi impulsowej BRIR, przedstawiono na rys. 8.29 i 8.30 odpowiednio dla rozdzielczości HRTF równej 1550 i 8802 punkty w przestrzeni.

Przy niższej rozdzielczości HRTF wynoszącej 1550 punktów, zastosowanie SPR powoduje mniejszy błąd w przypadku mikrofonu Zylia w zakresie częstotliwości poniżej 1 kHz sięgający 3 JND. W zakresie wysokich częstotliwości, nawet przy 4 kHz, błąd przy użyciu mikrofonu Zylia jest również nieduży (mniejszy niż JND), natomiast przy 4 kHz, większy błąd, sięgający 4 JND ma miejsce dla mikrofonu Ambeo w przypadku pomiaru w studio dźwiękowym.



Rys. 8.29. Metoda SPR upraszczania HRTF, rozdzielczość 1550. Różnice współczynnika międzyusznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR



Rys. 8.30. Metoda SPR upraszczania HRTF, rozdzielczość 8802. Różnice współczynnika międzyusznej korelacji skrośnej IACC zmierzone manekinem pomiarowym B&K 4100D i wyznaczone z użyciem przestrzennej odpowiedzi impulsowej BRIR

Wykorzystanie HRTF o wyższej rozdzielczości 8802 zwiększa błąd występujący w przypadku użycia mikrofonu Zylia tak w zakresie poniżej 1 kHz, jak i dla 4 kHz, natomiast zmniejsza błąd w przypadku mikrofonu Ambeo.

8.4. Propozycja rachunku błędu pomiaru mikrofonem ambisonicznym

W akustyce pomieszczeń zagadnienie niepewności pomiarowej jest problemem złożonym. Na wynik pomiaru ma wpływ nie tylko jakość użytego sprzętu pomiarowego i precyzja obliczeń, lecz także precyzyjne ustawienie źródła i odbiornika oraz warunki atmosferyczne wpływające na propagację dźwięku. Elementy te były wymienione w pracy Lundeby i in. [43], a szczegółowo opisywane są w pracy Witewa [105].

Norma PN-EN ISO 3382-1 [73] określa sposób wyznaczania niepewności pomiaru wyłącznie dla wskaźników czasu pogłosu T20 i T30. Norma zakłada, że metodą pomiaru jest technika szumu przerywanego i jedynie w odniesieniu do tej metody pozwala na wyznaczenie odchylenia standardowego wartości czasu pogłosu, biorąc pod uwagę liczbę uśrednień krzywych zaniku w danym punkcie pomiarowym przy określeniu krzywej zaniku na podstawie jej podobieństwa do linii prostej.

W technikach pomiarowych wyznaczania parametrów akustycznych pomieszczeń wykorzystujących pobudzenie przestrajającym tonem lub sygnałem pseudolosowym MLS krzywa zaniku wyznaczana jest z zarejestrowanej i obliczonej odpowiedzi impulsowej, a wynik jest uzyskiwany już w pojedynczym pomiarze. W związku z tym, na dokładność wyznaczenia krzywej zaniku nie ma wpływu losowość zastosowanego sygnału pobudzającego. Zgodnie z normą PN-EN-ISO 18233 [72] należy przeprowadzać analizę niepewności zmierzonej odpowiedzi impulsowej na podstawie Przewodnika Wyrażania Niepewności Pomiaru [61].

Bardzo dobry budżet czynników tworzących niepewność pomiaru parametrów akustycznych wewnątrz jest przedstawiony w pracy Witewa [105]. Podstawą przedstawionej w tym rozdziale propozycji jest rozszerzenie rachunku niepewności Witewa na przetwornik wielomikrofonowy w postaci mikrofonu ambisonicznego.

Pełny budżet niepewności pomiaru dla odpowiedzi impulsowej zgodnie z pracą Witewa [105] przedstawiono w tabeli 8.1.

Tabela 8.1. Budżet niepewności pomiarowej według pracy Witewa [105]

Symbol	Źródło niepewności	Udział w niepewności
b_t		$u(b_t)$, dB
b_{LS-dir}	Kierunkowość źródła	0,38
$b_{LS-level}$	Kalibracja poziomu źródła	0,26
$b_{Mic-cal}$	Wzorcowanie kalibratora	0,21
b_{LTI}	Stabilność długoterminowa	0,2
$b_{LS-spec}$	Charakterystyka częstotliwościowa głośnika	0,16
$b_{Mic-spec}$	Charakterystyka częstotliwościowa mikrofonu	0,15
b_{Filter}	Filtracja oktawaowa	0,12
$b_{Meteo-\theta}$	Zmiany tempertatury	0,12
$b_{Meteo-p}$	Zmiany ciśnienia atmosferycznego	0,07
$b_{Mic-dir}$	Kierunkowość mikrofonu	0,07
$b_{Mic-field}$	Zakłócenia pola akustycznego spowodowane obecnością mikrofonu	0,011
$b_{Meteo-rh}$	Zmiany wilgotności	$7,5 * 10^{-3}$
b_{LS-THD}	Zniekształcenia harmoniczne źródła dźwięku	Ok $1 * 10^{-3}$
$b_{Amp-THD}$	Zniekształcenia harmoniczne wzmacniacza	Ok $1 * 10^{-4}$
$b_{Mic-amp}$	Liniowość wzmacniacza	(0,17) Ok $1 * 10^{-4}$
$b_{Amp-SNR}$	Szumy własne wzmacniacza	Ok $1 * 10^{-6}$
$b_{Mic-SNR}$	Szumy własne mikrofonu	Ok $1 * 10^{-6}$
$b_{D/A-SNR}$	Szumy przetwornika C/A	Ok $1 * 10^{-8}$
$b_{D/A-THD}$	Zniekształcenia przetwornika C/A	Ok $1 * 10^{-8}$
$b_{clock jitter}$	Jitter zegara	Ok $1 * 10^{-8}$
$b_{Mic-THD}$	Zniekształcenia nieliniowe w mikrofonie	Ok $1 * 10^{-9}$
$b_{A/D-THD}$	Zniekształcenia przetwornika A/C	Ok $1 * 10^{-9}$
$b_{quant-SNR}$	Szumy kwantyzacji	Ok $1 * 10^{-12}$
$u_c(b_{Equip})$	Niepewność złożona	0,6
$U(b_{Equip})$	Niepewność rozszerzona	1,2 (k=2)

Z budżetu przedstawionego w tabeli 8.1 niepewność złożoną oblicza się zgodnie ze wzorem (8.1) biorąc pod uwagę wszystkie czynniki znaczące, tj. większe niż 10^{-2} dB. Obejmuje to pierwszych 11 źródeł niepewności podanych w tabeli, spośród których pięć wyróżnionych podkreśleniem odnosi się do mikrofonu.

$$\begin{aligned}
 u_c(b_{Equipment}) = & [u^2(b_{LS-dir}) + u^2(b_{LS-level}) + u^2(b_{LTI}) + \\
 & + u^2(b_{LS-spec}) + u^2(b_{Mic-spec}) + \\
 & + u^2(b_{Mic-cal}) + u^2(b_{Filter}) + \\
 & + u^2(b_{Meteo-\vartheta}) + u^2(b_{Meteo-p}) + \\
 & + u^2(b_{Mic-dir}) + u^2(b_{Mic-field})]^{1/2}
 \end{aligned} \tag{8.1}$$

Typowy wynik niepewności złożonej z tak opracowanego budżetu przy jednym mikrofonie pomiarowym wynosi 0,6 dB, co podaje tabela 8.1, a odnosi się do każdej wartości amplitudy odbić występujących w odpowiedzi impulsowej. Jest to więc niepewność pomiaru odnosząca się do amplitud w zarejestrowanym echogramie.

W przypadku oszacowania niepewności w pomiarach przy użyciu matrycy wielomikrofonowej wydaje się uzasadniona modyfikacja metody Wittewa polegająca na uwzględnieniu kolejnych czynników niezależnie dla każdej kapsuły mikrofonu oraz oddzielnie dla zakłócenia pola akustycznego spowodowanego obecnością mikrofonu ambisonicznego jako całości. Prowadzi to do modyfikacji wzoru (8.1) do postaci wzoru (8.2), w której ostatni składnik wzoru (8.1) został przeniesiony do wzoru (8.3) wyznaczającego sumaryczną niepewność pomiaru.

$$\begin{aligned}
 u_c(b_{Equipment}) = & [u^2(b_{LS-dir}) + u^2(b_{LS-level}) + u^2(b_{LTI}) + \\
 & + u^2(b_{LS-spec}) + u^2(b_{Mic-spec}) + \\
 & + u^2(b_{Mic-cal}) + u^2(b_{Filter}) + \\
 & + u^2(b_{Meteo-\vartheta}) + u^2(b_{Meteo-p}) + \\
 & + u^2(b_{Mic-dir})]^{1/2}
 \end{aligned} \tag{8.2}$$

Zakładając, że zespół mikrofonów tworzących matrycę mikrofonową wprowadza niezależną niepewność pomiarową – można ją oszacować na podstawie zależności w postaci

wzoru (8.3), stanowiącego uogólnienie wzoru (8.2) opisującego sumę niezależnych niepewności pomiarowych.

$$u_{c_n} = \sqrt{u_{c_{n1}}^2 + u_{c_{n2}}^2 + \dots + u_{c_{nn}}^2 + u^2(b_{Mic-field})} \quad (8.3)$$

Przy założeniu, że w matrycy występuje n identycznych mikrofonów, wzór X.3 może zostać uproszczony do postaci wzoru (8.4), który można przyjąć za podstawę wyznaczania niepewności pomiaru matrycy mikrofonowej:

$$u_{c_n} = \sqrt{n * u_c^2 + u^2(b_{Mic-field})}, \quad (8.4)$$

gdzie: u_{c_n} – niepewność złożona n -mikrofonowej matrycy,

n – liczba mikrofonów stanowiących matrycę,

u_c – niepewność złożona pomiaru za pomocą pojedynczego mikrofonu,

u $b_{Mic-field}$ - zakłócenia pola akustycznego spowodowane obecnością matrycy mikrofonowej.

Należy podkreślić, że przeprowadzona analiza jest uproszczona, ponieważ w przypadku matryc mikrofonowych, a więc również mikrofonów ambisonicznych w grę wchodzi niepewność związana z przesunięciem fazowym mikrofonów. Ponadto w przypadku mikrofonów MEMS stosowane są układy elektroniczne korygujące ich charakterystykę częstotliwościową

8.4.1. Przykład oszacowania błędu pomiarowego dla mikrofonu Sennheiser Ambeo i Zylia ZM1-3E

Do oszacowania niepewności pomiaru z użyciem matrycy mikrofonowej przyjęto dwa przypadki, zakładając, że:

1. Matryca składa się z n mikrofonów pomiarowych. W tym przypadku przyjęto dla każdego z mikrofonów niepewność złożoną $u_c = 0,6$ dB, wynikającą ze wzoru (8.2), mniejszą niż podana w tabeli 8.1 ze względu na pominięcie czynnika wpływu obecności mikrofonu na pole akustyczne.
2. Matryca składa się z n mikrofonów nie mających dokładności mikrofonów pomiarowych. W tym przypadku przyjęto dla każdego z mikrofonów niepewność złożoną

$u_c = 1,1$ dB. Jest to wartość wyznaczoną ze wzoru (8.2), ale po przyjęciu składowych niepewności związanych z mikrofonem dwukrotnie większych niż podane w tabeli 8.1.

Tabela 8.2. Oszacowanie niepewności wprowadzanej przez matrycę mikrofonową

Liczba mikrofonów stanowiących matrycę	Niepewność złożona matrycy, dB	
	$u_c = 0,6$ dB	$u_c = 1,1$ dB
4	1,2	2,3
19	2,7	5,0

Z przedstawionych w tabeli 8.2 szacunkowych obliczeń wynika, że w przypadku matryc mikrofonowych składających się wyłącznie z mikrofonów pomiarowych, niepewność pomiarowa wyniosłaby 1,2 dB dla mikrofonu ambisonicznego 1-go rzędu i 2,3 dB dla matrycy mikrofonowej składającej się z 19 przetworników. Zakładając, że pojedyncze mikrofony charakteryzowałyby się niepewnością pomiaru dwukrotnie większą – niepewność złożona wyniosłaby odpowiednio 2,3 i 5,0 dB.

9. Podsumowanie i dyskusja

Zgodnie z celem postawionym w pracy, elementy dyskusji przede wszystkim odnoszą się do różnic wartości parametrów akustycznych otrzymanych za pomocą obliczeń wynikających z pomiarów mikrofonami ambisonicznymi i mikrofonem laboratoryjnym (w przypadku parametrów uśrednianych przestrzennie T30, C50, C80 i Ts) oraz manekinem akustycznym (w przypadku parametru przestrzennego IACC).

Różnice w wynikach pomiarów czasu pogłosu T30 nie przekroczyły $\pm 4\%$, co stanowi wartość mniejszą od progowej różnicy postrzegania zmiany czasu pogłosu JND, przyjmowanej jako 5% zmiana (por. tabelę 2.1). W przypadku pomieszczeń o skrajnych warunkach pogłosowych, występują nieznaczne przekroczenia w granicach 8-9%. Największą różnicę zarejestrowano w przypadku mikrofonu Zylia ZM1-3E, w komorze bezechowej i wynosiła ona 11,3%. W szczególności w tym przypadku różnicę tę można powiązać z bardzo małymi referencyjnymi wartościami T30 występującymi w komorze bezechowej.

W przypadku pomiarów czasu wczesnego zaniku EDT, różnice względem mikrofonu pomiarowego obserwowane w pomieszczeniach użytkowych (studio dźwiękowe, korytarz, sala koncertowa) w większości przypadków mieszczą się w granicach $\pm 5\%$, czyli JND. Nieliczne przekroczenia tej wartości występują poniżej częstotliwości 1 kHz, jednak nie przekraczają 10%, czyli dwukrotności JND. Różnice występujące w komorze bezechowej nie mogą być uznane za reprezentatywne, z uwagi na bardzo małe wartości EDT mierzone w tym pomieszczeniu, które np. w paśmie 8 kHz wynoszą 0,3 ms.

Różnice w pomiarach mikrofonami ambisonicznymi i mikrofonem pomiarowym uzyskane dla czytelności dźwięku C50 i przejrzystości dźwięku C80 są zbliżone, co wiąże się z definicjami obydwu parametrów różniącymi się jedynie przedziałem całkowania wczesnej energii odpowiedzi impulsowej. Pomimo, że zauważalny jest wzrost wartości zmierzonych mikrofonem Zylia ZM1-3E w paśmie 4 kHz (co może być bezpośrednio związane z częstotliwością aliasingu wynoszącą dla tego mikrofonu 3,3 kHz), odstępstwa przy pomiarze mikrofonami ambisonicznymi nie przekroczyły wartości JND równej 1 dB prawie we wszystkich przypadkach (z wyjątkiem pomiarów mikrofonem Sennheiser Ambeo w studio dźwiękowym przy 125 Hz).

Różnice w pomiarach czasu środkowego Ts nie przekraczały 10 ms przyjmowanych jako JND. Znaczne różnice zaobserwowano w przypadku pomiarów w komorze pogłosowej, które sięgały dwukrotności JND w przypadku pomiarów mikrofonem Sennheiser Ambeo w paśmie 63 Hz oraz trzykrotności JND w przypadku mikrofonu Zylia ZM1-3E i pasma częstotliwości

8 kHz. Te różnice można traktować jako pomijalne, z uwagi na duże wartości czasu środkowego T_s zmierzone w tym pomieszczeniu. W przypadku pomiaru w komorze bezchowej, różnice te również można uznać za nieistotne, ponieważ czas środkowy T_s jest zbliżony do zera.

Wymaga komentarza brak znaczących różnic w pomiarach mikrofonami Sennheiser Ambeo i Zylia ZM1-3E występujący pomimo znacznych rozmiarów (średnica 10 cm) sfery stanowiącej powierzchnię położenia mikrofonów, a jednocześnie będącej przeszkodą w polu akustycznym. Mikrofon tych rozmiarów powodowałby już znaczące wzrosty mierzonego ciśnienia akustycznego przy częstotliwości 2,5 kHz. Brak tego efektu w przypadku mikrofonu Zylia w pomiarach z użyciem kanału W wynika zapewne z uśredniania, w którym ma miejsce kompensacja pomiaru po stronie frontowej i zacienionej tego mikrofonu, w stosunku do padającej fali akustycznej, Z tego powodu nie ma istotnych różnic między pomiarem w polu fali swobodnej w komorze bezchowej i w polu dyfuzyjnym komory pogłosowej.

Porównanie z literaturą omówioną w rozdziale 5.1 zwraca uwagę, że praktycznie w żadnej z prac [16, 17, 20, 26, 46, 47, 74] problem oceny dokładności użycia kanału W mikrofonu ambisonicznego przez porównanie z pomiarowym mikrofonem dookólnym nie jest poruszany. W pracy Clapp i in [16, 17], autorzy jedynie postulują dalsze badania, które miałyby ocenić zgodność pomiarów mikrofonem ambisonicznym z konwencjonalnym mikrofonem wszechkierunkowym, a także z uwzględnieniem dwukierunkowego mikrofonu ósemkowego.

Z kolei, w pracy Pantan [70] przeprowadzono pomiar parametru sal związanym z estradą, który nie jest mierzony w obecnej pracy, tzw. Stage Support. Pomiar w pracy [70] przeprowadzono mikrofonem ambisonicznym i wszechkierunkowym, a podane różnice zasadniczo nie przekraczają 0,5 dB. Można powtórzyć, że brak jednoznacznych wyników w wymienionej literaturze wydaje się wskazywać na sensowność podjęcia tego tematu w niniejszej pracy.

W tabeli 9.1 przedstawiono zwięzłe podsumowanie różnic wartości IACC wyznaczanych za pomocą odpowiedzi impulsowej BRIR przy użyciu czterech metod upraszczania HRTF i wartości IACC otrzymanych manekinem B&K4100-D lub sztuczną głową Neumann KU-100 (pomiar w sali koncertowej). Tabela wyróżnia wyłącznie te czynniki rozumiane jako kombinacja rodzaju mikrofonu ambisonicznego, rozdzielczości HRTF, metody redukcji HRTF i pomieszczenia, w których występowały różnice większe niż 1 JND, biorąc pod uwagę wszystkie pasma częstotliwości.

Tabela 9.1. Syntetyczne ujęcie różnic wartości parametru IACC zmierzonego z użyciem manekina B&K 4100-D lub sztucznej głowy Neumann KU-100 (sala koncertowa) i wyznaczonego z użyciem binauralnej przestrzennej odpowiedzi impulsowej BRIR.

Rozdzielczość HRIR		1550 punktów		8802 punkty	
Mikrofon		Ambeo	Zyilia ZM1-3E	Ambeo	Zyilia
Metoda upraszczania HRTF w procesie binauralizacji	LS	4 JND, $f = 4$ kHz <i>K, S, KP</i>	4 JND (poza 1 kHz) <i>K, S, KP</i>	2 JND, $f = 4$ kHz <i>KB</i>	4 JND, $f < 1$ kHz <i>K, S, KP</i>
	MagLS	2 JND, $f = 4$ kHz <i>KB, K</i>	4 JND, $f < 1$ kHz <i>K, S, KP</i> 2JND, $f = 4$ kHz <i>KB, S,</i>	—	4 JND, $f < 1$ kHz <i>K, S, KP</i>
	SPR	4 JND, $f = 4$ kHz <i>K, S, KP</i>	3 JND, $f < 1$ kHz <i>KB, K, S, KP</i>	—	4 JND, $f < 1$ kHz <i>KB, K, S, KP</i>
	TA	2 JND, $f = 4$ kHz <i>K, KP</i>	4 JND, $f < 1$ kHz <i>KB, K, S, KP</i> 2 JND, $f = 4$ kHz <i>KB, K, S, KP</i>	2 JND, $f = 500$ Hz <i>K, KP</i>	4 JND, $f < 1$ kHz <i>KB, K, S, KP</i>
Oznaczenia: <i>KB</i> – komora bezechowa, <i>K</i> – korytarz, <i>S</i> – studio dźwiękowe, <i>SK</i> – sala koncertowa, <i>KP</i> – komora pogłosowa					

Takie uproszczone uporządkowanie danych ułatwia wyciągnięcie jakichkolwiek syntetycznych wniosków. Na 72 pomiary, w przypadku 31 (nie licząc punktów częstotliwości) występowały różnice przekraczające 4 JND. Zauważyć jednak należy, że te przekroczenia dotyczyły pojedynczych określonych częstotliwości, a nie całości wyników.

W przypadku mikrofonu Sennheiser Ambeo i rozdzielczości HRTF 1550 punktów różnice sięgające 4 JND przy 4 kHz, jedynie w niektórych salach, które są wymienione w tabeli. Przy tej samej rozdzielczości HRTF, dla mikrofonu Zyilia ZM1-3E różnice w granicach 3-4 JND mają miejsce poniżej częstotliwości 1 kHz.

W przypadku mikrofonu Sennheiser Ambeo i rozdzielczości HRTF 8802 punktów różnica 2 JND ma miejsce przy 500 Hz i 4 kHz, jedynie w trzech podanych w tabeli salach. Natomiast w przypadku mikrofonu Zyilia ZM1-3E różnice sięgające 4 JND występują zawsze przy częstotliwościach mniejszych niż 1 kHz.

W podsumowaniu można stwierdzić, że niezależnie od gęstości HRTF, większe różnice wartości IACC otrzymanych z wyznaczanych BRIR w stosunku do wartości zmierzonych z użyciem manekina pomiarowego obserwowane są w odniesieniu do mikrofonu Zyilia ZM1-3E. Ponadto, w przypadku tego mikrofonu większe różnice występują przy wyższej (8802 punkty) rozdzielczości HRTF, a w przypadku mikrofonu Sennheiser Ambeo przy niższej (1550

punktów) rozdzielczości HRTF. Informacje zebrane w tabeli 9.1 korespondują z istotnością różnic na poziomie $p = 0,05$ przedstawionych w tabeli Z3.1 do Z3.3 w załączniku 3.

Podjmując próbę określenia najlepszej metody redukcji HRTF, można orientacyjnie przyjąć, że w przypadku mikrofonu Sennheiser Ambeo i rozdzielczości 1550 punktów, są to metody MagLS i TA, a przy rozdzielczości 8802 punktów, MagLS i SPR, przy przewadze rozwiązania z użyciem HRTF o rozdzielczości 8802 punktów. W odniesieniu do mikrofonu Zylia ZM1-3E wydaje się, że najlepszym wyborem jest metoda SPR i zastosowaniu rozdzielczości HRTF 1550 punktów. Przy rozdzielczości 8802 wszystkie metody upraszczania HRTF dały w przypadku tego mikrofonu równoważne wyniki.

Porównując uzyskane wyniki odnoszące się do IACC znane z literatury, można odnieść się do nielicznych prac, co było już wspomniane, ponieważ w większości prace te były poświęcone jakości obrazu przestrzennego uzyskiwanego przez tzw. *binaural rendering* (konstrukcję dwuosobnego przestrzennego obrazu dźwiękowego). Interesowano się odwzorowaniem określonych cech obrazu dźwiękowego, barwy dźwięku, ciągłością obrazu dźwiękowego i rozmieszczeniem obiektów dźwiękowych w przestrzeni badanymi na podstawie testów odsłuchowych [109]. Pörschmann [75] pisze o ocenie korelacji funkcji IACC wyznaczonych na podstawie odpowiedzi impulsowej BRIR oraz pomiarem z użyciem manekina wynoszącej od 0,6 do 1 w odniesieniu do wczesnych odbić dźwięku w sali i 0,1 do 0,6 w odniesieniu do części pogłosowej. Menzer [55] stwierdza dobrą korelację funkcji IACC (wyznaczonej i zmierzonej) w przedziale do 4 kHz i większe różnice powyżej, z użyciem własnej, proponowanej metody redukcji wymiarów HRTF.

W pracy Engela i in. [23] przeprowadzono szczegółową analizę przy ograniczeniu harmonik sferycznych reprezentujących HRTF do rzędu $N = 3$, co ma bezpośrednie odniesienie do danych uzyskanych mikrofonem trzeciego rzędu Zylia ZM1-3E. Engel podaje, że największy błąd oszacowania ILD, rzędu 2,5 do 3 dB ([23], rys. 5) występuje dla metody upraszczania HRTF SPR i LS, a dla metod TA i MagLS mniejszy, w granicach 0,5 do 1 dB. Analiza ITD w tym samym artykule wskazuje na największy błąd w procedurze MagLS (rzędu 0,3 ms), dla metod LS i SPR (od 0,1 do 0,2 ms), a dla metody TA jedynie 20 μ s.

Nie jest to w pełni zgodne z sumarycznym oszacowaniem wyników w obecnej pracy zrobionym na podstawie tabeli 9.1, ponieważ wskazuje ona, że metoda SPR przy niskiej rozdzielczości HRTF równej 1550 punktów pozwalała na uzyskanie najlepszych wyników w przypadku mikrofonu trzeciego rzędu Zylia ZM1-3E. Przy rozdzielczości 8802 żadna z metod nie warunkowała wyraźnie lepszych rezultatów. Rozbieżność może wynikać z tego, że w pracy Engela i in. [23] rozdzielczość stosowanego HRTF wynosiła prawie 12 tys. punktów.

10. Wnioski

Przedmiotem pracy było zastosowanie odpowiedzi impulsowej pomieszczenia rejestrowanej mikrofonami ambisonicznymi pierwszego i trzeciego rzędu do wyznaczenia parametrów akustycznych pomieszczenia. Ze względu na specyfikę mikrofonów ambisonicznych przeznaczonych do rejestracji przestrzennych właściwości obrazu dźwiękowego, najważniejszym elementem pracy były badania związane z wyznaczeniem międzyusznej korelacji skrośnej IACC, których wyniki przedstawiono w punkcie 8.3. Dla wyznaczania parametru IACC wykorzystuje się binauralną przestrzenną odpowiedź impulsowa sali BRIR pochodzącą połączenia jednopunktowej przestrzennej odpowiedzi impulsowej pomieszczenia SRIR z funkcjami przenoszenia głowy HRTF/HRIR. W związku z tym, w realizacji pracy wyznaczono odpowiedź impulsową SRIR z wykorzystaniem mikrofonu ambisonicznego Sennheiser Ambeo i Zylia ZM1-3E, a binauralną odpowiedź impulsową BRIR wyznaczono na podstawie splotu z funkcją HRTF manekinów akustycznych. Równolegle, z odpowiedzi impulsowej SRIR otrzymanej z użyciem wymienionych mikrofonów ambisonicznych dokonano wyznaczenia dookólnych parametrów sali: T30, EDT, C50, C80 i Ts, których wyniki przedstawiono w punkcie 8.2.

Badania przeprowadzono w pięciu salach o różnej pogłosowości, z czego komora bezechowa i komora pogłosowa należą do pomieszczeń laboratoryjnych skrajnych pod względem warunków akustycznych, gwarantujących pole fali swobodnej i pole rozproszone definiowane zgodnie z wymaganiami norm. Pozostałe trzy sale użytkowe reprezentowały różne właściwości akustyczne pod względem czasu pogłosu, z czego dwie sale (studio dźwiękowe i sala koncertowa) miały kwalifikowane warunki akustyczne.

Przeprowadzono ocenę dokładności wartości parametrów akustycznych sal wyznaczonych na podstawie przestrzennej odpowiedzi impulsowej SRIR i binauralnej odpowiedzi impulsowej BRIR poprzez odniesienie tych wartości do wartości zmierzonych z użyciem mikrofonu pomiarowego, manekina pomiarowego B&K4100-D i sztucznej głowy KU-100, która wykazała, że:

1. W odniesieniu do parametrów dookólnych mierzonych mikrofonem ambisonicznym kanałem W, różnice w stosunku do mikrofonu pomiarowego nie przekraczają w przypadku:
 - czasu pogłosu T30 i pomieszczeń użytkowych, 5 % (JND), a pomieszczeń laboratoryjnych – komory bezechowej i pogłosowej – w których występują skrajne warunki pola akustycznego 11%,
 - czasu wczesnego zaniku EDT, w pomieszczeniach użytkowych 8%,

- wskaźników C50 i C80 różnice nie przekraczają 1,5 dB, a jedynie w przypadku mikrofonu Zylia ZM1-3E w przy częstotliwości 8 kHz 2 dB, co może być związane z częstotliwością aliasingu 3,3 kHz,
 - wskaźnika Ts, w pomieszczeniach użytkowych 8 ms, w przypadku komory pogłosowej 28 ms.
2. W odniesieniu do międzyusznej korelacji skrośnej IACC, wyznaczanej na podstawie binauralnej odpowiedzi impulsowej pomieszczenia BRIR, różnice w stosunku do manekina akustycznego są najmniejsze dla mikrofonu:
- Sennheiser Ambeo, przy uproszczeniu rzędu HRTF o rozdzielczości 8802 punktów metodami Magnitude Least Squares (MagLS) lub Spatial Resampling (SPR),
 - Zylia ZM1-3E, przy uproszczeniu rzędu HRTF o rozdzielczości 1550 punktów metodą Spatial Resampling (SPR).

W świetle uzyskanych wyników wydaje się właściwym stwierdzenie, że wykazana została teza pracy zamieszczona na stronie 38, stwierdzająca, że *możliwe jest uzyskanie parametrów akustycznych pomieszczenia przy uwzględnieniu binauralnych parametrów przestrzennych, za pomocą sferycznej matrycy mikrofonowej przy uwzględnieniu funkcji HRTF.*

Nowością uzyskanych wyników jest wykazanie, że możliwe jest wykorzystanie matrycy mikrofonowej, czyli jednego przetwornika, do przeprowadzenia pomiarów różnego rodzaju parametrów akustycznych pomieszczenia w miejsce pomiarów z zastosowaniem mikrofonu pomiarowego dla grupy parametrów wszechkierunkowych i pomiaru dwuuszego z użyciem manekina akustycznego dla parametrów przestrzennych, jak to zaleca norma PN-EN ISO 3382-1.

W związku z powyższym należy uznać, że stosowanie mikrofonu ambisonicznego w celu uproszczenia pomiarów parametrów akustycznych pomieszczeń jest możliwe. Daje zastosowanie jednego urządzenia rejestrującego odpowiedź impulsową sali do wyznaczenia parametrów dookólnych i przestrzennych, w pierwszym przypadku z wykorzystaniem kanału W, a w drugim charakterystyki mikrofonu związanej z rzędem mikrofonu ambisonicznego, przynajmniej w zakresie rzędu pierwszego i trzeciego. Zastosowanie tej techniki umożliwia uniknięcie stosowania manekinu pomiarowego, a także umożliwia przyjęcie dowolnej orientacji ukierunkowania na źródło dźwięku bez konieczności powtarzania pomiaru.

Ponadto wyniki pracy pozwalają na następujące spostrzeżenia:

1. Wybór metody uproszczenia HRTF jest złożony, ponieważ zależy od takich czynników, jak rodzaj i rząd mikrofonu oraz wstępna rozdzielczość przestrzenna funkcji HRTF.
2. Wzrost rozdzielczości HRTF nie zawsze prowadzi do uzyskania wyników o wartościach bliższych do uzyskanych z użyciem mikrofonu i manekina pomiarowego.
3. Wykazano, że funkcje IACC mierzone manekinem B&K4100-D i KU-100 praktycznie nie różnią się od siebie, co jest ważne ze względu na fakt, że głowa KU-100 przeznaczona do nagrań, nie ma standardu urządzenia pomiarowego, a jest często używana w badaniach związanych z dźwiękiem przestrzennym rejestrowanym mikrofonami ambisonicznym.
4. Uproszczona próba oszacowania niepewności pomiaru mikrofonów ambisonicznych Sennheiser Ambeo i Zylia ZM1-3E wskazuje, że w odniesieniu do rejestrowanych amplitud odbić w sali, dokładność nie jest mniejsza niż 2,7 dB.

Wyniki i doświadczenia z realizacji pracy wskazują na dalsze kierunki badań, które powinny obejmować tak badania mikrofonów ambisonicznych rzędu wyższego niż trzy w pomiarach parametrów akustycznych sal, jak i wykorzystywania w tym celu operacji opartych na harmonikach sferycznych (ang. *upscaling*). Wskazana jest także systematyczna ocena uproszczenia rozdzielczości HRTF przy użyciu harmonik sferycznych różnych rzędów. Ponadto wydaje się istotne wypracowanie techniki dokładnej kalibracji mikrofonów ambisonicznych pod kątem pomiarów akustycznych, co mieści się w obecnie coraz bardziej istotnych kwestiach kalibracji nowoczesnych technologicznie mikrofonów typu MEMS.

Bibliografia

- [1] "Zylia Ambisonics Converter." <https://www.zylia.co/zylia-ambisonics-converter.html> (dostęp 19.06.2023).
- [2] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The cipc hrtf database," w *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*, 2001: IEEE, str 99-102.
- [3] Armstrong, C., Thresh, L., Murphy, D., and Kearney, G., "A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database," *Applied Sciences*, vol. 8, nr 11, s 2029, 2018.
- [4] Barron, M., *Auditorium acoustics and architectural design*. Routledge, 2009.
- [5] Ben-Hur, Z., Alon, D. L., Mehra, R., and Rafaely, B., "Efficient representation and sparse sampling of head-related transfer functions using phase-correction based on ear alignment," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, nr 12, str 2249-2262, 2019.
- [6] Ben-Hur, Z., Brinkmann, F., Sheaffer, J., Weinzierl, S., and Rafaely, B., "Spectral equalization in binaural signals represented by order-truncated spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 141, nr 6, str 4087-4096, 2017.
- [7] Beranek, L. L., *Concert halls and opera houses: music, acoustics, and architecture*. Springer, 2004.
- [8] Beranek, L. L., "Listener envelopment lev, strength g and reverberation time rt in concert halls," w *Proceedings of 20th international congress on acoustics, ica*, 2010.
- [9] Bernschütz, B., Giner, A. V., Pörschmann, C., and Arend, J., "Binaural reproduction of plane waves with reduced modal order," *Acta Acustica united with Acustica*, vol. 100, nr 5, str 972-983, 2014.
- [10] Bertet, S., Daniel, J., Parizet, E., and Warusfel, O., "Investigation on localisation accuracy for first and higher order ambisonics reproduced sound sources," *Acta Acustica united with Acustica*, vol. 99, nr 4, str 642-657, 2013.
- [11] Brandstein, M. and Ward, D., *Microphone arrays: signal processing techniques and applications*. Springer Science & Business Media, 2001.
- [12] Brinkmann, F. *et al.*, "Recent Advances in an Open Software for Numerical HRTF Calculation," *Journal Audio Engineering Society*, vol. (w druku), 2023.
- [13] Brinkmann, F. and Weinzierl, S., "Comparison of head-related transfer functions pre-processing techniques for spherical harmonics decomposition," w *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, 2018: Audio Engineering Society.
- [14] Bruel&Kjaer. "Specyfikacja techniczna manekina akustycznego Bruel&Kjaer typ 4100-D." <https://www.bksv.com/media/doc/bp1436.pdf> (dostęp 19.06.2023)
- [15] Capon, J., "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, nr 8, str 1408-1418, 1969.
- [16] Clapp, S., Guthrie, A., Braasch, J., and Xiang, N., "Investigating room acoustics using a 16-channel, 2nd-order ambisonic microphone," w *20th International Congress on Acoustics. International Congress on Acoustics*, 2010.

- [17] Clapp, S., Guthrie, A., Braasch, J., and Xiang, N., "Three-dimensional spatial analysis of concert and recital halls with a spherical microphone array," w *Proceedings of Meetings on Acoustics ICA2013*, 2013, vol. 19, nr 1: Acoustical Society of America, s 015020.
- [18] Courville, D. "Native B-Format Recording." https://courville.uqam.ca/ambisonic/native_b.html (dostęp 19.06.2023).
- [19] Cristensen, J., "Beamforming," *Bruel & Kjaer Technical Review*, vol. 1, str 1-48, 2004.
- [20] Czopek, D., Małecki, P., Piechowicz, J., and Wiciak, J., "Soundscape analysis of selected landforms on Spitsbergen," *Archives of Acoustics*, vol. 44, nr 3, str 511-519, 2019.
- [21] Deppisch, T., Ahrens, J., Garí, S. V. A., and Calamia, P., "Spatial Subtraction of Reflections from Room Impulse Responses Measured with a Spherical Microphone Array," w *2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2021: IEEE, str 346-350.
- [22] Dobrucki, A. B., Plaskota, P., and Pruchnicki, P., "System for High Speed Measurement of Head-Related Transfer Function," *Advanced Topics in Measurements*, s 1, 2012.
- [23] Engel, I., Goodman, D. F., and Picinali, L., "Assessing HRTF preprocessing methods for Ambisonics rendering through perceptual models," *Acta Acustica*, vol. 6, s 4, 2022.
- [24] Farina, A., "Simultaneous measurement of impulse response and distortion with a swept-sine technique," w *Audio engineering society convention 108*, 2000: Audio Engineering Society.
- [25] Farina, A., "Advancements in impulse response measurements by sine sweeps," w *Audio engineering society convention 122*, 2007: Audio Engineering Society.
- [26] Farina, A. and Ayalon, R., "Recording concert hall acoustics for posterity," w *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, 2003: Audio Engineering Society.
- [27] Farina, A., Capra, A., Chiesi, L., and Scopece, L., "A spherical microphone array for synthesizing virtual directive microphones in live broadcasting and in post production," w *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, 2010: Audio Engineering Society.
- [28] Gerzon, M. A., "Periphony: With-height sound reproduction," *Journal of the audio engineering society*, vol. 21, nr 1, str 2-10, 1973.
- [29] Gerzon, M. A., "The design of precisely coincident microphone arrays for stereo and surround sound," w *Audio Engineering Society Convention 50*, 1975: Audio Engineering Society.
- [30] Gerzon, M. A., "Recording concert hall acoustics for posterity," *Journal of the Audio Engineering Society*, vol. 23, nr 7, str 569-571, 1975.
- [31] Hammershøi, D. and Møller, H., "Sound transmission to and within the human ear canal," *The Journal of the Acoustical Society of America*, vol. 100, nr 1, str 408-427, 1996.
- [32] Hidaka, T., Okano, T., and Beranek, L., "Interaural cross correlation (IACC) as a measure of spaciousness and envelopment in concert halls," *The Journal of the Acoustical Society of America*, vol. 92, nr 4, str 2469-2469, 1992.
- [33] Hold, C., Gamper, H., Pulkki, V., Raghuvanshi, N., and Tashev, I. J., "Improving binaural ambisonics decoding by spherical harmonics domain tapering and coloration compensation," w *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019: IEEE, str 261-265.

- [34] Jarrett, D. P., Habets, E. A., and Naylor, P. A., "3D source localization in the spherical harmonic domain using a pseudointensity vector," w *2010 18th European Signal Processing Conference*, 2010: IEEE, str 442-446.
- [35] Jasiński, M. and Žera, J., "Variability of Head Related Transfer Functions across subjects," w *Proceedings of the 23rd International Congress on Acoustics: integrating 4th EAA Euroregio*, Aachen, Germany, M. Ochmann, M. Vorländer, and J. Fels, Eds., 2019, vol. 2019, str 3056-3062, doi: 10.18154/RWTH-CONV-239181.
- [36] Jasiński, M. and Žera, J., "Estimation of Asymmetry in Head Related Transfer Functions," *International Journal of Electronics and Telecommunications*, vol. 68, 2022.
- [37] Jordan, V. L., "Acoustical Criteria for Auditoriums and Their Relation to Model Techniques," *The Journal of the Acoustical Society of America*, vol. 47, nr 2A, str 408-412, 2005, doi: 10.1121/1.1911535.
- [38] Khaykin, D. and Rafaely, B., "Acoustic analysis by spherical microphone array processing of room impulse responses," *The Journal of the Acoustical Society of America*, vol. 132, nr 1, str 261-270, 2012.
- [39] Kulowski, A., *Akustyka sal*. Politechnika Gdańska, 2007.
- [40] Kuttruff, H., *Room acoustics*. Crc Press, 2016.
- [41] Lehman, P., "Über die Ermittlung raumakustischer Kriterien und deren Zusammenhang mit subjektiven Beurteilungen der Hörsamkeit" PhD Thesis, Technical University of Berlin, 1976.
- [42] Li, S. and Peissig, J., "Measurement of head-related transfer functions: A review," *Applied Sciences*, vol. 10, nr 14, s 5014, 2020.
- [43] Lundeby, A., Vigran, T. E., Bietz, H., and Vorländer, M., "Uncertainties of measurements in room acoustics," *Acta Acustica united with Acustica*, vol. 81, nr 4, str 344-355, 1995.
- [44] Majdak, P., Balazs, P., and Laback, B., "Multiple exponential sweep method for fast measurement of head-related transfer functions," *Journal of the Audio Engineering Society*, vol. 55, nr 7/8, str 623-637, 2007.
- [45] Malham, D., "Higher order ambisonic systems. space in music-music in space," Master, University of York, 2003.
- [46] Małecki, P., Czopek, D., Piechowicz, J., and Wiciak, J., "Acoustic analysis of the glacier caves in Svalbard," *Applied Acoustics*, vol. 165, s 107300, 2020.
- [47] Małecki, P., Wiciak, J., and Nowak, D., "Acoustics of the Orthodox Churches in Poland," *Archives of acoustics*, vol. 42, nr 4, str 579–590, 2017.
- [48] Masiero, B., Pollow, M., and Fels, J., "Design of a fast broadband individual head-related transfer function measurement system," 2011.
- [49] Masterson, C., "Binaural Impulse Response Rendering for Immersive Audio," Trinity College Dublin, 2011.
- [50] McCormack, L. and Politis, A., "SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods," w *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, 2019: Audio Engineering Society.
- [51] McCormack, L., Politis, A., Scheuregger, O., and Pulkki, V., *Higher-order processing of spatial impulse responses*. Universitätsbibliothek der RWTH Aachen, 2019.

- [52] McCormack, L., Pulkki, V., Politis, A., Scheuregger, O., and Marschall, M., "Higher-order spatial impulse response rendering: Investigating the perceived effects of spherical order, dedicated diffuse rendering, and frequency resolution," *Journal of the Audio Engineering Society*, vol. 68, nr 5, str 338-354, 2020.
- [53] McKenzie, T., McCormack, L., and Hold, C., "Dataset of spatial room impulse responses in a variable acoustics room for six degrees-of-freedom rendering and analysis," *arXiv preprint arXiv:2111.11882*, 2021.
- [54] Mckenzie, T., Murphy, D., and Kearney, G., "An evaluation of pre-processing techniques for virtual loudspeaker binaural ambisonic rendering," w *EAA Spatial Audio Signal Processing symposium*, 2019, str 149-154.
- [55] Menzer, F., Faller, C., and Lissek, H., "Obtaining binaural room impulse responses from b-format impulse responses using frequency-dependent coherence matching," *IEEE transactions on audio, speech, and language processing*, vol. 19, nr 2, str 396-405, 2010.
- [56] Merimaa, J. and Pulkki, V., "Spatial impulse response rendering I: Analysis and synthesis," *Journal of the Audio Engineering Society*, vol. 53, nr 12, str 1115-1127, 2005.
- [57] Meyer, E. and Thiele, R., "Raumakustische Untersuchungen in Zahlreichen Konzertsalen und Rundfunkstudios unter Anwendung neuerer Messverfahren," *Acustica*, vol. 6, s 431, 1956.
- [58] Meyer, J. and Agnello, T., "Spherical microphone array for spatial sound recording," w *Audio Engineering Society Convention 115*, 2003: Audio Engineering Society.
- [59] Meyer, J. and Elko, G., "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," w *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 2: IEEE, str II-1781-II-1784.
- [60] MH acoustics. "Specyfikacja techniczna mikrofonu Eigenmike em64." <https://mhacoustics.com/mh-team> (dostęp: 19.06.2023).
- [61] Miar, Główny Urząd., *Ewaluacja danych pomiarowych - Przewodnik wyrażania niepewności pomiaru*. 2008.
- [62] Mickiewicz, W., *Metrologia i przetwarzanie sygnałów w obrazowaniu wektorowego pola akustycznego*. Wydawnictwo Uczelniane Zachodniopomorskiego Uniwersytetu Technologicznego, 2019.
- [63] Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B., "Head-related transfer functions of human subjects," *Journal of the Audio Engineering Society*, vol. 43, nr 5, str 300-321, 1995.
- [64] Müller, S. and Massarani, P., "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, nr 6, str 443-471, 2001.
- [65] Nachbar, C., Zotter, F., Deleflie, E., and Sontacchi, A., "Ambix-a suggested ambisonics format," w *Ambisonics Symposium*, 2011, vol. 2011.
- [66] Neumann. "Sztuczna głowa Neumann KU-100." <https://www.neumann.com/en-en/products/microphones/ku-100/> (dostęp. 19.06.2023)
- [67] Okano, T., Beranek, L. L., and Hidaka, T., "Relations among interaural cross-correlation coefficient (IACC E), lateral fraction (LF E), and apparent source width (ASW) in concert halls," *The Journal of the Acoustical Society of America*, vol. 104, nr 1, str 255-265, 1998.
- [68] Oppenheim, A. V., Willsky, A. S., Nawab, S. H., and Ding, J.-J., *Signals and systems*. Prentice hall Upper Saddle River, NJ, 1997.

- [69] Palm, T., Koch, S., Brinkmann, F., and Alexa, M., "Curvature-adaptive mesh grading for numerical approximation of head-related transfer functions," materiały konferencyjne DAGA, Wiedeń, 2021.
- [70] Panton, L., Cabrera, D., and Holloway, D., "Using a spherical microphone array for stage acoustics: A preliminary case for a new spatial parameter," w *22nd International Congress on Acoustics*, 2016, str 1-10.
- [71] Perez-Lopez, A. and Stefanakis, N., "Analysis of spherical isotropic noise fields with an A-Format tetrahedral microphone," *The Journal of the Acoustical Society of America*, vol. 146, nr 4, str EL329-EL334, 2019.
- [72] *PN-EN ISO 18233:2006 Akustyka -- Zastosowanie nowych metod pomiarowych w akustyce budynku i pomieszczeń*, PKN, 2006.
- [73] *PN-EN ISO 3382-1:2009 Akustyka -- Pomiar parametrów akustycznych pomieszczeń -- Część 1: Pomieszczenia specjalne*, PKN, 2009.
- [74] Polack, J.-D., Figueiredo, F. L., and Liu, S., "Statistical analysis of a set of Parisian Concert Halls and Theatres," w *Acoustics 2012*, 2012.
- [75] Pörschmann, C. and Wiefeling, S., "Perceptual aspects of dynamic binaural synthesis based on measured omnidirectional room impulse responses," w *International Conference on Spatial Audio*, 2015.
- [76] Pulkki, V. and Merimaa, J., "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," *Journal of the Audio Engineering Society*, vol. 54, nr 1/2, str 3-20, 2006.
- [77] Rafaely, B., *Fundamentals of spherical array processing*. Springer, 2018.
- [78] Rafaely, B. and Avni, A., "Interaural cross correlation in a sound field represented by spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 127, nr 2, str 823-828, 2010.
- [79] Rafaely, B., Weiss, B., and Bachmat, E., "Spatial aliasing in spherical microphone arrays," *IEEE Transactions on Signal Processing*, vol. 55, nr 3, str 1003-1010, 2007, doi: 10.1109/TSP.2006.888896.
- [80] Reichardt, W., Alim, O. A., and Schmidt, W., "Abhängigkeit der grenzen zwischen brauchbarer und unbrauchbarer durchsichtigkeit von der art des musikmotives, der nachhallzeit und der nachhalleinsatzzeit," *Applied Acoustics*, vol. 7, nr 4, str 243-264, 1974.
- [81] Rife, D. D. and Vanderkooy, J., "Transfer-function measurement with maximum-length sequences," *Journal of the Audio Engineering Society*, vol. 37, nr 6, str 419-444, 1989.
- [82] Rowland, I. D. and Howe, T. N., *Vitruvius: Ten books on architecture*. Cambridge University Press, 2001.
- [83] Sabine, W. C., "Reverberation," *The American Architect*, vol. 4, 1900.
- [84] Sabine, W. C., *Collected papers on acoustics*. Harvard University Press Cambridge, MA, 1927.
- [85] Schmidt, R., "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, nr 3, str 276-280, 1986.
- [86] Schönstein, D. and Katz, B. F., "HRTF selection for binaural synthesis from a database using morphological parameters," w *International Congress on Acoustics (ICA)*, 2010.
- [87] Schörkhuber, C., Zaunschirm, M., and Höldrich, R., "Binaural rendering of ambisonic signals via magnitude least squares," w *Proceedings of the DAGA*, 2018, vol. 44, str 339-342.

- [88] Sennheiser. "Specyfikacja techniczna mikrofonu Sennheiser Ambeo." <https://sennheiser.pl/o/ambeo-vr-mic> (dostęp: 19.06.2023).
- [89] Shaw, E. A., "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," *The Journal of the Acoustical Society of America*, vol. 56, nr 6, str 1848-1861, 1974.
- [90] Shaw, E. A. G., "The external ear," in *Handbook of Sensory Physiology*. Berlin: Springer, 1974, str. 455-485.
- [91] Shu-Nung, Y., Collins, T., and Liang, C., "Head-related transfer function selection using neural networks," *Archives of Acoustics*, vol. 42, nr 3, str 365–373, 2017.
- [92] Shukla, R., Stewart, R., and Sandler, M., "User HRTF Selection for 3D Auditory Mixed Reality," 2021.
- [93] SMARTTECH3D. "Broszura produktowa skanera SMARTTECH3D med." <https://skaner3d.pl/urzadzenia/skanery-3d/smarttech3d-med/> (dostęp. 19.06.2023)
- [94] Sunder, K. and Woszczyk, W., "Investigation of Impulse Response Recording Techniques in Binaural Rendering of Virtual Acoustics," w *Audio Engineering Society Convention 141*, 2016: Audio Engineering Society.
- [95] Team, T. M. H. "Mesh2HRTF." <https://www.mesh2hrtf.org/> (dostęp: 22.05.2023).
- [96] Tervo, S., Pätynen, J., Kuusinen, A., and Lokki, T., "Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, nr 1/2, str 17-28, 2013.
- [97] Voong, T. M., Reuter, C., and Oehler, M., "Influence of individual HRTF preference on localization accuracy—a comparison between regular and bone conducting headphones," w *Audio Engineering Society Convention 148*, 2020: Audio Engineering Society.
- [98] Vorländer, M., *Auralization*. Springer, 2020.
- [99] Vorländer, M. and Kob, M., "Practical aspects of MLS measurements in building acoustics," *Applied Acoustics*, vol. 52, nr 3-4, str 239-258, 1997.
- [100] Weyna, S., *Rozpyły energii akustycznych źródeł rzeczywistych*. Wydawnictwa Naukowo-Techniczne, 2005.
- [101] Wierzbicki, J., Małecki, P., and Wiciak, J., "Localization of the Sound Source with the Use of the First-order Ambisonic Microphone," *Acta Physica Polonica A*, vol. 123, nr 6, str 1114-1117, 2013.
- [102] Wightman, F. L. and Kistler, D. J., "Headphone simulation of free-field listening. I: stimulus synthesis," *The Journal of the Acoustical Society of America*, vol. 85, nr 2, str 858-867, 1989.
- [103] Wightman, F. L. and Kistler, D. J., "Headphone simulation of free-field listening. II: Psychophysical validation," *The Journal of the Acoustical Society of America*, vol. 85, nr 2, str 868-878, 1989.
- [104] Williams, E. G. and Mann III, J. A., "Fourier acoustics: sound radiation and nearfield acoustical holography," Wydanie ed: Acoustical Society of America, 2000.
- [105] Witew, I., "Measurements in room acoustics: Uncertainties and influence of the measurement position."
- [106] Xie, B., *Head-related transfer function and virtual auditory display*. J. Ross Publishing, 2013.
- [107] Xie, B., *Spatial Sound: Principles and Applications*. CRC Press, 2022.

- [108] Zaunschirm, M., Frank, M., and Zotter, F., "Binaural rendering with measured room responses: First-order ambisonic microphone vs. dummy head," *Applied Sciences*, vol. 10, nr 5, s 1631, 2020.
- [109] Zaunschirm, M., Schörkhuber, C., and Höldrich, R., "Binaural rendering of Ambisonic signals by head-related impulse response time alignment and a diffuseness constraint," *The Journal of the Acoustical Society of America*, vol. 143, nr 6, str 3616-3627, 2018.
- [110] Ziegelwanger, H., Kreuzer, W., and Majdak, P., "Mesh2hrtf: Open-source software package for the numerical calculation of head-related transfer functions," w *22nd International Congress on Sound and Vibration*, 2015.
- [111] Zotter, F. and Frank, M., *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature, 2019.
- [112] Zylia. "Specyfikacja mikrofonu Zylia ZM1-3E." <https://www.zylia.co/zylia-zm-1-microphone.html> (dostęp: 19.06.2023).

Załącznik nr 1: Definicje parametrów akustycznych pomieszczeń

Definicje parametrów akustycznych [32, 40, 73], które są w rozprawie wyznaczane na podstawie przestrzennych odpowiedzi impulsowych SRIR i BRIR:

Czas pogłosu – czas po którym poziom ciśnienia akustycznego w pomieszczeniu po wyłączeniu źródła dźwięku maleje o 60 dB. Czas pogłosu wyznacza się z krzywej zaniku dźwięku w pomieszczeniu. Czas pogłosu T30 – wyznaczany jest w zakresie spadku o 30 dB, przeliczany na czas zaniku dźwięku o 60 dB. Jest stosowany z uwagi na ograniczoną dynamikę aparatury pomiarowej.

Czas wczesnego zaniku EDT – czas zaniku dźwięku od 0 do -10 dB po wyłączeniu źródła dźwięku. Parametr ten jest silnie skorelowany z postrzeganą pogłosowością pomieszczenia.

Czytelność dźwięku C50 – stosunek energii w pierwszych $t_e = 50$ ms odpowiedzi impulsowej pomieszczenia do energii zawartej w pozostałej części odpowiedzi impulsowej. Obliczany jest zgodnie z zależnością (Z1.1). Parametr ten jest wrażeniowo skorelowany z postrzeganą czytelnością mowy w pomieszczeniu.

Przejrzystość dźwięku C80 – jest to stosunek energii w pierwszych $t_e = 80$ ms odpowiedzi impulsowej pomieszczenia do energii zawartej w pozostałej części odpowiedzi impulsowej. Obliczany jest zgodnie z zależnością (Z1.1). Parametr ten jest wrażeniowo skorelowany z postrzeganą przejrzystością muzyki w pomieszczeniu.

$$C_{t_e} = 10 \log \frac{\int_0^{t_e} p^2(t) dt}{\int_{t_e}^{\infty} p^2(t) dt}, dB \quad (Z1.1)$$

Czas środkowy echogramu T_s – środek ciężkości energii odpowiedzi impulsowej (echogramu) pomieszczenia. Obliczany jest zgodnie ze wzorem (Z1.2). Duże wartości czasu środkowego związane są z dużą pogłosowością pomieszczenia i mniejszą zrozumiałością mowy.

$$T_s = \frac{\int_0^{\infty} t p^2(t) dt}{\int_0^{\infty} p^2(t) dt}, ms \quad (Z1.2)$$

Współczynnik międzyusznej korelacji skrośnej IACC – znormalizowana korelacja skrośna przebiegu czasowego sygnałów odbieranych przez lewe i prawe ucho słuchacza. Jest związany z subiektywnym wrażeniem przestrzenności dźwięku. Wyznacza się go na podstawie funkcji korelacji międzyusznej IACF (Z1.3), jako jej wartość maksymalna według wzoru (Z1.4).

$$IACF_{t_1, t_2}(\tau) = \frac{\int_{t_1}^{t_2} p_L(t) \cdot p_R(t + \tau) dt}{\sqrt{\int_{t_1}^{t_2} p_L(t) dt \int_{t_1}^{t_2} p_R(t) dt}}, \quad (Z1.3)$$

gdzie, t_1 i t_2 - przedziały czasu całkowania odpowiedzi impulsowej,

p_L – ciśnienie u wejścia lewego przewodu słuchowego,

p_P – ciśnienie u wejścia prawego przewodu słuchowego.

$$IACC_{t_1, t_2} = \max(IACF_{t_1, t_2}), \text{ dla } -1 \text{ ms} < \tau < 1 \text{ ms}. \quad (Z1.4)$$

W najbardziej ogólnym przypadku, jako przedziały całkowania przyjmuje się $t_1 = 0$ i $t_2 = \infty$ s. Wyróżnić można także wariant IACC określający przestrzenność we wczesnej części odpowiedzi impulsowej ($t_1 = 0$ i $t_2 = 0.08$ s) i w późnej jej części ($t_1 = 0.08$ i $t_2 = \infty$).

Załącznik nr 2: Schematy punktów pomiarowych w pomieszczeniach

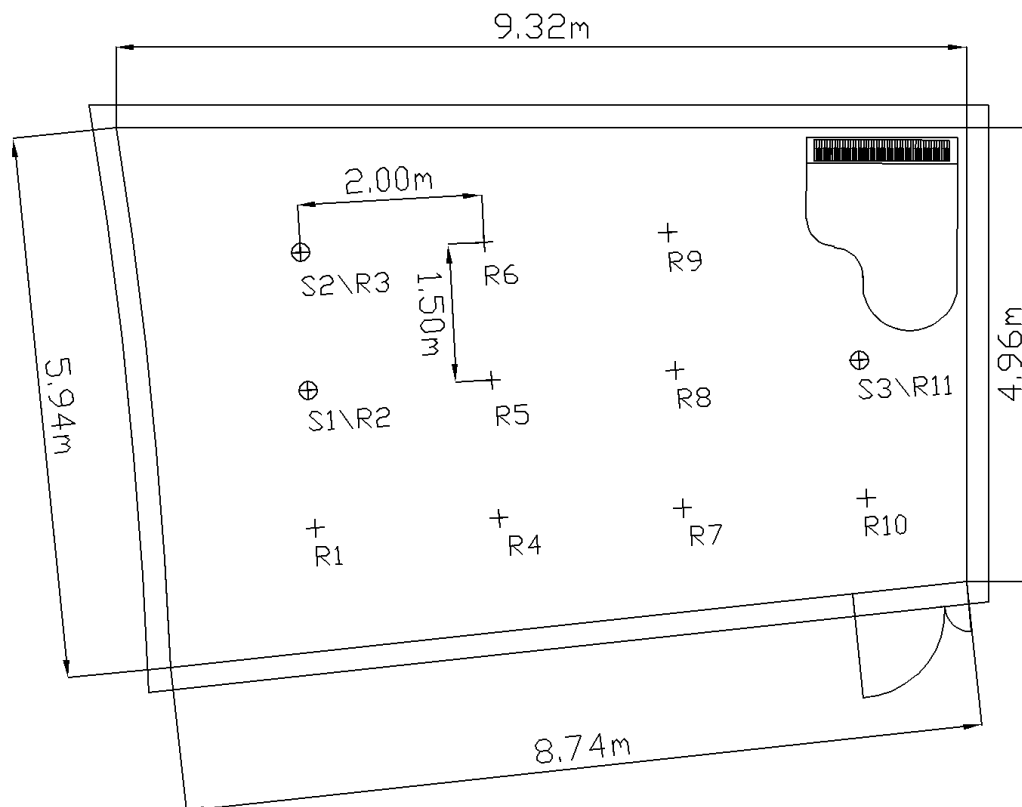
Schematy odnoszą się do rozmieszczenia źródeł dźwięku i mikrofonów pomiarowych na planach czterech pomieszczeń (studio dźwiękowego, korytarza, sali koncertowej i komory pogłosowej), których pomiary akustyczne opisano w pracy.

Komora bezechowa

W komorze bezechowej mikrofon pomiarowy i źródło dźwięku umieszczono w dwóch stałych punktach w odległości 6 metrów i w odległości nie mniejszej niż 2 m od ścian pomieszczenia. Pomiary stanowiły uśrednienie dwunastu pozycji kątowych wszechkierunkowego źródła dźwięku B&K 4292 wokół własnej osi.

Studio dźwiękowe

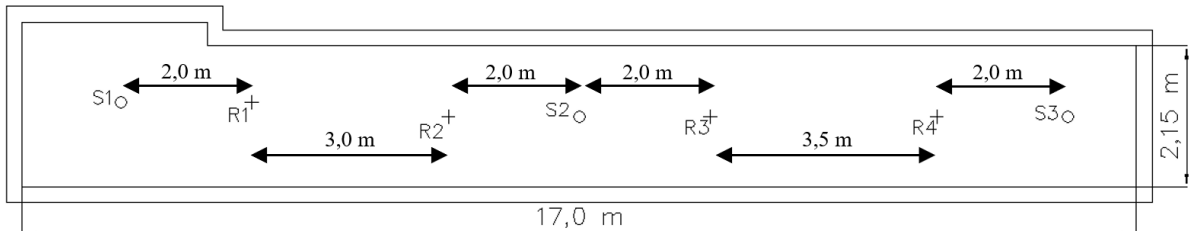
Pomiary w studio dźwiękowym Zakładu Elektroakustyki WEiTI PW były przeprowadzone przy trzech położeniach źródła dźwięku S1 do S3 przedstawionych na rys. Z2.1. Dla każdej pozycji źródła odpowiedzi impulsowe rejestrowano w dziesięciu punktach pomiarowych od R1 do R11. Punkty S1\R2 i S2\R3 oznaczają położenie źródła lub mikrofonu pomiarowego.



Rys. Z2.1. Rozmieszczenie punktów pomiarowych w studio dźwiękowym w Gmachu Elektroniki PW. S1, S2, S3 – położenie źródła dźwięku, R1-R11 – położenie mikrofonów pomiarowych

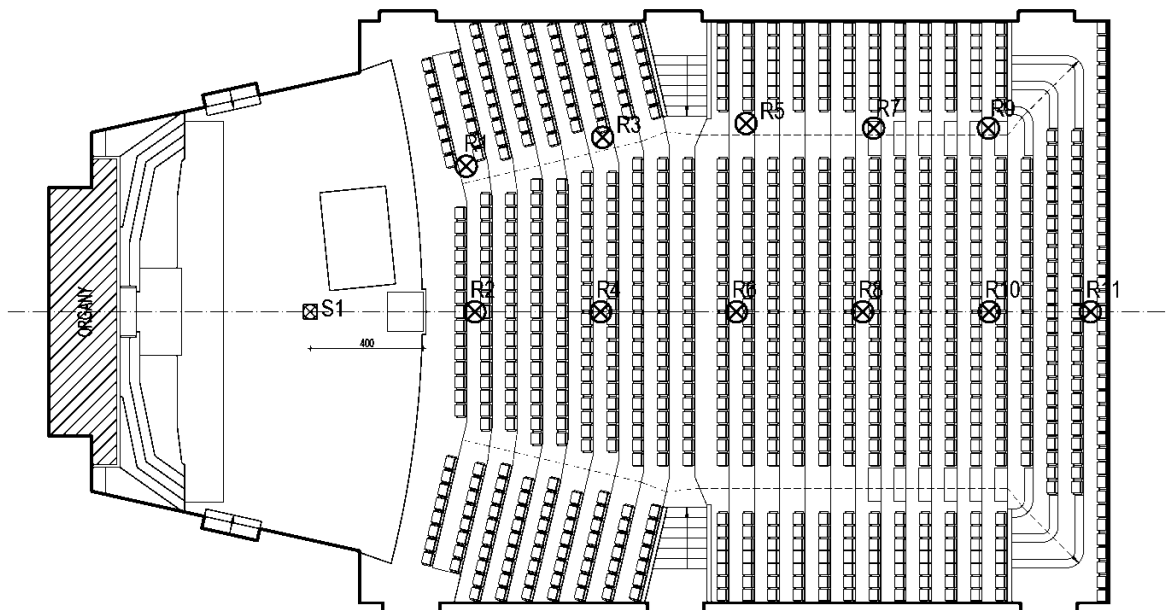
Korytarz w gmachu Elektroniki PW

Punkty pomiarowe w korytarzu Zakładu Elektroakustyki, WEiTl, PW przedstawiono na rysunku Z2.2. Odpowiedzi impulsowe rejestrowano w czterech punktach położenia mikrofonu pomiarowego oznaczonych jako R1 do R4 dla każdej z trzech pozycji ustawienia źródła dźwięku S1 do S3.



Rys. Z2.2. Rozmieszczenie punktów pomiarowych w korytarzu Zakładu Elektroakustyki, WEiTl, PW. S1, S2, S3 – położenie źródła, R1 – R4 – położenie mikrofonów pomiarowych

Sala koncertowa Filharmonii Pomorskiej im. Ignacego Paderewskiego w Bydgoszczy

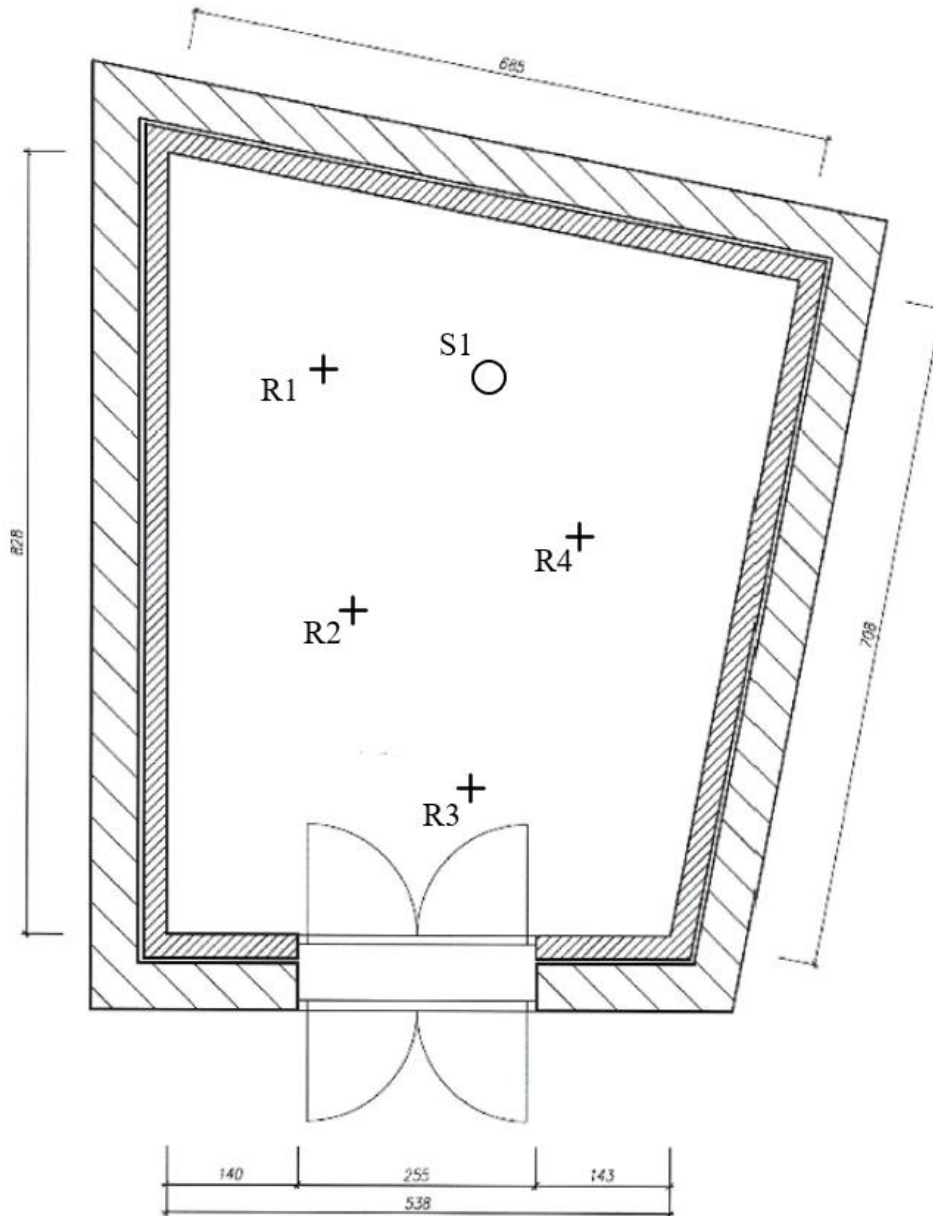


Rys. Z2.3. Rozmieszczenie punktów pomiarowych w sali koncertowej Filharmonii Pomorskiej im. Ignacego Paderewskiego w Bydgoszczy. S1 - położenie źródła dźwięku, R1-R11 – położenie mikrofonów pomiarowych

Rozmieszczenie punktów pomiarowych na planie sali koncertowej Filharmonii Pomorskiej im. Ignacego Paderewskiego w Bydgoszczy przedstawiono na rys. Z2.3. Pomiary przeprowadzono dla jednej lokalizacji, centralnie ustawionego źródła dźwięku S1 i jedenastu położzeń mikrofonu pomiarowego R1 – R11 w obszarze widowni.

Komora pogłosowa Instytutu Techniki Budowlanej

Pomiary w komorze pogłosowej Instytutu Techniki Budowlanej w Warszawie przeprowadzono przy jednym położeniu źródła dźwięku w punkcie S1 i czterech lokalizacjach mikrofonu pomiarowego R1 do R4, oddalonych od źródła o odpowiednio 1,5 m, 3,0 m, 3,5 m i 2 m, jak przedstawiono na rys. Z2.4.



Rys. Z2.4. Rozmieszczenie punktów pomiarowych w komorze pogłosowej Instytutu Techniki Budowlanej w Warszawie. S1 - położenie źródła dźwięku, R1-R4 – położenie mikrofonów pomiarowych.

Załącznik nr 3: Analiza wariancji danych

Załącznik przedstawia wyniki jednoczynnikowej analizy wariancji ANOVA przeprowadzonej z użyciem procedury `anova1` środowiska do MATLAB w wersji 2023b i wbudowanego testu post-hoc `multcompare`.

W tabeli Z3.1 podano pasma częstotliwości, w których wartości parametrów akustycznych sal wyznaczonych z przestrzennych odpowiedzi impulsowych SRIR różniły się od wartości zmierzonych z użyciem dookólnego mikrofonu pomiarowego B&K 4189 lub MK225 na poziomie istotności $p = 0,05$.

Tabela Z3.1. Pasma częstotliwości, w których rozkłady wyników parametrów wszechkierunkowych uzyskanych z wykorzystaniem mikrofonów Ambeo i Zylia istotnie różniły się od rozkładu tych parametrów uzyskanego przy użyciu mikrofonu pomiarowego.

	T30		EDT		C50		C80		Ts	
	Ambeo	Zylia	Ambeo	Zylia	Ambeo	Zylia	Ambeo	Zylia	Ambeo	Zylia
Komora Bezechowa	63, 125, 500, 4k	63, 125, 500, 8k	125, 250, 500,	125, 250, 500,	125, 250, 500,	63, 125, 250, 500, 8k	125, 250, 500	63, 125, 250, 500, 8k	63, 125, 250, 500,	125, 250, 500, 4k,
Studio dźwiękowe	2k,	1k, 8k	–	–	–	–	–	–	125	–
Korytarz	–	500, 8k	–	–	–	–	–	–	–	–
Studio	2k	1k, 4k, 8k	–	1k	–	–	–	–	125	–
Sala Koncertowa	4k, 8k	b/d	–	b/d	–	b/d	–	b/d	–	b/d
Komora Pogłosowa	500, 2k, 4k, 8k	500, 1k, 8k	–	–	–	–	–	–	–	–

W tabeli Z3.2 podano pasma częstotliwości, w których wartości współczynnika międzyuszyjnej korelacji skrośnej IACC są wyznaczonego z binauralnych odpowiedzi impulsowych BRIR przy uproszczeniu HRTF o rozdzielczości 1550 punktów różniły się od wartości zmierzonych z użyciem manekina pomiarowego B&K 4100-D lub sztucznej głowy Neumann KU-100 na poziomie istotności $p = 0,05$.

Tablica Z3.2. Pasma częstotliwości, w których wyniki parametru IACC wyznaczone na podstawie BRIR istotnie różniły się od wyników zmierzonych z użyciem referencyjnego manekina pomiarowego BK4100-D. Funkcja HRTF o rozdzielczości 1550 punktów

	LS		MagLS		SPR		TA	
	Ambeo	ZyLIA	Ambeo	ZyLIA	Ambeo	ZyLIA	Ambeo	ZyLIA
Komora Bezechowa	125, 4k, 8k	125, 250, 500, 1k, 4k	125	125, 250, 500, 1k, 4k, 8k	125, 500, 1k, 4k, 8k	125, 250, 500, 1k, 4k	125, 8k	125, 250, 500, 1k, 4k, 8k
Studio dźwiękowe	63, 1k, 4k, 8k	63, 125, 250, 500, 1k, 2k, 4k, 8k	63, 1k, 8k	63, 125, 250, 500, 1k, 4k	63, 4k, 8k	63, 125, 250, 500, 4k	63, 1k, 4k, 8k	63, 125, 250, 500, 1k, 4k
Korytarz	2k, 4k, 8k	63, 125, 250, 500, 1k, 2k, 4k, 8k	4k, 8k	63, 125, 250, 500, 1k, 8k	4k, 8k	63, 125, 250, 500	4k, 8k	63, 125, 250, 500, 1k, 8k
Studio	63, 1k, 4k, 8k	63, 125, 250, 500, 1k, 2k, 4k, 8k	63, 1k, 8k	63, 125, 250, 500, 1k, 4k	63, 4k, 8k	63, 125, 250, 500, 4k	63, 1k, 4k, 8k	63, 125, 250, 500, 1k, 4k
Sala Koncertowa	–	b/d	8k	b/d	–	b/d	–	b/d
Komora Pogłosowa	4k, 8k	63, 125, 250, 500, 2k, 4k, 8k	4k, 8k	63, 125, 250, 500,	500, 4k, 8k	63, 125, 250, 500,	4k, 8k	63, 125, 250, 500,

W tabeli Z3.3 podano pasma częstotliwości, w których wartości współczynnika międzyusznej korelacji skróśnej IACC są wyznaczonego z binauralnych odpowiedzi impulsowych BRIR przy uproszczeniu HRTF o rozdzielczości 8802 punktów różniły się od wartości zmierzonych z użyciem manekina pomiarowego B&K 4100-D lub sztucznej głowy Neumann KU-100 na poziomie istotności $p = 0,05$.

Tablica Z3.3. Pasma częstotliwości, w których wyniki parametru IACC wyznaczone na podstawie BRIR istotnie różniły się od wyników zmierzonych z użyciem referencyjnego manekina pomiarowego BK4100-D. Funkcja HRTF o rozdzielczości 8802 punktów

	LS		MagLS		SPR		TA	
	Ambeo	Zyilia	Ambeo	Zyilia	Ambeo	Zyilia	Ambeo	Zyilia
Komora Bezechowa	125, 500, 4k, 8k	125, 250, 500, 1k, 2k, 4k	125, 500, 1k	125, 250, 500, 1k, 4k, 8k	125, 500	125, 250, 500, 1k, 4k	125, 500, 1k, 4k, 8k	125, 250, 500, 1k, 4k
Studio dźwiękowe	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 2k, 4k, 8k	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 4k	63, 1k, 2k, 4k, 8k	63, 125, 250, 500, 1k, 8k	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 4k
Korytarz	–	63, 125, 250, 500, 1k, 8k	4k, 8k	63, 125, 250, 500, 1k, 8k	4k, 8k	63, 125, 250, 500, 4k, 8k	4k, 8k	63, 125, 250, 500, 1k,
Studio	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 2k, 4k, 8k	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 4k	63, 1k, 2k, 4k, 8k	63, 125, 250, 500, 1k, 8k	63, 1k, 2k, 8k	63, 125, 250, 500, 1k, 4k
Sala Koncertowa	–	b/d	63, 8k	b/d	–	b/d	2k	b/d
Komora Pogłosowa	500, 1k	63, 125, 250, 500,	500, 1k	63, 125, 250, 500,	500, 1k, 4k, 8k	63, 125, 250, 500, 4k, 8k	500, 1k	63, 125, 250, 500,