Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej Nr 60

XXVIII Seminarium

ZASTOSOWANIE KOMPUTERÓW W NAUCE I TECHNICE 2018

Oddział Gdański PTETiS

doi: 10.32016/1.60.15

EKSTERNALIZACJA W BINAURALNEJ AMBISONICZNEJ AURALIZACJI ŹRÓDEŁ KIERUNKOWYCH

Bartłomiej MRÓZ[1], Bożena KOSTEK[2]

1. Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki, Laboratorium Akustyki Fonicznej tel.: +48 58 347 16 36 e-mail: bartlomiej.mroz@pg.edu.pl
2. Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki, Laboratorium Akustyki Fonicznej tel.: +48 58 347 27 17 e-mail: bokostek@audioakustyka.org

Streszczenie: W artykule przedstawiono najważniejsze składniki procesu skutecznego renderowania trójwymiarowego obrazu dźwiękowego za pomocą słuchawek. W tym celu badany jest stopień oddziaływania poszczególnych czynników wpływających na eksternalizację dźwięku: śledzenie położenia głowy (ang. *head tracking*), indywidualne funkcje przenoszenia głowy (HRTF – *Head Related Transfer Function*, odnoszące się do matematycznej funkcji propagacji dźwięku wokół głowy), model pomieszczenia, jak również tzw. efekt zgodności pomieszczenia oraz indywidualne wyrównywanie charakterystyki przenoszenia słuchawek. Uzyskane wyniki wskazują, że śledzenie głowy, a także indywidualne funkcje przenoszenia głowy mają kluczowe znaczenie dla wiernej reprodukcji dźwięku. Z przeprowadzonych badań wynika również, że efekt zgodności pomieszczenia i wyrównywanie charakterystyki przenoszenia słuchawek są znaczącymi elementami procesu auralizacji.

Słowa kluczowe: ambisonia, eksternalizacja, technologia binauralna, auralizacja

1. WSTĘP

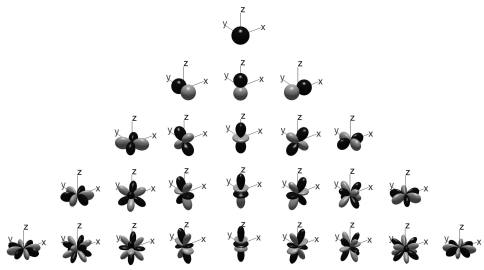
Rzeczywiste źródła dźwięku są zwykle postrzegane jako zlokalizowane poza głową. Problem pojawia się, gdy dźwięk przestrzenny (3D) jest odtwarzany na słuchawkach. W takim przypadku dynamiczne renderowanie binauralne (dwuuszne reprodukowanie odseparowanych treści dźwiękowych) prowadzi do błędów eksternalizacji, zatem źródła są postrzegane wewnątrz głowy odbiorcy [1]. Parametry wpływające na eksternalizację źródeł renderowanych binauralnie obejmują: (i) zindywidualizowane funkcje transferu głowy (ang. HRTFs – *Head-Related Transfer Functions*) [2], (ii) śledzenie ruchów (położenia) głowy (ang. *head tracking*) [3], (iii) charakterystykę czasu pogłosu [4], (iv) pomieszczenie odsłuchowe (np. efekt zgodności pomieszczenia odsłuchowego z dźwiękiem odbieranym w słuchawkach [5,6]) i (v) liniowość charakterystyk przenoszenia słuchawek [7,8].

W niniejszym artykule w pierwszej kolejności przedstawiono podstawowe pojęcia dotyczące dźwięku 3D [9,10]. Następnie skupiono się na problemie eksternalizacji w odsłuchu słuchawkowym oraz na zbadaniu zgodności pomieszczenia odsłuchowego z dźwiękiem w słuchawkach. Przedstawiono założenia eksperymentów oraz wnioski z uzyskanych wyników.

1.1. Ambisonia

Ambisonia jest techniką dźwięku przestrzennego w pełnym zakresie: oprócz płaszczyzny poziomej, obejmuje źródła dźwięku ponad i poniżej uszu słuchacza. Ambisonia w swojej zasadzie opiera się na syntezie pola falowego [11].

Wzrost rzędu ambisonii powoduje wzrost rozdzielczości odwzorowania we współrzędnych biegunowych, jak pokazano na rys. 1, a w konsekwencji, wyższej rozdzielczości kierunkowości dźwięku (bardziej „skupiona" lokalizacja) [12]. Co więcej, umożliwia dekodowanie pola akustycznego dla asymetrycznych ustawień głośników. Główną wadą jest wzrost liczby kanałów i zarazem głośników potrzebnych do odtworzenia odtwarzania HOA (*Higher Order Ambisonics*). Jeśli przyjmie się N jako rząd ambisonii, to liczba kanałów w odsłuchu 2D (pierścień horyzontalny) wynosi $2N + 1$, natomiast dla ambisonii 3D (pełna sfera) oznacza to $(2N + 1)^2$.



Rys. 1. Komponenty sferyczne ambisonii wyższego rzędu (HOA). Źródło: mathworks.com

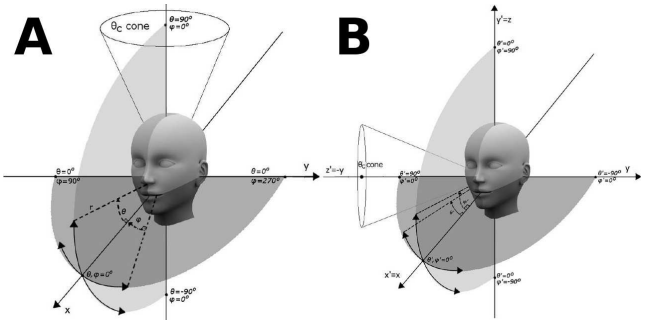
1.2. Dźwięk binauralny

Dźwięk binauralny ma za zadanie wywołanie trójwymiarowego wrażenia dźwięku u słuchacza. W ujęciu binauralnym zakłada się użycie słuchawek i zazwyczaj dwóch źródeł sygnału (chociaż w literaturze można znaleźć eksperymenty, które zakładają odstępstwa od tego podejścia [13,14]). Sygnały te zawierają szereg informacji pomagających stworzyć tzw. obraz słuchowy (ang. *auditory image*), tj. umiejscowić źródła dźwięku w przestrzeni. Lokalizacja odbywa się na dwóch płaszczyznach:

- wysokość, kierunek pionowy, płaszczyzna pionowa (ang. *elevation*, *sagittal direction*, *median plane*);
- szerokość, kierunek boczny/poziomy, płaszczyzna pozioma (ang. *lateral direction, lateral plane*).

Lokalizacja dźwięku obejmuje również takie aspekty jak dystans źródło-słuchacz, wielkość źródła dźwięku czy też jego kierunkowość.

W celu ułatwienia dokładnego opisu umiejscowienia wirtualnych źródeł dźwięku przekształcono sferyczny układ odniesienia tak, aby zamiast definiowania stożków na osi Z (rys. 2.), można je było rzutować na osi Y [15]. W konsekwencji dla różnych kątów azymutu (θ') można zdefiniować różne stożki pomieszania (ang. *cone of confusion*) – stożkowy zestaw punktów, promieniujący na zewnątrz od położenia w połowie drogi między uszami słuchacza, z którego źródło dźwięku generuje identyczne opóźnienia fazowe i przejściowe rozbieżności, sprawiając, że użycie informacji typowych dla słyszenia dwuusznego staje się nieużyteczne w lokalizowaniu dźwięku [16]). Oba systemy są równoważne, ponieważ układy współrzędnych mogą zostać przekształcone wzajemnie z obrotem o 90°.



Rys. 2. Standardowy układ sferyczny (A) oraz odwzorowany (B) [15]

1.3. Eksternalizacja

Pojęcie eksternalizacji oznacza zasadniczo lokalizację „poza głową". W eksternalizacji dźwięku w słuchawkach dźwięk przetwarzany jest w taki sposób, że postrzeganie jego położenia zmienia się z osi między uszami na zewnątrz głowy. Jednakże, nawet jeśli stosowane jest śledzenie ruchów głowy (ang. *head tracking*) w syntezie binauralnej, można zauważyć, że nadal występuje kilka problemów:

- eksternalizacja jest niżej oceniana dla kierunku frontalnego lub tylnego [17];
- percepcja odległości i lokalizacji może być zniekształcona;
- może być słyszalna interpolacja między kierunkami [18];
- scena dźwiękowa może nie brzmieć dość przekonująco.

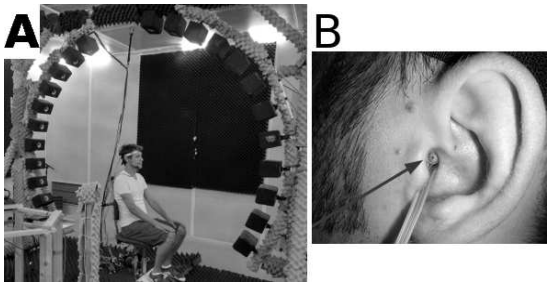
Paradoksalnie, w odsłuchu binauralnym równie istotnym czynnikiem jest akustyka pomieszczenia [19]. W badaniach wykonanych przez Völka i in. [20] zaobserwowano większą zgodność oceny eksternalizacji, jeśli funkcje przenoszenia głowy zostały zmierzone przed procesem renderowania, szczególnie w pomieszczeniu pogłosowym dla indywidualnych funkcji HRTF. Poprawa postrzegania odległości i lokalizacji w kontekście indywidualizacji HRTF jest również udokumentowana w badaniach Møllera [21] oraz Majdaka [22]. Ponadto można wprowadzić renderowanie ambisoniczne w celu zapewnienia płynnej/szybkiej rotacji źródeł. Jednakże, użycie ambisonii powoduje, że omawiane przez Bernschütza i in. [23], niedopasowanie modalne wynikające z przeskalowania w dół do zadanego rzędu (liczby kanałów) lub efekt filtracji grzebieniowej (zabarwienie), wydają się mieć wpływ na eksternalizację źródła. Potwierdzają to również badania Ben-Hura i in. [24]).

Kolejnym elementem procesu jest odpowiedź impulsowa słuchawek. Generalnie zakłada się, że charakterystyka częstotliwościowa słuchawek nie musi być płaska, to jednak ze względu na wrażliwość spektralną lokalizacji słuchowej, w przypadku aplikacji audio 3D, pożądana jest odpowiedź liniowa [25,26]. W tym celu przeprowadza się pomiar funkcji transferu słuchawek (*headphone transfer functions*, HpTFs). Teoretycznie, zmierzona odpowiedź impulsowa powinna w prosty sposób posłużyć do stworzenia filtra o charakterystyce odwrotnej do zmierzonych HpTF, dzięki któremu słuchawki otrzymają liniową charakterystykę. Jednakże, drobne zmiany geometryczne pomiędzy słuchawkami i indywidualnymi cechami anatomicznymi słuchacza powodują różnice w rezonansie zewnętrznego przewodu słuchowego (ułożenie słuchawek), które z kolei prowadzą do istotnych zmian wysokich częstotliwości funkcji przenoszenia słuchawek (HpTF) [27]. Z tego względu, podczas pomiaru HpTF słuchacz jest proszony o wielokrotne zdejmowanie i zakładanie słuchawek, a następnie oblicza się uśrednioną wartość tej wielkości. W tym celu, podobnie, jak w przypadku otwartych baz funkcji HRTF, została opracowana baza zbierająca funkcje HpTF – *Princeton Headphone Open Archive* (PHOnA) [28].

1.4. Funkcje przenoszenia głowy (ang. *Head-Related Transfer Functions*, HRTFs)

Binauralne odtwarzanie w słuchawkach wymaga wykorzystania funkcji HRTF w celu wywołania pełnego wrażenia słuchowego, tj. możliwości rozpoznania kierunku i eksternalizacji. HRTF składa się z odpowiedzi impulsowych mierzonych przy wejściu do kanału słuchowego, przy czym kanał słuchowy jest w takim przypadku zablokowany przez mikrofon. Ogólny (niezindywidualizowany) zestaw funkcji HRTF uzyskuje się za pomocą sztucznej głowy i manekina [21]. Ponadto istnieje wiele baz danych HRTF, w których zapisano uśrednione zestawy HRTF. Wśród głównych baz HRTF można wymienić następujące: LISTEN [29], CIPIC [30], FIU [31] i MIT-KEMAR [32].

Aby uzyskać indywidualny zestaw HRTF, małe mikrofony pomiarowe umieszcza się przy wejściu do kanałów słuchowych słuchacza. Wokół słuchacza są rozmieszczone w stałej konfiguracji głośniki. Z tego powodu typowo, to słuchacz obraca się w celu zebrania odpowiedzi impulsowych z każdego kierunku. Liczba głośników i zmierzonych pozycji określa rozdzielczość zestawu HRTF. Aby zapewnić precyzję i stabilność pomiaru, należałoby wykorzystywać np. śledzenie ruchów głowy, aby pominąć pomiar w przypadku zbyt dużego ruchu głowy [33]. Przykład takiego pomiaru przedstawiono na rys. 3.



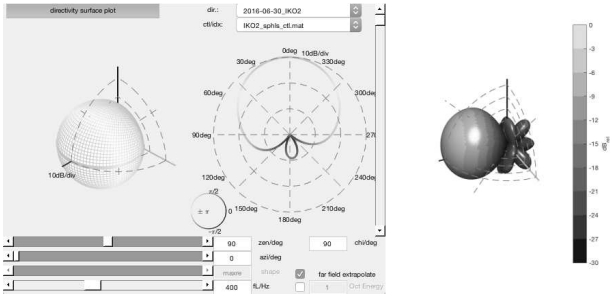
Rys. 3. Pomiar funkcji HRTF w ARI, Wiedeń (A). Mikrofon umieszczony w kanale usznym (B) (źródło: kfs.oeaw.ac.at).

MOST WIEDZY Pobrano z mostwiedzy.pl

Należy też wspomnieć o opracowanym standardzie SOFA (nr AES69-2015), określającym, jaką strukturę pliku należy zastosować w przypadku nagrań ambisonicznych, funkcjach HRTF czy też funkcjach HpTF. Celem standardu SOFA jest reprezentacja przestrzennie zorientowanych danych w sposób ogólny w celu zapewnienia wymienności plików. Ponadto, format SOFA został opracowany w sposób, który ma zapewnić kompatybilność wsteczną w przyszłości.

1.5. Kierunkowe źródła dźwięku

Współczynnik D/R (ang. *Direct-to-Reverberant Ratio*, stosunek energii dźwięku bezpośredniego do energii pogłosu) pozwala na określenie odległości źródła od słuchacza. W danym pomieszczeniu stosunek D/R jest odwrotnie proporcjonalny do odległości źródła dźwięku od słuchacza i odzwierciedla stosunek energii dźwięku bezpośredniego do odbitego. Wprowadzenie kierunkowego źródła dźwięku pozwala na precyzyjną kontrolę współczynnika D/R, co może się przyczynić do polepszenia lokalizacji źródła dźwięku. Idea ta została zaproponowana przez Laitinena i in. [34] oraz zbadana m.in. przez Wendta i in. [35]. Ponadto, IEM Graz (IEM – *Institut für Elektronische Musik und Akustik*) prowadzi pomiary w przestrzeni 2D oraz 3D wzorców kierunkowych rzeczywistych źródeł dźwięku, które są ogólnie dostępne w repozytorium DirPat [36]. Repozytorium jest również wyposażone w zestaw narzędzi do wizualizacji i analizy wzorców 3D/2D (rys. 4).



Rys. 4. Narzędzie do wizualizacji 3D wzorców kierunkowych źródeł z repozytorium DirPat (źródło: opendata.iem.at)

1.6. Wpływ zgodności pomieszczenia (ang. *Room Divergence Effect*)

Werner i in. [5] zaobserwowali pogorszenie eksternalizacji dla rozbieżności pomiędzy pomieszczeniem odsłuchowym a zsyntetyzowanym w porównaniu ze zbieżnymi uwarunkowaniami akustycznymi pomieszczeń. W opracowaniu tych badań omówiono nie tylko efekt braku zgodności pomieszczenia, ale także dezorientację przód/tył oraz dodanie wizualnych obiektów pomocniczych („pozornych" głośników).

2. EKSPERYMENT

2.1. Założenia eksperymentalne

Podobne obserwacje i wnioski zostały przedstawione w badaniu przeprowadzonym przez Mroza i in. [37]. Oprócz efektu zgodności pomieszczenia, w badaniu zestawiono i porównano ze sobą szereg aspektów, które bardzo rzadko są zestawiane w jednym eksperymencie, tj.: indywidualizacja funkcji HRTF, śledzenie ruchów głowy, model pomieszczenia, indywidualną equalizację słuchawek.

Kluczowe aspekty projektu obejmowały:

- dźwięk binauralny;
- Ambisonia 2D rzędu N=11 z wagami max-r_E;
- Wczesne odbicia do 3 rzędu (źródła pozorne);
- Dyfuzyjny charakter pogłosu, „przyczepiony" w tzw. *mixing time* – t_m, co odpowiada średniemu TOA (ang. *Time of Arrival*) odbić trzeciego rzędu;
- Kierunkowe źródło dźwięku (wiązka);
- Wirtualna konfiguracja koła 24-głośnikowego;
- Wirtualna auralizacja pomieszczenia IEM CUBE;
- Przeprowadzenie eksperymentu w dwóch różnych pomieszczeniach: komorze bezechowej i sali wykładowej;
- Indywidualne charakterystyki HRTF (Acoustics Research Institute, ÖAW);
- Uogólnione HRTF (MIT KEMAR);
- Różne modele pomieszczeń:
  - tylko dźwięk bezpośredni,
  - bezpośrednie + wczesne odbicia (do 3 rzędu),
  - pogłos rozproszony + wczesny + pogłos.

2.2. Symulacja pomieszczenia

Prostokątny pokój o wymiarach 1m × 12m jest symulowany za pomocą prostego modelu źródeł pozornych do trzeciego rzędu. Odpowiadające kierunki nadejścia dźwięku (ang. *Direction of Arrival*, DOA) czasy nadejścia (ang. *Time of Arrival*, TOA) oraz poziomy dla każdej ścieżki propagacji są zawarte w tab. 1. Pogłos rozproszony jest skonstruowany za pomocą szumu białego o rozkładzie normalnym wygenerowanym oddzielnie dla każdego kanału SRDRIR (ang. *Source and Receiver Directional Room Impulse Response*). Aby uwzględnić zależność częstotliwościową pogłosu rozproszonego, podzielono go na pasma oktaw (tab. 2), w której zastosowano zanikającą obwiednię odpowiadającą pożądanemu czasowi pogłosu (T_{60}).

Tab. 1. DOA, TOA i poziomy każdej ścieżki propagacji do 3. rzędu źródeł pozornych

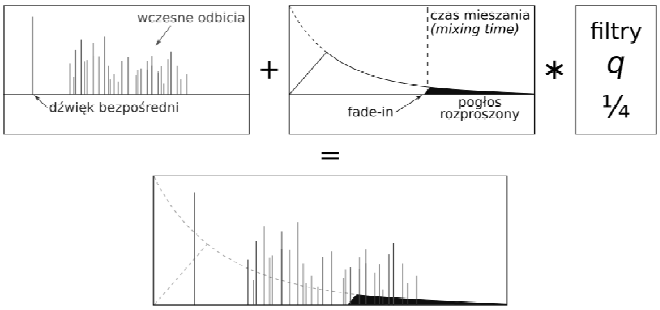
rząd	DOA (deg)	TOA (ms)	LVL (dB)	znak	rząd	DOA (deg)	TOA (ms)	LVL (dB)	znak
0	0	5	0	-					
1	-9	12	3,7	+	3	-94	60	-13,6	+
1	53	13	-7,4	+	3	64	63	-20	+
1	-99	48	-24,9	+	3	154	67	-10,7	+
1	164	58	-21,7	-	3	71	73	-17,7	+
2	31	17	0,2	+	3	-6	77	-25,5	-
2	-91	50	-17,1	+	3	-159	81	-14,8	+
2	-100	59	-16	+	3	-14	83	-12,7	+
2	153	60	-23,9	-	3	-150	83	-37,2	+
2	70	62	-26,3	+	3	118	85	-28	-
2	164	66	-10,8	+	3	-48	89	-25,7	-
2	-14	76	-23,9	-	3	-103	109	-30,2	+
2	-156	76	-25,5	-	3	164	129	-28,4	-

Tab. 2. Projekt filtrów dla pogłosu rozproszonego

Typ	T_{60}	Dolna częstotliwość odcięcia	Górna częstotliwość odcięcia
Dolnoprzepustowy	700 ms		250 Hz
Pasmowoprzepustowy	500 ms	250 Hz	500 Hz
Pasmowoprzepustowy	500 ms	500 Hz	1000 Hz
Pasmowoprzepustowy	500 ms	1000 Hz	2000 Hz
Pasmowoprzepustowy	415 ms	2000 Hz	4000 Hz
Pasmowoprzepustowy	335 ms	4000 Hz	8000 Hz
Górnoprzepustowy	250 ms	8000 Hz	

MOST WIEDZY Pobrano z mostwiedzy.pl

Generowanie modelowanego SRDRIR jest przedstawione na rys. 5 dla pojedynczego kanału, gdzie *q=0,0011*, a część rozproszona pojawia się (ang. *fade in*) w 20ms, począwszy od czasu mieszania (ang. *mixing time*) [38].



Rys. 5. Schemat modelowania RIR za pomocą modelu 2D źródeł pozornych 3. rzędu oraz pogłosu rozproszonego modelowanego w pasmach oktawowych.

2.3. Synteza binauralna, funkcje HRTF, equalizacja

Sygnały ambisoniczne (2D) do rzędu 11, wynikające z SRDRIR, są dekodowane na 24-kanałowy, równomiernie rozproszony pierścień z głośników za pośrednictwem pseudo-odwrotnego dekodera. Ruch głowy słuchacza jest śledzony przez kompaktowe urządzenie zamontowane na słuchawkach [39], a cała scena jest obracana za pomocą prostego, niezależnego od częstotliwości mnożenia macierzy przed dekodowaniem. Dekodowane sygnały binauralne są uzyskiwane przez operację splotu sygnałów głośnikowych z funkcjami HRTF (nie- lub zindywidualizowanymi) dla odpowiedniego kierunku (wybór wg najbliższego sąsiada). Zindywidualizowane funkcje HRTF (dla pola dalekiego), użyte w eksperymencie, zostały zmierzone Instytucie Akustyki Austriackiej Akademii Nauk w Wiedniu (dla 1550 kierunków), natomiast nieindywidualne HRTFy wybrano z bazy KEMAR. Zamiast pomiaru equalizacji słuchawek, wszyscy uczestnicy wykonali procedurę equalizacji wg pracy Griesingera [8], aby uwzględnić efekt filtracji grzebieniowej - zabarwienie wprowadzane przez słuchawki w eksperymencie.

2.4. Eksperyment psychoakustyczny

W teście wzięło udział 6 słuchaczy (mężczyźni, średni wiek 30 lat) z doświadczeniem w dziedzinie dźwięku przestrzennego. Ich zadaniem była ocena jakości eksternalizacji na ciągłej skali (od słabej do bardzo dobrej) w zadaniu o wielu pobudzeniach. W eksperymencie zostały zróżnicowane następujące uwarunkowania:

- złożoność modelowanej akustyki pomieszczenia: tylko dźwięk bezpośredni, dźwięk bezpośredni z wczesnymi odbiciami; bezpośredni dźwięk z wczesnymi odbiciami i pogłosem rozproszonym;
- dynamiczne zmiany: włączone i wyłączone śledzenie ruchów głowy;
- HRTF: indywidualne i nieindywidualizowane;
- equalizacja słuchawek: włączona lub wyłączona.

Dodatkowo założono dwa warunki zakotwiczenia, które modelują opóźnienia czasowe, symulując tylko ITD (z- i bez śledzenia głowy). Wszystkie warunki eksperymentalne są zawarte w tab. 3. Każdy słuchacz musiał wykonać testy w komorze bezechowej i sali wykładowej (porządek losowy). W każdym pokoju zadanie powinno być wykonane dwukrotnie, za każdym razem z losową kolejnością przedstawiania 14 warunków. Sygnałem źródłowym była mowa męska (EBU SQAM CD, tor ID 50), a powstałe bodźce odtwarzano w pętli przez słuchawki Stax Omega II (SR-007). Słuchacze mogli płynnie przełączać się między bodźcami i słuchać ich tak często, jak chcieli.

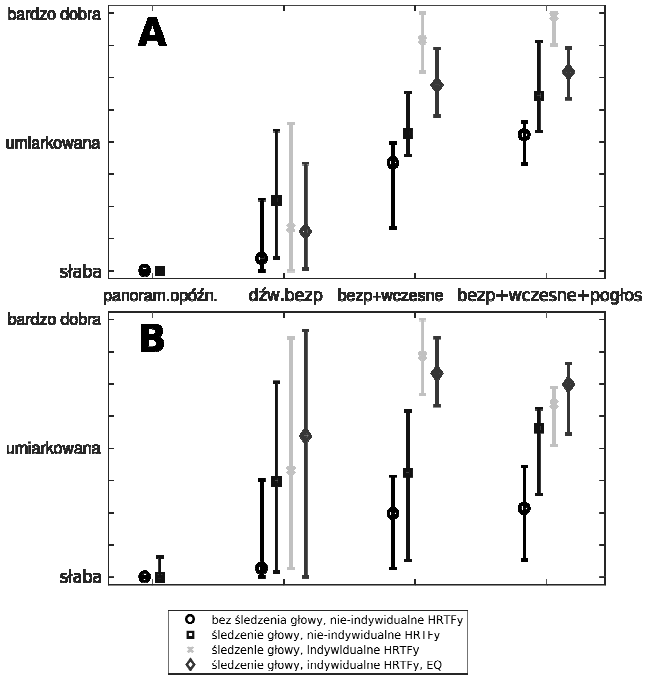
Warunek	długość RIR
1,2	dźwięk bezpośredni
3,6,9,12	dźwięk bezpośredni
4,7,10,13	bezpośredni+wczesne odbicia
5,8,11,14	bezpośredni+wczesne+pogłos rozproszony

Warunek	HRTF	śledzenie głowy	EQ
1	opóźnienia	-	-
2	opóźnienia	+	-
3-5	KEMAR	-	-
6-8	KEMAR	+	-
9-11	indywidualne	+	-
12-14	indywidualne	+	+

Tab. 3. Lista warunków ocenianych w eksperymencie

2.5. Wyniki

Wyniki ocen zostały pokazane na rys. 6 jako mediana i przedziały ufności 95%. Niezależnie od warunków odsłuchowych, warunki zakotwiczenia (stereofonia z opóźnieniami czasowymi) zostały ocenione jako znajdujące się wewnątrz głowy (słaba eksternalizacja). Zgodnie z założeniem, zwiększenie złożoności modelu pomieszczenia również zwiększa eksternalizację. Dodanie wczesnych odbić znacznie poprawia eksternalizację, podczas gdy dodanie części rozproszonej daje wyniki porównywalne do przypadku „dźwięk bezpośredni + wczesne odbicia". Można zatem stwierdzić, że śledzenie głowy i wykorzystanie zindywidualizowanych funkcji HRTF może dodatkowo poprawić eksternalizację, podczas gdy wyrównywanie charakterystyki słuchawek nie powoduje znaczącej poprawy. Podobne wyniki można znaleźć w literaturze [1,3,33].



Rys. 6. Wyniki oceny eksternalizacji: (A) w pomieszczeniu wykładowym, (B) w komorze bezechowej (panoram.opóźn. oznacza panoramowanie opóźnieniem)

MOST WIEDZY Pobrano z mostwiedzy.pl

3. WNIOSKI I PODSUMOWANIE

W eksperymentach najniższe oceny eksternalizacji zostały wskazane, niezależnie od wykorzystania możliwości śledzenia ruchów głowy, gdy nie uwzględniono modelu pomieszczenia i funkcji HRTF. Jednakże, jeśli wykorzystano opcję śledzenia ruchów głowy, auralizację pomieszczenia oraz zestaw funkcji HRTF, to poprawiło to ocenę lokalizacji dźwięku. Co więcej, funkcje HRTF mogą być dowolne, jeśli auralizowane pomieszczenie jest bezechowe (sygnał nie ma żadnych odbić). Niemniej jednak, przy bardziej złożonym sygnale (wczesne odbicia, rozproszony pogłos), kluczowe stają się indywidualne funkcje HRTF. Odpowiedź na pytanie, dla jakiego stopnia złożoności sygnału to następuje, nie jest nadal jasna i wydaje się być zależna od wystąpienia zjawiska odczucia zgodności pomieszczenia z odsłuchiwanym dźwiękiem. W warunkach komory bezechowej mniej złożony sygnał zyskiwał w ocenie słuchaczy wyższe wyniki, podczas gdy w zwykłym pomieszczeniu (sali wykładowej) eksternalizacja była zdecydowanie bardziej przekonująca dla auralizacji pomieszczenia z fazą bogatego pogłosu pełnego. Ten wynik świadczy o znaczeniu efektu zgodności pomieszczenia z odsłuchiwanym dźwiękiem. Jednakże, przypadek pośredni zyskał relatywnie wysokie oceny dla obu pomieszczeń. Z tego względu można zaryzykować twierdzenie, że wczesne odbicia (do 3. rzędu) mogą być uniwersalnym rozwiązaniem. Kreowany dźwięk 3D w odsłuchu słuchawkowym jest wystarczająco realistyczny nawet w przypadku dużych różnic pomiędzy pomieszczeniem wirtualnym a faktycznym pomieszczeniem odsłuchowym, a co za tym może poprawić wrażenie eksternalizacji. Z uwagi na niewielką złożoność obliczeniową w takim przypadku, ograniczenie modelu tylko do wczesnych odbić wydaje się być wystarczająco uzasadnione.

4. BIBLIOGRAFIA

1. Hartmann W. M., Wittenberg A.: On the externalization of sound images, The Journal of the Acoustical Society of America, 99, 6, 1996, 3768-3688.
2. Oberem J., Masiero B., Fels J.: Experiments on the Authenticity and Plausibility of Binaural Reproduction via Headphones employing different recording methods, Applied Acoustics, 114, 2016, 71-78.
3. Brimijoin W. O., Boyd A. W., Akeroyd M. A.: The contribution of head movement to the externalization and internalization of sounds, PLoS ONE, 8, 12, 2013, 1-12.
4. Lee H., Crawford-Emery R.: The subjective effect of BRIR length perceived headphone sound externalisation and tonal colouration, 136th Audio Eng. Soc. Convention, Paper no. 9044, Berlin, Niemcy, 2014.
5. Werner S., Klein F., Mayenfels T., Brandenburg K.: A summary on acoustic room divergence and its effect on externalization of auditory events, 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), 2016, 1-6.
6. Werner S., Götz G., Klein F.: Influence of Head Tracking on the Externalization of Auditory Events at Divergence between Synthesized and Listening Room Using a Binaural Headphone System, 142nd Audio Eng. Soc. Convention, Paper no. 9690, Berlin, Niemcy, 2017.
7. Ben-Hur Z., Brinkmann F., Sheaffer J., Weinzierl S., Rafaely B.: Spectral equalization in binaural signals represented by order-truncated spherical harmonics, The Journal of the Acoustical Society of America, 141, 6, 2017, 4087-4096.
8. Griesinger D.: Accurate Timbre and Frontal Localization without Head Tracking through Individual Eardrum Equalization of Headphones, 141st Audio Eng. Soc. Convention, Paper no. 9620, Los Angeles, USA, 2016.
9. Begault D. R., Wenzel E. M.: Headphone Localization of Speech, Human Factors, 35, 2, 1993, 361-376.
10. Völk F.: Externalization in data-based binaural synthesis, DAGA 2009, 35th Annual Conference on Acoustics, 2009.
11. Spors S., Ahrens J.: A comparison of wave field synthesis and higher-order Ambisonics with respect to physical properties and spatial sampling, 125th Audio Eng. Soc. Convention, Paper no. 7556, San Francisco, USA, 2008.
12. Mróz B.: Externalisation and Distance in Ambisonic Rendering on Headphones, Master thesis – Graz University of Technology and Gdańsk University of Technology, 2017.
13. Pomberger H., Sontacchi A., Frank M., Gmeiner T., Lucchi M.: Improved localization in the median plane with cue-preserving headphones, DAGA 2018, 44th Annual Conference on Acoustics, 2018, 948-951.
14. Erbes V., Schultz F., Lindau A., Weinzierl S.: An extraaural headphone system for optimized binaural reproduction, DAGA 2012, 38th Annual Conference on Acoustics, 2012.
15. Mattes S., Nelson P. A., Fazi F. M., Capp M.: Towards a human perceptual model for 3D sound localization, 28th Conference on Reproduced Sound: Auralisation: Designing With Sound, 2012.
16. Colman, A. M.: A Dictionary of Psychology, Oxford University Press, 3rd edition, 2014.
17. Kim S. M., Choi W.: On the externalization of virtual sound images in headphone reproduction: A Wiener filter approach, The Journal of the Acoustical Society of America, 117, 6, 2005, 3657-3665.
18. Katz B. F., Stitt P., Simon L., Hendrickx E., Andreopoulou A.: Contributions of head-related transfer function choice and head tracking to virtual loudspeaker binaural rendering, The Journal of the Acoustical Society of America, 141, 5, 2017, 3511-3512.
19. Mróz B.: Examination of the Factors Influencing Binaural Rendering on Headphones with the Use of Directivity Patterns, 144th Audio Eng. Soc. Convention, Paper no. 9953, Mediolan, Włochy, 2018.
20. Völk F., Heinemann F., Fastl H.: Externalization in binaural synthesis: effects of recording environment and measurement procedure, Acoustics '08, 2008, 6419-6424.
21. Møller H.: Fundamentals of binaural technology, Applied Acoustics, 36, 3, 1992, 171-218.
22. Majdak P., Masiero B., Fels J.: Sound localization in individualized and non-individualized crosstalk cancellation systems, The Journal of the Acoustical Society of America, 133, 4, 2013, 2055-2068.
23. Bernschütz B., Vázquez G. A., Pörschmann C., Arend J.: Binaural Reproduction of Plane Waves With Reduced Modal Order, Acta Acustica united with Acustica, 100, 2014, 972-983.

MOST WIEDZY Pobrano z mostwiedzy.pl

24. Ben-Hur Z., Brinkmann F., Sheaffer J., Weinzierl S., Rafaely B.: Spectral equalization in binaural signals represented by order-truncated spherical harmonics, The Journal of the Acoustical Society of America, 141, 6, 2017, 4087-4096.
25. Møller H., Jensen C., Hammershøi D., Sorensen M.: Design Criteria for Headphones, Journal of the Audio Engineering Society, 43, 4, 1995, 218-232.
26. Boren B., Roginska A.: The Effects of Headphones on Listener HRTF Preference, 131st Audio Eng. Soc. Convention, Paper no. 8537, Nowy Jork, USA, 2011.
27. Masiero B., Fels J.: Perceptually Robust Headphone Equalization for Binaural Reproduction, 130th Audio Eng. Soc. Convention, Paper no. 8388, Londyn, Wielka Brytania, 2011.
28. Boren B. B., Geronazzo M., Majdak P., Choueiri E.: PHOnA: A Public Dataset of Measured Headphone Transfer Functions, 137th Audio Engineering Convention, 2014.
29. http://recherche.ircam.fr/equipes/salles/listen/ (wyświetlono: 19 IX 2018).
30. Algazi V. R., Duda R. O., Thompson D. M., Avendano C.: The CIPIC HRTF database, Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575), 2001, 99-102.
31. Gupta N., Barreto A., Joshi M., Agudelo J. C.: HRTF database at FIU DSP Lab, 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, 2010, 169-172.
32. Gardner B., Martin K.: HRTF Measurements of a KEMAR Dummy-Head Microphone, MIT Media Lab Perceptual Computing - Technical Report #280, 1994.
33. Hendrickx E., Stitt P., Messonnier J. C., Lyzwa J. M., Katz B., de Boishéraud C.: Influence of head tracking on the externalization of speech stimuli for non-individualized binaural synthesis, The Journal of the Acoustical Society of America, 141, 3, 2017, s. 2011-2023.
34. Laitinen M. V., Politis A., Huhtakallio I., Pulkki V.: Controlling the perceived distance of an auditory object by manipulation of loudspeaker directivity, The Journal of the Acoustical Society of America, 137, 6, 2015, 462-468.
35. Wendt F., Zotter F., Frank M., Höldrich R.: Auditory Distance Control Using a Variable-Directivity Loudspeaker, Applied Sciences, 7, 7, 2017.
36. Brandner M., Frank M., Rudrich D.: DirPat – Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers, 144th Audio Eng. Soc. Convention, eBrief no. 425, Mediolan, Włochy, 2018.
37. Mróz B.: Experiment on externalization in binaural directional-source auralization, DAGA 2018, 44th Annual Conference on Acoustics, 2018, 336-338.
38. Lindau A., Kosanke L., Weinzierl S.: Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses, 128th Audio Eng. Soc. Convention, Paper no. 8089, Londyn, Wielka Brytania, 2010.
39. Romanov M., Berghold P., Frank M., Rudrich D., Zaunschirm M., Zotter F.: Implementation and evaluation of a low-cost headtracker for binaural synthesis, 142nd Audio Eng. Soc. Convention, Paper no. 9689, Berlin, Niemcy, 2017.

EKSTERNALIZATION IN BINAURAL AMBISONIC AURALIZATION OF DIRECTIONAL SOURCES

The article presents the most important components of the process of effectively rendering a three-dimensional sound image using headphones. To this end, the impact of a set of factors affecting sound externalisation is examined: head tracking, individual head transfer functions (HRTF – Head Related Transfer Function, related to the mathematical function of sound propagation around the head), room model, as well as a so-called room divergence effect and individual headphone alignment. Furthermore, the directivity aspect of the sound source is discussed as a convenient way to control Direct-to-Reverberant ratio and thus provides a robust control of the auditory distance. The results obtained indicate that the tracking of the head position as well as the individual functions of the head transfer are crucial for faithful sound reproduction. The conducted research also shows that the effect of room compliance and headphone alignment prove to be significant factors throughout the auralisation process.

Keywords: Ambisonics, externalisation, binaural technology, auralisation.

MOST WIEDZY Pobrano z mostwiedzy.pl