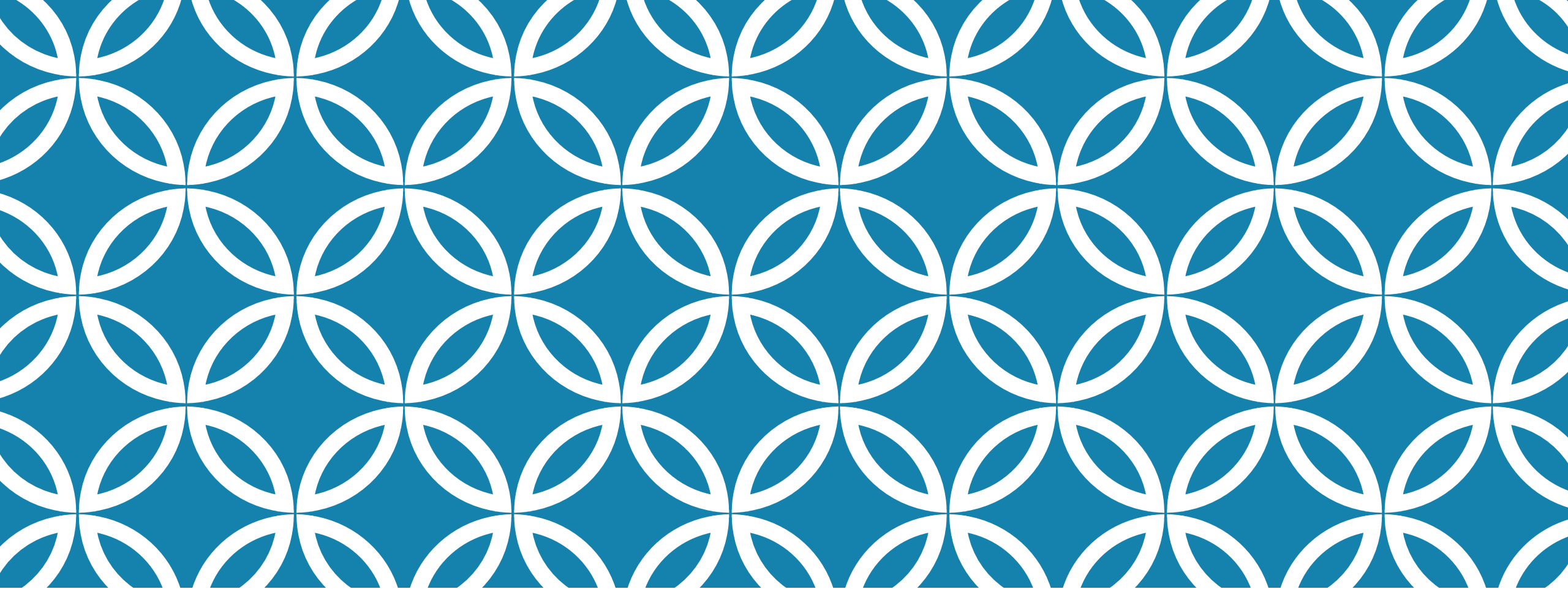




WSTĘP DO MULTIMEDIÓW

WYKŁAD 2 — LOKALIZACJA ŹRÓDŁA DŹWIĘKU, METODY OCENY JAKOŚCI DŹWIĘKU

A. Paula Pietrzak

LOKALIZACJA DŹWIĘKÓW

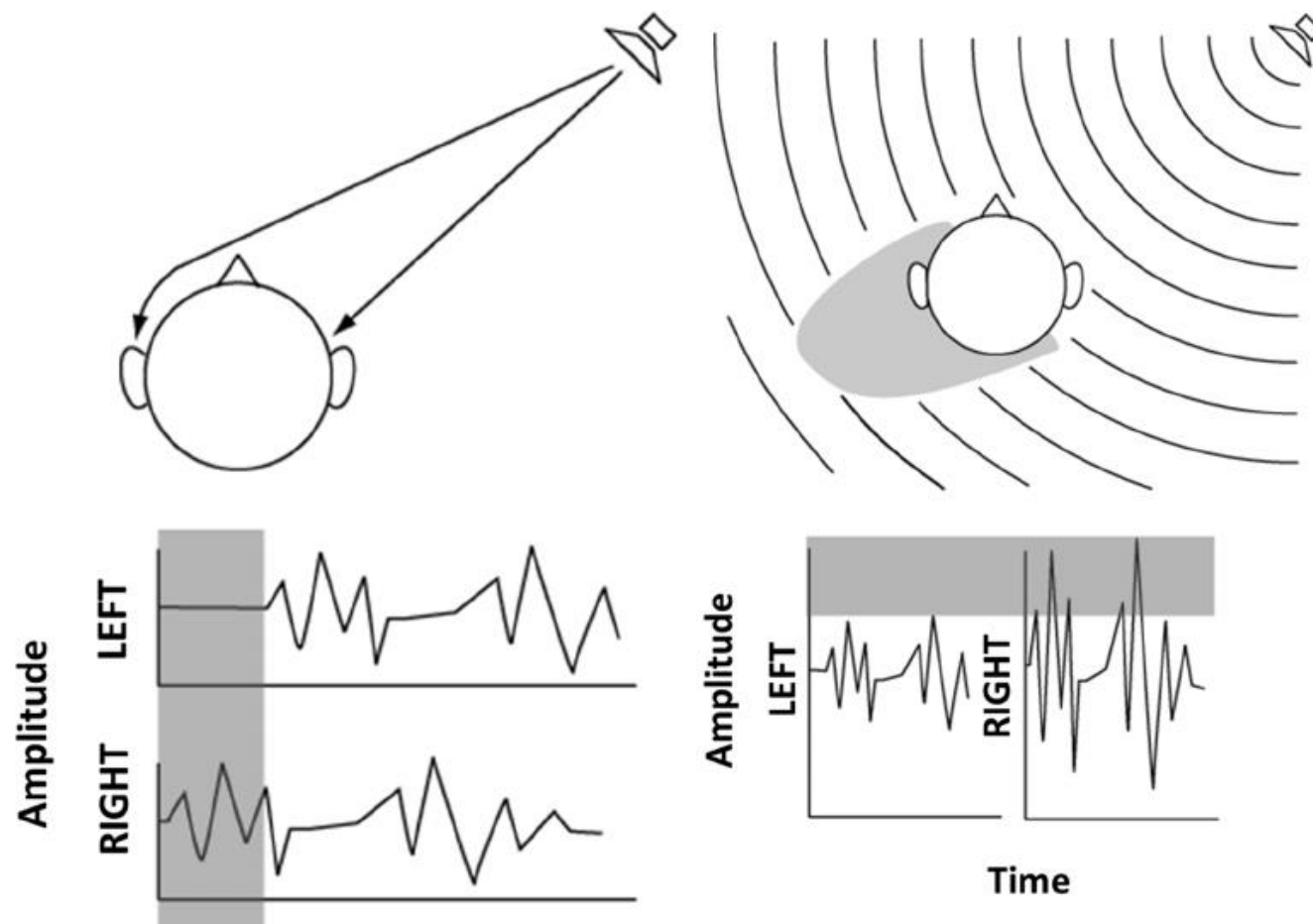
Zdolność oceny odległości od źródła dźwięku i kierunku z jakiego dociera dźwięk.

ITD – Interaural Time Difference

(dominuje dla niskich częstotliwości)

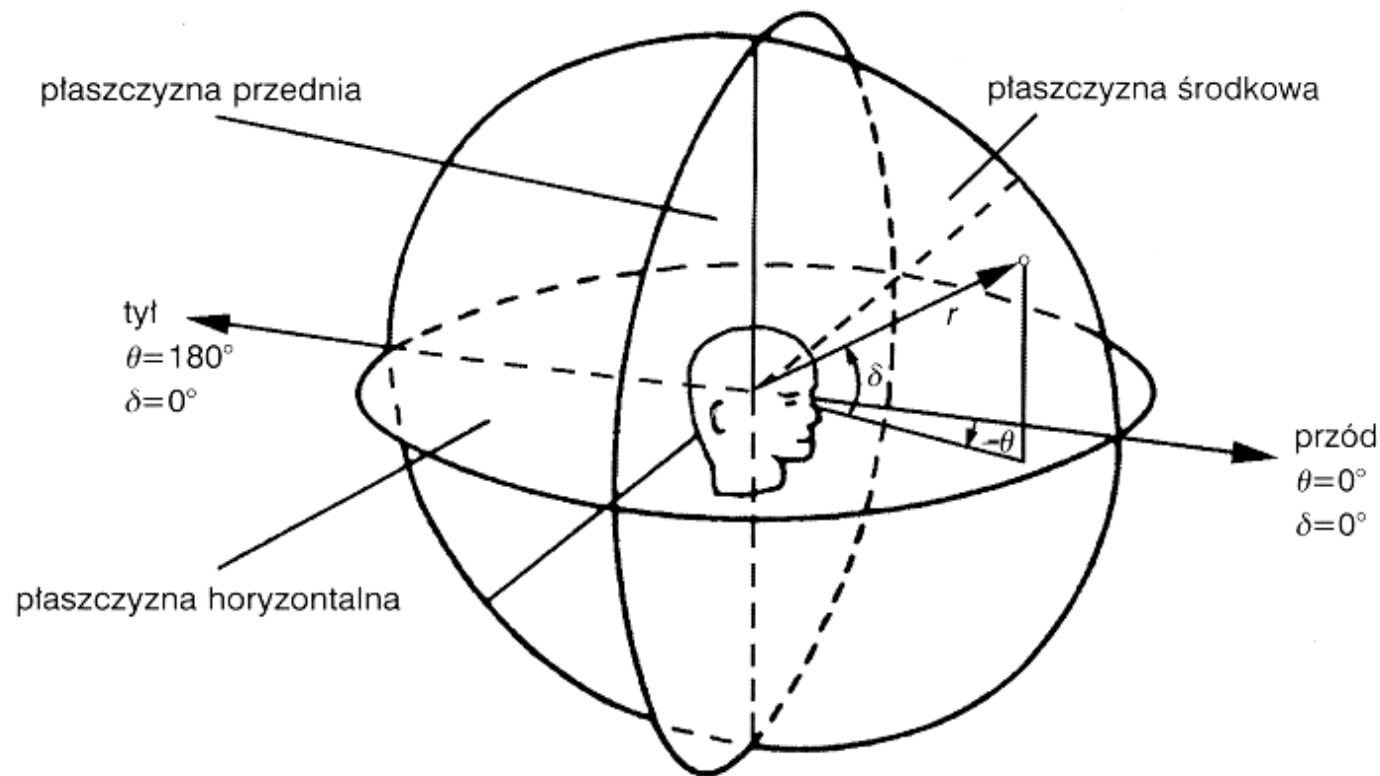
ILD – Interaural Level Difference

(dominuje dla wysokich częstotliwości)



LOKALIZACJA W PŁASZCZYŹNIE PIONOWEJ

- kąt elewacji (biegunowy)
- lokalizacja na podstawie częstotliwościowych cech dźwięku związanych z odbiciami od małżowiny usznej, itp.
- rozdzielczość znacznie gorsza niż w płaszczyźnie horyzontalnej



CZYNNIKI LOKALIZACYJNE

- ITD
- ILD (IID)
- Efekt precedensu
- Efekt „cocktail-party”
- MAA/MAMA
- lokalizacja odległości
- rola ruchów głowy
- HRTF
- doświadczenie słuchowe i wiązanie zjawisk akustycznych z obserwacjami wzrokowymi

DŹWIĘK PRZESTRZENNY

- Dźwięk binauralny
 - diotyczny
 - dychotyczny
- Lateralizacja
- Eksternalizacja
- Auralizacja – „wizualizacja” dźwięku
- Ambisonia

MONO
SOUND



STEREO
SOUND



5.1
DISCRETE CHANNEL
CONFIGURATION



3D AUDIO
AMBISONICS

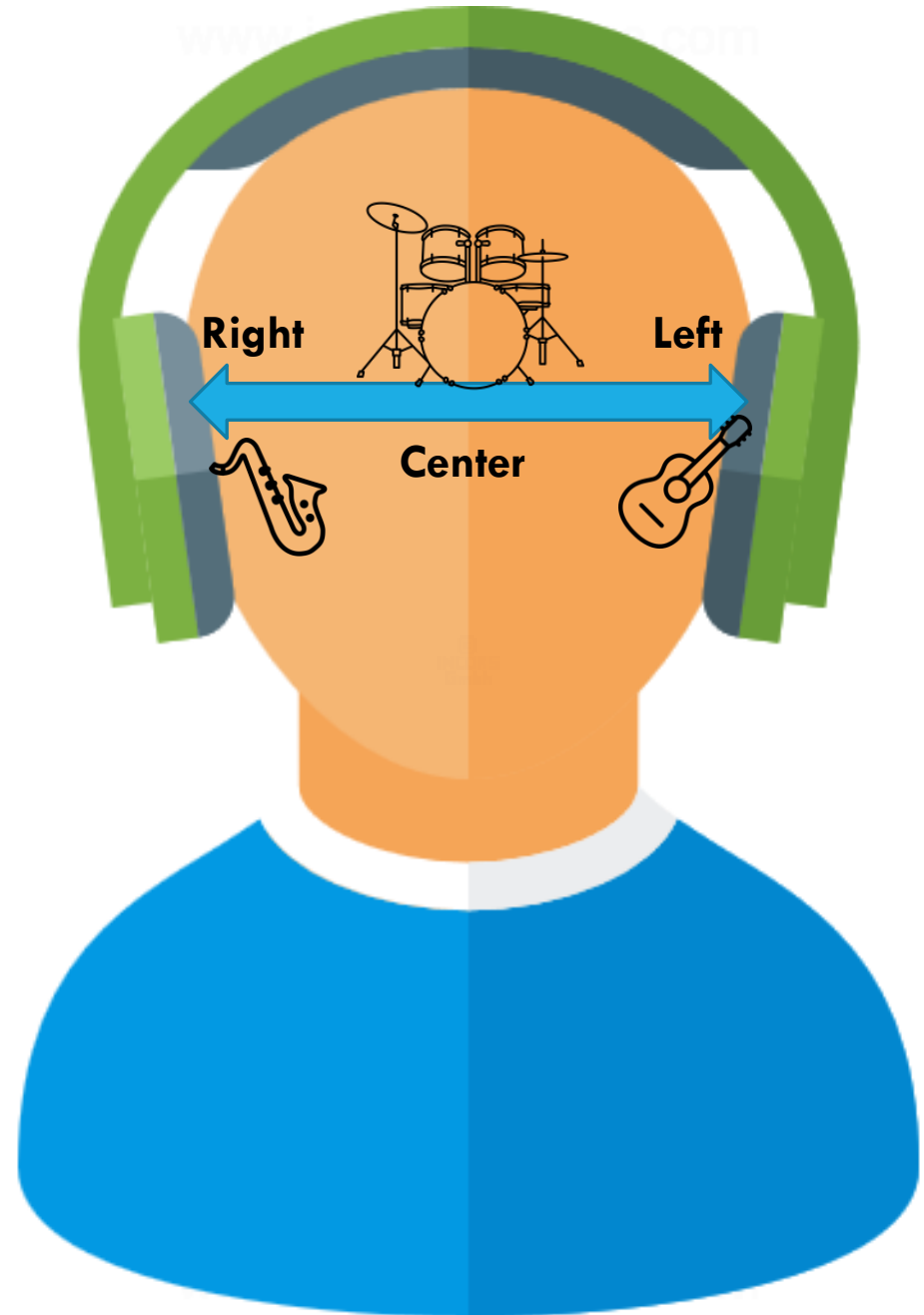


LATERALIZACJA

Występuje przy odsłuchu słuchawkowym stereo.

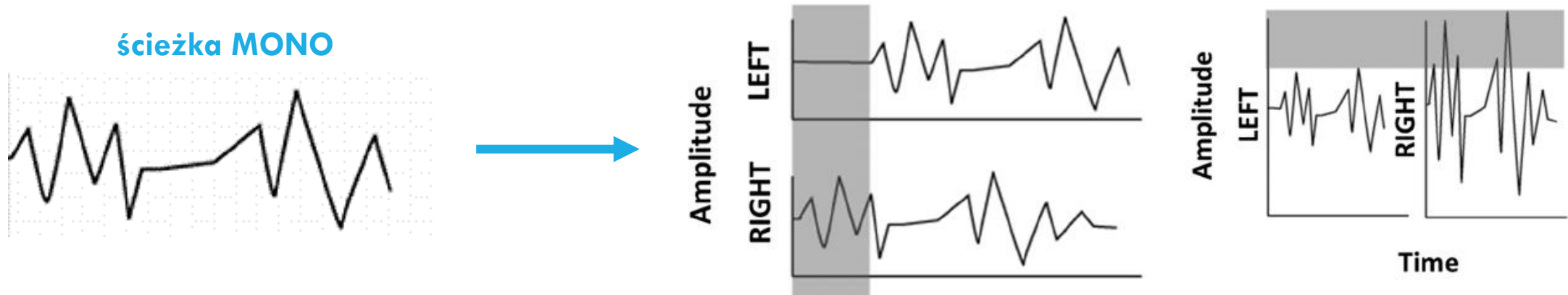
Źródło dźwięku lokalizowane jest w środku głowy, na osi prawe-lewe ucho.

Nie ma wrażenia dźwięku „poza” głową.

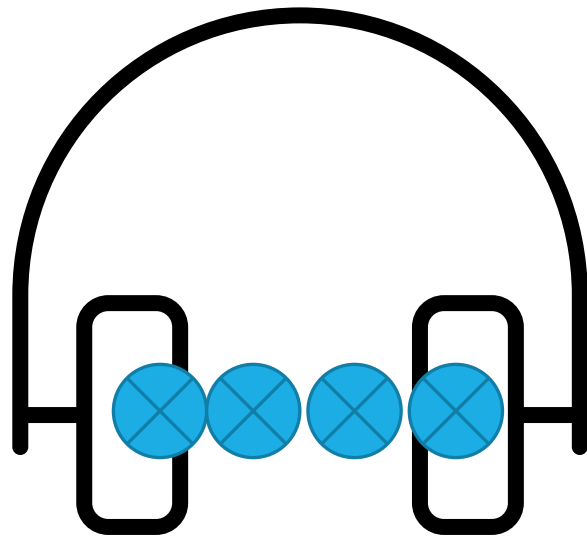


LATERALIZACJA

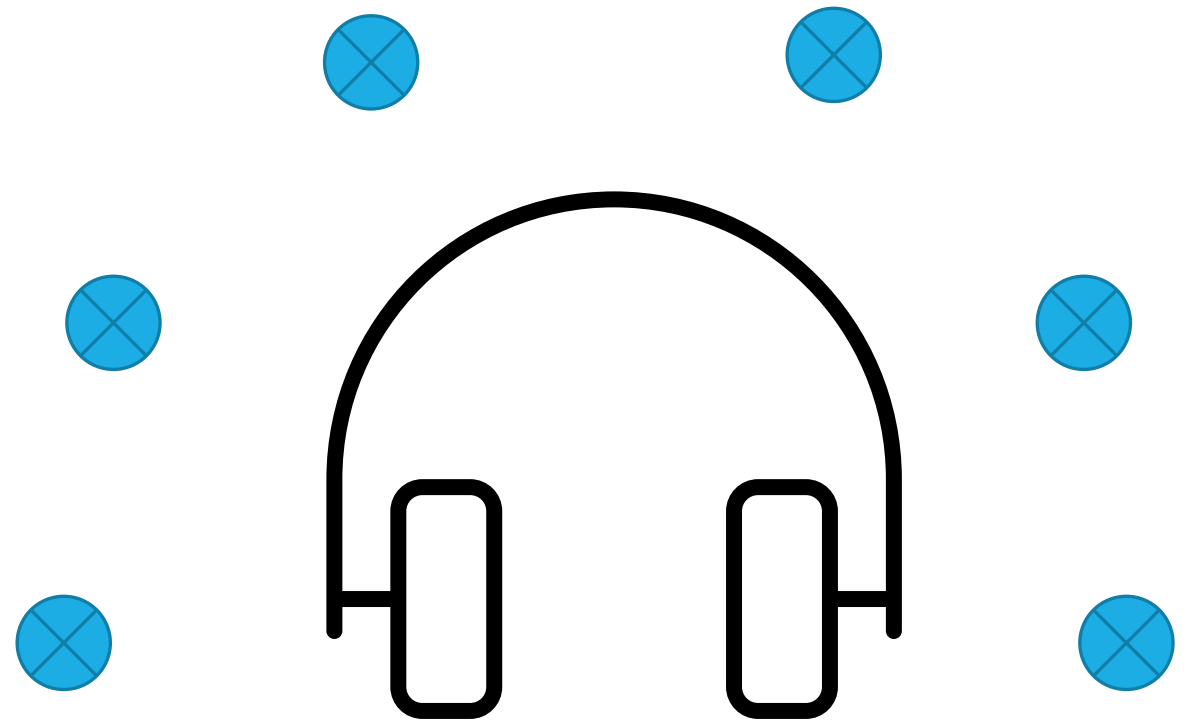
- Wprowadzanie przesunięcia czasowego (ITD) i różnicy natężenia dźwięku (ILD) do ścieżek dźwiękowych docierających do prawego i lewego ucha
- Uzyskanie obrazu przestrzennego w środku głowy



EKSTERNALIZACJA DŹWIĘKU

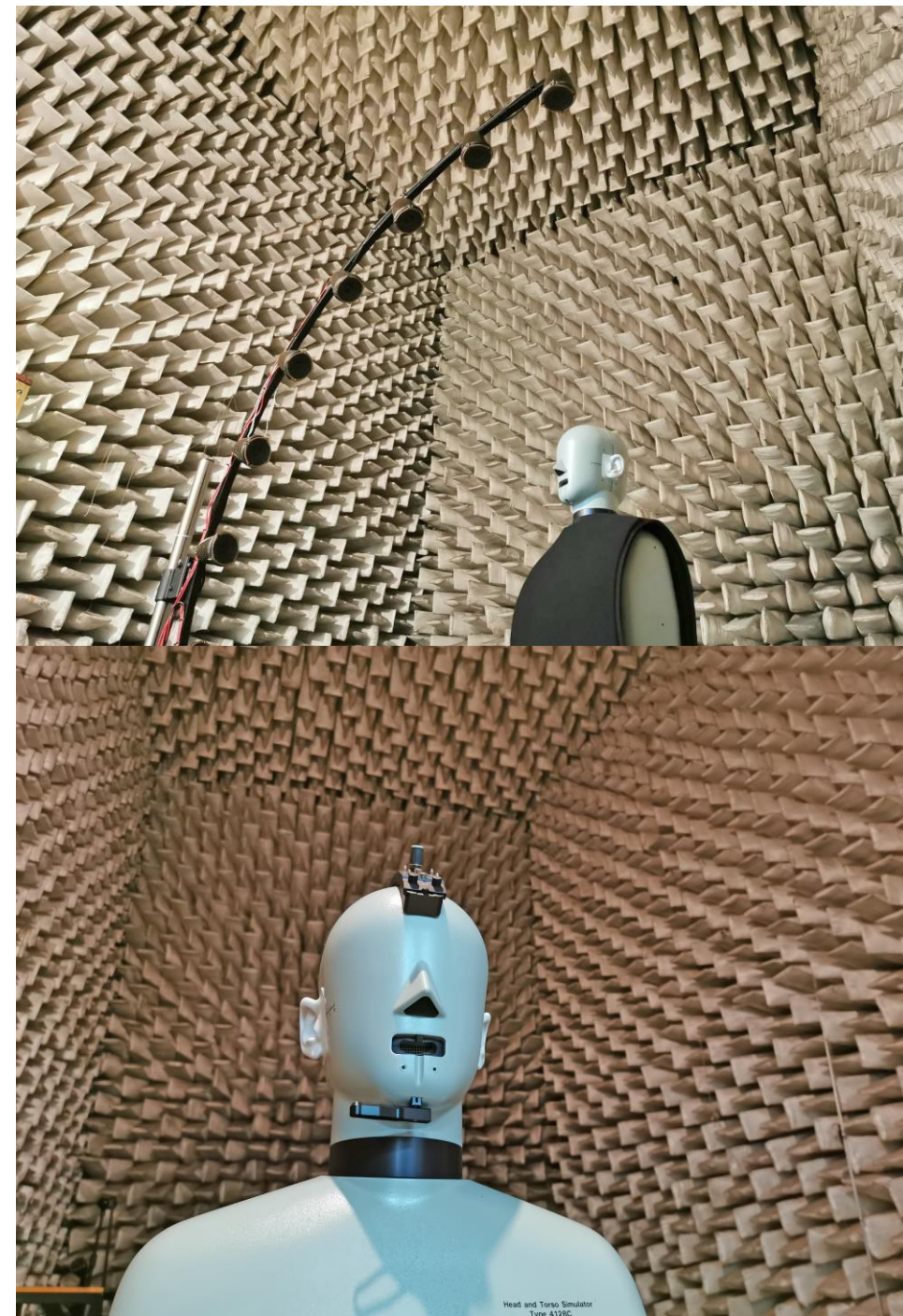


LATERALIZACJA



EKSTERNALIZACJA

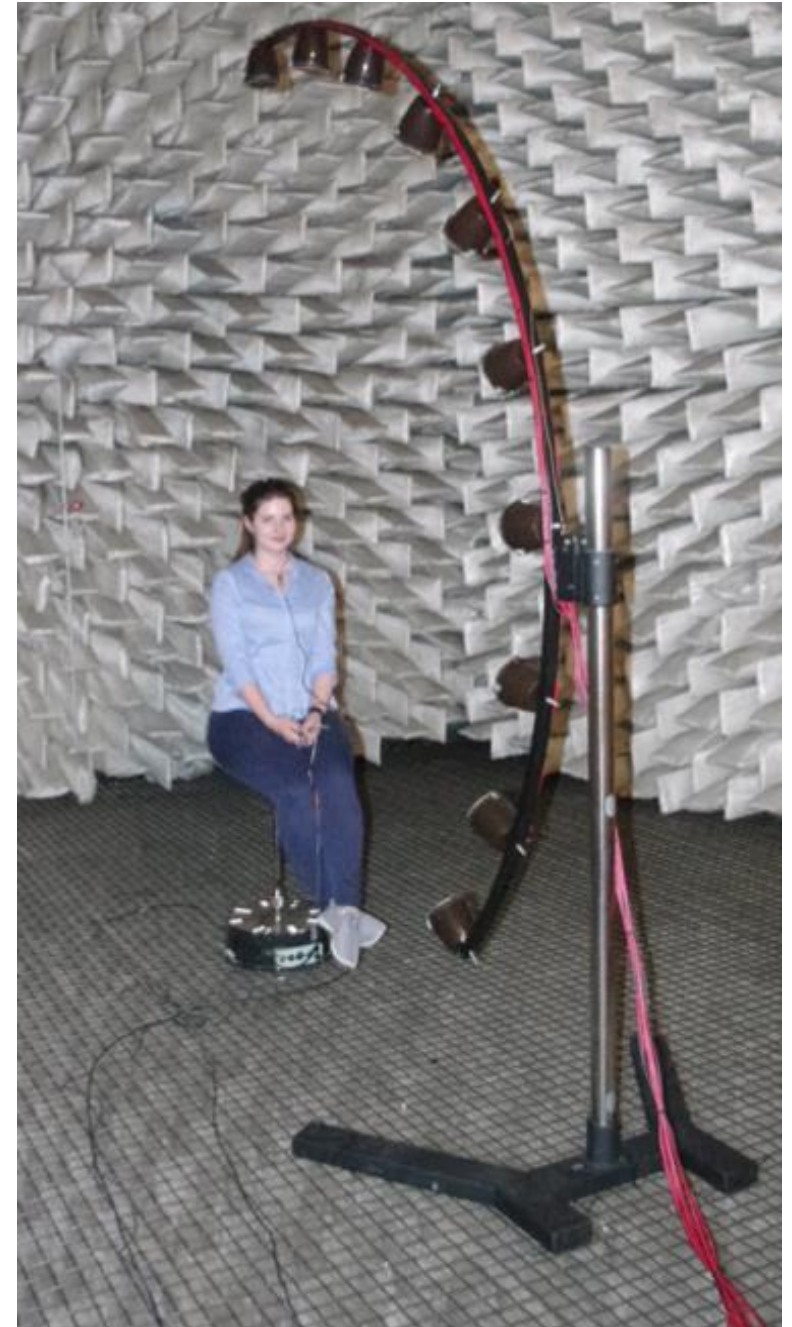
NAGRANIA BINAURALNE





HEAD RELATED TRANSFER FUNCTION

- **HRTF** - „funkcja przenoszenia głowy”
- określa wpływ głowy i ciała na dźwięk o danej częstotliwości docierający do ucha z danego kąta w przestrzeni



HRTF

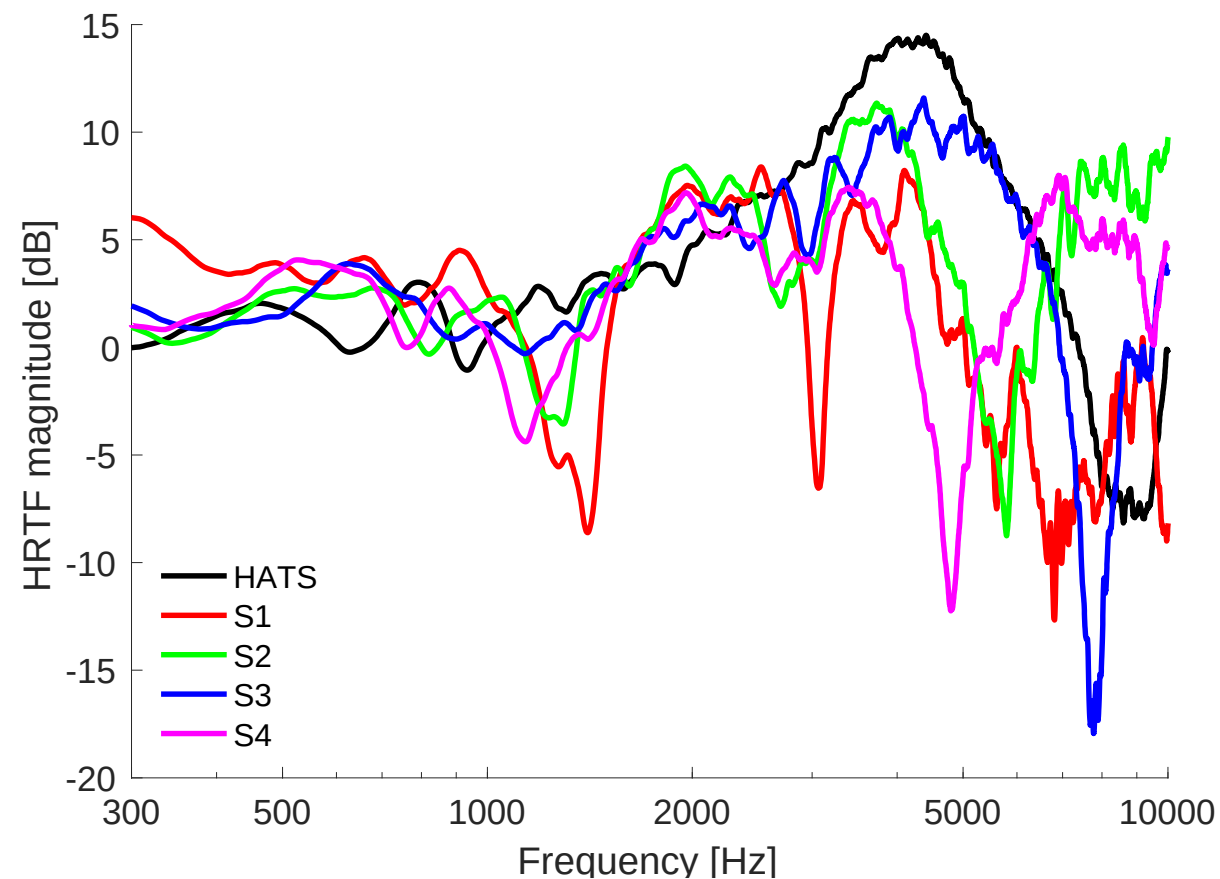
Zbiór funkcji dla poszczególnych kątów, przykładowo dla stanowiska w ZEA:

370 pozycji źródła dźwięku:

- 10 kątów elewacji ($-45^\circ \div 90^\circ$)
- 37 kątów w płaszczyźnie poziomej ($5^\circ \div 355^\circ + 0^\circ$)

Ogólnodostępna baza:

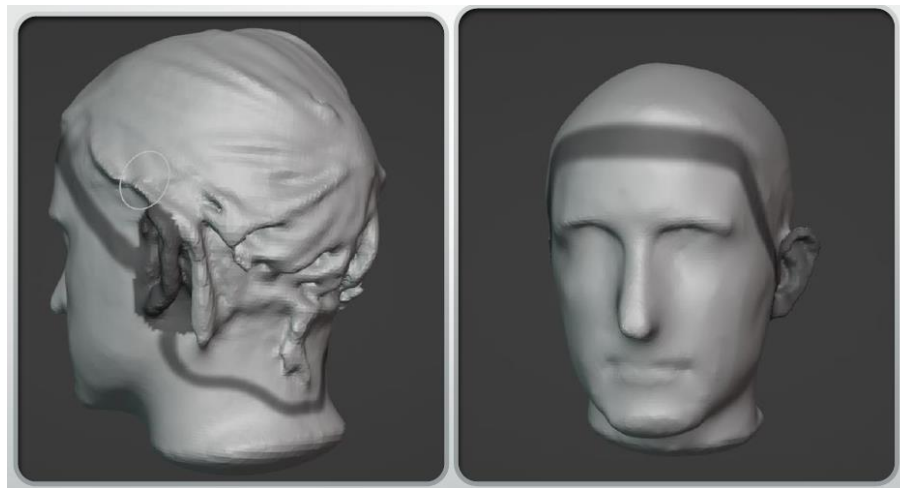
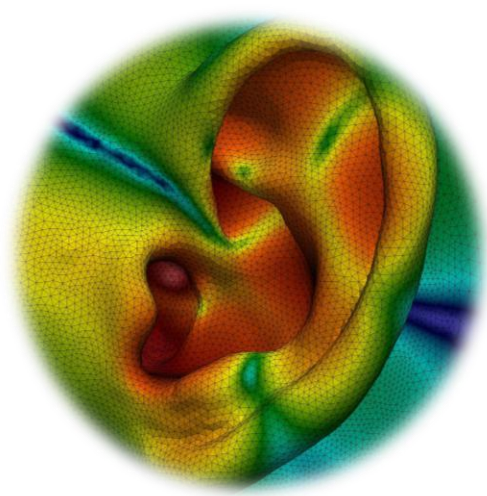
www.sofaconventions.org/mediawiki/index.php/Files



TEMATYKA HRTF

- Symulacja HRTF
 - na podstawie skanu 3D
 - na podstawie cech antropomorficznych
- Symulacja dźwięku docierającego z określonego kąta w przestrzeni
- Mechanizm i algorytmy śledzenia ruchu i pozycji głowy w celu przełączania funkcji HRTF na podstawie pozycji głowy

MODELOWANIE ZE SKANU 3D



MESH-2-HRTF

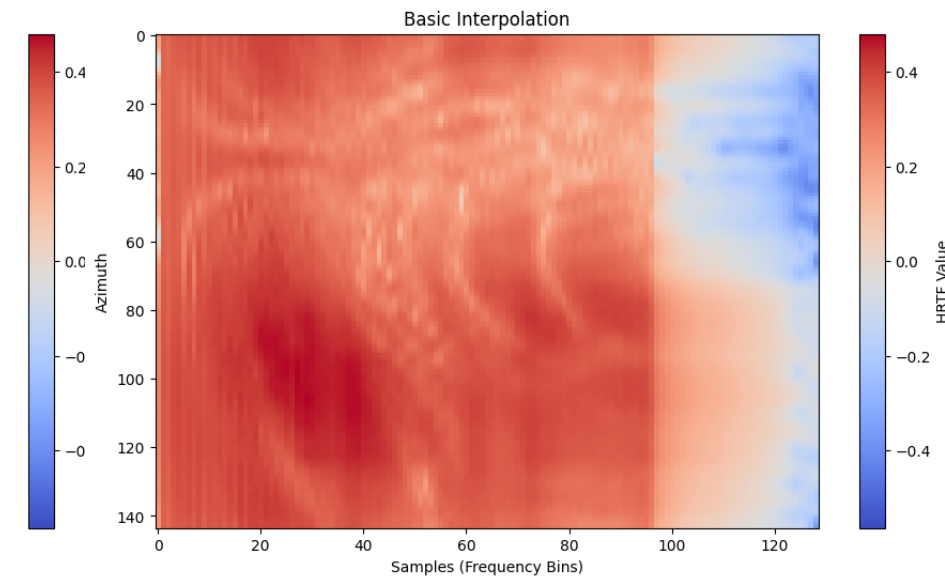
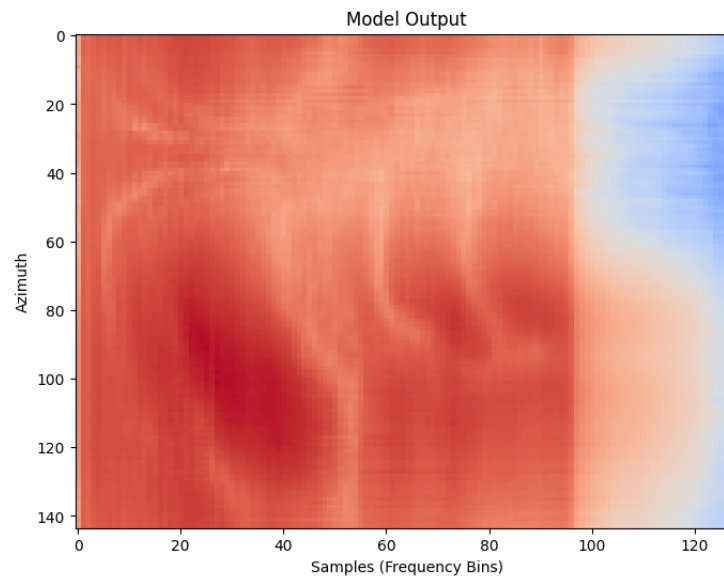
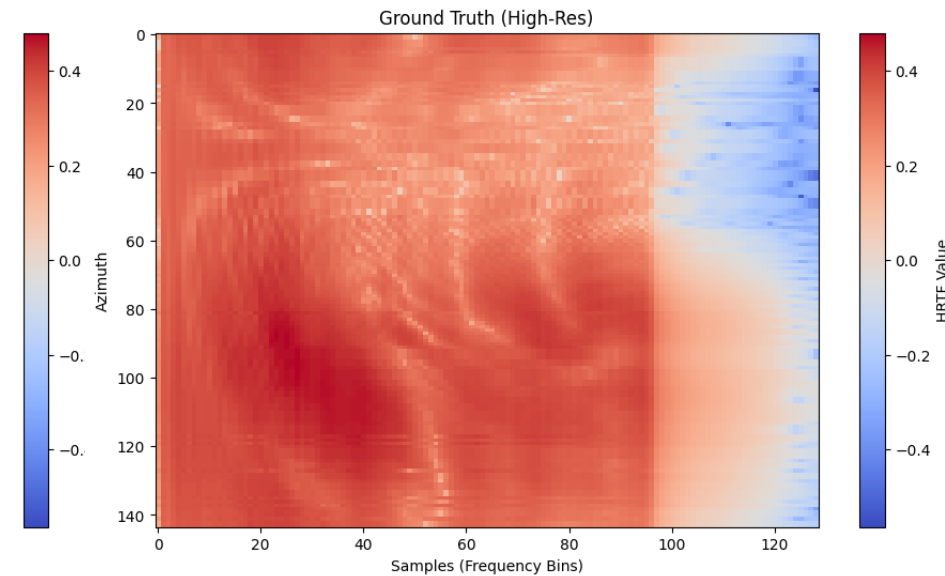
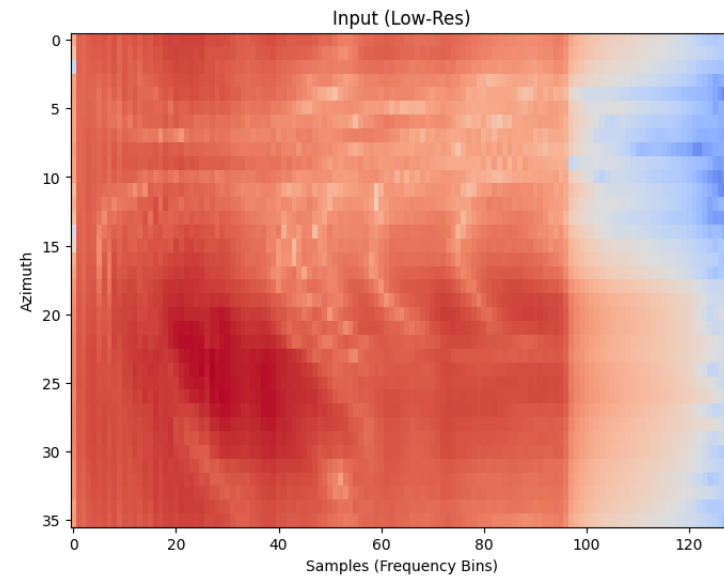
- Ziegelwanger, H., Kreuzer, W., and Majdak, P. (2015). "Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions," in *Proceedings of the 22nd International Congress on Sound and Vibration*, Florence, Italy.

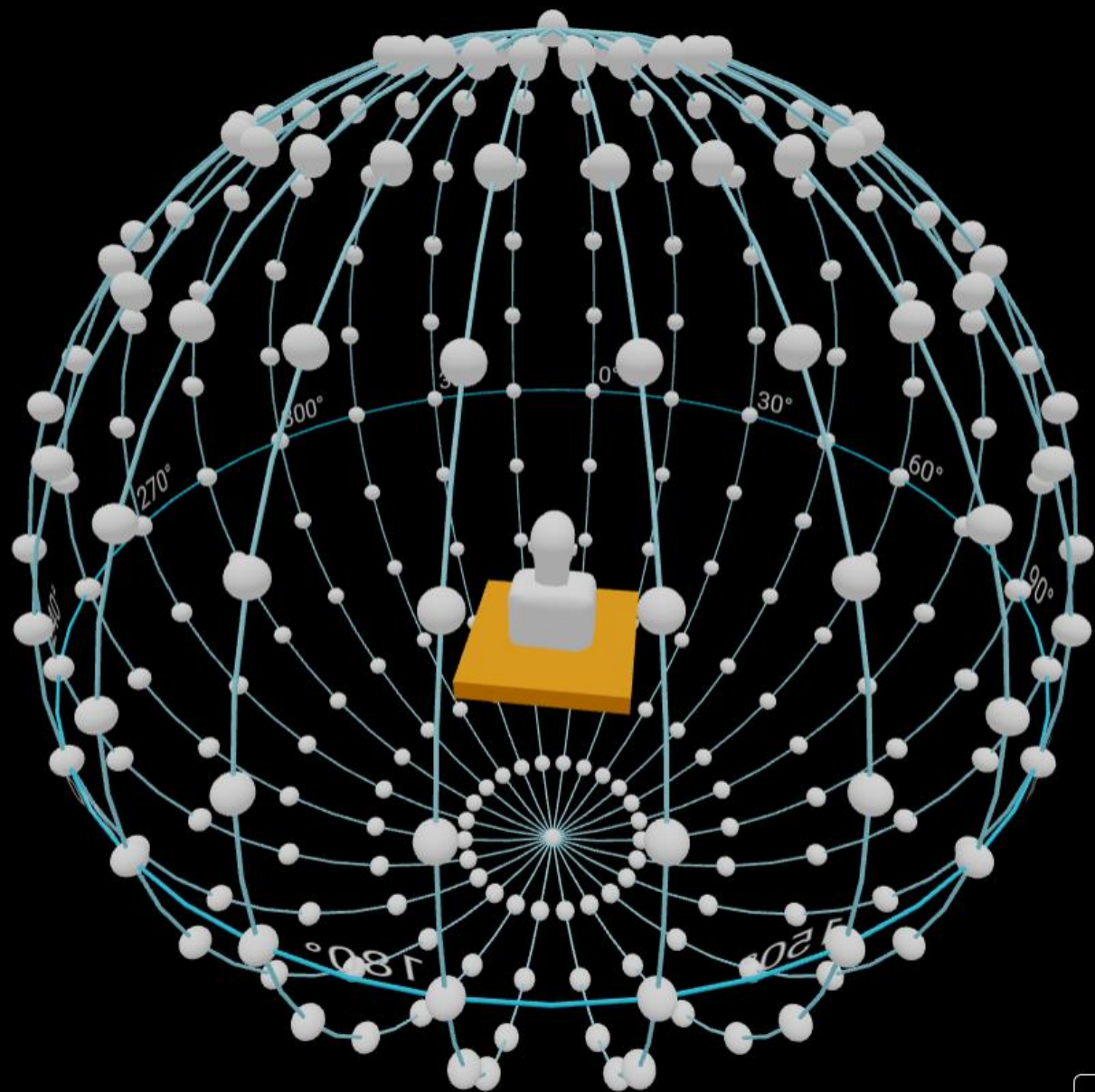


inż. Andrzej Budny, Stworzenie bazy pHRTF na podstawie skanów 3D

HRTF / AI

- wykorzystanie sztucznej inteligencji do poprawy jakości/rozdzielczości HRTF





Samples

V-60

Azimuth: 0
Elevation: -60

V30

Azimuth: 0
Elevation: 30

H15

Azimuth: 15
Elevation: 0

H60

Azimuth: 60
Elevation: 0

H270

Azimuth: 270
Elevation: 0

H120

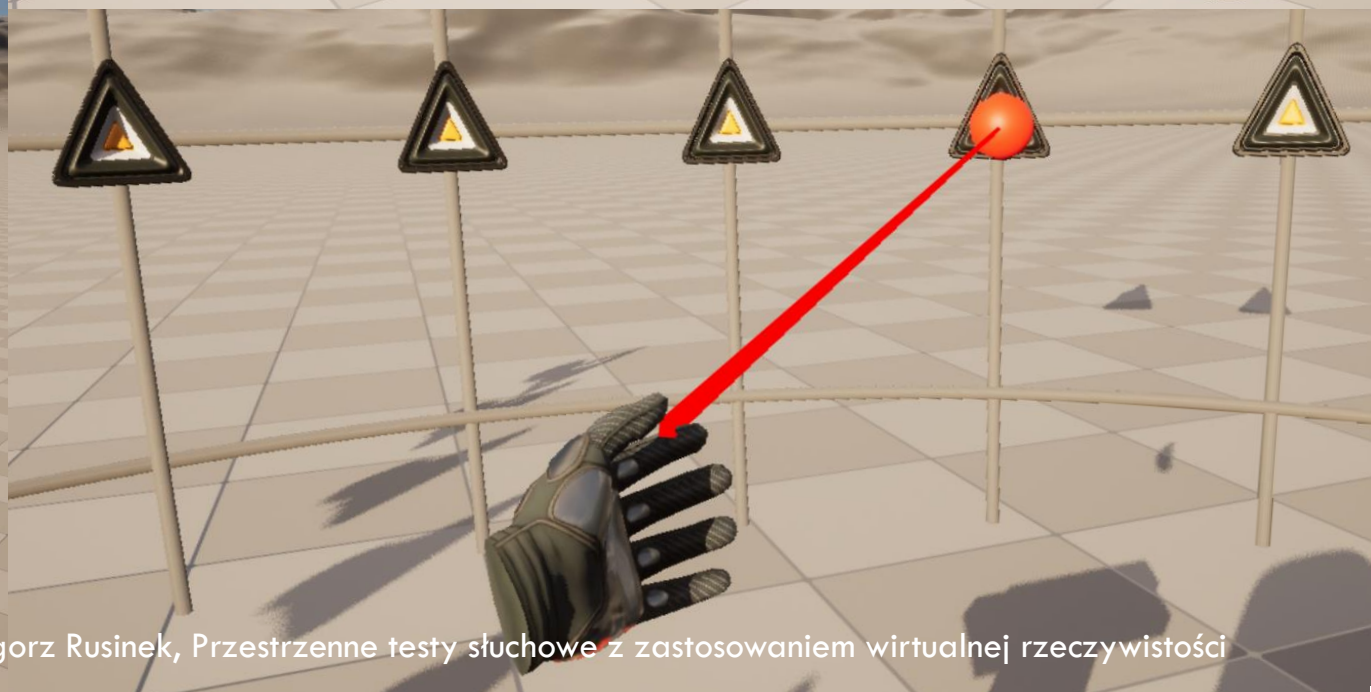
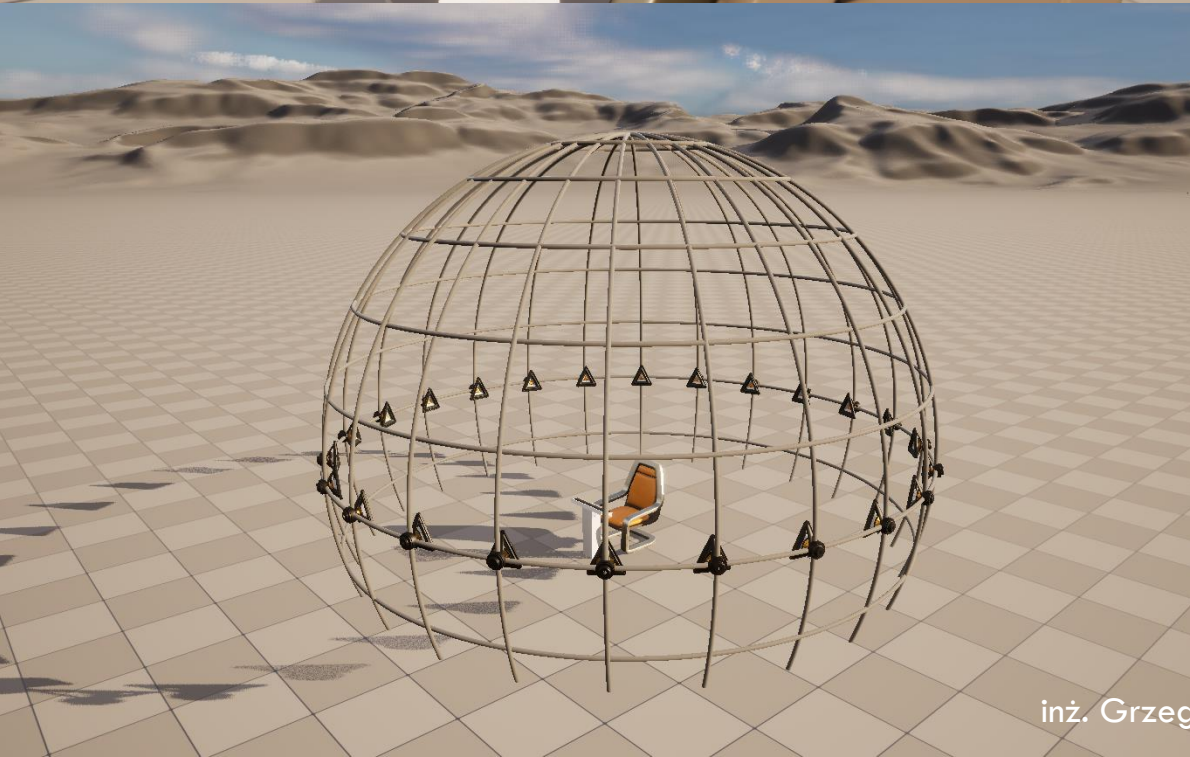
Azimuth: 120
Elevation: 0

H195

Azimuth: 195
Elevation: 0

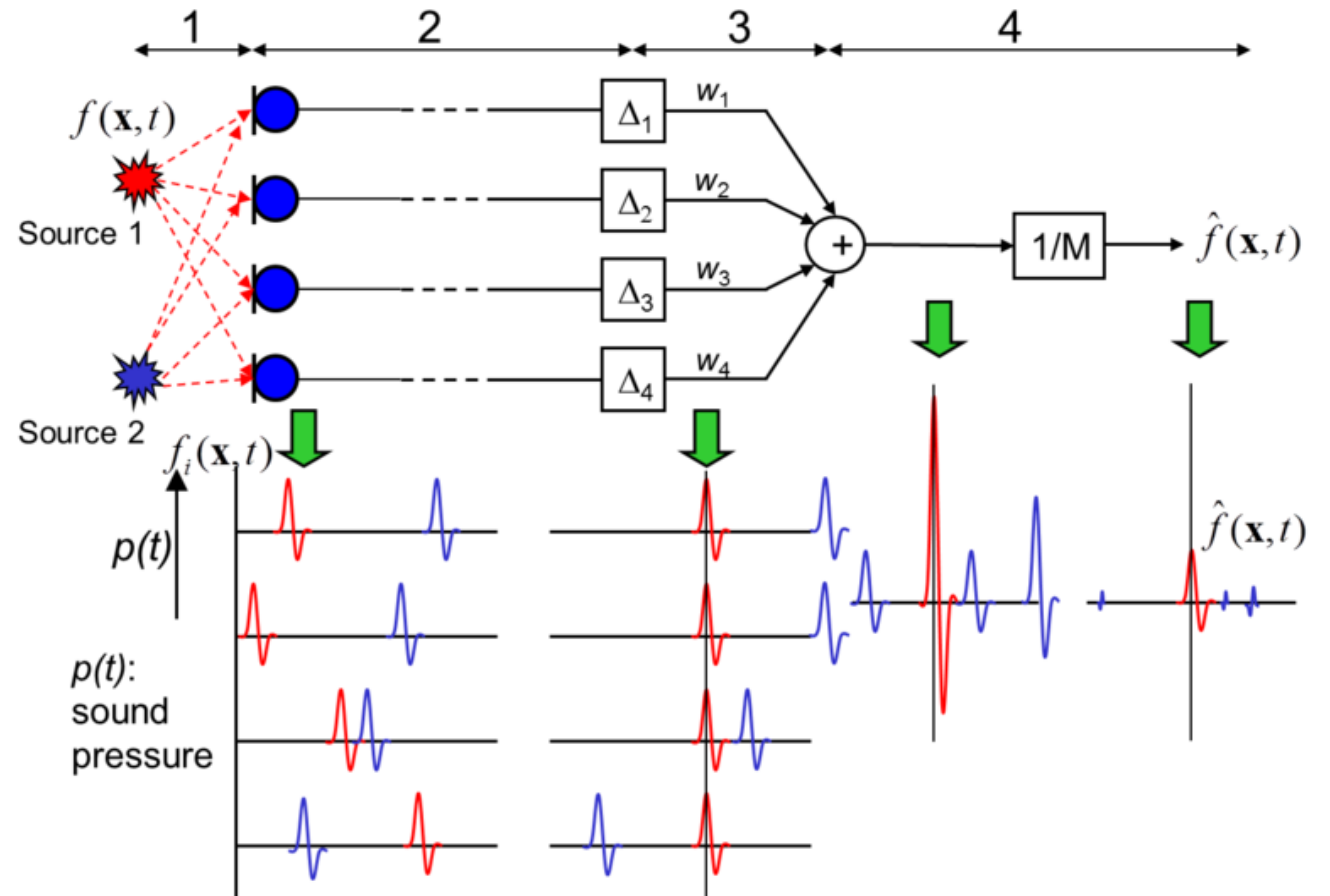
V-45

Azimuth: 0
Elevation: -45[Add new sample](#)



BEAMFORMING

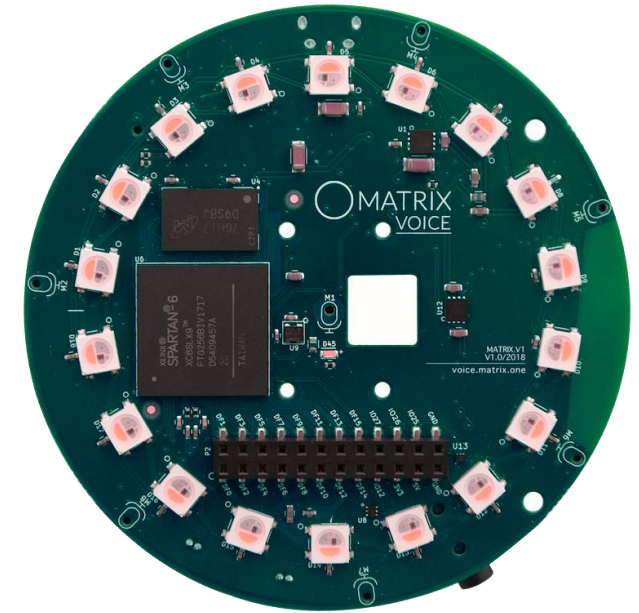
- „kształtowanie wiązki”
- wykorzystanie matrycy mikrofonowej
- przetwarzanie sygnałów z poszczególnych mikrofonów
- kształtowanie charakterystyki kierunkowej matrycy
- możliwość uzyskania silnie kierunkowej charakterystyki



BEAMFORMING

Zastosowania:

- usuwanie wpływów zakłóceń na sygnał użyteczny
- filtracja adaptacyjna w celu usunięcia niepożądanych sygnałów
- neurobeamforming – wykorzystanie sieci neuronowej i uczenia maszynowego do znalezienia wag



ZYLIA

- matryca 19-mikrofonowa, 48 kHz, 24 bit



AMBISONIA | AURALIZACJA

- ambisonia – rejestracja dźwięku w wielu płaszczyznach
- auralizacja – „wizualizacja dźwięku”

Przykład:

- ambisonia 3-rzędu
- matryca 19-mikrofonowa, 48 kHz, 24-bit



<https://www.youtube.com/watch?v=m5YgcCSvLp4>

<https://www.youtube.com/watch?v=I2thYPkBsPI> (nagranie 360°)

OCENA JAKOŚCI DŹWIĘKU

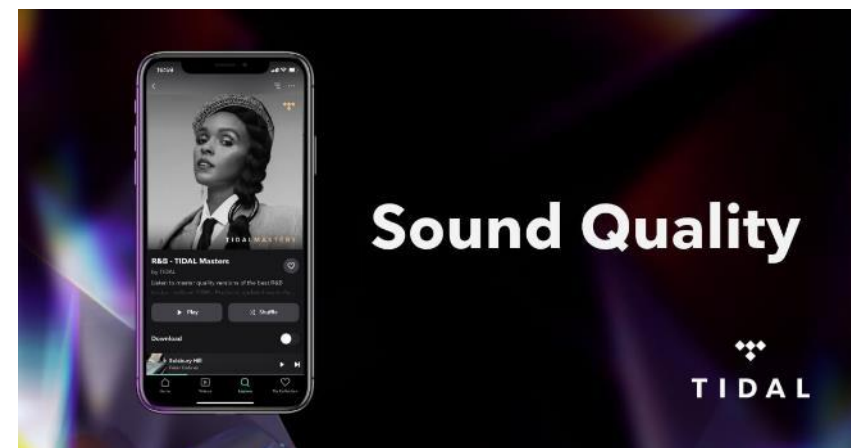
Problem:

w jaki sposób ocenić jakość dźwięku?

- Jakość urządzeń audio (głośników, odtwarzaczy)
- Jakość transmisji (zrozumiałość mowy lub muzyki)
- Jakość kodowania (mowy lub muzyki)
- Jakość dźwięku w pomieszczeniu

„Czy A brzmi dobrze, czy źle?”

„Czy lepiej brzmi A, czy B?”



METODY OCENY JAKOŚCI DŹWIĘKU

Pomiar parametrów fizycznych

tzw. „obiektywne”

Często nie dają wystarczającej informacji o jakości dźwięku.

Percepcyjne modele oceny

Wykorzystują modele zjawisk psychoakustycznych do programowego obliczania oceny jakości dźwięku.

Testy słuchowe

tzw. „subiektywne”

Opierają się na indywidualnym wrażeniu słuchaczy.

METODY OCENY JAKOŚCI DŹWIĘKU

Pomiar parametrów fizycznych

- SNR
- Zniekształcenia harmoniczne
 - THD
 - THD+N
- Zniekształcenia intermodulacyjne
- Przepływność

Percepcyjne modele oceny

- MOS
- PEAQ
- PESQ

Testy słuchowe

- Test AB
- Test ABC
- Test ABX
- MUSHRA

MOS - MEAN OPINION SCORE

- „subjective” MOS
- „objective” MOS

Ocena

Jakość

Zniekształcenia

5

Doskonała (Excellent)

Niesłyszalne (Imperceptible)

4

Dobra (Good)

Słyszalne, niedokuczliwe (Perceptible, but not annoying)

3

Średnia (Fair)

Lekko dokuczliwe (Slightly annoying)

2

Słaba (Poor)

Dokuczliwe (Annoying)

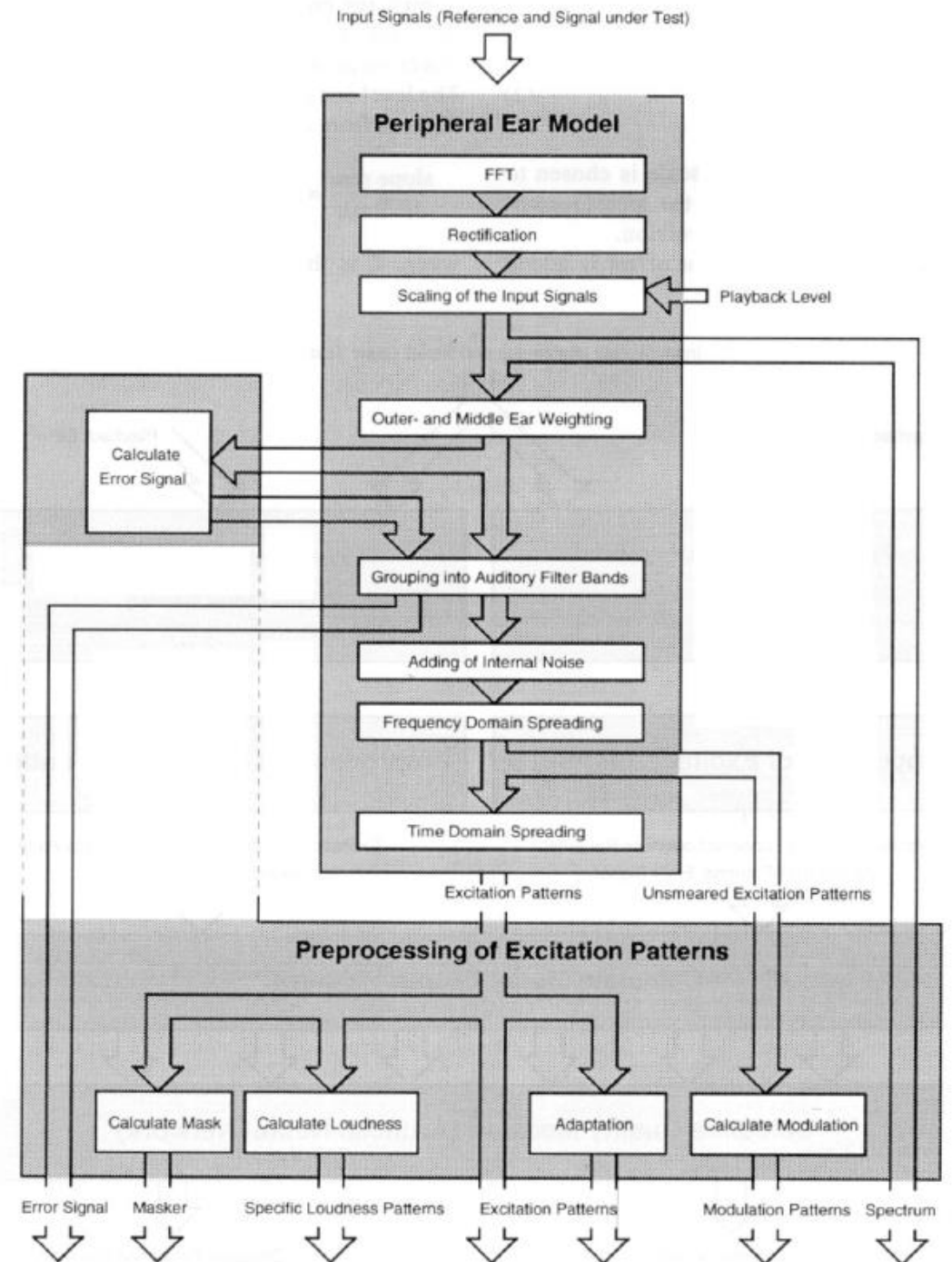
1

Zła (Bad)

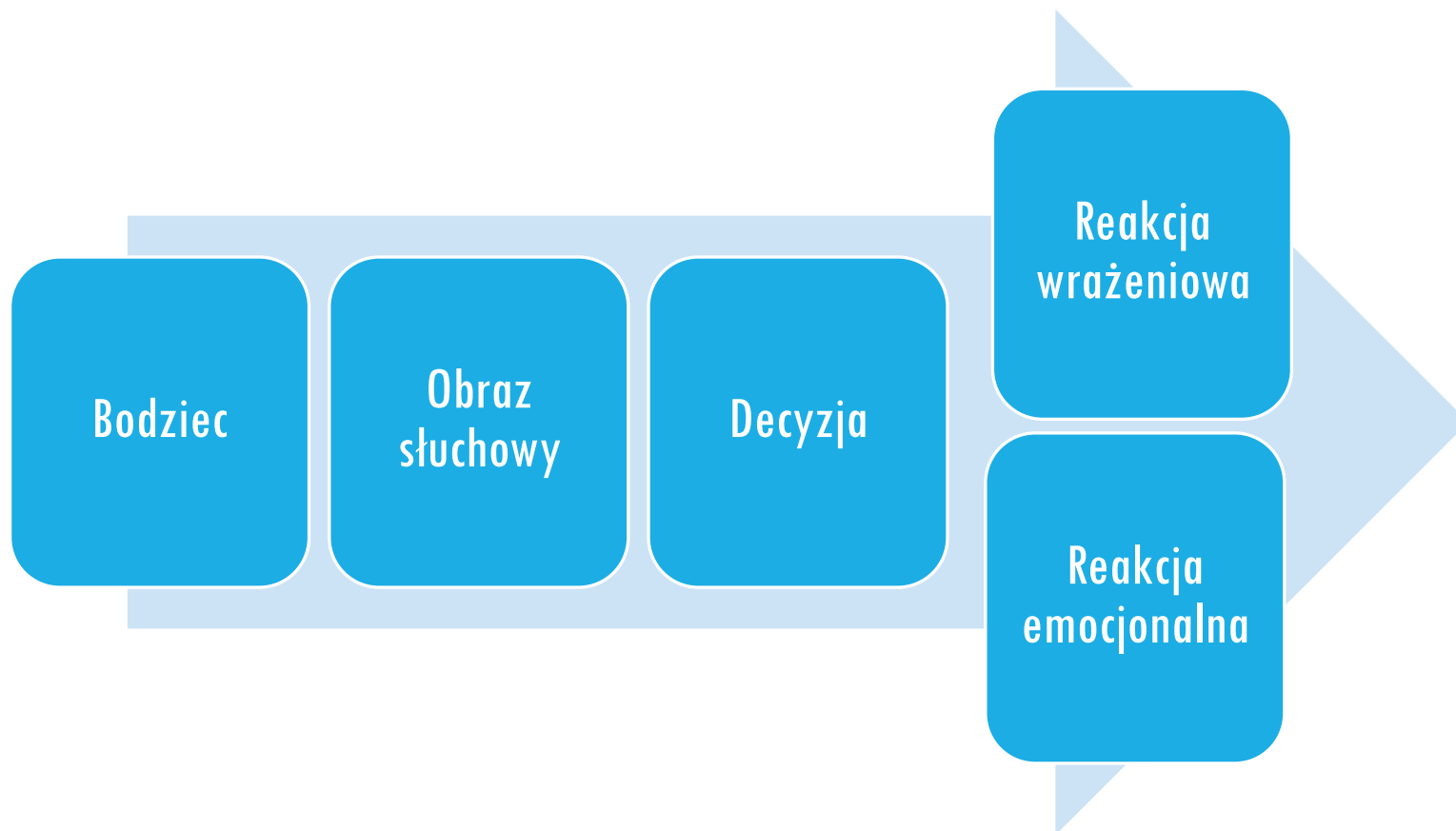
Bardzo dokuczliwe (Very annoying)

PERCEPTUAL EVALUATION OF AUDIO/SPEECH QUALITY (PEAQ/PESQ)

- Kompresja stratna dźwięków opiera się na wykorzystaniu modeli psychoakustycznych,
- To powoduje, że zniekształcenia, które wprowadza kompresja, są niesłyszalne dla słuchacza
- Zniekształcenia występują jednak w sygnale, więc obiektywne pomiary zniekształceń mogą nie oddawać właściwej informacji o wrażeniowej jakości.



SCHEMAT PERCEPCJI SŁUCHOWEJ



REAKCJE I PODSTAWA WOBEC DECYZJI

Reakcja wrazeniowa – efekt procesów fizjologicznych zachodzących w trakcie słuchania, związanych z cechami bodźca

Reakcja emocjonalna – odbicie nastawienia człowieka do obrazu słuchowego, związana z jego gustem i przyzwyczajeniami

Postawa wobec decyzji – niektórzy mówią „tak”, tylko gdy są absolutnie pewni, inni częściej

OBIEKTYWIZACJA OCEN

W celu **obiektywizacji** oceny słuchowej należy minimalizować wpływ **reakcji emocjonalnej** oraz **postawy wobec decyzji** na wyniki, poprzez:

- odpowiednie formułowanie zadań testowych
- dobór i szkolenie słuchaczy
- statystyczną obróbkę wyników

PLANOWANIE TESTU

Merytorycznie:

1. Sprecyzowanie, co podlega ocenie i jakiego zjawiska dotyczy:

- **detekcja**
- **dyskryminacja**
- **identyfikacja**
- **skalowanie**

2. Wybór

- **rodzaju,**
- **zakresu**
- **metody oceny**

Technicznie:

3. Konstrukcja testu

4. Wybór pomieszczenia
odsluchowego

5. Skompletowanie urządzeń toru
odsluchowego

6. Zorganizowanie grupy sluchaczy

7. Zebranie i analiza wynikow

OCENA ZESTAWU GŁOŚNIKOWEGO

Przykładowy test oceny **bezwzględnej**:

- Prezentujemy sygnały testowe
- Słuchacz musi ocenić jakość sygnału
- Musi sam znaleźć punkt odniesienia
- Nie uzyskamy dobrych wyników od laików, potrzebni eksperci

OSZACOWANIE LICZBOWE

PRZYKŁADOWE SKALE OCEN

Bezwzględna

Ocena

Jakość

Zniekształcenia

5

Doskonała

Niesłyszalne

4

Dobra

Słyszalne, niedokuczliwe

3

Średnia

Lekko dokuczliwe

2

Słaba

Dokuczliwe

1

Zła

Bardzo dokuczliwe

OCENA STOPNIA KOMPRESJI

Przykładowy test oceny **względnej**:

- Oceniana jest różnica między sygnałem oryginalnym a badanym
- Prostsze do oceny, podany sygnał odniesienia
- Test odpowiedni dla laików, ale trzeba sprawdzić ich wiarygodność

OSZACOWANIE LICZBOWE

PRZYKŁADOWE SKALE OCEN

Względna (porównawcza)

Ocena	Porównanie jakości
3	Znacznie lepsza
2	Lepsza
1	Nieco lepsza
0	Taka Sama
-1	Nieco gorsza
-2	Gorsza
-3	Znacznie gorsza

MUSHRA

Multiple **S**timuli with **H**idden **R**eference and **A**ncor

