

Lokalizacja i śledzenie źródeł dźwięku (I)

Tomasz Mąka <tmaka@wi.zut.edu.pl>

- 1876: L. Rayleigh, "On our Perception of the Direction of a Source of Sound"
- 1878: S.P. Thompson, "Phenomena of Binaural Audition"
- 1907: L. Rayleigh, "On our Perception of Sound Direction"
- 1934: S.S. Stevens, E.B. Newman, "The Localization of Actual Sources of Sound"
- 1940: H. Wallach, "The Role of Head Movements and Vestibular and Visual Cues in Sound Localization"
- 1948: L.A. Jeffress, "A Place Theory of Sound Localization"
- 1948: Masking Level Difference (J.C.R. Licklider / I.J. Hirsh)
- 1949: H. Wallach, E.B. Newman, M.R. Rosenzweig, "The Precedence Effect in Sound Localization"
- 1958: A.W. Mills, "On the Minimum Audible Angle"

[Yost2017]

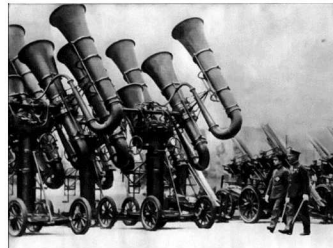
Lokalizacja i śledzenie źródeł dźwięku (I)

Katedra Architektury Komputerów i Telekomunikacji, Wydział Informatyki ZUT

Historia



https://en.wikipedia.org/wiki/Acoustic_location

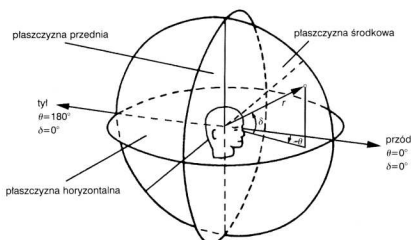


Przed wynalezieniem radaru (przed drugą wojną światową) stosowano takie systemy w celach obronnych przed nalotami z powietrza. Były to systemy pasywne, których zaletą było wykrywanie obiektów poza polem widzenia dzięki zjawisku dyfrakcji fal dźwiękowych.

Podstawowe pojęcia

Lokalizacja	Ocena kierunku i odległości do źródła dźwięku.
Słyszenie binauralne	Słyszenie dwuuszne.
Lateralizacja	Opis pozornego położenia źródła dźwięku wewnątrz głowy.
Stan diotyczny	Jednakowe bodźce docierające do obydwu uszu.
Stan dychotyczny	Różne bodźce docierające do obydwu uszu.

Układ współrzędnych do określania kierunków



[Moore1999]

- Płaszczyzna przednia przecina górne krawędzie wlotów do kanałów słuchowych zewnętrznych.
- Punkty należące do płaszczyzny środkowej są w jednakowej odległości od obydwu uszu.
- Miejsce przecięcia się wszystkich płaszczyzn znajduje się w środku głowy.
- Kierunek źródła dźwięku definiuje się za pomocą pary kątów: azymutalnego i biegunowego.

Blauert zaproponował układ współrzędnych zawierający trzy płaszczyzny (przódnią horyzontalną i środkową) do określania kierunków źródeł dźwięku względem głowy.

Teoria dupleksowa

- Zaproponowana w roku 1907 przez Lorda Rayleigha.
- Lokalizacja źródeł dźwięku o małych częstotliwościach realizowana jest na podstawie zmian międzysusznych różnic czasu, a źródeł o wysokich częstotliwościach z użyciem międzysusznych różnic natężenia.
- Teoria ta sprawdza się dobrze dla źródeł będącymi tonami prostymi.
- Dotyczy jedynie lokalizacji źródeł dźwięku w płaszczyźnie poziomej dookoła głowy.
- Nie jest ścisła dla dźwięków złożonych

Lokalizacja źródła dźwięku

- Zakres wartości różnic czasu może zmieniać się od 0μs (przód) do ~690μs (bok).
- Dla tonów, których długość fali jest porównywalna z odległością między uszami, różnice czasowe nie wpływają na zdolności lokalizacyjne.
- Jeśli długość fali dźwięków jest porównywalna z wymiarami głowy, to dochodzi do wystąpienia efektu dyfrakcji. W wyniku tego może dojść do zaniku cienia akustycznego.
- W przypadku lokalizacji tonów często występują problemy w ocenie kierunku przód/tył lub góra/dół.

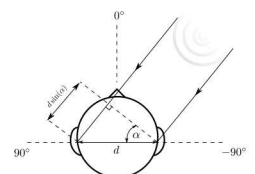
Lokalizacja źródła dźwięku w płaszczyźnie poziomej

Znając przesunięcie czasowe sygnałów można określić kąt, pod jakim dociera czoło fali płaskiej emitowanej przez źródło dźwięku:

$$\Delta t \cdot c = d \cdot \sin(\alpha)$$

gdzie:

- Δt – opóźnienie dźwięku docierające do uszu,
- c – prędkość dźwięku,
- d – odległość między uszami.



Kąt źródła dźwięku w płaszczyźnie poziomej określa zależność:

$$\alpha = \arcsin\left(\frac{\Delta t \cdot c}{d}\right)$$

Atrybuty lokalizacji binauralnej

Międzysusna różnica czasów – ITD (ang. Interaural time difference)

Atrybut ten wyznacza się na podstawie porównania sygnałów docierających do uszu. Do porównania sygnałów i określenia opóźnień między nimi stosuje się funkcję autokorelacji wzajemnej (lub jej wariant liczony w dziedzinie częstotliwości nazywany GCC-PHAT).

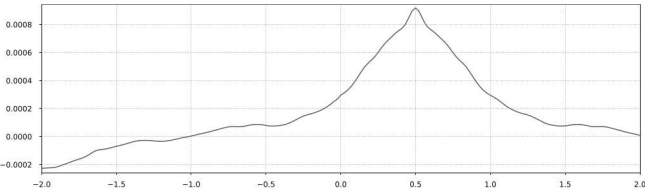
g(k) = \sum_{n=0}^{N-1-k} x_L(n) \cdot x_R(n - k)

W celu określenia opóźnienia poszukuje się największego pików w sygnale wynikowych, którego pozycja określa opóźnienie:

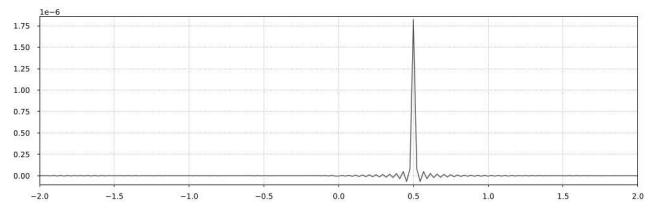
\Delta t = \underset{k}{\operatorname{argmax}}[g(k)]

Atrybuty lokalizacji binauralnej

Generalized Cross Correlation



Generalized Cross Correlation with Phase Transform (GCC-PHAT)



Atrybuty lokalizacji binauralnej

Międzysusna różnica poziomów – ILD (ang. Interaural level difference)

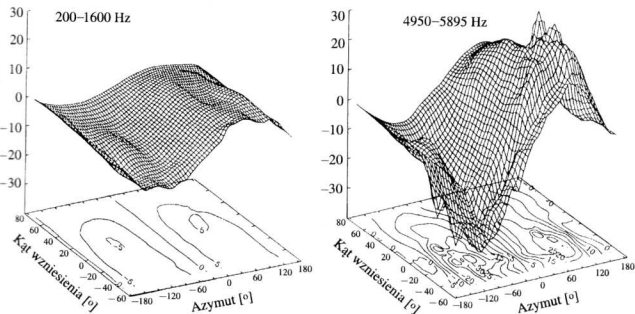
Atrybut ten jest obliczany na podstawie średniej mocy sygnałów docierających do uszu.

ILD = 10 \cdot \log_{10} \left(\frac{\sum_{k=0}^{K-1} x_L^2(k)}{\sum_{k=0}^{K-1} x_R^2(k)} \right)

Czasami ILD jest obliczane w dziedzinie częstotliwości.

Atrybuty lokalizacji binauralnej

Międzysusne różnice poziomu w dB.



[Ozimek2018]

Atrybuty lokalizacji binauralnej

Generalized Cross Correlation with Phase Transform (GCC-PHAT)

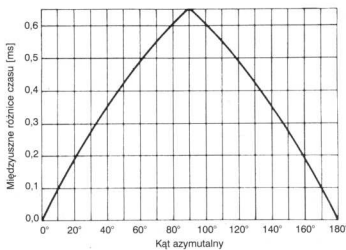
g(k) = \mathcal{F}^{-1} \left(\frac{X_L(f) \cdot X_R^*(f)}{|X_L(f) \cdot X_R^*(f)|} \right)

gdzie:

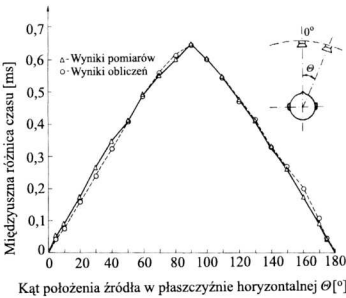
- \mathcal{F}^{-1}(\cdot) - odwrotna transformata Fouriera,
- X_L, X_R - transformaty Fouriera sygnałów x_L oraz x_R,
- X_R^* - wektor wartości sprzężonych do X_R.

Atrybuty lokalizacji binauralnej

Międzysusne różnice czasu.



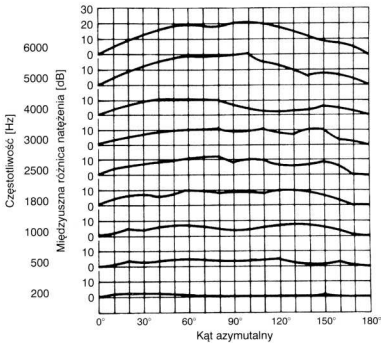
[Moore1999]



[Ozimek2018]

Atrybuty lokalizacji binauralnej

Międzysusne różnice natężenia dla tonów prostych.

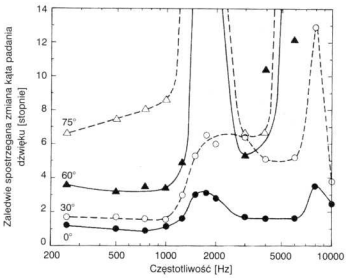


[Moore1999]

- Różnice międzysusne natężeń dla częstotliwości mniejszych niż 500Hz są prawie niezauważalne.
- Dla częstotliwości powyżej 5kHz różnice natężeń mogą wynosić powyżej 20dB.

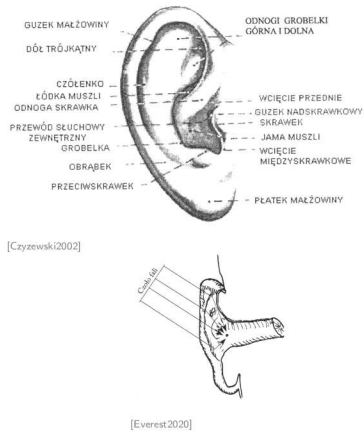
Zaledwie spostrzegana zmiana kąta padania

MMA (ang. Minimum Audible Angle)



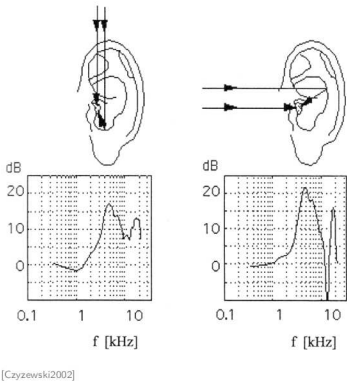
[Moore1999]

- Tony proste w płaszczyźnie horyzontalnej.
- Cztery kierunki odniesienia (0°, 30°, 60° i 75°)
- MMA jest najmniejsza dla dźwięków dochodzących z przodu słuchacza.

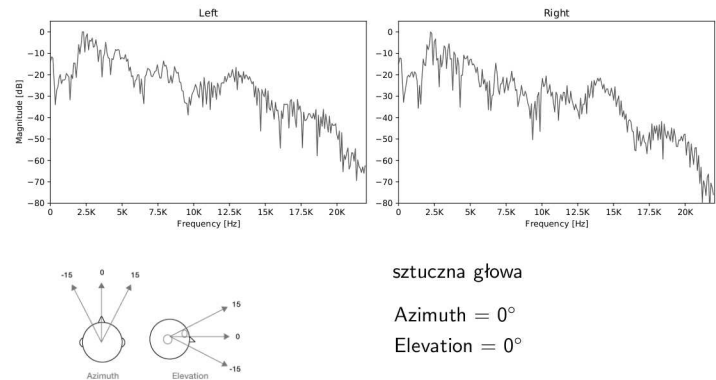


- Małżowina uszna jest źródłem informacji wykorzystywanych do oceny położenia w pionie.
- Stanowi również źródło informacji na temat rozróżnienia położenia tył/przód.
- Dostarcza więcej informacji dla sygnałów szerokopasmowych.
- Dla fali dźwiękowej o częstotliwości większej niż 6kHz występuje oddziaływanie z małżowiną uszną jako przeszkodą akustyczną.

- Zamknięcie przestrzeni uformowanych przez małżowinę zmniejsza zdolność lokalizacji (3, 8 i 10kHz).
- Małżowina modyfikuje widma docierających dźwięków. Modyfikacja ta zależy od kąta padania dźwięku względem głowy.
- Stosunek widma dźwięku emitowanego i odbieranego nazywa się funkcją przeniesienia głowy (HRTF).
- HRTF jest różne dla każdego kierunku w przestrzeni.

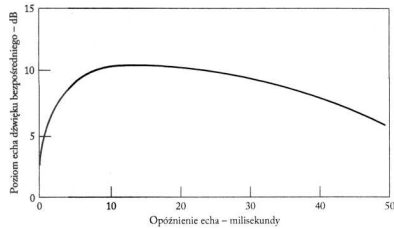


HRTF

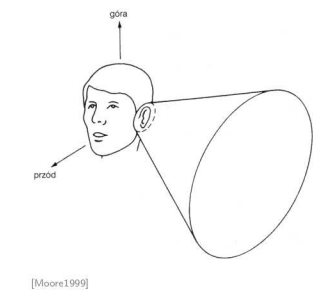


Efekt pierseństwa

- Gromadzenie odbić dochodzących w czasie do 35ms od dotarcia dźwięku bezpośredniego w układzie słuchowym.
- Dźwięk, który dociera pierwszy decyduje o percepcji kierunku kolejnych dźwięków.
- Opóźnienie echa w zakresie od 5 do 35ms musi mieć poziom wyższy ~10dB od poziomu dźwięku bezpośredniego żeby echo było rozróżnialne.



Stożek niepewności



- W przypadku braku małżowiny usznej oraz braku ruchu głowy, międzysusna różnica czasu nie wystarcza do jednoznacznego określenia położenia źródła dźwięku.
- Występuje zjawisko **stożka niepewności**: dowolne źródło dźwięku znajdujące się na powierzchni tego stożka wywołuje taką samą międzysusną różnicę czasu.
- Niejednoznaczności związane ze stożkiem niepewności lub z położeniem źródła dźwięku w kierunku pionowym można rozstrzygnąć poruszając głowę.

Efekt pierwszeństwa

- Do słuchacza oprócz dźwięku bezpośredniego, docierają fale odbite od obiektów.
- W roku 1949 Wallach i in. sformułowali **efekt pierwszeństwa** i opisali w jaki sposób układ słuchowy przetwarza dźwięki odbite.
- Jeśli czas pomiędzy krótkimi następującymi po sobie dźwiękami jest wystarczająco krótki, to są one odbierane jako pojedynczy dźwięk.
- Jeżeli dwa krótkie dźwięki są słyszalne jako jeden, to położenie całego dźwięku zależy głównie od położenia pierwszego z nich.
- Efekt ten występuje jedynie dla dźwięków o nieciągłym charakterze lub transjentów.

Efekt pierwszeństwa

- Dźwięk następujący ma mały, ale zauważalny wpływ na obserwowane zjawisko (wpływ na cały obraz dźwiękowy)
- Efekt nie występuje, gdy czas między dwoma dźwiękami jest mniejszy lub równy 1ms, a także gdy dźwięk następujący ma wyższe natężenie od poziomu dźwięku prowadzącego (~10dB). Również nie ma tego efektu przy zmianie warunków akustycznych.
- Najbardziej wyraźny efekt pierwszeństwa jest w sytuacji, gdy oba dźwięki są podobne jakościowo.
- Efekt ten przypisuje się rezultatowi przetwarzania ITD oraz ILD.
- Wystąpienie tego efektu następuje w skończonym czasie.

Estymacja odległości źródła dźwięku

Czynniki wpływające na estymację odległości:

- Charakterystyka źródła dźwięku.
- Poziom dźwięku.
- Ruchy głowy.
- Stosunek poziomu dźwięku bezpośredniego do odbitego.

Dla źródeł bliskich małe ruchy głową spowodują duże zmiany kąta w płaszczyźnie poziomej. W przypadku odległych źródeł zmiany te będą znikome.

Wraz ze wzrostem stosunku poziomu dźwięku bezpośredniego do odbitego wywołuje efekt oddalania się źródła dźwięku.

[Everest2020]	F. A. Everest, K. C. Pohlman, "Podręcznik akustyki", Wydawnictwo Sonia Draga, Katowice, 2020
[Kleczkowski2013]	P. Kleczkowski, "Percepcja dźwięku", Wydawnictwa AGH, Kraków, 2013
[Moore1999]	B. C.J. Moore, "Wprowadzenie do psychologii słyszenia", Wydawnictwo PWN, Warszawa-Poznań, 1999
[Ozimek2018]	E. Ozimek, "Dźwięk i jego percepcja - aspekty fizyczne i psychoakustyczne", Wydawnictwo PWN, Warszawa, 2018
[Czyżewski2002]	A. Czyżewski, B. Kostek, H. Skarżyński, "Technika komputerowa w audiologii, foniatryi i logopedii", Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2002
[Yost2017]	W. A. Yost, "History of sound source localization: 1850-1950", 173rd Meeting of Acoustical Society of America and 8th Forum Acusticum, Boston, Massachusetts, 25-29 June 2017