

Sluchové priestorové vnímanie a jeho neurálny kód

Norbert Kopčo

Katedra kybernetiky a umelej inteligencie FEI TU, 040 01 Košice
Hearing Research Center, Boston University, Boston, MA
E-mail: kopco@tuke.sk

Abstrakt. Schopnosť človeka detekovať zvuk maskovaný šumom je ovplyvnená priestorovou konfiguráciou zdrojov zvuku a šumu. Táto štúdia prezentuje výsledky psychofyzikálneho experimentu, ktorý meral túto priestorovú závislosť pre širokospektrálne a nízkofrekvenčné stimuly. Na analýzu výsledkov je použitý jednoduchý model, ktorý predpokladá, že centrálny sluchový systém používa na detekciu zvukov vždy ten periférny sluchový kanál, ktorý poskytuje najlepší odstup signál šum; ale nepredpokladá, že by centrálny sluchový systém bol schopný informácie z viacerých kanálov kombinovať. Výsledky naznačujú, že takýto model generuje správne predikcie v prípade, že sa spektrálna šírka signálov nemení, ale nedokáže popísať závislosť výsledkov na spektrálnej šírke stimulov. Preto je možné usudzovať, že centrálny sluchový systém je schopný pri priestorovom počúvaní kombinovať informácie z viacerých periférnych sluchových kanálov.

1. Úvod

Schopnosť človeka porozumieť reči v komplikovanom (napr. hlučnom) prostredí je veľmi dobrá. Tento fenomén je označovaný aj ako „efekt koktailovej párty“ (angl. „cocktail-party effect“, Blauert, 1997). Predpokladá sa, že efekt koktailovej párty súvisí s efektom „priestorového odmaskovania“ (angl. „spatial release from masking“, SRM), t.j. s tým, že človek dokáže veľmi presne detekovať zvuk v prítomnosti šumu za predpokladu, že zdroj zvuku je priestorovo oddelený od zdroja šumu (Saber et al., 1991). Cieľom tejto štúdie je charakterizovať niektoré vlastnosti neurálnych mechanizmov, na ktorých je priestorové odmaskovanie založené.

Predošlé psychofyzikálne štúdie priestorového odmaskovania ukázali, že SRM je určené dvomi faktormi (pozri napr. Kopčo a Shinn-Cunningham, 2000 pre SRM čistých tónov, alebo Good, Gilkey a Ball, 1997 pre SRM širokospektrálnych zvukov): SRM vysokofrekvenčných stimulov je determinované primárne energetickými faktormi, t.j. tým, že pri zmene polohy jedného zo zdrojov zvuku sa zmení pomer signál šum (SNR) počutého zvuku. Odmaskovanie nízkofrekvenčných stimulov je okrem energetických faktorov spôsobené aj binaurálnymi faktormi, t.j. schopnosťou sluchového systému porovnať čas a intenzitu, s ktorými bol zvuk zaznamenaný ľavým a pravým uchom, a na základe tohto porovnania zvuky ďalej separovať. Žiadna z predošlých štúdií neanalyzovala relatívnu dôležitosť energetických a binaurálnych faktorov pre SRM širokospektrálnych podnetov.

Zvukové podnety sa v periférnej časti sluchovej dráhy spracúvajú v paralelných frekvenčných kanáloch, ktoré vytvárajú distribuovaný neurálny kód zvukového podnetu. Existujú dve alternatívy ako môže centrálny sluchový systém spracovávať informácie z periférnych kanálov pri prezentácii širokospektrálnych zvukov: 1) centrálny sluchový systém si zvolí jediný periférny sluchový kanál (ten s najpriaznivejším SNR) a SNR v tomto úzkospektrálnom kanáli určuje SNR potrebné na detekovanie širokospektrálneho zvuku; 2) centrálny sluchový systém kombinuje

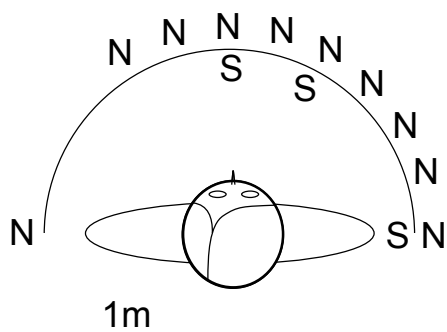
informácie z viacerých kanálov, čo spôsobuje, že prahové SNR potrebné pre detekciu širokospektrálneho zvuku je nižšie ako SNR potrebné na detekciu v jednotlivých (úzkospektrálnych) kanáloch. Výsledky porovnanie elektrofyziologických meraní neurálnej aktivity v mačacom collicule inferior s ľudskými behaviorálnymi dátami (Lane et al., 2003) naznačujú, že alternatíva o použití jediného periférneho sluchového kanálu je pravdepodobnejšia.

Cieľom tejto štúdie je overiť hypotézu o použití jediného periférneho kanálu centrálnym sluchovým systémom pre širokospektrálne a nízkofrekvenčné stimuly. Štúdia porovnáva experimentálne merané hodnoty priestorového odmaskovania širokospektrálnych a nízkofrekvenčných zvukov (Kopčo et al., 2003) s hodnotami predikovanými modelom založeným na jedinom periférnom sluchovom kanáli.

2. Experimentálne metódy

Kopčo et al. (2003) merali priestorové odmaskovanie nízkofrekvenčných a širokospektrálnych stimulov u piatich poslucháčov so zdravým sluchom. Priestorová poloha signálu (t.j. zvuku, ktorý má poslucháč detekovať) a šumu (t.j., maskovacieho zvuku) bola simulovaná vo virtuálnom bežeckom priestore vytvorenom použitím priestorových prenosových funkcií (head-related transfer functions, HRTF). Simulované polohy zdroja signálu a šumu sú zobrazené na Obrázku 1.

Experiment mal tri samostatné časti, v prvej sa meralo priestorové odmaskovanie širokospektrálnych, v druhej nízkofrekvenčných zvukov. V tretej časti sa postupne menila šírka pásma pre signál a šum na 0° s cieľom priamo určiť schopnosť človeka kombinovať informácie z viacerých periférnych kanálov. Signál bol tvorený sériou frekvenčných prechodov (angl. chirp) s periodicitou 40 Hz a dĺžkou 200 ms. Širokospektrálny signál mal frekvenčné pásmo 300-12000 Hz, nízkofrekvenčný signál mal pásmo 300-1500 Hz. Maskovací šum bol biely šum s frekvenčným pásmom 200-14000 Hz (širokospektrálny) alebo 200-2000 Hz (nízkofrekvenčný). Na určenie prahovej hodnoty SNR sa v experimente adaptívne menila hlasitosť šumu použitím „3-down, 1-up, 2-interval 3-alternative forced choice“ procedúry. Táto procedúra určuje prahové SNR, pri ktorom pravdepodobnosť, že poslucháč správne detekuje signál je 79.4%. Stimuly boli prezentované cez slúchadlá v zvukovo



Obrázok 1. Simulované polohy signálu (S) a šumu (N)

izolovanej miestnosti.

3. Model

Na analýzu dát a na stanovenie, či centrálny sluchový systém kombinuje informácie z viacerých periférnych sluchových kanálov alebo používa len jeden kanál, bol implementovaný jednoduchý model. Tento model (Obrázok 2) pozostáva zo 60 logaritmicke rozmiestnených gamatónových filtrov (Johannesma, 1972) na ucho. Pre každú priestorovú konfiguráciu signálu a šumu model vypočíta efektívnu energiu na výstupe každého filtra (zvlášť pre signál a zvlášť pre šum). Model potom predpokladá, že prah je stanovený kanálom (filtrom) s najväčším pomerom signál šum (SNR) z pomedzi všetkých 120 filtrov. Model má jeden voľný parameter, konštantu, ktorá sa nastaví tak, aby predikcia zodpovedala nameranému prahovému SNR pre konfiguráciu keď je signál a šum na tej istej pozícii.

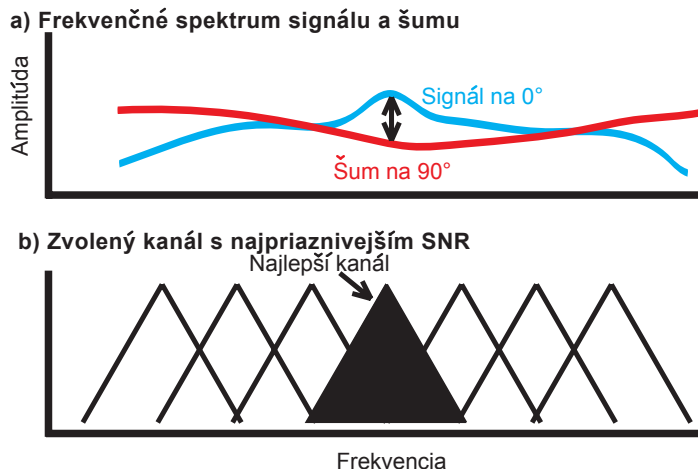
4. Výsledky

Experimentálne meranie priestorového odmaskovania pozostávalo z troch častí. V prvých dvoch častiach sa meralo prahové SNR pre 21 priestorových konfigurácií signálu a šumu (viď Obr. 1), v prvej pre nízkofrekvenčné a v druhej pre širokospektrálne stimuly. V tretej časti boli signál aj šum umiestnené na 0° , a merala sa závislosť prahového SNR na spektrálnej šírke signálu.

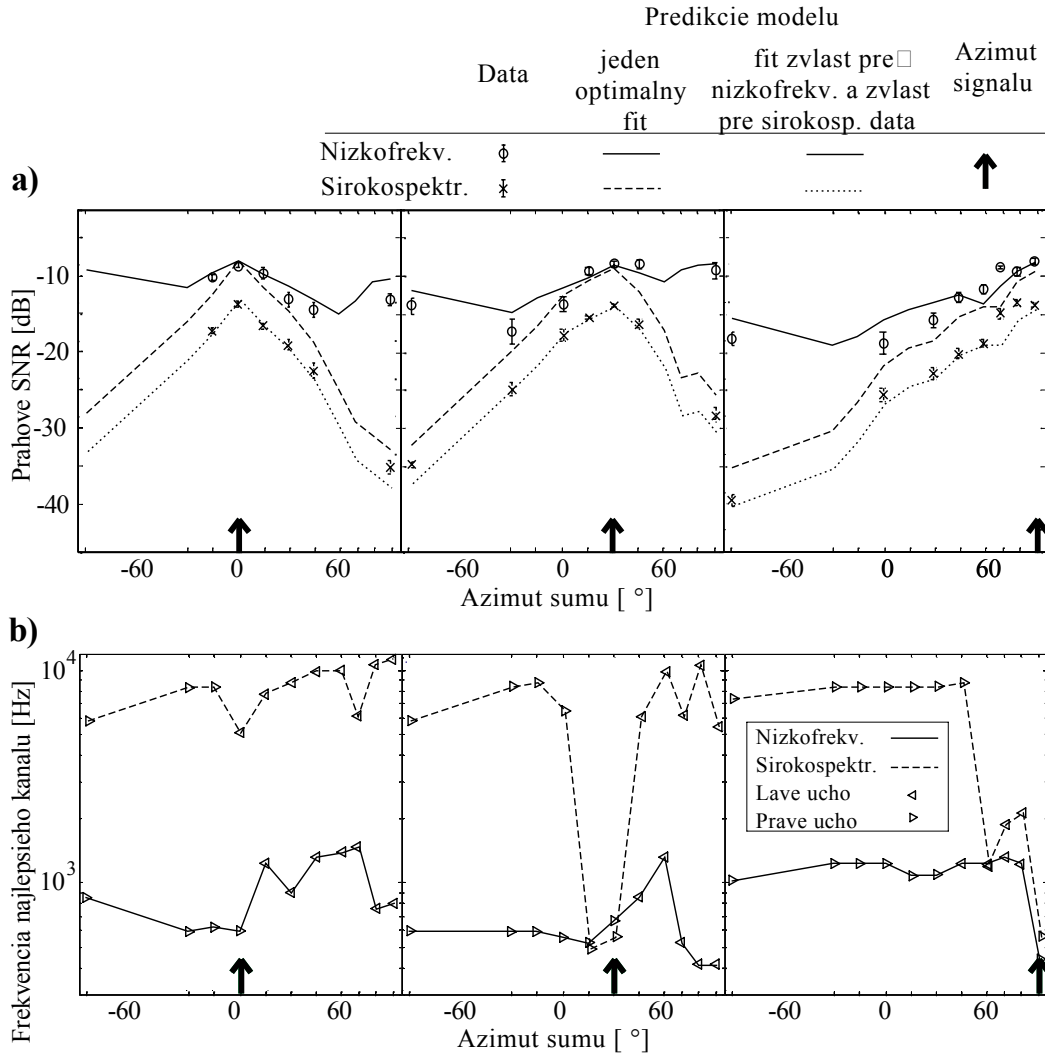
4.1 Širokospektrálne a nízkofrekvenčné stimuly

Obrázok 3a ukazuje merané a predikované prahové SNR ako funkciu azimutu šumu pre tri rôzne nemenné polohy signálu. Priestorové odmaskovanie je viditeľné pre širokospektrálne (x) aj pre nízkofrekvenčné (o) stimuly. Veľkosť priestorového odmaskovania (pokles prahového SNR v porovnaní s prahom keď je signál a šum na rovnakom mieste) je väčší pre širokospektrálne (30 dB) než nízkofrekvenčné stimuly (12 dB). Obrázok zároveň ukazuje, že detekovateľnosť širokospektrálnych signálov je lepšia ako u nízkofrekvenčných signálov pre všetky priestorové konfigurácie signálu a šumu.

Čiary v obrázku 3a reprezentujú dve sady predikcií vygenerovaných testovaným modelom V prvej sade bol voľný parameter modelu nastavený tak, aby model presne predpovedal priemernú hodnotu prahových SNR pre konfigurácie s nízkofrekvenčným



Obrázok 2. Schéma modelu použitého na analýzu dát. Model zvolí z pomedzi 120 kanálov kanál s najpriaznivejším pomerom signál šum (SNR)



Obrázok 3. Priestorové odmaskovanie širokospektrálnych a nízkofrekvenčných stimulov. a) Merané (stredná hodnota a štandardná chyba) a predikované prahové SNR ako funkcia azimutu šumu. b) Frekvencia a ucho (ľavé alebo pravé) najlepšieho kanálu.

signálom a šumom na rovnakom azimute. Následne boli vygenerované predikcie pre nízkofrekvenčné (plné čiary) aj širokospektrálne (prerušované čiary) dáta. V druhej sade bol parameter modelu nastavený zvlášť pre nízkofrekvenčné dáta (plné čiary) a zvlášť pre širokospektrálne dáta (bodkované čiary).

Keď sa parameter modelu nastaví zvlášť pre nízkofrekvenčné dáta (plná čiara a symboly „o“) a zvlášť pre širokospektrálne dáta (bodkovaná čiara a symboly „x“), predikcie modelu sú veľmi presné. To znamená, že model založený na voľbe najvhodnejšieho kanálu je dostatočný na presné predikovanie priestorového odmaskovania, ak sa prezentované stimuly nemenia.

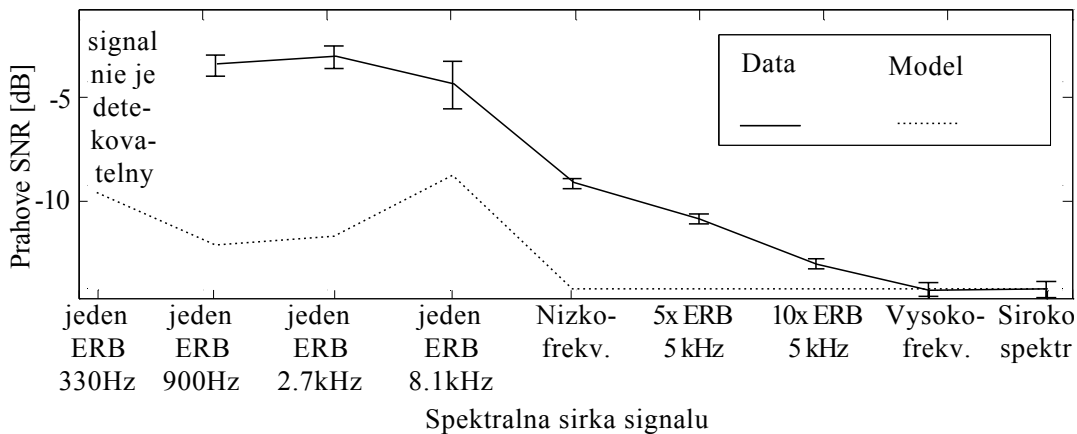
Ku každej predikcii z Obr. 3a zodpovedá bod na Obr. 3b, ktorý ukazuje ucho (ľavé/pravé) a stredovú frekvenciu najlepšieho kanálu, ktorý bol modelom zvolený na vygenerovanie predikcie. Najlepší kanál pre širokospektrálne dáta má väčšinou stredovú frekvenciu okolo 8 kHz (prerušovaná čiara na Obr. 3b), t.j. v oblasti, kde binaurálne

faktory neovplyvňujú detektabilitu zvukov. Naviac, uvažovaný jednoduchý model predikuje širokospektrálne dáta presne napriek tomu, že nepočíta s binaurálnymi faktormi. Tieto výsledky naznačujú, že priestorové odmaskovanie širokospektrálnych signálov je určené hlavne energetickými faktormi na vysokých frekvenciách. Pre nízkofrekvenčné stimuly sa najlepší kanál nachádza v oblasti, kde binaurálny systém môže napomáhať detekcii (plná čiara na Obr. 3b). Naviac, model často predikuje menšie množstvo priestorového odmaskovania než je viditeľné z nízkofrekvenčných dát, čo naznačuje že energetické aj binaurálne faktory ovplyvňujú priestorové odmaskovanie nízkofrekvenčných signálov.

Keď sa na predikciu nízkofrekvenčných a širokospektrálnych dát použije jediný parameter (plné a prerušované čiary v Obr. 3a), model predikuje, že prahové SNR by mali byť v konfiguráciách so signálom a šumom blízko seba približne rovnaké, nezávisle na spektrálnej šírke stimulov. Táto predikcia je v rozpore s dátami, ktoré ukazujú, že detektabilita nízkofrekvenčných signálov je vždy horšia ako u širokospektrálnych signálov. To znamená, že model uvažujúci len s jedným (najlepším) kanálom nedokáže predikovať priestorové odmaskovanie stimulov s rozdielnou spektrálnou šírkou.

4.2 Závislosť prahového SNR na spektrálnej šírke signálu

Tretia časť experimentu bola zameraná na bližšie objasnenie závislosti prahového SNR na spektrálnej šírke signálu. V tejto časti bolo odmeraných niekoľko prahových SNR pre signál a šum na azimute 0° a pre rôzne spektrálne šírky signálu. Spektrálna šírka signálu sa menila v násobkoch ekvivalentnej pravouhlej šírky (equivalent rectangular bandwidth, ERB, Moore, 1997), kde 1 ERB zodpovedá šírke ľudského periférneho sluchového filtra. Merané prahové SNR (plná čiara v Obr. 4) vykazujú sústavný pokles so zväčšujúcou spektrálnou šírkou signálu. V protiklade k tomu, model predikuje takmer nemennú hodnotu prahového SNR, s malým nárastom len pre stimuly s najužším spektrom (jeden ERB alebo nízkofrekvenčné dáta). Tento výsledok ďalej podporuje hypotézu, že centrálny sluchový systém je schopný kombinovať informácie z viacerých periférnych sluchových kanálov.



Obrázok 4. Závislosť prahového SNR na spektrálnej šírke signálu pre signál a šum umiestnené na 0° : dáta a predikcie modelu

5. Záver a diskusia

Prezentované dáta naznačujú, že priestorové odmaskovanie môže viesť k zlepšeniu detekovateľnosti širokospektrálnych signálov až o 30 dB a nízkospektrálnych signálov až o 12 dB. Priestorové odmaskovanie širokospektrálnych signálov je určené energetickými faktormi na vysokých frekvenciách, zatiaľčo odmaskovanie nízkofrekvenčných signálov je determinované nízkofrekvenčnými energetickými faktormi ako aj malým príspevkom binaurálnych faktorov. Ďalej dáta ukazujú, že detekovateľnosť maskovaných signálov sa zlepšuje s ich spektrálnou šírkou.

Jednoduchý model založený na nájdení najlepšieho periférneho sluchového kanálu je schopný presne predikovať priestorové odmaskovanie širokospektrálnych alebo nízkofrekvenčných stimulov, ale nedokáže predikovať odmaskovanie pre oba typy stimulov súčasne. Vo všeobecnosti tento model nedokáže popísať závislosť prahového SNR na spektrálnej šírke stimulov.

Tieto výsledky naznačujú, že centrálny sluchový systém je schopný kombinovať informácie z viacerých periférnych kanálov, a preto model, ktorý by mal byť schopný predikovať priestorové odmaskovanie stimulov s rôznym spektrom musí obsahovať mechanizmus pre toto kombinovanie.

Fyziologické výsledky (Lane et al., 2003) naznačujú, že aktivita neurónov v collicule inferior zodpovedá jednoduchému modelu, ktorý nepredpokladá kombináciu informácií z viacerých periférnych kanálov. Preto sa dá predpokladať, že táto kombinácia sa uskutočňuje v časti sluchovej dráhy, ktorá je centrálnejšia než je colliculus inferior.

Literatúra

- Good MD, Gilkey RH, Ball JM (1997) "The relation between detection in noise and localization in noise in the free field." In RH Gilkey and TR Anderson (Eds, Binaural and Spatial Hearing in Real and Virtual Environments. Lawrence Erlbaum, Mahawah, NJ, pp. 349-376.
- Johannesma PIM (1972). "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in IPO Symposium on Hearing Theory, edited by B. L. Cardozo, E. de Boer, & R. Plomp (IPO, Eindhoven, The Netherlands), pp. 58-69.
- Kopčo N, Shinn-Cunningham BG (2000). "Spatial unmasking of nearby pure-tone sources in a simulated anechoic environment," J. Acoust. Soc. Am. 107(5 Pt. 2), 2849 (Presented at the 139th meeting of the Acoustical Society of America).
- Kopčo N, Lane CC, Shinn-Cunningham BG. (2003) "Spatial unmasking of chirp trains in a simulated anechoic environment: Behavioral results and model predictions." Association for Research in Otolaryngology Abstract #935
- Lane CC, Delgutte B, Colburn HS (2003) "A Population of ITD-Sensitive Units in the Cat Inferior Colliculus Shows Correlates of Spatial Release from Masking," Association for Research in Otolaryngology Abstract #706
- Moore BCJ (1997) An Introduction to the Psychology of Hearing (4e). San Diego, CA: Academic Press.