

GUI PRO DEMONSTRACI PRINCIPŮ BINAURÁLNÍ LOKALIZACE ZDROJŮ ZVUKU

O. Glaser, F. Rund, D. Štorek

Katedra radioelektroniky, Fakulta elektrotechnická,
České vysoké učení technické v Praze

Abstrakt

Tento článek představuje uživatelská rozhraní vytvořená v prostředí Matlab, která umožňují ilustraci základních principů binaurální lokalizace. První z popisovaných rozhraní umožňuje subjektivně sledovat změnu lokalizace zvukového podnětu při intenzitních a/nebo časových rozdílech mezi signály reprodukoványými do jednotlivých uší posluchače. Druhé rozhraní ilustruje komplexnější přístup při simulaci filtrace zvuku směrově závislou přenosovou funkcí vztaženou k hlavě (HRTF - Head Related Transfer Function).

1 Úvod

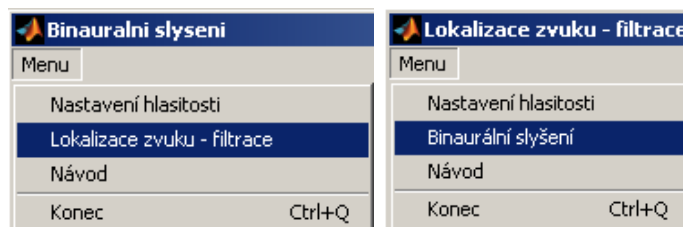
Lokalizace zdroje zvuku patří mezi často studované jevy, jejichž využití má význam nejen v oblasti multimediální techniky, zejména hudebního a herního průmyslu, ale např. i v oblasti asistivních technologií, jako pomůcka při orientaci zrakově znevýhodněných občanů. Detailní studium principů lokalizace má samozřejmě význam i pro oblast zpracování a záznamu zvuku, kdy umožňuje např. vývoj efektivnějších algoritmů pro kompresi zvuku.

Pro rozvoj zmíněných oblastí a vývoj nových aplikací je vhodné disponovat nástroji, které umožňují demonstrovat a ověřovat principy binaurální lokalizace zdrojů zvuku. V dalším textu budou představena dvě uživatelská rozhraní, která umožňují seznámení i další studium zmíněných principů.

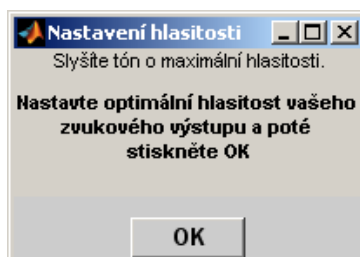
2 Principy binaurální lokalizace

Problematiku binaurální lokalizace přehledně shrnuje např. publikace [1]. *Lokalizaci* zdroje zvuku definuje jako určení jeho směru a odhad vzdálenosti, dále zavádí pojem *lateralizace* jako vnímání zvuku na spojnici uší uvnitř hlavy. Pokud není rozdíl mezi signály v obou uších ani v intenzitě, ani v čase (resp. je pod určitým prahem), je zvuk *lateralizován* doprostřed hlavy. Je zřejmé, že pokud zdroj signálu je např. vlevo od posluchače, dosáhne zvukový signál levého ucha dříve než pravého, tj. existuje určitý časový rozdíl (tzv. *ITD* – *Interaural Time Difference*), a zvuk je vnímán zleva. Stejně tak dochází k rozdílu v intenzitě (*ILD* – *Interaural Level Difference*), zvuk v přivráceném uchu je hlasitější než v odvráceném. ITD se uplatňuje především na nízkých kmitočtech, ILD spíše na kmitočtech vyšších, kde se více uplatňuje akustický stín hlavy a časový rozdíl již není rozlišitelný. Demonstrací vlivu ILD a ITD na lateralizaci zdroje zvuku se zabývá první z prezentovaných GUI, obr. 3.

Výše popsané rozdíly ovšem vysvětlují zejména problematiku lokalizace v horizontální rovině. Vjem výšky (elevace) zdroje, rozlišení mezi zvukem přicházejícím zepředu a zezadu, a vjem vzdálenosti, se nedá vysvětlit jen s použitím ILD a ITD. Tyto jevy závisí na ovlivnění příchozího zvuku zejména odrazy od různých částí těla subjektu (např. ramen, hlavy a ušního boltce), které je interpretováno převážně na základě zkušenosti subjektu s konkrétním typem zvuku, např. řečového signálu. Dále je lokalizace umožněna také bezděčnými pohyby hlavy subjektu a v neposlední řadě i porovnáním se zrakovým vjemem.



Obrázek 1: Detail menu obou programů.



Obrázek 2: Dialog pro nastavení hlasitosti (společné pro oba programy).

Zmíněné ovlivnění příchozího zvuku odrazy od různých částí hlavy subjektu lze popsat směrově závislou filtrací. Pro každý směr je tedy možné určit příslušnou přenosovou funkci (*HRTF - Head Related Transfer Function*), popř. impulzní odezvu (*HRIR - Head Related Impulse Response*) – viz např. publikace [2]. Pro účely modelování příchodu zvuku z určitého směru tedy stačí provést konvoluci tohoto zvuku s HRIR odpovídající příslušnému směru. Tento přístup demonstruje druhé z prezentovaných GUI, obr. 4, které využívá HRIR z databáze [3].

3 Implementace v prostředí Matlab

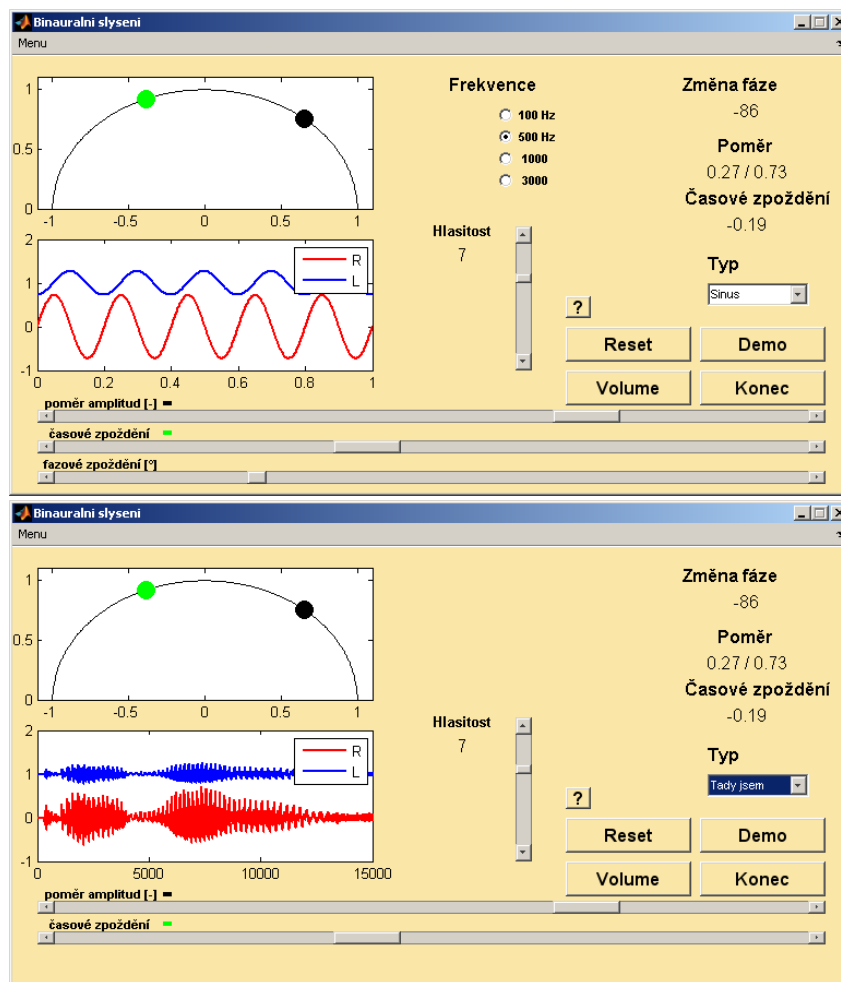
Obě zmiňovaná GUI byla implementována v prostředí Matlab a budou detailněji popsána níže. Obě rozhraní předpokládají poslech na sluchátka.

Společné pro obě rozhraní je menu (obr 1). Z tohoto menu je možno přejít z jednoho do druhého rozhraní, dále vyvolat nastavení hlasitosti, stručný manuál, případně program ukončit. Po volbě položky „Nastavení hlasitosti“ je zobrazen dialog z obr. 2 a přehráván harmonický signál o kmitočtu 1 kHz. Při přehrávání tohoto zvuku má uživatel možnost nastavit úroveň výstupu do sluchátek tak, aby hlasitost signálu byla ještě přijatelná.

4 GUI pro demonstraci základních principů lokalizace

První program umožňuje ovlivňovat signál do každého ucha a sledovat výsledný efekt na lateralizaci zdroje zvuku. Uživatelské rozhraní je zobrazeno na obr. 3. Program umožňuje dva typy signálů – harmonický signál (obr. 3 nahoře) a přehrávání krátkého zvukového souboru (např. řečové ukázky - obr. 3 dole)).

V případě harmonického signálu si může ještě uživatel vybrat z několika kmitočtů, a tak testovat kmitočtovou závislost procesu lateralizace (lokalizace). Základním ovládacím prvkem tohoto rozhraní jsou tři horizontální posuvníky. První posuvník slouží k ovládní poměru signálů levého a pravého ucha (ILD). Další dva posuvníky slouží k nastavování ITD – prostředním se ovládá posun počátku testovacích signálů (jde o krátké úseky harmonického signálu), druhým lze dosáhnout jemného posunu fáze signálu. Hodnoty nastavených parametrů se vypisují v pravé části rozhraní. V levé části jsou umístěna dvě zobrazovací okna. V horním okně se zobrazuje odhad pozice vnímaného zvuku na základě ILD (černě) a ITD (zeleně), v dolním ilustrace průběhu použitých signálů.



Obrázek 3: Uživatelské rozhraní programu pro demonstraci vlivu ILD (intenzitní rozdíly) a ITD (časové rozdíly) na lateralizaci zdroje zvuku. Nahoře buzení harmonickým signálem, dole vstup z externího signálu (řečová ukázka)

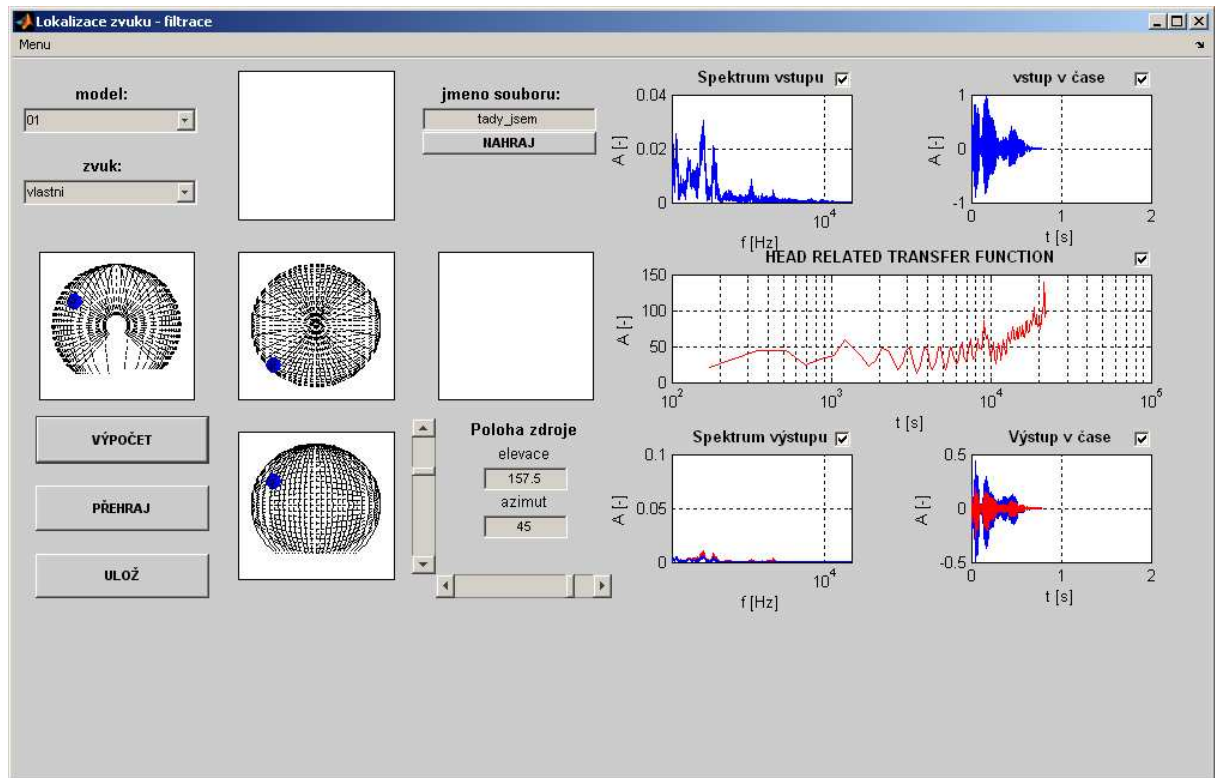
V případě přehrání krátkého externího souboru (.wav), např. řečové ukázky (obr. 3 dole), má uživatel k dispozici pouze dva posuvníky – nastavení poměru amplitud a posun počátku signálu.

Dále je uživateli dostupný režim *Demo*, po stisku tohoto tlačítka je přehrána sekvence harmonických signálů s různým kmitočtem (výškou) a různě nastavenými parametry ILD a ITD.

5 GUI pro lokalizaci zdroje zvuku pomocí HRTF

Druhý program umožňuje simulaci lokalizace zvukového zdroje s využitím HRTF. Jak bylo zmíněno výše, využívá HRIR z databáze [3]. První volbou je tedy volba příslušné sady HRIR (jak je popsáno v [2], zmiňovaná databáze obsahuje HRIR od řady subjektů), která probíhá z rozbalovací nabídky *model*. Je tedy možné testovat individuální rozdíly mezi jednotlivými subjekty, jejichž HRIR je k dispozici.

Druhou volbou je volba vstupního signálu (nabídka *zvuk*). V programu jsou tři možnosti, první je šumový signál, druhá harmonický signál a třetí umožňuje načíst uživatelem definovaný signál z externího souboru .wav. Podle vybrané možnosti se upraví vzhled prostřední části rozhraní, v případě šumu je možno nastavit jen jeho délku, v případě harmonického signálu jeho délku a kmitočet a v případě externího souboru vybrat tento soubor.



Obrázek 4: Uživatelské rozhraní programu pro demonstraci lokalizace zdroje zvuku pomocí HRTF.

Nejdůležitějším nastavením je nastavení směru příchodu zvuku. Toto probíhá pomocí posuvníků *azimut* a *elevace* – po jejich nastavení se uživateli zobrazí nastavená poloha formou půdorysu, bokorysu a nárysu. Podle aktuální polohy je automaticky vybrán vhodný pohled.

Po nastavení všech vstupních parametrů se stisknutím tlačítka *výpočet* provede vytvoření vstupního signálu (popř. načtení ze souboru), načtení HRIR odpovídající příslušnému směru a konvoluce vstupního signálu a HRIR. V pravé části rozhraní jsou okna, kde se zobrazuje vstupní signál v časové i kmitočtové oblasti, použité HRTF a výstupní (upravený) signál v časové i kmitočtové oblasti. Z důvodů snížení náročnosti výpočtu lze vykreslování všech průběhů potlačit zrušením zaškrtnutí v příslušném boxu u každého okénka.

Po provedení výpočtu lze výstupní signál přehrát, popř. uložit do výstupního souboru.

6 Závěr

Tento příspěvek prezentoval dvě GUI umožňující demonstraci a studium základních principů lokalizace zdroje zvuku. Tato uživatelská rozhraní se vhodně doplňují, první umožňuje uživateli seznámení se základními principy, druhé nabízí komplexnější pohled (a věrnější výsledek) pomocí metody HRTF.

Využití představených rozhraní je možné pro demonstrační a výukové účely, ale na našem pracovišti se předpokládá jejich využití v projektech zaměřených na asistivní technologie. Pomocí zmíněných rozhraní lze provést základní ověření, zda u osob ze zrakovým znevýhodněním probíhá lokalizace zdroje zvuku stejným způsobem jako u populace bez tohoto znevýhodnění. Podle výsledků tohoto ověření bude možné např. přikročit k implementaci rozhraní, které bude usnadňovat orientaci zrakově znevýhodněných osob.

Poděkování

Tento příspěvek byl podpořen MŠMT ČR jako součást specifického výzkumu na ČVUT v Praze. Prezentované programy využívají data z CIPIC HRTF databáze [3].

Reference

- [1] Syrový, V., Hudební akustika. Praha: AMU, 2003, pp. 427. ISBN 80-7331-901-2
- [2] Algazi, V. R., Duda, R. O., Thompson, D. M., Avendano, C., The CIPIC HRTF Database, Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics, pp. 99-102, Mohonk Mountain House, New Paltz, NY, Oct. 21-24, 2001.
- [3] The CIPIC HRTF Database [online], c2001, [cit. 21. 10. 2009]. Dostupné z <http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm>.

František Rund

Katedra radioelektroniky, FEL ČVUT v Praze, Technická 2, 166 27, Praha 6
tel. 22435 2108, e-mail: xrund@fel.cvut.cz

Dominik Štorek

student FEL ČVUT v Praze, e-mail: storedom@fel.cvut.cz

Ondřej Glaser

student FEL ČVUT v Praze, e-mail: glaseond@fel.cvut.cz