

AUDIO DISPLEJ PRO NAVIGAČNÍ ÚČELY

A. Štolc

Katedra radioelektroniky, fakulta elektrotechnická,
České vysoké učení technické v Praze, Česká republika

Abstrakt

Práce se zabývá vytvořením jednoduchého navigačního audio displeje v grafickém uživatelském prostředí GUI (Graphical User Interface) použitelného se stereofonními sluchátky. Audio displej klade důraz na intuitivnost, přehlednost a jednoduchost ovládání díky přehlednému grafickému vstupu. Pomocí grafického a zvukového výstupu si lze vytvořit základní přehled o vybraném zvukovém stimulu, a to i při jeho signálové manipulaci v obou dvou kanálech. V práci je obsaženo vysvětlení postupů a principů použitých pro ovládání celého systému.

1 Úvod

Audio displeje se využívají k přeměně dat na zvukový signál, což je možné využít pro široké spektrum aplikací. Používají se již více než sto let a s příchodem nových technologií se jejich možnosti neustále rozšiřují. Jsou stále více používány i v navigačních systémech, ve kterých vytvářejí při využití stereo signálu zcela nový rozměr. Ty pak ve výsledku poskytují přesnější informaci o směru uživatelům, kterými mohou být lidé se zrakovým postižením či lidé pracující v prostředí omezující zrakové pole (hasiči, piloti, atd.). Tato práce obsahuje popis vytvoření audio displeje založeného právě na základních polohovacích a sonifikačních principech v prostředí GUI programu Matlab, který bude virtuálně polohovat vybraný stimul prostým kliknutím myši. Audio displej je koncipován pro použití se stereofonními sluchátky.

2 Návrh audio displeje v grafickém prostředí

Audio displej by měl mít co nejintuitivnější ovládání. Jako vstupní informaci pro přehrání monofonního či stereofonního zvuku z požadované polohy je potřeba znát pouze elevaci a azimut. To lze docílit prostým kliknutím myši do prostoru okna `figure`, kterým dosaneme při použití správného příkazu souřadnice X a Y. V Matlabu pak tento příkaz, který by měl být umístěn do funkce `axes1.ButtonDownFcn()` volané při kliknutí myši do prostoru `axes1`, vypadá následovně

```
pos=get(handles.axes1,'CurrentPoint');
```

Tímto příkazem dostaneme matici `pos`, jejíž dva prvky obsahují informaci o x-ové a y-ové poloze kurzoru v grafu. Před přehráním vzorku je nutné jej napolohovat pomocí známých polohovacích principů. Vzhledem k tomu, že je právě postřehnutelný rozdíl (JND - Just Notable Difference) v lokalizaci směru v horizontální rovině cca 5° , stačí nám simulovat azimut právě s touto kvantizací na x-ové ose. Z důvodu využití diatonické durové stupnice, která bude popsána v článku později, použijeme na vertikální y-ové ose kvantizaci po 6.4° . Pro simulaci elevace použijeme také strukturovaný model boltce [2], který nám virtuálně napolohuje zvuk ve vertikální rovině. Pokud tedy uživatel klikne do prostoru, audio displej přehraje zvuk virtuálně polohovaný do nejbližšího bodu, ve kterém se kvantizační kroky vertikální a horizontální roviny prolínají. Audio displej by měl také obsahovat prvky pro věrnější vjem přehrávaného stimulu, jelikož při přehrání stimulu samotného jej budeme slyšet uvnitř hlavy. Prostým přičtením stimulu k tomutož signálu zpožděnému a utlumenému můžeme vytvořit efekt odrazu od podlahy, čímž docílíme reálnější vjem stimulu.

3 Principy virtuálního polohování

K lokalizaci zdroje zvuku v prostoru je potřeba bilaterálního (oboustranného) slyšení. Jinými slovy potřebujeme minimálně dva signály, jeden pro levé a druhý pro pravé ucho [1]. První principy lokalizace směru zvuku byly představeny již v roce 1907 Lordem Rayleighem v jeho Duplexní teorii. Jedná se konkrétně o rozdíl v intenzitě zvukových vln mezi levým a pravým uchem ILD (Interaural Level Difference – meziušní rozdíl hladin) a o rozdíl v čase dopadu ITD (Interaural Time Difference – časový meziušní rozdíl) [1]. Oba dva principy se navzájem doplňují, z důvodu poměru vlnových délek zvukových vln k velikosti hlavy je každý však účinný v jiném frekvenčním pásmu. Jako způsob pro virtuální polohování zvukových zdrojů v horizontální rovině použijeme algoritmy založené právě na principech ITD a ILD, ILD pomocí panoramování, neboli zeslabování jednoho a zesilování druhého kanálu, či pomocí hlavového stínu. Pro virtuálního polohování elevace, tedy polohy zvukového stimulu ve vertikální rovině, bude využito strukturálního modelu boltce [2].

Bylo vypořádáno [4], že si lidský mozek spojuje výšku tónu s prostorovým vnímáním, konkrétně že s rostoucí frekvencí roste vertikálně i vnímaná poloha zvuku. Jako alternativu k strukturálnímu modelu boltce při virtuálním polohování ve vertikální rovině byla přiložena i sonifikační metoda pitch shifting [3]. Pro ulehčení adaptace uživatele na akustický displej budou mít jednotlivé skoky ve vertikální rovině diferencí jednoho půltónu na diatonické durové stupnici při základním tónu frekvence nahraného stimulu, což se při počtu sedmi tónů při oktávové elevaci 45° rovná kvantování po 6.4°.

4 Virtuální polohování zvuku v horizontální rovině

Pro simulaci ILD jsou použity rovnice 1 a 2, které lineárně aproximují výpočet mezi zesílením levého a pravého kanálu a požadovaným azimutem.

$$y_L = x \cdot g_L \quad (1)$$

$$g_L + g_R = 1 \quad (2)$$

Tento způsob virtuálního polohování zvuku v horizontální rovině je také znám jako amplitudové panorama [1]. Pokud bychom použili skutečný fyzikální model principu ILD, musíme podle velikosti úhlu filtrovat i spektrálně. Hlavový stín vytvoříme pomocí rovnic [2]

$$H_{HS}(\omega, \theta) = \left(\frac{1 + j \frac{\alpha \omega}{2\omega_0}}{1 + j \frac{\alpha \omega}{\omega_0}} \right), \quad 0 \leq \alpha(\theta) \leq 2. \quad (3)$$

ITD využívá fázového posunu mezi levým a pravým kanálem při nenulovém azimutu. Předpokládáme, že ucho na straně zvukového zdroje uslyší dopadající zvukovou vlnu dříve. V základním algoritmu nás tedy pouze zajímá časová diference mezi pravým a levým uchem. Jeho implementaci provedeme tak, že přehrajeme jeden kanál zpožděný o n vzorků od druhého kanálu podle rovnic

$$n_{vz} = \Delta t \cdot f_{vz} \quad (4)$$

$$\Delta t(\theta) = -\frac{r}{c} \cos \theta. \quad (5)$$

5 Virtuální polohování zvuku ve vertikální rovině

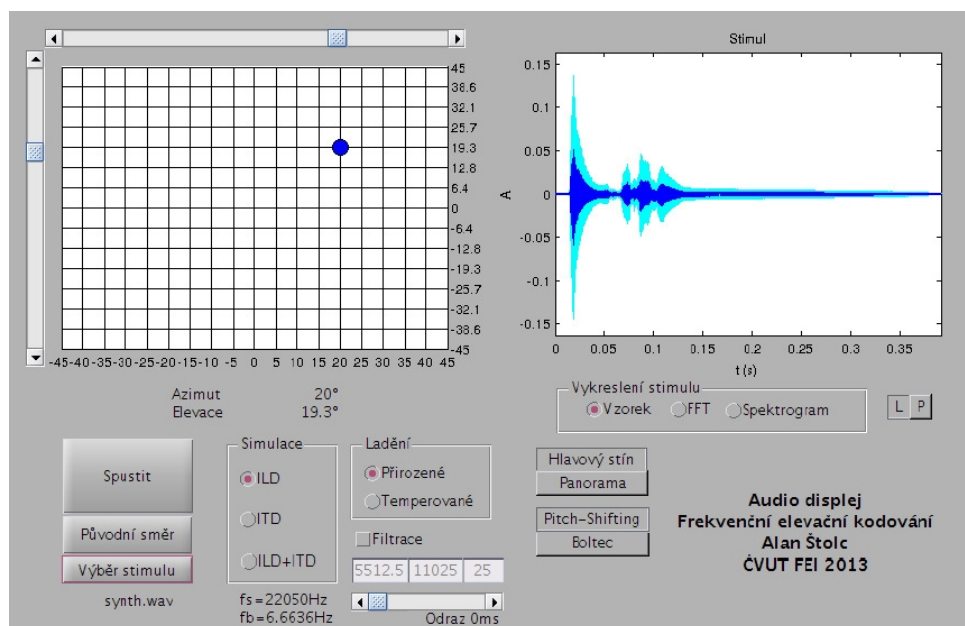
Hlavní podíl na vnímání elevace má ušní boltec. Impulsní odezvou bylo prokázáno [2], že největší aktivita ve frekvenční filtraci zvukové vlny probíhá v prvních 0.7 ms, což při standardní vzorkovací frekvenci 44.1 kHz znamená prvních 32 vzorků. Stačí nám tedy pro simulaci elevace vytvořit například vhodný FIR filtr 32 řádu a simulovat pouze prvních 5 odrazů. Boltec nám filtrací při každém z odrazů vytvoří útlum a zpoždění. Odražený signál vyjádříme odrazovým koeficientem ρ_{odraz} a časové zpoždění pomocí τ_{odraz} , které vypočteme podle rovnice [2]

$$\tau_{odraz}(\theta, \Phi) = A_n \cos\left(\frac{\theta}{2}\right) \sin[D_n(90^\circ - \Phi)] + B_n, \quad (6)$$
$$-90^\circ \leq \theta \leq 90^\circ, -90^\circ \leq \Phi \leq 90^\circ,$$

kde A_n je amplituda, B_n ofset a D_n škálovací faktor.

6 Implementace

Pro jednotnou implementaci všech algoritmů a principů uvedených výše využijeme GUI programu Matlab, konkrétně jeho rozhraní GUIDE (Graphical User Interface Development Environment - Vývojové prostředí GUI). Při implementaci byl také použit Signal Processing Toolbox obsahující například funkce `spectrogram` pro jednoduché vytvoření spektrogramu či `fir` pro vytvoření FIR filtru (Finite Impulse Response - konečná impulsová odezva). Pro zadání vstupních hodnot elevace a azimutu slouží mřížkovaná mapa v levém horním okně programu. Program je ovládán dvěma způsoby. Jedním z nich je napolohování elevace a azimutu pomocí dvou posuvných sliderů vedle tohoto okna, jedním pro horizontální a druhým pro vertikální rovinu. Tímto se posouvá v okně modré kolečko, které značí virtuální polohu stimulu, ten se posléze přehraje pomocí tlačítka *Spustit* v levém dolním rohu spolu s ostatním nastavením.



Obrázek 1: Vytvořený audio displej

Druhým způsobem ovládání je kliknutí kurzoru myši přímo do okna s mapou azimutů a elevací, stimul se poté posune do zvolené polohy a přehraje automaticky, pro opakované přehrání lze stisknout stejné místo v okně či jednodušeji znovu tlačítko *Spustit*. Pokud chceme rychlé přehrání původního směru (nulový azimut i elevace) pro srovnání s aktuální virtuální polohou zvuku, stiskneme tlačítko *Původní směr*.

Pro změnu stimulu slouží tlačítko *Výběr stimulu*, u kterého se zobrazí i název vybraného stimulu a jeho vzorkovací a základní frekvence. V pravém horním okně se nachází grafická informace o aktuálně používaném stimulu. Na výběr je vykreslení časového průběhu, FFT v logaritmickém měřítku a vykreslení spektrogramu. Vedle výběru těchto hodnot si u časového průběhu a FFT lze navolit i sledování levého či pravého kanálu tlačítky *L* a *R*. Kanály se vždy vykreslí oba, zvolený kanál má však výraznější barvu a překrývá druhý nezvolený kanál. Stimul lze také filtrovat pásmovou propustí a zvolit střední frekvenci, šířku pásma a řád FIR filtru [3]. Stimulu lze přidat jeho zpožděný signál o 0-5ms pro simulaci odrazu od podlahy, aby nezněl uvnitř hlavy pro věrnější vjem.

7 Závěr

V práci byl uveden postup vytvoření implementace jednoduchého audio displeje pomocí programu Matlab a jeho prostředí GUI pomocí rozhraní GUIDE. Jeho hlavní výhodou je jednoznačně intuitivní ovládání, nabízí možnost nahrání libovolného stimulu, přehledný grafický vstup a výstup, k dispozici je řada možností k jejich otestování a typů simulací pro polohování. Implementace audio displeje bude použita při studiu možností navigace osob se zaneprázdněným či zhoršeným zorným polem prostřednictvím zvukových stimulů za pomoci stereofonních sluchátek.

Poděkování

Tato práce byla podpořena grantem Studentské grantové soutěže ČVUT v Praze č. SGS11/159/OHK3/3T/13

Reference

- [1] Blauert, J.: *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*
The MIT Press, 1996, ISBN 0-262-02413-6.
- [2] Brown, C. P., Duda, R. O.:
A Structural Model for Binaural Sound Synthesis
IEEE Transaction on Speech and Audio Processing, vol. 6, no. 5, p. 476 - 489, 1998.
- [3] Zölzer, U.: *DAFX Digital Audio Effects*
John Wiley & Sons, Ltd. New York, 2011. ISBN 948-0-470-66599-2.
- [4] Pratt, Carroll C.: *The Spatial Character of High and Low Tones*
Harvard University, 1929

Bc. Alan Štolc

Katedra radioelektroniky, FEL ČVUT v Praze, Technická 2, 166 27 Praha 6, Česká republika
stolcala@fel.cvut.cz