



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

**Fakulta elektrotechnická
Katedra radioelektroniky**

Využití sonifikace pro navigační účely

Application of Sonification for Navigation Purposes

Diplomová práce

Studijní program: Komunikace, multimédia a elektronika
Studijní obor: Multimediální technika

Vedoucí práce: Ing. František Rund, Ph.D.

Bc. Alan Štolc

Praha 2013

Vložit originální zadání

Seznamte se s metodami pro polohování virtuálních zdrojů zvuku (VAS - Virtual Auditory Space) pro účely navigace s využitím metod sonifikace pro vylepšení vjemu polohy. Navrhněte způsob zpřesnění lokalizace virtuálních zdrojů pomocí sonifikačních metod. Navrženou metodu ověřte poslechovým testem a výsledky porovnejte s výsledky dosaženými jinými metodami.

Čestné prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne _____

Podpis

Poděkování

Rád bych tímto poděkoval především Ing. Františku Rundovi, PhD. za jeho trpělivost, cenné rady a připomínky při vedení mé diplomové práce. Dále bych rád poděkoval svým rodičům za finanční podporu při studiu.

Abstrakt

Tato diplomová práce se zabývá problematikou virtuálního polohování zvuku pro navigační účely s využitím sonifikace pro vylepšení vjemu polohy. Jsou v ní popsány základní typy virtuálního polohování zvukových stimulů, dále jsou popsány základní typy sonifikace, audio displejů a sonifikačních metod. Je v ní také obsažen popis vytvoření audio displeje kombinující základní techniky virtuálního polohování zvuku v horizontální rovině a sonifikační techniku přeladování základního tónu stimulu v rovině vertikální. Audio displej je otestován a výsledky okomentovány.

Klíčová slova: *virtuální polohování zvuku, sonifikace, audio displej*

Abstract

This thesis contains problematics of virtual sound positioning for navigation purposes with the use of sonification for enhancing localization perception. It describes basic types of virtual sound positioning techniques and also, basic types of sonification, auditory displays and sonification techniques are described. It also includes description of creation of the auditory display, that combines basic techniques of virtual sound positioning in horizontal plane and pitch shifting sonification technique in vertical plane. Auditory display is tested and the results are commented.

Keywords: *virtual sound positioning, sonification, auditory display*

Obsah

1	Úvod	1
2	Principy lokalizace	2
2.1	Souřadný systém	2
2.2	Lokalizace zvukového zdroje	3
2.3	Lokalizace směru v horizontální rovině	3
2.3.1	ITD	4
2.3.2	ILD	4
2.4	Lokalizace směru ve vertikální rovině	5
3	Virtuální polohování zvuku	6
3.1	ILD	6
3.2	ITD	7
3.3	Strukturální model boltce	7
3.4	HRIR	8
3.5	HRTF	9
4	Sonifikace	10
4.1	Rozdělení sonifikace	11
4.2	Kategorizace sonifikace	12
4.3	Audio displeje	13
4.3.1	Výhody audio displejů	14
4.3.2	Rozdělení audio displejů	15
4.3.3	SWAN	16
4.3.4	vOICe	17
4.3.5	Ostatní navigační audio displeje	18
4.4	Prostorové vnímání kmitočtů	19
4.5	Diatonická durová stupnice	20
4.6	Pitch shifting	22
4.6.1	Základní tón signálu	22
4.6.2	Time Stretching/Resampling	24
4.6.3	Fázový vokodér	26

5 Implementace audio displeje	28
5.1 Využití diatonické durové stupnice	29
5.2 Audio displej s elevačním kódováním	29
5.3 Příložený program	30
5.3.1 Nastavení a ovládání simulačního programu	31
5.3.2 Virtuální polohování stimulu	31
5.3.3 Úprava stimulu	31
5.3.4 Ovládání grafického vstupu	32
5.3.5 Grafický výstup	32
5.3.6 Vývojový diagram zkušebního rozhraní	33
5.4 Testovací rozhraní	34
5.4.1 Ovládání myší	34
5.4.2 Ovládání joystickem	35
5.5 Použité algoritmy	37
5.6 Zpracování výsledků	38
6 Testování	39
6.1 Testovací podmínky	39
6.2 Metodika testování	40
6.3 Výsledky měření	42
7 Závěr	45
8 Literatura	46
9 Soubory na příloženém CD	49
10 Přílohy	50

Seznam obrázků

1	Sférická soustava souřadnic [3]	2
2	ITD - Časový meziušní rozdíl	4
3	Cone of Confusion [6]	5
4	HRIR - elevace 0, azimut 60	8
5	Frekvenční odezva ideální tuhé koule [5]	9
6	Definice sonifikace [11]	10
7	Audifikace	11
8	Parametricky mapovaná sonifikace	12
9	Rozdělení organizovaných zvuků a poloha sonifikace [11]	12
10	Navigační systém SWAN [17]	16
11	Navigační systém vOICe	17
12	Tactical Sound Garden - Přístupové body v části Curychu [24]	18
13	Výsledky měření C.C.Prattové [1]	19
14	Solmizační slabiky a jejich přirozené poměry	20
15	Spektrogram řečového signálu	22
16	Reálné kepstrum řečového signálu	23
17	Princip zvýšení základní frekvence	24
18	Princip snížení základní frekvence	25
19	Princip STFT [25]	26
20	Princip fázového vokodéru	27
21	Srovnání JND elevace HRTF a diatonické durové stupnice	28
22	Testovací software	30
23	Nastavení testovacího softwaru	31
24	Grafický výstup	32
25	Vývojový diagram zkušebního rozhraní	33
26	Testovací prostředí programu - ovládání myší	34
27	Testovací prostředí programu - ovládání joystickem	35
28	Vývojový diagram testovacího rozhraní	36
29	Použitý syntetický stimul	39
30	Průběh testování	40
31	Měření audio displeje	41

32	Výsledky měření elevace - Použití myši v učebním módu	50
33	Odchylka měření elevace - Použití myši v učebním módu	50
34	Výsledky měření azimutu - Použití myši v učebním módu	51
35	Odchylka měření azimutu - Použití myši v učebním módu	51
36	Výsledky měření elevace - Použití myši	52
37	Odchylka měření elevace - Použití myši	52
38	Výsledky měření azimutu - Použití myši	53
39	Odchylka měření azimutu - Použití myši	53
40	Výsledky měření elevace - Použití joysticku	54
41	Odchylka měření elevace - Použití joysticku	54
42	Výsledky měření azimutu - Použití joysticku	55
43	Odchylka měření azimutu - Použití joysticku	55
44	Výsledky měření elevace - Použití joysticku naslepo	56
45	Odchylka měření elevace - Použití joysticku naslepo	56
46	Výsledky měření azimutu - Použití joysticku naslepo	57
47	Odchylka měření azimutu - Použití joysticku naslepo	57
48	Výsledky měření elevace - Strukturální model boltce	58
49	Odchylka měření elevace - Strukturální model boltce	58
50	Výsledky měření azimutu - Strukturální model boltce	59
51	Odchylka měření azimutu - Strukturální model boltce	59

Seznam zkratek

ILD	Interaural Level Difference - Rozdíl hladin mezi ušima
ITD	Interaural Time Difference - Časový rozdíl mezi ušima
HRIR	Head-Related Impulse Response Impulsová odezva vztažená k hlavě
HRTF	Head-Related Transfer Function Přenosová funkce vztažená k hlavě
IIR	Infinite Impulse Response - Nekonečná impulsová odezva
FIR	Finite Impulse Response - Konečná impulsová odezva
FFT	Fast Fourier Transform - Rychlá Fourierova transformace
STFT	Short-Time Fourier Transform Krátkodobá Fourierova transformace
ISTFT	Inverse Short-Time Fourier Transform Inverzní krátkodobá Fourierova transformace
DFT	Discrete Fourier Transform - Diskrétní Fourierova transformace
SWAN	System for Wearable Audio Navigation Systém pro přenosnou osobní audio navigaci
GPS	Global Positioning System - Globální polohovací systém
RFID	Radio Frequency Identification Identifikace na bázi rádiových frekvencí
JND	Just Notable Difference - Právě postřehnutelný rozdíl
GUI	Graphical User Interface - Grafické uživatelské rozhraní

1 Úvod

V práci je v úvodní teoretické části zpracována problematika základních principů binaurální lokalizace zvukového zdroje spolu s možnostmi jeho virtuálního polohování pomocí stereofonních sluchátek, a to jak v horizontální, tak ve vertikální rovině.

Druhá polovina teoretické části se zaměřuje na základní přehled problematiky a technik sonifikace pro její následné využití k vytvoření audio displeje použitelného pro navigační účely. Je zde v jednoduchosti popsána i teorie diatonické durové stupnice, techniky pro zjištění základní frekvence zvuku a pitch shifting.

V praktické části se nachází popis navrženého a vytvořeného audio displeje. Využívá frekvenčního elevačního kódování, které je založeno na fenoménu, že lidské ucho vyhodnocuje vyšší frekvence prostorově ve vyšších polohách, nežli frekvence nižší [1]. Jeho jednotlivé kroky v elevaci jsou pak od sebe vzdáleny stejnými frekvenčními poměry, jako tóny diatonické durové stupnice. Díky těmto poměrům a metodě pitch shifting je možné použít téměř libovolný zvukový stimul. Tyto vlastnosti výrazně usnadňují adaptaci uživatelů na audio displej samotný. V horizontální rovině se pak řídí známými polohovacími principy využívajíc stereofonního zvukového signálu.

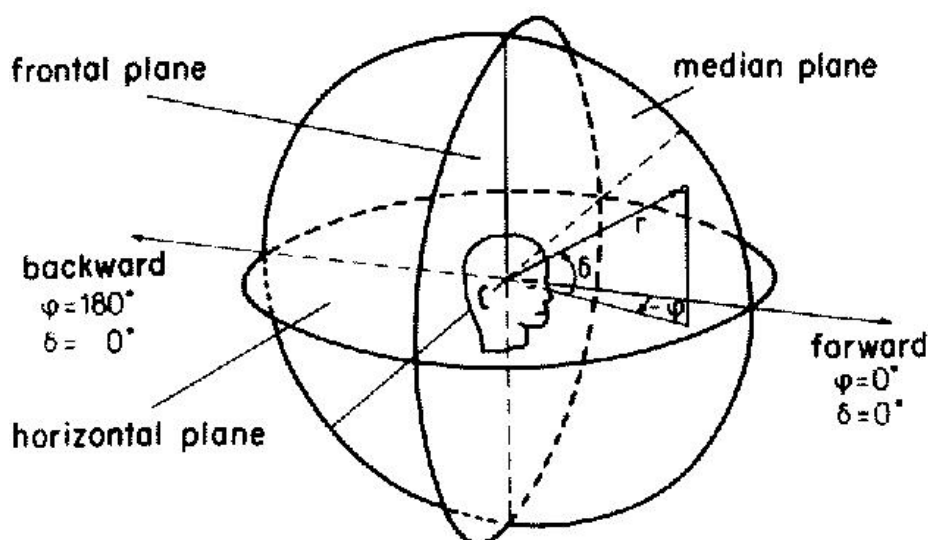
V poslední části práce jsou popsány podmínky, metodika a výsledky testování audio displeje. Výsledky měření jsou rovněž vykresleny do grafů a zhodnoceny.

Cílem práce, a tedy i důvodem vytvoření tohoto audio displeje, je prokázání použitelnosti sonifikačních technik k navigačním účelům a také snaha o zpřesnění elevačního polohování virtuálního zdroje, které je prozatím pomocí jiných známých polohovacích technik neuspokojivé. Při jeho zlepšení se poté dá tento audio displej využít například pro navigační účely ve virtuálním prostoru či v prostoru reálném tam, kde je důraz kladen i na lokalizaci ve vertikální rovině. Práce je součástí grantu č. SGS11/159/OHK3/3T/13.

2 Principy lokalizace

2.1 Souřadný systém

Ze všeho nejdříve je nejdůležitější popsat systém, ve kterém budeme pracovat. Jako nejvíce praktický se pro popis polohy v horizontální (azimutu) a vertikální rovině (elevace) u problematiky prostorového slyšení jeví sférická soustava souřadnic [2], v jejímž středu budeme předpokládat hlavu posluchače, a to přesně v polovině spojnice levého a pravého ucha. Za těchto podmínek poté definujeme úhel δ jako úhel elevace, neboli pro uživatele nahore a dole, a φ jako úhel azimutu, neboli vlevo či vpravo. Uši se tedy nachází v nulové elevaci na pozici $\pm 90^\circ$ a vše, co se odehrává přímo před hlavou, se nachází v nulové elevaci a na nulovém azimutu. Dále definujeme vzdálenost R jakožto vzdálenost zdroje od pozorovatele. Ilustrační popis celého systému je na obrázku 1.



Obrázek 1: Sférická soustava souřadnic [3]

Dalším důležitým faktem při práci v tomto systému je, že hodnota -90° v elevaci či v azimutu je rovna 270° a naopak, ale pro zjednodušení pracujeme s hodnotami do $\pm 180^\circ$. Taktéž například poloha 180° v azimutu a 0° v elevaci je totožná s 0° v azimutu a 180° v elevaci, vždy záleží na konkrétním případě použití.

2.2 Lokalizace zvukového zdroje

K lokalizaci zdroje zvuku v prostoru je potřeba bilaterálního, neboli oboustranného slyšení. Jinými slovy potřebujeme minimálně dva signály, jeden pro levé a druhý pro pravé ucho [4]. Z jejich srovnávání poté vyhodnocujeme směr, ze kterého se daný zvuk šíří [2]. Tyto signály dělíme na:

- monotické - signál pouze v jednom kanálu
- diotické - stejný signál v každém kanálu
- dichotické - v každém kanálu jiný signál

Úspěšné lokalizace v prostoru je možné docílit pomocí signálů dichotických, pomocí signálů diotických je možná částečná lokalizace také, ale pouze ve vertikální rovině bez lokalizace azimutu. Vzdálenost zdroje zvuku vnímáme pomocí síly signálu, čím se zdroj zvuku nachází dál, tím je kvůli útlumu zvukové vlny ve vzduchu intenzita slabší, a proto jej vnímáme z větší vzdálenosti. Zároveň mají vysoké frekvence větší útlum se zvyšující se vzdáleností nežli frekvence nízké, tento jev lze aproximovat dolní propustí. Je zde však předpoklad, že zdroj zvuku předem známe v jeho původní nefiltrované podobě. Dalšími podmínkami jsou také minimální hlasitost zvuku, která by měla být pro úspěšnou lokalizaci cca 50 dB, délka trvání alespoň 250 ms, spektrální obsah a šířka pásma položené i do vyšších kmitočtů pro úspěšnou lokalizaci elevace a také vizuální obsah spojený s akustickou informací. V prvním cca 50 ms lokalizujeme zdroj ze směru, ze kterého přijde první zvukový impuls. Tento jev se nazývá precedence effect a vyskytuje se zejména v uzavřených prostorech.

Lokalizace směru, respektive elevace a azimutu, funguje na dalších principech, které se od sebe z důvodu umístění a tvaru uší ve vertikální a horizontální rovině liší.

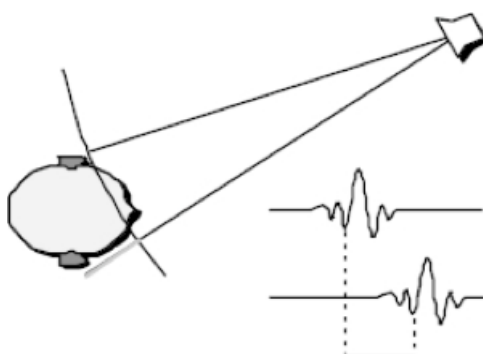
2.3 Lokalizace směru v horizontální rovině

První principy lokalizace směru zvuku byly představeny již v roce 1907 Lordem Rayleighem v jeho Duplexní teorii [4]. Jedná se konkrétně o dva principy, a to o rozdílu v intenzitě zvukových vln mezi levým a pravým uchem ILD (Interaural Level Difference) a o rozdílu v jejich čase dopadu ITD (Interaural Time Difference).

Polovina hlavy v horizontální rovině blíže ke zvukovému zdroji se nazývá ipsilaterální, zatímco opačnou polovinu nazýváme kontralaterální. Na tuto stranu nemůže zvuková vlna dopadnout po přímce, a dochází zde k jevům popsaných níže.

2.3.1 ITD

Prvním principem detekce směru zvukového vjemu v horizontální rovině je ITD, který předpokládá zpožděný dopad zvukové vlny na kontralaterální stranu, jedná se tedy o fázový posuv. Příklad ITD je vidět a obrázku 2. Pokud například použijeme vzdálenost levého a pravého ucha 21.5 cm, vznikne mezi nimi při azimutu $\pm 90^\circ$ největší časový rozdíl cca 625 μs . U frekvencí menších než cca 800 Hz je posun fáze velmi výrazný, jelikož je vzdálenost uší pro tyto kmitočty menší než polovina vlnové délky. Vzniká zde takové fázové zpoždění, které má za důsledek výrazné zjednodušení lokalizace zvukového zdroje. Proto je tento princip nejvýraznější u frekvencí menších než cca 1.5 kHz a s klesající frekvencí roste současně i jeho účinnost.



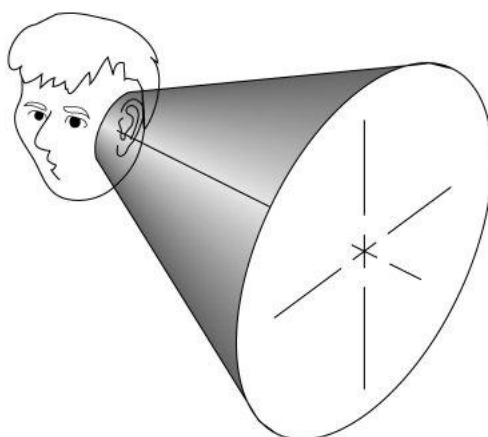
Obrázek 2: ITD - Časový meziušní rozdíl

2.3.2 ILD

Dalším principem detekce směru zvuku v horizontální rovině je ILD, který předpokládá rozdíl v intenzitě dopadající zvukové vlny do levého a pravého ucha. Ucho na kontralaterální straně hlavy zachytí dopadající zvukovou vlnu o menší intenzitě. Důvodem je hlavně tzv. akustický stín hlavy (head shadow). Hlavový stín vzniká tím, že zvukové vlny nedopadají na kontralaterální straně hlavy po přímce, ale po povrchu hlavy, která má přibližný tvar koule. S klesající délkou zvukové vlny, která je srovnatelná či menší než průměr hlavy dochází k odrazům od povrchu. Při průměru hlavy 21.5 cm je tedy tento jev nejvýraznější při frekvencích větších než 1.5 kHz. U frekvencí menších než cca 80 Hz je pak kvůli velikosti hlavy lokalizace pomocí ILD prakticky nemožná, výjimkou je přímý poslech například pomocí stereofonních sluchátek.

2.4 Lokalizace směru ve vertikální rovině

Pokud v prostoru zvolíme libovolnou polohu zvukového zdroje, pak při aplikaci principů vysvětlených výše dojdeme k závěru, že existují i další polohy zvukových zdrojů, které mají stejné vlastnosti jako zdroj zvolený. Konkrétně bychom tak mohli lokalizovat mnoho poloh jednomu zvukovému zdroji, a to na kuželu protínajícím zvolenou polohu zvukového zdroje se špičkou ve středu soustavy. Tento jev se nazývá Cone of Confusion, neboli kužel neurčitosti [5].



Obrázek 3: Cone of Confusion [6]

Lidské sluchové ústrojí je však pomocí specifického tvaru boltce a odrazů od ramen a hlavy schopno lokalizovat zvukový zdroj i ze zadní poloviny a vertikální roviny. Pro každý směr vytváří přirozený filtr, spektrální složka přijatého zvuku se pak projeví na jeho následné lokalizaci. Při poloze zvukového zdroje v zadní polovině je zvuk stíněn zadní stranou boltce, zatímco z přední poloviny dopadá zvuková vlna přímo, což se projeví na celém rozsahu spektra. Při lokalizaci elevace pak slouží boltce primárně pro filtraci vysokofrekvenční energetické složky zvukové vlny a ramena s hlavou pak pro složku nízkofrekvenční. Každý člověk má jiný tvar a velikost boltce a ramen a proto má individuální přenosovou funkci, neboli vlastní systém filtrů pro různé směry, na který je zvyklý. Pro úspěšnou lokalizaci je zde ale stejně jako u vyhodnocování vzdálenosti zvukového zdroje předpoklad, že posluchač podobný zvuk již slyšel v jeho nefiltrované podobě.

3 Virtuální polohování zvuku

3.1 ILD

Jak bylo zmíněno v druhé kapitole, princip ILD předpokládá rozdíly v intenzitě dopadajícího signálu do levého a pravého ucha. Při simulaci ILD můžeme použít rovnice 1 a 2, které lineárně aproximují výpočet mezi zesílením levého a pravého kanálu jednoduchým vztahem

$$y_{L(P)} = x \cdot g_{L(P)} \quad (1)$$

$$g_L + g_P = 1. \quad (2)$$

Toto jsou základní rovnice, které předpokládají, že pokud např. $\varphi = 90^\circ$, pak $g_L = 0$ a $g_P = 1$, požadovaný azimut tedy dostaneme např. pomocí $\varphi = 90(1 - 2 \cdot g_L)$. Další rovnicí aplikovatelnou pro simulaci ILD je např.

$$\tan(\varphi) = \left(\frac{g_L - g_P}{g_L + g_P} \right), \quad (3)$$

kteřá má tangenciální závislost a platí pro případ, kdy reproduktory svírají s posluchačem úhel 45° . Tyto způsoby virtuálního polohování zvuku v horizontální rovině jsou také známy jako amplitudové panorama [4] a jsou hojně využívány v nahrávacím průmyslu, pro PC hry, film, rozhlas atd. Existuje mnoho typů, rovnice uvedené na této stránce patří mezi ty nejjednodušší na implementaci. Mezi přednosti těchto algoritmů patří právě snadná implementace a relativně věrný vjem polohy zvuku.

Pokud bychom použili skutečný fyzikální model principu ILD, musíme podle velikosti úhlu filtrovat i spektrálně. Hlavový stín můžeme vytvořit následujícím způsobem [7]

$$H_{HS}(\omega, \theta) = \frac{1 + j \frac{\varphi \omega}{2\omega_0}}{1 + j \frac{\varphi \omega}{\omega_0}}, \quad 0 \leq \varphi(\theta) \leq 2, \quad (4)$$

kde ω_0 vyjadřuje poměr mezi rychlostí zvuku c a poloměrem hlavy r . Poloha nuly v rovnici 4 se pak mění s azimutem podle rovnice

$$\varphi(\theta) = \left(1 + \frac{\varphi_{min}}{2} \right) + \left(1 - \frac{\varphi_{min}}{2} \right) \cos \left(\frac{\theta}{\theta_{min}} 180^\circ \right). \quad (5)$$

Hodnoty $\varphi_{min} = 0.1$ a $\theta_{min} = 150$ jsou stanoveny pro ideální frekvenční odezvu kulového modelu hlavy [5]. Těmito rovnicemi tedy můžeme simulovat útlum vysokých frekvencí u hlavového stínu. Nevýhodou všech těchto algoritmů je nemožnost simulace zadních 180° a elevace, pro tyto polohy je nutné použít ještě další spektrální filtraci.

3.2 ITD

ITD využívá fázového posunu mezi levým a pravým kanálem při nenulovém azimutu. Předpokládáme, že ucho na kontralaterální straně uslyší dopadající zvukovou vlnu později. V základním algoritmu nás tedy pouze zajímá časový rozdíl mezi pravým a levým uchem.

$$\Delta t(\theta) = -\frac{r}{c} \cos \varphi \quad (6)$$

Jeho implementaci provedeme tak, že přehrajeme jeden kanál zpožděný o n vzorků od druhého kanálu, poloha, respektive vychýlení, pak bude záviset na vzorkovací frekvenci.

$$n_{vzorku} = \Delta t \cdot f_{vz} \quad (7)$$

3.3 Strukturální model boltce

Hlavní podíl na vnímání elevace má ušní boltce. Bylo prokázáno [7], 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 FIR filtr 32 řádu a simulovat pouze prvních 5 odrazů. Boltce 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 [7]

$$\tau_{odraz}(\varphi, \delta) = A_n \cos\left(\frac{\varphi}{2}\right) \sin[D_n(90^\circ - \delta)] + B_n, \quad (8)$$

$$-90^\circ \leq \varphi \leq 90^\circ, -90^\circ \leq \delta \leq 90^\circ,$$

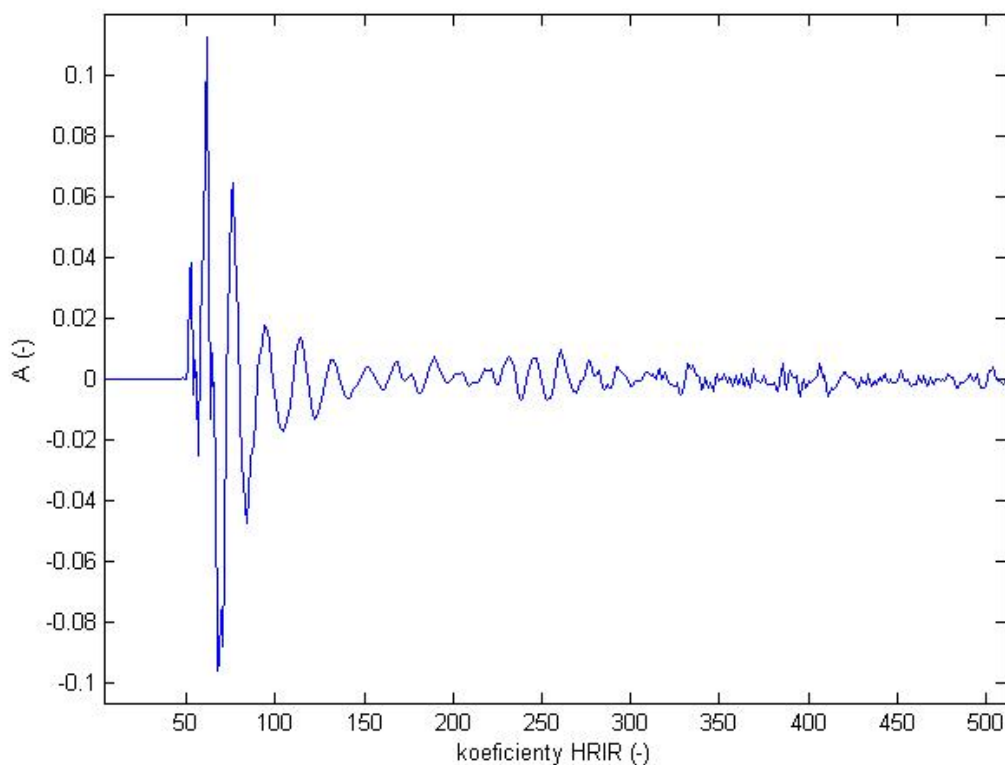
kde A_n je amplituda, B_n offset a D_n škálovací faktor. Tyto hodnoty můžeme pro každý z odrazů podle strukturálního modelu [7] doplnit podle tabulky 1.

n	ρ_{boltce}	A_n	B_n	D_n
2	0.5	1	2	1
3	-1	5	4	0.5
4	0.5	5	7	0.5
5	-0.25	5	11	0.5
6	0.25	5	13	0.5

Tabulka 1: Koeficienty pro strukturální model boltce

3.4 HRIR

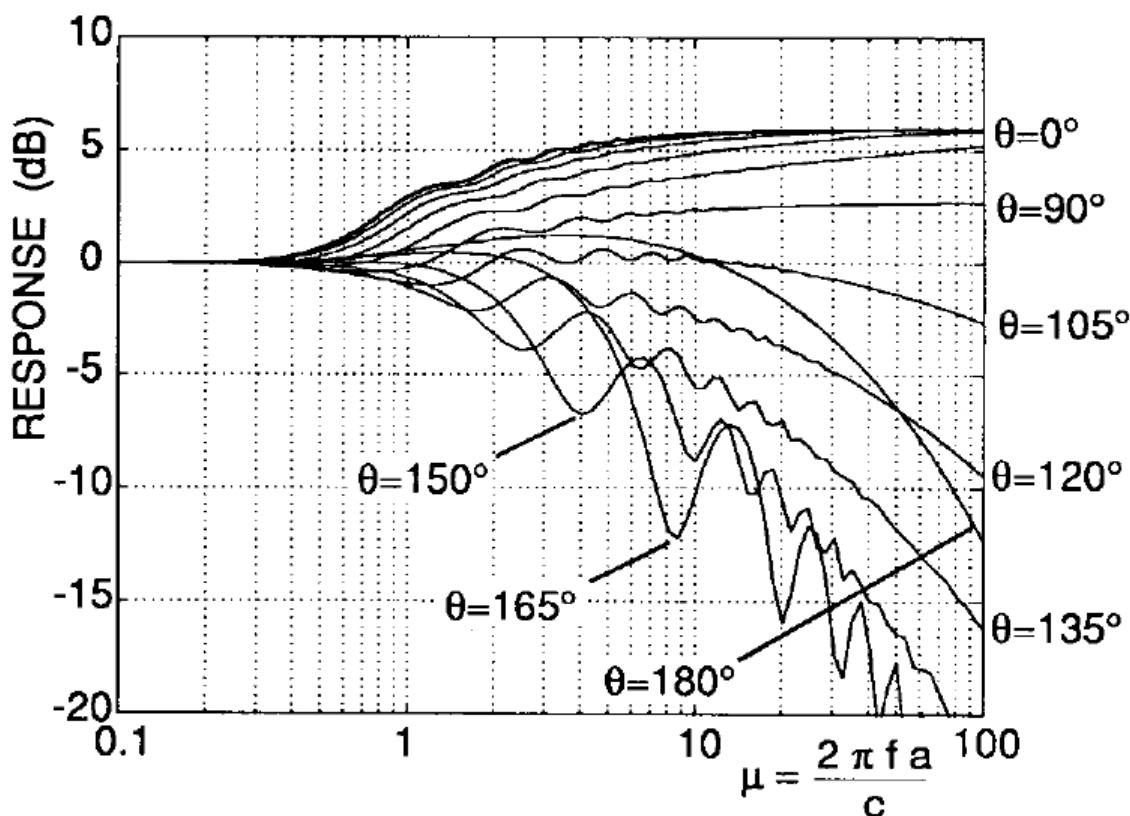
HRIR (Head Related Impulse Response - impulsová odezva vztažená k hlavě) je odezvou soustavy zvukový zdroj - sluchová soustava na Diracův impuls [5]. Na obrázku 4 níže je vidět příklad impulsové odezvy zvukového zdroje v horizontální rovině a azimutu 60° . Impulsová odezva je měřena párem mikrofonních sond na umělé hlavě či reálném uživateli, mění se zejména v závislosti na azimutu, elevaci, vzdálenosti a na fyzických rozměrech hlavy, ramen a boltců. Pro každou polohu máme tedy dvojici impulsových odezev, jeden pro levé a druhý pro pravé ucho. Pokud provedeme konvoluci těchto signálů s monofonním či stereofonním zvukovým signálem, uslyšíme jej ze směru původního impulsu. Impulsová odezva změřená mikrofonní sondou je vlastností lineární soustavy, na kterou můžeme nahlížet jako na IIR filtr (Infinite Impulse Response - nekonečná impulsová odezva), ze kterého se při konečném počtu vzorků stane FIR filtr (Finite Impulse Response - konečná impulsová odezva). Ten se nám postará o simulaci jak ve vertikální, tak i horizontální rovině. Nevýhodou však je, že má každý člověk individuální HRIR a její měření je značně zdlouhavé, jelikož každý směr a vzdálenost potřebuje svoje individuální měření (svůj vlastní pár FIR filtrů).



Obrázek 4: HRIR - elevace 0, azimut 60

3.5 HRTF

Pokud převedeme HRIR do spektrální oblasti pomocí Fourierovy transformace, dostaneme HRTF (Head Related Transfer Function – přenosová funkce vztažená k hlavě). Jedná se o přenosovou funkci, která nám udává útlum zvukového signálu z daného směru a vzdálenosti v závislosti na frekvenci. Stejně jako HRIR se v ní mísí veškeré polohovací principy zmíněné výše.



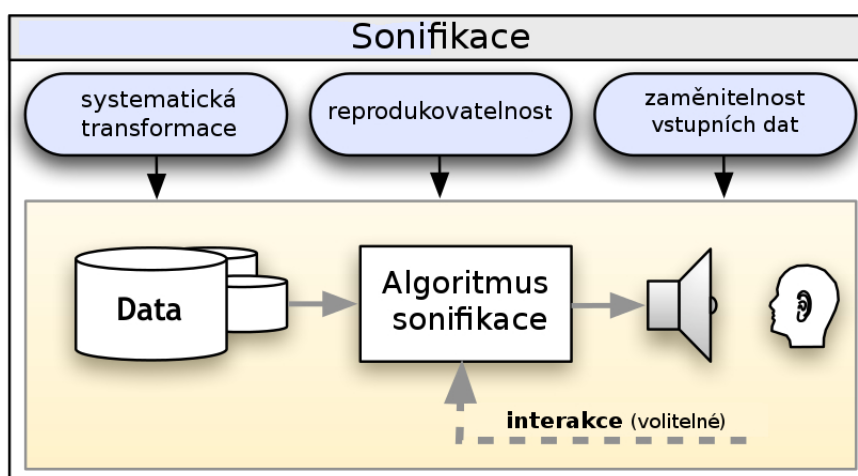
Obrázek 5: Frekvenční odezva ideální tuhé koule [5]

Pro účely virtuálního polohování zvukového zdroje můžeme změřit vlastní naměřené HRIR, HRIR umělé hlavy či použít HRIR některé z dostupných databází, ale odhad směru se v jiném než v případě vlastní HRIR či HRTF zhorší. Další nevýhodou HRIR a HRTF je poměrně nízká kvalita simulace elevace, rozlišení ve vertikální rovině se pohybuje kolem 15° , a to pouze u přesnějších individuálních HRTF [10]. V reálných situacích se vertikální JND (Just Notable Difference - právě postřehnutelný rozdíl) pohybuje v hodnotách přes 20° [8].

4 Sonifikace

Zvuk je používán v rozhraní člověk-stroj již mnoho let, ačkoliv zpočátku se využívaly zejména varovné zvuky. Jako je vizualizace převedením dat do vizuální formy, tak je sonifikace technika transformující vstupní datové relace na výstup v podobě neřečového akustického signálu za účelem zjednodušení komunikace či interpretace za těchto podmínek [11]:

- Zvukový signál na výstupu objektivně znázorňuje vlastnosti či vazby k vstupním datům.
- Transformace vstupních dat na zvukový signál je systematická a s přesnou definicí, jakým způsobem je provedena.
- Sonifikace je reprodukovatelná, při stejných vstupních datech musí být vždy ten samý zvukový výstup.
- Systém použitý se zaměněnými vstupními daty po návratu na původní vstupní data přehrává původní akustický výstup.



Obrázek 6: Definice sonifikace [11]

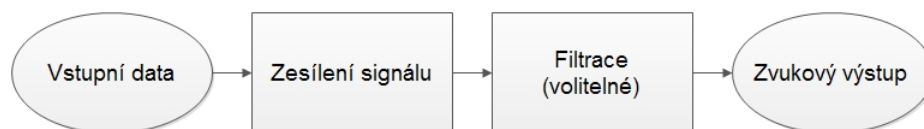
Všechny tyto podmínky mají své odůvodnění, zejména aby byla sonifikace chápána jako exaktní vědní obor. Tyto podmínky oddělují sonifikaci například od hudby, ve které není hlavním účelem zjišťovat, jak a proč hudebník hýbe prsty, aby rozezněl nástroj, ale vychutnat si hudební zážitek. Rozdíl je obdobný jako u vizualizace, po které požadujeme co nejjednodušší zobrazení informací, zatímco u namalovaného obrazu se nám jedná o výše zmíněný umělecký prožitek. I přes tyto poznámky, hudba i sonifikace jsou organizovanými zvuky, čímž k sobě mají blízko. Sonifikační techniky tedy mohou znít i jako hudba, avšak pouze za zmíněných podmínek.

4.1 Rozdělení sonifikace

Sonifikaci rozdělujeme na 5 základních skupin, které splňují 4 základní požadavky pro sonifikaci, uvedené na předchozí stránce [13]:

- *Audifikace*

Nejjednodušší sonifikační metodou je audifikace, ve které uskutečňujeme přímý převod vstupních dat do zvukového signálu. Audifikace většinou probíhá v reálném čase a předpokládá šumové či amplitudové změny ve vstupním signálu, které označují určitou událost. Signál je pouze zesílen na slyšitelné hodnoty a předpokládá šířku pásma vstupních dat ve slyšitelném pásmu 20 Hz až 20 kHz, filtrování signálu pro zaměření se na určitý úsek je možné. Používá se například v lékařství a seismologii.



Obrázek 7: Audifikace

- *Audio ikony*

Zvuky vyjadřující obdobnou událost. Zaměřují se tedy na typ dat než na jejich hodnoty, většinou nemají velkou informační hloubku. Mohou být nahrané či uměle vytvořené. Jako dobrý příklad poslouží ikonka koš, která po vysypání vydá zvuk zmačkaného papíru.

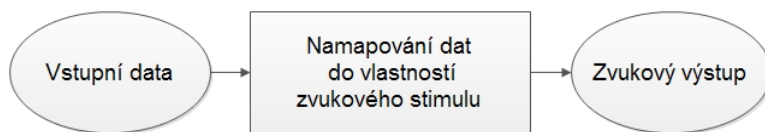
- *Earcony*

Zvuky přiřazené událostem, s jejímiž daty nemají nic společného. Často se jedná o synteticky vytvořené zvuky, které dokreslují prostředí či něco oznamují. Jsou hojně využívané například v grafických operačních systémech. Příkladem může být pípání mikrovlnky či zvuk stisknutí tlačítka ve výběrovém menu. Stejně jako audio ikony nemají větší informační hloubku, ale najdou se i výjimky, například v podobě oznamovacích tónů nedostupnosti telefonu a neexistujícího čísla.

- *Parametricky mapovaná sonifikace*

Vlastnosti zvukových signálů jsou závislé na vlastnostech vstupních dat, podle kterých jsou mapovány. S jejich změnou se mění rychlost, základní frekvence či síla signálu. Také je možné změnit celý stimul nebo provést spektrální filtrace. Tvoří se zde přímá závislost mezi vstupními daty a výstupním signálem. Tato závislost modifikuje jednu či více vlastností zvukového stimulu. Tato kategorie má již na rozdíl od předchozích informační hloubku, která může mít až několik rozměrů.

Do kategorie parametricky mapované sonifikace spadá i sonifikace zvukového toku, ve kterém modifikujeme základní vlastnosti kontinuálního zvukového signálu v závislosti na vlastnostech vstupních dat, například jeho rychlost.



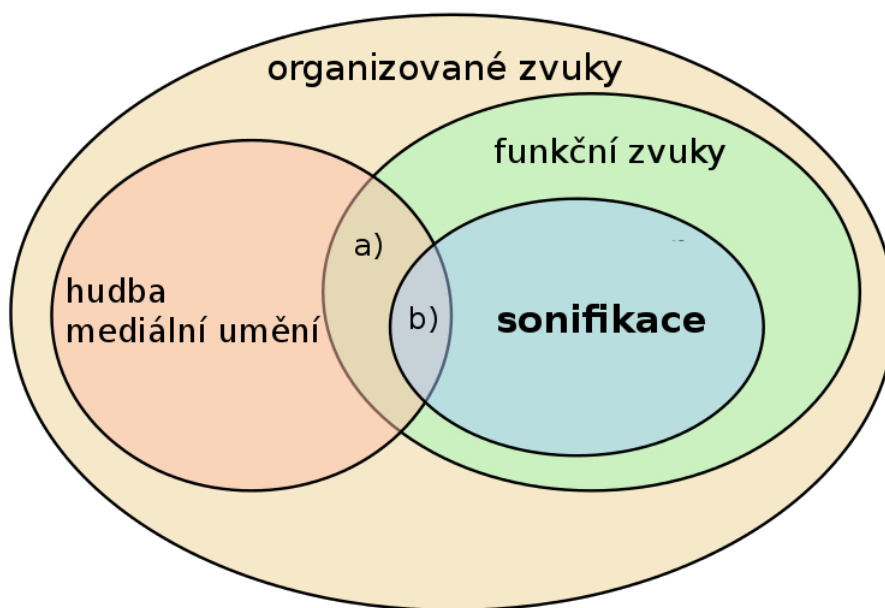
Obrázek 8: Parametricky mapovaná sonifikace

- *Modelově založená sonifikace*

Tato kategorie je velmi podobná parametricky mapované sonifikaci, avšak data nejsou mapována přímo, ale podle určitého modelu, který je pro uživatele logičtější či jednodušší pro adaptaci. Předpokládá se informační hloubka signálu.

4.2 Kategorizace sonifikace

Kam přesně spadá sonifikace v audio signálech lze vidět na obrázku 9. Je z něj zřejmé, že se jedná o organizované a funkční zvuky. Prolnutí s hudební množinou nastává v případě, kdy využíváme hudebních vlastností při mapování, jako například tempo při pravidelné frekvenci stimulu, kdy jeho změna znamená událost, či frekvenční mapování podle hudebních stupnic, kdy jsou kvantizovaná data namapována na jednotlivé tóny.



Obrázek 9: Rozdělení organizovaných zvuků a poloha sonifikace [11]

4.3 Audio displeje

Audio displeje mají za úkol přeměnit data na akustický signál ve směru počítač-uživatel [13], jedná se tedy o nadřazenou skupinu sonifikace, která je jednou z možností použití. Byly používány již od začátku 20. století a v dnešní době se staly běžnou součástí lidského života.

Jeden z prvních audio displejů můžeme znát v podobě detektoru ionizačního záření (Geigerův-Müllerův detektor [15]), který má svou referenční frekvenci akustického signálu v pravidelných intervalech při nulové či zanedbatelné radiaci a při výskytu gama záření se frekvence signálu zvyšuje a intervaly se stávají kratší. Uživatel si tedy zvykne na frekvenci při nulové hodnotě radiace a poté pohotově reaguje při případných změnách signálu. V tomto konkrétním případě není dán důraz na přesnost, ale na intuitivnost a jednoduchost ovládání. Uživatel se také nemusí nic učit a má volné ruce i zrakové pole.

Konkrétní a velmi podobná aplikace detektoru ionizačního záření je pro běžného člověka, například jednoduchý akustický indikátor volného místa při couvání auta, při kterém se nevyskytuje při zvětšení frekvence stimulu radioaktivní záření, ale například hrana chodníku.

Při tvoření audio displeje máme možnost při změně vstupních dat měnit vlastnosti stimulu, tj. frekvenci, fázi, amplitudu, spektrální vlastnosti či stimul samotný. Je vhodné vytvořit takové mapování, které bude intuitivní a přitom dostatečně přesné.

Při tvoření audio displeje pouze s jedním rozměrem je vhodné klást si otázky, jak by se měla změnit výstupní vlastnost zvukového stimulu, pokud se vstupní hodnota zvětší/zmenší dvojnásobně atp. Geiger-Müllerův detektor mění hned dvě vlastnosti, frekvenci stimulu a jeho intervaly, čímž se docílilo přesvědčivého oznámení výskytu nebezpečí. Při dvou či více vstupních hodnotách je nutné brát v úvahu vzájemnou kooperaci výstupních vlastností stimulu tak, aby byly zřejmě oddělitelné, či se v případě detektoru ionizačního záření vhodně doplňovaly.

4.3.1 Výhody audio displejů

Zajímavý je výčet výhod audio displejů při srovnání očí a uší obecně:

- Uši mají oproti očím tu výhodu, že se na rozdíl od nich nedají zavřít a unaví se poslechem méně nežli oči pozorováním, tudíž je teoreticky můžeme využívat neustále. Tento fakt však závisí u obou dvou případů na síle signálu.
- Uši, oproti očím, díky své citlivosti na dočasné změny odkrývají některé informace, které mohou být pro zrakové pole zamaskované [13].
- Uši vnímají všesměrově, zatímco oči pouze z jednoho směru pod určitým prostorovým úhlem.
- Lidský sluch nám umožňuje poslech více kanálů najednou, a to i ve více frekvenčních pásmech. Jednotlivé složky pak můžeme selektivně vyfiltrovat uvnitř hlavy podle toho, která nás právě zajímá. Tomuto jevu se odborně říká Coctail Party effect [16], neboli volně přeloženo jako fenomén koktejlové párty, na které hraje současně hudba a hovoří více lidí najednou. I tak jsme schopni při soustředěném poslechu slyšet osobu, se kterou právě hovoříme a utlumit ty, které zrovna neposloucháme. Podobně se můžeme soustředit při poslechu muziky na jednotlivé nástroje.
- Pokud srovnáme bitový tok signálu audia a videa, audio jej má podstatně menší. Při standardním rozlišení nekomprimovaného stereofonního audiosignálu (kvantizace 16ti bity, vzorkovací frekvence 44.1 kHz) dostaneme rychlost 0.1764 Mbps, tuto rychlost dosahují pouze vysoce komprimovaná videa.

K omezujícím vlastnostem audio displejů naopak patří paradoxně skutečnost, že omezují sluchové pole. Je tedy vhodné volit takový stimul, který bude interferovat co nejméně s okolním prostředím při plnohodnotném plnění své funkce. Kombinování různých audio displejů ve stejný moment se taktéž nedoporučuje, jelikož se člověk může dokonale soustředit pouze na jeden a při vzájemné interferenci by se mohly rušit.

4.3.2 Rozdělení audio displejů

Sonifikace a audio displeje jsou velmi úzce propojené. Sonifikaci ale spíše chápeme jako vědní obor, zatímco audio displeje jako jeho konkrétní aplikaci. Audio displeje můžeme podle typu funkce rozdělit do 4 základních kategorií [13]:

- *Alarmy, upozornění a varování*

Indikace něčeho, co se stalo či se stane v blízké budoucnosti. Tento typ akustického signálu je informačně velmi chudý a často oznamuje pouze jednu událost. Příkladem může být domovní zvonek či oznámení mikrovlnky o ohřátém jídle.

- *Status, proces, monitorující zprávy*

Na rozdíl od předchozí skupiny je zvukový signál opakován v určitých intervalech a jeho charakteristiky se mění v závislosti na vstupních datech. Příkladem může být Geigerův-Müllerův detektor zmíněný výše či nemocniční přístroj pro monitorování životní činnosti pacienta.

- *Průzkum dat*

Tyto audio displeje jsou v podstatě ty, kterých se sonifikace týká nejvíce. Zpracovávají vlastnosti dat a převádějí je podle určité rovnice či modelu do akustického signálu. Typ a účinnost záleží hlavně na konkrétním modelu, v každém případě podává výstup větší informační hloubku, nežli dvě předchozí skupiny.

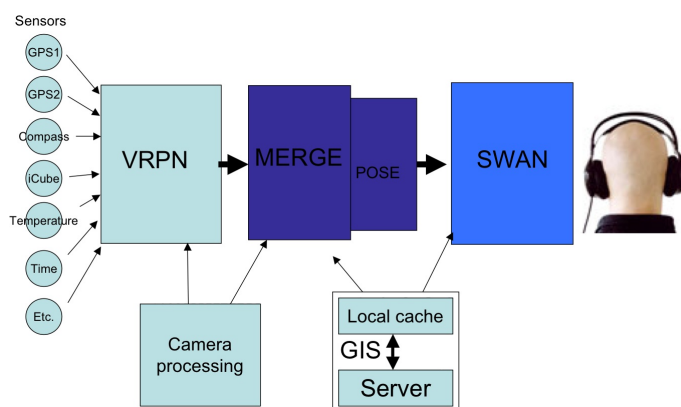
K těmto kategoriím byla postupně přidána i kategorie poslední, a to

- *Umění, zábava, sport a cvičení*

Je zřejmé, že v dnešní době multimédií se sonifikace uplatní i v těchto odvětvích. Možností využití je nespočetné množství právě tam, kdy se chtějí uživatelé soustředit na jiný než vizuální vjem. Příkladem mohou být počítačové hry.

4.3.3 SWAN

Jeden z příkladů audio displeje vyvíjeného technickou univerzitou v Georgii (Georgia Institute of Technology) je systém SWAN (System for Wearable Audio Navigation - systém pro přenosnou osobní audio navigaci) [17]. Navigační systém využívá mnoho technologií spojených v jednu funkční jednotku, např. GPS (Global Positioning System - globální polohovací systém), infračervený vysílač, pedometr, RFID (Radio Frequency Identification - identifikace na rádiové frekvenci), akcelerometry, kompas a v neposlední řadě databázi HRTF spolu s navigačními stimuly pro orientační body a ostatními zvuky pro identifikaci okolních objektů. Nutno dodat, že systém GPS není pro navigační účely chodců co do preciznosti příliš vhodný, a jeho chyba se v horizontální rovině pohybuje kolem 3 m. Systém navíc mění svou polohu skokově [18]. I tak je systém SWAN díky vzájemné kooperaci všech těchto technologií podle slov autorů pro účely navigace dostatečně spolehlivý. Tento systém tedy pro zvukovou navigaci využívá HRTF databáze. Využívá také různých stimulů pro identifikaci navigačních bodů, dále objektů, které se vyskytují po cestě a zvuky pro signalizaci změny povrchu.

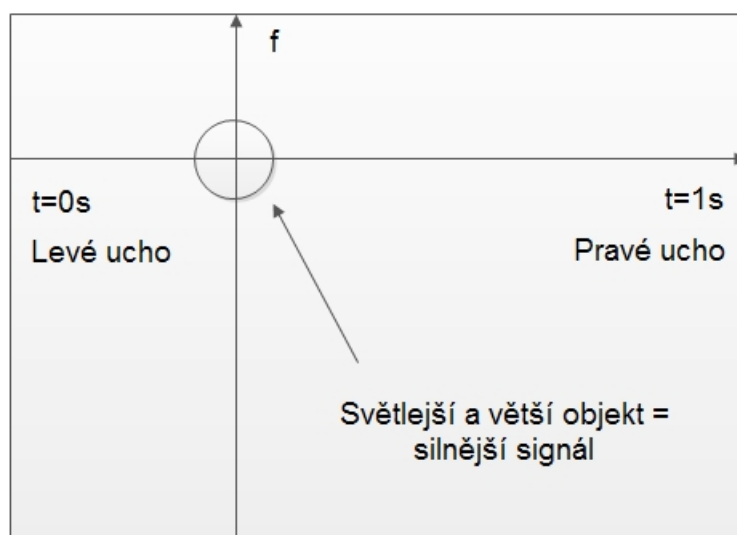


Obrázek 10: Navigační systém SWAN [17]

Funguje na jednoduchém principu: Nejdříve nalezne nejjednodušší cestu bez překážek a vytvoří tak virtuální mapu s vyznačenou cestou s lineárními úseky. Naviguje uživatele na nejbližší navigační bod dělící dva úseky a poté, co jej dosáhne, respektive nachází se v přijatelném radiusu, změní lokaci navigačního bodu na bod následující. Navigace s využitím stereofonních sluchátek funguje jako audio displej s polohováním díky databázi HRTF, zvukových stimulů a jednoduchému signálovému processingu. Pro navigační body jsou v rámci testování [19] užity stimuly sinus 1 kHz (úzké spektrum), krátký úsek růžového šumu (široké spektrum) a zvuk sonaru, přičemž při uskutečněném pokusu autoři v jejich závěru upřednostňují zvuk sonaru pro jeho spektrální vlastnosti a i z hlediska příjemnosti pro uživatele, v testování má i decentně lepší výsledky než ostatní stimuly.

4.3.4 vOICE

Dalším zajímavým navigačním audio displejem, který je již plně funkční, je systém vOICE [20]. Byl primárně vyvinut pro navigaci nevidomých. Princip audio displeje je jednoduchý, využívá převodu obrazového záznamu kamery na akustickou informaci, která se mění v závislosti na jeho obsahu. Černobílý 2D obraz s přednastaveným rozlišením 176x64 bodů je předán jako sekundový zvukový stereo signál, který připomíná skener. Začátek obrazu je oznámen earconem v podobě pípnutí, obraz se vyčítá po jednotlivých sloupcích zleva doprava rovnoměrně v rozmezí jedné sekundy. Výraznější světlé objekty se poté projeví jako akustický signál, jehož frekvence závisí na vertikální poloze objektu v obraze, tedy čím se nachází výše, tím je i vyšší frekvence. Jedná se tedy o jednoduché frekvenční elevační mapování. Poloha objektu v horizontální rovině je mapována pomocí algoritmů založených na duplexní teorii, zároveň se tento virtuálně napolohovaný stimul spouští v moment, který je lineárně závislý na poloze objektu v horizontální rovině. Objekt nacházející se v jedné čtvrtině šířky obrazu se tedy přehrává vždy za 250 ms po úvodním pípnutí. Síla signálu pak mapuje světlost a velikost objektu. Systém je pak použitelný jako samostatná jednotka s integrovanou kamerou ve slunečních brýlích tak, aby přibližně zachycovala obraz, který by za normálních podmínek vnímaly oči, tak i např. jako aplikace pro mobilní telefon, což výrazně snižuje náklady a umožňuje vyzkoušet systém komukoliv. Výhody v podobě jednoduchosti aplikace jsou zřejmé, zůstává však otázkou, jak si na tento systém zvyknou samotní uživatelé, kteří v praxi nemanipulují pouze se světlými objekty.



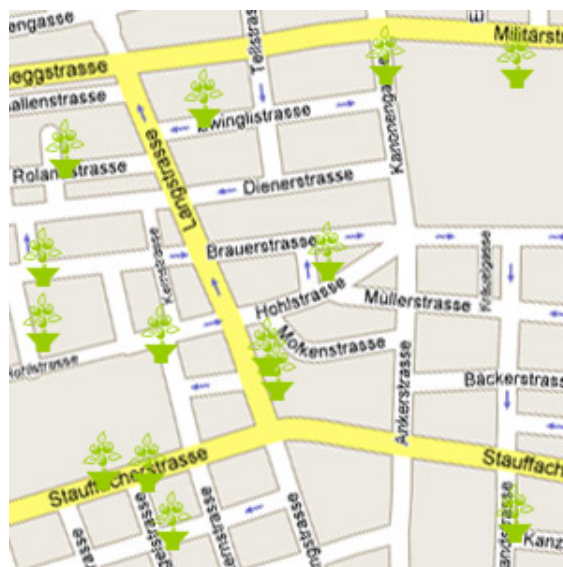
Obrázek 11: Navigační systém vOICE

4.3.5 Ostatní navigační audio displeje

Audio displejů možných k použití za účelem navigace je celá řada. Ty pak většinou využívají k navádění uživatele polohování určitého stimulu podle principů duplexní teorie či databáze HRTF, liší se však přístupem ke zmapování trajektorie. Ten bývá založený buď na základě GPS [17, 21], či zmapování jiným systémem, například pomocí kompasu a akustické informace o světové straně [21]. Jako společné pak mají vytvoření trasy po rovných úsecích, v jejichž spojích se nachází virtuálně napolohovaný stimul, většinou earcon, který se přesune do dalšího bodu poté, co se k němu uživateli podaří přiblížit na určitou vzdálenost.

Pokud nemají předem vyznačenou trajektorii, spoléhají se na GPS polohu uživatele. The roaring navigator [23] je audio displej kombinovaný s audio průvodcem navržený pro prostředí zoologických zahrad. Pokud se uživatelé dostanou do oblasti v blízkosti výběhu zmapovaného na virtuální mapě, ozve se z jeho směru typický zvuk pro zvíře ve výběhu. Hlavní výhodou tohoto systému je jeho použitelnost v chytrých mobilních telefonech, které jsou dnes již standardně vybaveny lokalizačním systémem. Tímto není zapotřebí výroby dalšího specializovaného hardwaru. Na podobném principu funguje i projekt Tactical Sound Garden [24] s tou výjimkou, že se pro lokalizaci uživatele nevyužívá systému GPS, ale síla signálu přístupových bodů strategicky rozmístěných Wi-Fi sítí.

Navigační audio displej lze také využít k navigaci po menu tam, kde se potřebujeme soustředit na vizuální složku jiné činnosti, například při řízení auta [22]. Pro pohyb po menu se zde využívá jednoduchých earconů, které se mění v závislosti na poloze uživatele v menu.



Obrázek 12: Tactical Sound Garden - Přístupové body v části Curychu [24]

4.4 Prostorové vnímání kmitočtů

V roce 1929 americká psycholožka Carroll C. Prattová vydala publikaci, ve které popisuje lidské prostorové vnímání různě vysokých tónů [1]. Provedla následující pokus. Skupině deseti lidí pouštěla v náhodném pořadí pět sinusových signálů s oktávovými rozestupy (odstup dvojnásobné frekvence), konkrétně 256 Hz až 4096 Hz. Testovaní měli do dotazníku vyplnit, jak vysoko na dvoumetrové škále (rozdělené do 15 úseků) umístěné na stěně před nimi se jim daný tón na stupnici od jedné do patnácti jevil. Tóny se pouštěly z šesti různých pozic v horizontální rovině za touto stěnou, patrně se jednalo o plátno. I když ze začátku trvalo testovaným osobám, než si na systém zvykly, později vyplňovaly dotazníkovou formou tabulky takřka automaticky. Výsledky jsou vidět na originální tabulce převzaté z původní publikace, písmena A-F označují horizontální pozice reproduktorů za plátnem a čísla pod nimi průměrný výsledek (vertikální pozici) každého tónu. Ačkoliv není zcela zřejmé, jaký postup Prattové volila pro uskutečnění testu, z jejích výsledků můžeme předpokládat, že člověk prostorově vnímá vyšší tóny výše než tóny nižší.

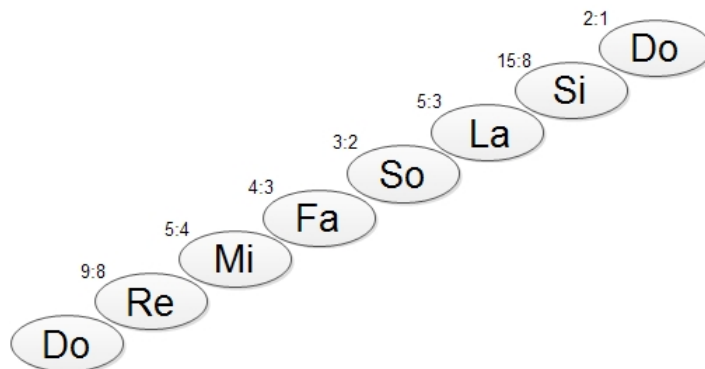
Pitch	os					
	A	B	C	D	E	F
4096.....	12.4	10.4	13.6	13.4	10.0	14.4
2048.....	9.4	9.4	10.7	11.0	9.1	11.8
1024.....	7.8	8.3	8.8	8.0	8.4	9.7
512.....	6.4	7.1	7.2	7.4	6.8	6.9
256.....	4.6	6.2	6.4	5.4	5.8	1.9

Obrázek 13: Výsledky měření C.C.Prattové [1]

Tuto skutečnost lze pravděpodobně vysvětlit tak, že jsou lidé navyklí na spektrální filtraci při vnímání elevace, přičemž při elevacích vyšších má přijatý signál filtrovaný sluchovou soustavou větší podíl vysokofrekvenční složky nežli při elevacích nižších. Svoji roli hraje pravděpodobně ale i psychologická stránka.

4.5 Diatonická durová stupnice

Diatonická durová stupnice je založená na kombinaci intervalu malé sexty (8:5) na šestém a velké septimy (15:8) na sedmém stupni [29]. Toto platí pro čisté, didymické ladění, které je také označováno za přirozené ladění. Přirozené proto, že ostatní tóny ladíme na intervalech odvozených z poměrů přirozených čísel vůči tónu základnímu. Ladit hudební nástroj můžeme však i temperovaně. Pokud se budeme zabývat konkrétními frekvencemi, přirozené ladění je nejčastěji založeno na intervalech mezi oktávou (2:1), kvintou (3:2) a velkou tercií (5:4).



Obrázek 14: Solmizační slabiky a jejich přirozené poměry

Frekvence jednotlivých tónů se však v jednotlivých stupnicích liší v závislosti na základním tónu, pokud bychom měli přesně dané frekvence, některé stupnice by zněly rozladěně. Proto pro zjednodušení používáme i ladění rovnoměrné temperované. Vycházíme z toho, že oktáva je dvojnásobek frekvence a na škále od základního tónu k oktávě se nachází 12 tónů s půltónovými intervaly. Abychom skočili právě o půltón nahoru, musíme daný tón vynásobit dvanáctou odmocninou oktávy, tj. dvanáctou odmocninou ze 2. Následující rovnice ukazuje výpočet frekvence tónu

$$f_{ton} = f_{base} \cdot (\sqrt[12]{2})^{pozice}. \quad (9)$$

V tabulce 2 jsou vyznačené tóny, se kterými budeme pracovat (při stupnici C-dur jsou to CDEFGA_Hc, tzn. vycházíme z komorního A = 440 Hz) a jejich vzájemné poměry frekvencí. Jsou hlavně vidět rozdíly mezi temperovaným a přirozeným laděním v centech, ten je definován jako 1/100 temperovaného tónu, tedy 1/1200 oktávy (temperovaných tónů je 12), vzdálenost mezi dvěma půltóny je sto centů, přičemž vzdálenost o velikosti jednoho centu je tedy

$$1c = \sqrt[1200]{2} = 1,0005777895. \quad (10)$$

Tato vzdálenost je lidskému uchu nepostřehnutelná, JND se pohybuje podle vytrénovanosti od 4 centů výše. Vzhledem k tomu, že vzdálenost mezi jednotlivými půltóny budou dosahovat pro obě dvě ladění kolem 100 centů, dostáváme se na spolehlivou hranici rozeznání dvou stimulů s různou základní frekvencí. V C durové diatonické stupnici má největší rozdíl mezi jednotlivými ladeními tón A, který se liší 15.71 centu v absolutní hodnotě mezi laděním přirozeným a temperovaným, to bude mít za následek i rozdílné vnímání stupnice. V tabulce 2 nalezneme poměry vůči základní frekvenci prvního tónu, přičemž interval se mezi frekvencemi s každým tónem zvětšuje (zmenšuje, pokud jdeme po stupnici směrem dolů). Základní frekvence stimulu pro naše účely měla být proto dostatečně vysoká. Je známo, že člověk slyší lépe rozdíly mezi tóny vysokými než mezi tóny nízkými, proto se například lépe ladí kytara než kytara basová.

Přirozené ladění zlomkem	Tón při základním tónu C	Přirozené ladění	Temperované ladění	Odchylka v centech
1	C	1	1	0
16/15	Cis	1,0666	1,05946	11,72
9/8	D	1,125	1,12246	3,91
6/5	Dis	1,2	1,20921	15,64
5/4	E	1,25	1,25992	13,69
4/3	F	1,3333	1,33484	2
45/32	Fis	1,40625	1,41421	9,78
3/2	G	1,5	1,49831	1,96
8/5	Gis	1,6	1,5874	13,69
5/3	A	1,6666	1,68179	15,71
16/9	B	1,7777	1,7818	3,99
15/8	H	1,875	1,88775	11,73
2	c	2	2	0

Tabulka 2: Poměry mezi hudebními tóny

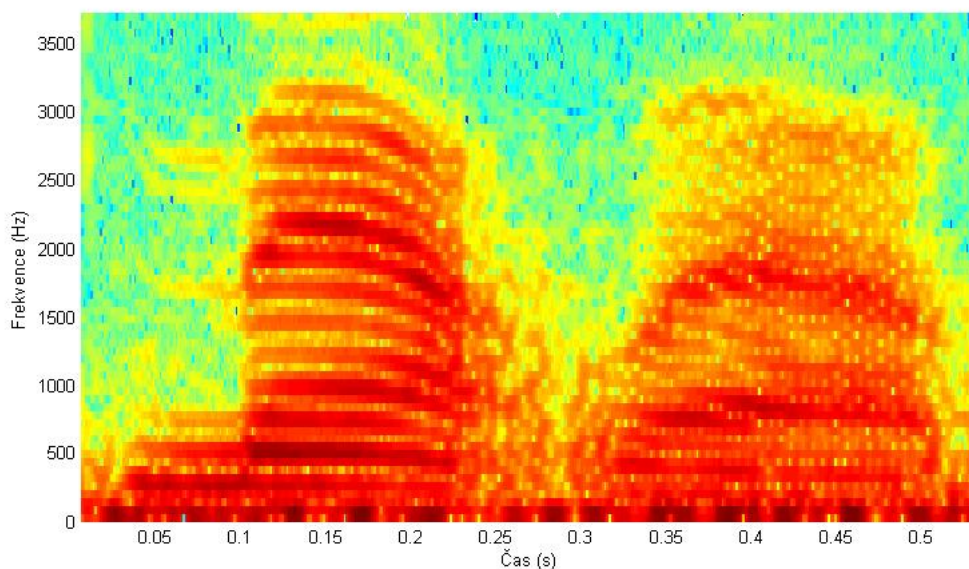
Frekvenčně závislá JND vychází pro frekvence menší než 500 Hz zhruba 3 Hz a pro frekvence větší než 1000 Hz cca 1 Hz [8]. Pokud si však uvědomíme, že například rozdíl mezi tónem a a $a1$ je 220 Hz a mezi tóny $a2$ a $a3$ 880 Hz, je zřejmé, že čím vyšší tón, tím větší frekvenční rozdíl vůči dalšímu půltónu. Pokud uvažujeme použití přeladování stimulů pomocí diatonické durové stupnice, je lepší vybírat stimuly se základní frekvencí vyšší, ideálně kolem 1 kHz, čímž se spodní oktávou dostaneme na nejméně 500 Hz. Tímto způsobem dostaneme lepší odhad tónů u uživatelů, jelikož s větší frekvencí základního tónu budou větší intervaly mezi tóny po celé stupnici. Při přeladování může vzniknout problém s přeladováním dolů, kdy rozdíly v základních frekvencích nejsou tak znatelné, jako mezi vysokými tóny.

4.6 Pitch shifting

Pitch shifting (volně přeloženo jako posuv výšky) je technika, která přeladí základní frekvenci zvukového signálu, aniž by změnila jeho délku [14]. To lze nejjednodušeji provést umělým natáhnutím či zkrácením signálu pomocí přidání či odebrání vzorků (Time Stretching - natahování času) a poté jeho převzorkováním (Resampling), složitější případy pak pomocí fázového vokodéru či jiných technik. Používá se zejména v hudebním průmyslu.

4.6.1 Základní tón signálu

K detekci základního tónu harmonických signálů se využívá jejich periodicity, v signálech se nachází okamžiky s maximální amplitudou, jejichž převrácená perioda je považována za základní tón signálu. Většina reálných zvukových signálů však nejsou harmonické, ale kvaziperiodické. Navíc se základní tón v průběhu signálu může měnit, v některých signálech navíc ani nemusí být k nalezení. Na obrázku níže je spektrogram hlasového signálu, konkrétně slova "Matlab", ze kterého není základní perioda signálu zřejmá.



Obrázek 15: Spektrogram řečového signálu

Informaci o této periodě získáme například pomocí reálného keprtra, které nám zvýrazňuje změny a umožňuje detekovat periodicity [14]. Kepstrum získáme podle rovnice

$$c(n) = \text{Re} \{ DFT^{-1} [\ln |DFT(s(n))|] \}. \quad (11)$$

DFT (Discrete Fourier Transform - diskretní Fourierova transformace) z předchozí rovnice získáme podle

$$DFT(s(n)) = S(k) = \sum_{n=0}^{N-1} s(n) \cdot e^{-i2\pi kn/N} \quad (12)$$

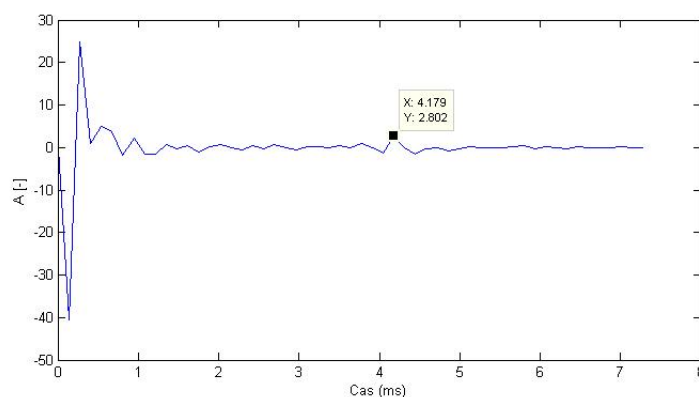
a inverzní DTF transformaci pomocí rovnice

$$DFT^{-1}(S(k)) = s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) \cdot e^{i2\pi kn/N}. \quad (13)$$

Z reálného spektra nás poté zajímá první výrazná špička za počátečním skokem, která značí periodicitu. Po zjištění jeho polohy, která udává periodu, získáme frekvenci jejím převrácením podle

$$f_0 = 1/T_0. \quad (14)$$

Pro řečový signál vykreslený ve spektrogramu výše vypadají první koeficienty jako na obrázku 16. První výrazná špička za klesajícím spektrem se nachází na 4.179 ms, z čehož nám vyjde základní frekvence signálu 239.29 Hz.



Obrázek 16: Reálné kepstrum řečového signálu

Další jednoduchou možností zjištění základního tónu signálu je pomocí detektoru nul. Harmonický signál projde za periodu dvakrát nulou, pokud vhodným způsobem odfiltrujeme vyšší harmonické a šumovou složku, můžeme vypočítat základní frekvenci podle

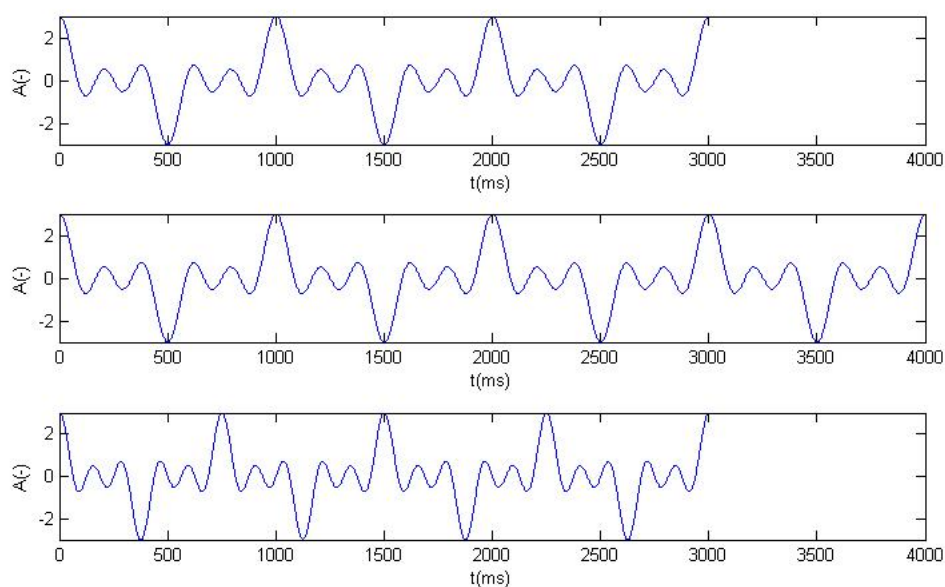
$$f_0 = \frac{1}{2} f_{vz} \cdot n_{nuly} / n_{vzorku}. \quad (15)$$

Poslední často používaná metoda pro zjištění základní frekvence je pomocí autokorelace. Cílem je najít takový fázový posun, aby korelace byla co nejvyšší. Tento posun odpovídá jedné periodě, na kterou můžeme znovu aplikovat rovnici 14.

Ačkoliv je dobré znát základní frekvenci vzorku, ve skutečnosti ji pro metodu pitch shifting, která si vystačí se zadaným poměrem vůči původnímu signálu, nepotřebujeme znát. Slouží pouze pro představu a pro srovnání se základními tóny ostatních vzorků.

4.6.2 Time Stretching/Resampling

Tato metoda je založena na základním principu vysvětleném na obrázku 17, který popisuje zvýšení základní frekvence. První graf znázorňuje původní signál. Pokud chceme jeho frekvenci posunout o $4/3$, musíme vzorek nejdříve roztáhnout (Time Stretching), v tomto konkrétním příkladě přidat jednu periodu, což je vidět na prostředním grafu. Poté jej musíme zkrátit na délku původního signálu, což zajistíme podvzorkováním (Downsampling - snížení vzorkovací frekvence). Výsledek je vidět na spodním grafu obrázku.



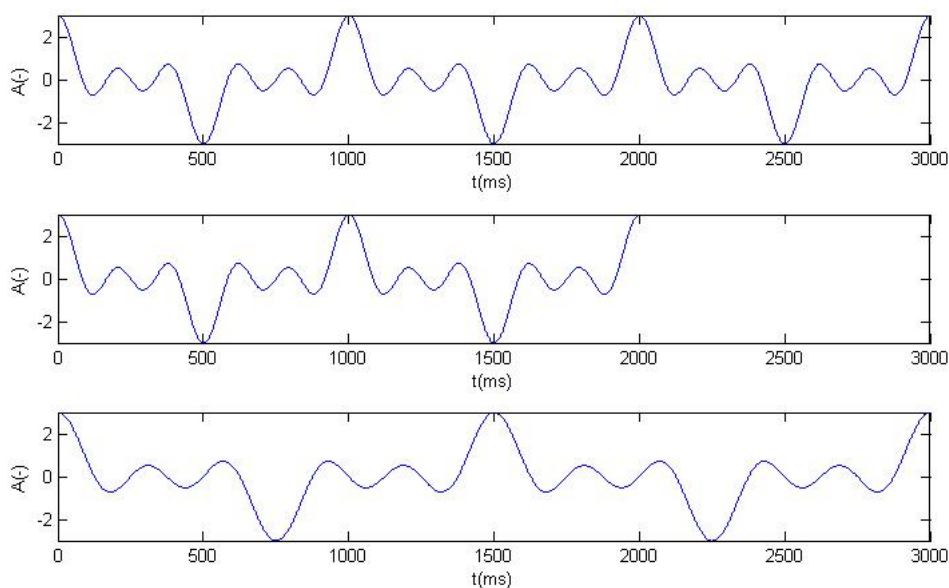
Obrázek 17: Princip zvýšení základní frekvence

Na obrázku 18 vidíme opačný postup pro snížení základní frekvence. V prvním grafu se nachází vzorek signálu, jehož frekvenci chceme zmenšit o $1/3$. Vzorek nejdříve zkrátíme, v tomto konkrétním příkladě ubereme jednu periodu, jak je tomu na prostředním grafu, a poté jej pomaleji vyčíst, tedy zvýšit vzorkovací frekvenci (Upsampling). Výsledný signál je vidět na posledním grafu.

Ukázali jsme si ale tuto metodu na syntetickém sudém periodickém signálu, ve kterém jsme ubírali a přidávali celé periody, čímž jsme si značně ulehčili práci. Pro signály složitější by se nám tento postup podařil jen velmi těžko. Je nutné provést frekvenční analýzu a vzorky přidávat či ubírat nikoliv v oblasti časové, ale v oblasti frekvenční. Také je důležité zvolit správný postup. Při změně všech frekvencí o konstantní hodnotu docílíme kmitočtového posunu (Frequency Shifting), ale jednotlivé harmonické mezi sebou již nebudou mít stejné poměry a výsledný charakter zvuku bude disharmonický. Je proto třeba posouvat podle předpisu

$$f_i \text{ preladena} = v \cdot f_i \text{ puvodni}, \quad (16)$$

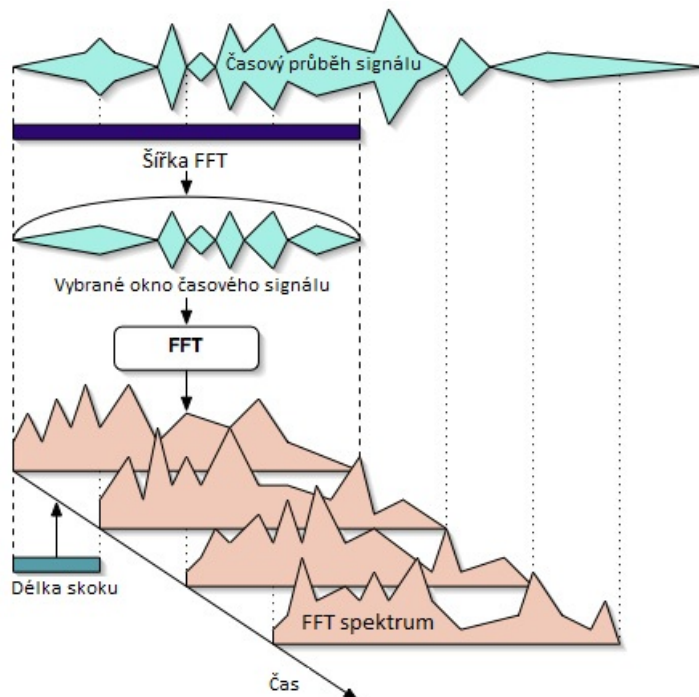
kde i značí číslo harmonické složky a v poměr původní a přeladěné základní frekvence signálu, kdy pro $v > 1$ se frekvence zvětší a pro $v < 1$ zmenší. Tímto se zachovají jednotlivé poměry všech harmonických a signál si zachová svoji charakteristiku. V tomto případě se však mění i délka signálu spolu s poměrem v a změnou frekvenční osy změním stejným poměrem i osu časovou, čehož můžeme docílit i změnou rychlosti přehrávání.



Obrázek 18: Princip snížení základní frekvence

4.6.3 Fázový vokodér

Princip fázového vokodéru si lze představit jako spektrogram signálu, kterému natahujeme časovou osu, čímž měníme časové charakteristiky, zatímco krátkodobé časové charakteristiky zůstávají zachovány. Signál nejdříve zanalyzujeme převedením na segmenty, na které aplikujeme FFT (Fast Fourier Transform - rychlá Fourierova transformace), posléze provedeme příslušnou transformaci upravující fáze jednotlivých segmentů a nakonec provedeme syntézu jednotlivých segmentů zpět do časové oblasti [14].



Obrázek 19: Princip STFT [25]

Nejdříve ze všeho je nutné převést reálný signál z časové do frekvenční oblasti. Na to použijeme krátkodobou Fourierovu transformaci (Short-Time Fourier Transform - STFT), která nám na výstup vrátí v podobě matice v podobě segmentů s informací o amplitudě a fázi podle rovnice

$$X(n, k) = \sum_{m=0}^{N-1} x[m] w[n-m] e^{-jk \frac{2\pi}{N} m} = |X(n, k)| e^{j\varphi(n, k)}, \quad (17)$$

kde $X(n, k)$ je časově proměnným spektrem signálu $x[m]$ a udává modul $|X(n, k)|$ a fázi $\varphi(n, k)$ pro časový okamžik n jednotlivých segmentů k v rozmezí celých čísel v intervalu $\langle 0, N-1 \rangle$.

Signálem $w[n - m]$ se rozumí okno, neboli časový interval, ve kterém se STFT počítá. Toto okno musíme zvolit uvážlivě, neboť při příliš krátkém okně budeme mít rozmazané spektrum a naopak při dlouhém okně menší možnost lokalizace v čase. Okno použijeme libovolné nepravoúhlé. Bude-li mít vstupní signál $x[m]$ přesně úhlový kmitočet Ω_k , fáze $\varphi(n, k)$ bude v čase konstantní. Její změna v čase značí odchylku od středního úhlového kmitočtu Ω_k k -tého pásma. Důležitým faktem je, že změna fáze $\varphi(n, k)$ mezi dvěma časovými okamžiky určuje okamžitý kmitočet.



Obrázek 20: Princip fázového vokodéru

Pokud chceme změnit výšku signálu, určíme nejdříve změnu fáze vůči předchozímu segmentu

$$\Delta\varphi(n, k) = \varphi(n, k) - \varphi(n - 1, k) \quad (18)$$

a tuto změnu vynásobíme poměrem požadované změny výšky v vůči původnímu signálu a přičteme k aktuální hodnotě fáze celkového výstupního vzorku

$$\Psi(n + 1, k) = \Psi(n, k) - v\Delta\varphi(n, k). \quad (19)$$

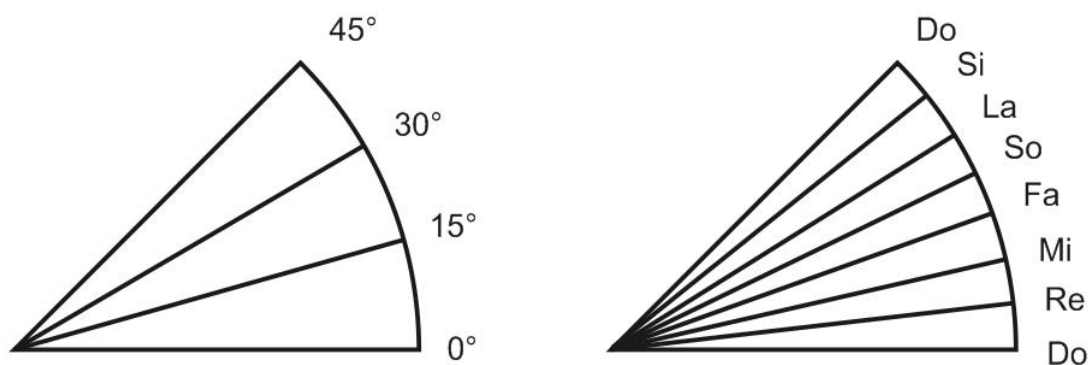
Zpětnou ISTFT (Inverse STFT - Inverzní STFT) dostaneme podle vzorce

$$y[n] = \frac{1}{Nw(0)} \sum_{k=0}^{N-1} X(n, k)e^{j\Psi(n, k)} \quad (20)$$

Touto rovnicí získáme zpátky signál fázově posunutý ve spektru, čímž se z původního signálu stal signál přeladěný, aniž by se výrazně změnila charakteristika signálu.

5 Implementace audio displeje

Na obrázku 21 je vidět jeden z důvodů vytvoření navigačního audio displeje na principu prostorového vnímání frekvencí zvukového stimulu. Pro virtuální simulaci elevace je zapotřebí spektrální filtrace, buďto pomocí HRTF nebo fyzikálního modelu, který předpokládá odrazy od ramen, ušních boltců etc. Přesnost těchto metod však dává ve vertikální rovině špatné výsledky, JND je v nejlepším případě kolem 15° , zatímco JND reálné lidské lokalizace je pro elevaci rovno kolem 10° [10]. Při použití HRTF musíme nejdříve změřit impulsové odezvy pro veškeré polohy (elevace a azimut), ve kterých chceme simulovat virtuální zvukový zdroj, případně využít některou z již naměřených databází, což však vertikální rozlišení zhorší. Strukturální model ucha je na tom obdobně, spíše hůře.



Obrázek 21: Srovnání JND elevace HRTF a diatonické durové stupnice

Pro zpřesnění lokalizace zvukového stimulu za účelem navigace tedy můžeme vytvořit alternativní mapování elevace. Předpokládejme elevaci 0° pro základní frekvenci stimulu. Tu zvolíme i jako základní tón diatonické durové stupnice popsané v kapitole 4.5. Při přeladování pomocí pitch shiftingu [26] a hodnot z tabulky 2 výše se každý následující vyšší tón rovná zvětšení elevace. Pro náš účel, zvětšení či zmenšení základní frekvence znamená i zvětšení či zmenšení elevace, tento jev nám výrazně pomůže při adaptaci uživatelů na samotný systém vycházející z psychologického předpokladu, že člověk akusticky vnímá zvětšování vstupní veličiny jako zvětšování frekvence tónu [32]. Pokud použijeme audio displej s diatonickou durovou stupnicí, rozdíl mezi jednotlivými úrovněmi elevací (při použití oktáva = 45°) je cca 6.4° ($45^\circ/7$), což znamená, že pokud se při lokalizaci polohy či navigaci směru pomocí audio displeje spletete o jeden tón, máme stále lepší rozlišení než u metod virtuálního polohování elevace při jejich nejlepších výsledcích. Pokud se uživatel netrefí o dva tóny, JND elevací jsou pak srovnatelné, tento příklad počítáme s nejlepším možným JND pro individuálně naměřenou HRTF.

5.1 Využití diatonické durové stupnice

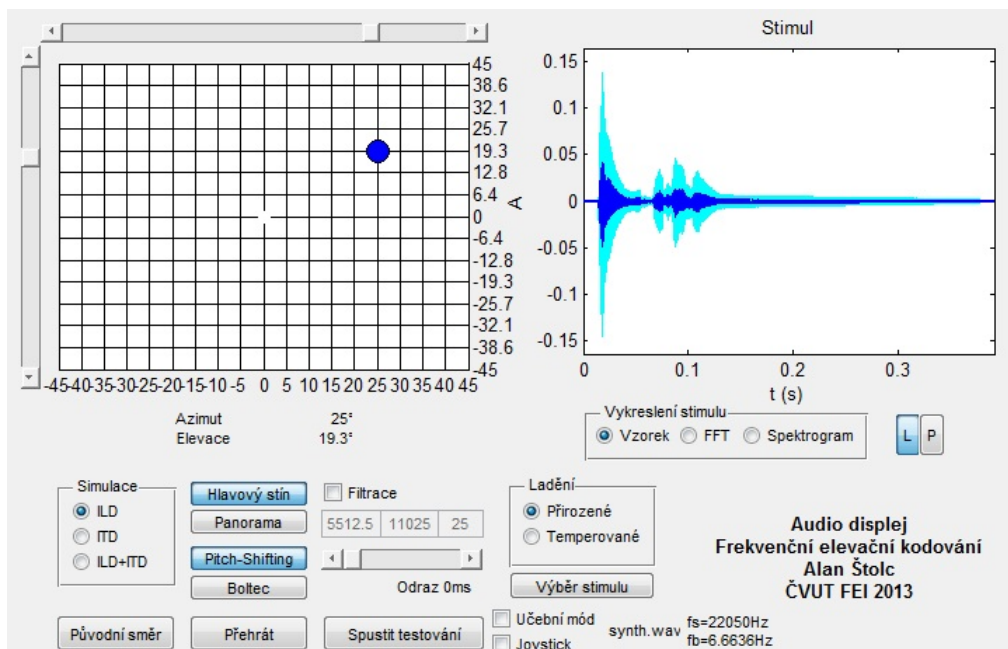
Při mapování jednotlivých elevačních kroků je důležité rozmyslet jejich frekvenční vzdálenost. Pokud bychom přeladřovali po oktávách, po kterých například přeladřovala ve svém pokusu C.C.Prattová [1], subjektivní kvalita stimulu přeladřěného o 4 oktávy by výrazně klesla a stimul by se mohl změnit v nepříjemné pištění. U tónů či earconů nám to nevádí, ale pokud chceme použít uživatelsky příjemný stimul či stimuly, který si vybere sám uživatel, musíme zvolit menší intervaly. Jako ideální se jeví diatonická durová stupnice, kterou má v mysli zafixovanou většina uživatelů (solmizační slabiky Do-Re-Mi-Fa-So-La-Si-Do), jejíž intervaly mezi jednotlivými tóny jsou dostatečně velké na rozeznání dvou poloh, ale zároveň nejsou tak daleko od sebe, aby přeladřěný stimul zněl uživatelsky nepříjemně ve vyšších polohách. Durová (veselá) stupnice byla zvolena z důvodu uživatelské přívětivosti, navíc je vžitější nežli stupnice mollová (smutná), jedná se (při základním tónu C) o stejnou stupnici, kterou můžeme zahrát na bílých klávesách klavíru. Jako základní tón vezmeme základní frekvenci stimulu a přeladřujeme podle potřeby. Předpokládáme standardní stupnici o 8mi tónech, zvolený úhel 45° se jeví pro oktávu jako nejvhodnější a nejjednodušší, uživatel v tomto rozmezí lokalizuje zvuk nejlépe. Elevace může jít samozřejmě i do záporných hodnot, základní frekvence stimulu se však musí vydělit vhodným poměrem v opačném pořadí (Do-Si-La-So-Fa-Mi-Re-Do). Tento princip můžeme spojit se známými polohovacími metodami azimutu uvedenými v kapitole 3, čímž nám vznikne jednoduchý audio displej, který bude polohovat zdroj zvuku v horizontální rovině pomocí ILD a ITD a v rovině vertikální pomocí změny základní frekvence stimulu, implementace je velmi jednoduchá.

5.2 Audio displej s elevačním kódováním

Vytvoříme jednoduchý audio displej, který bude přeladřovat základní frekvenci stimulu metodou pitch shifting v závislosti na elevaci. Metoda pitch shifting bude provedena pomocí fázového vokodéru a jako mapování bude použita diatonická durová stupnice s oktávou ve 45° , jednotlivé elevační kroky tedy budou mít odstup 6.4° . Pro tento účel, zvětšení či zmenšení základní frekvence znamená i zvětšení či zmenšení elevace, tento jev výrazně pomůže při adaptaci uživatelů na samotný systém. Výše vytvořený audio displej bude fungovat pouze pro polovinu před uživatelem pro horizontální i vertikální rozmezí -45° až 45° , jelikož zorné pole uživatele se nachází zhruba v této oblasti a výsledky měření v ní budou pravděpodobně nejzajímavější. Pro virtuální polohování v horizontální rovině bude využito základních principů ILD a ITD, které jsou pro navigační účely dostatečně přesné. Jednotlivé kroky v horizontální rovině budou mít odstup přibližné azimutální JND uživatelů, tedy 5° .

5.3 Přiložený program

Pro jednoduchou implementaci testovacího prostředí v GUI (Graphical User Interface - grafické uživatelské rozhraní) byl využit program Matlab, konkrétně jeho rozhraní GUIDE (Graphical User Interface Development Environment - Vývojové prostředí GUI). Byl zvolen pro nenáročnost na programovací zručnost a pro implementované výpočetní knihovny. Při implementaci byly použita knihovna Signal Processing Toolbox, obsahující například funkce `spectrogram` pro jednoduché vytvoření spektrogramu či `fir` pro vytvoření FIR filtru (Finite Impulse Response - konečná impulsová odezva). Dále byly implementovány převzaté algoritmy strukturálního modelu boltce [27], fázového vokodéru [33] a podpůrné programy pro ovládání joysticku [34], z důvodu časové úspory.

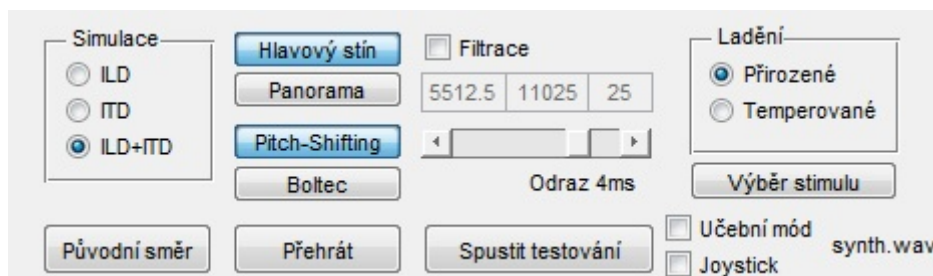


Obrázek 22: Testovací software

Vytvořené grafické rozhraní programu má dvě hlavní funkce. První z nich je vyzkoušení různých stimulů pro navigační potřeby, je koncipován pro nahrání a jednoduchou manipulaci se stimulem při současném pozorování a poslechu vzorku z různých poloh za použití algoritmů zmíněných v kapitole 3 pro simulaci ILD a ITD v horizontální rovině či strukturálního modelu boltce a frekvenčního elevačního kódování v rovině vertikální. Druhou funkcí je pak samotné testování lokalizace stimulů virtuálně polohovaných různými technikami a nabízí tedy nepřímou i možnost srovnání audio displeje s přeladovaným stimulem a stimulem polohovaným pomocí strukturálního modelu boltce.

5.3.1 Nastavení a ovládání simulačního programu

Program umožňuje celou řadu nastavení a variací virtuálního polohování zvukového zdroje, je možné si vyzkoušet různé algoritmy a vidět na grafickém výstupu, co přesně daný algoritmus s použitým stimulem dělá.



Obrázek 23: Nastavení testovacího softwaru

5.3.2 Virtuální polohování stimulu

Program umí simulovat azimut i elevaci v rozmezí od -45° do 45° , azimut s intervaly 5° a elevaci s rozdílem 6.4° z důvodů vysvětlených v kapitole 5.2. Jednotlivá nastavení virtuálního polohování jsou k vidění na obrázku 23. Pro simulaci azimutu jsou na výběr principy ILD, ITD a jejich kombinace, u principů ILD je pak možné zvolit virtuální polohování v horizontální rovině za pomoci panoramy nebo hlavového stínu. Elevaci simuluje pomocí zvětšování či zmenšování základní frekvence metodou pitch shifting za použití poměrů diatonické durové stupnice mezi tóny (na výběr je ladění temperované i přirozené), nebo lze zvolit simulaci pomocí strukturálního modelu boltece.

5.3.3 Úprava stimulu

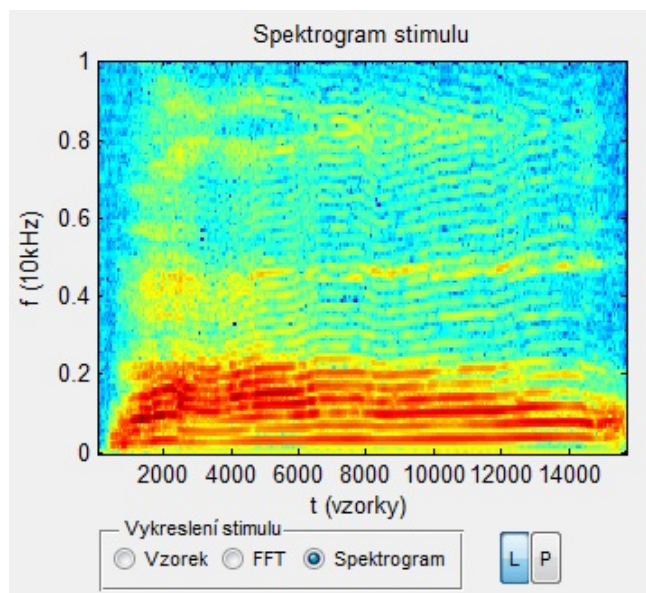
Stimul lze filtrovat pásmovou propustí, použitelnou i jako horní či dolní propust, a nastavit střední frekvenci (první pole pod kolonkou filtrace), šířku pásma (druhé pole) a řád FIR filtru (třetí pole), uživatel si tak může vyzkoušet, jak zní jednotlivé části spektra stimulu. Pomocí slideru pod nastavením filtrace je možné vytvořit hřebenový filtr, respektive přidat zpožděný signál stimulu o 0-5 ms pro simulaci odrazu od podlahy pro lepší vjem, aby neznělo uživateli uvnitř hlavy, což je častý problém virtuálního polohování pomocí duplexních principů. Tímto se nepodaří tento problém zcela úplně odstranit, vjem je pro uživatele však věrnější. Pro výběr stimulu slouží tlačítko *Výběr stimulu*, pod kterým se zobrazí název vybraného stimulu a informace o jeho vzorkovací a základní frekvenci.

5.3.4 Ovládání grafického vstupu

Pro zadání vstupních hodnot elevace a azimutu slouží mřížkovaná mapa v levém horním okně programu. Zkušební 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. Druhým způsobem ovládání je kliknutím 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*.

5.3.5 Grafický výstup

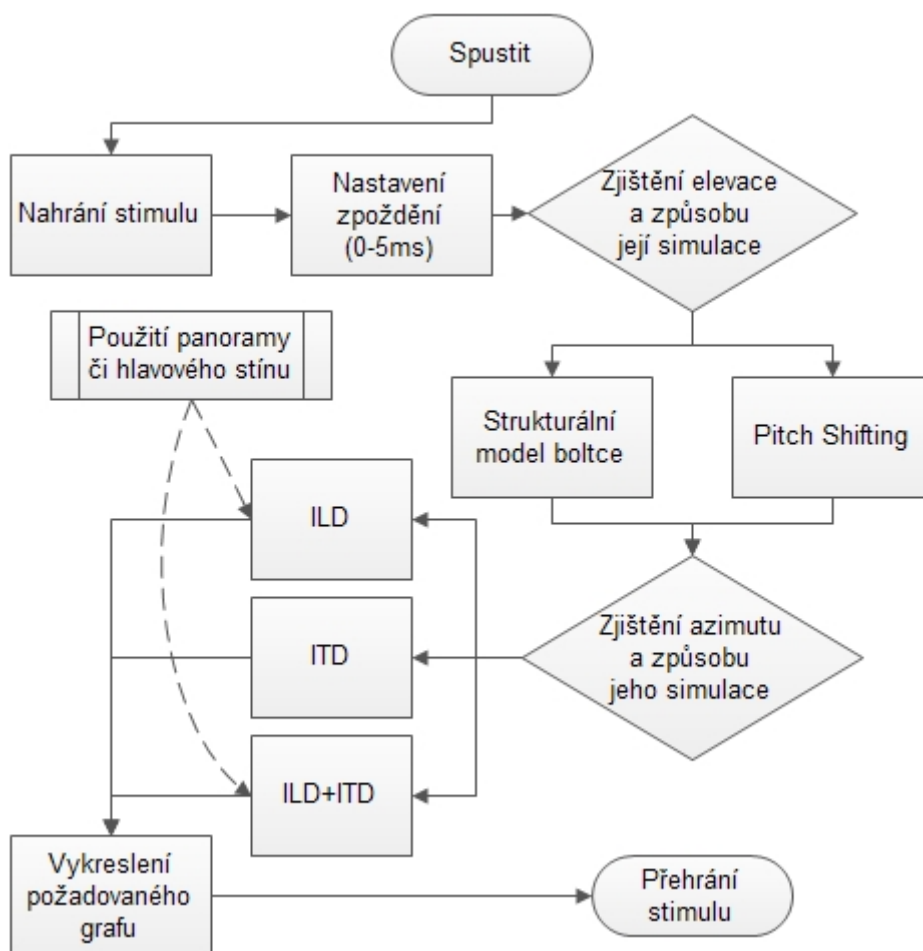
V levém horním okně se nachází grafická informace o aktuálně používaném stimulu. Na výběr je vykreslení časového průběhu, FFT (Fast Fourier Transform - rychlá Fourierova transformace) v logaritmickém měřítku a vykreslení spektrogramu. Vedle výběru těchto hodnot si u časového průběhu a FFT lze navolit i sledování levého či pravého kanálu tlačítky *L* a *R*. V případě časového průběhu a FFT se kanály vždy vykreslí oba, zvolený kanál má však výraznější barvu a překrývá druhý nezvolený kanál.



Obrázek 24: Grafický výstup

5.3.6 Vývojový diagram zkušební rozhraní

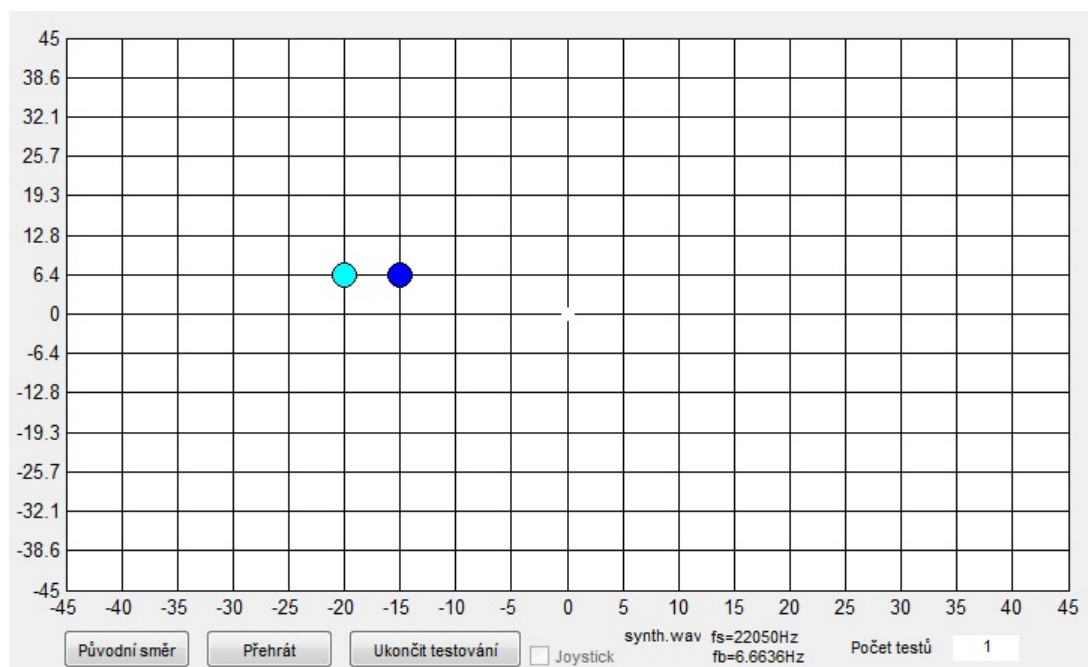
Na obrázku 25 je vidět vývojový diagram programu. Jako první se nastaví zpoždění signálu, tedy simulace odrazu od podlahy, které se i v reálném prostředí odehraje na prvním místě. Zároveň se v tomto bodě provede filtrace, pokud je nějaká nastavena. Program poté z grafického vstupu zjistí, jak a jaký azimut a elevaci chceme simulovat. Posléze se stimul naplňuje do požadované elevace a azimutu. Použité virtuální metody polohování v horizontální a vertikální rovině jsou na sobě nezávislé, není tedy nutné postupovat přesně v tomto pořadí.



Obrázek 25: Vývojový diagram zkušební rozhraní

5.4 Testovací rozhraní

Při stisku tlačítka *Spustit test* v hlavním programu zmizí grafický výstup s vykreslením informace o vzorku a veškeré rozhraní pro nastavení simulace virtuální polohy stimulu a zvětší se okno s mřížkou s mapováním elevace a azimutu. Testovací rozhraní lze ovládat buďto myší či joystickem, k nastavení ovládání slouží checkbox vedle tlačítka *Spustit*, který se po spuštění rozhraní zablokuje, uživatel se tedy pro ovládací zařízení musí rozhodnout ještě před a nikoliv během testování, důvodem je následná separace výsledků stanovených při ovládání myší a joysticku do samostatných textových souborů. Druhý checkbox *Učební mód* slouží k zobrazení polohy náhodně polohovaného zvukového stimulu po vlastním odhadu uživatele, čímž ulehčuje navyknutí si na mapování audio displeje.



Obrázek 26: Testovací prostředí programu - ovládání myší

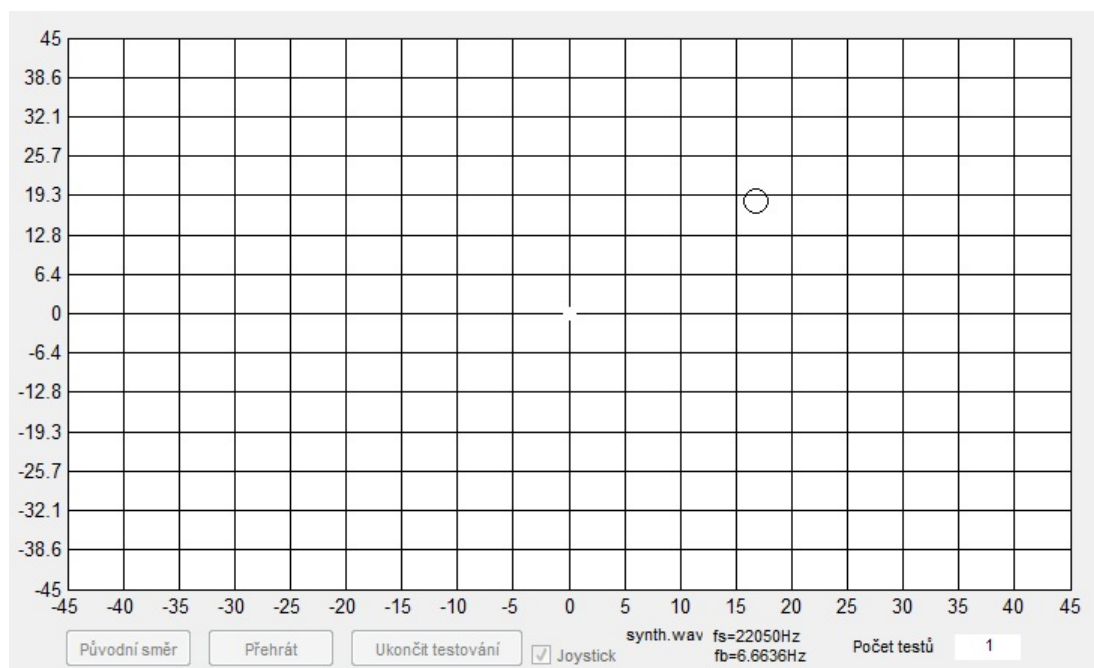
5.4.1 Ovládání myší

Při ovládání myší vyčká program po kliknutí na grafický vstup mřížky okamžik v řádu stovek milisekund, a poté pustí do sluchátek daný zvukový stimul s nastavením, které jsme zvolili. Směr vybere rovnoměrně náhodným zvolením elevace a azimutu, aniž by na mapě ukázal odkud. Na testované osobě je, aby odhadla směr, ze kterého virtuálně polohovaný zvuk přišel. Stisknutím tlačítka *Spustit* se mu zvuk z daného náhodného směru přehraje znovu, tlačítkem *Původní směr* si může připomenout, jak zní stimul ze směru před ním.

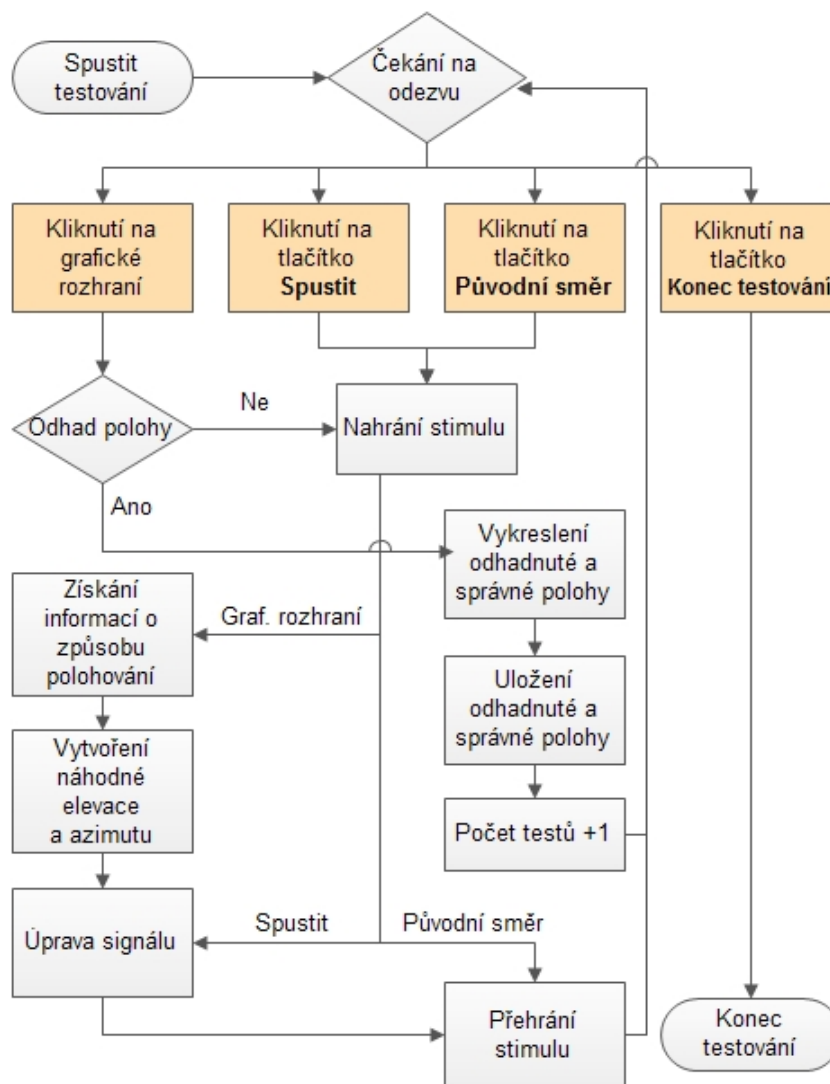
Poté co znovu klikne na okno s mřížkou a zvolí tak polohu, kterou považuje za správnou a která se mu zaokrouhlí na polohu nejbližšího členění os grafu a vykreslí na testovacím rozhraní, se mu s malým časovým odstupem vykreslí i směr skutečný, pokud má nastavený učební mód.

5.4.2 Ovládání joystickem

Ovládání grafického testovacího rozhraní pomocí joysticku je velmi podobné ovládání pomocí myši. Vzhledem k tomu, že nelze využít kurzoru myši, při prvním kliknutí hlavního tlačítka joysticku se po přehrání zvuku náhodné polohy objeví na obrazovce prázdné kolečko - kurzor joysticku, který označuje jeho aktuální polohu. Ta se mění v reálném čase spolu s nakloněním ovládací páky, tato poloha se do opětovného zmáčknutí hlavního tlačítka nijak nezaokrouhluje. Při testování jsou zablokovány veškerá tlačítka stisknutelná myší, tlačítka na joysticku ale plní stejnou úlohu. Hlavní tlačítko slouží jako alternativa levého tlačítka myši při kliknutí na grafický vstup, dvě tlačítka pro opětovné přehrání napolohovaného a původního stimulu a poslední tlačítko k ukončení testování. Postup je posléze stejný jako u ovládání rozhraní pomocí myši, hlavní rozdíl a výhoda joysticku je ta, že tímto ovládacím prvkem lze ovládat testovací rozhraní bez náhledu na grafický vstup, je tedy vhodné i pro testy bez možnosti přímé zpětné vazby či zrakově postižené osoby.



Obrázek 27: Testovací prostředí programu - ovládání joystickem



Obrázek 28: Vývojový diagram testovacího rozhraní

S každým testem se zvětší indikátor počtu testů. Po dosažení požadovaného počtu testů testování ukončíme tlačítkem *Ukončit testování* nebo tlačítkem joysticku k tomu předurčenému. Program uloží veškerá data do souboru ve formátu *YYYYMMDDHHMM.txt* do složky *vysledky* či *vysledky_jst*, podle použitého ovládacího prvku. Po uložení výsledků se vrátí program zpět do simulačního módu a zjeví se okno s procentuální úspěšností uživatele. Tato úspěšnost je průměrem vertikální a horizontální roviny a byla dodána pouze pro povzbuzení uživatele k dalšímu měření, nemá tedy žádnou důležitou hodnotu. Testovací rozhraní tímto způsobem připomíná formu hry a testující osoby se tak ani při desítkách pokusů neunaví, ovládání v podobě klikání kurzoru myši na grafické rozhraní je také velmi intuitivní.

5.5 Použité algoritmy

Věškeré použité algoritmy jsou okomentovány na CD příloze, za zmínku stojí například funkce nastavující grafický vstup, tedy spuštění žádané funkce při kliknutí myši na objekt. Tato funkce musí být vložena do spouštěcí funkce programu a pokaždé, pokud se na graf klikne pro její opětovné nastavení.

```
set(handles.axesX, 'ButtonDownFcn', @axesX_ButtonDownFcn, handles);
```

Funkce, která vrací polohu kurzoru myši v objektu vrátí vektor s x-ovou a y-ovou polohou.

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

Vzhledem k tomu, že se v GUI nachází více grafů, je nutné označit, se kterým chceme pracovat. Matlab si jej zapamatuje a všechny příkazy týkající se grafů provede na posledním takto určeném grafu.

```
axes(handles.axesX);
```

Při vykreslení bodu do prázdného grafického vstupu je důležité vždy zachovat velikost os, Matlab by si je jinak automaticky upravil podle sebe. Vykreslení černě ohraničeného modrého bodu se souřadnicemi X a Y o velikosti 12-ti pixelů do grafu se zachováním původních os provedeme pomocí příkazu

```
set(handles.axesX, 'grid', 'on', 'GridLineStyle','-', 'XMinorGrid','off'); hold on;  
Y,X,'o', 'MarkerEdgeColor','k', 'MarkerFaceColor','c', 'MarkerSize',12);.
```

Pro ovládání grafického vstupu pomocí joysticku, byly přebrány funkce `mat_joy.mex` a `mat_joy.c`, které ovládání umožňují, v Matlabu se pak informace o poloze joysticku a stavu jednotlivých tlačítek dostane pomocí funkce

```
[pos, but] = mat_joy(0);,
```

kde vektor `pos` obsahuje x-ovou a y-ovou polohu joysticku a vektor `but` stavy jednotlivých tlačítek. Ty je nutné si předem odzkoušet, každý joystick má své individuální nastavení, výchozí program je přednastavený pro ovládání prvních čtyř tlačítek joystickem Genius MaxFighter F-17.

Pro ukládání výsledků slouží příkaz

```
dlmwrite(['vysledky\pitch\' num2str(datum) '.txt'], vysledky);,
```

který v tomto konkrétním případě uloží matici `vysledky` do textového souboru `datum.txt`. Tento soubor má pak tvar textu s informacemi mezi sloupci matice oddělenými čárkou a mezi řádky pomocí zalomení do nového řádku.

5.6 Zpracování výsledků

Výsledky se během testování ukládají do matice `vysledky`, která má v prvním sloupci informaci o čísle měření aktuálního testování, ve druhém a ve třetím sloupci informaci o x-ové a y-ové poloze náhodně napolohovaného zvukového stimulu a v posledních dvou sloupcích se pak nachází odhad uživatele, tedy x-ová a y-ová poloha odhadnutého místa. Matice se posléze uloží do textového souboru do složky podle typu testování, jednotlivá měření od sebe dělí čas měření, neboli název textového souboru. Tyto složky se musí před následujícím testováním vždy uložit do jiné složky, například složky se jménem testovaného subjektu, aby v naměřených datech nenastal chaos. Výsledky všech měřených subjektů poté sloučíme do 5 textových souborů, rozdělené podle typu měření. Dále je nutné vytvořit algoritmus, pro vyčítání výsledků a jejich vykreslení do grafu. Pro tento úkol slouží program `average.m` zabalený spolu s ostatními výsledky v souboru `mereni.zip` na přiloženém CD. Průměrování zajistíme funkcí `mean`, která vrací aritmetický průměr

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (21)$$

Vytvoříme šestý sloupec s absolutním rozdílem mezi skutečným a odhadnutým azimutem a sedmý sloupec s absolutním rozdílem mezi skutečnou a odhadnutou elevací a nalezneme celkový průměr odchylky elevace (či azimutu) pomocí

```
elev_avg=mean(vysledky(:,7));
```

Dále nalezneme vektor s polohou jednotlivých elevací (azimutů)

```
elevX=find(vysledky(:,3)==X);
```

kde `X` v tomto případě značí požadovanou elevaci. Tu posléze zprůměrujeme pomocí příkazu

```
elevX_avg=mean(vysledky(elevX,7));
```

Tímto dostaneme průměrné odchylky pro jednotlivé elevace a azimuty. Výsledky jsou zpracovány do grafů v závislosti na typu měření a v závislosti na elevaci a azimutu.

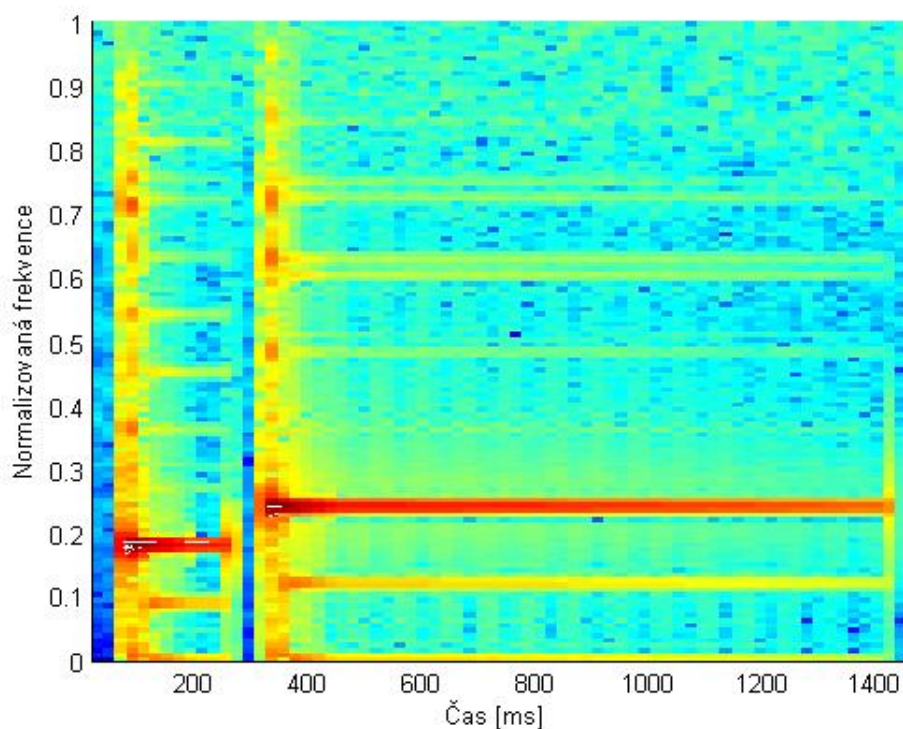
6 Testování

6.1 Testovací podmínky

Testování proběhlo za použití standardní výpočetní techniky, sluchátek KOSS MV1, joysticku Genius MaxFighter F-17 a externí zvukové karty M-Audio Fast Track. Proběhlo v domácích podmínkách v klidném a ničím nerušeném prostředí.

K testování byl využit vytvořený program v prostředí Matlab popsáný v páté kapitole, jako stimul byl využit syntetický signál *synth.wav*, který připomíná zvuk sonaru a obsahuje dostatečnou vysokofrekvenční složku. Jeho spektrogram je vidět na obrázku 29, stimul se nachází na CD příloze.

Každé z měření trvalo celkem i s vysvětlením funkcí a ovládání programu zhruba 45 minut. Nastavení programu, konkrétně polohování zvukového stimulu v horizontální rovině, bylo provedeno za použití obou dvou algoritmů simulující principy ILD a ITD popsáných ve třetí kapitole, ILD bylo simulováno pouze za použití hlavového stínu. Simulaci odrazu od podlahy si mohl každý uživatel nastavit podle sebe. Ladění stupnice, podle které se polohovalo v elevačním frekvenčním kódování bylo ponecháno po celou dobu přirozené.

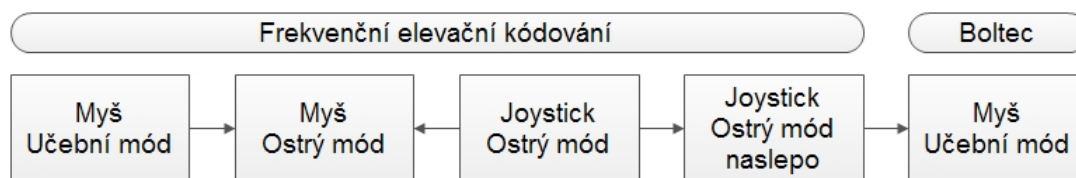


Obrázek 29: Použitý syntetický stimul

6.2 Metodika testování

Bylo testováno celkem 6 jedinců ve věkovém rozmezí kolem 25ti let, někteří jedinci byli hudebně založení a někteří nikoliv. Testovaným subjektům byl vysvětlen smysl testování, princip virtuálního polohování v horizontální i vertikální rovině a před testováním jim také bylo umožněno vyzkoušet si simulační prostředí programu. Při předchozím zkoušení betaverze programu na neznalých uživatelích se prokázalo, že uživatelé s postupem času sami lokalizovali vyšší frekvence do vyšších poloh, čímž se potvrdilo i prvotní měření C.C.Prattové [1] a opětovným postupem by se pouze ztrácel čas.

Smyslem měření je primárně zjistit přibližnou přesnost navrženého frekvenčního elevačního kódování, o jehož funkčnosti se na rozdíl od virtuálního polohování pomocí strukturálního modelu boltce či použití HRIR neví.



Obrázek 30: Průběh testování

Jako první proběhlo testování frekvenčního elevačního kódování s ovládáním pomocí myši. Tyto první pokusy proběhly v učebním módu, kdy se uživatelům zobrazovala po jejich odhadu polohy virtuální polohy zvukového stimulu poloha skutečná. To napomáhalo k navyknutí si na rozhraní a zlepšování se v odhadování skutečné polohy.

Poté, co si navykli na rozhraní, což bylo zhruba po třiceti měřeních, se učební mód odstranil spolu s napomáhajícím kolečkem označujícím skutečnou polohou a uživatelé se tak museli spolehnout na svůj vlastní odhad. Neměli již tedy nic ke srovnání kromě původního zvukového stimulu vycházejícího z nulové polohy.

Dále pokračovalo měření s joystickem v ostrém módu s náhledem na obrazovku bez pomocného kolečka se skutečnou polohou. Toto měření je provedeno z toho důvodu, aby si uživatel navykl na ovládání joysticku. Podmínky byly však kromě ovládacího rozhraní identické s předchozím měřením, výsledky by tedy měly být velmi podobné.

Následovalo identické měření, avšak bez náhledu na obrazovku, kdy uživatelé odhadovali polohu pouze s joystickem. Toto měření mělo ověřit nezávislost uživatele na vizuálním kontaktu s navigační mřížkou.

Následovala poslední série měření, a to strukturálního modelu boltce v učebním módu pouze pomocí myši, které mělo za úkol hrubě odhadnout přesnost tohoto typu virtuálního polohování ve vertikální rovině. Toto měření slouží pouze ke srovnání s měřením prvním, které mělo stejné podmínky, avšak bylo teoreticky znevýhodněno svou počáteční polohou v plánu testování, jelikož uživatelé ještě nebyli úplně navyklí na testovací rozhraní. Změřené hodnoty budou tedy srovnány a okomentovány v závěru. Testování strukturálního modelu boltce pomocí joysticku se již nedělo z časových a praktických důvodů, uživatelé byli již po tolika měřeních unaveni.



Obrázek 31: Měření audio displeje

V každém módu měření bylo cílem naměřit u všech testovaných kolem 30ti odhadů, tedy kolem $5 \times 30 = 150$ měření na jednu osobu, celkem bylo provedeno přes $150 \times 6 = 900$ měření. Na jednu polohu elevace v každé části měření by tedy celkově mělo vyjít průměrně 10 měření, vzhledem k 15ti možným elevačním polohám.

Uživatelé si mohli kdykoliv přehrát originální zvuk i zvuk náhodně napolohovaný, testování nebylo nijak časově omezeno.

Každý uživatel měl při každém měření unikátní náhodnou sérii poloh zvukového stimulu.

6.3 Výsledky měření

Měření probíhalo téměř bez komplikací a uživatelé neměli problémy s navyknutím si na ovládací rozhraní. Jediná komplikace nastala při měření s joystickem, který z mechanických důvodů nedosahoval krajních poloh a proto byl algoritmus výpočtu náhodné polohy ošetřen tak, aby tyto polohy negeneroval, ve výsledných grafech tedy bohužel nejsou zahrnuty, stejně jako hodnota -40° v azimutu, která se během měření při použití joysticku (třetí metoda měření) ani jednou nevyskytla, za což může algoritmus náhodně generovaného azimutu. I to však nebránilo uživatelům krajní polohy během měření zvolit, zvláště patrné jsou u měření s joystickem naslepo, viz. grafy na obrázcích 44 a 46.

Bylo vyzorováno, že uživatelé s hudebním zájemem do principu audio displeje pronikli rychleji. Pro srovnání jejich výsledků by však bylo třeba udělat více měření a zaměřit se přímo na tento typ měření.

Z celkových vypočtených průměrných hodnot je vidět, že nejlépe ze všech dopadlo ve virtuálním polohování elevace testování frekvenčního elevačního kódování pomocí joysticku. Naopak nejhůře dopadlo virtuální polohování pomocí strukturálního modelu boltce, ve kterém měli často uživatelé i problém určit, zdali se zvukový zdroj nachází v horní či dolní polorovině. Z tabulky je vidět, že průměrná odchylka azimutu byla ve všech případech podobná, kolem 9.5° . Výjimkou byl joystick použitý naslepo, uživatelé si patrně dostatečně nezmapovali jeho krajní polohy. Výsledek byl jinak vzhledem k použití stejných polohovacích technik v horizontální rovině předpokládán.

Testovací mód	Prům. chyba elevace	Prům. min. chyba elev.	Prům. max. chyba elev.
Myš - učební mód	8.55°	1.5°	23.2°
Myš	9.42°	0.5°	23.9°
Joystick	7.44°	1°	19.9°
Joystick slepý mód	10.64°	1°	26.7°
Strukt. model boltce	19.82°	3°	43°

Tabulka 3: Odchylky odhadnutých hodnot ve vertikální rovině

Data z tabulek 3 a 4 jsou převzatá z výsledků programu `average.m` na přiloženém CD. Tento program vrací kromě údajů v tabulce i odchylky pro horní a dolní polorovinu, přičemž rozdíl v odhadu je až 3° pro měření s myší v ostrém módu, což může svědčit fakt o měnícím se JND s rostoucí frekvencí tónu. Jediná výjimka nastala u měření joystickem, kdy má dolní polorovina o 0.5° lepší odhad. Odhad elevace u strukturálního modelu boltce vyšel pro obě dvě poloroviny přibližně stejně.

V příloze práce jsou vidět dva typy grafů. První typ obsahuje naměřené (velká prázdná kolečka, čím tmavší, tím větší hustota), skutečné (světle zelená kolečka, tyto hodnoty tvoří diagonálu) a průměrné (modrá kolečka) elevace či azimuty. Jednotlivé elevační polohy jsou namapovány horizontálně vedle sebe, aby mohly být naměřené polohy pro nich vynesené vertikálně a graf byl přehledný, u azimutů pak platí opačné pravidlo, tj. namapování vertikálně a vnesení horizontálně. Čím blíže jsou modrá kolečka diagonále, tím úspěšnější měření.

Testovací mód	Prům. chyba azimutu	Prům. min. chyba azim.	Prům. max. chyba azim.
Myš - učební mód	8.87°	1.2°	23.8°
Myš	10.13°	1.8°	23.2°
Joystick	9.95°	1.6°	28.1°
Joystick slepý mód	16.96°	3.8°	50.9°
Strukt. model boltce	9.12°	2.4°	17.9°

Tabulka 4: Odchyly odhadnutých hodnot v horizontální rovině

Druhý typ grafu obsahuje přehled naměřených odchylek, neboli rozdílu mezi skutečnou náhodnou polohou a polohou odhadnutou uživatelem pro jednotlivé metody měření. Elevace a azimuty jsou namapovány horizontálně, vertikálně jsou pak vyneseny odchylky. V grafech jsou vyznačeny minima, tedy například správný odhad v případě nulové odchylky (0 na y-ové ose), maxima, tedy odhad, který byl nejvíce mimo skutečnou polohu a uprostřed těchto rozsahů aritmetický průměr pro jednotlivé elevace a azimuty. První typ grafu se vykreslí pospuštění programu `data.m` a druhý po spuštění programu `average.m`, oba dva zabalené ve složce `mereni.zip` na přiloženém CD.

Ačkoliv bylo ve většině případů zřejmé, zdali jde zvuk zleva, zprava, zezdola či ze shora, i tak se najde měření, ve kterých se uživatel netrefil o celých 85°, viz. graf na obrázku 47. Jedná se o extrém, daný uživatel pravděpodobně zazmatkoval a spletl si strany. Průměrné maximální odchylky elevace a azimutu jsou vidět v tabulkách 3 a 4.

V grafech měření frekvenčního elevačního kódování, ve kterých mohl uživatel přesně určit polohu myši (obrázky 32-39) je skutečně vidět, že horní polorovina má větší úspěšnost odhadu nežli polorovina dolní, což potvrzuje i jeden z algoritmů programu `average.m` na přiloženém CD. Při použití joysticku měli uživatelé zhoršené podmínky, jelikož nebyli na citlivé ovládání joysticku zvyklí. Mnohdy určovali pouze polohu v dané oblasti a nikoliv polohu přesně chtěnou z důvodu třesu páky joysticku. Výsledky jsou proto pravděpodobně méně přesné, nežli u ovládání myši.

Zajímavý jev nastal u frekvenčního elevačního kódování, kdy v horní polorovině uživatelé lokalizovali stimul nejdříve výše, nežli jeho simulovaná poloha a s vyššími simulovanými polohami začínali ubírat až lokalizovali stimul níže, tento jev nastává kolem elevace 25.7°. V dolní polovině měli uživatelé tendenci elevaci s jejím klesáním zvyšovat.

Obecně můžeme z grafů vyčíst, že střední polohy jak azimutu tak elevace mají větší úspěšnost lokalizace nežli polohy okrajové. Je to pravděpodobně dáno subjektivním prahem kolem středu, za kterým uživatel hýbe hruběji s lokalizací stimulu.

Ačkoliv nebyl odhad azimutu hlavním účelem měření, není od věci se podívat i na výsledky měření horizontální roviny. Velmi zajímavé je velká odchylka v horizontálních krajních polohách audio displeje, ve kterém se uživatelé mnohdy netrefili ani jednou do správné polohy, jako je viděn například na obrázku 39 či 43. Je zapříčiněná patrně špatným soustředěním na prostorovou informaci audio displeje v horizontální rovině.

Z grafů i numerických výsledků měření je zřejmé, že implementace audio displeje s frekvenčním elevačním kódováním bylo úspěšné a že se podařilo touto sonifikační metodou zlepšit elevační vjem uživatele.

7 Závěr

Práce pojednává o polohování virtuálních zdrojů zvuku pro účely navigace s využitím metod sonifikace. V úvodní části byl podán stručný přehled lokalizačních metod a algoritmů pro virtuální polohování zdroje zvuku pomocí stereofonních sluchátek. Dále byla vysvětlena základní problematika sonifikace a s ní spojená problematika audio displejů.

Cílem práce bylo navržení metody pro zpřesnění lokalizace virtuálních zdrojů zvuku pomocí sonifikačních metod. Za tímto účelem byl vytvořen jednoduchý audio displej využívající frekvenčního elevačního kódování, který byl otestován ve vytvořeném programu v prostředí Matlab s uspokojivými výsledky. Jako frekvenční poměry pro jednotlivé kroky v elevaci byly využity poměry diatonické durové stupnice, která je známá většině uživatelů a která nepřeladuje základní frekvenci použitého stimulu do zbytečně vysokých hodnot.

Z měření byly prokázány úspěšné polohovací vlastnosti vytvořeného audio displeje ve vertikální rovině, které předčily i výsledky strukturálního modelu boltce. V nejlepším případě byla změřena průměrná odchylka určení elevace 7.44° , v nejhorším pak 10.64° , což však mohlo být zapříčiněno chybou metody měření. V souladu se zadáním byly srovnány výsledky metody frekvenčního elevačního kódování s výsledky metody virtuálního elevačního polohování pomocí strukturálního modelu boltce. Vypočtená průměrná odchylka této metody byla 19.82° , což je výrazně více, nežli nově implementovaná metoda pomocí sonifikace.

Sonifikační technika použitá pro vytvořený audio displej je funkční a úspěšně vylepšuje elevační lokalizaci, respektive její přesnost. Kombinace sonifikačních metod s již známými polohovacími metodami tedy dopadla úspěšně.

Audio displej by dále mohl být vylepšen pro použití v celém prostoru kolem uživatele pro všesměrovou navigaci s již pevně danými stimuly pro různé směry.

Práce je součástí grantu č. SGS11/159/OHK3/3T/13.

8 Literatura

- [1] Pratt, Carroll C.: *The spatial character of high and low tones*
Harvard University, 1929
- [2] Begault, D.R.: *3D Sound for Virtual Reality and Multimedia*
Academic Press, 1994
ISBN 0-120-84735-3
- [3] Sima S.: *HRTF Measurements and Filter Design for a Headphone-Based 3D-Audio System* [online],
dostupné z WWW: http://edoc.sub.uni-hamburg.de/haw/volltexte/2009/697/pdf/BA_Sylvia_Sima.pdf
- [4] Blauert, J.: *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*
The MIT Press, 1996
ISBN 0-262-02413-6
- [5] Duda, R., Martens W.L.: *Range-Dependence of the HRTF for a Spherical Head Model*
J. Acoust. Soc. Am. vol. 104, 1998, pp. 3048-3058
- [6] Psychology 3203: *Sound Localization* [online],
dostupné z WWW: <http://www1.appstate.edu/kms/classes/psy3203/>
- [7] 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
- [8] Kollmeier, B., Brand, T., Meyer, B.:
Perception of Speech and Sound
Springer, 2008, ISBN 978-3-540-49125-5
- [9] Trutt, J. W.:
On Our Perception of the Direction of a Source of Sound
Proc. Musical Assoc., 1875, vol. 2, pp. 75–84
- [10] Sodnik, J., Susnik, R., Stular, M., Tomazic, S.:
Spatial sound resolution of an interpolated HRIR library
Appl Acoust vol.66, 2005, 1219–34
- [11] Hermann, T.: *Taxonomy and Definitions for Sonification and Auditory Display*
Proceedings of the 14th International Conference on Auditory Display,
Paris, France June 24 - 27, 2008

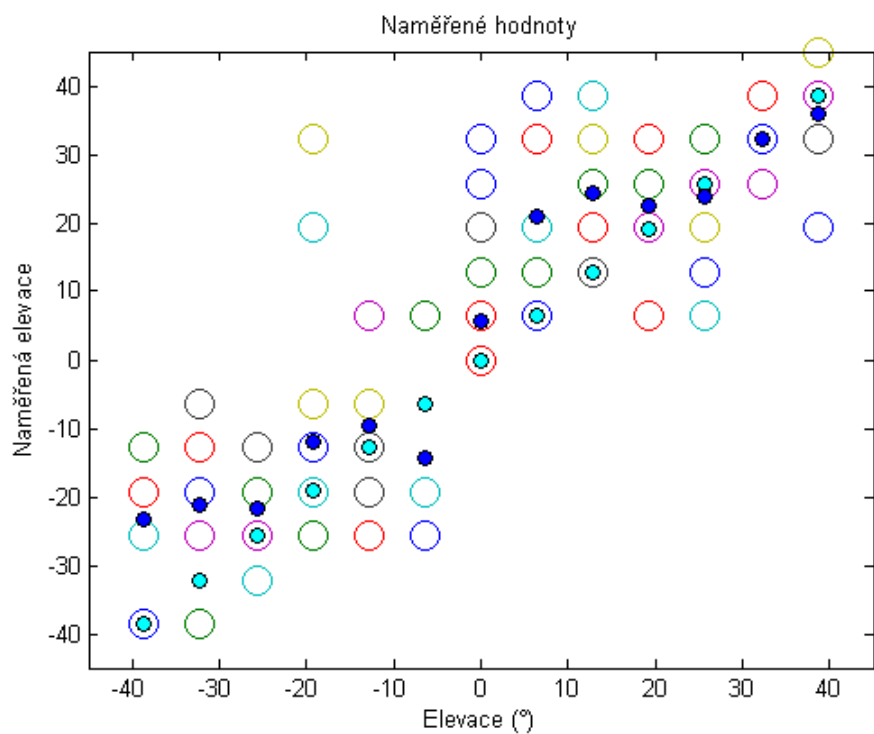
- [12] Hermann, T., Ritter, H.: *Listen to your Data: Model-Based Sonification for Data Analysis*
Department of Computer Science University of Bielefeld, 1999
- [13] Hermann, T.: *The Sonification Handbook*
1 st. ed., 2011, Berlin: Logos Publishing House
ISBN 978-3-8325-2819-5
- [14] Zölzer, U.: *DAFX Digital Audio Effects*
John Wiley & Sons, Ltd. New York, 2011
ISBN 948-0-470-66599-2
- [15] Knoll, G.F.: *Radiation Detection and Measurement*
John Wiley and sons, 2000
ISBN 0-471-07338-5
- [16] Crispian, K, Ehrenberg, T.: *Evaluation of the cocktail-party effect for multiple speech stimuli within a spatial auditory display*
J. Acoust. Soc. Am., 11, 932-941
- [17] SWAN [online],
dostupné z WWW: <http://sonify.psych.gatech.edu/research/swan/>
- [18] Rapant, P.: *Družicové polohové systémy*
VŠB-TU Ostrava, 2002
ISBN 80-248-0124-8
- [19] Walker, B.N., Lindsay, J.:
Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice
Georgia Institute of Technology, 2006
- [20] *vOICe* [online], dostupné z WWW: <http://www.seeingwithsound.com/>
- [21] Loomis J.M., Golledge R.G., Klatzky R.L.:
Navigation System for the Blind: Auditory Display Modes and Guidance
dostupné z WWW: <http://monet.cs.columbia.edu/courses/mobwear/resources/>
- [22] Sikström E.: *Persistent audio cues in auditory display interfaces* [online]
dostupné z WWW: <http://epubl.ltu.se/1402-1552/2009/089/>
- [23] Stahl C.: *The roaring navigator*
MobileHCI '07, p.383-386, ISBN: 978-1-59593-862-6
- [24] *Tactical Sound Garden* [online],
dostupné z WWW: <http://www.tacticalsoundgarden.net/>
- [25] Music 102: *Music, Information, Neuroscience* [online]
dostupné z WWW: <http://eamusic.dartmouth.edu/mcasey/m102/>

- [26] Rabiner, L.R., Gold, B.:
Theory and Application of Digital Signal Processing
 Prentice Hall, 1975, ISBN 0-13-914101-4
- [27] Middlebrooks, J.C., Green, D.M.:
Sound localization by human listeners
 Annual Rev. Psychology vol. 42, 1991, 135–59
- [28] Raykar V.C., Duraiswami R., Yegnanarayana B.:
*Extracting the frequencies of the pinna spectral notches
 in measured head-related impulse responses*
 J. Acoust. Soc. Am., 2005, 118, 364–74
- [29] Balzano, G.J.: *The Group-Theoretic Description of 12-Fold
 and Microtonal Pitch Systems*
 The MIT Press, 1980, Computer Music Journal vol.4, pp. 66-84
- [30] Glaser O.: *Implementace binaurální lokalizace zdroje zvuku*
 ČVUT FEL, 2010, Diplomová práce
- [31] Gotzen, A.D., Bernardini, N., Arfib, D.:
*Traditional Implementations of a Phase-Vocoder:
 The Tricks of the Trade*
 Conference on Digital Audio Effects (DAFX-00), 2000
- [32] Walker B.N., Ehrenstein A.:
Pitch and Pitch Change Interact in Auditory Displays
 Journal of Experimental Psychology, 2000, Vol. 6, No.1,15-30
- [33] Ellis D.: *A Phase Vocoder in Matlab* [online]
 dostupné z WWW: <http://www.ee.columbia.edu/dpwe/resources/>
- [34] *mat_joy.zip* [online],
 dostupné z WWW: <http://alcoholic.eu/category/matlab-2/>

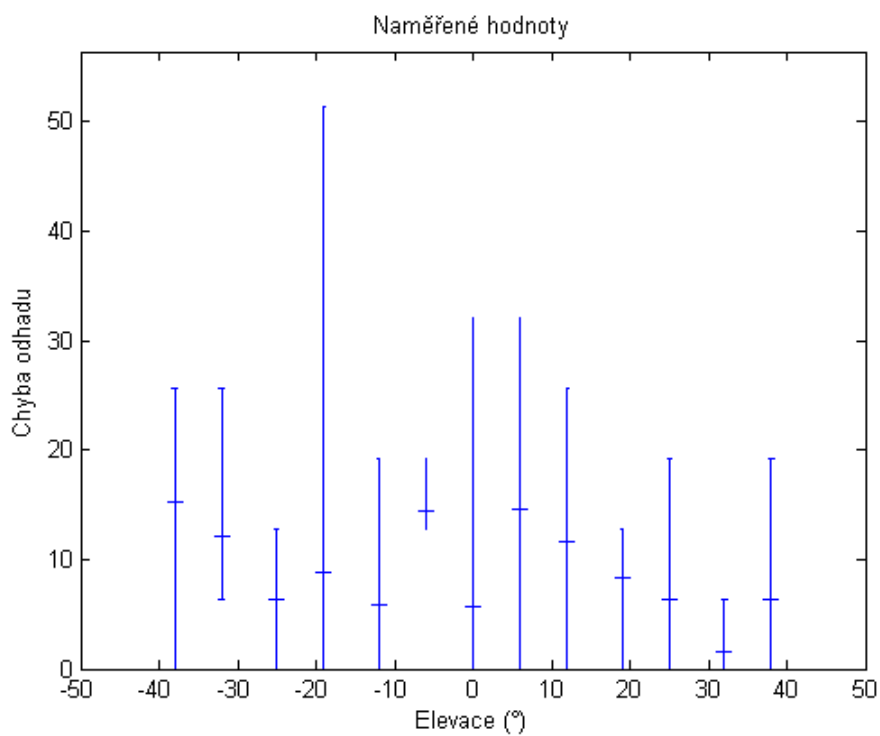
9 Soubory na přiloženém CD

• main.m	Hlavní program
• main.fig	Skript s GUI grafikou
• ild.m	Skript se simulací ILD
• itd.m	Skript se simulací ITD
• boltec.m	Strukturální model boltce [30]
• get_sample.m	Úprava a polohování stimulu
• low_pass.m	Pásmová propust
• shadow.m	Hlavový stín [30]
• base_f.m	Výpočet základní frekvence [4]
• stft.m	Short-Time Fourier Transform [33]
• istft.m	Inverzní STFT [33]
• pvoc.m	Skript na zrychlení signálu [33]
• pvsample.m	Skript interpolující signál [33]
• plot_fft.m	Funkce pro vykreslování grafů
• load_create_plot.m	Modifikace a vykreslení stimulu
• testing.m	Testovací část v GUI
• testing_joy.m	Testovací část v GUI s joystickem
• filter_comb.m	Simulace odrazu stimulu
• visibility.m	Viditelnost prvků v GUI
• mat_joy.mex	Ovládání joysticku [34]
• mat_joy.c	Ovládání joysticku [34]
• uspesnost.m	Výpočet úspěšnosti lokalizace [34]
• meow.wav, synth.wav	Přiložené stimuly
• mereni.zip	Naměřené hodnoty, výpočty a vykreslení
• grafy.zip	Grafy naměřených hodnot

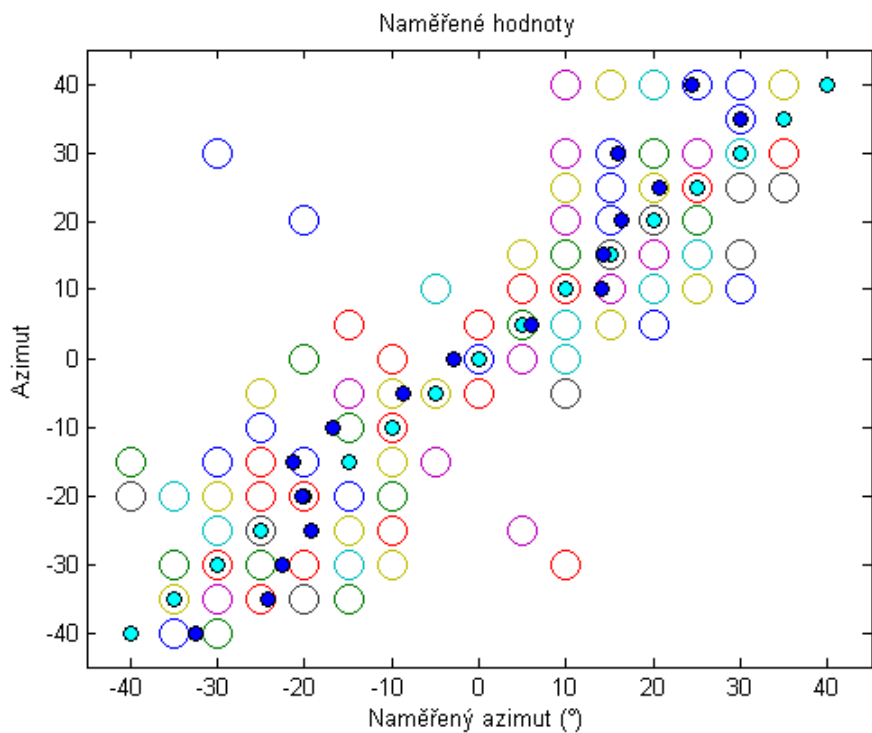
10 Přílohy



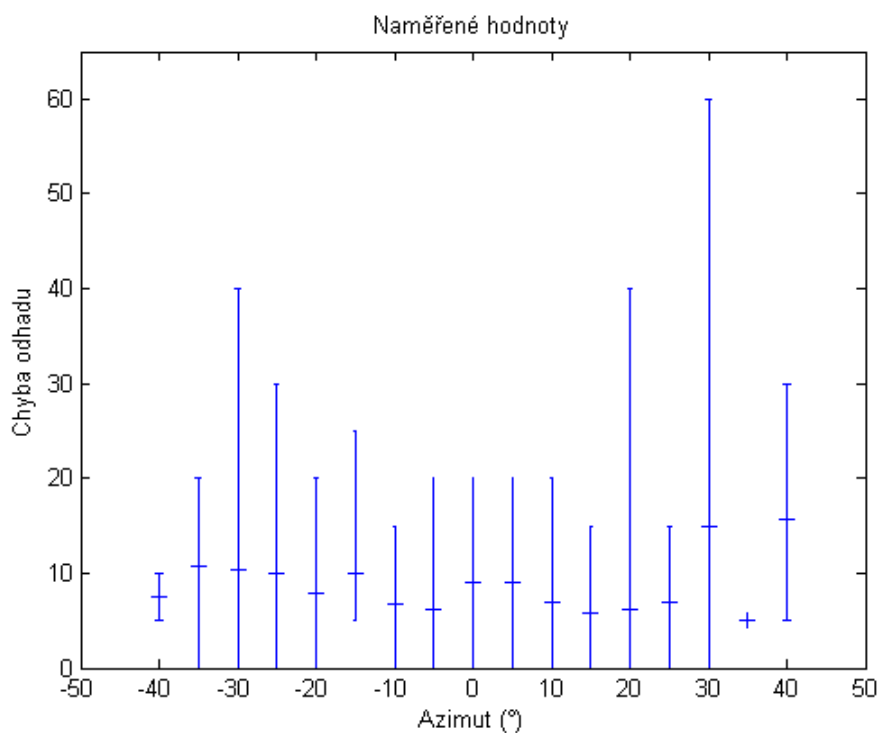
Obrázek 32: Výsledky měření elevace - Použití myši v učebním módu



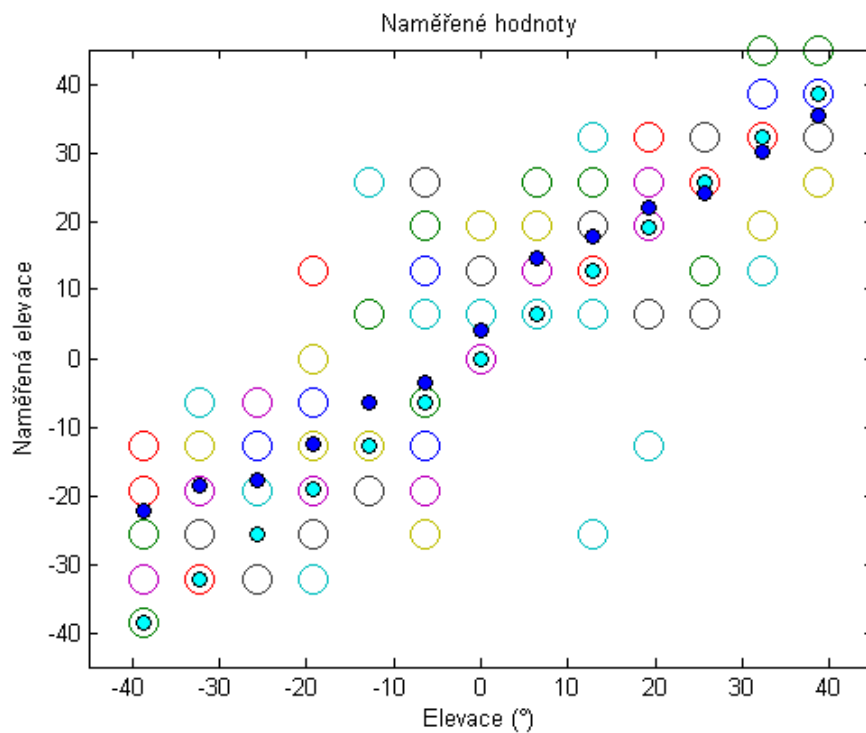
Obrázek 33: Odchylka měření elevace - Použití myši v učebním módu



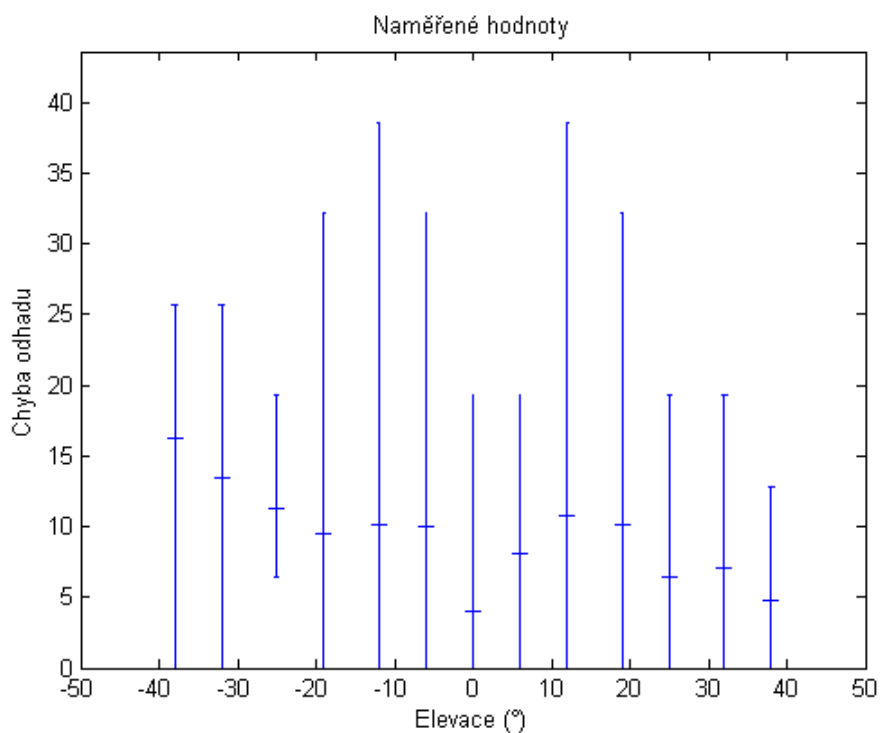
Obrázek 34: Výsledky měření azimutu - Použití myši v učebním módu



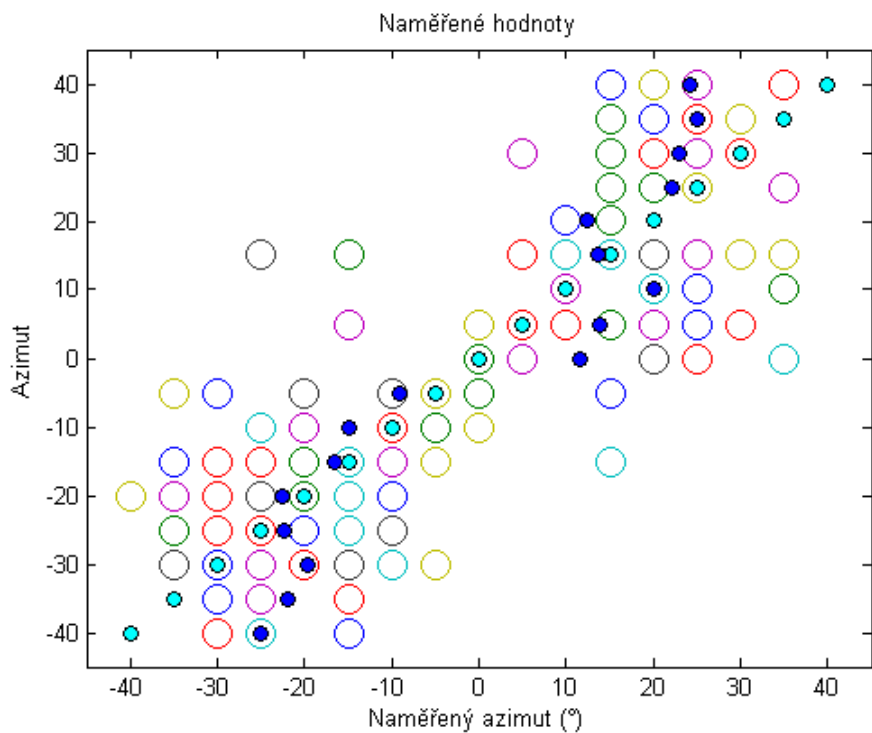
Obrázek 35: Odchylka měření azimutu - Použití myši v učebním módu



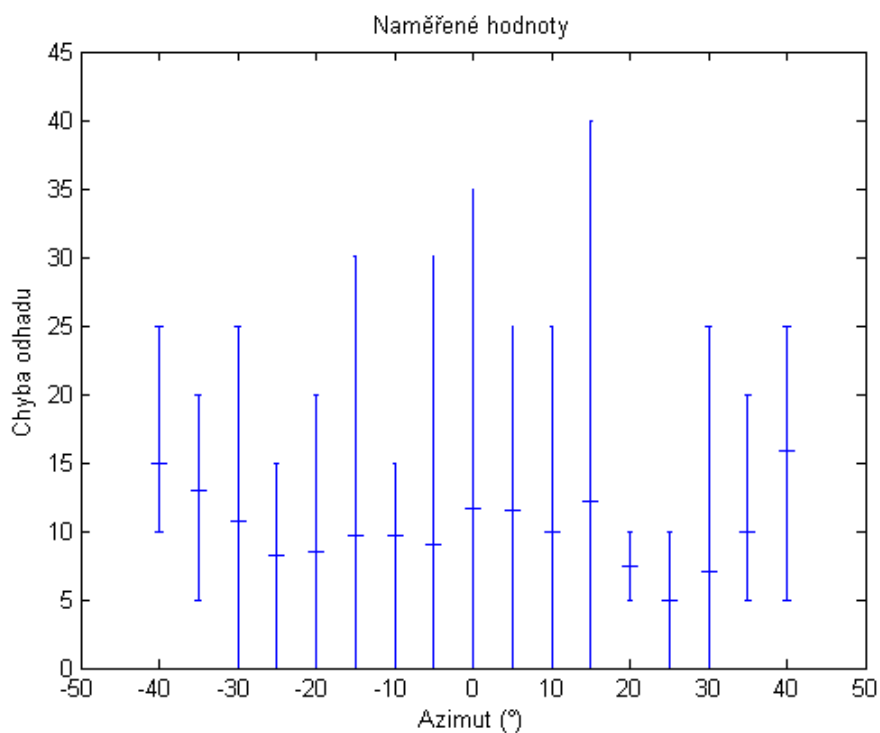
Obrázek 36: Výsledky měření elevace - Použití myši



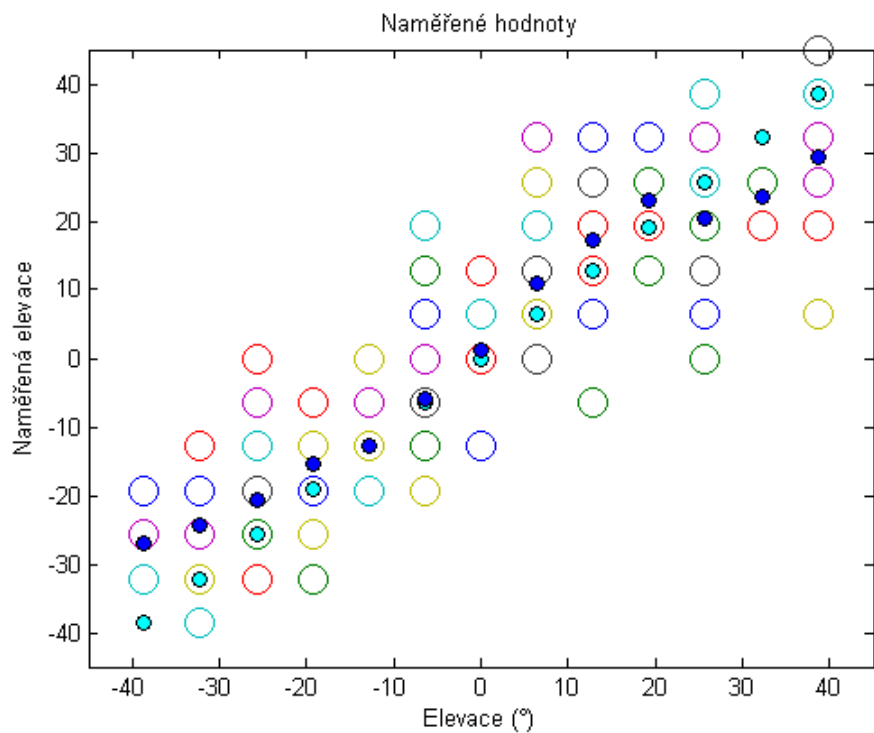
Obrázek 37: Odchylka měření elevace - Použití myši



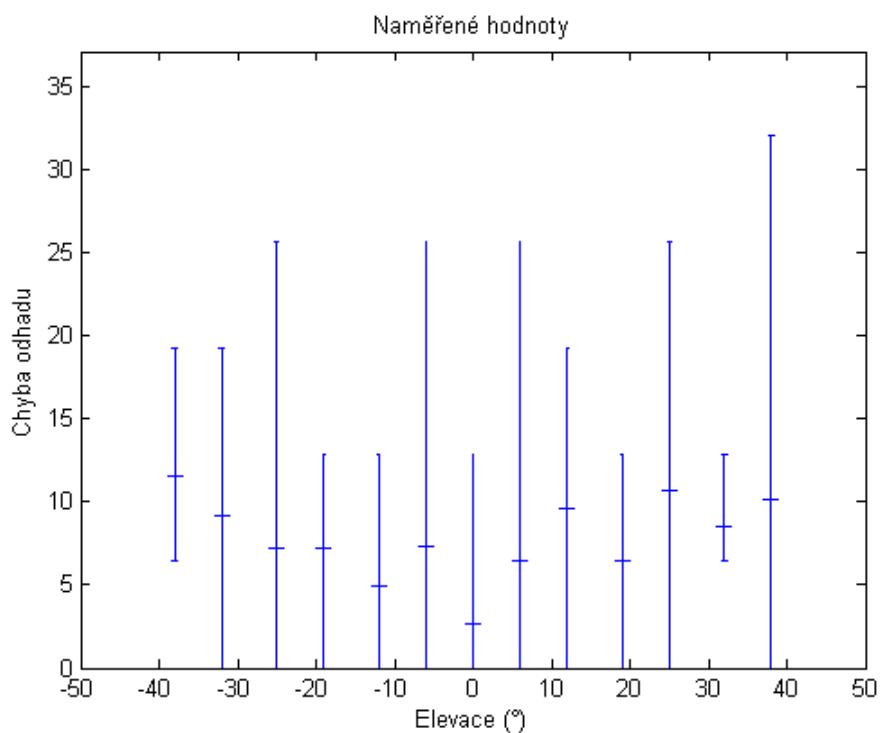
Obrázek 38: Výsledky měření azimutu - Použití myši



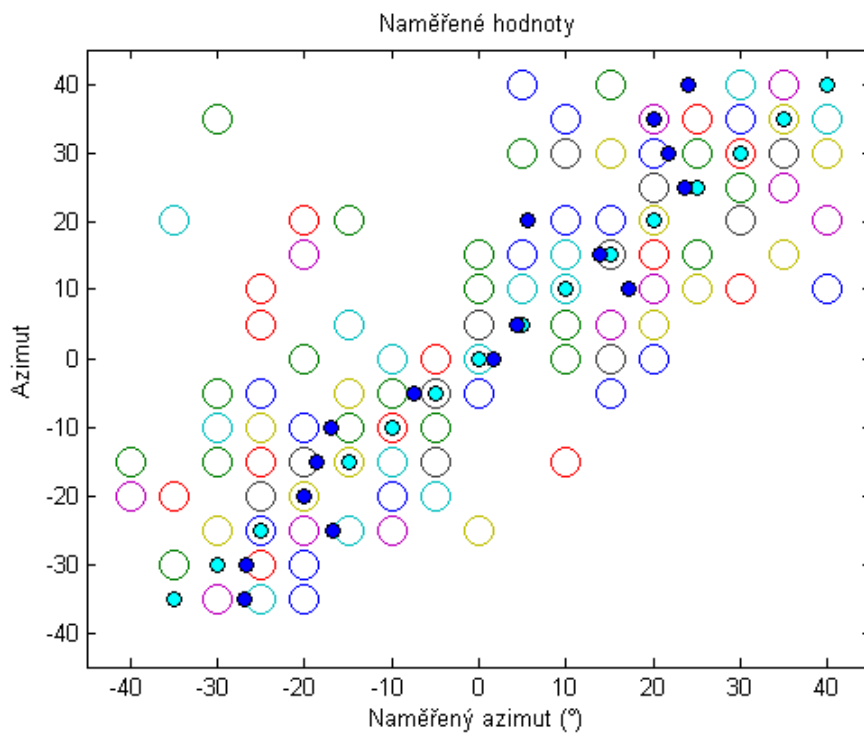
Obrázek 39: Odchylka měření azimutu - Použití myši



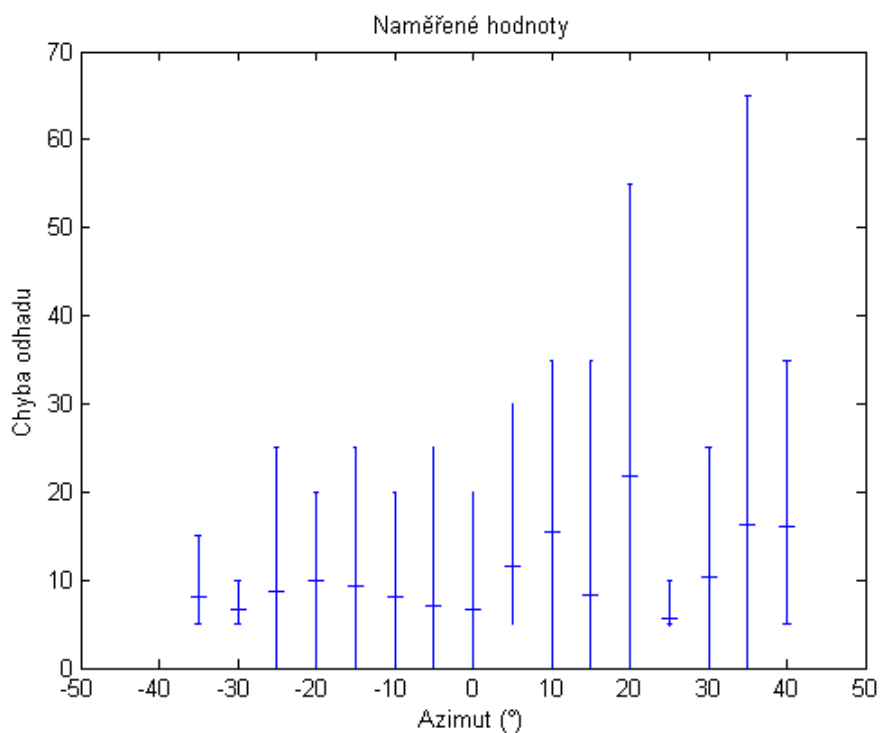
Obrázek 40: Výsledky měření elevace - Použití joysticku



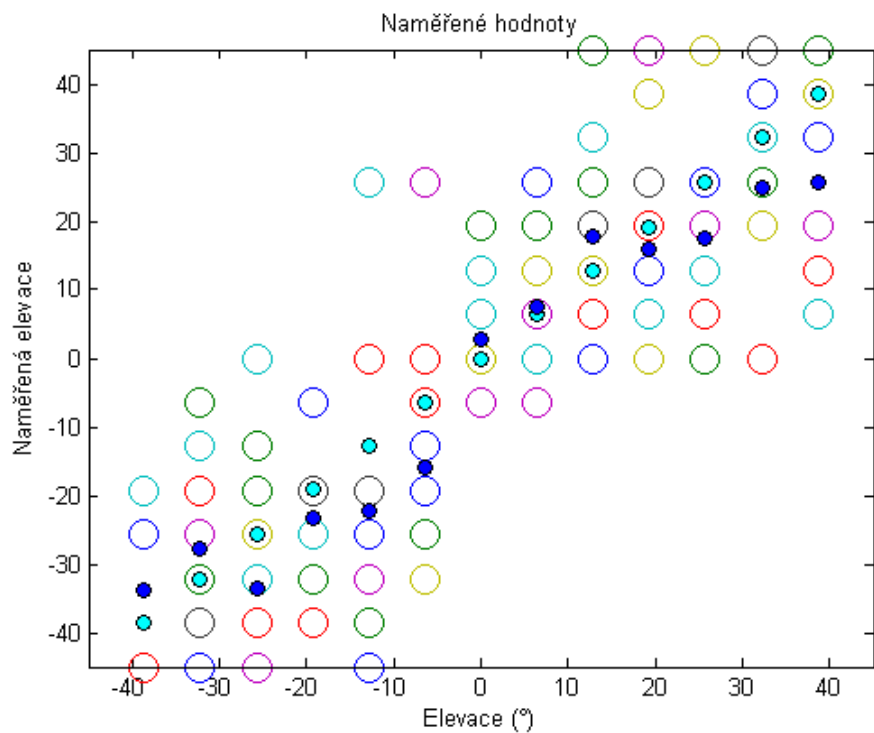
Obrázek 41: Odchylka měření elevace - Použití joysticku



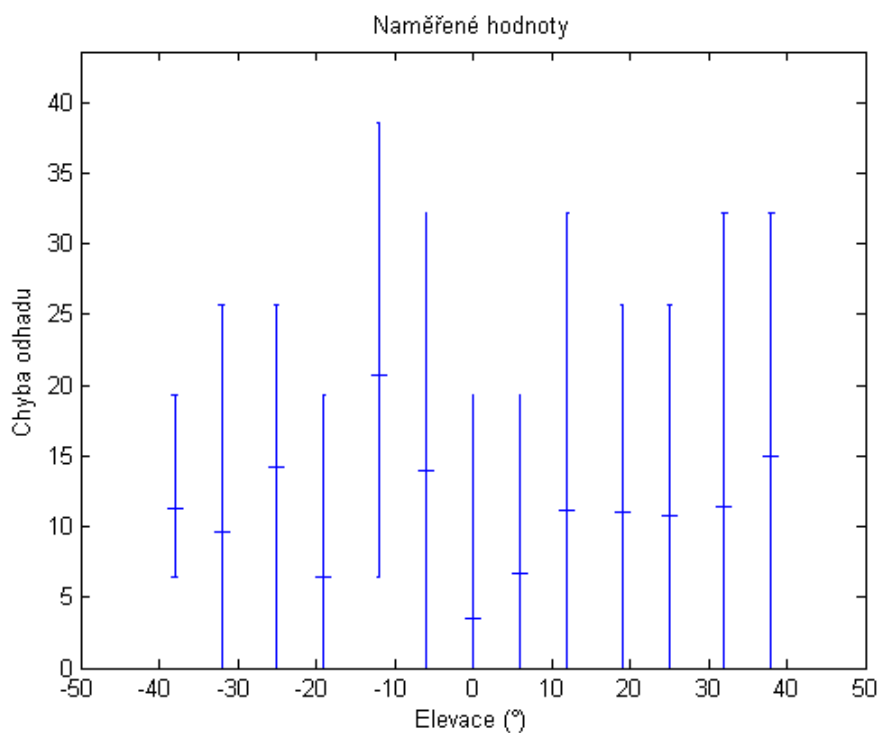
Obrázek 42: Výsledky měření azimutu - Použití joysticku



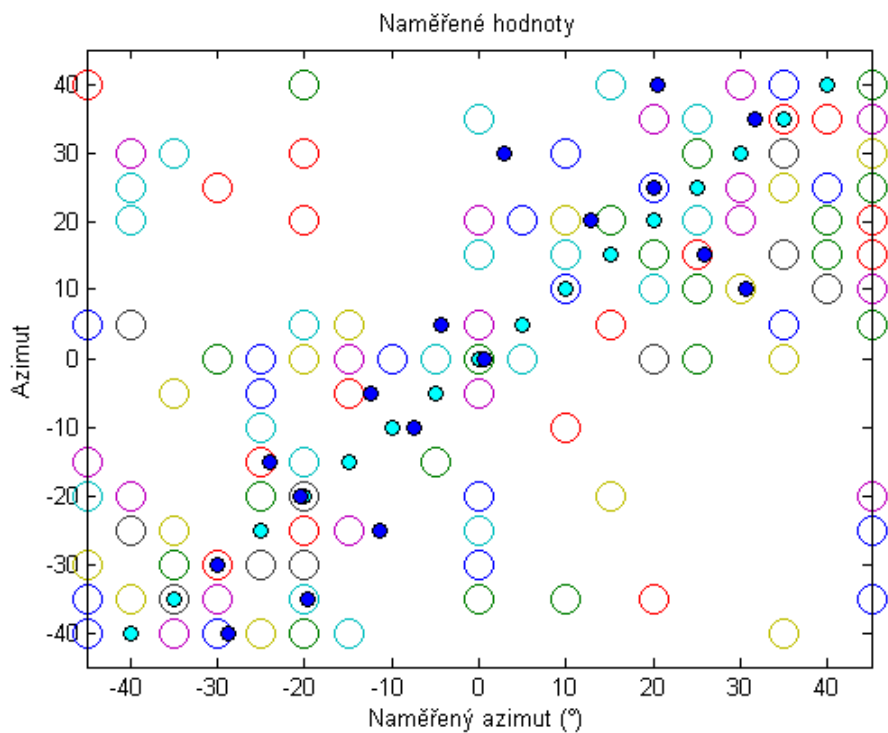
Obrázek 43: Odchylka měření azimutu - Použití joysticku



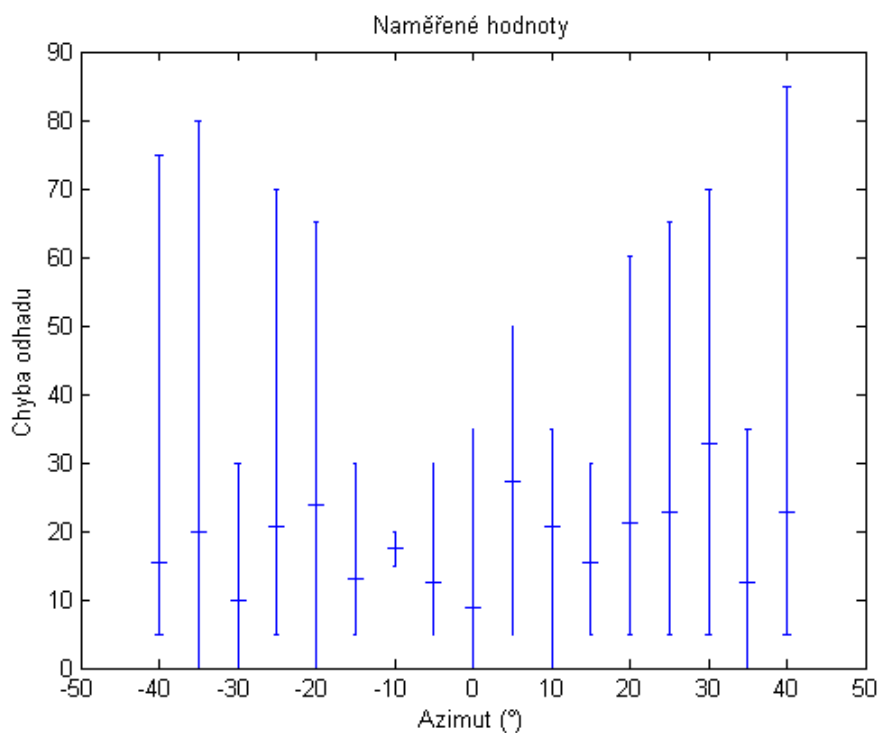
Obrázek 44: Výsledky měření elevace - Použití joysticku naslepo



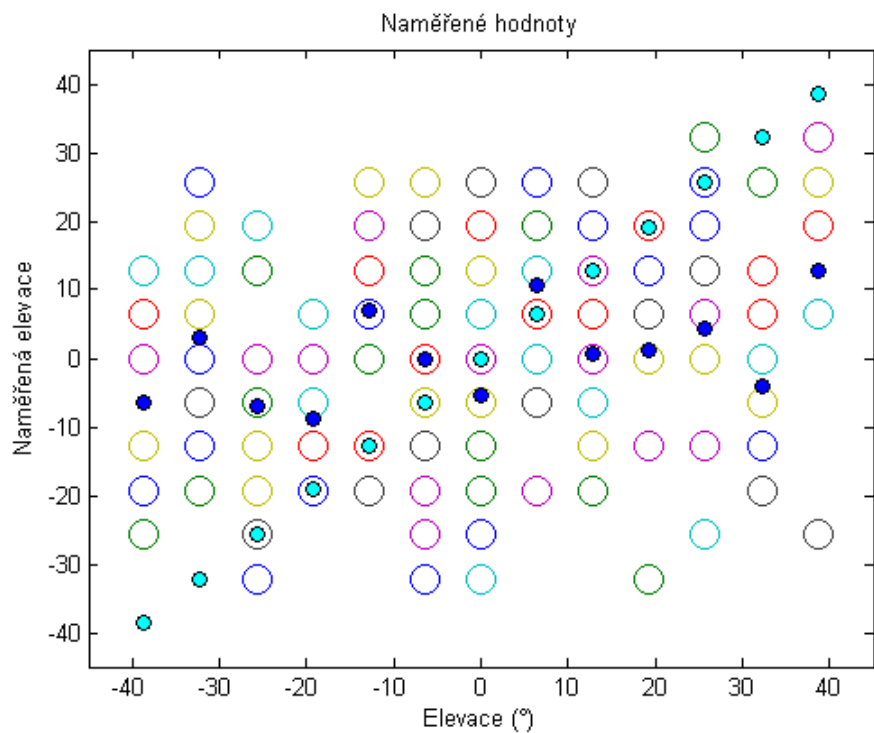
Obrázek 45: Odchylka měření elevace - Použití joysticku naslepo



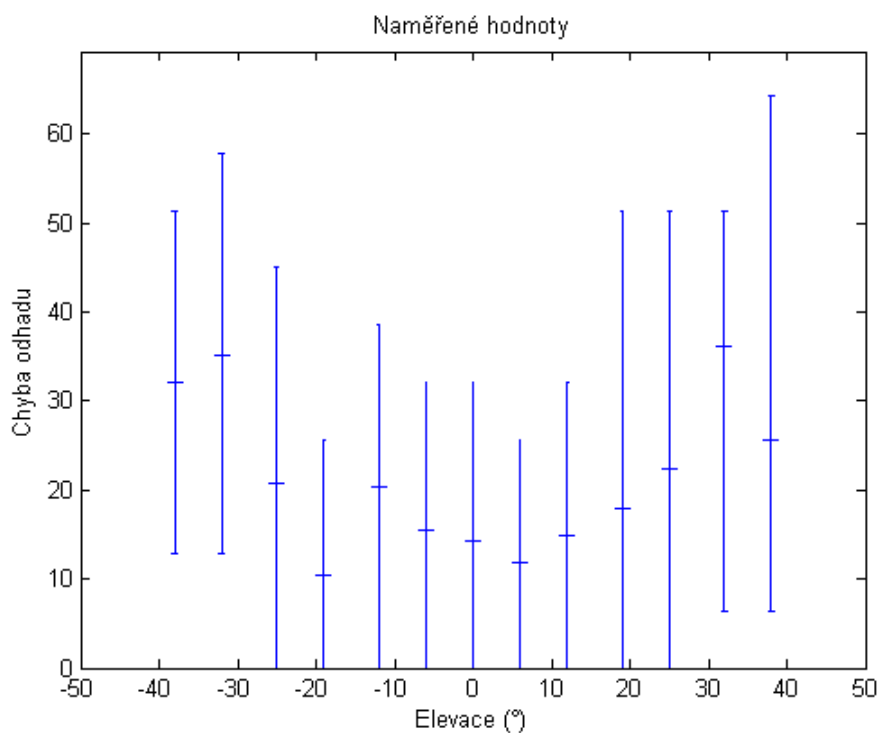
Obrázek 46: Výsledky měření azimutu - Použití joysticku naslepo



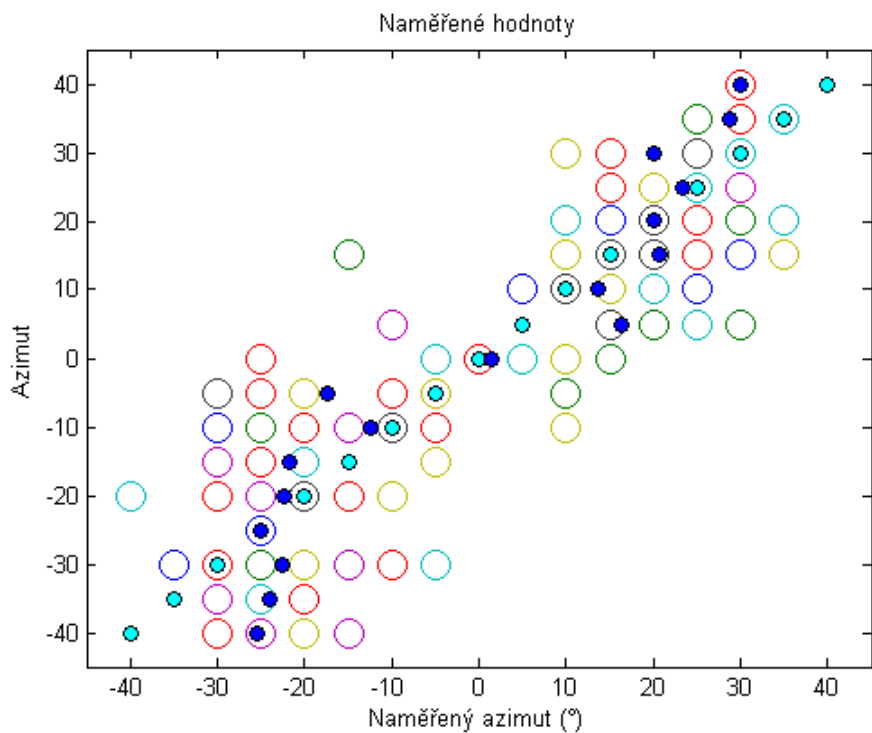
Obrázek 47: Odchylka měření azimutu - Použití joysticku naslepo



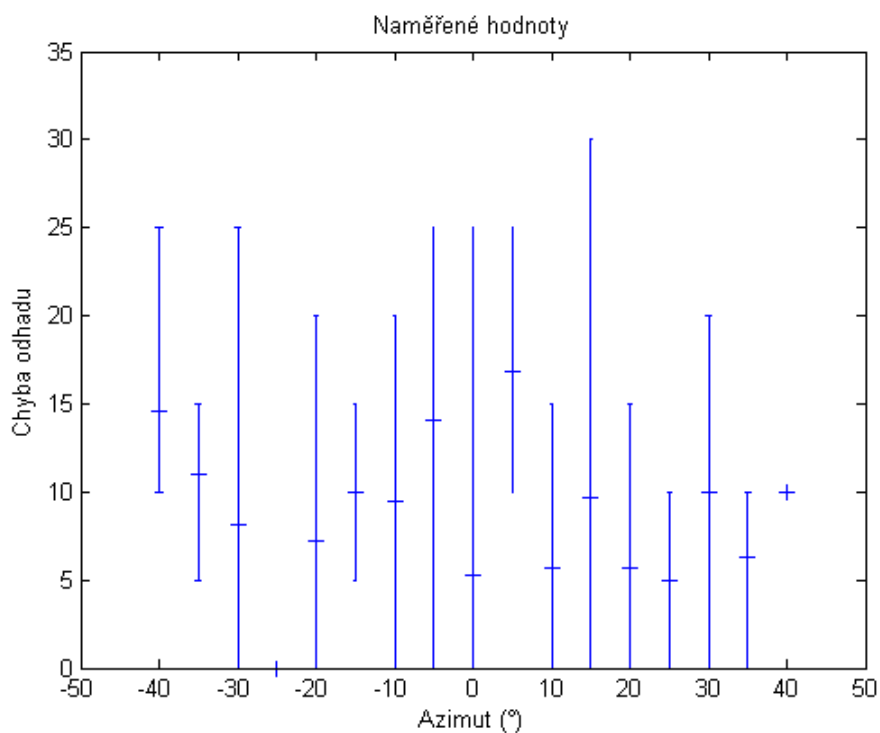
Obrázek 48: Výsledky měření elevace - Strukturální model boltce



Obrázek 49: Odchylka měření elevace - Strukturální model boltce



Obrázek 50: Výsledky měření azimutu - Strukturální model boltce



Obrázek 51: Odchylka měření azimutu - Strukturální model boltce