



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ**

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**AKUSTICKÁ DETEKCE POZICE ŘEČNÍKA POMOCÍ  
MIKROFONNÍHO POLE**

ACOUSTIC DETECTION OF SPEAKER POSITION USING MICROPHONE ARRAY

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**YEN NHI CHU**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. IGOR SZÖKE, Ph.D.**

BRNO 2019

## Zadání bakalářské práce



20489

Studentka: **Chu Yen Nhi**  
Program: Informační technologie  
Název: **Akustická detekce pozice řečníka pomocí mikrofonního pole**  
**Acoustic Detection of Speaker Position Using Microphone Array**  
Kategorie: Zpracování signálů

### Zadání:

1. Seznamte se s mikrofonními poli a s odhadem pozice řečníka (zdroje zvuku) vůči poli.
2. Nastudujte a zvolte vhodný algoritmus pro odhad pozice řečníka v místnosti na základě jeho řeči zachycené mikrofonním polem. Navrhněte vhodnou testovací metriku a datovou sadu.
3. Implementujte algoritmus a otestujte jeho přesnost. Vypočítejte polární souřadnice řečníka vůči libovolnému bodu.
4. Vylepšete a zefektivněte navržený algoritmus tak, aby mohl běžet na ARM procesorech.
5. Zhodnoťte výsledky a navrhněte směry dalšího vývoje.
6. Vytvořte A2 plakátek a cca 30 vteřinové video prezentující výsledky vašeho projektu.

### Literatura:

- Podle pokynů školitele

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 až 3 ze zadání.

Podrobné závazné pokyny pro vypracování práce viz <http://www.fit.vutbr.cz/info/szz/>

Vedoucí práce: **Szóke Igor, Ing., Ph.D.**  
Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.  
Datum zadání: 1. listopadu 2018  
Datum odevzdání: 15. května 2019  
Datum schválení: 10. května 2019

## Abstrakt

Cílem této práce je vytvořit systém umožňující detekci pozice řečníka v uzavřené místnosti opatřené mikrofonním polem. Přičemž se zabývá dvěma algoritmy pro lokalizaci zdroje zvuku, jmenovitě metodou tvarování přijímací charakteristiky a metodou časových zpoždění. Dané metody byly implementovány na poskytnutý hardware, spolehlivost metod otestována a výsledky byly vyhodnoceny. Výsledky ukázaly, že obě použité metody jsou poměrně přesné, pokud jsou použity v malém prostoru (přibližně 3×3 m). Nedostatky použitého řešení jsou podrobně rozebrány a jsou navržena možná vylepšení. Dané řešení je použitelné jako základ pro tvorbu vestavěných systémů, které budou moci určit polohu zdroje zvuku.

## Abstract

The purpose of this thesis is to create a system for detection of position in a closed room equipped with microphone array. It deals with two algorithms for sound source localization, namely beamforming and time difference of arrival. The methods were implemented on the provided hardware, reliability of the methods was tested and results were evaluated. The results have shown that both used methods were relatively accurate if used in small space (approx. 3×3 m). Shortcomings of the solution were analyzed in detail, possible improvements were proposed. This work can be used as a basis for creating embedded systems for sound source localization.

## Klíčová slova

mikrofonní pole, lokalizace zdroje zvuku, delay and sum metoda, metoda časových zpoždění

## Keywords

microphone array, sound source localization, delay and sum, time difference of arrival

## Citace

CHU, Yen Nhi. *Akustická detekce pozice řečníka pomocí mikrofonního pole*. Brno, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Igor Szöke, Ph.D.

# Akustická detekce pozice řečníka pomocí mikrofonního pole

## Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně pod vedením pana Ing. Igora Szökeho, Ph.D. Další informace mi poskytl Jan Havran z firmy Audified. Uvedla jsem všechny literární prameny a publikace, ze kterých jsem čerpala.

.....

Yen Nhi Chu  
15. května 2019

## Poděkování

Chtěla bych poděkovat svému vedoucímu doktorovi Szökemu za odborné vedení, cenné rady a připomínky, které mi pomohly práci dokončit. Dále bych chtěla poděkovat i Janu Havranovi z firmy Audified, který mi poskytl nutné informace pro práci se zapůjčeným hardwarem.



# Obsah

<b>1</b>	<b>Úvod</b>	<b>3</b>
<b>2</b>	<b>Problematika lokalizace zdroje zvuku</b>	<b>4</b>
2.1	Mikrofonní pole . . . . .	5
2.2	Prostředí šíření . . . . .	6
2.3	Zjednodušující předpoklady . . . . .	7
2.4	Určení úhlu dopadu signálu ze zpoždění . . . . .	8
2.5	Metoda časových zpoždění . . . . .	9
2.6	Metoda tvarování přijímací charakteristiky . . . . .	11
<b>3</b>	<b>Měření a testování</b>	<b>14</b>
3.1	Hardware . . . . .	14
3.1.1	Jack . . . . .	16
3.2	Zpracování signálů . . . . .	16
3.3	Testovací prostředí . . . . .	18
3.4	Průběh testování . . . . .	19
3.5	Testování určování úhlů ze zpoždění . . . . .	20
<b>4</b>	<b>Výsledky testování</b>	<b>22</b>
4.1	Metoda časových zpoždění . . . . .	22
4.1.1	Určování úhlů ze zpoždění . . . . .	22
4.1.2	Určení souřadnic zdroje . . . . .	26
4.2	Metoda Delay and Sum . . . . .	29
4.2.1	Určování úhlů ze zpoždění . . . . .	29
4.2.2	Určení souřadnic zdroje . . . . .	33
4.3	Porovnání použitých metod . . . . .	36
4.3.1	Lokalizace zdroje v prostoru se dvěma zdroji zvuku . . . . .	39
<b>5</b>	<b>Budoucí postup a možná rozšíření</b>	<b>43</b>
5.1	Zpracování ticha . . . . .	43
5.2	Použití více mikrofonních polí . . . . .	43
5.3	Osvětlení pozice zdroje reflektorem . . . . .	43
<b>6</b>	<b>Závěr</b>	<b>45</b>
	<b>Literatura</b>	<b>47</b>
<b>A</b>	<b>Další grafy výsledků testování</b>	<b>48</b>
A.1	Určování úhlu ze zpoždění . . . . .	48

A.2 Lokalizace v prostoru se dvěma zdroji zvuku . . . . .	50
<b>B Obsah přiloženého paměťového média</b>	<b>53</b>
<b>C Plakát</b>	<b>54</b>

# Kapitola 1

## Úvod

Obsah této práce se dotýká několika oblastí. Jednou z nich je zpracování signálů. Signálem je v této práci myšlen zvuk, který s sebou nese informace. Množství informací, které lze ze signálu získat závisí na přijímači.

Zpracování zvukového signálu může zahrnovat syntézu, modifikaci a analýzu signálu. Toho může být docíleno aplikací matematických algoritmů s účelem upravit signál tak, aby jeho informační hodnota byla co nejvyšší. Takto dokážeme zjistit mimo jiné lokaci zdroje signálu, získat obsah zprávy obsažené v signálu a informace o vlastnostech přenosového média (tj. prostředí šíření signálu) [4].

Další oblastí, které se lehce dotkneme je fyzika. Z fyzikálního hlediska považujeme zvuk za mechanické vlnění v látkovém prostředí. Tímto prostředím může být například vzduch, ale i fyzické těleso. Prostředí šíření zvuku má velký dopad na to, jak výsledný zachytávaný signál vypadá.

Cílem práce je implementovat algoritmus pro lokalizaci zdroje zvuku na zapůjčený hardware a vyhodnotit dosažené výsledky. Nedílnou součástí řešení je tedy i samotné ověření skutečné aplikovatelnosti v reálném prostředí. Pro vyhodnocení úspěšnosti použitých metod se využijí statistické veličiny.

Samotná lokalizace zdroje, ať už zvuku či jiného signálu, nachází uplatnění v celé škále oblastí. V některých z nich se využívá již delší dobu, například v radarové a sonarové technologii. Stále však vznikají nové příležitosti, kde se může lokalizace zdroje využít. Jedním z takových novějších uplatnění je při videokonferencích, kde se používají kamery, které se dokáží otáčet směrem ke zdroji zvuku.

Tato práce může sloužit jako základ pro tvorbu vestavěných systému, které vyžadují detekci pozice řečníka. V práci se tedy zaměřuji především na určení toho, na kolik je zvolený algoritmus přesný při zpracovávání lidské řeči. Konkrétně se práce zabývá dvěma algoritmy umožňující lokalizaci. S tím, že výsledné zhodnocení by mělo taktéž zahrnovat informaci o tom, který z vybraných algoritmů je vhodnější.

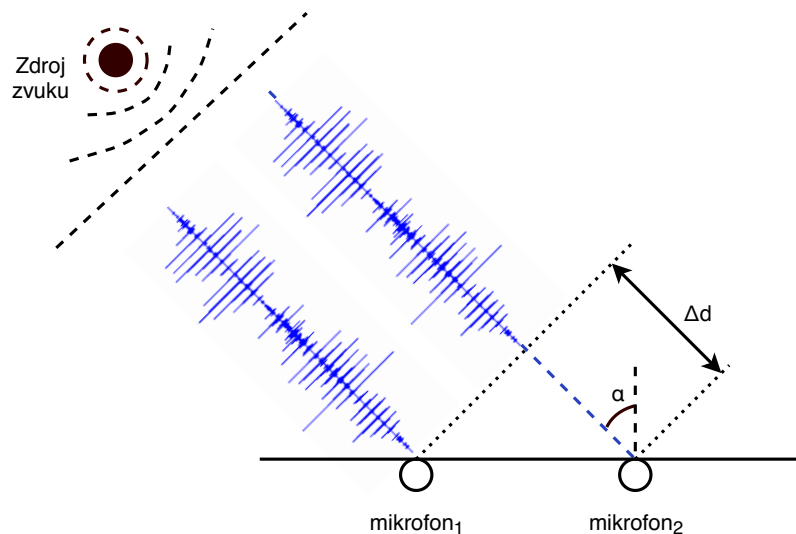
Tato práce se v kapitole 2 zabývá teorií za samotnou lokalizací zvuku, to zahrnuje stanovení předpokladů nutných pro lokalizaci a představení používaných algoritmů. Poté následuje specifikace testování v kapitole 3, včetně popisu měřícího aparátu a určení testovacích metrik. Kapitola 4 s výsledky testování obsahuje vyhodnocení samotných metod i jejich vzájemné porovnání. V předposlední kapitole navrhuji určitá vylepšení do budoucna.

## Kapitola 2

# Problematika lokalizace zdroje zvuku

Cílem práce je prokázat aplikovatelnost algoritmů pro lokalizaci zdroje zvuku v kombinaci s hardwarem poskytnutým fakultou FIT VUT. Pro splnění tohoto cíle je nutné, si představit různé metody používané k lokalizaci.

Princip lokalizace zdroje zvuku spočívá v určení úhlu, pod kterým dorazí zvukové vlny na mikrofony (*DOA - Direction of Arrival*). K tomu se používá mikrofonní pole, které dokáže vytvořit vícekanálový záznam zvuku. Ten je následně zpracován jednou z metod pro určení zdroje zvuku.



Obrázek 2.1: Šíření zvukového signálu k mikrofonom,  $\alpha$  je úhel dopadu,  $\Delta d$  je dráha, kterou musí signál urazit navíc k druhému mikrofonomu

Existují tři hlavní skupiny metod, které se používají pro určování polohy zdroje zvuku.

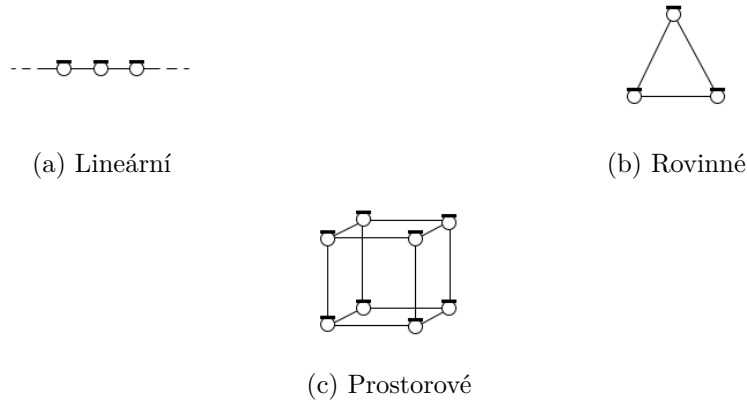
- Metoda časových zpoždění (*Time Difference of Arrival - TDOA*)
- Metoda tvarování přijímací charakteristiky (*Beamforming*)
- Metody spektrálních odhadů

Práce se detailněji zaměřuje na první dvě metody (sekce 2.5 a 2.6), obě tyto metody lze totiž provádět v časové oblasti, čímž se liší od metod spektrálních odhadů. Třetí možnost se od prvních dvou podstatně liší a není vhodná pro zpracování řeči kvůli její značně vysoké výpočetní náročnosti [5].

## 2.1 Mikrofonní pole

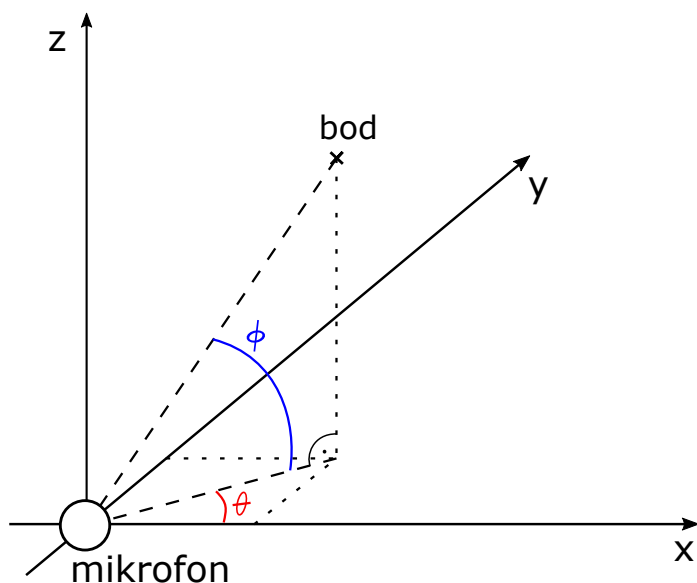
Mikrofonní pole je sensorové pole tvořené dvěma či více mikrofony. Šířený zvuk, který se dostane k mikrofonnímu poli je zachycen všemi senzory (mikrofony). Každý mikrofon tedy poskytuje vlastní výstupní signál. V nejjednodušších případech se tyto výstupní signály příliš neliší a jedná se v podstatě o stejný signál, jen zpožděn vůči ostatním a to kvůli tomu, že se mikrofony nachází v různých vzdálenostech od zdroje. V horších případech jsou signály různě ovlivňovány vnějším prostředím, nejeví se jako stejné a jejich zpracování se tímto komplikuje [4].

Podle prostorového rozložení jednotlivých mikrofonů můžeme mikrofonní pole dělit na lineární, rovinné a prostorové (obrázek 2.2).



Obrázek 2.2: Možná uspořádání mikrofonních polí

Lineární a rovinná pole se používají pro lokalizaci zdroje v rovině. Lineární pole dokáže rozlišit úhel dopadu do rozmezí až  $180^\circ$ , rovinné pokryje celých  $360^\circ$ . Pomocí prostorového mikrofonního pole lze odhadnout pozici zdroj zvuku v trojrozměrném prostoru. K tomu je třeba získat hodnoty azimutu a elevace. Uvažujeme-li klasický trojrozměrný souřadnicový systém (obrázek 2.3). Azimut ( $\theta$ ) je úhel, který svírá obraz zdroje, promítnutý na rovinu  $xy$ , s osou  $x$  v rovině  $xy$ . Může nabývat hodnot  $0^\circ$  až  $360^\circ$ . Elevace ( $\phi$ ) je úhel mezi osou  $z$  a zdrojem zvuku v rovině, která prochází zdrojem zvuku a osou  $z$ . Elevace nabývá hodnot  $-90^\circ$  až  $90^\circ$ . Pro zjištění jednotlivých úhlů se využívá jen ta část mikrofonního pole, která leží ve stejné rovině jako je hledaný úhel. Chceme-li tedy získat azimut uvažujeme jen ty mikrofony, které jsou spolu spojeny horizontálně, podobně u elevace uvažujeme jen vertikálně spojené mikrofony.



Obrázek 2.3: Úhly pro určení bodu v prostoru –  $\theta$  je azimut (leží v rovině  $xy$ ),  $\phi$  je elevace (leží v rovině kolmé na rovinu  $xy$ )

Další možností dělení mikrofonního pole je podle vzdáleností mezi jednotlivými mikrofony. Takové dělení rozlišuje uniformní a neuniformní mikrofonní pole. Uniformní pole je tvořeno mikrofony, které jsou ve stejné vzdálenosti od sebe. Naopak v neuniformním poli se mikrofony mohou od sebe nacházet v různých vzdálenostech.

Nejčastěji se lze setkat s uniformními mikrofonními poli. Jedná se především o uniformní lineární pole (*ULA – Uniform Linear Array*), uniformní kruhové pole (*UCA – Uniform Circular Array*) a uniformní rovinné pole (*UPA – Uniform Planar Array*). UPA nachází uplatnění především u velkých vojenských systémů s fázovacími poli [4].

Kolik mikrofونů je třeba pro určení polohy zdroje je závislé na velikosti a tvaru prohledávaného prostoru, zvolené metodě, aj. Obecně však platí, čím více mikrofونů, tím přesnější je odhad.

## 2.2 Prostředí šíření

V reálném prostředí existuje mnoho faktorů, které mohou ovlivnit vlastnosti výsledného zachytávaného signálu. Například dozvuk a interference.

Zvuk se šíří od zdroje všemi směry ve formě vln, narazí-li na libovolnou překážku odrazí se od ní, tím vznikají další vlny. Takto dochází ke vzniku dozvuku a zvukové interference [1].

Zvuková interference je jev, kdy odražená vlna postupuje proti šířené vlně a dochází ke skládání jejich vlnění. Podle toho jakou fázi má odražená vlna, může výsledný složený signál být silnější či slabší. Tohoto jevu se využívá u aktivního potlačování zvuku (*active noise cancellation*).

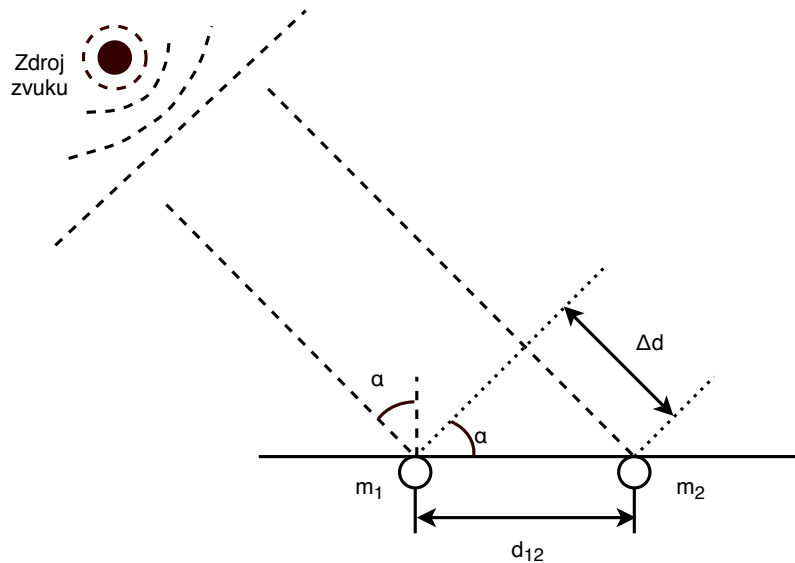
Odražená vlna cestuje déle než vlna přímá, dorazí tedy oproti ní se zpožděním. Při každém odrazu také dochází k tomu, že část energie vlnění je pohlcena překážkou, což má za následek snížení intenzity zvuku. Takto vzniká dozvuk (ozvěna). Dozvuk může být zachycen mikrofonom, pro lokalizaci však není podstatný, spíše má negativní efekt v tom, že zkresluje zpracovávaný signál.

Výběrem vhodného prostředí lze tyto vlivy do určité omezit, např. bezodrazové komory.

## 2.3 Zjednodušující předpoklady

Při zpracování dat z reálného světa není možné vždy zahrnout všechny vnější vlivy do výpočtu. Proto si zde definujeme určité zjednodušující předpoklady a podmínky, za kterých se budou výpočty provádět.

Metody zmiňované v sekcích 2.5 a 2.6 předpokládají, že zdroj zvuku se nachází v dostatečně velké vzdálenosti od přijímačů tak, že se dopadající vlny jeví jako rovinné, díky tomuto předpokladu lze tvrdit, že dopadá-li zvuková vlna na vícero mikrofonů pod úhlem jiným než nulovým (tzn. zdroj se nachází přímo před mikrofonem), dorazí zvuková vlna ke každému mikrofonu s různým zpožděním. Dále to také zajišťuje, že vlna dopadá na každý mikrofon (jsou-li v řadě za sebou) pod stejným úhlem.



Obrázek 2.4: Dopad zvuku na mikrofonní pole o 2 mikrofonech (označeny  $m_1$  a  $m_2$ ),  $\alpha$  je úhel dopadu,  $\Delta d$  je rozdíl v dráze, kterou musí signál urazit navíc k mikrofonu  $m_2$ ,  $d_{12}$  je vzdálenost mikrofonů

Z obrázku lze vidět, že zvuková vlna, která dorazí k mikrofonu  $m_2$  musela oproti  $m_1$  urazit ještě navíc dráhu  $\Delta d$ . Taktéž lze vidět úhel dopadu  $\alpha$  a fakt, že se počítá vzhledem ke kolmici k mikrofonom.

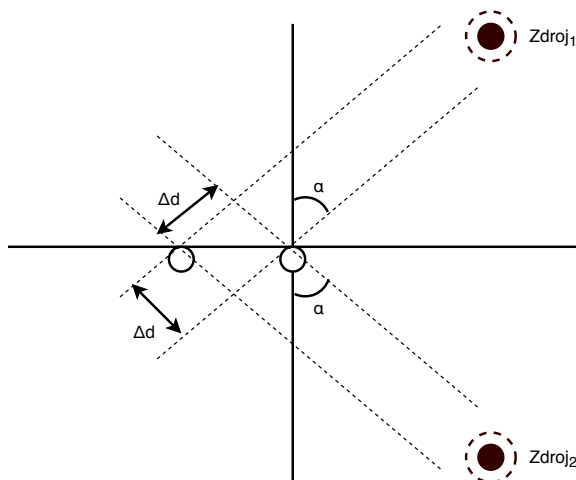
Dále platí, že prostředí, ve kterém se šíří zvuk, považujeme za homogenní (stejnorodé). V takovém případě rychlost šíření zvuku bude stálá. Z vzorce rychlosti zvuku ve vzduchu lze vidět, její závislost na teplotě okolního prostředí.

$$c = (331,57 + 0,607 \cdot t) \text{ m s}^{-1} \quad (2.1)$$

Při výpočtu neuvažujeme žádné vnější vlivy typu dozvuk, avšak při výsledném zhodnocení by bylo vhodné zhodnotit na kolik mohly ovlivnit výsledky měření.

Úhel dopadu zvuku může nabývat pouze hodnot  $-90^\circ$  až  $90^\circ$ , kde  $0^\circ$  odpovídá kolmici k mikrofonom. A to protože u ULA (2.2a) tvořeného mikrofony, které dokáží zachycovat zvuk ze všech směrů, dochází k tzv. předozadní nejednoznačnosti. Ta je dána tím, že mikrofon nedokáže rozlišit, zda se zdroj zvuku nachází před nebo za ním. Úhel dopadu je především závislý na zpoždění zvuku, a protože zpoždění zvuku přicházejícího zezadu mi-

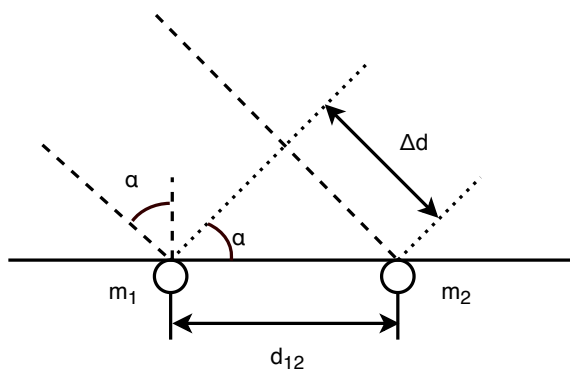
krofonního pole je stejné jako zpoždění zvuku přicházející zepředu, nelze rozlišit tyto dvě situace od sebe.



Obrázek 2.5: Předozadní nejednoznačnost, samotné určení úhlu  $\alpha$  nevypovídá nic o tom, zda se zdroj nachází před nebo za mikrofonom

## 2.4 Určení úhlu dopadu signálu ze zpoždění

Jednou ze způsobů, jak určit pozici zdroje zvuku je zjistit pod jakým úhlem dopadají na mikrofón přicházející signály. Daný úhel se dá spočítat se znalostí toho, jaké zpoždění je mezi signály, které dorazily ke dvěma sousedním mikrofónům.



Obrázek 2.6: Detailní obrázek dopadu signálu na mikrofóny,  $\alpha$  je úhel dopadu,  $\Delta d$  je rozdíl drah,  $d_{12}$  je vzdálenost mezi mikrofóny  $m_1$  a  $m_2$

Použité vzorce pro výpočet úhlu:

$$\Delta d = \frac{N}{F_s} \cdot c \quad (2.2)$$

$$\varphi = \arcsin \frac{\Delta d}{d_{12}} \quad (2.3)$$

kde  $N$  je zpoždění signálů ve vzorcích,  $F_s$  je vzorkovací frekvence,  $c$  je rychlost zvuku,  $d_{12}$  je vzdálenost mezi dvěma mikrofóny v mikrofónním poli a  $\varphi$  je úhel dopadu odpovídající danému zpoždění.

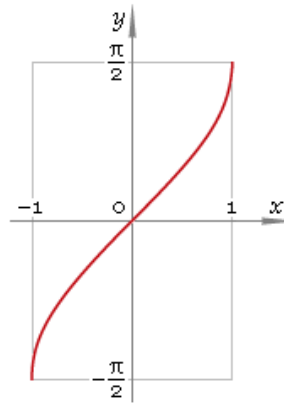


Je nutné zmínit, že vypočítaný úhel nabývá hodnot z intervalu  $\langle -90^\circ; 90^\circ \rangle$ , to je dáno tím, že zpoždění může nabývat kladných i záporných hodnot. Vzorce (i obrázek!2.6) počítají s tím, že jeden mikrofon je od zdroje dál, ale v případě, že se orientace obrátí a vzdálenější mikrofon je ve skutečnosti blíže ke zdroji, tak je výsledkem záporné zpoždění.

Při určování úhlu při lokalizaci je nutné si uvědomit, že se používá funkce arcsin, což má dopad na jemnost úhlů, které je možné dosáhnout s omezeným intervalem možných zpoždění.

zpoždění	úhel [°]
-13	-69.70
-12	-59.97
-11	-52.52
-10	-46.18
⋮	⋮
-2	-8.3
-1	-4.13
0	0
1	4.13
2	8.3
⋮	⋮
11	52.52
12	59.97
13	69.70

(a) Tabulka některých hodnot zpoždění a jim odpovídající úhly získané aplikací vzorců 2.2 a 2.3



(b) Průběh funkce arcsin

Obrázek 2.7: Znázornění omezené přesnosti při určování úhlů

Na obrázku 2.7b s průběhem funkce arcsin, lze vidět, že čím blíže k hodnotě 1 tím je funkční hodnota arcsin strmější. V podstatě to znamená, že blíže k  $90^\circ$  (resp.  $\frac{\pi}{2}$ , stejně platí i zrcadlově k  $-90^\circ$  a  $-\frac{\pi}{2}$ ) se úhly (hodnota osy  $y$ ) mění hodně i s malým krokem na ose  $x$ . Konkrétně zpoždění 1 prvku u úhlů okolo  $0^\circ$  odpovídá  $4^\circ$  avšak zpoždění 1 vzorku u úhlů na krajích ( $90^\circ$  a  $-90^\circ$ ) odpovídá více než  $10^\circ$ , jak lze vidět v tabulce 2.7a. Takže s ohledem na to, že nedokážeme rozlišovat zpoždění jemnější než 1 vzorek, je tímto přesnost určení úhlu u krajních hodnot (tj. dál od  $0^\circ$ ) podstatně omezená.

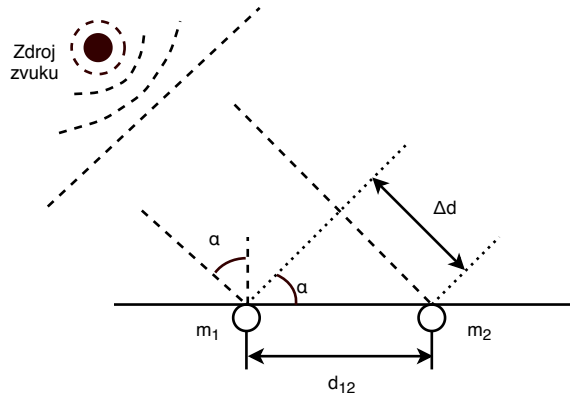
Za podmínek specifikovaných v sekci o testování 3.3 může zpoždění nabývat jen hodnot z intervalu  $\langle -13; 13 \rangle$ , což vede k tomu, že určení úhlu je značně omezeno. Z tabulky 2.7a lze vidět jakých konkrétních hodnot může hledaný úhel nabývat.

## 2.5 Metoda časových zpoždění

Tato podkapitola vychází primárně z následujících zdrojů: [6], [2] a [3].

Jedná se o často používanou metodu, protože je výpočetně nenáročná. Nevýhodou této metody je, že nepředpokládá výskyt vícero zdrojů zvuků. Metoda se skládá z dvou hlavních kroků, proto se jí také někdy říká dvoukroková metoda. V prvním kroku dochází k určení

zpoždění signálů mezi zvolenými dvojicemi mikrofoni pomocí vzájemné korelační funkce. Druhý krok pak s pomocí znalosti určených zpoždění a geometrie mikrofonního pole zjišťuje pozici zdroje zvuku.



Obrázek 2.8: Metoda časových zpoždění -  $\alpha$  je úhel dopadu,  $\Delta d$  je rozdíl drah,  $d_{12}$  je vzdálenost mezi mikrofony  $m_1$  a  $m_2$

Z obrázku lze vidět že, zvuková vlna, která dorazí k mikrofonu  $m_1$  musí urazit ještě dodatečnou dráhu k mikrofonu  $m_2$ . Tuto dráhu navíc lze vyjádřit dvěma způsoby

$$\Delta d = d_{12} \cdot \sin \alpha \quad (2.4)$$

$$\Delta d = \tau_{12} \cdot c \quad (2.5)$$

kde  $d_{12}$  je vzdálenost mezi mikrofony,  $\alpha$  je úhel dopadu,  $c$  je rychlost šíření zvuku a  $\tau_{12}$  je časové zpoždění zvuku mezi mikrofony  $m_1$  a  $m_2$ .

Aplikací vzájemné korelační funkce na signály z mikrofoni lze získat zpoždění v počtech vzorků. Pro následující vzorec platí, že  $x_1$  a  $x_2$  jsou zpracovávané signály z mikrofoni  $m_1$  a  $m_2$ ,  $k$  je posun ve vzorcích, pro který se počítá korelační koeficient  $R$  a  $N$  je celkový počet vzorků.

$$R[k] = \frac{1}{N} \sum_{n=0}^{N-1} x_1[n] \cdot x_2[n+k] \quad (2.6)$$

Výstupem této funkce je graf podobnosti mezi korelovanými vstupy. Index  $k$ , na kterém se nachází globální maximum grafu, odpovídá počtu vzorků, o které se musí posunout druhý signál, aby podobnost byla nejvyšší. Se znalostí  $k$  a vzorkovací frekvence  $F_{vz}$  lze zjistit časové zpoždění.

$$\tau = \frac{k}{F_{vz}} \quad (2.7)$$

Druhý krok spočívá v kombinaci spočteného zpoždění a znalosti geometrie mikrofonního pole pro sestavení vzorce hyperboly, jehož výsledkem bude množina bodů, na kterých se může vyskytovat zdroj zvuku při daném zpoždění. Pro zjištění konkrétní polohy zdroje zvuku je třeba zkonstruovat aspoň 2 různé hyperboly a najít jejich průsečík. To znamená, že je nutné mít aspoň 2 mikrofonní pole.

Na začátku druhého kroku je třeba si stanovit souřadný systém a v něm vyjádřit polohy obou mikrofoni. Řekneme-li, že zdroj se nachází na souřadnicích  $s_x, s_y$  a mikrofony mají souřadnice  $m_{1x}, m_{1y}$  a  $m_{2x}, m_{2y}$ . Pak přímá dráha od zdroje k mikrofoni se dá vyjádřit

následovně:

$$d_{m_1} = \sqrt{(s_x - m_{1x})^2 + (s_y - m_{1y})^2} \quad (2.8)$$

$$d_{m_2} = \sqrt{(s_x - m_{2x})^2 + (s_y - m_{2y})^2} \quad (2.9)$$

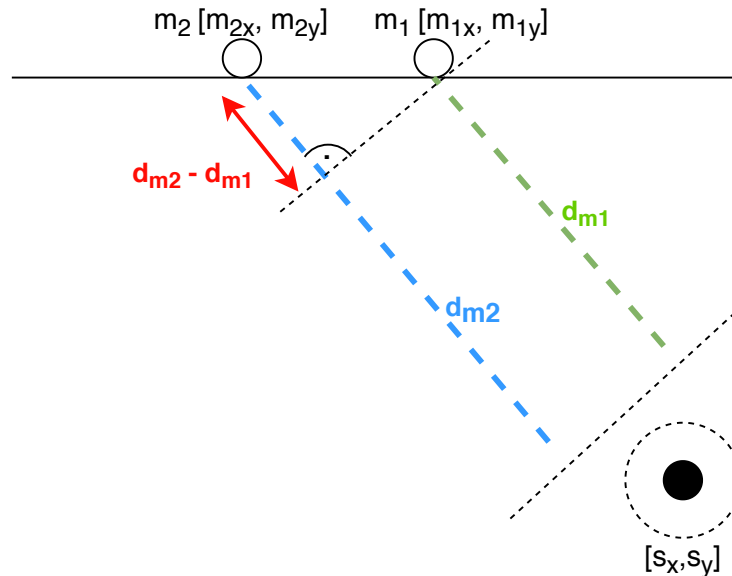
Rozdíl v drahách lze spojit s dobou zpoždění, s jakou dorazí zvuková vlna ke vzdálenějšímu mikrofonu.

$$\tau = \frac{d_{m_2} - d_{m_1}}{c} \quad (2.10)$$

Po dosazení rovnic 2.8 a 2.9 do 2.10

$$\tau = \frac{\sqrt{(s_x - m_{2x})^2 + (s_y - m_{2y})^2} - \sqrt{(s_x - m_{1x})^2 + (s_y - m_{1y})^2}}{c} \quad (2.11)$$

kde  $s_x$  a  $s_y$  představují proměnné v rovnici. Takto vypadající vzorec je pak nápadně podobný vzorci pro popis kuželosečky, konkrétně hyperboly.



Obrázek 2.9: Výpočet rozdílu drah signálů,  $[s_x, s_y]$  jsou souřadnice zdroje zvuku,  $d_{m_N}$  jsou dráhy od zdroje k příslušným mikrofonům  $m_N$  na souřadnicích  $[m_{Nx}, m_{Ny}]$ ,  $N$  je číslo mikrofonu

## 2.6 Metoda tvarování přijímací charakteristiky

Tato podkapitola vychází primárně z následujících zdrojů: [6], [2] a [3].

Podstata této metody spočívá v tom, že se zužuje prostor, na který mikrofony zaměřují své vnímání. Tímto lze i redukovat vlivy okolí, které chceme zanedbat. Oblast, na kterou se zužuje vnímání mikrofonů je určeno silou signálu, které vychází z daného místa. Pro tuto metodu se někdy také používá označení tvarovač (*beamformer*).

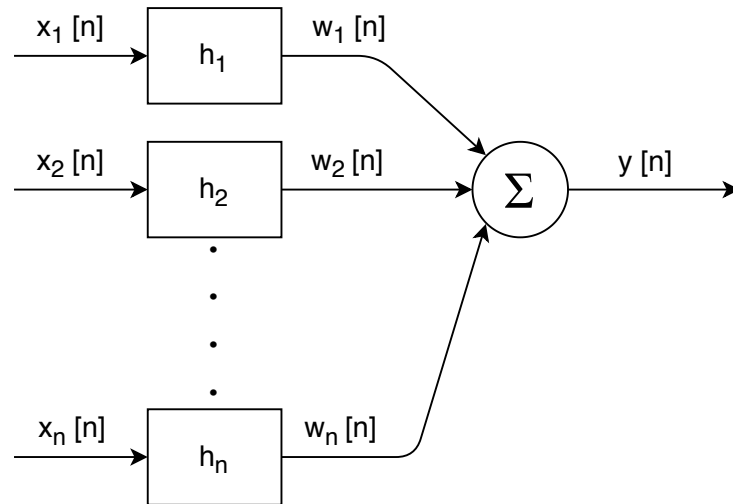
Pro tvarování charakteristiky nepředstavuje problém, pokud se v prohledávaném prostoru nachází více zdrojů zvuků, protože se vždy vybere jen ten s nejsilnějším signálem.

V této kapitole se budeme zaměřovat hlavně na Delay and Sum techniku (dále jen jako DAS), která je nejjednodušší variantou této metody a je označována jako konvenční

tvarovač. Jednoduchost této varianty spočívá ve faktu, že zpracovávané signály nejsou nijak pozměňovány a provádí se nad nimi jen součet a posun, kdežto jiné varianty tvarovačů mohou na signály aplikovat filtry a měnit jejich charakteristiky.

Hlavní myšlenka DAS zakládá na myšlence, že signály snímané jednotlivými mikrofony jsou totožné jen vždy opožděné vůči sobě o určitou dobu. V předchozí podkapitole o metodě časových zpoždění byl představen vzorec 2.4, který spojuje úhel dopadu s dobou zpoždění. Lze tedy tvrdit, že určitému zpoždění vždy odpovídá určitý úhel dopadu.

U DAS se zpoždění nevypočítává explicitně jako tomu bylo u metody časových zpoždění. Jednotlivé zpracovávané signály jsou vůči sobě posouvány a sčítány spolu ve snaze získat co nejvyšší energetickou hodnotu jejich součtu. Tato nejvyšší hodnota je dána tím, že sčítané signály byly vůči sobě posunuty, tak že se zpoždění eliminovalo, a tím vznikl nejsilnější možný složený signál.



Obrázek 2.10: Blokové schéma pro Delay and Sum,  $x_i$  jsou vstupní signály, blok  $h_i$  aplikuje zpoždění na dané signály,  $w_i$  posunutý výstupní signály a  $y$  je signál vytvořený součtem všech  $w_i$  signálů

Blokové schéma na obrázku 2.10 popisuje DAS pro obecný počet mikrofonů. Mikrofony zachytávají signály  $x_i$ , bloky  $h_i$  představují aplikaci zpoždění na signály, výstupní  $w_i$  signály jsou tedy vůči  $x_i$  posunuty o  $h_i$ . Nakonec jsou všechny sečteny do jednoho složeného signálu  $y$ .

$$y[k] = \sum_{i=0}^n x_i[k - h_i] \quad (2.12)$$

kde  $n$  je počet mikrofonů.

Energetická hodnota složeného signálu se nakonec může vypočítat vzorcem RMS (*Root Mean Square*).

$$V_{RMS} = \sqrt{\frac{1}{N} \sum_{k=0}^N y^2[k]} \quad (2.13)$$

kde  $N$  je délka zkoumaného signálu.

Výpočetní složitost DAS spočívá ve faktu, že je třeba  $V_{RMS}$  počítat pro každé možné zpoždění mezi signály. S rostoucím počtem vstupních signálů se tato složitost ještě násobně zvyšuje.

Maximální získaná  $V_{RMS}$  udává ideální posun mezi signály, z něj lze pak lehce spočítat úhel dopadu zvuku.

Pro správné fungování DAS je nutné se zaměřit také na vhodné rozmístění mikrofonů. Protože může docházet k tzv. prostorovému aliasingu, který má za následek nejednoznačnost při určování úhlu (výskyt několik maximálních hodnot). Prostorovému aliasingu lze zabránit, když fázový rozdíl mezi dvěma prostorově vzorkovanými signály nepřesáhne interval  $(-\pi, \pi)$  [2]. Toho je zajištěno, když vzdálenost dvou sousedních mikrofonů bude splňovat následující podmínku.

$$d < \frac{c}{2 \cdot f_{max}} \quad (2.14)$$

kde  $f_{max}$  je maximální frekvence zachytávaného signálu (v našem případě by to měla být převážně lidská řeč).

Takové omezení by v případě lidské řeči, která může dosahovat  $f_{max}$  až 10 kHz, umožňovalo stavět mikrofony řádově jen v jednotkách centimetrů od sebe. A to i v případě uvažování nejnižšího přijatelného maxima 3 kHz, při kterém je zajištěna aspoň nejzákladnější srozumitelnost řeči. Avšak protože lidská řeč není nutně harmonická, negativní vliv prostorového aliasingu by nemusel být příliš závažný, proto jsem se rozhodla toto omezení zanedbat při testování.

## Kapitola 3

# Měření a testování

Praktická část této bakalářské práce se zaměří na ověření metod TDOA<sup>1</sup> a DAS<sup>2</sup> a vyhodnocení jejich aplikovatelnosti v praxi.

Hlavní rozdíly mezi těmito dvěma metodami lze shrnout do dvou bodů. Prvním je rozdíl ve způsobu, kterým je určováno časové zpoždění (korelační funkce vs. ohodnocení součtů posunutých signálů). Další odlišnost spočívá v tom, že TDOA ve své konečné fázi vrací pozici zdroje zvuku vůči zvolenému souřadnicovému systému, oproti tomu DAS takovou fázi neobsahuje a končí jen určením zpoždění, ze kterého lze nanejvýš určit směr přicházejícího zvuku (tzn. úhel dopadu). Avšak vzhledem k tomu, že jediné, co je nutné k tomu, aby šlo určit souřadnice zdroje pomocí poslední fáze TDOA, je zpoždění signálů a známá geometrie prostoru a mikrofونů, lze tyto kroky provést i pro zpoždění získané DAS metodou. Určení souřadnic zdroje v námi známém souřadnicovém systému je totiž způsob, jakým budeme ve většině experimentů ověřovat přesnost těchto metod.

První část testování se však bude zabývat jen ověřením, zda vypočtené hodnoty zpoždění odpovídají přibližné poloze zdroje. A to z toho důvodu, že ne vždy je v praxi nutné znát přesnou polohu zdroje. Také je vhodné ověřit, že první část algoritmů (tj. určení zpoždění signálů) vrací adekvátní výsledky.

### 3.1 Hardware

Pro testovací účely mi bylo od fakulty informačních technologií VUT v Brně zapůjčen hardware, který byl vytvořen firmou Audified.

Hlavním účelem tohoto hardware je poskytnout přísun synchronizovaných signálů z mikrofونů. To je zajištěno několika vstupními kartami, které umožňují připojit až 16 mikrofонů. Signály z mikrofонů jsou kartami zasílány do hlavní desky hardware. Na té se nachází čip ADSP se dvěma DSP jádry SHARC.

Dále se na hlavní desce nachází ARM jádro, na kterém běží zjednodušená verze Linux operačního systému. Další součástí hardwaru jsou dva AD převodníky, SD karta, ze které bootuje Linux, aj. Pro více informací se podívejte na stránku výrobce čipu ADSP<sup>3</sup>.

---

<sup>1</sup>TDOA - Time Difference of Arrival

<sup>2</sup>DAS - Delay and Sum

<sup>3</sup>ADSP čip - informace o produktu: <https://www.analog.com/en/products/adsp-sc589.html#product-overview>

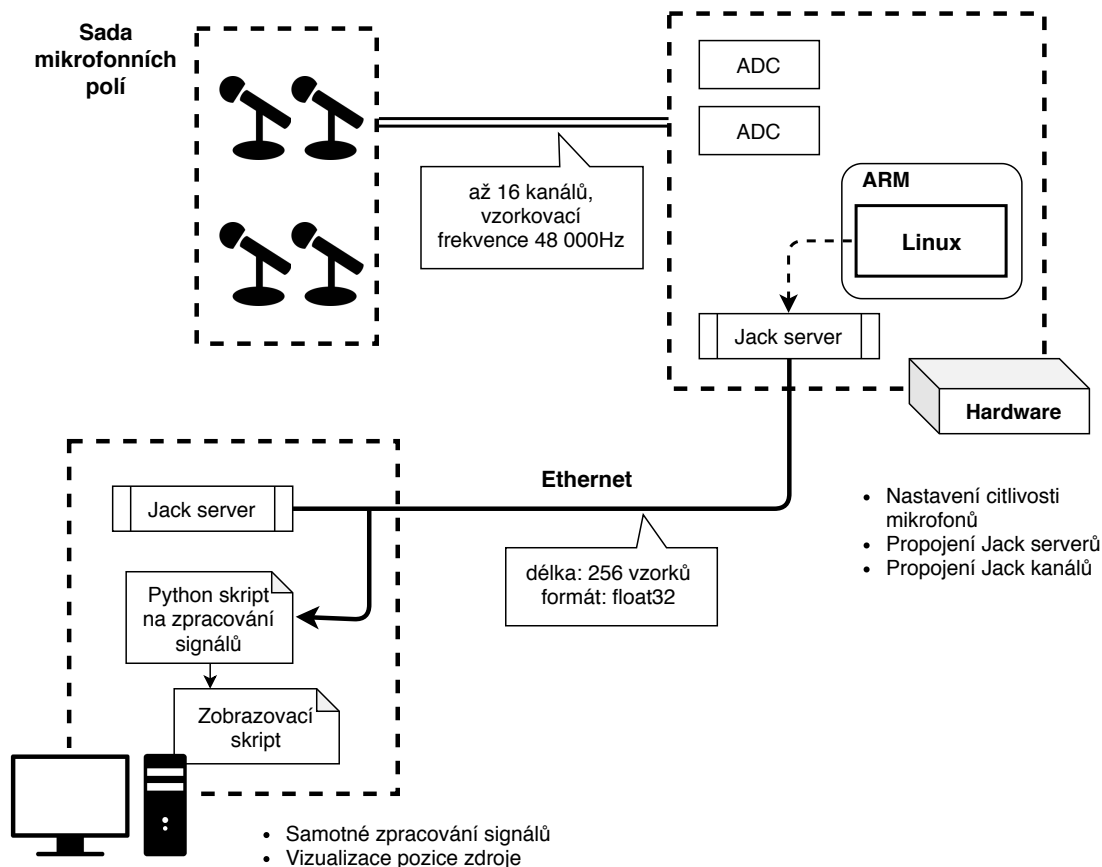


Obrázek 3.1: Hardware - rozhraní pro zapojení 16 mikrofonů

Součástí poskytnutého vybavení byly i mikrofony. Citlivost mikrofonů šlo nastavovat softwarově ve stupnici od 0 po 6 (odpovídající 0 až 60 *dB*).



Obrázek 3.2: Hardware - mikrofony



Obrázek 3.3: Schéma zapojení hardware k počítači se základní funkcionalitami

### 3.1.1 Jack

Hlavní mechanika přenosu signálů z hardware do počítače, kde jsou následně zpracovány, spočívá ve využití Jack API<sup>4</sup>. Linux, který běží na hardwaru, již Jack nainstalovaný má - konkrétně verzi `jack2`, která poskytuje vyšší propustnost oproti verzi `jack1` a také umožňuje připojení s Windows operačním systémem.

Aby z hardware mohly proudit signály do počítače bylo nutné připojit obě tato zařízení přes síťový kabel. Komunikace přes Jack funguje na základě architektury klient-server. Jak na počítači, tak na hardwaru musely tedy běžet jack servery, klient na hardware tyto servery propojil a klient na počítači byl vlastní program ke zpracování signálů.

Signály byly vzorkovány frekvencí 48 000 Hz a posílány po oddílech o 256 vzorcích.

## 3.2 Zpracování signálů

Samotné zpracování signálů bylo prováděno Python skriptem. Ten mimo záležitostí nutné pro funkčnost Jack rozhraní (tj. připojení se k severu, propojení vstupů a výstupů atd.) dále také provádí nad signály operace diktované metodou TDOA nebo DAS.

<sup>4</sup>JACK - Audio Connection Kit <http://www.jackaudio.org/>



Ke zpracování signálů byly použity mimo jiné následující knihovny:

- `numpy` - umožňuje efektivní práci s číselnými poli (v tomto případě signály)
- `scipy.optimize` - konkrétně pro funkci zajišťující výpočet průsečíku hyperbolických funkcí
- `multiprocessing` - výpočet zpoždění probíhá v novém vlákně, tím se neblokuje vstup nových vzorků signálů

Výpočet zpoždění signálů využívalo mechaniku plovoucího okna. Samotný výpočet byl prováděn nad bufferem vzorků o velikosti 48 000 vzorků (odpovídá jedné vteřině zvukového záznamu), přičemž 11904 (přibližně 0,25 sekund) vzorků z toho tvořilo překryv s předchozím bufferem. Zpoždění tedy bylo zjišťováno přibližně každých 0,75 sekund (36096 vzorků – tj.  $141 \times 256$ ), přičemž se počítalo ze vzorků celé poslední vteřiny.

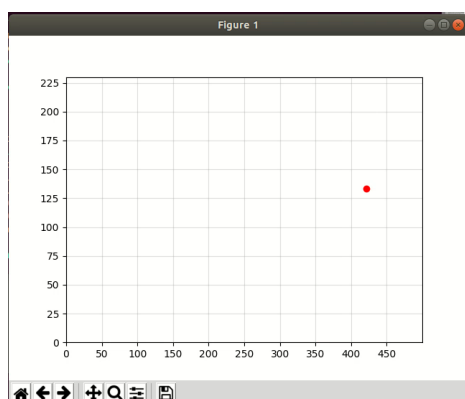
Délka plovoucího okna byla zvolena s ohledem na to, že okno nesmí být příliš dlouhé, kvůli narůstajícímu množství vzorků, s kterým je pak nutné počítat, ale nesmí být ani příliš malé. První testování jsem prováděla nad menším plovoucím oknem – o velikosti 20 000 vzorků, výpočet probíhal přibližně každých 0,3 s. Ukázalo se, že metody v takovém případě nejsou moc stabilní a vypočtené hodnoty značně kolísaly od očekávaných hodnot. Také jsem usoudila, že není nutné zobrazovat pozici zdroje tak často jako je 0,3 s, zcela dostačující je změnu zobrazovat s menší frekvencí.

Zpracování ticha se ukázalo být problematické. Vzhledem k tomu, že výsledné zpoždění je zcela v režii jednotlivých operací prováděných jako součást TDOA nebo DAS metody, v případě velmi slabého signálu, který tedy značí ticho, je chování těchto metod nepředvídatelné. Proto jsem se během testování snažila vyhnout se pauzám s tichem.

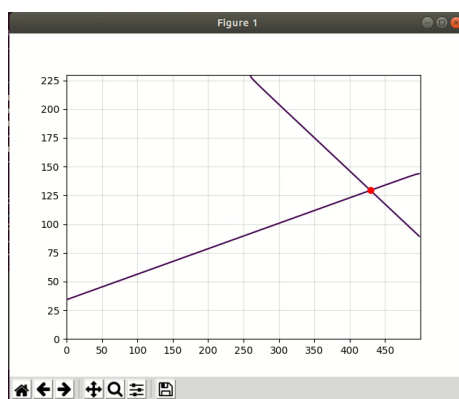
Výstupem skriptu jsou souřadnice zdroje zvuku v daném souřadnicovém systému, respektive hodnota zpoždění ve vzorcích, pokud se testuje pouze výpočet první fáze algoritmů. Tento výstup lze zpracovat dalším Python skriptem, který vykreslí vypočtenou polohu v grafu znázorňujícím prohledávaný prostor.

Důvodem, proč je vykreslení prováděné druhým skriptem, je ten, aby se zabránilo tomu, že vykreslování zablokuje skript a tím dojde k desynchronizaci vstupních dat. Ze stejného důvodu je součástí prvního skriptu i modul `multiprocessing`, který veškeré výpočty směruje na vedlejší proces, který nebude omezovat přísun nových dat z mikrofonů.

Vizuální výstup zobrazovacího skriptu lze vidět na obrázcích 3.4a a 3.4b (výstup odpovídá geometrii testovacího prostředí z následující sekce 3.3). Na obrázku 3.4b, lze vidět, jak funguje lokalizace TDOA pomocí hyperbol. Fialové linie na grafu totiž představují úsek hyperbol, které jsou oříznuté jen na prohledávaný prostor. Avšak na obrázku 3.5, kde je graf rozšířen na dvojnásobný rozsah, lze vidět celý tvar hyperbol. Ze stejného obrázku lze i vypořadovat efekt předozadní nejednoznačnosti zmíněné v sekci 2.3. Také si lze všimnout, že střed hyperbol odpovídá pozici, kde se nachází mikrofony.

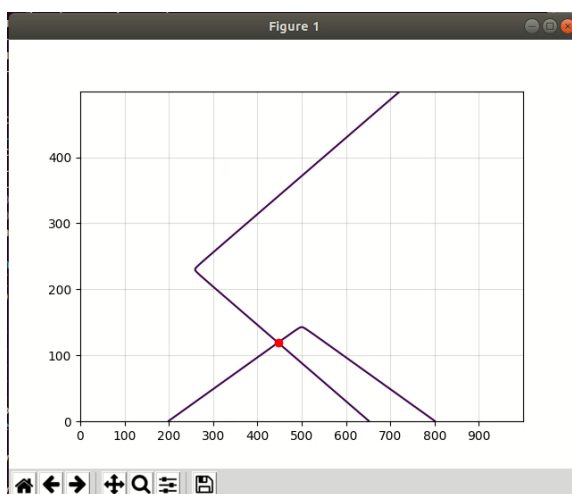


(a) Výstup zobrazující polohu jen bodem



(b) Výstup zobrazující polohu průsečíkem hyperbol

Obrázek 3.4: Výstup zobrazovacího python skriptu

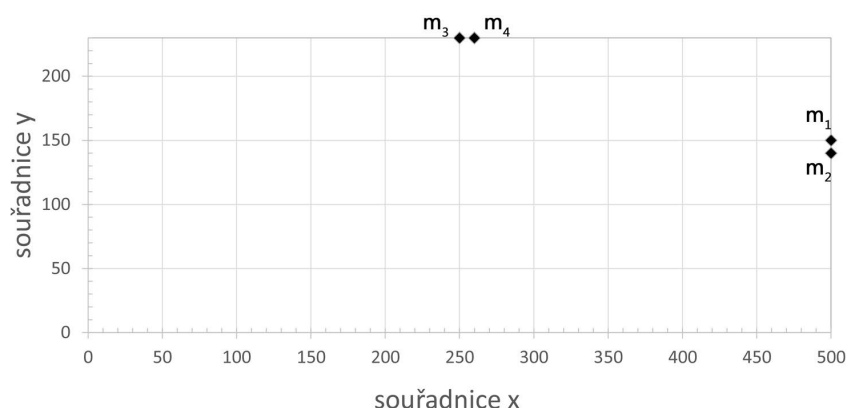


Obrázek 3.5: Celý tvar hyperbol pro určení pozice zdroje v grafu s dvojnásobným rozsahem

### 3.3 Testovací prostředí

Testování probíhalo v uzavřené místnosti o velikosti  $5 \times 2.3$  metrů.

Byly využity 4 mikrofony rozdělené do párů tak, aby tvořily dvě separátní mikrofonní pole. Mikrofony byly rozmístěny přibližně doprostřed jednotlivých stěn, tak aby mohly pokrýt celých  $180^\circ$  bez uvažování předozadní nejednoznačnosti. Nákres prohledávaného prostoru lze vidět na obrázku 3.6, jednotky v obrázku odpovídají centimetrům v realitě v poměru 1:1.



Obrázek 3.6: Geometrie prohledávaného prostoru s pozicemi mikrofonů

Rozestup mezi mikrofony činil 10 centimetrů. Vzdálenosti mikrofonů hraje důležitou roli nejen při určování zpoždění, ale také pro přesnost lokalizace zdroje. Vzdálenost mikrofonů má vedlejší efekt na to, jakých hodnot může zpoždění nabývat. To je dáno tím, že při výpočtu úhlu dopadu zvuku se používá goniometrická funkce arcsin, jejíž definiční obor je omezen na interval  $\langle -1; 1 \rangle$ , toto má za následek fakt, že i hodnota zpoždění ve vzorcích je omezena na specifický interval. Šířka tohoto intervalu je závislá na vzdálenosti mikrofonů, pro hodnoty mimo tento interval není lokalizace zdroje možná. Znalost tohoto intervalu lze využít i při výpočtu zpoždění. Například korelace se může počítat pouze v rozmezí tohoto intervalu – tímto lze ušetřit mnoho času. Pro konkrétní polohy mikrofonů mohlo zpoždění nabývat pouze hodnot z intervalu  $\langle -13; 13 \rangle$ .

Tento interval zpoždění má taktéž vliv na přesnost s jakou lze lokalizovat zdroj. Neboť ze zpoždění se určuje úhel dopadu, musí se  $180^\circ$  rozdělit mezi interval hodnot zpoždění. Více o této problematice bylo zmíněno v podsekcí 2.4 zabývající se problematikou určování úhlu dopadu ze zpoždění.

Výška mikrofonů odpovídala předpokládané výšce zdroje zvuku, tak aby se mikrofony a zdroj nacházely ve stejné rovině. Vzhledem k tomu, že se předpokládá lokalizace řečníka, jsem usoudila, že není třeba zahrnout vertikální souřadnici, a protože výška, ze které bude zvuk přicházet, by se tedy za normálních okolností měnit neměla.

Další důvod pro tuto volbu je však více pragmatický. Bylo důležité zajistit, aby mikrofony a zdroj byly ve stejné rovině, která je zároveň i rovnoběžná s podlahou místnosti, především protože souřadnicový systém s výslednou lokalizovanou polohou se nachází v rovině s mikrofony a zdrojem zvuku. A pro možnost porovnání skutečné pozice řečníka s výslednou vypočtenou pozicí je nutné, abychom měli souřadný systém shodný s geometrií místnosti. V případě, že by mikrofony se zdrojem tvořili rovinu, která by však byla šikmá vzhledem k podlaze místnosti, neodpovídaly by vypočtené souřadnice vzdálenostem ve skutečné místnosti a vyhodnocení úspěšnosti by vyžadovalo buď složitý přepočítání z jednoho souřadného systému do druhého, nebo úplně jinou metriku úspěšnosti.

### 3.4 Průběh testování

Z důvodu zachování konzistence a omezení zkreslení výsledků, byla zdrojem zvuku ve všech případech stejná hlasová nahrávka (kromě testování v sekci 4.3.1). Každý experiment trval po určitou dobu (většinou přibližně jednu minutu, nebo 30 vteřin) – během této doby došlo

mnohokrát k odhadu pozice zdroje, to je dáno tím, že výpočet zpoždění signálů je prováděno každých 0,75 sekundy. Experimenty byly dělány tak, aby proběhl výpočet pozice zdroje aspoň  $50\times$  nebo  $100\times$ . Takto získaná sada hodnot byla zpracována a vyhodnocena.

Vzhledem k tomu, že výsledky testování jsou ve formě sady naměřených hodnot, nejvhodnější metrikou je nějaká statistická veličina. U většiny případů se bude jednat o výpočet průměrné odchylky od očekávané pozice. Někdy je také zahrnutá hodnota modu a rozptylu.

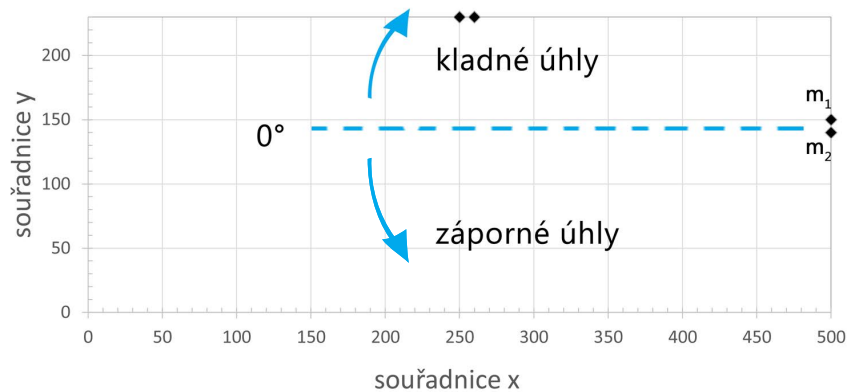
### 3.5 Testování určování úhlů ze zpoždění

Jeden okruh testování se zaměří na schopnost metod určit úhel dopadu signálu. Tímto se zároveň i ověří jak přesně dokáží algoritmy určit zpoždění mezi dvěma signály. Cílem je vyhodnotit přesnost zjištěných úhlů oproti skutečnému úhlu, dále také zjistit zda vzdálenost od mikrofonů má vliv na tuto přesnost či nikoliv.

Určování úhlů testuji především, protože v kapitole 5 o budoucím rozšíření se zmiňuji o možné aplikaci samotného určování úhlu v praxi.

Testování probíhalo ve stejné místnosti, která je popsána v sekci 3.3, byly však využity jen mikrofony  $m_1$  a  $m_2$  (viz. obrázek 3.7).

Vypočítaný úhel nabývá hodnot z intervalu  $\langle -90^\circ; 90^\circ \rangle$  s tím, že záporné úhly jsou směrem k mikrofonu  $m_2$  a kladné úhly k mikrofonu  $m_1$ , jak je zobrazeno na obrázku 3.7.



Obrázek 3.7: Orientace úhlů

<b>zpoždění</b>	<b>úhel [°]</b>
<b>-13</b>	-69.70
<b>-12</b>	-59.97
<b>-11</b>	-52.52
<b>-10</b>	-46.18
⋮	⋮
<b>-2</b>	-8.3
<b>-1</b>	-4.13
<b>0</b>	0
<b>1</b>	4.13
<b>2</b>	8.3
⋮	⋮
<b>11</b>	52.52
<b>12</b>	59.97
<b>13</b>	69.70

Tabulka 3.1: Tabulka hodnot zpoždění a jim odpovídající úhly získané aplikací vzorců 2.2 a 2.3

Tabulka 3.1 obsahuje hodnoty zpoždění, které jsou platné pro zvolený rozestup mikrofonů (tj. 10 cm), které je specifikováno v sekci 3.3 o testovacím prostředí. Z tabulky vyplývá, že za daných podmínek testování, nelze určit úhly vyšší jak  $70^\circ$  od středu. To protože zpoždění nemůže nabývat vyšší hodnoty než 13.

I přes zmíněné nedostatky určování úhlů ze zpoždění, si stále myslím, že má smysl tímto zabývat. To protože omezenou přesnost lze snížit například navýšením vzorkovací frekvence. Pokud zpoždění může nabývat většího intervalu hodnot, tak je možné rozlišovat jemnější úhly.

A vzhledem k tomu, že určení zpoždění nezahrnuje složitý výpočet průsečíku hyperbolických funkcí, lze určování úhlů použít v situacích, kdy je vyžadováno časté zjišťování pozice a počítání souřadnic by to příliš zdržovalo.

# Kapitola 4

## Výsledky testování

V této kapitole jsou zahrnuty výsledky testování metod TDOA<sup>1</sup> a DAS<sup>2</sup>. Obě tyto metody byly otestovány na přesnost pro určení úhlu dopadu signálu a souřadnic zdroje. Kapitola je zakončena porovnáním výsledků obou metod, v porovnání je zahrnut i scénář, kdy se v prohledávaném prostoru nacházely dva zdroje zvuku.

### 4.1 Metoda časových zpoždění

Tato sekce obsahuje výsledky testování metody TDOA. Podsekcce 4.1.1 se zabývá využitím pouze prvního kroku metody, tj. bez počítání hyperbolických funkcí, ostatní testování však hyperbolické funkce využívají.

#### 4.1.1 Určování úhlů ze zpoždění

První testování se zaměřilo na určování úhlů dopadu zvuku pomocí zjištěných dob zpoždění. K tomuto bylo zapotřebí jeden pár mikrofonů. Výpočet zpoždění ze signálů byla prováděna funkcí vzájemné korelace.

Proběhlo 11 dílčích experimentů, při nichž se zdroj zvuku nacházel v různých vzdálenostech od mikrofonů. Zdroj se po dobu jednoho měření nepřesouval. Každé měření se skládalo z 100 odhadů úhlu (tzn. proběhlo 100 kalkulací), tudíž doba jednoho měření byla něco přes jednu minutu.

Tabulky 4.1 obsahují statistické vyhodnocení jednotlivých měření. V první tabulce 4.1a s průměrem, rozptylem a směrodatnou odchylkou lze vidět, že průměr získaných úhlů se pohybuje většinou v rozmezí  $\pm 4^\circ$  od skutečného úhlu. U případů kdy je hodnota průměru vzdálené očekávané hodnotě (více než  $6^\circ$ ), to je způsobené tím, že některé spočítané úhly vychýlily zcela mimo ostatní získané úhly, což potvrzuje i značně vysoká hodnota rozptylu pro daný případ (např. pro test č. 6, 8 v tabulce 4.1a) v porovnání s ostatními případy. Tento trend lze vyzorovat i na grafech zahrnutých v příloze A.1, které zobrazují všechny naměřené hodnoty v rámci daného měření. Z tabulky lze také zaznamenat to, že se snižující se vzdáleností od mikrofonů se spočítané úhly více shlukují a nemají tendenci příliš se vychylovat od ostatních hodnot, což lze poznat z nízkých hodnot rozptylů u testů ve vzdálenosti 1 m od mikrofonů. Oproti tomu, u větších vzdáleností se objevují mnohem vyšší

---

<sup>1</sup>TDOA - Time Difference of Arrival

<sup>2</sup>DAS - Delay and Sum

hodnoty rozptylu. Lépe lze tento jev pozorovat na obrázcích v příloze, kde jsou zobrazeny konkrétní získané hodnoty.

	vzdálenost od mikrofonů [m]	$\varphi_s [^\circ]$	$\overline{\varphi_z}$	$\sigma^2$	$\sigma$
1	<b>5</b>	<b>-18</b>	-14.75	394.18	19.85
2		<b>0</b>	-1.78	155.31	12.46
3	<b>4</b>	<b>-18</b>	-11.5	8.93	2.99
4		<b>-9</b>	-4	46.6	6.83
5		<b>0</b>	-2.81	87.01	9.33
6		<b>12</b>	4.44	175.47	13.25
7	<b>3</b>	<b>-18</b>	-11.85	32.32	5.69
8		<b>-10</b>	-3.24	166.89	12.92
9		<b>0</b>	-2.16	65	8.06
10	<b>2</b>	<b>-18</b>	-12.18	27.6	5.25
11		<b>0</b>	0.5	74.97	8.66
12	<b>1</b>	<b>-35</b>	-31.73	8.90	2.99
13		<b>-18</b>	-19.15	9.14	3.03
14		<b>-5</b>	-3.80	47	6.86
15		<b>0</b>	1.74	5.2	2.28

(a) Tabulka s výsledky č. 1 - skutečný úhel ( $\varphi_s$ ), průměr ze získaných úhlů ( $\overline{\varphi_z}$ ), rozptyl ( $\sigma^2$ ), směrodatná odchylka ( $\sigma$ )

	vzdálenost od mikrofonů [m]	$\varphi_s [^\circ]$	modus $[^\circ]$	$\overline{\Delta\varphi} [^\circ]$	$\overline{\Delta\varphi}$ [cm]
1	<b>5</b>	<b>-18</b>	-16.77	8.44	77.65
2		<b>0</b>	0	5.32	46.43
3	<b>4</b>	<b>-18</b>	-12.5	6.5	45.38
4		<b>-9</b>	-4.13	5.14	35.88
5		<b>0</b>	0	4.99	34.84
6		<b>12</b>	8.3	9.75	68.07
7	<b>3</b>	<b>-18</b>	-12.5	6.71	35.13
8		<b>-10</b>	-8.3	16.03	83.93
9		<b>0</b>	0	3.85	20.16
10	<b>2</b>	<b>-18</b>	-12.5	6	20.94
11		<b>0</b>	0	2.74	9.56
12	<b>1</b>	<b>-35</b>	-30.33	3.64	6.35
13		<b>-18</b>	-21.15	2.78	4.85
14		<b>-5</b>	-4.13	1.90	3.32
15		<b>0</b>	0	1.99	3.47

(b) Tabulka s výsledky č. 2 - skutečný úhel ( $\varphi_s$ ), modus získaných úhlů (tj. nejčastěji získaný úhel), průměr odchylek od skutečného úhlu ( $\overline{\Delta\varphi}$ ) v jednotkách úhlů i délky (cm)

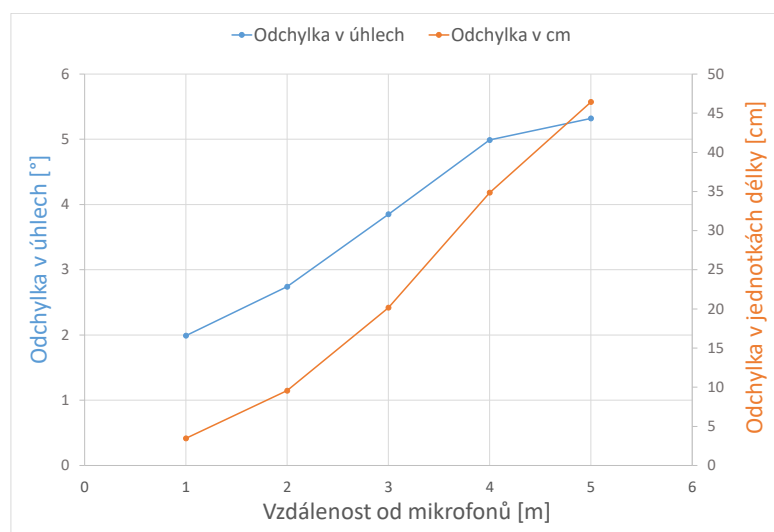
Tabulka 4.1: Výsledky pro určování úhlů pomocí metody TDOA

Druhá tabulka výsledků (4.1b) k průměrné odchylce ve stupních přiřazuje odpovídající délku v centimetrech (tato hodnota je spočítána zjednodušeně pomocí obvodu kružnice, je tudíž jen orientační). Porovnáním těchto dvou hodnot lze vidět, že ačkoliv se odchylka ve

stupních nezdá být velká (např. pouhých  $5^\circ$ ), její ekvivalent v jednotkách délky může být až půl metru rozdíl mezi skutečnou pozicí zdroje a zjištěnou pozicí (např. u testu č. 2).

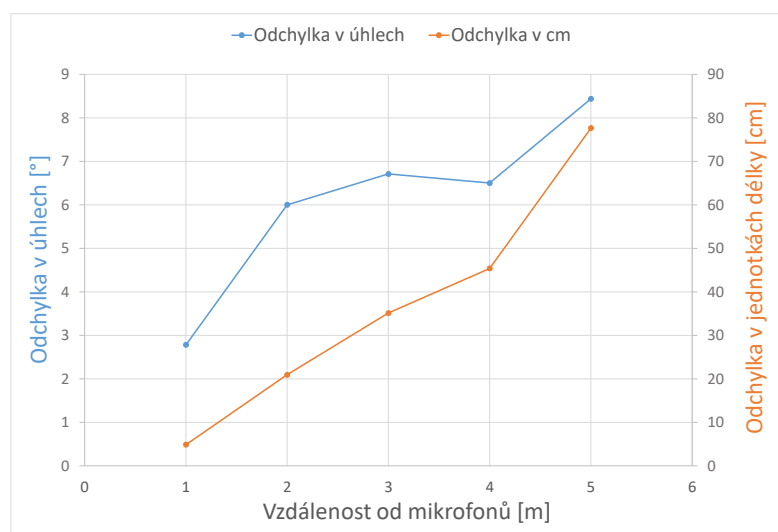
Absolutní chyba v jednotkách délky tedy stoupá implicitně se vzdáleností od mikrofonů a bohužel se s tím nedá nic dělat. Tohle lze sledovat i na grafech 4.1 a 4.2, kde vidíme, že hodnota odchyly v jednotkách délky má tendenci stoupat strměji než jí odpovídající ekvivalent v jednotkách úhlů. Každý graf ukazuje hodnoty odchylek při zjišťování úhlů  $0^\circ$  (4.1) a  $18^\circ$  (4.2) pro různé vzdálenosti od mikrofonů. Čím dále se zdroj nachází od mikrofonů tím se určování pozice pomocí úhlů stává nepřesnější. To lze poznat také z grafu na základě toho, že s rostoucí vzdáleností od mikrofonů se průměrná odchylka (zjištěných úhlů od skutečného úhlu) zvětšuje.

Lze tedy s jistotou tvrdit, že určení úhlů je omezeno nejen co se týče šířky prohledávaného prostoru (viz. teorie v sekci 2.4), ale i ve vzdálenosti od mikrofonů.



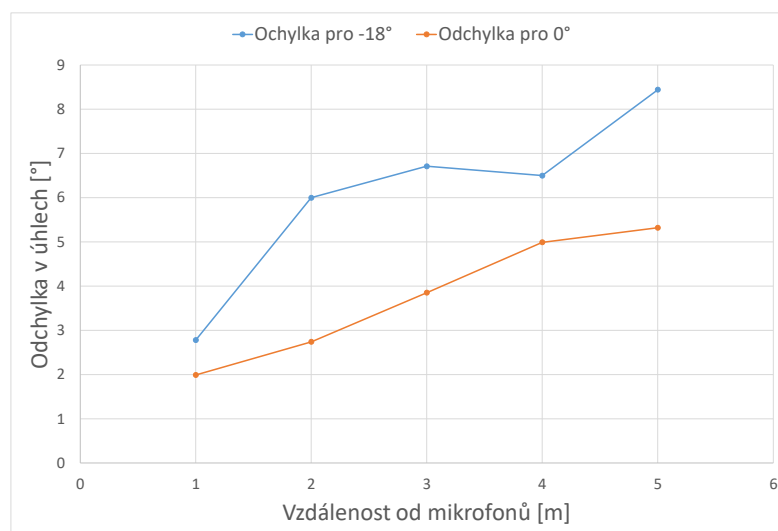
Obrázek 4.1: Průměrné odchylky od skutečného úhlu (ve stupních a centimetrech) z měření při úhlu  $0^\circ$  se vzrůstající vzdáleností od mikrofonů (testy č. 2, 5, 9, 11 a 15 v tabulce 4.1b)





Obrázek 4.2: Průměrné odchylky od skutečného úhlu (ve stupních a centimetrech) z měření při úhlu  $-18^\circ$  se vzrůstající vzdáleností od mikrofonů (testy č. 1, 3, 7, 10 a 13 v tabulce 4.1b)

Dále lze z výsledků obsažených v tabulce vyvodit fakt, že ve stejné vzdálenosti od mikrofonů je přesnější určení úhlu, který se nachází blíže středu mikrofonů (tj. kolem úhlu  $0^\circ$ ). Tohle lze lehce poznat i z grafu 4.3, který porovnává průměrné odchylky (ve stupních) získané z měření u úhlů  $0^\circ$  a  $-18^\circ$  pro různé vzdálenosti. Je vidět, že odchylky získané kolem úhlu  $0^\circ$  jsou vždy menší než odchylky u úhlu  $-18^\circ$ .



Obrázek 4.3: Graf porovnání průměrných odchylek od skutečné hodnoty pro testy u úhlů  $0^\circ$  a  $-18^\circ$

Ze získaných hodnot bych usoudila, že určování uhlů pomocí TDOA metody je použitelné a ještě poměrně přesné do vzdálenosti 3 metrů od mikrofonů. Do této vzdálenosti se odchylka pohybuje maximálně do 30 cm, s tím, že čím blíž k mikrofonům se zdroj nachází, tím je chyba lokalizace menší.

#### 4.1.2 Určení souřadnic zdroje

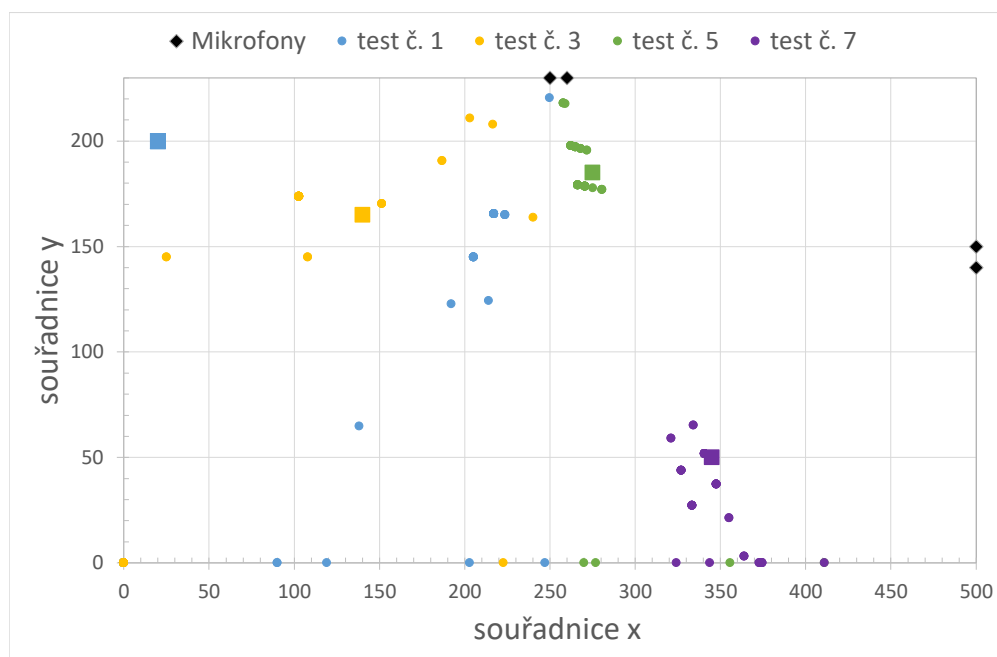
Proběhlo 8 dílčích testů, při nichž se zdroj zvuku nacházel v různých pozicích po dobu 30 vteřin. Během této doby došlo k 50 odhadům pozice zdroje.

Nutno podotknout, že pokud nějaká souřadnice nabývá hodnoty 0, jedná se většinou o případ, kdy vypočtená souřadnice přesahovala vymezený prostor a tudíž musela být fixně nastavena na 0. Takové hodnoty jsem brala jako nevalidní - počet nevalidních hodnot je zahrnut v tabulce výsledků ke každému měření.

Tabulka 4.2 obsahuje, pro každý provedený dílčí test, skutečnou pozici zdroje, průměrnou odchylku zjištěných pozic od skutečné, průměr ze zjištěných pozic, rozptyl a počet nevalidních pozic (tj. aspoň jedna souřadnice byla nulová). Je nutné podotknout, že pro výpočet statistických veličin (průměr, rozptyl atd.) v tabulce, byly nevalidní hodnoty ignorovány – neboť značí, že vypočtená pozice se nenachází v prohledávaném prostoru. Oproti tomu na obrázcích 4.4 a 4.5 jsou nevalidní hodnoty ukázány, protože z nich lze aspoň poznat jakým směrem je natočena aspoň jedna hyperbola.

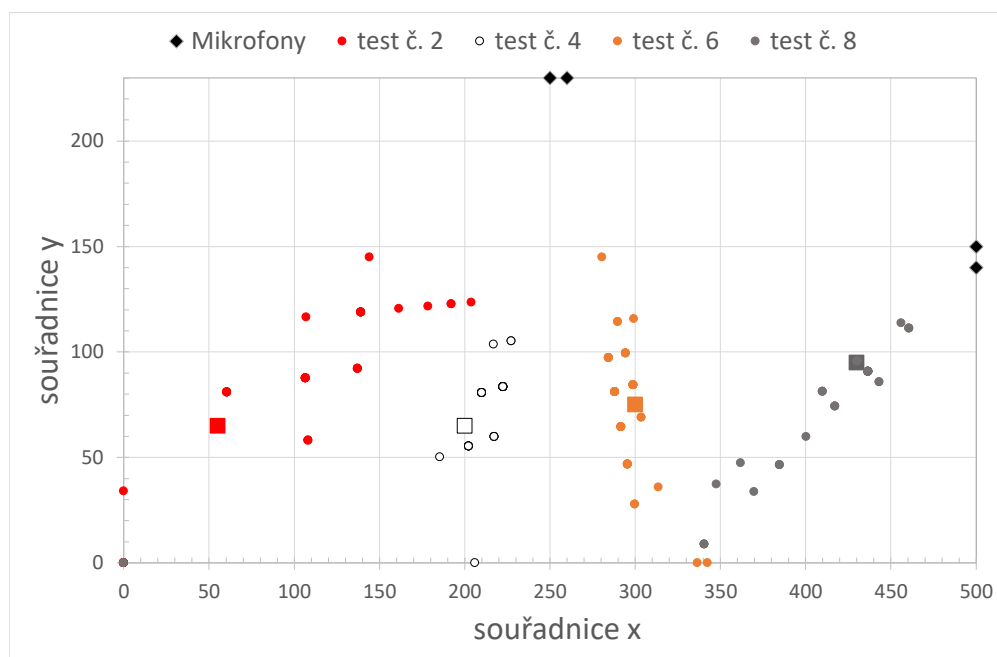
	skutečné		průměrná odchylka		průměr		rozptyl		nevalidní
	$x$	$y$	$x$	$y$	$x$	$y$	$x$	$y$	
1	<b>20</b>	<b>200</b>	193	44	213	157	268	539	18
2	<b>55</b>	<b>65</b>	65	33	120	97	1460	443	14
3	<b>140</b>	<b>165</b>	43	12	116	173	1791	172	15
4	<b>200</b>	<b>65</b>	14	16	213	77	68	161	1
5	<b>275</b>	<b>185</b>	8	11	267	189	27	156	4
6	<b>300</b>	<b>75</b>	8	20	293	78	40	576	2
7	<b>345</b>	<b>50</b>	9	34	340	41	104	227	8
8	<b>430</b>	<b>95</b>	20	18	420	80	965	671	1

Tabulka 4.2: Výsledky zjišťování souřadnic přes TDOA - tabulka se skutečnou pozicí zdroje, průměrnou odchylku získaných pozic od skutečné, průměr z získaných hodnot, rozptyl (pro výpočet jsou nevalidní hodnoty ignorovány) a počet nevalidních hodnot

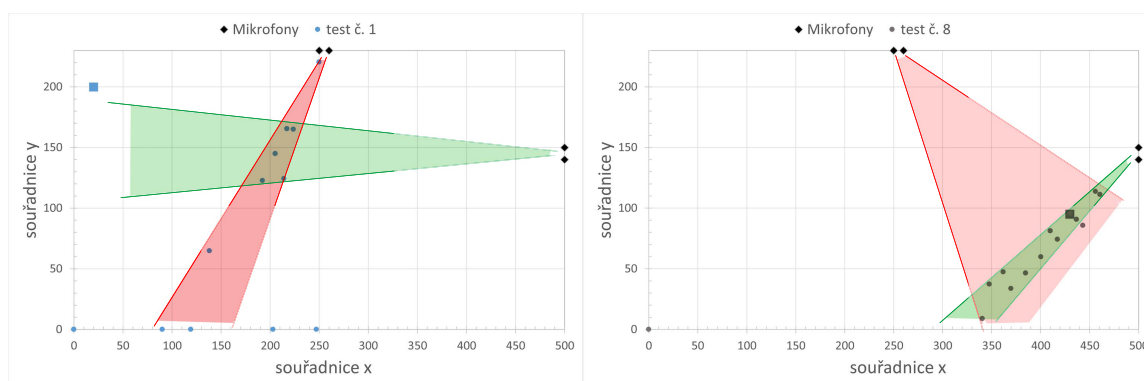


Obrázek 4.4: Spočítané pozice zdroje v testech č. 1, 3, 5 a 7, skutečná pozice je zobrazena čtvercem, zjištěné pozice kolečkem

Z testu č. 1 (viz. obr. 4.4) – tj. s nejvzdálenější pozicí zdroje od mikrofonů  $m_1$  a  $m_2$  (mikrofonní pole vpravo na obrázcích) a téměř u zdi s mikrofony  $m_3$  a  $m_4$  (mikrofonní pole nahoře) – lze vidět, že určení polohy zdroje naprosto selhává. Z tabulky však jde vidět, že pokud uvažujeme jen validní hodnoty souřadnice  $y$ , pak je průměrná odchylka pro  $y$  poměrně malá. Co nejspíš způsobuje to, že vypočítané pozice nejsou ani zdaleka tak blízko skutečné pozici je, že mikrofonní pole z mikrofonů  $m_3$  a  $m_4$  nedokáže správně určit, že se zdroj nachází téměř u zdi. Když si představíme hyperboly a kde se protínají, aby ve výsledku dávaly dané body, můžeme vidět, že hyperbola vycházející z mikrofonů  $m_1$  a  $m_2$  se v některých případech blížila skutečné poloze, avšak druhá hyperbola nedokázala skutečnou polohu vůbec určit a místo toho mířila zcela jinam. Proto jsou body roztroušeny daleko od očekávané hodnoty, ale celkově se udržují ve shluku u sebe. Konkrétně tohle vystihuje obrázek 4.6a.



Obrázek 4.5: Spočítané pozice zdroje v testech č. 2, 4, 6 a 8, skutečná pozice je zobrazena čtvercem, zjištěné pozice kolečkem



(a) Test č. 1 pro TDOA

(b) Test č. 8 pro TDOA

Obrázek 4.6: TDOA - ukázky orientací hyperbol pro určení pozice zdroje

Je dost možné, že pro pro určení přes hyperbolické funkce platí stejná omezení, která platila pro určování úhlů, tj. omezená přesnost na krajích způsobené tím, že zpoždění může nabývat jen omezeného počtu hodnot. Nicméně, bez ohledu na to zda to platí či nikoliv, ze získaných dat lze rozhodnout, že určení pozice přes TDOA není přesné, pokud se zdroj nachází na okrajích  $180^\circ$  rozsahu nějakého mikrofonního pole.

Další jev, který má za následek nepřesnost výsledku je pozorovatelný u testů č. 6 a 8 (obr. 4.5), kde spočítané pozice tvoří shluk do tvaru linky. To je dáno tím, že jedna hyperbola

měla stabilnější charakter než hyperbola druhá. Na obrázku 4.6b lze vidět, že hyperbola vycházející z horního mikrofonního pole se kolísá v širší oblasti, kdežto druhá hyperbola vycházející z bočního mikrofonního pole má stabilnější orientaci směrem ke zdroji zvuku. Z toho bych usoudila, že čím blíže je zdroj zvuku k mikrofonnímu poli, tím je spočítaný tvar hyperboly stabilnější a bližší skutečné pozici zdroje zvuku.

Z provedených experimentů lze vidět, že TDOA je spolehlivé a přesné pro pozice, které se nachází blízko středů mikrofonních polí (myšleno blízko kolmice na střed mikrofonních polí) a zároveň jsou i blízko k mikrofonním polím – nejvíce to jde vidět na testech č. 4, 5 a 7.

Z tabulky 4.2 lze dále také vyčíst to, že čím dále od mikrofonního pole, tím větší je šance, že vypočítaná pozice je nevalidní – tj. nenachází se v prohledávaném prostoru. Taková situace nastává, když se hyperboly neprotínají a řešení nelze najít, v takovém případě, jsem souřadnici, která přesahovala přes okraj místnosti nastavila na hodnotu 0. Od 4. testu dál je vidět značný pokles v množství nevalidních hodnot, zjišťované pozice jsou totiž blíže k mikrofonním polím.

Ve výsledku bych řekla, že určení pozice přes TDOA metodu je použitelné jen v malých místnostech, kde prohledávaný prostor není příliš široký. Možným řešením by bylo použití více mikrofonních polí, kde by se vždy určovala pozice pomocí mikrofonního pole, které je mu nejbližší – to by se dalo zjistit tím, že se spočítá síla zachyceného signálu pro každé mikrofonní pole a vybere se to mikrofonní pole, pro které je signál nejsilnější. Pro to by ovšem nejspíš byla vhodnější metoda DAS, která zjišťuje sílu signálu jako součást výpočtů pro určení zpoždění mezi signály.

## 4.2 Metoda Delay and Sum

Ačkoliv je sumační vzorec využívaný metodou DAS časově náročnější v porovnání s funkcí vzájemné korelace, kterou využívá TDOA, v našem případě, kdy pracujeme pouze se dvěma signály z dvou mikrofónů je použití tohoto vzorce ještě poměrně dost efektivní. Proto jsem se rozhodla otestovat i tuto metodu.

### 4.2.1 Určování úhlů ze zpoždění

Testování probíhalo stejně jako tomu bylo v sekci 4.1.1 s metodou TDOA (tj. 11 testů, v každém proběhlo 100 odhadů), jediný rozdíl byl v použitém algoritmu pro výpočet zpoždění. Informace jako geometrie testovacího prostoru, použité vzorce pro výpočet úhlu i předpokládaná omezení přesnosti při určování okrajových úhlů, specifikovaná dříve (sekce 2.4) pro TDOA, jsou tudíž platná i pro toto testování. Zkoumané úhly byly taktéž stejné a to především, aby se mohly metody porovnat s udržením určité konzistence.

	vzdálenost od mikrofonů [m]	$\varphi_s [^\circ]$	$\overline{\varphi_z}$	$\sigma^2$	$\sigma$
1	<b>5</b>	<b>-18</b>	-5.21	247.28	15.73
2		<b>0</b>	-4.15	240.99	15.52
3	<b>4</b>	<b>-18</b>	-8.58	288.06	16.97
4		<b>-9</b>	1.3	108.95	10.44
5		<b>0</b>	-5.9	125.5	11.2
6		<b>12</b>	4.41	517.13	22.74
7	<b>3</b>	<b>-18</b>	-9.19	188.78	13.74
8		<b>-10</b>	-8.67	56.79	7.54
9		<b>0</b>	-3.95	164.56	12.83
10	<b>2</b>	<b>-18</b>	-14.16	196	14
11		<b>0</b>	-1.03	127.5	11.3
12	<b>1</b>	<b>-35</b>	-30	43.6	6.6
13		<b>-18</b>	-16.88	9.77	3.12
14		<b>-5</b>	-3.23	28.06	5.3
15		<b>0</b>	-1.6	90.1	9.5

(a) Tabulka s výsledky č. 1 - skutečný úhel ( $\varphi_s$ ), průměr ze získaných úhlů ( $\overline{\varphi_z}$ ), rozptyl ( $\sigma^2$ ), směrodatná odchylka ( $\sigma$ )

	vzdálenost od mikrofonů [m]	$\varphi_s [^\circ]$	modus $[^\circ]$	$\overline{\Delta\varphi} [^\circ]$	$\overline{\Delta\varphi} [cm]$
1	<b>5</b>	<b>-18</b>	-12.5	12.85	112.14
2		<b>0</b>	4.13	8.82	76.97
3	<b>4</b>	<b>-18</b>	-12.5	9.6	67.02
4		<b>-9</b>	-4.13	10.44	72.88
5		<b>0</b>	0	8.25	57.6
6		<b>12</b>	8.3	11.13	77.7
7	<b>3</b>	<b>-18</b>	-12.5	8.9	46.6
8		<b>-10</b>	-8.3	3.88	20.32
9		<b>0</b>	0	6.1	31.94
10	<b>2</b>	<b>-18</b>	-16.77	10.7	37.35
11		<b>0</b>	0	5.5	19.2
12	<b>1</b>	<b>-35</b>	-30.33	5.11	8.92
13		<b>-18</b>	-16.77	2.74	4.78
14		<b>-5</b>	-4.13	4.55	7.94
15		<b>0</b>	0	2.87	5.01

(b) Tabulka s výsledky č. 2 - skutečný úhel ( $\varphi_s$ ), modus získaných úhlů (tj. nejčastěji získaný úhel), průměr odchylek od skutečného úhlu ( $\overline{\Delta\varphi}$ ) v jednotkách úhlů i délky (cm)

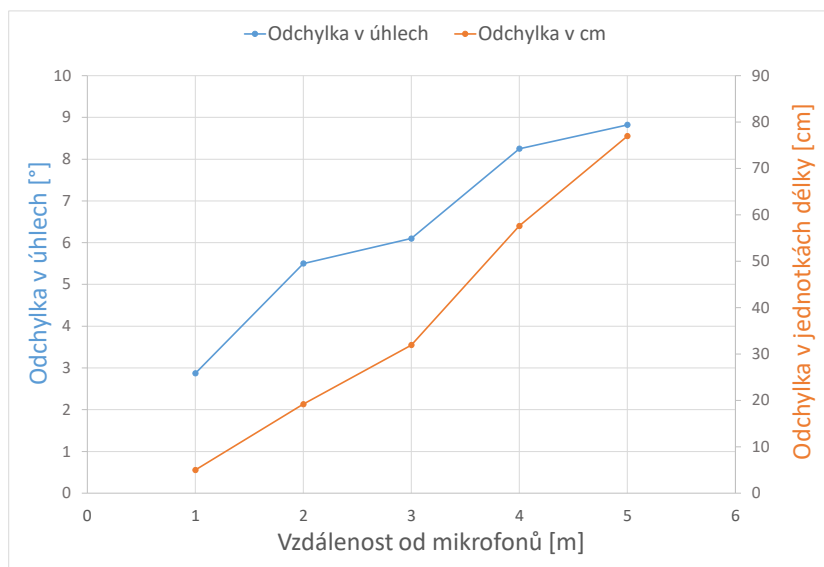
Tabulka 4.3: Tabulky výsledků pro určování úhlů pomocí metody DAS

Tabulka 4.3 obsahuje výsledky testování. V mnohém se závěry vyvozené z této tabulky shodují s těmi u metody TDOA.

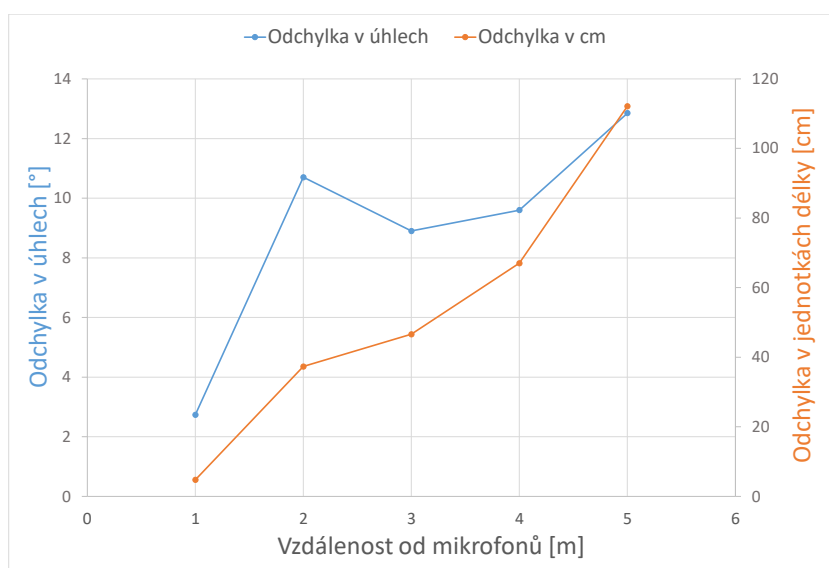
Mody se nachází blízko hledané hodnotě, což značí, že v získaných hodnotách je určitá konzistence a zároveň, že se shlukují u správné hodnoty. Průměr spočítaných úhlů a skutečného úhlu se ale ve vyšších vzdálenostech poměrně liší. To je způsobené častým kolísáním

spočítaných hodnot. Ačkoliv se většina zjištěných hodnot shlukovala u očekávané hodnoty, některé se velmi vychylovaly daného shluku. Lépe to jde vidět na grafech zahrnutých v příloze A.1. Zároveň to však lze pozorovat ze značně vysokých hodnot rozptylu u vyšších vzdáleností. Rozptyl klesá, čím blíže se zdroj nachází k mikrofonním polím. Algoritmus se stává stabilnější a hodnoty se tak často nevychylují od normy.

Grafy 4.7 a 4.8 zobrazují, jak se odchylka zvětšuje s narůstající vzdáleností od mikrofonů při zjišťování stejného úhlu. Přičemž můžeme vidět v grafu pro úhel  $-18^\circ$ , že odchylka v centimetrech se zvětšuje i když odchylka ve stupních se snížila u přechodu z 2 m na 3 m což jen značí, že se nevyplatí určovat úhly pro větší vzdálenosti, neboť odchylka v jednotkách délky roste implicitně.



Obrázek 4.7: Průměrné odchylky od skutečného úhlu (ve stupních a centimetrech) z měření při úhlu  $0^\circ$  se vzrůstající vzdáleností od mikrofonů (testy č. 2, 5, 9, 11 a 15 v tabulce 4.3b)



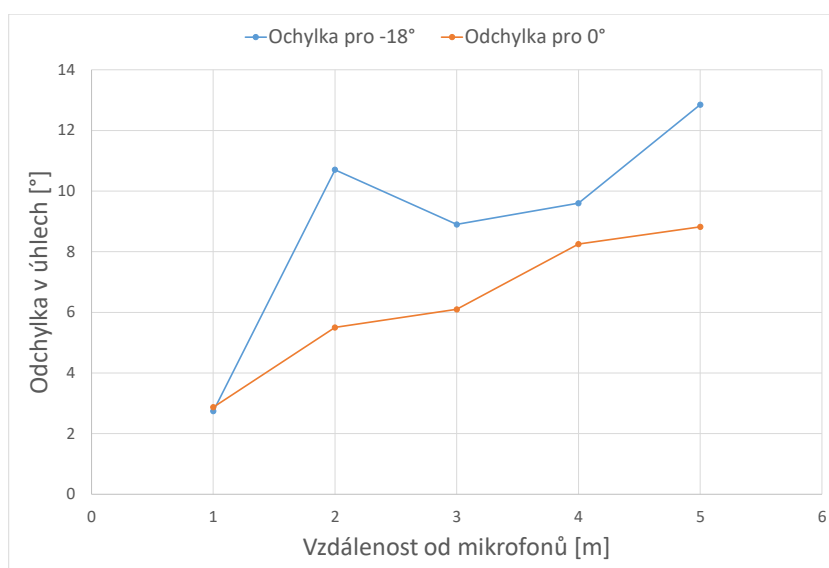
Obrázek 4.8: Průměrné odchylky od skutečného úhlu (ve stupních a centimetrech) z měření při úhlu  $-18^\circ$  se vzrůstající vzdáleností od mikrofonů (testy č. 1, 3, 7, 10 a 13 v tabulce 4.3b)

Dále si můžeme všimnout, že pro úhel  $0^\circ$  (obr. 4.7) roste odchylka lineárněji než pro úhel  $-18^\circ$  (obr. 4.8). Pro úhel  $-18^\circ$  se u 2 m se objevuje lokální maximum grafu. Nutno zmínit, že proběhlo několik opakování testů pro vzdálenost 2 m pro ověření, že daný vzrůst hodnoty odchylky nebyl způsoben nějakou chybou měření, ale z 6 dalších pokusů, všechny vykazovaly podobné chování, a to, že odchylka byla vyšší než ve vzdálenostech 3 m a 4 m (pro tyto vzdálenosti byly taktéž opakovány testy několikrát, výsledky se příliš nelišily).

Je možné, že ve vzdálenostech 3 m a 4 m pro úhel  $-18^\circ$  byl algoritmus přesnější, protože se daná poloha nacházela blízko zdi místnosti. Vzhledem k tomu, že DAS určuje úhel přicházejícího signálu na základě toho, kde se nachází nejsilnější signál, je možné, že signál byl zesílen ozvěnou.

Stejně jako tomu bylo u metody TDOA, můžeme vidět v tabulce 4.3b a následně v grafu 4.9, že odchylka pro měření při vyšších úhlech je obecně vyšší než u  $0^\circ$ .





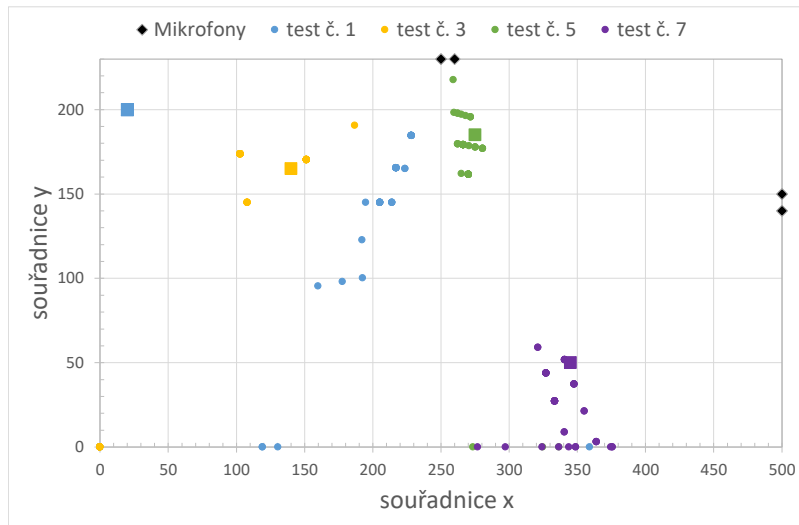
Obrázek 4.9: Graf porovnání průměrných odchylek od skutečné hodnoty pro testy u uhlů 0° a -18°

#### 4.2.2 Určení souřadnic zdroje

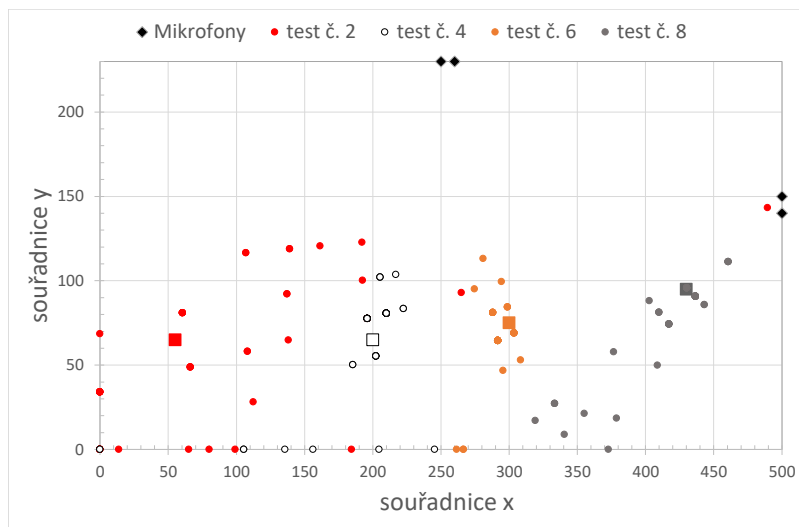
Testování probíhalo stejně jako u metody TDOA, taktéž se pro určení pozic používá stejných hyperbolických vzorců, s tím rozdílem, že spočtené zpoždění je získáno jinou metodou.

	skutečné		průměrná odchylka		průměr		rozptyl		nevalidní
	$x$	$y$	$x$	$y$	$x$	$y$	$x$	$y$	
1	<b>20</b>	<b>200</b>	193	42	213	157	207	540	9
2	<b>55</b>	<b>65</b>	84	27	139	84	7726	761	17
3	<b>140</b>	<b>165</b>	25	9	126	169	637	91	9
4	<b>200</b>	<b>65</b>	7	16	204	77	68	171	8
5	<b>275</b>	<b>185</b>	7	12	269	177	26	168	3
6	<b>300</b>	<b>75</b>	8	10	293	70	40	135	3
7	<b>345</b>	<b>50</b>	12	18	339	33	134	246	10
8	<b>430</b>	<b>95</b>	23	19	417	78	1221	690	1

Tabulka 4.4: Výsledky zjišťování souřadnic přes DAS - tabulka se skutečnou pozicí zdroje, průměrnou odchylku získaných pozic od skutečné, průměr z získaných hodnot, rozptyl (pro výpočet jsou nevalidní hodnoty ignorovány) a počet nevalidních hodnot



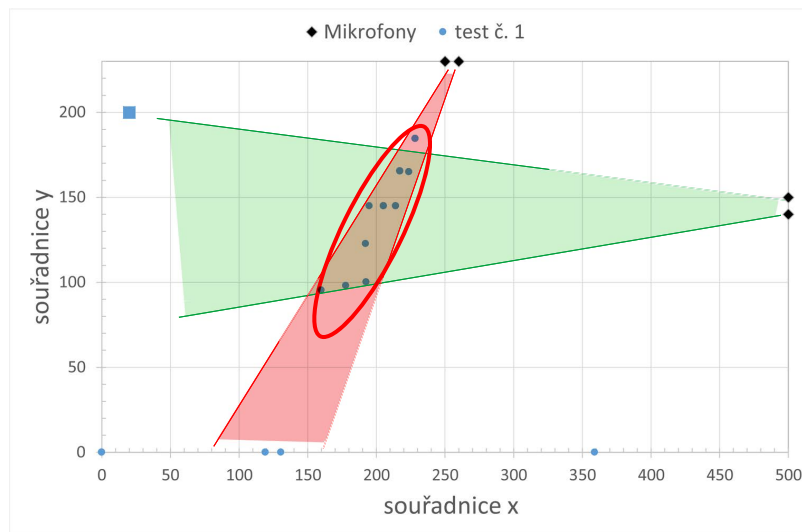
Obrázek 4.10: DAS - Spočítané pozice zdroje v testech č. 1, 3, 5 a 7, skutečná pozice je zobrazena čtvercem, zjištěné pozice kolečkem



Obrázek 4.11: DAS - Spočítané pozice zdroje v testech č. 2, 4, 6 a 8, skutečná pozice je zobrazena čtvercem, zjištěné pozice kolečkem

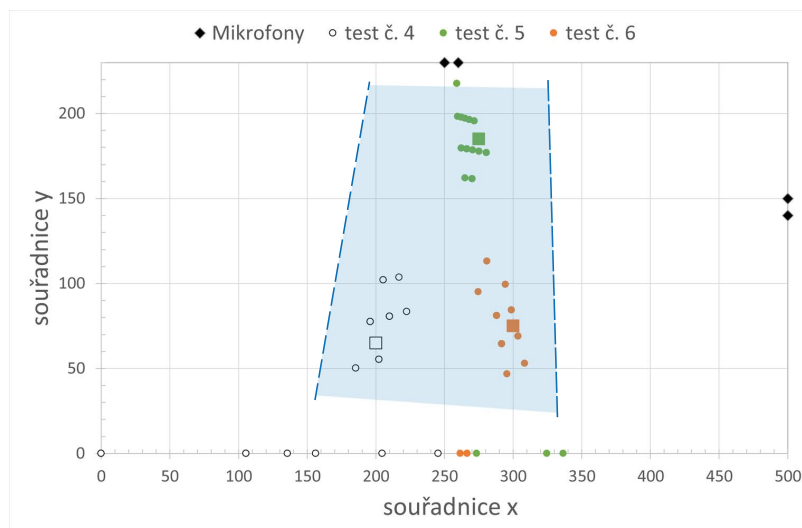
U nejobtížněji zjistitelné pozice (test č. 1) lokalizace opět selhává a příčina je stejná jako u TDOA, tj. neschopnost horního mikrofonního pole lokalizovat zdroj, který se nachází na okrajích jeho úhlového rozsahu. To lze poznat z toho, že všechny zjištěné pozice pro test

č. 1 jsou tvoří společně protáhlý útvar, tento útvar v podstatě kopíruje orientaci hyperboly horního mikrofonního pole (viz. obrázek 4.12).



Obrázek 4.12: Test č. 1 pro DAS - orientace hyperbol

Z hodnot rozptylů v tabulce můžeme soudit, že algoritmus je nejstabilnější a zároveň i nejpřesnější u testů č. 4, 5 a 6. Co tyto testy mají společného je, že se jedná o pozice, blízko ke komici k hornímu mikrofonnímu poli, zároveň se nachází přibližně mezi druhým a třetím metrem od bočního mikrofonního pole.



Obrázek 4.13: Oblast s nejlepšími výsledky pro metodu DAS

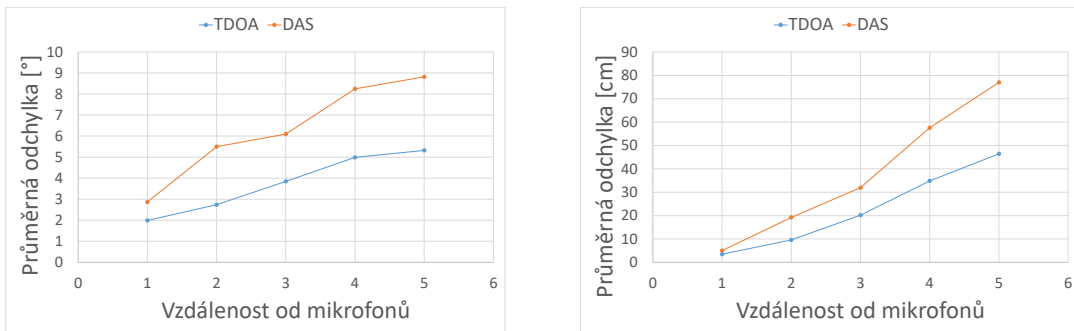
Z výsledků testů s pozicemi zdroje, které byly blíže bočnímu mikrofonnímu poli, můžeme vypořadovat, že jak rozptyl tak i průměrná odchylka se zvyšují (nejvíc je to zjevné u testu č. 8). Což je s největší pravděpodobností způsobené tím, že se to nachází příliš stranou od horního mikrofonního pole, a to je opět neschopno stabilně lokalizovat zdroj, který je na okrajích jeho úhlového rozsahu.

Další věcí, kterou si lze z výsledků odnést je fakt, že metoda DAS je použitelná i v kombinaci s hyperbolickými funkcemi, časová náročnost výpočtů přitom není na tolik vysoká, že by způsobovala problém pro lokalizaci.

### 4.3 Porovnání použitých metod

Tato sekce se zaměří na porovnání výsledků z testování obou metod. Jako první porovnáme výsledky získané z určování úhlů.

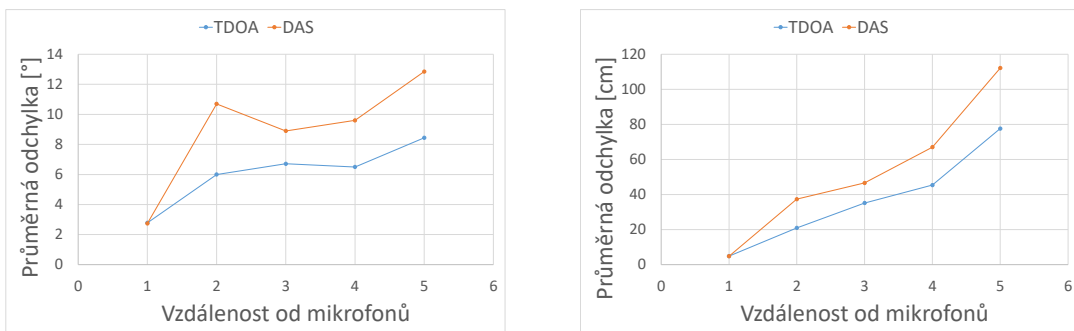
Na grafech 4.14 je porovnání průměrných odchylek získaných při úhlu  $0^\circ$ . Můžeme poznat, že TDOA metoda je výrazně přesnější než metoda DAS, a to většinou aspoň o  $2^\circ$ . Tyto odchylky převedeny orientačně na jednotku délky (cm) ukazují, že TDOA metoda je průměrně o 10 cm přesnější než DAS do vzdálenosti 3 metrů s tím, že na větších vzdálenostech se DAS metoda stává ještě více nepřesná. Pokud bychom brali odchylku 30 cm ještě jako přijatelnou, pak se lze spolehnout na obě metody do vzdálenosti 3 m s tím, že TDOA bude poskytovat lepší výsledky.



(a) Porovnání prům. odchylek ve stupních

(b) Porovnání prům. odchylek v centimetrech

Obrázek 4.14: Průměrné odchylky při určování úhlu  $0^\circ$  oběma metodami



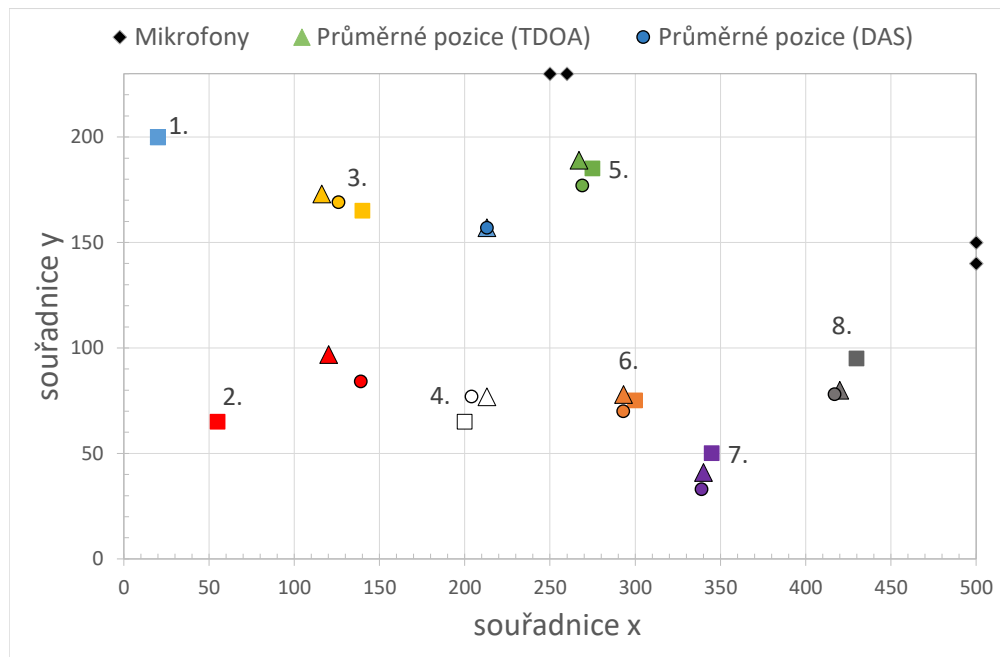
(a) Porovnání prům. odchylek ve stupních

(b) Porovnání prům. odchylek v centimetrech

Obrázek 4.15: Průměrné odchylky při určování úhlu  $-18^\circ$  oběma metodami

Pro úhly dále od kolmice na mikrofonní pole (jako příklad udávám výsledky z určování úhlu  $-18^\circ$ ) obecně platí, že průměrné odchylky jsou horší než u určování úhlů blíže ke kolmici na mikrofonní pole. Opět se ukazuje, že TDOA je přesnější než DAS. Obě metody ukazují výrazné zvětšení odchylek pro vzdálenosti vyšší než 4 m s tím, že do 3 m se průměrná odchylka pohybuje do maxima 30 cm.

Co se týče určování souřadnic zdroje přes obě metody, vzhledem k tomu, že DAS mělo poněkud horší výsledky při určování úhlů, jsem předpokládala, že i pro souřadnice tomu bude stejně. Z grafu však lze vidět, že průměry ze spočítaných pozic z obou metod se nachází velmi blízko u sebe.



Obrázek 4.16: Určování souřadnic - porovnání průměrných pozic z TDOA (znázorněny jako  $\Delta$  trojúhelníky) a DAS (znázorněny jako  $\circ$  kolečka), čísla značí pořadí testu

	průměr ze spočítaných pozic		rozptyl $x$		rozptyl $y$	
	TDOA	DAS	TDOA	DAS	TDOA	DAS
1	[213, 157]	[213, 157]	<u>268</u>	207	539	<u>540</u>
2	[120, 97]	[139, 84]	1460	<u>7726</u>	443	<u>761</u>
3	[116, 173]	[126, 169]	<u>1791</u>	637	<u>172</u>	91
4	[213, 77]	[204, 77]	68	68	161	<u>171</u>
5	[267, 189]	[269, 177]	<u>27</u>	26	156	<u>168</u>
6	[293, 78]	[293, 70]	40	40	<u>576</u>	135
7	[340, 41]	[339, 33]	104	<u>134</u>	217	<u>246</u>
8	[420, 80]	[417, 78]	965	<u>1221</u>	671	<u>690</u>

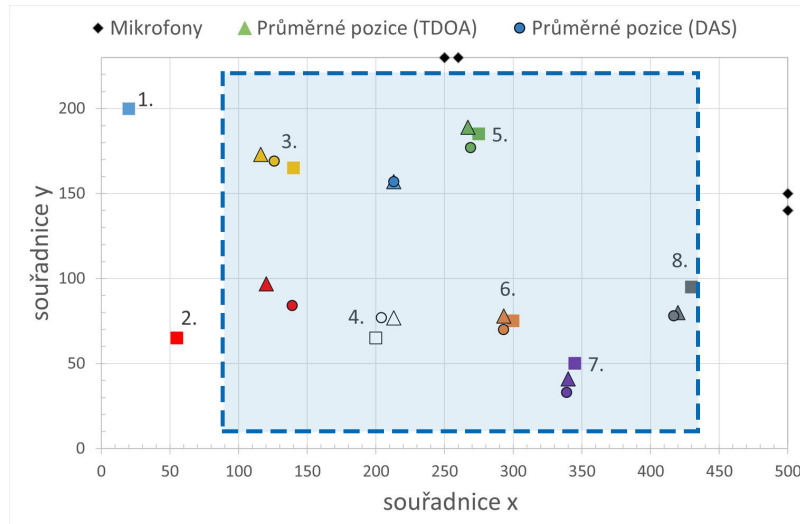
Tabulka 4.5: Tabulka porovnání hodnot výsledků metod, podtržené jsou hodnoty rozptylů, které jsou větší než rozptyl u druhé metody

Tabulka 4.5 shrnuje výsledky získané z testování metod. Pro porovnání jsou udány hodnoty průměrů ze získaných pozic a rozptyl pro každou souřadnici. Ve většině případů je rozptyl u metody DAS vyšší než rozptyl u TDOA, což značí, že DAS poskytuje méně stabilnější výsledky a více kolísá, ale z grafu 4.16 lze poznat, že průměrná pozice je poměrně blízko k hledané pozici.

Zdá se, že ačkoliv se DAS jeví méně přesná pro určování úhlů, pro určování konkrétních souřadnic se její výsledky v mnohém neliší od výsledků, které poskytuje metoda TDOA. Obě metody poskytovaly spolehlivé výsledky pro oblast znázorněnou na obrázku 4.17 (přibližně o velikosti  $3 \times 2$  m). To lze také poznat z tabulky 4.6, kde pro testy v dané oblasti (tj. především pro testy č. 4, 5, 6 a 7) jsou odchylky jen v řádech několika centimetrů. Ze stejné tabulky taky můžeme vidět, že hodnoty souřadnice  $x$  mají většinou vyšší odchylku než hodnoty souřadnic  $y$ . Vzhledem k tomu že souřadnice se získávají jako průsečík hyperbol, myslím si, že odchylky souřadnic  $x$  jsou způsobené nepřesností hyperboly vycházející z horního mikrofonního pole. Za to hyperbola vycházející z bočního mikrofonního pole bývá přesnější a proto jsou odchylky pro souřadnici  $y$  nižší. To je pravděpodobně způsobené tím, že hledané pozice se nachází většinou blízko kolmice na boční mikrofonní pole.

	odchylky $x$		odchylky $y$	
	TDOA	DAS	TDOA	DAS
1	193	193	44	42
2	65	84	33	27
3	43	25	12	9
4	14	7	16	16
5	8	7	11	12
6	8	8	20	10
7	9	12	34	18
8	20	23	18	19

Tabulka 4.6: Tabulka průměrných odchylek z testování obou metod



Obrázek 4.17: Oblast s nej přesnějšími výsledky pro obě metody

Ve výsledku jsou obě metody použitelné a spolehlivé v oblasti přibližně  $3 \times 3$  m s tím, že odchylky budou průměrně v řádech několika centimetrů. Oblast musí být však vycentrována tak, aby se nenacházela příliš bokem ke kolmicím na mikrofonní pole neboť u nich jsou výsledky nej přesnější.

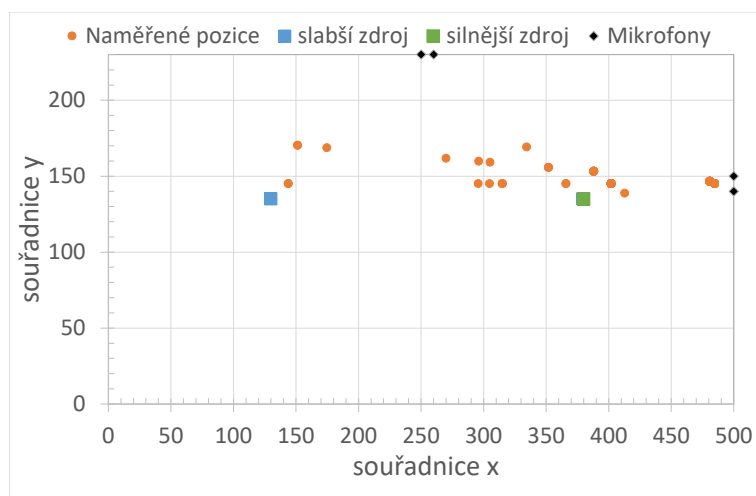
#### 4.3.1 Lokalizace zdroje v prostoru se dvěma zdroji zvuku

Další část testování se zaměřila na lokalizaci silnějšího zdroje zvuku, když se v prohledávaném prostoru nacházely dva zdroje zvuku. Zdroje zvuku představovaly dvě různé nahrávky řeči, zdroje byly postaveny tak, aby byly stejně vzdálené od horního mikrofonního pole, a přitom měly stejnou  $y$  souřadnici – tímto tedy testujeme schopnost jednoho mikrofonního pole (toho horního) identifikovat silnější zdroj.

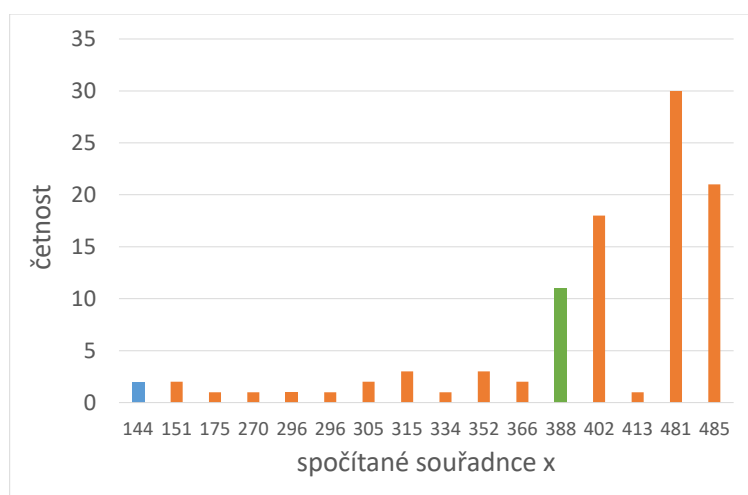
Nahrávky byly nastaveny tak, aby jedna víc hlasitá než druhá. Cílem bylo ověřit schopnost obou metod lokalizovat silnější zdroj. Při vyhodnocování zanedbáváme souřadnici  $y$  neboť oba zdroje by měli mít tuto souřadnici stejnou. Přesné pozice zdrojů jsou [130, 135] a [380, 135]. V grafech s četností je barevně vyznačena vypočtená souřadnice, která je nejbližší pozicím zdrojů.

Na obrázcích (4.18, 4.19) můžeme vidět výsledky testování daných metod. Z grafů zobrazujících všechny spočítané pozice (4.18a, 4.19a) lze vidět, že obě metody často ukazují směrem k silnějšímu zdroji a žádná z nich není vyloženě lepší než ta druhá. Grafy četností však ukazují, že ještě častěji než blízko k silnějšímu zdroji ukazují metody polohy na souřadnicích  $x = 480$ . Dá se předpokládat, že je to způsobeno skládáním zvukových signálů od obou zdrojů. A ačkoliv zjištěné pozice nejsou tak přesné ani tak blízko k očekávané hodnotě, stále se daří metodám identifikovat přibližný směr silnějšího signálu. Rozhodně se však nedá tvrdit, že lokalizace je přesná.

Testovaly se dva případy, rozdíl mezi nimi byl, že se pozice silnějšího a slabšího zdroje přehodily. Druhý test vynesl obdobné výsledky. V příloze A.2 jsou zahrnuty grafy popisující výsledky druhého testu.



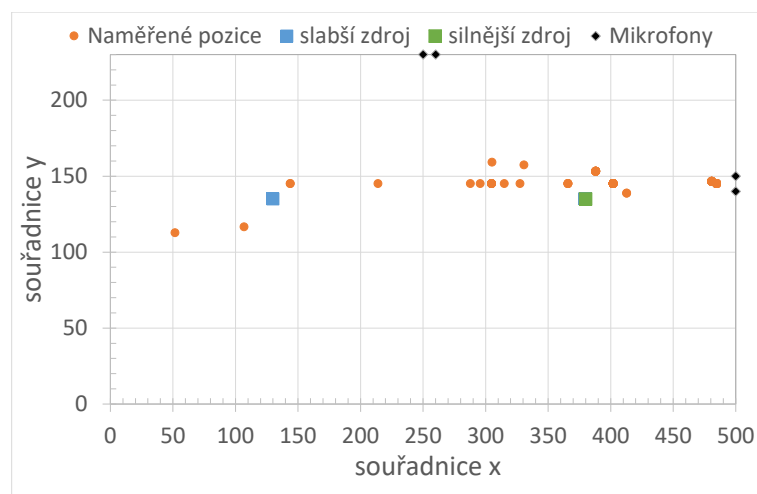
(a) TDOA - Graf se spočítanými pozicemi zdroje (oranžové) silnější zdroj (zelená), slabší zdroj (modrá)



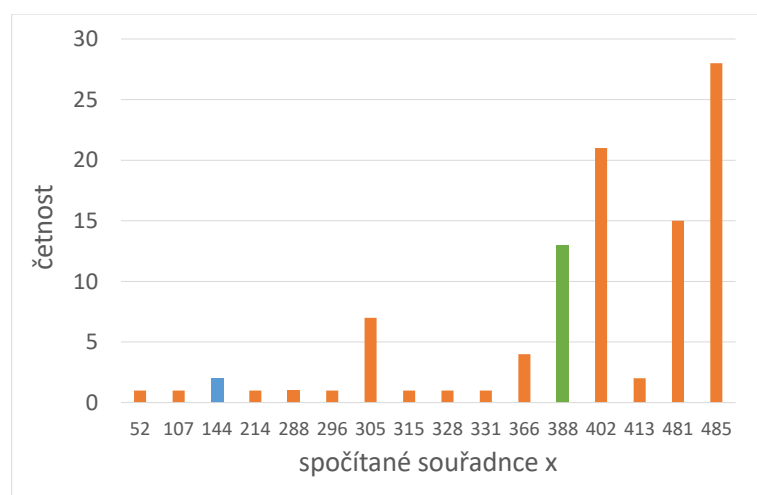
(b) TDOA - Četnost zjištěných souřadnic  $x$ , zeleně je vyznačena hodnota nejbližší k silnějšímu zdroji, modře hodnota nejbližší slabšímu zdroji

Obrázek 4.18: Test 1: TDOA - výsledky testu se dvěma zdroji





(a) DAS - Graf se spočítanými pozicemi zdroje (oranžové) silnější zdroj (zelená), slabší zdroj (modrá)



(b) DAS - Četnost zjištěných souřadnic  $x$ , zeleně je vyznačena hodnota nejbližší k silnějšímu zdroji, modře hodnota nejbližší slabšímu zdroji

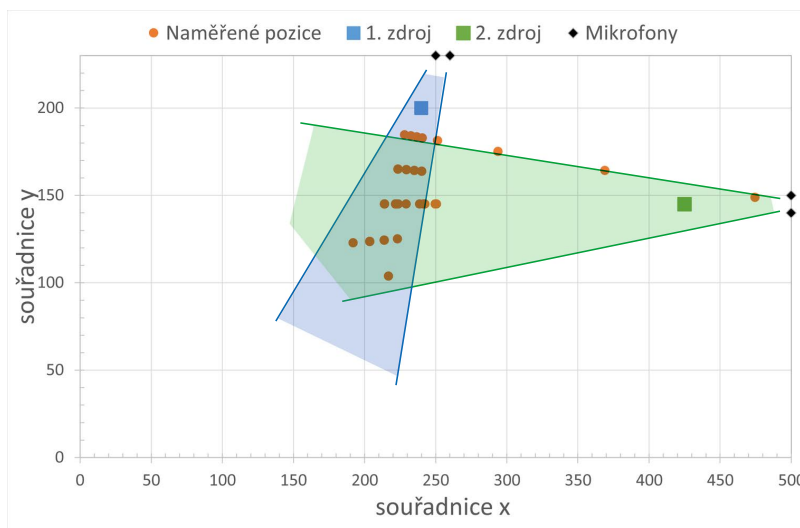
Obrázek 4.19: Test 1: DAS - výsledky testu se dvěma zdroji

Ačkoliv podle teorie TDOA není vhodnou metodou, když se v prohledávaném prostoru nachází více zdrojů zvuků, výsledky testování ukázaly, že TDOA není v tomto ohledu nijak horší než metoda DAS. V teorii by měla být DAS metoda vhodnější, protože se zaměřuje na směr odkud přichází silnější signál, ale v případě zpracovávání řeči se nedá zaručit, že během řeči nenastane odmlka, během které se bude mít slabší zdroj větší vliv na výsledek. A vzhledem k tomu, že se nedá zaručit aby hledaný zdroj zvuku měl v každém okamžiku

výpočtu vždy silnější vliv na mikrofonní pole, bych s troufala říci, že pro přesné určování pozice řečníka, když se v místnosti nachází i jiný zdroj zvuku, není ani jedna metoda vhodná.

Pokud se však požaduje, aby se určil jen přibližný směr, ze kterého přichází silnější zvuk, tak by na to měly být obě metody dostačující. Je však nutné, aby se ignorovaly náhlé jednorázové výchyly, které mohou nastat vlivem toho, že v jednu chvíli se slabší zdroj bude zdát silnější.

Dále je nutné poznamenat, že lokalizaci silnějšího zdroje má smysl uvažovat jen pokud obě mikrofonní pole dokáží rozpoznat stejný silnější zdroj. Pokud by se jeden zdroj nacházel blíže k jednomu mikrofonnímu poli a naopak druhý zdroj by byl blízko druhému mikrofonnímu pole, bude výsledná pozice zkreslená tím, že každé mikrofonní pole se bude zaměřovat na jiný zdroj a to ten, který je mu blíže. Toto lze vidět na obrázku 4.20, který zobrazuje výsledky z testu, kdy se dva stejně silné zdroje nacházely každý blíže jinému mikrofonnímu poli. Z obrázku lze vidět, že hyperbola z horního mikrofonního pole směřovala převážně k bližšímu zdroji (tj. zdroj č. 1), naopak hyperbola z bočního mikrofonního pole směřovala na zdroj druhý, protože ten byl zase blíž jejímu mikrofonnímu poli. To mělo za následek, že spočítané hodnoty ležely mimo oba zdroje.



Obrázek 4.20: Test se dvěma zdroji, každý je blíž jinému mikrofonnímu poli - Spočítané pozice

Lokalizace zdroje zvuku, když se v prohledávaném prostoru nachází víc než jeden zdroj zvuku není moc přesné, v nejlepších případech lze určit přibližný směr silnějšího zdroje. Dále bych řekla, že kvůli tomu, že se v lidské řeči objevují běžně odmlky a pauzy, nelze zaručit, že řečník bude vždy nejsilnějším zdrojem zvuku v místnosti. Dále je problém s tím, že pokud se bude uvažovat lokalizace pomocí více než jednoho mikrofonního pole, může nastat situace, kdy každé mikrofonní pole bude považovat jiný zdroj jako nejsilnější.

Ve výsledku bych tedy řekla, že lokalizace pomocí daných metod, pokud se v místnosti nachází více než jeden zdroj zvuku, není spolehlivá. Algoritmy se nedokáží vypořádat s problémy, které v takových situacích vyvstávají. Zdroj by musel být konstantně a o mnoho silnější než všechny ostatní zdroje zvuku, aby se dalo uvažovat o lokalizaci pomocí TDOA nebo DAS.

## Kapitola 5

# Budoucí postup a možná rozšíření

Metody byly otestovány a z výsledků mohu s určitostí tvrdit, že je ještě lze vylepšit pro lepší použitelnost v praxi.

### 5.1 Zpracování ticha

Jedním z problémů těchto metod je, že nedokáží detekovat ticho a zpracovávají ho normálně, což může mít za následek to, že vypočítaná poloha je nepředvídatelná.

Během testování jsem se takovým situacím mohla vyhnout, proto jsem se příliš nezaobírala jeho řešením tohoto problému. Avšak budou-li se metody používat v praxi, bude třeba tento problém vyřešit.

Existuje však řada algoritmů zabývajících se touto problematikou. Například různé způsoby prahování a filtrování.

### 5.2 Použití více mikrofonních polí

Testováním jsem zjistila, že metody pro lokalizaci zdroje zvuku jsou spolehlivé jen v určitém rozsahu – tj. když je zdroj blízko mikrofonnímu poli a blízko kolmici na střed mikrofonního pole.

Jedním ze způsobů, jak lze tento problém vyřešit, je zavedení více mikrofonních polí, tím se prohledávaný prostor lépe pokryje a měření bude přesnější. Je však nutné vyřešit koordinaci mezi mikrofony, tak by se vždy počítalo s hodnotou zpoždění z nevhodnějšího mikrofonního pole. To protože hyperbolické funkce předpokládají 2 mikrofonní pole.

Nejjednodušší by tedy bylo použít metodu DAS<sup>1</sup>, neboť ta už počítá sílu signálu pro určení zpoždění. Tyto hodnoty sil by šlo porovnat i mezi vícero mikrofonními poli a tak zvolit to nevhodnější a vůči němu určit úhel nebo pozici.

### 5.3 Osvětlení pozice zdroje reflektorem

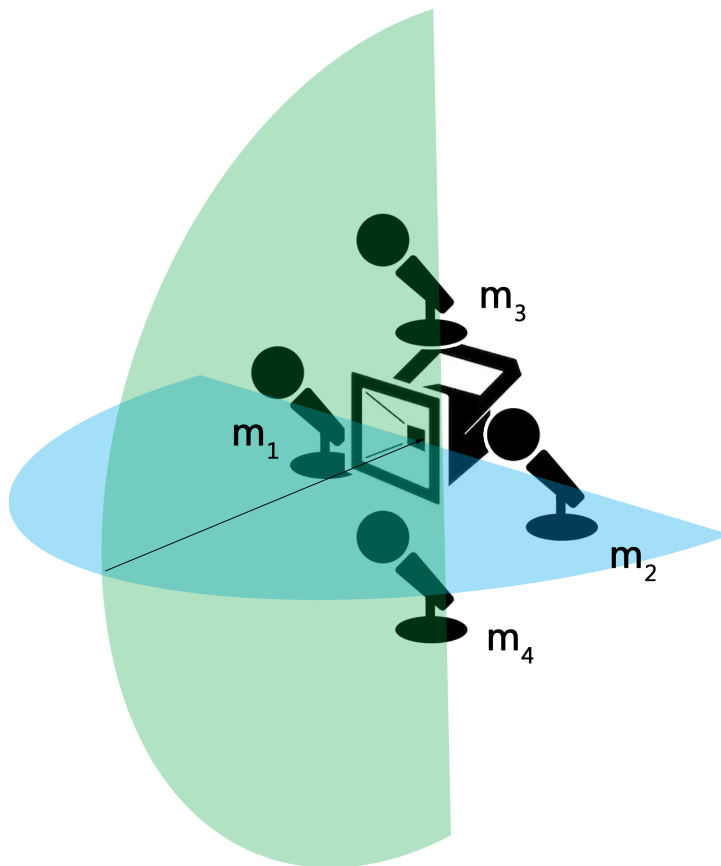
V této sekci je vysvětlení toho, jak bych pokračovala s prací za účelem docílit osvětlení pozice zdroje pomocí světelného reflektoru. Věřím, že by šlo by aplikovat několik zjednodušení oproti tomu, jak se metody testovali v předešlé části práce.

---

<sup>1</sup>DAS – Delay and Sum

Uvažujeme-li reflektor, který se bude moci otáčet vhodně do 4 směrů - nahoru, dolů, doprava a doleva, lze soudit, že k osvětlení místa zdroje zvuku stačí pouze úhel dopadu, nikoliv přesné souřadnice tohoto místa. Tím pádem odpadá nutnost používání hyperbolických funkcí z TDOA<sup>2</sup> metody a lze určit úhel jednoduchým aplikováním arcsin funkce (viz. testování v sekcích 4.1.1 a 4.2.1).

Bude-li měřicí aparát společně s reflektorem sestaven podle obrázku 5.1



Obrázek 5.1: Sestava s reflektorem a 4 mikrofony

Mikrofony  $m_1$  a  $m_2$  určí natočení reflektoru doprava (azimut) nebo doleva a mikrofony  $m_3$  a  $m_4$  zase natočení nahoru dolů (elevace). Určení azimutu a elevace by mělo být dostačující pro lokalizaci řečníka v prostoru.

Vzhledem k tomu, že není nutné určit souřadnice zdroje může se mikrofonní pole nacházet kdekoli nad řečníkem. Pro určení úhlu dopadu zvuku není nutné, aby mikrofony byly ve stejné výšce jako zdroj. Ideální pozice pro tuto sestavu by nejspíše byl horní roh místnosti, tak aby reflektor mohl možnost osvětlit všechny části místnosti. Zároveň kladu důraz, aby se využil roh místnosti, takto se plocha, ve které se určuje úhel omezí pouze kolem kolmice na mikrofony, takto se neprojeví problém zvětšujícího se rozdílů úhlu připadající na zpoždění 1 vzorku (zmíněné v podsekcí 2.4). Pokud se však využije horní roh, bylo by vhodné natočit osu mikrofonního pole pro určení elevace (tj.  $m_3$  a  $m_4$ ), tak aby kolmice na osu mikrofonního pole směřovala do místnosti dolů. Tohle bude mít za následek to, že se plně využije oblast, ve které je určování úhlu nejpřesnější (tj. okolo  $0^\circ$ ).

<sup>2</sup>TDOA – Time Difference of Arrival

## Kapitola 6

# Závěr

V několika posledních kapitolách jsem se zabývala zkoumáním, na kolik jsou dané algoritmy (DAS<sup>1</sup> a TDOA<sup>2</sup>) použitelné v kombinaci s dostupným hardwarem. Zároveň jsem v daných kapitolách porovnávala výsledky získané oběma použitými metodami.

Testování probíhalo v poměrně malém prostoru (3×5 m), převážně protože použité mikrofony nebyly dostatečně citlivé pro větší vzdálenosti. Ukázalo se však, že i takto zvolený prostor byl dostačující pro prokázání podstatných vlastností metod. Z dosažených výsledků lze tvrdit, že obě použité metody jsou schopné lokalizovat zdroj a i to poměrně přesně, pokud budou splněny určité podmínky (vhodné vzdálenosti). Tato spolehlivost a přesnost metod však klesá výrazně s rostoucí vzdáleností od mikrofonních polí. Do vzdálenosti přibližně 3 metrů lze metody používat s tím, že odchylky budou maximálně v řádech několika centimetrů.

Práce se zabývala dvěma způsoby určení pozice zdroje a to přes úhel nebo konkrétní souřadnici. Určení úhlů má svoje nedostatky především v rámci přesnosti na větších vzdálenostech, ale jedná se o rychlý způsob lokalizace. Určování souřadnic zdroje zvuku, za cenu delšího výpočtu, zase poskytuje možnost určit konkrétní pozici vůči libovolnému bodu, který je dopředu stanoven. Tento způsob má také však svá omezení (nutnost znát geometrii místnosti dopředu, stanovení vhodného souřadného systému).

Určení úhlu, ze kterého přichází zvuk je dostačující pro případ, že by bylo třeba natočit něco směrem ke zdroji zvuku. Například automatické natočení kamery, mikrofonu, reflektoru a podobných přístrojů. Naopak určení konkrétní pozice řečníka by se mohlo hodit například v kombinaci s jinými automatizovanými vestavěnými přístroji, například automatické spuštění přístrojů na základě toho, odkud se ozve povel.

Ačkoliv jsou metody vyhodnocovány především ze statistického hlediska (pro dokázání jejich spolehlivosti), metody jsou použitelné i pro lokalizaci v reálném čase. Pozice zdroje je vypočítávána a aktualizována v intervalech kratších než jedna vteřina, což by pro lidské poměry mělo být dostačující. Je tedy možné lokalizovat i polohu zdroje, který je v pohybu.

Dané řešení bohužel není dělané pro velké prostory. V kapitole 5 však navrhuji možné vylepšení tak, aby algoritmy šlo aplikovat i pro větší prostory se zachováním přesnosti. Další vývoj by se měl zaměřit především na rozšíření plochy, ve které lze algoritmy použít, tak aby jejich výsledky byly ještě spolehlivé.

Práce může sloužit jako základ pro tvorbu vestavěných systémů s cílem vyznačit pozici zdroje zvuku. Další možností pro budoucí vývoj může být i rozšíření o rozpoznávání řeči.

---

<sup>1</sup>DAS – Delay and Sum

<sup>2</sup>TDOA – Time Difference of Arrival

Kombinace lokalizace zdroje zvuku a rozpoznávání různých povelů může najít uplatnění například v chytrých domácnostech.

# Literatura

- [1] Bernat, P. *Akustika, vznik a šíření zvuku, frekvenční analýza a syntéza, sluchový vjem zvukového signálu* [online]. 15.01.2019 [cited]. Dostupné z: <[https://homen.vsb.cz/~ber30/texty/varhany/anatomie/pistaly\\_akustika.htm](https://homen.vsb.cz/~ber30/texty/varhany/anatomie/pistaly_akustika.htm)>.
- [2] Dostálek, P. : *Embedded systém pro lokalizaci zdroje zvuku*. Dizertační práce, Zlín, 2009, Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky.
- [3] Grénar, M. *Tvarování přijímací charakteristiky mikrofonového pole* (Bakalářská práce). Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 42 s. Vedoucí práce Ing. Ivan Míča.
- [4] Prabhakar S.Naidu. *Sensor array signal processing*. CRC Press LLC, 2001. ISBN 0-8493-1195-0.
- [5] Tkadlec, J. *Lokalizace vzdáleného zdroje zvuku polem mikrofonů* (Bakalářská práce). Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2011. 39 s. Vedoucí práce Ing. Zdeněk Havránek, PhD.
- [6] Tuma, J.; Janečka, P.; Vala, M.; aj. *Sound Source Localization. Proceedings of the 13th International Carpathian Control Conference (ICCC)*. 2012. Dostupné z: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6228744&tag=1>>. ISSN 978-1-4577-1868-7.

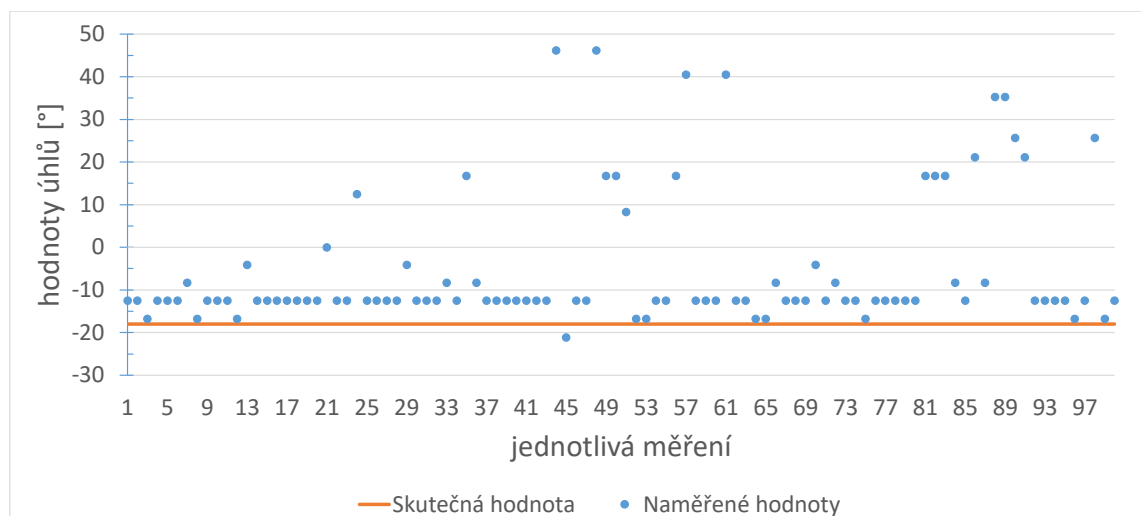
# Příloha A

## Další grafy výsledků testování

Většina výsledků testování byla statisticky vyhodnocena v textu práce, zde přikládám grafy zobrazující konkrétní získané hodnoty pro určité provedené experimenty.

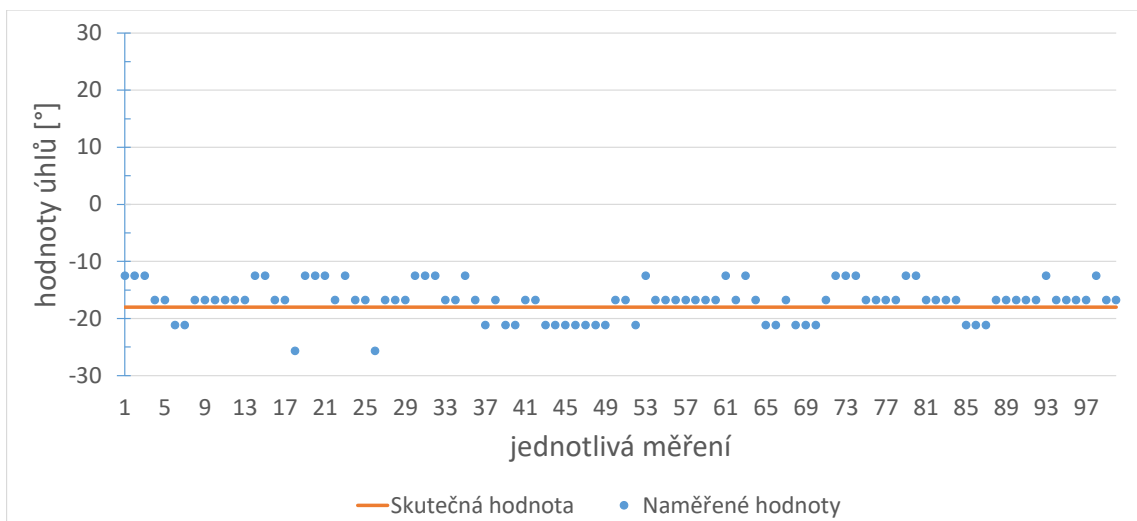
### A.1 Určování úhlu ze zpoždění

Zde jsou uvedeny konkrétní spočtené hodnoty úhlů pro metodu TDOA a DAS. Tyto výsledky se váží k tabulkám 4.1 a 4.3. Uvedené grafy jsou ukazují, že ve větších vzdálenostech od mikrofonního pole měly výsledky tendenci se často vychylovat od očekávané hodnoty, což mělo za následek vysoké hodnoty rozptylu, naopak u menších vzdáleností tato tendence mizí.

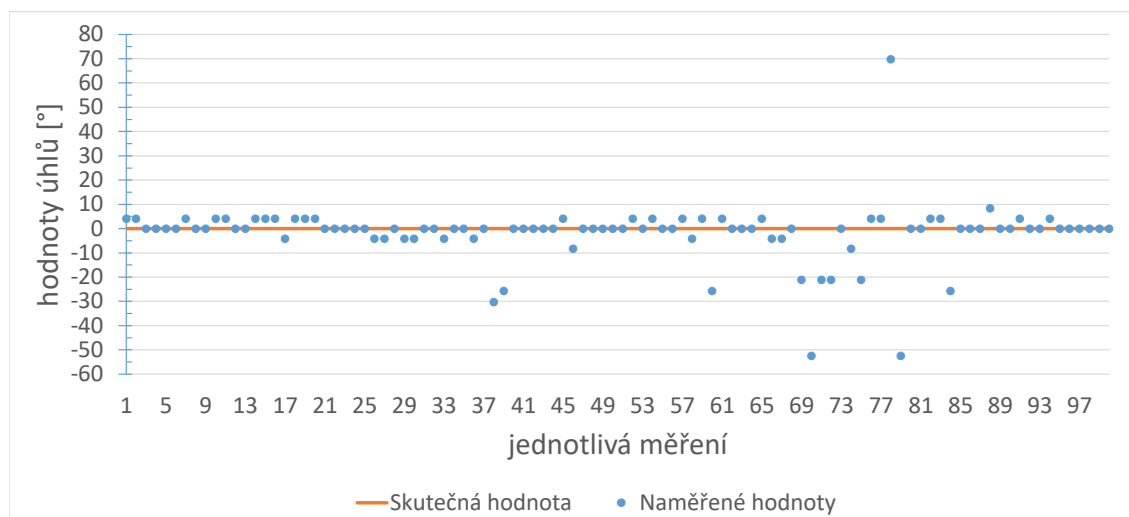


Obrázek A.1: Získané úhly metodou DAS, pro vzdálenost 5 m od mikrofónů a při úhlu  $-18^\circ$

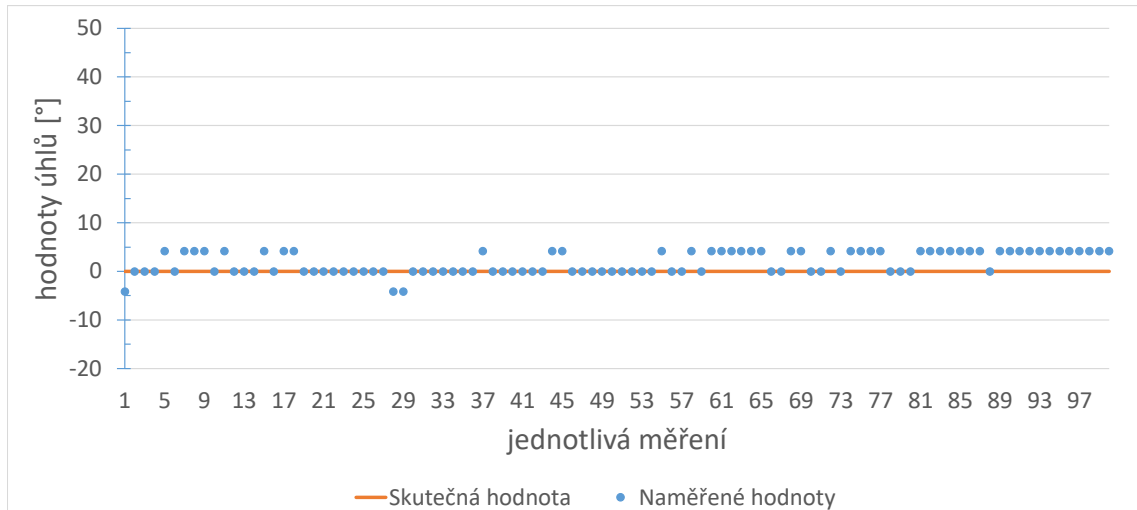




Obrázek A.2: Získané úhly metodou DAS, pro vzdálenost 1 m od mikrofonů a při úhlu  $-18^\circ$



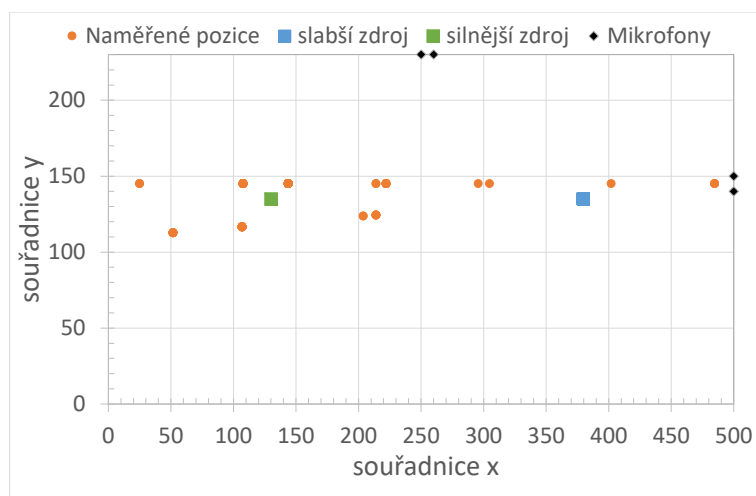
Obrázek A.3: Získané úhly metodou DAS, pro vzdálenost 1 m od mikrofonů a při úhlu  $-0^\circ$



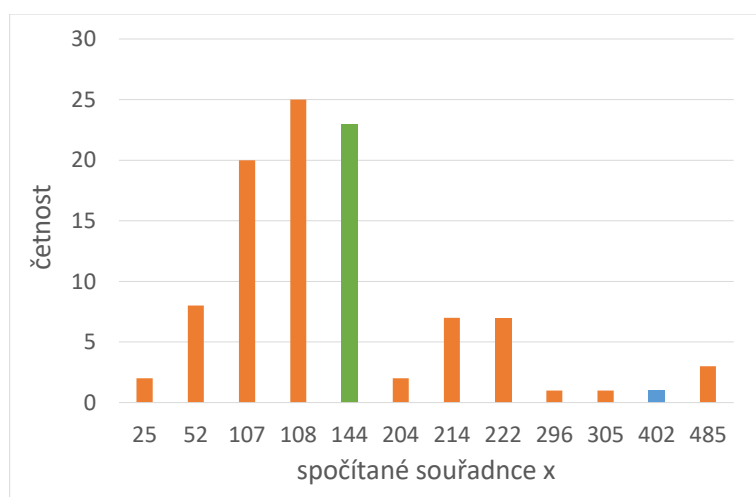
Obrázek A.4: Získané úhly metodou DAS, pro vzdálenost 1 m od mikrofonů a při úhlu 0°

## A.2 Lokalizace v prostoru se dvěma zdroji zvuku

Zde jsou zahrnuty grafy výsledků při použití metod TDOA a DAS pro lokalizaci v místnosti s více než jedním zdrojem zvuku. Jedná se o výsledky 2. testu, který je zmíněn v podsekcí [4.3.1](#).

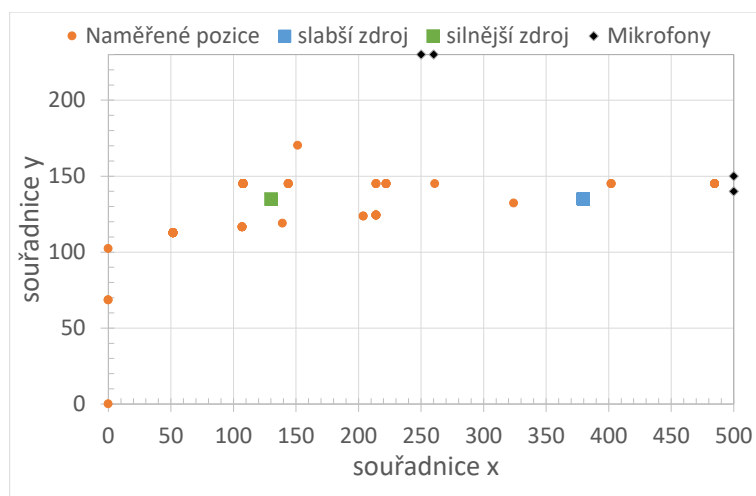


(a) TDOA - Graf se spočítanými pozicemi (oranžové), silnějším zdrojem (zelená), slabším zdrojem (modrá)

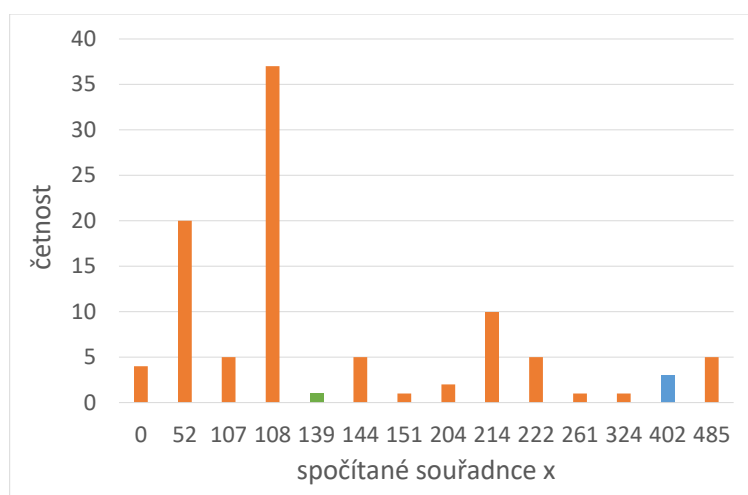


(b) TDOA - Četnost zjištěných souřadnic  $x$  z předchozího grafu A.5a, zeleně je vyznačena hodnota nejbližší k silnějšímu zdroji, modře hodnota nejbližší slabšímu zdroji

Obrázek A.5: Test 2: TDOA - výsledky testu se dvěma zdroji



(a) DAS - Graf se spočítanými pozicemi (oranžové), silnějším zdrojem (zelená), slabším zdrojem (modrá)



(b) DAS - Četnost zjištěných souřadnic  $x$  z předchozího grafu A.6a , zeleně je vyznačena hodnota nejbliže k silnějšímu zdroji, modře hodnota nejbliže slabšímu zdroji

Obrázek A.6: Test 2: DAS - výsledky testu se dvěma zdroji

## Příloha B

# Obsah přiloženého paměťového média

Přiložené paměťové médium obsahuje následující adresářovou strukturu:

```
/
├── scripts.....složka se skripty pro zpracování signálů
├── latex.....složka s LATEX zdrojovými texty a dalšími náležitostmi
├── xchuye00-detekce-pozice.pdf..... písemná zpráva v PDF formátu
├── README
├── plakat.pdf.....propagační plakát
└── video.mp4.....propagační video
```

Soubor README blíže popisuje, jak se skripty spouštějí a co je nutné k tomu, aby fungovaly správně. Také je zahrnut postup, jak se pracovat s hardwarem od Audified.

## Příloha C

# Plakát

Součástí zadání bylo vytvořit plakát velikosti A2. Na další straně je plakát zobrazen, na tuto stranu se nevešel vzhledem k tomu, že nadpis kapitoly zabírá tolik místa na stránce.

# AKUSTICKÁ DETEKCE POZICE ŘEČNÍKA pomocí mikrofonního pole

Fakulta informačních technologií VUT v Brně  
Ústav počítačové grafiky a multimedií

Chu Yen Nhi  
vedoucí: Ing. Igor Szóke PhD.

## Mikrofonní pole

... je senzorové pole, které ke tvoření dvěma nebo více mikrofony a kterým je možné vytvořit vícekanalový záznam zvuku

## Lokalizace zdroje zvuku

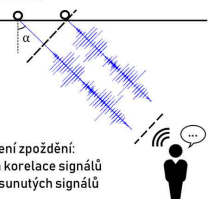
... spočívá v určení zpoždění mezi signály a určení úhlu dopadu signálu na mikrofonní pole. Pro získání přesné pozice je nutné mít aspoň 2 mikrofonní pole.

Využívají se 2 algoritmy lokalizace zdroje:

- metoda časových zpoždění (TDOA)
- metoda tvarování přijímací charakteristiky (DAS)

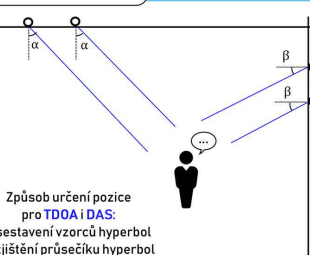
1.

určení úhlu dopadu signálu na základě zjištěného zpoždění

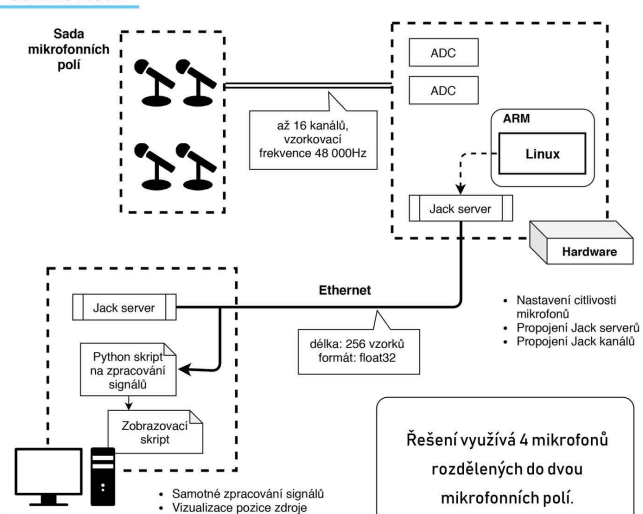


2.

určení přesné pozice řečníka pomocí dvou mikrofonních polí



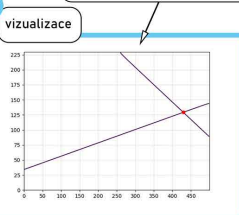
## Schéma řešení



Signály zpracovává Python skript, jeho výstupem jsou souřadnice zdroje zvuku ve zvolené soustavě souřadnic. Je nutné předem znát geometrii prohledávaného prostoru.

3.

Výstup zobrazovacího skriptu



## Dosažené výsledky

Na základě provedených testů bylo prokázáno, že oba použité algoritmy jsou poměrně přesné v rozsahu **3x3 metrů**. Do tohoto rozsahu se odchylka vypočítané pozice od skutečné pozice pohybuje v řádech několika desítek centimetrů (do max. 30 cm). S rostoucí vzdáleností od mikrofónů se nepřesnost lokalizace zvyšuje.