

**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

---

**Fakulta elektrotechnická  
Katedra radioelektroniky**

**Akustické zaměřování  
Acoustic Source Localization**

Diplomová práce

**Studijní program:** (MP2) Komunikace, multimedia a elektronika (magisterský)

**Studijní obor:** (2602T015) Multimediální technika (magisterský)

**Vedoucí práce:** Dr. Ing. Libor Husník

**Bc. Josef Zitko**

---

**Praha 2014**

## Čestné prohlášení

Prohlašuji, že jsem zadanou diplomovou prací zpracoval sám s přispěním vedoucího práce a používal jsem pouze literaturu v práci uvedenou. Dále prohlašuji, že nemám námitek proti půjčování nebo zveřejňování textové a programové části mé diplomové práce se souhlasem katedry.

V Praze dne .....

.....

podpis studenta

## Poděkování

Chtěl bych poděkovat svým rodičům za podporu při studiu a tudíž za možnost zpracování této diplomové práce. Stejně tak bych rád poděkoval vedoucímu diplomové práce Dr. Ing. Liboru Husníkovi za odbornou pomoc při jejím zpracování. A v neposlední řadě děkuji za částečné financování z projektu Metody monitorování a modelování v akustice SGS13/193/OHK3/3T/13.

V Praze dne .....

.....

podpis studenta

## **Abstrakt**

Na začátku této práce popisuji základní vlastnosti lokalizačních metod. Především se zaměřuji na metody pracující s odhadem časového zpoždění (*metody TDOA*) a na metodu využívající akustickou překážku. V následující části diplomové práce se věnuji zpracování digitálního signálu v algoritmech lokalizačních metod TDOA (*CC a GCC*) a seznamuji čtenáře s grafickým uživatelským rozhráním lokalizačního programu. Dále popisuji jednotlivé součásti akustického zaměřovače (*zkonstruovaného v rámci této práce*) a uvádím jeho parametry. V poslední části diplomové práce se věnuji vyhodnocení naměřených hodnot. Z naměřených výsledků jsou potvrzeny teoretické předpoklady, které jsou uvedeny v rešerši akustických lokalizačních metod. Na základě zprůměrování všech naměřených hodnot je stanovena výsledná přesnost akustického zaměřovače využívající lokalizační metodu TDOA GCC.

*Klíčová slova: Akustické zaměřování, akustický zaměřovač, TDOA, GCC, PHAT, SCOT, RP*

## **Abstract**

At the beginning of this thesis I describe basic properties of localization methods. Primarily I focus on methods working with an estimation of time delay (*TDOA methods*) and method using acoustic barrier. In the following part of the diploma thesis I deal with digital signal processing in algorithms of TDOA localization methods (*CC and GCC*) and graphical user interface of application for localization. In next part of the work I describe individual components of acoustic locator (*constructed in this thesis*) and present his parameters. In the last part of the diploma thesis I deal with evaluating of measured values. Theoretical assumptions which are shown in search of acoustic localization methods are confirmed from measured results. Accuracy of the acoustic locator using GCC TDOA method is determined from averaged all measurements results.

*Key words: Acoustic localization, acoustic locator, TDOA, GCC, PHAT, SCOT, RP*

## Obsah

1.	Šíření zvuku .....	9
1.1.	Popis zvukové vlny.....	9
1.2.	Šíření zvuku ve volném a uzavřeném prostoru .....	13
2.	Základní principy akustického zaměřování .....	14
2.1.	Geometrie mikrofonních polí, jejich vlastnosti a omezení .....	18
2.1.1.	Maximální vzdálenost mikrofonů .....	20
2.1.2.	Prostorový aliasing mikrofonního pole .....	21
2.1.3.	Princip triangulace .....	22
2.2.	Metody TDOA – Time Difference Of Arrival .....	23
2.2.1.	TDOA CC – Cross Correlation.....	24
2.2.2.	TDOA MSC – Magnitude Squared Coherence .....	31
2.2.3.	Metody s tvarováním směrových charakteristik .....	33
2.3.	Metody využívající spektrální odhad vysokého rozlišení.....	36
2.4.	Metoda s akustickou překážkou .....	37
3.	Číslíkové zpracování signálů.....	40
3.1.	Kalibrace .....	41
3.2.	Filtrace .....	42
3.3.	Grafické rozhraní - GUI .....	45
4.	Konstrukce akustického zaměřovače.....	47
4.1.	Geometrie mikrofonního pole akustické zaměřovače .....	47
4.2.	Mikrofonní předzesilovače .....	49
4.3.	Mikrofonní vložky .....	51
4.4.	Zvuková karta .....	52
5.	Měření.....	53
5.1.	Výsledky měření .....	55
6.	Závěr.....	58
7.	Literatura.....	59
8.	Příloha A – Naměřené výsledky.....	63
9.	Příloha B – Obsah příloženého CD .....	67

## Seznam obrázků

Obr. 1: Šíření rovinné vlny .....	10
Obr. 2: Šíření kulové vlny od bodového zdroje.....	12
Obr. 3: Metoda zrcadlení zdrojů .....	13
Obr. 4: Princip triangulace.....	15
Obr. 5: Závislost úhlového rozlišení na velikosti zpoždění .....	17
Obr. 6: Zleva: Lineární, plošné a prostorové rozložení mikrofonního pole .....	18
Obr. 7: Zleva: uniformní a neuniformní lineární mikrofonní pole .....	18
Obr. 8: Stranová neurčitost lineárního mikrofonního pole .....	19
Obr. 9: Periodicita signálů .....	20
Obr. 10: Princip triangulace.....	23
Obr. 11: Model s impulsní odezvou .....	25
Obr. 12: Vzájemná korelační funkce se zpožděním a útlumem .....	26
Obr. 13: Model s aditivním šumem.....	26
Obr. 14: Vzájemná korelační funkce se zpožděním, útlumem a šumem.....	27
Obr. 15: Model vícecestného šíření .....	27
Obr. 16: Vzájemná korelační funkce (více cest).....	28
Obr. 17: Blokové schéma výpočtu MSC.....	32
Obr. 18: Blokové schéma tvarovače DAS .....	34
Obr. 19: Směrová charakteristika .....	35
Obr. 20: Návrh a reálná konstrukce zaměřovače .....	38
Obr. 21: Rozložení akustického tlaku okolo překážky .....	38
Obr. 22: Blokové schéma výpočetního algoritmu .....	40
Obr. 23: Frekvenční maska .....	43
Obr. 24: Funkce fft.m MATLAB .....	43
Obr. 25: Grafické rozhraní lokalizačního programu.....	45
Obr. 26: Mikrofonní pole 3D poloprostor .....	47
Obr. 27: Mikrofonní pole 2D .....	48
Obr. 28: Trubička s aretací.....	48
Obr. 29: Konstrukční prvek umožňující přestavbu mikrofonního pole .....	49
Obr. 30: Pozice mikrofonů.....	49

Obr. 31: Přenosová frekvenční charakteristika (simulace) .....	50
Obr. 32: Schéma zapojení mikrofonního předzesilovače .....	50
Obr. 33: Zkonstruovaný blok předzesilovačů .....	51
Obr. 34: Zvuková karta ESI MAYA 44 USB .....	52
Obr. 35: Frekvenční charakteristika YAMAHA MSP5 STUDIO .....	53
Obr. 36: Blokové schéma zapojení YAMAHA MSP5 STUDIO .....	53
Obr. 37: Měřicí soustava .....	54

## Úvod

Důvodem, proč se lidé zabývají akustickým zaměřováním, je především snaha zkonstruovat akustické zaměřovače, které jsou následně využívány v mnoha oblastech (*například vývoje nebo asistenčních služeb*). Mezi již úspěšně využívané akustické zaměřovače mohou zařadit některé přístroje pro asistenční služby fyzicky postižených lidí nebo zaměřovače využívané k nalezení mechanických závad ( *které se projevují mechanickými vibracemi*) a mnohé další.

Podoba zvukového signálu je ovlivněna mnoha faktory, které následně rozhodují o způsobu zpracování signálu a přesnosti zaměřování. Faktory, které nejvíce ovlivňují podobu zvukového signálu, jsou typ prostoru (*volný nebo uzavřený*), teplota prostředí (*změna rychlosti zvuku*), šířka pásma, počet zvukových zdrojů a pohyb zvukového zdroje. Tyto faktory jsou zároveň i základními podmínkami pro výběr vhodné lokalizační metody. Akustické lokalizační metody rozdělujeme na tři základní skupiny. První skupinou jsou metody založené na detekci časového zpoždění signálu (*TDOA – Time Difference Of Arrival*). Druhou skupinou jsou lokalizační metody využívající spektrálního odhadu vysokého rozlišení (*pracují pouze ve frekvenční oblasti*). A poslední třetí skupina využívá charakteristických vlastností zvukové vlny dopadající na překážku. Metody z první skupiny, tedy ze skupiny využívající časového zpoždění, můžeme dále dělit na metody zpracovávající signál v časové oblasti (*TDOA CC*) a na skupinu pracující v oblasti frekvenční (*TDOA MSC*). Třetí metoda z této skupiny využívá směru s maximálním ziskem ve směrové charakteristice. Tuto metodu můžeme opět rozdělit na dvě podskupiny. V první podskupině jsou zaměřovače s pohyblivým akustickým měničem (*rotační měření směrové charakteristiky*) a ve druhé podskupině jsou zaměřovače se statickým mikrofonním polem (*tvarovač DAS*), zde se zachycený signál uměle zpožďuje a nahrazuje mechanickou rotací akustického měniče. Třetí skupina ze základního rozdělení je založená na charakteristickém chování zvukové vlny dopadající na překážku. Tato zaměřovací metoda pracuje na principu akustického stínu, který se za určitých podmínek vytváří za překážkou (*mikrofonní pole je zároveň i překážkou*).

Jelikož existuje mnoho akustických zaměřovacích metod a každá z nich dosahuje nejlepších výsledků za různých podmínek, budu se v této práci zabývat zařazením zaměřovacích metod z hlediska jejich využitelnosti. V závěru této práce ověřím platnost teoretických vlastností vybraných lokalizačních metod TDOA pomocí zkonstruovaného akustického zaměřovače.



# 1. Šíření zvuku

Zvuk definujeme jako mechanické kmity kontinua. Tento druh vlnění můžeme popsat pomocí obecné vlnové rovnice, viz vztah č. 1. V plynných a kapalných látkách se zvuk šíří ve formě podélného vlnění, v pevných látkách ve formě příčného i podélného vlnění. Zabývat se budu pouze šířením zvuku v plynném prostředí (*podélné vlnění*), konkrétně se zaměřím na vlastnosti kulové a rovinné vlny a následně pohovořím o problematice šíření zvuku v uzavřeném a volném prostoru.

## 1.1. Popis zvukové vlny

(kapitola převzata z mé práce [27])

Jak již bylo řečeno v předchozím odstavci, jakékoliv vlnění můžeme popsat pomocí obecné vlnové rovnice a proto i zvukovou vlnu. V této kapitole se zaměřím na odvození základních vlastností kulové a rovinné vlny.

Kulovou vlnu (ve vzdálené zóně od zdroje zvuku) často aproximujeme vlnou rovinnou. Tuto aproximaci si můžeme dovolit právě tehdy, pokud můžeme zanedbat zakřivení vlnoplochy (u kulové vlny). Základním výpočetním principem v oblasti akustického zaměřování je princip triangulace (kapitola *Princip triangulace*). Zde je nutné uvažovat zvukovou vlnu jako vlnu rovinnou. Tuto aproximaci děláme ve všech vzdálenostech od zdroje zvuku. A proto v případě zaměřování v blízkém poli vnášíme aproximací chybu do výpočtu. Odvození základních vlastností kulové a rovinné vlny:

$$\frac{1}{c^2} \cdot \frac{\partial^2 \Phi}{\partial t^2} = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \cdot \Phi \quad (1)$$

- $\Phi$  – rychlostní potenciál
- $c$  – rychlost zvuku

V případě že  $\Phi$  je v čase harmonicky proměnné s kruhovou frekvencí  $\omega$ , můžeme jej vyjádřit:

$$\Phi = \Psi e^{j\omega t} \quad (2)$$

Nyní provedeme druhou derivaci vztahu **č. 2** podle času:

$$\frac{\partial^2}{\partial t^2} \Phi = -\omega^2 \Psi e^{j\omega t} = -\omega^2 \Phi \quad (3)$$

Po dosazení a úpravě vztahu **č. 3** do obecné vlnové rovnice dostaneme tuto rovnici:

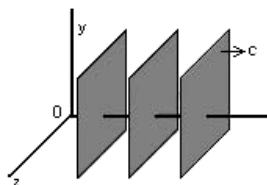
$$\Delta \Psi + k^2 \Psi = 0 \quad (4)$$

kde definujeme vlnové číslo  $k$ , jako:

$$k = \frac{\omega}{c} \quad (5)$$

## Rovinná vlna

Tento typ vlny je charakteristický svou rovinnou vlnoplochou (*spojíme-li místa se shodnou fází, dostáváme rovinu*). Odvození charakteristických vlastností rovinné vlny provedeme na základě rozboru obecné vlnové rovnice.



Obr. 1: Šíření rovinné vlny

Rovinná vlna postupuje pouze v ose  $X$  a z toho vyplývá, že členy obsahující souřadnice  $Y$  a  $Z$  jsou nulové, a proto je můžeme zanedbat.

$$\frac{d^2 \Psi}{dx^2} + k^2 \Psi = 0 \quad (6)$$

Řešení rovnice č. 6, má tvar:

$$\Psi = Ae^{\alpha x} \quad (7)$$

Po dosazení vztahu č. 7 do rovnice č. 6 a jejího vyřešení dostáváme charakteristickou rovnici:

$$\alpha^2 + k^2 = 0 \quad (8)$$

Ze které plyne řešení:

$$\Psi(x) = A_1 e^{-jkx} + A_2 e^{jkx} \quad (9)$$

Protože

$$\Phi = \Psi e^{j\omega t} \quad (10)$$

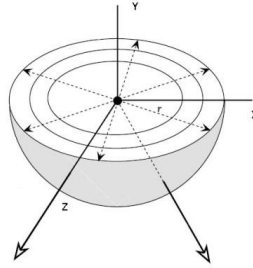
Můžeme po úpravě napsat:

$$\Phi = A_1 e^{j(\omega t - kx)} + A_2 e^{j(\omega t + kx)} \quad (11)$$

Vztah č. 11 vyjadřuje šíření vlny ve směru osy  $X$  s amplitudou  $A_1$  a šíření druhé vlny s amplitudou  $A_2$  ve směru osy  $-X$ .

## Kulová vlna

Kulová vlna je charakteristická svou vlnoplochou, která má tvar koule. V případě, že chceme matematicky popsat kulovou vlnu, musíme uvažovat zdroj tohoto vlnění jako obecnou pulzující kouli. Od tohoto zdroje se vlna šíří všemi směry stejně rychle (*pro případy homogenního a izotropního prostředí*) a ve vzdálenosti  $r$  od zdroje má ve všech bodech stejnou fázi.



Obr. 2: Šíření kulové vlny od bodového zdroje

Po úpravě obecné vlnové rovnice můžeme psát:

$$\frac{d^2(\Psi r)}{dr^2} + k^2(\Psi r) = 0 \quad (12)$$

- $r$  – vzdálenost od zdroje vlnění

Vlnové číslo  $k$ :

$$k = \frac{\omega}{c} \quad (13)$$

A tudíž:

$$\Psi r = A_1 e^{-jkr} + A_2 e^{jkr} \quad (14)$$

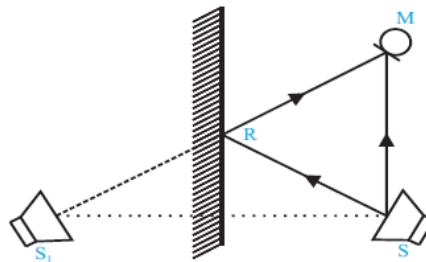
$$\Phi r = A_1 e^{j(\omega t - kr)} + A_2 e^{j(\omega t + kr)} \quad (15)$$

Rovnice **č. 15** zahrnuje řešení vlny postupující od zdroje i řešení šíření vlny do zdroje. V praxi se téměř vždy využívá pouze řešení pro šíření vlny od zdroje, z tohoto důvodu položíme konstantu  $A_2 = 0$ .

$$\Phi = \frac{A_1}{r} e^{j(\omega t - kr)} \quad (16)$$

## 1.2. Šíření zvuku ve volném a uzavřeném prostoru

Volné prostory se vyznačují tím, že zde nedochází k odrazům zvukových vln. Odražené zvukové vlny přicházejí téměř vždy z jiných směrů, než se nachází zdroj zvuku. Jsou však i specifické případy odrazů kdy se zvuková vlna šíří ze směru zdroje zvuku (*tato vlna bude zpožděna a utlumena ekvivalentně vzdálenosti a charakteru cesty, kterou musela urazit*). Obrázek č. 3 převzat z [22].



Obr. 3: Metoda zrcadlení zdrojů

Za běžných podmínek (*uzavřené, polouzavřené prostory*) dochází k odrazům zvukových vln. Charakter odražených vln závisí na jejich časovém zpoždění a přirozeném útlumu prostředí. Aby nedocházelo při zaměřování k chybám (*vlivem odražených vln*), je nutné používat lokalizační metody, které jsou schopny tyto odražené vlny (*zrcadlové zdroje*) eliminovat nebo alespoň určitým způsobem potlačit. Vlastnosti jednotlivých lokalizačních metod pro případy vícecestného šíření vln, jsou popsány v kapitole *Metody TDOA – Time Difference Of Arrival*.

## 2. Základní principy akustického zaměřování

Volba akustické zaměřovací metody a její přesnost jsou závislé na podrobném popisu prostoru, ve kterém k zaměřování dochází. Základní rozdělení prostorů je na prostory uzavřené a prostory volné. Volné prostory jsou charakteristické tím, že v nich nedochází k odrazu zvukové vlny. To má za následek, že se v zachyceném signálu neobjevují zpožděné kopie signálů a jiné s tím související artefakty, vnášející do naměřených výsledků chyby. Především metody založené na časovém zpoždění jsou choulostivé na přítomnost odražených zvukových vln. Důvodem je princip, na kterém tyto lokalizační metody pracují (*časová podobnost signálů – korelační metody a frekvenční podobnost signálů – koherenční metody*). V následujícím seznamu jsou vypsány nejčastěji využívané akustické lokalizační metody a odkazy na práce, které se jimi zabývaly:

### 1. TDOA – Time Difference Of Arrival (*časový rozdíl příchodu*)

1.1. TDOA CC – Cross Correlation (*vzájemná korelační funkce*) – [23],[22],[7],[5]

1.2. TDOA MSC – Magnitude Squared Coherence (*normovaná koherence*) – [11],[9]

### 1.3. Metody s tvarováním směrových charakteristik

1.3.1. TDOA DAS – rotační akustický měnič

1.3.2. TDOA DAS – statické pole akustických měničů – [26],[23],[22],[20]

### 2. Metody využívající spektrálního odhadu vysokého rozlišení – [28]

### 3. Metody s akustickou překážkou – [27]

Další důležitou podmínkou pro korektní výsledky je přesné stanovení rychlosti zvuku  $c$ . Pro naše případy je dostačující počítat rychlost ze vztahu odvozeného pro šíření zvuku ve vzduchu (*vztah č. 17*). V tomto vztahu je jediná proměnná a tou je teplota daného plynu (*vzduchu*). Z tohoto důvodu je v následující tabulce znázorněno, jak velký vliv může mít nesprávné určení teploty, respektive rychlosti zvuku  $c$ .

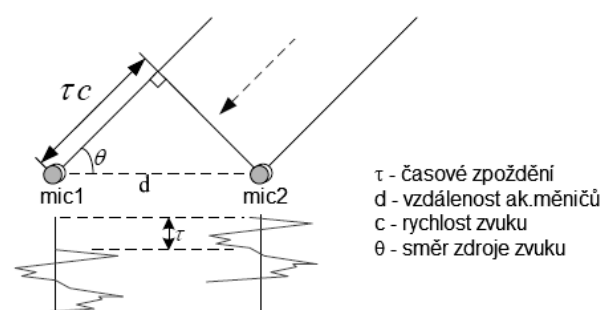
$$c = 331.57 + 0.607 \cdot \vartheta \text{ [ms}^{-1} \text{]} \quad (17)$$

- $\vartheta$  – teplota vzduchu

$\vartheta$ [°C]	$c$ [m/s]	úhel [°]
0.0	331.6	0.0
$\Delta\vartheta$ [°C]	$\Delta c$ [m/s]	$\Delta\text{úhel}$ [°]
2	1.2	0.2
4	2.4	0.4
6	3.6	0.6
8	4.9	0.7
10	6.1	0.9
15	9.1	1.4
20	12.1	1.9

Tab. 1: Úhlová chyba

Jako referenční hodnota byla zvolena rychlost zvuku  $c$  při teplotě  $0^\circ\text{C}$ . Pro konkrétní odchylky od této referenční teploty je vypočítána změna rychlosti zvuku  $c$  a následně i tomu odpovídající úhlová chyba. Tato chyba může dosahovat hodnot od několika desetin až po jednotky stupňů. Úhel byl vypočten z časového zpoždění pomocí triangulace (*triangulace je jedním ze základních způsobů, jak stanovit časové zpoždění zvukové vlny pomocí geometrie pravoúhlého trojúhelníku a bude podrobně popsán v kapitole Princip triangulace*). Úhlová chyba zapříčiněná nepřesným určením rychlosti zvuku  $c$  není chybou konečnou. Tato chyba se může dále zvětšovat z důvodu konečného úhlového rozlišení daného akustického zaměřovače. Rozlišení zaměřovače, využívajícího digitálního zpracování signálu, je závislé především na jeho vzorkovacím kmitočtu  $f_s$  a vzájemné vzdálenosti mikrofونů  $d$ . Čím vyšší součin  $f_s \cdot d$ , tím je vyšší rozlišovací schopnost zaměřovače. K názornému příkladu využijeme výpočet pomocí triangulace.



Obr. 4: Princip triangulace

Na obrázku **č. 4** jsou dva akustické měniče ve vzájemné vzdálenosti  $d$ . Z jednoduché geometrie pravoúhlého trojúhelníka lze vypočítat úhel  $\theta$ . Tento úhel bude závislý na vzdálenosti  $d$  a vzdálenosti  $\tau \cdot c$  ( $c$  – rychlost zvuku,  $\tau$  – časové zpoždění zvukové vlny mezi  $mic_1$  a  $mic_2$ ). Časové zpoždění zde musí být vždy pouze celistvým násobkem periody vzorkovacího kmitočtu  $f_s$ , kterým je vzorkován signál z  $mic_1$  a  $mic_2$ . Po dosažení konkrétních hodnot vypočítáme maximální zpoždění ve vzorcích (*round* je příkaz v prostředí MATLAB pro zaokrouhlení k nejbližšímu celému číslu):

$$\Delta n = +- \text{round}\left(\frac{\tau}{T_s}\right) = +- \text{round}\left(\frac{d}{c} \cdot \frac{1}{T_s}\right) = +- \text{round}\left(\frac{d \cdot f_s}{c}\right) \quad (18)$$

$$\Delta n = +- \text{round}\left(\frac{d \cdot f_s}{c}\right) = +- \text{round}\left(\frac{0.301 \cdot 48000}{344}\right) = +- 42 = 84 \text{ vzorků} \quad (19)$$

- $d$  = 0.301 m
- $c$  = 344 m/s
- $f_s$  = 48 kHz ( $T_s = 1/f_s$ )

Průměrnou rozlišovací schopnost zaměřovače (pro dané parametry) vypočteme takto:

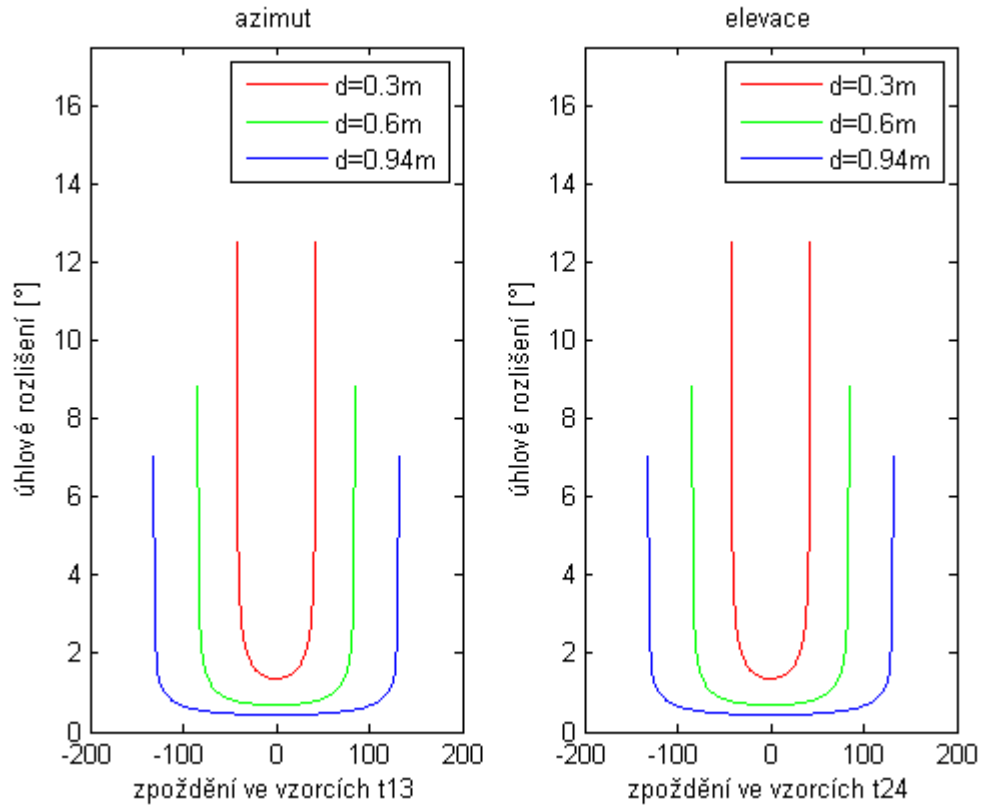
$$AV \text{ rozlišení} = \frac{180}{\Delta n} = \frac{180}{84} = 2.14^\circ \quad (20)$$

		AV rozlišení [°]			AV rozlišení [°]
$f_s = 48 \text{ kHz}$	$d = 0.301 \text{ m}$	2.14	$f_s = 96 \text{ kHz}$	$d = 0.301 \text{ m}$	1.07
	$d = 0.602 \text{ m}$	1.07		$d = 0.602 \text{ m}$	0.54
	$d = 0.94 \text{ m}$	0.68		$d = 0.94 \text{ m}$	0.34

Tab. 2: Průměrná rozlišovací schopnost zaměřovače

Rozlišovací schopnost v celém lokalizačním rozsahu není konstantní. V úhlových pozicích, které odpovídají maximálnímu kladnému nebo zápornému zpoždění je rozlišovací schopnost minimální. Maximální úhlové rozlišení je v oblastech nulového časového zpoždění.





Obr. 5: Závislost úhlového rozlišení na velikosti zpoždění

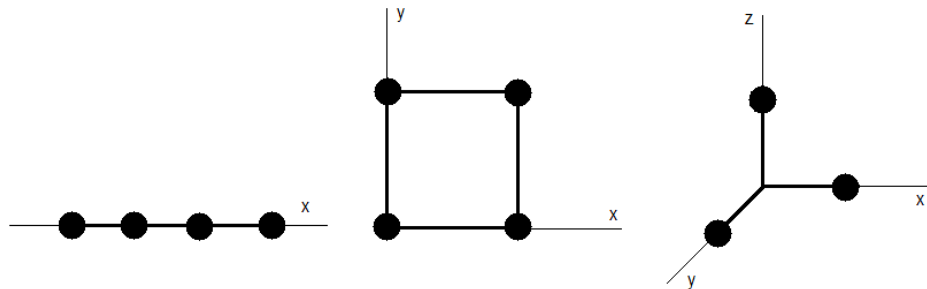
Celková úhlová chyba akustického zaměřovače, způsobená konečným rozlišením a dosazením nesprávné hodnoty rychlosti zvuku  $c$ , může dosahovat tak velkých hodnot, že zaměřovač již nemusí vyhovovat z hlediska přesnosti. Právě proto je důležité dodržovat tyto základní pravidla:

- Přesně stanovit rychlost zvuku  $c$ .
- Používat vysoké vzorkovací kmitočty  $f_s$  (možnost využít převzorkování).
- Zvětšovat vzdálenost  $d$  mezi mikrofonními páry.

## 2.1. Geometrie mikrofonních polí, jejich vlastnosti a omezení

Mikrofonní pole rozdělujeme do skupin dle prostorového uspořádání mikrofonů. Rozložení mikrofonního pole určuje maximální úhlové meze, ve kterých bude akustický zaměřovač schopen pracovat. V praxi je velmi často využíváno pravidelné (*uniformní pole*) 2D (*plošné*) nebo 3D (*prostorové*) uspořádání mikrofonů. Počet mikrofonů se určuje v závislosti na požadovaném úhlovém rozlišení, efektivitě a samozřejmě i ceně. Základní uspořádání mikrofonních polí rozdělujeme do následujících kategorií:

- Lineární (*mikrofony rozmístěny na společné přímce*)
- Plošné (*mikrofony rozmístěny na společné ploše*)
- Prostorové (*mikrofony rozmístěny v prostoru*)



Obr. 6: Zleva: Lineární, plošné a prostorové rozložení mikrofonního pole

Tato pole se dále rozdělují dle jejich pravidelnosti:

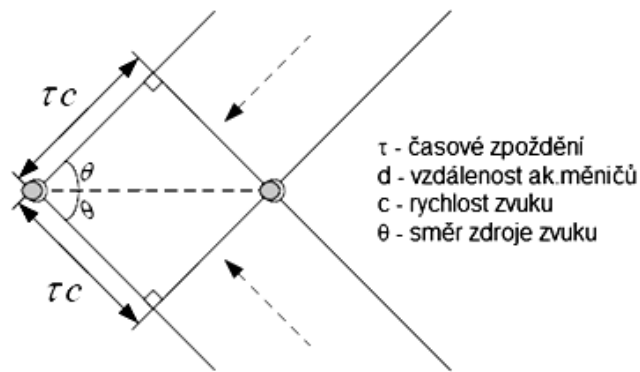
- Pravidelné (*uniformní*)
- Nepravidelné (*neuniformní*)



Obr. 7: Zleva: uniformní a neuniformní lineární mikrofonní pole

S lineárním mikrofonním polem jsme schopni lokalizovat pouze v jedné polorovině. To znamená, že azimut  $\theta$  může dosahovat hodnot  $0^\circ$ - $180^\circ$ , případně  $180^\circ$ - $360^\circ$ . Nejsme schopni určit, který z těchto dvou případů je ten správný, jelikož se jedná pouze o

lineární pole. U lineárních polí nastává problém se stranovou neurčitostí, viz následující obrázek.



Obr. 8: Stranová neurčitost lineárního mikrofonního pole

Pro každé vypočtené časové zpoždění  $\tau$  existují dvě polohy. Polohy jsou symetrické vůči mikrofonní ose a nacházejí se vždy v opačných polorovinách. Zde nejsem schopni rozlišit, o kterou polorovinu se jedná. Důvodem této neurčitosti jsou shodná časová zpoždění  $\tau$ , která by takto polohované zdroje vytvářely.

Tento problém vyřešíme přestavbou lineárního pole na plošné. V plošném poli se mikrofony nejčastěji rozmisťují na osy pomyslného kříže. Toto rozložení je plošné a zároveň i uniformní. Pomocí takto rozložených mikrofونů jsme už schopni přesně určit azimut  $\theta$  v rozmezí  $0^\circ$ -  $360^\circ$ . Zde využíváme jeden mikrofonní pár pro určení azimutu  $\theta$  a druhý pár pro určení poloroviny ( $0^\circ$ - $180^\circ$  nebo  $180^\circ$ - $360^\circ$ ).

Posledním případem je prostorové rozložení mikrofonního pole. Zde se snažíme lokalizovat zdroj zvuku v úhlovém rozmezí  $0^\circ$ - $360^\circ$  (azimut  $\theta$ ) a  $\pm 90^\circ$  (elevace  $\phi$ ). Nejčastěji se mikrofony umisťují do rohů pomyslné krychle, toto rozložení je prostorové a opět uniformní. Pro určení elevace  $\phi$  jsou dostačující dva mikrofony. Dochází zde k polokulové neurčitosti. V případě, že známe azimut  $\theta$ , je známá i polokoule, ve které se zdroj zvuku nachází.

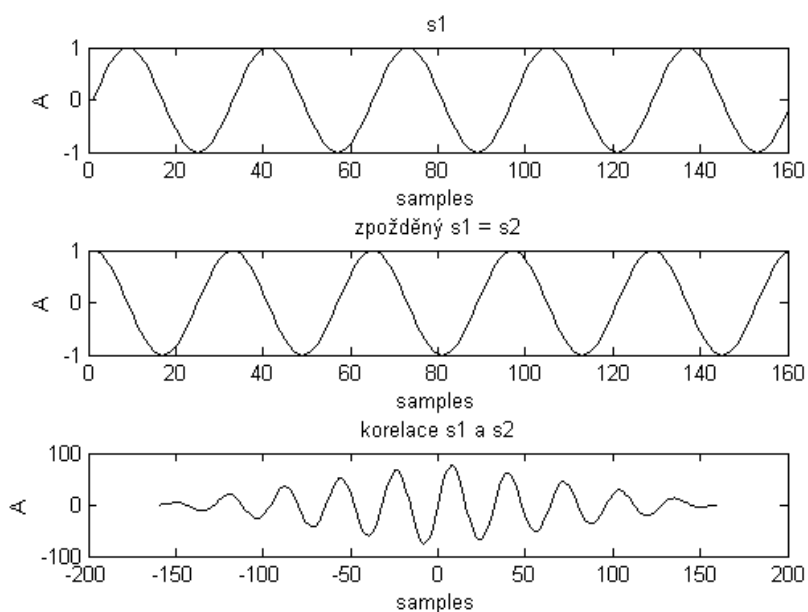
### 2.1.1. Maximální vzdálenost mikrofonů

Maximální vzdálenost mikrofonů (*vzdálenost mikrofonů, mezi kterými se určuje časové zpoždění  $\tau$* ), se určuje z požadavků na průměrnou rozlišovací schopnost zaměřovače a dále z frekvenčního a časového charakteru zaměřovaného zvukového signálu. Stanovení maximální vzdálenosti mikrofonů (*dále jen  $d_{max}$* ) vycházejícího z požadavků na průměrnou rozlišovací schopnost zaměřovače bylo popsáno v kapitole *Základní principy akustického zaměřování*. Viz následující vztah:

$$\Delta n = + - \text{round}\left(\frac{d \cdot fs}{c}\right) \quad (21)$$

Ze vztahu č. 21 je zřejmé, že maximální úhlové rozlišení zaměřovače je závislé na vzorkovacím kmitočtu  $fs$ , vzdálenosti  $d$  a na rychlosti zvuku  $c$ . Rychlost zvuku  $c$  a vzorkovací kmitočet  $fs$  jsou parametry, které většinou neměníme nebo někdy ani nemůžeme. Proto se nabízí možnost změny vzdálenosti  $d$ .

Dalšími faktory, které určují velikost  $d_{max}$ , jsou časové a spektrální vlastnosti zaměřovaných signálů. Konkrétně se jedná o specifické vlastnosti periodických signálů.



Obr. 9: Periodicita signálů

Na obrázku č. 9 je zobrazena vzájemná korelační funkce signálů  $s_1$  a  $s_2$  (*poslední graf na obrázku*). Zde je viditelné maximum, které odpovídá časovému zpoždění  $\tau$  mezi signály. Posuneme-li signál  $s_1$  nebo  $s_2$  o celistvou periodu  $T$ , signály budou mít stále stejný charakter. Vzájemná korelační funkce těchto posunutých signálů bude shodná, a tudíž i časové zpoždění  $\tau$ , které z ní je vypočteno. Z toho vyplývá, že u periodických signálů jsme schopni určit časové zpoždění  $\tau$  o maximální velikosti, které je dáno periodou  $T$ . Z tohoto důvodu je vhodné volit  $d_{max}$  tak, aby maximální časové zpoždění  $\tau$  mezi signály bylo menší, než perioda  $T$ . Z této podmínky můžeme odvodit velikost maximální vzdálenosti  $d_{max}$ :

$$d_{max} < \lambda \quad (22)$$

- $\lambda$  – vlnová délka

$$\lambda = \frac{c}{f} \quad (23)$$

- $f$  – frekvence

Pokud zmenšujeme vzdálenost  $d$  mezi mikrofony a zároveň chceme zachovat rozlišovací schopnost zaměřovače, je nutné zvýšit vzorkovací kmitočet  $f_s$  (*kolikrát zmenšíme vzdálenost  $d$ , tolikrát zvýšíme vzorkovací frekvenci  $f_s$* ).

### 2.1.2. Prostorový aliasing mikrofonního pole

Charakter rozložení mikrofonního pole se vždy navrhuje podle toho, jaké vlastnosti (*rozlišovací schopnost, úhlové rozmezí*) akustického zaměřovače požadujeme. Jedním z dalších kritérií, které ovlivňují návrh mikrofonního pole, je takzvaný prostorový aliasing. Prostorový aliasing je obdobou frekvenčního aliasingu při vzorkování signálů. Připomeňme si Shannon-Nyquistovu vzorkovací podmínku:

$$f_s \geq 2f_m \quad (24)$$

Tato podmínka říká, že vzorkovací kmitočet  $f_s$  musí být větší nebo roven dvojnásobku maximální frekvence  $f_m$ , která je obsažena ve spektru vzorkovaného signálu. Pokud by nebyla podmínka dodržena, docházelo by na vyšších frekvencích k vzájemnému

překrývání spektrálních kopií (*aliasing*), která vznikají v důsledku vzorkování.

Jak již bylo řečeno, prostorový aliasing je obdobou aliasingu při vzorkování signálů. Zde se mikrofony chovají jako prostorové vzorkovače, u kterých musíme dodržet podmínku pro jejich maximální vzájemnou vzdálenost  $d$ . V případech kdy vzájemná vzdálenost mikrofونů  $d$  je větší, než polovina vlnové délky  $\lambda$ , dochází ve směrové charakteristice k tvorbě vedlejších (*nechtěných*) laloků.

$$d > \frac{\lambda}{2} \quad (25)$$

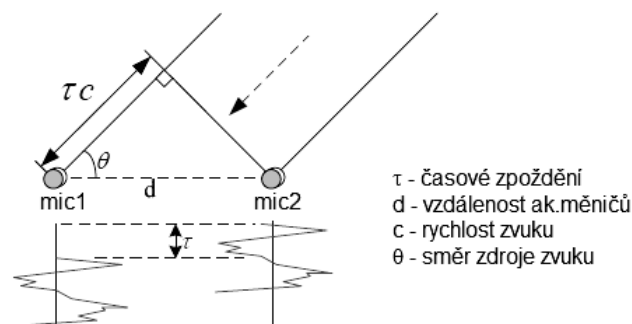
Odvození maximální vzdálenosti  $d$ , při které nedochází k tvorbě vedlejších laloků (ve směrové charakteristice):

$$d \leq \frac{\lambda_{\min}}{2} = \frac{c}{2f} \quad (26)$$

Tato práce se primárně zabývá zaměřovacími metodami TDOA CC. U těchto metod se nevyužívá směrových charakteristik, a proto se tato podmínka může zanedbat. Nelze ji však zanedbat u zaměřovacích metod založených na tvarování směrových charakteristik (*například metoda TDOA DAS*).

### 2.1.3. Princip triangulace

Základní způsob výpočtu směru zdroje zvuku z odhadnutého časového zpoždění  $\tau$  je založen na jednoduché geometrii pravoúhlého trojúhelníka. Tento způsob výpočtu se nazývá triangulace.



Obr. 10: Princip triangulace

Na obrázku č. 10 je znázorněna dvojice mikrofónů  $mic_1$  a  $mic_2$  (lineární mikrofonní pole), která je ozařována akustickým zdrojem ze směru  $\theta$ . Abychom mohli využít princip triangulace, musíme zvukovou vlnu prohlásit za vlnu rovinnou (tuto aproximaci si můžeme dovolit až ve vzdálené zóně od zdroje zvuku). Poté pomocí goniometrických funkcí a časového zpoždění  $\tau$  vypočítáváme úhel  $\theta$ .

$$\cos(\theta) = \frac{\tau \cdot c}{d} \quad (27)$$

- $\tau$  – časové zpoždění
- $c$  – rychlost zvuku
- $d$  – vzdálenost mezi mikrofony

A následně:

$$\theta = \arccos\left(\frac{\tau \cdot c}{d}\right) \quad (28)$$

## 2.2. Metody TDOA – Time Difference Of Arrival

Metody založené na určení časového zpoždění  $\tau$  příchozího signálu se obecně nazývají TDOA. Časové zpoždění  $\tau$  mezi jednotlivými mikrofonními páry bývá nejčastěji vypočítáváno pomocí tří základních metod. První metoda je založená na odhadu zpoždění  $\tau$  v časové oblasti (TDOA CC), tedy na výpočtu vzájemných korelačních funkcí. Druhá velice často využívaná metoda je založena na odhadu časového

zpoždění  $\tau$  ve frekvenční oblasti (*TDOA MSC*). Základní princip metody TDOA MSC je založen na výpočtu podobnosti daných signálů v jednotlivých frekvenčních pásmech (*případně na jednotlivých frekvencích*). Poslední metoda ve skupině TDOA je založena na tvarování směrových charakteristik. Tato metoda (*TDOA DAS*) využívá směrovou charakteristiku naměřenou pomocí rotačního akustického měniče, nebo pomocí statického pole akustických měničů. V případě měření směrové charakteristiky statickým polem nahrazujeme rotaci definovaným zpožděním signálu. Z naměřené směrové charakteristiky se posléze určí směr  $\phi$  s největším ziskem a v tomto směru se nachází i zdroj zvuku.

### 2.2.1. TDOA CC – Cross Correlation

Zaměřovací metoda TDOA CC využívá pro určení časového zpoždění  $\tau$  výpočet vzájemné korelační funkce  $R_{xy}(\tau)$ . Výstupní hodnoty z této funkce odpovídají sumě/integrálu součinů dvou vstupních signálů, které jsou vzájemně posunuty o konkrétní časový úsek  $\tau$  (*v případě diskrétních signálů o počet vzorků k*). Vzájemnou korelační funkci  $R_{xy}(\tau)$  a autokorelaci  $R_{xx}(\tau)$  definujeme pomocí následujících vztahů:

Vzájemná korelační funkce spojitých signálů  $x(t)$  a  $y(t)$ :

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} x^*(t) \cdot y(t + \tau) dt \quad (29)$$

Vzájemná korelační funkce diskrétních signálů  $x[n]$  a  $y[n]$ :

$$R_{XY}[k] = \sum_{n=-\infty}^{\infty} x^*[n] \cdot y[n + k] \quad (30)$$

Autokorelační funkce spojitého signálu  $x(t)$ :

$$R_{XX}(\tau) = \int_{-\infty}^{\infty} x^*(t) \cdot x(t + \tau) dt \quad (31)$$



Autokorelační funkce diskrétního signálu  $x[n]$ :

$$R_{xx}[k] = \sum_{n=-\infty}^{\infty} x^*[n] \cdot x[n+k] \quad (32)$$

Vychýlený odhad vzájemné korelační funkce (*MATLAB funkce xcorr - biased*):

$$R_{xy}[k] = \frac{1}{N} \sum_{n=0}^{N-1-|k|} x^*[n] \cdot y[n+k] \quad (33)$$

Nevychýlený odhad vzájemné korelační funkce (*MATLAB funkce xcorr - unbiased*):

$$R_{xy}[k] = \frac{1}{N-|k|} \sum_{n=0}^{N-1-|k|} x^*[n] \cdot y[n+k] \quad (34)$$

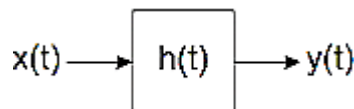
Výpočet vzájemné korelační funkce ze spektrálních charakteristik:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} X(f) \cdot Y^*(f) \cdot e^{j2\pi f\tau} df = \int_{-\infty}^{\infty} S_{xy}(f) \cdot e^{j2\pi f\tau} df \quad (35)$$

- $X(f)$ ,  $Y(f)$  – frekvenční spektra (*komplexní spektrum*) signálů  $x(t)$  a  $y(t)$

### Šumové vlastnosti korelačních funkcí (vztahy převzaty z [17]):

Nyní si odvodíme, jak šumové pozadí ovlivní výsledek vzájemné korelační funkce  $R_{xy}(\tau)$ . Představme si situaci, kdy signál ze zdroje zvuku prochází nedisperzním prostředím (*disperzní prostředí = prostředí ve kterém je rychlost šíření vlny závislá na její frekvenci*) o známé impulsové odezvě (*obrázek č. 11*), a následně se k tomuto výstupnímu signálu přičte aditivní šum (*obrázek č. 13*).



Obr. 11: Model s impulsní odezvou

$$h(j\omega) = K_0 e^{j\omega \tau_0} \quad (36)$$

- $h(j\omega)$  – přenosová funkce soustavy (*frekvenční oblast*)

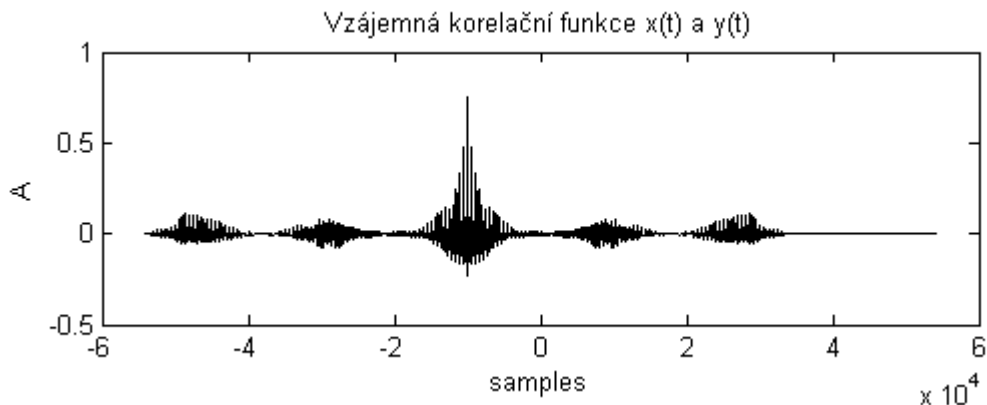
Signál po průchodu nedisperzním prostředím je utlumen (*konstanta*  $K_0$ ) a zpožděn o čas  $\tau_0$ :

$$y(t) = K_0 x(t - \tau_0) \quad (37)$$

Vzájemná korelační funkce  $R_{xy}(\tau)$  mezi vstupním  $x(t)$  a výstupním signálem  $y(t)$ :

$$\begin{aligned} R_{xy}(\tau) &= E[x(t)y(t + \tau)] = \\ &= E[x(t)K_0 x(t - \tau_0 + \tau)] = \\ &= K_0 E[x(t)x(t - \tau_0 + \tau)] = \\ &= K_0 R_{xx}(\tau - \tau_0) \end{aligned} \quad (38)$$

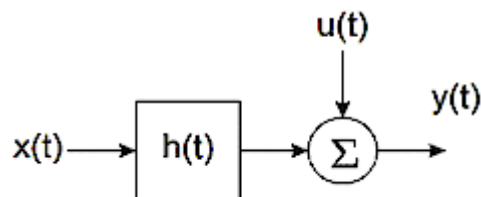
- $E[\ ]$  – operátor střední hodnoty



Obr. 12: *Vzájemná korelační funkce se zpožděním a útlumem*

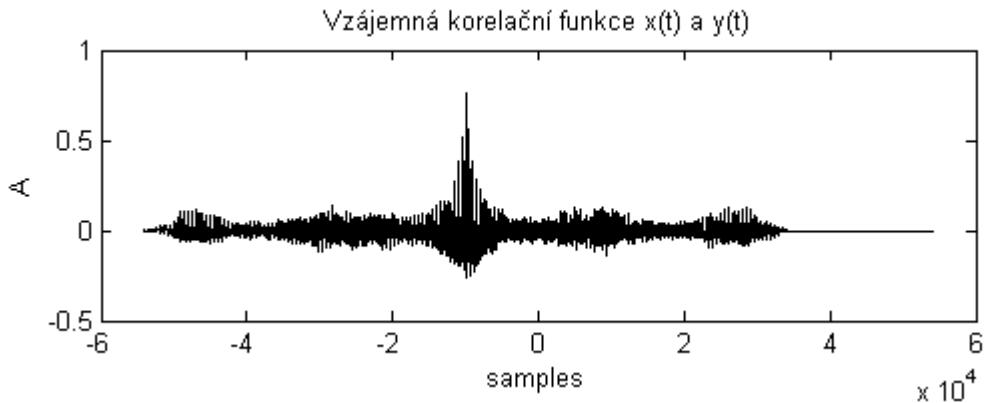
Je patrné, že výsledná vzájemná korelační funkce  $R_{xy}(\tau)$  mezi vstupním signálem  $x(t)$  a výstupním signálem  $y(t)$ , je rovna autokorelační funkci signálu  $x(t)$  vynásobenou konstantou  $K_0$  a posunutou o  $\tau_0$ , kde  $\tau_0$  odpovídá časovému zpoždění mezi vysláním  $x(t)$  a přijmutím signálu  $y(t)$ .

Předchozí odvození využijeme pro výpočet vzájemné korelační funkce  $R_{xy}(\tau)$  pro případ, kdy se k výstupnímu signálu  $y(t)$  přičte šum  $u(t)$ .



Obr. 13: *Model s aditivním šumem*

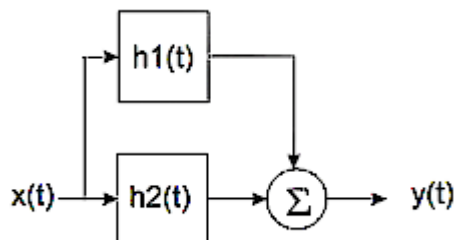
$$\begin{aligned}
R_{xy}(\tau) &= E[x(t)y(t+\tau)] = \\
&= E[x(t)[K_0x(t-\tau_0+\tau)+u(t)]] = \\
&= K_0R_{xx}(\tau-\tau_0) + E[x(t)u(t+\tau)]
\end{aligned}
\tag{39}$$



Obr. 14: Vzájemná korelační funkce se zpožděním, útlumem a šumem

V tomto případě se vzájemná korelační funkce  $R_{xy}(\tau)$  skládá ze dvou částí. První část je modifikovaná (útlum a časové zpoždění) autokorelační funkce signálu  $x(t)$  a druhá část odpovídá vzájemné korelační funkci signálů  $x(t)$  a šumu  $u(t)$ . V případě kdy signál  $x(t)$  a šum  $u(t)$  jsou vzájemně nekorelované, budou následně v korelační funkci  $R_{xy}(\tau)$  dobře rozeznatelné. V našem případě uvažujeme za šum veškeré nevhodné signály (šum pozadí, další výrazný akustický zdroj, indukovaný šum, tepelný šum, aj.).

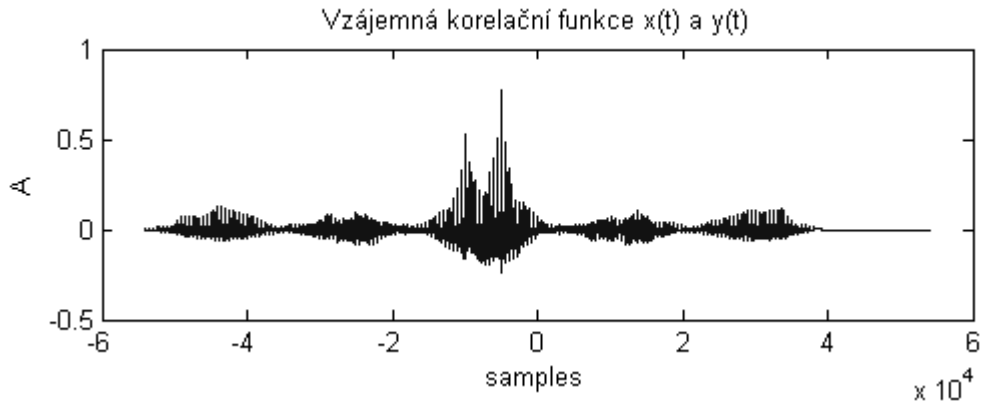
Při šíření zvukové vlny od vysílače k přijímači často dochází k tomu, že se vlna nešíří pouze cestou přímou, ale i cestami nepřímými (vlivem odrazů). To má za následek interferenci mezi vlnami přímými a odraženými. Nyní si odvodíme, jak vícecestné šíření signálu ovlivňuje vzájemnou korelační funkci  $R_{xy}(\tau)$  (pro jednoduchost předpokládáme pouze jeden odraz –  $y_1(t)$  je odražená vlna,  $y_2(t)$  je přímá vlna).



Obr. 15: Model vícecestného šíření

$$y(t) = y_1(t) + y_2(t) = K_1x(t - \tau_1) + K_2x(t - \tau_2) \quad (40)$$

$$R_{xy}(\tau) = K_1R_{xx}(\tau - \tau_1) + K_2R_{xx}(\tau - \tau_2) \quad (41)$$



Obr. 16: Vzájemná korelační funkce (více cest)

Vzájemná korelační funkce  $R_{xy}(\tau)$  je nyní součtem jednotlivých autokorelačních funkcí signálů  $x(t)$ , které se liší amplitudou  $K_{1,2}$  a zpožděním  $\tau_{1,2}$ . Na obrázku č. 16 jsou velice dobře rozeznatelná dvě maxima. První maximum odpovídá časovému zpoždění přímého signálu  $\tau_1$  a druhé odpovídá časovému zpoždění odraženého signálu  $\tau_2$ . V tomto případě jsou maxima velice dobře rozlišitelná. Avšak může se stát i to, že nebude možné je takto jednoduše rozlišit (tzv. široká korelace). Ostrost maxim závisí na frekvenčních vlastnostech signálů, ze kterých se počítá vzájemná korelační funkce  $R_{xy}(\tau)$ . Jedná-li se o širokopásmové signály, budou tato maxima ostrá (tzv. úzká korelace) a naopak. V praxi se využívá vztah pro výpočet minimální šířky pásma signálu, který vychází z poměru amplitud a časových zpoždění jednotlivých korelačních maxim:

$$B\Delta\tau \approx 3 \frac{R_{xy}(\tau_1)}{R_{xy}(\tau_2)} \quad (42)$$

- B – šířka pásma
- $\Delta\tau$  – časové zpoždění mezi jednotlivými maximy
- $R_{xy}(\tau_1)/R_{xy}(\tau_2)$  – poměr korelačních maxim

## Základní vlastnosti a podmínky, za kterých je možné využít metodu TDOA CC:

- Využitelná pouze pro nedisperzní prostředí (v disperzním prostředí dochází k deformaci tvaru vlnového balíku, kvůli frekvenční závislosti rychlosti vlny).
- Jsme schopni zaměřit pouze jeden zdroj zvuku.
- V případě nekorelovaného signálu  $x(t)$  a šumu  $u(t)$  lze tyto dva signály úspěšně rozlišit.
- Pro možnost rozlišit korelační maxima (při vícecestné šíření) je nutné zpracovávat signál s minimální šířkou frekvenčního pásma.

## TDOA GCC – Generalized Cross Correlation

Generalized cross correlation (v překladu obecná křížová/vzájemná korelace) s váhovacími funkcemi je jednou z nejčastěji využívaných metod pro akustické zaměřování. V této kapitole si vysvětlíme princip výpočetního algoritmu a rozdíl proti klasické (neváhované) korelační funkci  $R_{xy}(\tau)$ . Algoritmus GCC je založen na výpočtu vzájemné korelační funkce ve spektrální oblasti:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} X(f) \cdot Y^*(f) \cdot e^{j2\pi f\tau} df = \int_{-\infty}^{\infty} S_{xy}(f) \cdot e^{j2\pi f\tau} df \quad (43)$$

- $X(f)$  – komplexní frekvenční spektrum signálu  $x(t)$
- $Y(f)$  – komplexní frekvenční spektru signálu  $y(t)$

Signály  $X(f)$  a  $Y(f)$  jsou Fourierovy transformace (frekvenční spektrum) signálů  $x(t)$  a  $y(t)$ . Symbol  $*$  označuje komplexní sdružení. Výpočet  $R_{xy}(\tau)$  pomocí *fft* (fast Fourier transform) je efektivnější než výpočet  $R_{xy}(\tau)$  v časové oblasti. Především při implementaci v signálových procesorech (DSP) je tato implementace velice žádaná (z důvodu rychlých algoritmů výpočtu *fft*). Váhování korelační funkce  $R_{xy}(\tau)$  probíhá ve spektrální oblasti. Existuje mnoho variant váhovacích funkcí  $\psi(f)$ , které mají různé vlastnosti (vlastnosti váhovacích funkcí budou v rámci měření porovnávány):

- **Rothův procesor RP** (*Roth Processor*):

$$\psi(f) = \frac{1}{X(f) \cdot X^*(f)} = \frac{1}{S_{xx}(f)} \quad (44)$$

- **Vyhlazená koherenční transformace SCOT** (*Smoothed Coherence Transform*):

$$\psi(f) = \frac{1}{\sqrt{S_{xx}(f) \cdot S_{yy}(f)}} \quad (45)$$

- **Fázová transformace PHAT** (*Phase Transform*):

$$\psi(f) = \frac{1}{|X(f) \cdot Y^*(f)|} = \frac{1}{|S_{xy}(f)|} \quad (46)$$

Výsledný vztah pro výpočet vzájemné korelační funkce  $R_{xyGCC}(\tau)$  s váhovací funkcí  $\psi(f)$ , vypočtené ze spektrálních charakteristik signálů  $x(t)$  a  $y(t)$ , můžeme zapsat takto:

$$R_{xyGCC}(\tau) = \int_{-\infty}^{\infty} \psi(f) \cdot X(f) \cdot Y^*(f) \cdot e^{j2\pi f\tau} df = \int_{-\infty}^{\infty} \psi(f) \cdot S_{xy}(f) \cdot e^{j2\pi f\tau} df \quad (47)$$

### **Základní vlastnosti a podmínky, za kterých je možné využít metodu TDOA GCC:**

- *Váhovací funkce RP a SCOT je vhodné použít při zaměřování v prostředí s velkým šumovým pozadím a vysokou koncentrací odražených vln.*
- *Efektivnější výpočet než výpočet v časové oblasti.*
- *Relativně jednoduchá implementace v DSP (využití fft).*
- *Při použití váhovací funkce má takto vypočtená korelace ostřejší maxima, než korelace bez váhovací funkce.*

## 2.2.2. TDOA MSC – Magnitude Squared Coherence

Zaměřovací metody TDOA CC a TDOA DAS jsou metody, kde se zpoždění  $\tau$  vypočítává v časové oblasti. U metody TDOA MSC se toto zpoždění vypočítává z frekvenčních charakteristik, a to konkrétně z jejich výkonových hustot. Zde se využívá charakteristických vlastností MSC funkce. MSC funkci definujeme dle následujících vztahů:

$$C_{XY}(f) = \frac{S_{XY}(f)}{\sqrt{S_{XX}(f) \cdot S_{YY}(f)}} \quad (48)$$

- $S_{XY}(f)$  – vzájemná spektrální výkonová hustota signálu  $x(t)$  a  $y(t)$
- $S_{XX}(f), S_{YY}(f)$  – spektrální výkonová hustota signálu  $x(t)$  a  $y(t)$

$C_{XY}$  je vzájemná koherenční funkce dvou signálů  $x(t)$  a  $y(t)$ . Výstupní hodnoty z koherenční funkce jsou obecně komplexní a určují vzájemnou podobnost vstupních signálů v jednotlivých frekvenčních pásmech. V algoritmech pro výpočet časového zpoždění  $\tau$  (z frekvenční oblasti) se často využívá modifikovaná vzájemná koherenční funkce MSC. Jedná se o koherenční funkci, která je normovaná a tudíž výstupní hodnoty jsou v rozmezí  $\langle 0, 1 \rangle$ .

$$MSC(f) = \frac{|S_{XY}(f)|^2}{S_{XX}(f) \cdot S_{YY}(f)} \quad (49)$$

MSC opět obsahuje informaci o vzájemné frekvenční podobnosti vstupních signálů. Není však komplexní, a proto z ní nelze přímo vypočítat časové zpoždění  $\tau$ . Časové zpoždění  $\tau$  vypočteme z imaginární části vzájemné výkonové spektrální hustoty (CPSD). Výkonové spektrální hustoty použité pro výpočet MSC jsou vyhlazené odhady těchto charakteristik. Pro diskrétní signál  $x[n]$  a  $y[n]$  definujeme vyhlazené spektrální výkonové hustoty dle následujících vztahů:

$$S_{XY}(f) = \frac{1}{L} \sum_{k=1}^L X_k(f) \cdot Y_k^*(f) \quad (50)$$

$$S_{XX}(f) = \frac{1}{L} \sum_{k=1}^L |X_k(f)|^2 \quad (51)$$

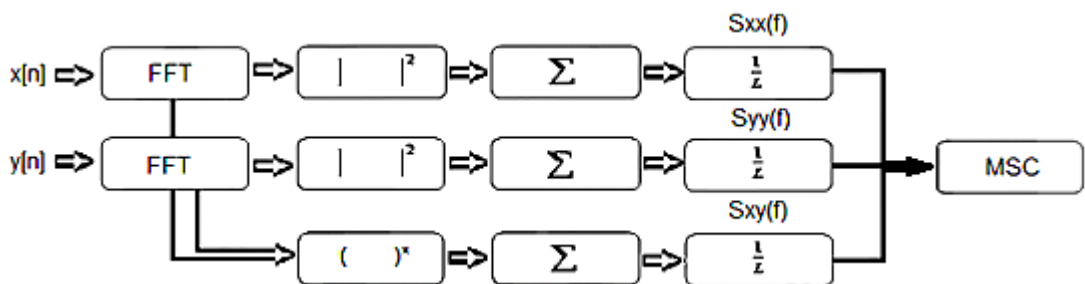
$$S_{YY}(f) = \frac{1}{L} \sum_{k=1}^L |Y_k(f)|^2 \quad (52)$$

- L – počet realizací spektrálního segmentu  $X_k, Y_k$ , použitých při průměrování

### Vlastnosti funkce MSC:

- Nabývá hodnot z intervalu  $\langle 0, 1 \rangle$ .
- Udává lineární závislost vstupních signálů.
- Vypovídá o podobnosti signálů v jednotlivých frekvenčních pásmech.
- Blíží-li se k jedné, je podobnost signálů v daném frekvenčním pásmu vysoká a naopak.
- Je závislá na délce zpracovaného signálu (délce okna) a počtu segmentů.

Blokové schéma výpočtu MSC funkce pomocí rychlé Fourierovy transformace je zobrazeno na obrázku č. 17.



Obr. 17: Blokové schéma výpočtu MSC



Posun mezi signály určíme na základě stanovení frekvenční oblasti, kde MSC dosahuje maximálních hodnot (*v tomto frekvenčním pásmu je podobnost signálu nejvyšší*). Nalezením tohoto pásma máme informaci o frekvenční oblasti, ve které budeme počítat posun mezi signály. Nyní je třeba spočítat vzájemnou spektrální výkonovou hustotu  $S_{XY}(f)$ , která je komplexní funkcí a tudíž nese informaci i o fázi. Posun mezi signály  $g_o(f)$  zde odpovídá derivaci její fázové charakteristiky  $\phi_{xy}(f)$  ve vymezeném frekvenčním pásmu. (*výpočet  $g_o(f)$  převzat z [17]*)

$$[A, f_\phi] = \max\{MSC(f)\} \quad (53)$$

$$g_o(f) = -\frac{d\phi_{xy}(f_\phi)}{df} \quad (54)$$

### **Základní vlastnosti a podmínky, za kterých je možné využít metodu TDOA MSC:**

- *Za určitých podmínek (odlišnost spektrálních vlastností zdrojů) jsme schopni lokalizovat více zdrojů zvuku.*
- *Využitelná pro disperzní i nedisperzní prostředí.*
- *Nevýhodná pro zpracování širokopásmových signálů v reálném čase.*
- *Dobrá rozlišitelnost maxim určujících posun mezi signály.*

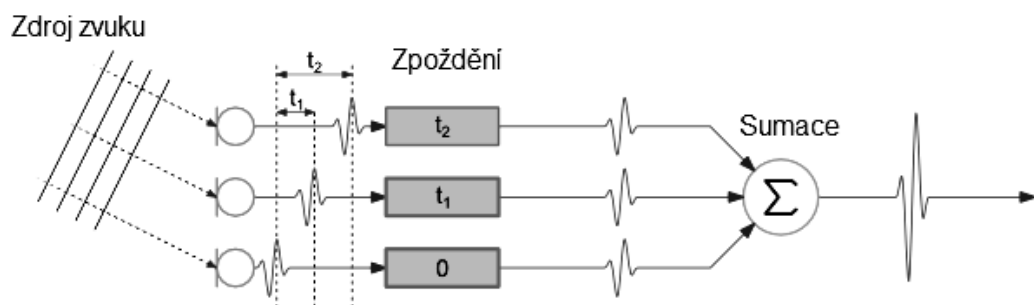
### **2.2.3. Metody s tvarováním směrových charakteristik**

Metody založené na tvarování směrových charakteristik byly jedny z prvních využívaných akustických zaměřovacích metod. Pracují na relativně jednoduchém principu, kdy záměrně měníme vlastnosti parametrů nahraných zvukových signálů tak, aby tato změna odpovídala určitému časovému zpoždění  $\tau$ . Pomocí metod, založených na tvarování směrové charakteristiky, jsme schopni lokalizovat více zdrojů zvuku.

Nevýhodou těchto metod je relativně velká náročnost na výpočetní výkon a tomu odpovídající doba výpočtu. Z tohoto důvodu nejsou tyto metody vhodné pro lokalizaci rychle se pohybujících zdrojů zvuku. Tyto metody byly využívány především v době analogového zpracování signálů. Nyní v době digitálního zpracování signálu jsou z důvodu lepší efektivity a přesnosti použity jiné metody (*TDOA CC/MSD, metody založené na spektrálním odhadu vysokého rozlišení*).

### Tvarovač DAS (Delay and Sum)

Základním funkčním blokem těchto metod je tvarovač směrových charakteristik. Nejčastěji využívaný je tvarovač DAS (*v překladu zpozdí a sečti*). Jak už z názvu vyplývá, při zpracování signálu dochází ke zpoždění a součtu nahraných signálů. Tvarování charakteristik můžeme rozdělit právě na tyto dvě základní fáze zpracování signálů (*zpoždění a sumaci*). Základní blokové schéma je znázorněno na obrázku č. 18.



Obr. 18: Blokové schéma tvarovače DAS

Na obrázku č. 18 jsou tři akustické měniče, na které dopadá zvuková vlna. V případě že zvuková vlna dopadá kolmo k ose lineárního mikrofonního pole, je mezi jednotlivými signály nulové časové zpoždění  $\tau$ . V případě že vlna dopadá v jiném směru (*než v kolmém*), je mezi nahranými signály určité časové zpoždění  $\tau$ . Toto časové zpoždění se snažíme vykompenzovat umělým časovým zpožděním. Pokud by bylo splněno nulové zpoždění mezi všemi signály, dosáhli bychom po sumaci maximálního zisku, který hledáme.

Pro stanovení maximálního zisku a tudíž i směru zvukového zdroje, je třeba naměřit nebo dopočítat směrovou charakteristiku pole akustických přijímačů pro

všechna reálná časová zpoždění  $\tau$ . Pokud máme rotační akustický měnič, určíme směrovou charakteristiku tak, že si v konkrétních úhlových bodech zaznamenáváme úroveň signálu. Nevýhoda tohoto přístupu je ta, že po celou dobu měření (*jedné rotace*) je nutné udržet stejné podmínky (*neměnné vlastnosti zdroje zvuku*). Proto se častěji využívá statické pole akustických měničů, kde se signál zaznamená a posléze uměle zpožďuje a sčítá.

Předpokládejme pole s  $N$  mikrofony, kde  $n$  je číslo mikrofonu ( $n=1,2,3,\dots, N$ ) [14].

$$F_n(\tau) = (n-1)\tau \quad (55)$$

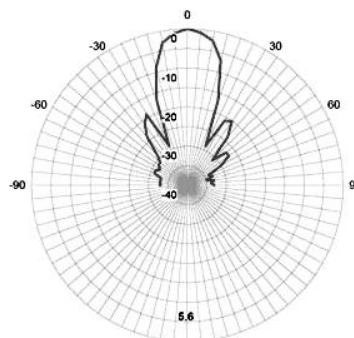
Kde  $\tau$  je časové zpoždění mezi sousedními mikrofony, které se vypočítává triangulační metodou. Pro daný úhel  $\theta$  se určí, o kolik vzorků mají být signály zpožděny.

$$x_n[k] = x_n[k + F_n(\tau)] \quad (56)$$

- $x_n[k]$  – uměle posunutý signál z mikrofonu  $n$

$$y[k] = \frac{1}{N} \sum_{n=1}^N x_n[k] \quad (57)$$

Výstupní signál z tvarovače DAS je součtem vzájemně posunutých signálů o celistvé násobky časového zpoždění. Takto se vypočte celková suma vzájemně posunutých signálů pro každý bod ve směrové charakteristice a poté se vyhodnotí její maximální hodnota. V bodě (*amplituda  $A$ , úhel  $\theta$* ) maximální hodnoty směrové charakteristiky došlo k vykompenzování časového zpoždění  $\tau$  sčítaných signálů.



Obr. 19: Směrová charakteristika

Na obrázku **č. 19** je maximální hodnota směrové charakteristiky v úhlové pozici  $0^\circ$ . Tento směr označíme jako směr lokalizovaného zdroje zvuku.

Směrová charakteristika DAS tvarovače (s *lineárním mikrofonním polem*) je vypočtena pro úhlový rozsah  $0 - 180^\circ$ , druhá polovina směrové charakteristiky s ní bude vždy symetrická. Symetrie dílčích kusů směrových charakteristik vychází ze stranové neurčitosti, která nastává při měření s lineárním polem mikrofonů. Zde pro každou platnou hodnotu vypočteného časového zpoždění  $\tau$  náleží právě dva úhlové směry. Při návrhu mikrofonního pole je dále nutné dodržet podmínku maximální vzdálenosti mikrofonů  $d_{max}$ .

### **Základní vlastnosti a podmínky, za kterých je možné využít metodu TDOA DAS:**

- *Možnost zaměřování více akustických zdrojů najednou.*
- *Využitelná pouze pro nedisperzní prostředí.*
- *Relativně vysoká výpočetní náročnost a k tomu odpovídající výpočetní čas (proto není vhodná pro lokalizaci rychle se pohybujícího zdroje zvuku).*
- *Zpracování signálu v časové oblasti (obdobná metoda, jako vzájemná korelace), tudíž méně výrazná korelační maxima než u metod TDOA MSC a GCC.*

### **2.3. Metody využívající spektrální odhad vysokého rozlišení**

Jedná se o metody pracující ve spektrální oblasti. Metody využívající spektrálního odhadu vysokého rozlišení (*neboli High Resolution Spectral Estimation, dále jen HRSE*) dělíme na tři základní skupiny:

- **AR** – Autoregressive Modeling (*autoregresivní modelování*)
- **MV** – Minimum Variance (*minimální rozptyl*)
- **MUSIC** – Multiple Signal Classification (*vícenásobná klasifikace signálu*)

Tyto metody využívají spektrální fázovou korelační matici, která je odvozená pro všechny elementy v poli. Pokud není přesně zadána, určuje se na základě přijímaných

dat, konkrétně souborem průměrování signálu přes interval, ve kterém je zvukový signál a šum považován za stacionární [18].

Metoda AR je limitována pouze pro lokalizaci vzdálených zdrojů zvuku. Metody MV a MUSIC jsou použitelné pro lokalizaci blízkých i vzdálených zdrojů zvuku. Metody HRSE jsou velice přesné a účinné i v prostředí s velkým rušením (*dobré šumové vlastnosti*). Často jsou využívány pouze pro zaměřování signálů s úzkou šířkou frekvenčního pásma. Důvodem je rostoucí výpočetní náročnost při růstu šířky pásma zaměřovaných signálů. Vzhledem k tomu, že je nutné prohlásit zvukový signál a šum za stacionární a dále je nutné znát charakter zdroje zvuku, nejsou tyto metody vhodné pro detekci řečových signálů.

**Základní vlastnosti a podmínky, za kterých je možné využít metody se spektrálním odhadem vysokého rozlišení:**

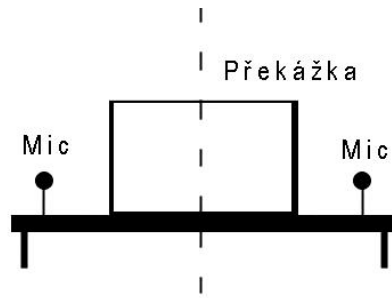
- *Možnost zaměřování více akustických zdrojů (spektrální odlišnost).*
- *Závislost výpočetní náročnosti na šířce pásma.*
- *AR vhodná pouze pro zvukové zdroje v blízkém poli.*
- *MV a MUSIC vhodná pro zvukové zdroje v blízkém i vzdáleném poli.*

## **2.4. Metoda s akustickou překážkou**

Tato metoda využívá tvorby akustického stínu za překážkou. Akustický stín se za překážkou nevytváří vždy, záleží na poměru velikosti  $d$  překážky a vlnové délky  $\lambda$ . Zvuková vlna se na překážce může chovat takto:

- *Odraz ( $d \gg \lambda$  - rozměr překážky je řádově větší než vlnová délka)*
- *Ohyb ( $d \approx \lambda$  - rozměr překážky je srovnatelný s vlnovou délkou)*
- *Překážka neovlivňuje šíření vlny ( $d \ll \lambda$ )*

Zkonstruovaný zaměřovač pro metodu s akustickou překážkou je na obrázku č. 20 (obrázek převzat z mé práce [27]).



Obr. 20: Návrh a reálná konstrukce zaměřovače

U této metody je nutné dodržet podmínku pro odraz zvukové vlny od překážky, protože právě tehdy se za překážkou vytváří akustický stín. V případě že by podmínka dodržena nebyla, můžeme za překážkou naměřit stejné nebo za určitých podmínek i větší hodnoty akustického tlaku, než před překážkou.

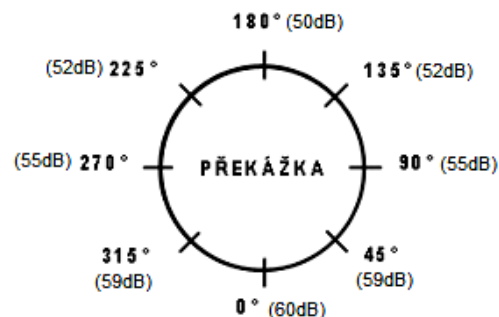
Rozlišovací schopnost akustického zaměřovače pro tuto metodu je přímo závislá na počtu použitých mikrofonů.

$$\text{rozlišení} = \frac{360}{N} [^\circ] \quad (58)$$

- N – počet mikrofonů

Proto je rozlišovací schopnost těchto akustických zaměřovačů výrazně menší, než v případě kterékoliv metody, která zde byla uvedena. Další vlastností je schopnost zaměřování pouze jednoho dominantního akustického zdroje.

Směr zdroje zvuku odpovídá mikrofonnímu páru (ležící na společné přímce) s největším rozdílem naměřených akustických tlaků, viz obrázek č. 21.



Obr. 21: Rozložení akustického tlaku okolo překážky

Největší rozdíl akustických tlaků je mezi mikrofony na pozicích  $0^\circ$  a  $180^\circ$  ( $\Delta L_p = 10\text{dB}$ ). To znamená, že zdroj zvuku se nachází ve směru od středu k úhlové značce  $180^\circ$ , nebo od středu k úhlové značce  $0^\circ$ . Který směr je správný, rozhodne znaménko rozdílu akustických tlaků (*zde se zdroj zvuku nachází ve směru od středu k úhlové značce  $0^\circ$* ).

**Základní vlastnosti a podmínky, za kterých je možné využít metodu s akustickou překážkou:**

- Možnost zaměřování pouze jednoho dominantního akustického zdroje.
- Rozlišení zaměřovače je závislé na počtu použitých mikrofonů.
- Nenáročná na výpočetní výkon.
- Nutnost dodržet podmínku pro odraz zvukové vlny.

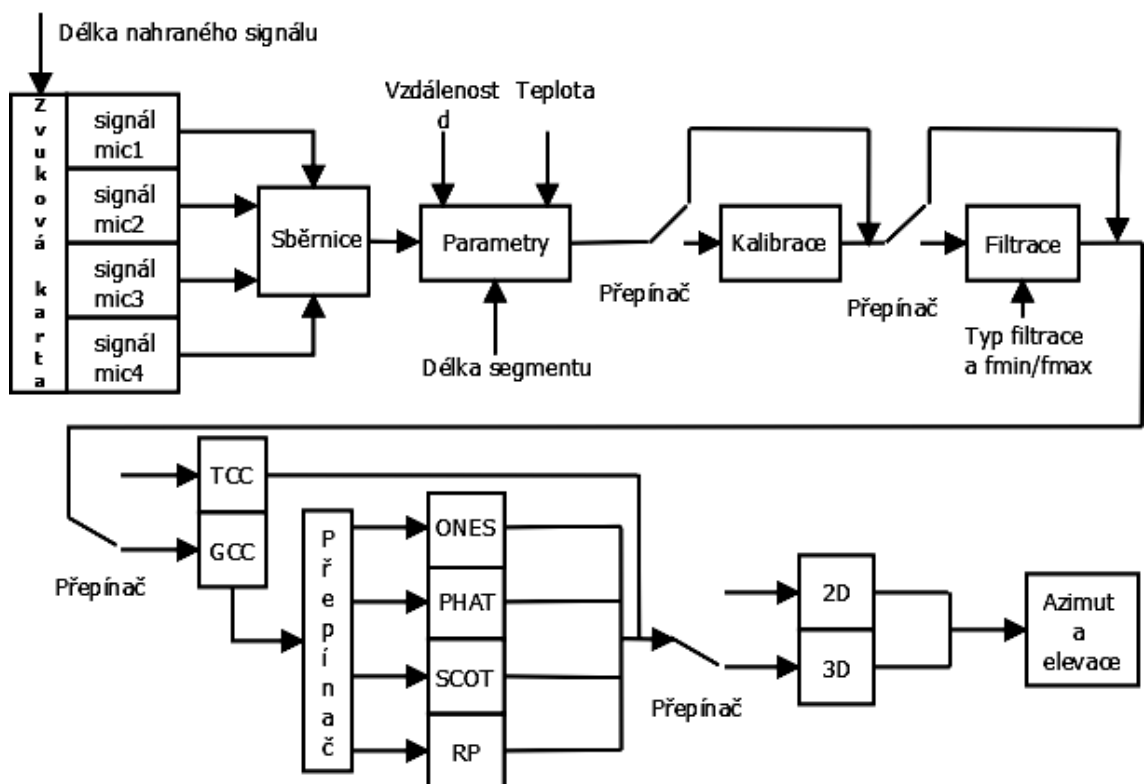
**Souhrn základních vlastností jednotlivých lokalizačních metod:**

metoda	ŠP signály	ÚP signály	Zaměřování pohyblivých zdrojů	Zaměřování více zdrojů zvuků	Závislost výpočetní náročnosti na šířce pásma	Zdroj zvuku (vzdálenost)	Ostatní
TDOA CC	ano	ne	ano	ne	ne	vzdálený	málo ostrá maxima
TDOA GCC	ano	ne	ano	ne	ne	vzdálený	výpočetně nenáročné a přesné
TDOA MSC	ne	ano	ano	ano	ano	vzdálený	dobře rozlišitelná maxima
TDOA DAS	ano	ano	ano	ano	ne	vzdálený	výpočetní náročnost
HRSE AR	ne	ano	ano	ano	ano	vzdálený	s rostoucí šířkou pásma roste výpočetní náročnost
HRSE MV	ne	ano	ano	ano	ano	vzdálený i blízký	
HRSE MUSIC	ne	ano	ano	ano	ano	vzdálený i blízký	
Metoda s akustickou překážkou	ano	ano	ano	ne	ne	vzdálený i blízký	nízká rozlišovací schopnost
ÚP - úzkopásmové signály							
ŠP - širokopásmové signály							

Tab. 3: Vlastnosti lokalizačních metod

### 3. Číslicové zpracování signálů

V této kapitole se zaměřím na nejdůležitější principy zpracování číslicového signálu, které jsem ve výpočetním algoritmu použil. Dále zobrazím blokové schéma výpočetního algoritmu a popíši jednotlivé bloky. Závěr kapitoly obsahuje popis grafického rozhraní (*dále jen GUI*) programu pro lokalizaci zdroje zvuku pomocí metody GCC a TCC. Kvůli snazší orientaci čtenáře v textové části diplomové práce a souvisejícím programu popíši výpočetní algoritmy pomocí funkcí z prostředí MATLAB (*\* takto zapsané symboly označují násobení vektorů. Násobí se pouze prvky vektorů na shodných pozicích.*).



Obr. 22: Blokové schéma výpočetního algoritmu

Základní bloky schématu jsou kalibrace, filtrace a výpočetní algoritmy TCC a GCC (*blok GCC a TCC vrací odhad časového zpoždění*). Z časových zpoždění je následně vypočítán úhel azimutu a elevace. Bloky kalibrace, filtrace, GCC a TCC jsou vytvořeny jako jednotlivé funkce. Toto rozdělení usnadňuje testování výpočetního algoritmu a umožňuje tak snadnou implementaci funkcí do jiných výpočetních algoritmů. Na začátku programu využívám funkce *pa\_wavrecord.m* a *pa\_wavplayrecord.m*. Jsou to funkce z balíčku, který slouží k nahrávání a přehrávání signálů pomocí zvukových karet



pracujících s ASIO, MME nebo WDM ovladači. Tyto funkce jsou převzaty z [8]. Algoritmus následně pokračuje k podmínce rozhodující o použití kalibrace.

### 3.1. Kalibrace

Kalibrace slouží k vyrovnání odlišných frekvenčních vlastností jednotlivých mikrofonů a jejich předzesilovačů. Kalibrace je provedena ve frekvenční oblasti vynásobením spektra signálu kalibrační křivkou příslušné dvojice mikrofon-předzesilovač. Kalibrační křivky můžeme nalézt v textových souborech *calibr1.txt* a *calibr10.txt*. Označení *fr1* a *fr10* vyjadřuje frekvenční rozlišení kalibrační křivky ( $fr10 = 10\text{Hz}$ ,  $fr1 = 1\text{Hz}$ ). Snížením frekvenčního rozlišení snížíme výpočetní nároky, a tudíž i výpočetní čas a přesnost volby frekvenčních mezí pásmové propusti (*blok filtrace*).

$$f_R = \frac{f_S}{nfft} \quad (59)$$

- $f_R$  – frekvenční rozlišení
- $f_S$  – vzorkovací kmitočet
- $nfft$  – počet vzorků, ze kterého se vypočítává spektrum

Program umožňuje volbu počtu vzorků v časovém segmentu, a proto je nutné tyto segmenty doplňovat nulami na potřebnou délku *nfft*. V případě, že by délka segmentu byla zadána větší než *nfft*, program zahlásí „Wrong values“ v jeho informačním okně.

Kalibrační křivky jsou vytvořeny z naměřených přenosových frekvenčních charakteristik jednotlivých párů mikrofon-předzesilovač. Tyto charakteristiky si označme jako  $f_{Pi}$ . Kalibrační křivka pro danou dvojici  $i$  je následně počítána dle vztahu č. 60.

$$calibr_{-i} = \frac{AVf_{Pi}}{f_{Pi}} = \frac{1}{N} \sum_{i=1}^N f_{Pi} \quad (60)$$

- $AVf_{Pi}$  – průměrná přenosová charakteristika
- $f_{Pi}$  – přenosová charakteristika mikrofonu  $i$
- $calibr_{-i}$  – kalibrační křivka mikrofonu  $i$

Kalibrace se provede vynásobením modulu spektra časového segmentu a kalibrační křivky:

$$SP_i = abs(fft(S_i)) \quad (61)$$

$$Ph_i = angle(fft(S_i)) \quad (62)$$

- $S_i$  – časový segment signálu z mikrofonu  $i$
- $Ph_i$  – fázová složka spektra
- $SP_i$  – modulová složka spektra

$$SP_{iC} = SP_i \cdot *calibfr_{-i} \quad (63)$$

- $SP_{iC}$  – kalibrovaná modulová složka spektra

Návrat kalibrovaného spektra do časové oblasti provedeme zpětnou Fourierovou transformací *ifft*. K tomu je potřeba kalibrovaný modul spektra a jeho fázová složka (*fázovou složku neměníme*). Z těchto dvou složek vytvoříme opět komplexní signál, který musí být symetrický (*shodné hodnoty záporných i kladných složek spektra*). Symetrii spektra musíme dodržovat, jelikož se jedná o reálné signály.

$$S_{iC} = ifft(SP_{iC} \cdot *exp(1i * Ph_i)) \quad (64)$$

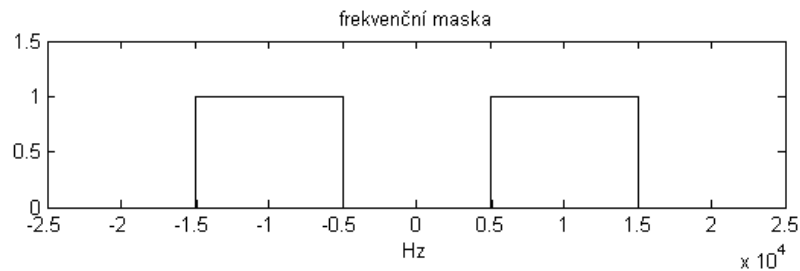
- $S_{iC}$  – kalibrovaný časový segment signálu z mikrofonu  $i$

Následujícím blokem je filtrace signálu, která pracuje na velice podobném principu jako kalibrace.

### 3.2. Filtrace

Pomocí filtrace odstraňujeme nežádoucí frekvenční části signálu. Filtraci můžeme provést v časové nebo frekvenční oblasti. Zvolil jsem filtraci ve frekvenční oblasti pro její snadné nastavení frekvenčních mezí ( $f_{min}$  a  $f_{max}$ ) a celkovou implementaci. Stejně jako u

kalibrace je nutné, aby bylo frekvenční rozlišení  $f_R$  konstantní. Z tohoto důvodu opět zařazujeme doplňování časových segmentů nulami na délku  $nfft$ . V případě  $f_R=1\text{Hz}$  je možné nastavení  $f_{min}$  a  $f_{max}$  s přesností na 1Hz. Základem filtrování ve frekvenční oblasti je frekvenční maska. Frekvenční maska je vektor obsahující pouze jedničky a nuly dle zadaných mezí  $f_{min}$  a  $f_{max}$  (pásmová propust).



Obr. 23: Frekvenční maska

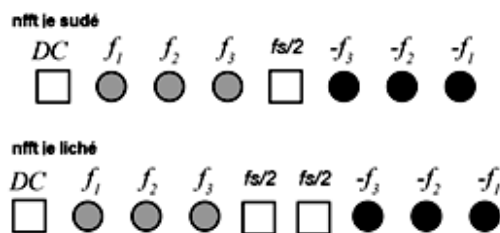
Filtrace se provede vynásobením frekvenční masky a modulu spektra časového segmentu:

$$SP_{iF} = SP_i \cdot f_{mask} \quad (65)$$

$$S_{iF} = ifft(SP_{iF} \cdot \exp(1i \cdot Ph_i)) \quad (66)$$

- $S_{iF}$  – filtrovaný časový segment signálu z mikrofonu  $i$
- $SP_{iF}$  – filtrovaná modulová složka spektra
- $f_{mask}$  – frekvenční maska

Frekvence nacházející se na pozicích nul ve frekvenční masce jsou vynulovány a frekvence na pozicích jedniček zůstávají nezměněny. Přesná podoba frekvenční masky je vždy závislá na konkrétním výpočetním systému. MATLAB a jeho funkce `fft.m` vrací vypočítané spektrum signálu v následující podobě:



Obr. 24: Funkce `fft.m` MATLAB

Při vzorkovacím kmitočtu  $f_s=48kHz$  a frekvenčním rozlišení  $f_R=1Hz$  je  $nfft=48000$  vzorků, analogicky  $f_R=10Hz$  je  $nfft=4800$  vzorků. Obě dvě hodnoty  $nfft$  jsou sudé, a proto využíváme jen jeden typ frekvenční masky. V GUI lokalizačního programu je možnost filtrování pomocí frekvenční masky, nebo kombinace spektrálního odečítání a frekvenční masky.

## Frekvenční odečítání

Frekvenční odečítání je druh filtrování ve frekvenční oblasti. Od modulové spektrální složky časového segmentu se odečte odhadnuté vyhlazené spektrum šumu. Spektrální složky, které jsou po odečtení záporné, nahrazujeme nulami (*jednocestné usměrnění*). Lze však použít i dvoucestné usměrnění. V tom případě u záporných koeficientů změnímme znaménko. Osobně jsem zvolil jednocestné usměrnění, které z hlediska SNR dosahovalo lepších výsledků.

$$SP_{FSi} = SP_i - SP_N \quad (67)$$

- $SP_N$  – vyhlazený modul spektra šumu
- $SP_{FSi}$  – filtrovaná modulová složka spektra

$$SP_{FSi}(SP_{FSi} < 0) = 0 \quad (68)$$

Po odečtení a vynulování záporných členů dochází opět k převodu komplexního frekvenčního spektra do časové oblasti:

$$S_{FSi} = \text{ifft}(SP_{FSi} \cdot \exp(1i * Ph_i)) \quad (69)$$

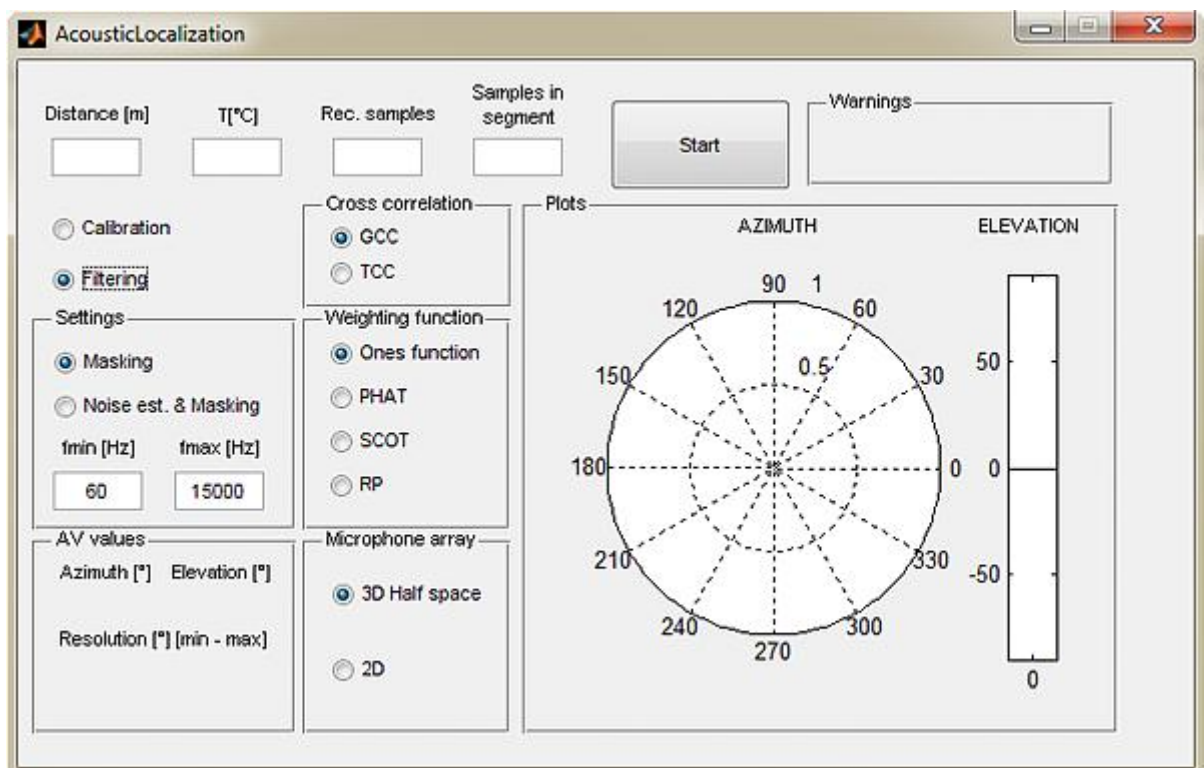
- $S_{FSi}$  – filtrovaný (*frekvenční odečítání*) časový segment signálu z mikrofonu  $i$

### 3.3. Grafické rozhraní - GUI

GUI lokalizačního programu dovoluje uživateli nastavovat velké množství parametrů. Tato volnost nastavování parametrů usnadňuje měření a následné ověřování teoretických předpokladů. Obrázek č. 25 převzat z mé práce [32].

GUI je rozděleno na několik základních bloků:

- Nastavení základní parametrů (*Distance, T, Rec. samples, Samples in segment*)
- Kalibrace mikrofonů (*Calibration*)
- Filtrování nahraných signálů (*Filtering*)
- Volba výpočtu korelační funkce (*TCC, GCC*)
- Volba váhovací funkce (*Ones function, PHAT, SCOT, RP*)
- Volba mikrofonního pole (*3D Half space, 2D*)
- Informační okno (*Warnings*)
- Tlačítko start (*spouští lokalizační program*)



Obr. 25: Grafické rozhraní lokalizačního programu

## Zásady práce s grafickým uživatelským rozhraním (GUI)

Lokalizační program obsahuje sadu podmínek hlídajících správné zadání všech parametrů. V případě, že jsou parametry zadané ve špatném formátu (*text místo čísel, záporné hodnoty a jiné*), zobrazí se v informačním okně hláška „Wrong values“. Nejsou-li vyplněny všechny požadované parametry, zobrazí se hláška „Fill all text boxes“. Hláška „Noise estimation“ oznamuje, že program právě nahrává signál pro odhad vyhlazeného spektra šumu. Po dobu zobrazení této hlášky je nutné, aby byl naprostý klid a program mohl nahrát signál obsahující pouze informaci o nežádoucím přirozeném šumu. Poslední varovnou hláškou je „No sound“. Tato hláška oznamuje uživateli, že nahraný signál má příliš malou úroveň na to, aby byl lokalizován (*nastavením parametru  $p_{vl}=0$  v kódu programu vyrušíme podmínku na hlídání úrovně signálu*).

Parametr *Distance* označuje vzdálenost mezi mikrofony  $mic_1 - mic_3$  a  $mic_2 - mic_4$ . Pro vzdálenost mezi mikrofony se zadává pouze jeden parametr. To znamená, že program je přednastaven pro geometrie, ve kterých jsou vzájemné vzdálenosti mikrofonů shodné (*geometrii mikrofonního pole popíše v kapitole Geometrie mikrofonního pole akustického zaměřovače*). Parametr *T* uvádí teplotu prostředí, ve kterém se šíří zvuková vlna. Z teploty se vypočte rychlost šíření zvuku *c*. Parametr *Rec. samples* určuje, kolik vzorků je nahráno v rámci jedné lokalizační smyčky. Parametr *Samples in segment* nastaví velikost časového segmentu, ze kterého se počítá korelační funkce.

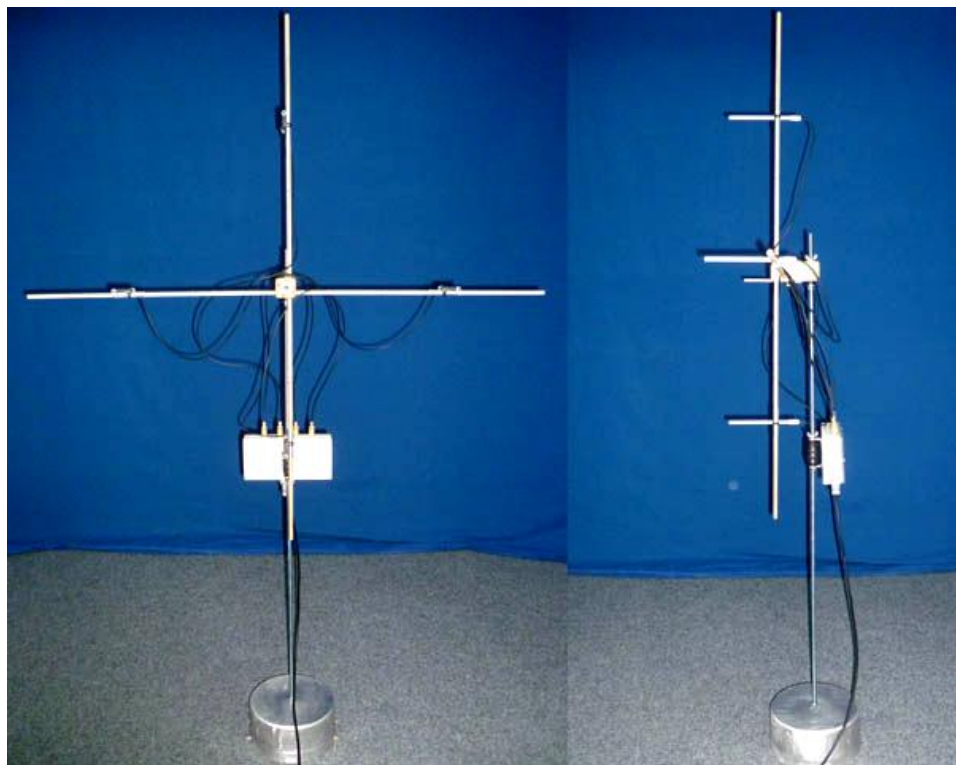
Zaškrtnutím volby *Calibration* (*kalibrace*) nebo *Filtering* (*filtrování*) zavádíme do výpočetního algoritmu příslušnou operaci. V rámci filtrování si můžeme vybrat ze dvou možností (*Masking a Noise est. & Masking*). První ze zmíněných označuje filtrování pomocí frekvenční masky, druhá filtrování pomocí frekvenční masky a frekvenčního odečítání (*u obou možností je nutné vyplnit pole  $f_{min}$  a  $f_{max}$ , která určují meze pásmové propusti*). Při výběru způsobu výpočtu korelační funkce máme opět dvě možnosti (*GCC a TCC*). *GCC* je váhovaná korelace s možností volby váhovací funkce *Ones function* (*jednotková váhovací funkce*), *PHAT*, *SCOT* nebo *RP*. *TCC* je způsob výpočtu korelace v časové oblasti pomocí vzájemného posuvu dvou vektorů (*TCC není vhodné pro dlouhé časové segmenty z hlediska výpočetního nároku*). Poslední možností je volba mikrofonního pole. Zde můžeme zvolit 3D Half space (*azimut  $\theta <0-180^\circ>$  a elevace  $\phi <+-90^\circ>$ ) nebo 2D (*azimut  $\theta <0-360^\circ>$* ).*

## 4. Konstrukce akustického zaměřovače

V této kapitole popíšeme zkonstruovaný akustický zaměřovač lokalizující v rovině a poloprostoru. V rámci popisu zaměřovače uvedu vlastnosti předzesilovačů, geometrii mikrofonního pole, parametry použité zvukové karty a mikrofonních vložek.

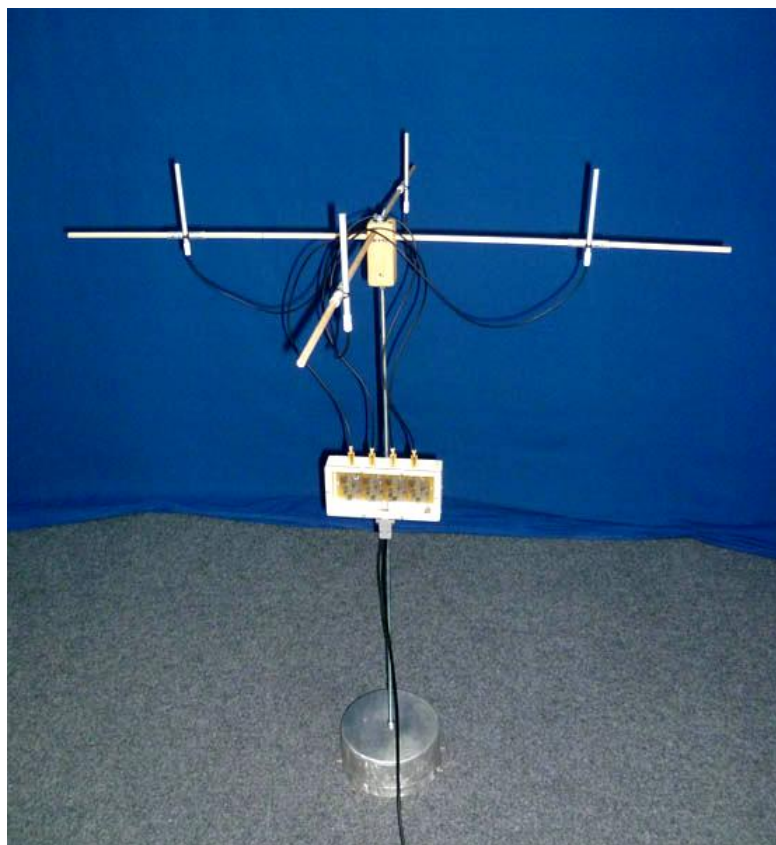
### 4.1. Geometrie mikrofonního pole akustického zaměřovače

První podmínka pro konstrukci mikrofonního pole byla schopnost zaměřování v rovině i v poloprostoru. Druhou podmínkou byla možnost měnit vzájemnou vzdálenost mikrofonů  $d$ . Jedním z hlavních omezení při návrhu mikrofonního pole bylo použití čtyř mikrofonů (z důvodu čtyř vstupů zvukové karty). Konečný návrh mikrofonního pole splňujícího dané podmínky je na obrázku č. 26 a č. 27 (obrázky č. 26 a č. 27 převzaty z mé práce [32]).



Obr. 26: Mikrofonní pole 3D poloprostor

Jednoduchou přestavbou mikrofonního pole pro zaměřování v poloprostoru jsme schopni vytvořit zaměřovač lokalizující v celé rovině (obrázek č. 27).



*Obr. 27: Mikrofonní pole 2D*

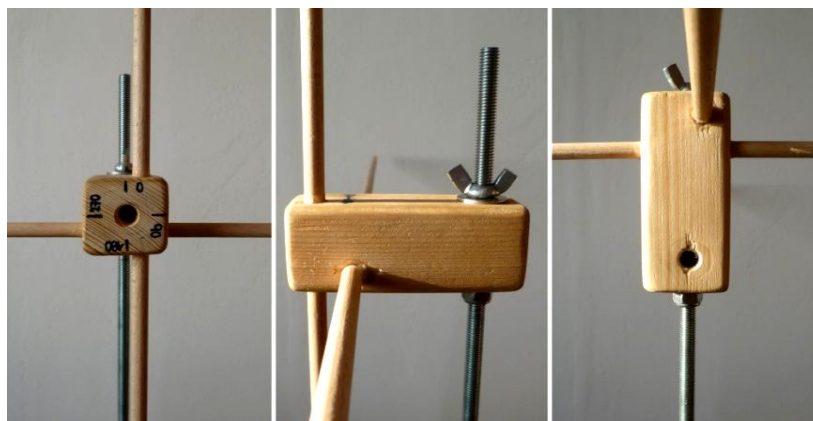
Mikrofony jsou přidělány na nosné tyče pomocí trubiček s aretací (obrázek č. 28), kterými se dá plynule nastavit vzájemná vzdálenost mikrofonů  $d$ . Obrázek č. 28 převzat z mé práce [32].



*Obr. 28: Trubička s aretací*

Konstrukční prvek, který umožňuje přestavbu mikrofonního pole, je zobrazen na obrázku č. 29 (obrázek převzat z mé práce [32]).

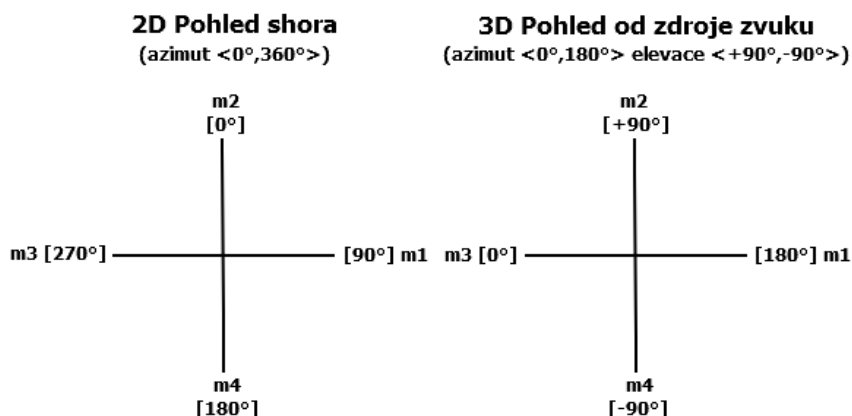




Obr. 29: Konstrukční prvek umožňující přestavbu mikrofonního pole

Při změně typu mikrofonního pole (2D, 3D Half space) není nutné měnit pozice mikrofonů. Pro konkrétní mikrofonní pole se v GUI zaměřovacího programu zvolí příslušná volba *Microphone array*.

Na obrázku č. 30 je znázorněno umístění jednotlivých mikrofonů a označení jejich úhlových pozic v mikrofonních polích.

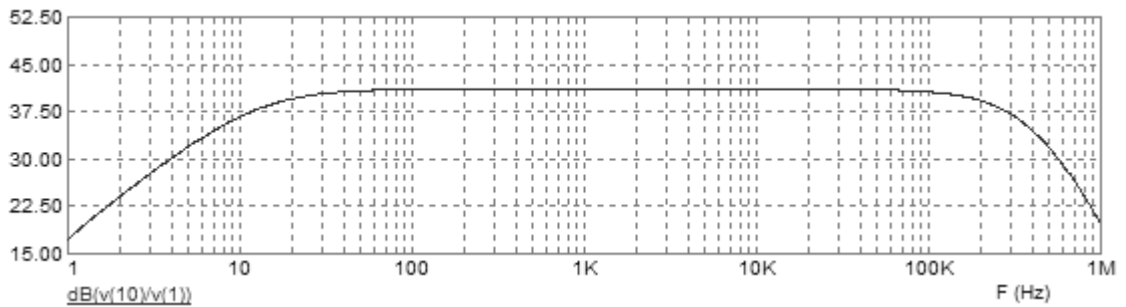


Obr. 30: Pozice mikrofonů

## 4.2. Mikrofonní předzesilovače

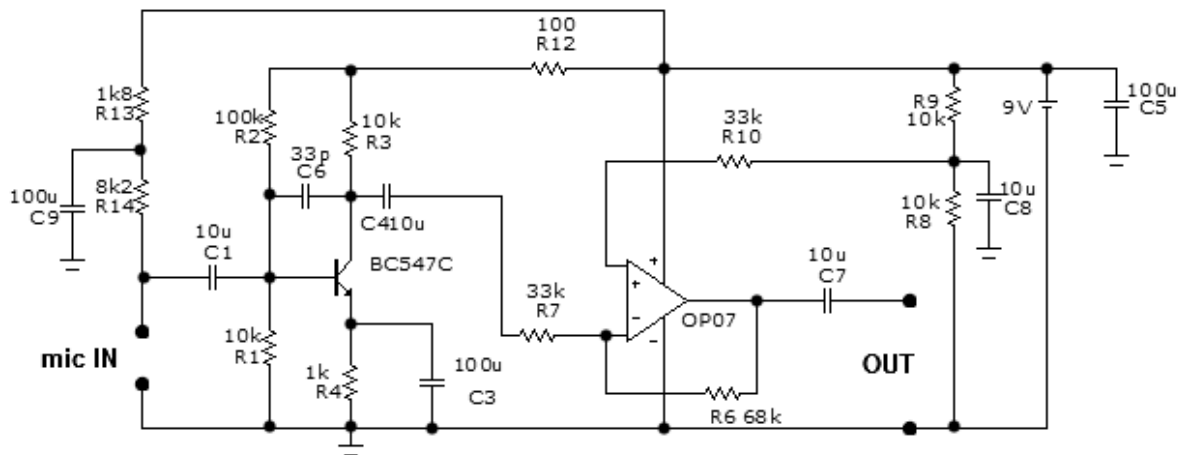
Základními požadavky na mikrofonní předzesilovače jsou shodné vlastnosti jednotlivých kanálů ( $mic_1$ ,  $mic_2$ ,  $mic_3$ ,  $mic_4$ ), konstantní zesílení ve vymezeném frekvenčním pásmu ( $f_{min}=100\text{Hz}$ ,  $f_{max}=20\text{ kHz}$ ) a dostatečné zesílení. Potřebné zesílení bylo stanoveno na základě jednoduchého měření, kdy jsem umístil mikrofon do vzdálenosti 4m, a poté jsem mluvil směrem k němu. Zesílení bylo nastaveno tak, aby při běžné řeči (*hlasitost*)

docházelo k vybuzení 0.5 – 0.6x rozsahu A/D převodníku zvukové karty. Ze simulačního programu vychází zesílení předzesilovačů  $A=41.1\text{dB}$ . Na zhotoveném předzesilovači bylo naměřeno zesílení  $A=40.3\text{dB}$  (v pásmu 100 Hz - 20 kHz). Přenosová frekvenční charakteristika ze simulačního programu je zobrazena na obrázku č. 31.



Obr. 31: Přenosová frekvenční charakteristika (simulace)

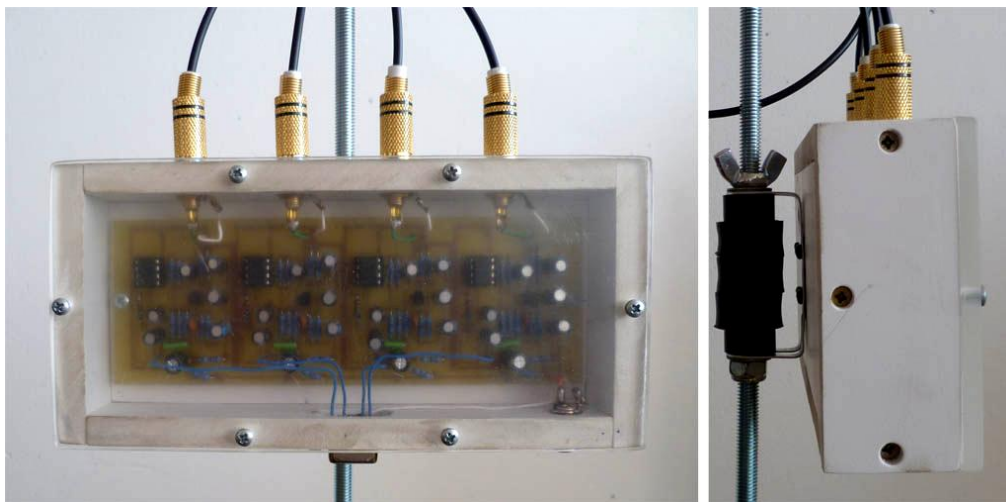
Schéma zapojení mikrofonního předzesilovače je na obrázku č. 32. Toto zapojení můžeme rozdělit na dva stupně. První stupeň je zapojení se společným emitorem a druhý stupeň je invertující zapojení operačního zesilovače.



Obr. 32: Schéma zapojení mikrofonního předzesilovače

V rámci nastavení základních parametrů výpočetního algoritmu je nutné zadat hodnotu prahovací podmínky  $p_{vl}$ . Jelikož se jedná o relativně citlivé předzesilovače, byla nastavena podmínka  $p_{vl}=0.0015$  (úroveň střední hodnoty dvoucestně usměrněného zvukového signálu). Tato podmínka musí být nastavena pro každý použitý typ

předzesilovačů zvlášť. Na obrázku č. 33 (obrázek převzat z mé práce [32]) je zobrazen zkonstruovaný blok předzesilovačů. Pro vstupní kanály byly použity konektory RCA a pro výstupní kanály společný konektor D-sub DE-9.



Obr. 33: Zkonstruovaný blok předzesilovačů

Blok předzesilovačů je připevněn přímo na konstrukci akustického zaměřovače. Takto minimalizujeme délku vodičů propojující mikrofonní vložky a předzesilovače. Minimalizace délky vodičů je velice důležitá z hlediska velikosti šumu v užitečném signálu. Nejvhodnější umístění předzesilovačů je přímo k mikrofonním vložkám. Toto propojení jsem zavrhnul, jelikož by na vysokých frekvencích docházelo k ovlivňování zvukové vlny plošným spojem daného předzesilovače.

### 4.3. Mikrofonní vložky

Při konstrukci akustického zaměřovače byly použity mikrofonní vložky MCE-2000. V katalogovém listu jsou uvedeny tyto parametry:

- Frekvenční rozsah: 38Hz – 15kHz
- Citlivost: (5.6+-1.2) mV/Pa
- SNR>58dB

Z hlediska citlivosti a frekvenčních vlastností jsou dostačující pro účely akustického zaměřování a byly již odzkoušeny v rámci bakalářské práce [27].

## 4.4. Zvuková karta

Pro účely převodu analogového signálu na digitální jsem použil zvukovou kartu ESI Maya 44 USB. Zvuková karta komunikuje s PC přes ASIO drivery, a tudíž umožňuje přímé ovládání z prostředí MATLAB funkcemi *pa\_wavplayrecord.m* a *pa\_wavrecord.m*. Výrobce v katalogovém listu uvádí tyto parametry:

- Vzorkovací kmitočet:  $f_s=44.1\text{kHz}$  nebo  $f_s=48\text{kHz}$
- $U_{inmax} = -10\text{dBV}$  a  $Z_{in} 10\text{k}\Omega$
- $U_{outmax} = -10\text{dBV}$  a  $Z_{out} 100\Omega$
- ADC: 18bit, dynamický rozsah 85dB, frekvenční rozsah 20Hz-20kHz( $f_s=48\text{kHz}$ )
- DAC: 20bit, dynamický rozsah 87dB, frekvenční rozsah 20Hz-20kHz( $f_s=48\text{kHz}$ )

Při použití ASIO 2.0 driverů je možné využít všechny 4 vstupní i výstupní kanály (*MME a WDM umožňují použít pouze 2 vstupní a 2 výstupní kanály*). V mém případě využívám 4 vstupní kanály pro snímání signálů z mikrofónů a jeden výstupní kanál při měření.

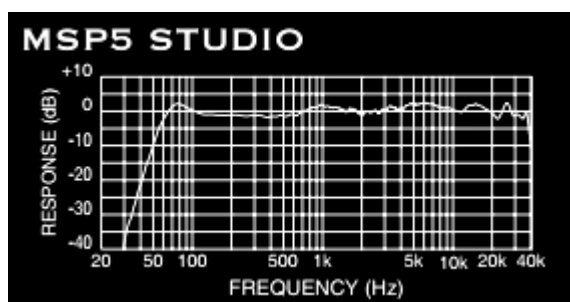
Při ověřování katalogových parametrů ( $U_{inmax}$  a *frekvenční rozsah*) jsem byl velice nemile překvapen. Frekvenční rozsah ADC neodpovídá katalogovému popisu  $f_{min}=20\text{Hz}$  a  $f_{max}=20\text{kHz}$  při vzorkovacím kmitočtu  $f_s=48\text{kHz}$ . Zvuková karta byla schopna korektně převádět signál v rozsahu  $f_{min}=10\text{Hz}$  až  $f_{max}=7\text{kHz}$ . Frekvenční rozsah DAC již odpovídá katalogovému listu a maximální vstupní napětí je rovno  $U_{pp}=0.535\text{V}$ . Vzorkovací frekvence byla zvolena  $f_s=48\text{kHz}$  pro větší úhlové rozlišení. Obrázek č. 34 převzat z [www.esi-audio.com](http://www.esi-audio.com).



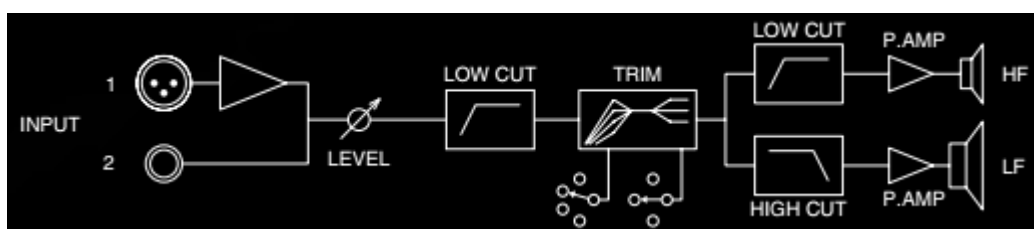
Obr. 34: Zvuková karta ESI MAYA 44 USB

## 5. Měření

Měření bylo provedeno v akusticky upravené místnosti. V této místnosti jsou instalovány akustické pohlcovací materiály, konkrétně SONIT D30 (*porézní materiál*). Z tohoto důvodu jsou zde dobře utlumeny střední a vysoké kmitočty. Místnost nebyla prázdná, a proto docházelo k elementárním odrazům od jednotlivých překážek (*skříňe, židle a jiné*). Akustický zaměřovač byl pro všechna měření umístěn ve vzdálenosti  $l=3\text{m}$  od zdroje zvuku. V této vzdálenosti se akustický zaměřovač nachází ve vzdálené zóně pro všechny kombinace frekvencí harmonických signálů a nastavení vzájemné vzdálenosti mikrofonů  $d$ . Zvuková karta generovala měřicí signály (*harmonické, řečové a šumové*). Tento signál byl zesílen a reprodukován aktivním dvoupásmovým reproduktorem YAMAHA MSP5 STUDIO (*obrázky č. 35 a č. 36 převzaty z [www.yamaha.com](http://www.yamaha.com)*).



Obr. 35: Frekvenční charakteristika YAMAHA MSP5 STUDIO



Obr. 36: Blokové schéma zapojení YAMAHA MSP5 STUDIO

Dělicí kmitočet je dle výrobce na frekvenci  $f_d=2.5\text{kHz}$ . Žádný harmonický signál nebyl nad touto frekvencí, a proto určování vzájemné pozice zaměřovače a zdroje zvuku bylo dle reproduktoru pro dolní kmitočtové pásmo (*pro  $f < 2.5\text{kHz}$* ).

V rámci měření byla určena přesnost akustického zaměřovače pro jednotlivé volby váhovacích funkcí GCC, typy akustických signálů a pozice zdroje zvuku (*nerovnoměrné úhlové rozlišení*). Délka signálu (*ze kterého se počítala korelační*

*funkce*) byla nastavena na délku 8192 vzorků. V případě, že docházelo ke špatnému zaměření, byl tento časový segment prodloužen na 16384 vzorků, případně 32768 vzorků. Pro měření byly použity čisté harmonické signály se specifickými frekvencemi  $f=500\text{Hz}$ ,  $f=250\text{Hz}$  a  $f=100\text{Hz}$ . Frekvence byly voleny na základě podmínky pro maximální vzájemnou vzdálenost mikrofonů  $d$  (vždy tak, aby tato podmínka byla pro konkrétní vzdálenosti  $d$  splněna).



Obr. 37: Měřicí soustava

Dále byly použity řečové signály, konkrétně dvě namluvená slova „brambora“ a „zpěv“. Slovo *zpěv* bylo nahráno tak, že řečník slovo zašeptal a slovo *brambora* bylo nahráno bez šeptání. Slovo *zpěv* mluveno šepotem má charakter šumu a slovo *brambora* mluvené bez šeptání má charakter periodického signálu (s frekvencí  $f_0=108\text{Hz}$ ). Poslední měřicí signál je generovaný šum funkcí *randn.m* z prostředí MATLAB. Měřicí signály byly zvoleny tak, aby se prokázaly vlastnosti akustického zaměřovače pro šumové, řečové, periodické a neperiodické signály. Poslední parametr, který byl při měření měněn, je vzájemná vzdálenost mikrofonů  $d$ . Byly určeny tři základní hodnoty  $d=0.3\text{m}$ ,  $d=0.6\text{m}$  a  $d=0.94\text{m}$ . Poslední hodnota ( $d=0.94\text{m}$ ) je maximální rozpětí konstruovaného zaměřovače. Při zvýšení vzájemné vzdálenosti  $d$  dochází dle teoretických předpokladů ke zvýšení rozlišovací schopnosti a tudíž i přesnosti zaměřování.

## 5.1. Výsledky měření

V rámci měření jsem měnil vzdálenost  $d$ , úhlovou pozici zdroje zvuku a typ zvukového signálu (vzdálenost  $l=3m$  byla pro všechny měření stejná). Nyní vysvětlím, jak tyto parametry ovlivňují přesnost lokalizace zdroje zvuku.

V kapitole *Základní principy akustického zaměřování* jsem popisoval zvyšování úhlového rozlišení při rostoucí vzdálenosti  $d$ . Tento předpoklad se během měření potvrdil. Závislost úhlové odchylky na vzdálenosti  $d$  je uvedena v tabulce č. 4.

$d$ [m]	úhlová odchylka [°]
0.3	2.43
0.6	1.63
0.94	1.44

Tab. 4: Závislost úhlové odchylky na vzdálenosti  $d$

Při změně vzdálenosti  $d$  se projevovала další závislost. A to nutnost prodloužení časových segmentů při zmenšování vzdálenosti  $d$  (pro zachování přesnosti lokalizace). Při zaměřování se vzdáleností  $d=0.94m$  nebylo potřeba u žádné z váhovacích funkcí prodlužovat časový segment. Pro vzdálenosti  $d=0.3m$  a  $d=0.6m$  již bylo někdy nutné časové segmenty prodlužovat na délku 16384 nebo 32768 vzorků. U váhovací funkce PHAT a SCOT byla tato závislost téměř totožná. S funkcí RP nebylo prodlužování téměř nutné a při použití Ones function se prodlužování segmentů nikterak zásadně neprojevilo na přesnosti lokalizace. Další důležitou vlastností je závislost úhlové odchylky při změně váhovací funkce. Ve většině prací zabývajících se akustickým zaměřováním na základě GCC je pro lokalizaci zvolena váhovací funkce PHAT. Při měření měla funkce PHAT nejvyšší stabilitu výsledků (*nedocházelo k velkým výkyvům naměřených hodnot*). Závislost úhlové odchylky na volbě váhovací funkce je zobrazena v tabulce č. 5.

váhovací funkce	úhlová odchylka [°]
Ones function	3.19
PHAT	1.47
SCOT	1.28
RP	1.4

Tab. 5: Závislost úhlové odchylky na volbě váhovací funkce

Za běžných podmínek (*částečně utlumená místnost, slabé šumové pozadí*) jsou váhovací funkce PHAT, SCOT a RP z hlediska přesnosti srovnatelné. Při použití váhovací funkce Ones function se sníží přesnost i stabilita naměřených výsledků.

Nyní se podívejme na závislost přesnosti lokalizace na typu zvukového signálu. Mikrofonní pole jsem ozvučoval signály šumovými, harmonickými a řečovými (*znělé a neznělé*). Šumové signály (*šum a slovo „zpěv“*) byly obecně velice dobře zaměřeny a stabilita výsledků byla také výborná. Což dokazuje teoretický předpoklad, že čím má signál větší šířku pásma, tím dochází k přesnější a stabilnější lokalizaci. Naopak zaměřování čistých harmonických signálů se nepodařilo. Tuto vlastnost přisuzuji velice malé šířce pásma a vlivům odražených vln. Poslední zaměřovaný signál byl znělý řečový signál (*slovo „brambora“*). Tento zvukový signál byl zaměřován s menší přesností než šumové signály, ale i přes to s velice dobrými výsledky. Celková závislost přesnosti lokalizace na typu signálu je zobrazena v tabulce č. 6 a č. 7.

signál	úhlová odchylka [°]
harmonický signál	22.07
slovo "zpěv"	1.34
slovo "brambora"	2.80
šum	1.36

Tab. 6: Závislost úhlové odchylky na volbě zvukového signálu

	PHAT	SCOT
signál	úhlová odchylka [°]	úhlová odchylka [°]
harmonický signál	23.95	17.21
slovo "zpěv"	1.45	1.26
slovo "brambora"	1.49	1.29
šum	1.48	1.29
	RP	Ones function
signál	úhlová odchylka [°]	úhlová odchylka [°]
harmonický signál	22.27	24.84
slovo "zpěv"	1.44	1.23
slovo "brambora"	1.46	6.95
šum	1.28	1.40

Tab. 7: Závislost úhlové odchylky na volbě zvukového signálu jednotlivých funkcí



Poslední vyšetřovanou závislostí je přesnost lokalizace na úhlové pozici zdroje zvuku. Změna přesnosti je zapříčiněna nerovnoměrným úhlovým rozlišením. Zaměřovač má největší rozlišovací schopnost v oblasti nulového časového zpoždění. V případě geometrie 3D Half space je největší rozlišovací schopnost v oblasti azimutu  $\theta = 90^\circ$  a elevace  $\phi = 0^\circ$ . Závislost úhlové odchylky na úhlové pozici je uvedena v tabulce č. 8.

	úhlová pozice [°]	úhlová odchylka [°]
Azimut	90	1.87
	60	2.36
	120	1.93
Elevace	0	0.92
	10	1.37
	-30	2.56

*Tab. 8: Závislost úhlové odchylky na úhlové pozici*

Ze zkušenosti s polohováním zdroje zvuku a akustického zaměřovače si musím položit otázku, zda je člověk s běžnými prostředky schopen umístit zdroj zvuku přesně do úhlových pozic vůči zaměřovači. Při polohování zdroje zvuku je do měření vnášena polohovací chyba minimálně  $\pm 0.5^\circ$ , a proto musí být celková přesnost zaměřovače uvedena s touto polohovací chybou. Celková přesnost zaměřovače je rovna **1.84 $\pm$ 0.5°**.

## 6. Závěr

V rámci této práce jsem vytvořil rešerši lokalizačních metod, naprogramoval algoritmus a GUI pro akustické zaměřování (*metoda TDOA GCC*), zkonstruoval akustický zaměřovač (*předzesilovače a mikrofonní pole*). V poslední části práce jsem ověřil teoretické předpoklady na základě měření. Čtenář si dle rešerše akustických lokalizačních metod může snadno zvolit metodu, které se bude posléze věnovat. Výpočetní algoritmus a jeho grafické uživatelské rozhraní je univerzální nástroj, který dovoluje nastavování mnoha parametrů. V případě že by si čtenář chtěl zhotovit akustický zaměřovač (*využívající tento výpočetní algoritmus*), musí dodržet tyto zásady:

- Použít zvukovou kartu pracující s ASIO, MME nebo WDM ovladači.
- Zvuková karta musí mít čtyři vstupní kanály.
- Ve výpočetním algoritmu nastavit vzorkovací kmitočet  $f_s$  zvukové karty.
- Zkonstruovat mikrofonní pole uvedené v kapitole *Geometrie mikrofonního pole pro akustický zaměřovač*.
- Dle vlastností mikrofonních předzesilovačů nastavit výkonovou podmínku  $p/vl$  nebo ji zakázat ( $p/vl=0$ ).
- Změnit kalibrační data v textových souborech *calibr1.txt* a *calibr10.txt*

Zkonstruovaný akustický zaměřovač dosahuje přesnosti **1.84+-0.5°**. Tato hodnota je spočítána ze zprůměrovaných výsledků ze všech měření (*všechny kombinace vzdálenosti  $d$ , váhovací funkce a měřící signály*). Zaměřovač dosahuje nejlepších výsledků při zaměřování širokopásmových signálů (*řeč zahrnuji do širokopásmových signálů*), za použití metody TDOA GCC a délky časového segmentu v rozmezí 8192-32768 vzorků (*pro vzorkovací kmitočet  $f_s=48kHz$* ). Váhovací funkci volíme dle velikosti SNR. V prostředí se silným šumovým pozadím a velkou koncentrací odražených vln volíme funkce RP a SCOT s délkou segmentu 16384-32768 vzorků. V prostředí se slabým šumovým pozadím a malou koncentrací odražených vln je vhodné použít váhovací funkci PHAT s délkou segmentu 8192 vzorků. Pro maximální úhlové rozlišení volíme velkou vzdálenost  $d$  (*s ohledem na  $d_{max}$  - periodické signály*), vysoký vzorkovací kmitočet  $f_s$  a umísťujeme zaměřovač tak, aby se nejčastěji lokalizované pozice nacházely na pozicích s největším úhlovým rozlišením.

## 7. Literatura

- [1] Uhlíř, J., Sovka, P.: *Číslicové zpracování signálů*. ČVUT, Praha 2002
- [2] Sovka, P., Pollák, P.: *Vybrané metody číslicového zpracování signálů*. ČVUT, Praha 2003
- [3] Škvor, Z.: *Elektro-akustika a akustika*. ČVUT, Praha 2012
- [4] Kolmer, F. Kyncl, J.: *Prostorová akustika*. SNTL/ALFA, Praha 1980
- [5] Jean-Marc Valin, Francois Michaud, Jean Rouat, Dominic Létourneau.: *Sound Source Localization Using a Microphone Array on a Mobile Robot*. [online] Quebec, Canada [vid. 19. 10. 2013] Dostupné z:  
<<http://people.xiph.org/~jm/papers/iros.pdf>>
- [6] Kevin Murphy, Matthew Daigle, Matthew Gates, Vadiraj Hombal.: *Sound Tracking Pan-Tilt Machine*. [online] Rensselaer polytechnic institute [vid. 19. 10. 2013] Dostupné z:  
<<http://cats-fs.rpi.edu/~wenj/ECSE4962S04/progress/progressreportgroup3.pdf>>
- [7] Jin Fu Liou, Kunal Joshi, Gaurang Vador.: *Real-Time Sound Source localization* [online] [vid. 19. 10. 2013] Dostupné z:  
<[http://www.eeweb.com/project/kunal\\_joshi/real-time2-d-audio-source-localization-on-tis-320c6713dsk-in-embedded-c](http://www.eeweb.com/project/kunal_joshi/real-time2-d-audio-source-localization-on-tis-320c6713dsk-in-embedded-c)>
- [8] Matt frear.: *Play and record multi-channel audio using either an ASIO, MME, WDM*. [online] [vid. 19. 10. 2013] Dostupné z:  
<<http://www.mathworks.com/matlabcentral/fileexchange/4017&watching=4017>>
- [9] Youn, D. H., Mathews, V. J.: *On using the short-time unbiased spectral estimation algorithm for estimating time delays and magnitude squared coherence fiction*. [online] The university of Iowa [vid. 19. 10. 2013] Dostupné z:  
<<http://content.lib.utah.edu/utils/getfile/collection/uspace/id/4822/filename/5148.pdf>>
- [10] Cobos, M., Lopez, J., Spors, S.: *A sparsity-Based approach to 3D binaural sound synthesis using time frequency array processing*. [online] Polytechnic university of Valencia [vid. 19. 10. 2013] Dostupné z:  
<<http://asp.eurasipjournals.com/content/2010/1/415840/>>

- [11] Govindan, R. B., Raethjen, J., Kopper, F., Claussen, J.C., Deuschl, G.: Estimation of time delay by coherence analysis. [online] Cornell University [vid. 19. 10. 2013] Dostupné z:  
<<http://arxiv.org/abs/physics/0411141>>
- [12] Taghizadeh, S. R.: *Digital signal processing, Part 3 Discrete- time sinals & systems case studies*. [online] University of Nort London [vid. 19. 10. 2013] Dostupné z:  
<<http://aparate.elth.ucv.ro/POPA/Matla/matlabsignal.pdf>>
- [13] Orság, P.: *Měření základních charakteristik mikrofonů*. [online] VUT Brno [vid. 19. 10. 2013] Dostupné z:  
<[https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=58358](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=58358)>
- [14] Tkadlec, J.: *Lokalizace vzdáleného zdroje zvuku polem mikrofonů*. [online] VUT Brno [vid. 19. 10. 2013] Dostupné z:  
<[https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=39440](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=39440)>
- [15] Eksler, V.: *Lokalizace zdroje vlnění polem mikrofonů v trojrozměrném prostoru*. [online] FEKT VUT Brno [vid. 19. 10. 2013] Dostupné z:  
<<http://www.elektrorevue.cz/clanky/04051/index.html#KAPpojmy>>
- [16] Ekola group, spol. s.r.o.: *Akustická kamera*. [online] Ekola group, Praha [vid. 19. 10. 2013] Dostupné z:  
<<http://www.ekolagroup.cz/cs/akusticka-kamera/>>
- [17] Sovka, P., Turoň, V.: *(A2M31SMU) Přednáška 2: Měření zpoždění mezi signály*. ČVUT, Praha 2012.
- [18] Novotný, P.: *Lokalizace řečníka prostřednictvím mikrofonního pole*. Diplomová práce, České vysoké učení v Praze, Fakulta elektrotechnická, leden 2009.
- [19] Vít, T.: *Aplikace mikrofonního pole*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<[https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=65729](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=65729)>
- [20] Kurc, D.: *Identifikace zdrojů hluku pomocí beamformingu*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<[https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=65729](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=65729)>
- [21] Klinský, T.: *Navigace robota na základě lokalizace zdroje zvuku*. [online] ČVUT FIT [vid. 30. 11. 2013] Dostupné z:  
<<https://dip.felk.cvut.cz/browse/details.php?f=F8&d=K104&y=2012&a=klinstom&t =bach>>

- [22] Bezdíček, M.: *Lokalizace pohyblivých akustických zdrojů*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<[https://www.vutbr.cz/studium/zaverecne-prace?zp\\_id=32033](https://www.vutbr.cz/studium/zaverecne-prace?zp_id=32033)>
- [23] Mikeš, P.: *Lokalizace statických akustických zdrojů*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<<https://dspace.vutbr.cz/handle/11012/5799>>
- [24] Grobelný, P.: *Mikrofonová pole pro prostorovou separaci akustických signálů*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<<https://dspace.vutbr.cz/handle/11012/3728>>
- [25] Varma, K.: Time delay estimate based direction of arrival estimation for speech in reverberant environments. [online] Virginia polytechnic institute [vid. 30. 11. 2013] Dostupné z:  
<<http://scholar.lib.vt.edu/theses/available/etd-10302002-220938/unrestricted/Thesis.pdf>>
- [26] Bartoň, Z.: *Tvarování přijímací charakteristiky mikrofonních polí*. [online] VUT Brno [vid. 30. 11. 2013] Dostupné z:  
<<http://hdl.handle.net/11012/3728>>
- [27] Zitko, J.: *Akustické zaměrování*. ČVUT FEL, Praha 2012
- [28] Ishi, C., Chatot, O., Hiroshi, I., Norihiro, H.: *Evaluation of a MUSIC-based real-time sound localization of multiple sound sources in real noisy environments*. [online] MIT- Massachusetts Institute of Technology [vid. 29. 1. 2014] Dostupné z:  
<<http://dspace.mit.edu/handle/1721.1/59423>>
- [29] Chen, L., Yongchun, L., Fancheng, K.: *Acoustic Source Localization Based on Generalized Cross-correlation Time-delay Estimation*. [online] China University of Mining and Technology School of Information and Electrical Engineering [vid. 29. 1. 2014] Dostupné z:  
<<http://www.sciencedirect.com/science/article/pii/S1877705811024167>>
- [30] Humphrey, A., R.: *Automatic Loudspeaker Location Detection for use in Ambisonic Systems*. [online] The University of York [vid. 29. 1. 2014] Dostupné z:  
<<http://www.speech.kth.se/prod/publications/files/1693.pdf>>

- [31] Murray, J., C., Erwin, H., Wermter, S.: *Robotic Sound-source Localization and Tracking Using Interaural Time Difference and Cross-Correlation*. [online] University of Sunderland [vid. 29. 1. 2014] Dostupné z:  
<[http://www.informatik.uni-hamburg.de/WTM/ps/murray\\_ulm04.pdf](http://www.informatik.uni-hamburg.de/WTM/ps/murray_ulm04.pdf)>
- [32] Zitko, J.: *Acoustic localization*. POSTER 2014. 18 th International Student Conference on Electrical Engineering. Czech technical university, faculty of Electrical Engineering, Prague 2014

## 8. Příloha A – Naměřené výsledky

PHAT							
d = 0.3m, l = 3m, $f_{mFF} = 5666\text{Hz}$ , $f_{md} = 1133\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 500Hz</b>		90.0	0.0	90.0	-25.4	90.0	-1.4
<b>řečové signály</b>	"zpěv"	91.4	1.4	61.6	12.4	121.6	-26.9
	"brambora"	92.7	1.4	63.1	12.4	123.2	-28.4
<b>šumový signál</b>		91.4	1.4	63.1	12.4	121.6	-26.9

PHAT							
d = 0.6m, l = 3m, $f_{mFF} = 1417\text{Hz}$ , $f_{md} = 567\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 250Hz</b>		55.2	34.8	50.0	77.5	90.0	0.0
<b>řečové signály</b>	"zpěv"	90.7	0.7	62.3	11.0	121.6	-27.7
	"brambora"	91.4	0.7	60.8	10.3	121.6	-28.4
<b>šumový signál</b>		90.7	0.7	61.6	11.0	121.6	-27.7

PHAT							
d = 0.94m, l = 3m, $f_{mFF} = 577\text{Hz}$ , $f_{md} = 362\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 100Hz</b>		90.0	0.0	90.0	0.0	90.0	0.0
<b>řečové signály</b>	"zpěv"	90.9	0.9	61.0	10.5	121.0	-28.0
	"brambora"	90.9	0.9	61.0	10.9	121.0	-28.5
<b>šumový signál</b>		90.9	0.0	61.0	10.5	121.5	-28.0

### Legenda:

barva pole	počet vzorků
	8192
	16384
	32768

$c = 340\text{ms}^{-1}$  (rychlost zvuku)

$f_{mFF}$  - mezní frekvence pro podmínku vzdálené zóny (FAR FIELD)

$f_{md}$  - mezní frekvence pro podmínku maximální vzdálenosti mikrofonů  $d_{max}$

d - vzdálenost mezi mikrofony (mezi kterými dochází k odhadu  $\tau$ )

l - vzdálenost zdroje zvuku od akustického zaměřovače

SCOT							
d = 0.3m, l = 3m, $f_{mFF} = 5666\text{Hz}$ , $f_{md} = 1133\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 500Hz</b>		90.0	0.0	90.0	-30.0	90.0	-1.4
<b>řečové signály</b>	"zpěv"	92.7	1.4	61.6	11.0	121.6	-26.9
	"brambora"	92.7	0.0	63.1	11.0	123.2	-28.4
<b>šumový signál</b>		90.0	1.4	61.6	12.4	121.6	-26.9

SCOT							
d = 0.6m, l = 3m, $f_{mFF} = 1417\text{Hz}$ , $f_{md} = 567\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 250Hz</b>		89.3	0.7	70.5	0.7	90.0	0.0
<b>řečové signály</b>	"zpěv"	91.4	-0.7	61.6	11.0	120.8	-28.4
	"brambora"	91.4	-0.7	62.3	11.0	120.8	-28.4
<b>šumový signál</b>		91.4	-0.7	61.6	11.0	122.4	-28.4

SCOT							
d = 0.94m, l = 3m, $f_{mFF} = 577\text{Hz}$ , $f_{md} = 362\text{Hz}$		$az_1 = 90^\circ$	$el_1 = 0^\circ$	$az_2 = 60^\circ$	$el_2 = 10^\circ$	$az_3 = 120^\circ$	$el_3 = -30^\circ$
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 100Hz</b>		90.0	0.0	90.0	0.0	90.0	0.0
<b>řečové signály</b>	"zpěv"	90.9	0.4	61.0	10.0	121.0	-29.0
	"brambora"	90.0	0.0	60.5	10.5	122.0	-29.0
<b>šumový signál</b>		90.9	0.4	61.0	10.5	121.0	-29.0

**Legenda:**

barva pole	počet vzorků
	8192
	16384
	32768

$c = 340\text{ms}^{-1}$  (rychlost zvuku)

$f_{mFF}$  - mezní frekvence pro podmínku vzdálené zóny (FAR FIELD)

$f_{md}$  - mezní frekvence pro podmínku maximální vzdálenosti mikrofonů  $d_{max}$

d - vzdálenost mezi mikrofony (mezi kterými dochází k odhadu  $\tau$ )

l - vzdálenost zdroje zvuku od akustického zaměřovače



RP							
d = 0.3m, l = 3m, f <sub>mFF</sub> = 5666Hz, f <sub>md</sub> = 1133Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 500Hz</b>		90.0	0.0	90.0	-34.8	90.0	-11.0
<b>řečové signály</b>	"zpěv"	88.6	1.4	58.4	12.4	118.4	-26.9
	"brambora"	91.4	1.4	60.0	11.0	121.6	-26.9
<b>šumový signál</b>		88.6	1.4	58.4	12.4	118.4	-28.4

RP							
d = 0.6m, l = 3m, f <sub>mFF</sub> = 1417Hz, f <sub>md</sub> = 567Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 250Hz</b>		15.4	25.4	67.6	0.7	90.0	0.0
<b>řečové signály</b>	"zpěv"	91.4	0.7	59.2	11.7	121.6	-27.7
	"brambora"	91.4	1.4	61.6	11.7	121.6	-27.7
<b>šumový signál</b>		91.4	0.7	61.6	11.7	120.0	-27.7

RP							
d = 0.94m, l = 3m, f <sub>mFF</sub> = 577Hz, f <sub>md</sub> = 362Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 100Hz</b>		90.0	0.0	90.0	0.0	90.0	0.0
<b>řečové signály</b>	"zpěv"	90.9	0.9	61.0	10.9	121.0	-28.5
	"brambora"	90.9	1.3	61.0	11.4	121.0	-27.5
<b>šumový signál</b>		90.9	0.9	61.0	10.9	120.5	-28.5

**Legenda:**

barva pole	počet vzorků
	8192
	16384
	32768

c = 340ms<sup>-1</sup> (rychlost zvuku)

f<sub>mFF</sub> - mezní frekvence pro podmínku vzdálené zóny (FAR FIELD)

f<sub>md</sub> - mezní frekvence pro podmínku maximální vzdálenosti mikrofonů d<sub>max</sub>

d - vzdálenost mezi mikrofony (mezi kterými dochází k odhadu τ)

l - vzdálenost zdroje zvuku od akustického zaměřovače

Ones function							
d = 0.3m, l = 3m, f <sub>mFF</sub> = 5666Hz, f <sub>md</sub> = 1133Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 500Hz</b>		84.5	-4.1	61.6	-47.6	81.8	-51.8
<b>řečové signály</b>	"zpěv"	91.4	1.4	60.0	12.4	123.2	-31.6
	"brambora"	94.1	-4.1	80.4	4.1	112.4	-18.0
<b>šumový signál</b>		91.4	1.4	60.0	12.4	123.2	-31.6

Ones function							
d = 0.6m, l = 3m, f <sub>mFF</sub> = 1417Hz, f <sub>md</sub> = 567Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 250Hz</b>		21.8	72.2	50.0	77.5	99.6	-26.1
<b>řečové signály</b>	"zpěv"	90.7	0.0	61.6	10.3	120.0	-26.1
	"brambora"	92.0	-2.0	69.8	6.2	113.1	-24.6
<b>šumový signál</b>		90.7	0.0	62.3	10.3	119.2	-24.6

Ones function							
d = 0.94m, l = 3m, f <sub>mFF</sub> = 577Hz, f <sub>md</sub> = 362Hz		az <sub>1</sub> = 90°	el <sub>1</sub> = 0°	az <sub>2</sub> = 60°	el <sub>2</sub> = 10°	az <sub>3</sub> = 120°	el <sub>3</sub> = -30°
signály		azimut[°]	elevace[°]	azimut[°]	elevace[°]	azimut[°]	elevace[°]
<b>harmonický signál f = 100Hz</b>		100.5	-3.9	81.7	-2.2	107.2	-14.9
<b>řečové signály</b>	"zpěv"	90.9	0.0	60.5	10.5	120.5	-26.6
	"brambora"	112.3	-0.4	70.1	9.6	127.3	-30.5
<b>šumový signál</b>		90.9	0.4	60.5	10.0	120.5	-26.6

**Legenda:**

barva pole	počet vzorků
	8192
	16384
	32768

c = 340ms<sup>-1</sup> (rychlost zvuku)

f<sub>mFF</sub> - mezní frekvence pro podmínku vzdálené zóny (FAR FIELD)

f<sub>md</sub> - mezní frekvence pro podmínku maximální vzdálenosti mikrofonů d<sub>max</sub>

d - vzdálenost mezi mikrofony (mezi kterými dochází k odhadu τ)

l - vzdálenost zdroje zvuku od akustického zaměřovače

## 9. Příloha B – Obsah příloženého CD

<b>Složka „MATLAB“</b> .....	Funkce, knihovny a kalibrační data
<b>Složka „Měření“</b> .....	Materiály, použité při měření
Složka „Měřicí signály“ .....	Signály použité při měření
Složka „Obrázky z měření“ .....	Obrázky z průběhu měření
Výsledky měření .....	Zpracované výsledky z měření
<b>Složka „Text“</b> .....	Textová podoba diplomové práce
Složka „Obrázky“ .....	Obrázky ve formátu PNG a JPEG
Složka „Akustický zaměřovač“ .....	Obrázky akustického zaměřovače
Složka „Obrázky“ .....	Obrázky z textu diplomové práce
Složka „Zdroje“ .....	Použitá literatura ve formátu PDF