

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

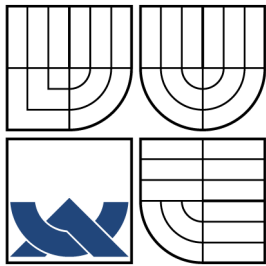
LOKALIZACE STATICKÝCH AKUSTICKÝCH ZDROJŮ

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

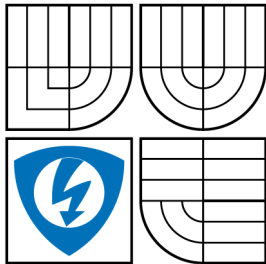
AUTOR PRÁCE
AUTHOR

PETR MIKEŠ

BRNO 2010



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND
COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

LOKALIZACE STATICKÝCH AKUSTICKÝCH ZDROJŮ STATICAL ACOUSTICAL SOURCES LOCALIZATION

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

PETR MIKEŠ

VEDOUCÍ PRÁCE
SUPERVISOR

ING. IVAN MÍČA

BRNO 2010



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Bakalářská práce

bakalářský studijní obor
Teleinformatika

Student: Petr Mikeš

ID: 106639

Ročník: 3

Akademický rok: 2009/2010

NÁZEV TÉMATU:

Lokalizace statických akustických zdrojů

POKYNY PRO VYPRACOVÁNÍ:

Z existujících metod lokalizace akustických zdrojů, popsaných v dostupné technické literatuře vyberte nejméně dvě metody vhodné pro širokopásmové signály, např. pro řeč. Tyto metody implementujte pomocí zvoleného programovacího jazyka a měřením v reálných podmínkách, případně simulacemi ověřte jejich přesnost.

DOPORUČENÁ LITERATURA:

[1] EKSLER, V. Prostorová lokalizace a separace naslepo zdrojů akustických signálů polem mikrofonů. Doktorská disertační práce. Vysoké učení technické v Brně, 2006.

[2] BRANDSTEIN, M., WARD, D. Microphone Arrays: Signal Processing Techniques and Applications. Springer-Verlag Berlin, 2001. 389 s. ISBN 3-540-41953-5.

Termín zadání: 29.1.2010

Termín odevzdání: 2.6.2010

Vedoucí práce: Ing. Ivan Míča

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ZDE VLOŽIT PRVNÍ LIST LICENČNÍ
SMOUVY

Z důvodu správného číslování stránek

ZDE VLOŽIT DRUHÝ LIST LICENČNÍ
SMOUVY

Z důvodu správného číslování stránek

ZDE VLOŽIT podepsané Prohlášení o
shodě listinné a elektronické formy

Z důvodu správného číslování stránek

ABSTRAKT

Mikrofonové pole se může vyskytovat u mnoha důležitých aplikací. Primární použití, ze kterých vychází většina dalších aplikací, jsou pouze dvě: lokalizace akustických zdrojů a separace jednotlivých signálů z jejich směsice. Mikrofonové pole již bývá zakomponováno v běžných mobilních telefonech, kde redukuje šumy a ruchy, které přicházejí z okolí a degradují tím řečníkův hlasový projev. Další použití je například při rozpoznávání řeči, lokalizaci více řečníků nebo pro izolaci signálu pouze od jednoho řečníka.

KLÍČOVÁ SLOVA

Mikrofonové pole, senzor, lokalizace zdroje zvuku, prostorová lokalizace, časové zpoždění

ABSTRACT

The microphone array can be used in many important applications. The primary usage, which most of other applications are based on, is only: acoustic sources localization and separation of each signal from their compound. The microphone array is usually included in ordinary mobile phones, where it reduces noise and sound of its surrounding that otherwise degrades vocal speech. Other microphone array utilization is in for example speech identification, speakers' localization or signal isolation of one speaker.

KEYWORDS

Microphone array, sensor, signal localization, dimensional localization, time delay

MIKEŠ P. *Lokalizace akustických zdrojů polem mikrofonů*. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2010. 63 s., 4 s. příloh. Bakalářská práce. Vedoucí práce byl Ing. I. Míča.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Lokalizace statických akustických zdrojů“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji vedoucímu semestrální práce Ing. Ivanu Míčovi za velmi užitečnou metodickou, odbornou a pedagogickou pomoc a další cenné rady při zpracování práce.

V Brně dne

.....

(podpis autora)

OBSAH

Úvod	15
1 Úvod do problematiky	16
1.1 Historie a současný stav problematiky sensorových polí	16
1.2 Uspořádání, struktura a geometrie pole	17
1.3 Odraz zvuku, šíření zvuku v bezodrazovém i v reálném prostředí . . .	18
1.3.1 Bezodrazové prostředí	18
1.3.2 Odraz zvuku	20
1.3.3 Ohyb zvuku	23
2 Metody lokalizace DOA	24
2.1 Metoda tvarování přijímací charakteristiky	24
2.1.1 Tvarovač typu Delay And Sum	26
2.1.2 Současná lokalizace více zdrojů	27
2.2 Spektrální odhad vysokého rozlišení	28
2.3 Metoda odhadu časových zpoždění	28
2.3.1 Určení posuvů signálu mezi senzory	30
2.3.2 Stanovení úhlu DOA	33
2.3.3 Nejednoznačnost určení směru příchodu signálu	34
2.3.4 Metoda GCC PHAT	34
3 Dosažené výsledky při lokalizaci zdrojů	38
3.1 Měřicí aparatura	38
3.1.1 Kalibrace měřicí aparatury	39
3.2 Výsledky metod v bezodrazovém prostředí	41
3.2.1 Metoda tvarování přijímací charakteristiky	41
3.2.2 Metoda odhadu časových zpoždění	43
3.3 Výsledky metod v reálném prostředí	44
3.3.1 Metoda tvarování přijímací charakteristiky	44
3.3.2 Metoda odhadu časových zpoždění	45
3.3.3 GCC PHAT	48
3.4 Shrnutí výsledků lokalizace DOA	49
4 Závěr	53
Literatura	54
Seznam zkratk	56

Seznam symbolů a veličin	57
Seznam příloh	59
A Měřicí aparatura	60
B Tabulky	61
C Obsah přiloženého DVD	63

SEZNAM OBRÁZKŮ

1.1	Bezodrazová komora s podlahou tvořenou kovovým roštem.	19
1.2	Semi-bezodrazová komora s pevnou podlahou.	20
1.3	Šíření ve volném prostředí.	21
1.4	Odraz zvuku a rozložení intenzit.	22
1.5	Odraz zvuku a rozptyl.	22
1.6	Ohyb zvuku.	23
2.1	Lineární uspořádání pole.	24
2.2	Prostorová odezva tvarovače typu DAS.	25
2.3	Tvarovač (Beamformer) typu Delay And Sum.	26
2.4	Znázornění azimutu a elevace DOA.	29
2.5	Rozmístění senzorů v rovině.	30
2.6	Křížová korelace.	31
2.7	Zobrazení vektorů.	32
2.8	Rozdíl vzdáleností signálu dopadajícího na dvojici senzorů.	33
2.9	Nejednoznačnost určení směru příchodu signálu.	35
2.10	Porovnání korelační funkce při zpracování metodou TDOA a GCC PHAT: a) <i>GCC PHAT (1 024 vzorků)</i> , b) <i>křížová korelační funkce TDOA (1 024 vzorků)</i> , c) <i>GCC PHAT (10 240 vzorků)</i> , d) <i>křížová korelační funkce TDOA (10 240 vzorků)</i>	36
3.1	Audio interface Event Layla, Mikrofon Behringer ECM800 a mikrofonní předzesilovač Behringer MIC2200.	38
3.2	Konstrukce pro uchycení mikrofonů na točně.	39
3.3	Lineární uspořádání pole s naznačeným DOA.	40
3.4	Výsledný 3-D graf lokalizace metodou tvarování přijímací charakteristiky v bezodrazovém prostředí.	41
3.5	Výsledný 2-D graf lokalizace metodou tvarování přijímací charakteristiky v bezodrazovém prostředí pro frekvenci $f=1\ 900$ Hz.	42
3.6	Časový průběh testovacích signálů a) výstřel b) „five“	42
3.7	Kmitočtové spektrum testovacích signálů a) výstřel b) „five“	43
3.8	Výsledný 2-D graf lokalizace metodou tvarování přijímací charakteristiky v reálném prostředí pro frekvenci $f = 508$ Hz.	45
3.9	Porovnání lokalizace metodou TDOA a GCC PHAT v reálných podmínkách se zpracovávaným blokem o velikosti $k = 1\ 024$	47
3.10	Porovnání lokalizace metodou TDOA a GCC PHAT v reálných podmínkách se zpracovávaným blokem o velikosti $k = 512$	49

3.11	Lokalizace přesného umístění metodou GCC PHAT.	50
A.1	Zapojení měřicí aparatury.	60

SEZNAM TABULEK

3.1	Skutečný úhel a vypočítané úhly metodou TDOA.	44
3.2	Závislost správného určení DOA na S/N pro testovací signál „Base“ pro geometrii pole do rovnoramenného trojúhelníku.	46
3.3	Závislost správného určení DOA na S/N pro testovací signál „Base“ geometrii pole do rovnostranného trojúhelníku.	48
B.1	Skutečný úhel a vypočítané úhly metody TDOA s rozložením do pravoúhlého trojúhelníku v reálných podmínkách se zpracovávaným blokem vzorků o velikosti $k = 10\ 240$	61
B.2	Skutečný úhel a vypočítané úhly metody TDOA s rozložením do rovnostranného trojúhelníku v reálných podmínkách se zpracovávaným blokem vzorků o velikosti $k = 10\ 240$	62

ÚVOD

Mikrofonové pole má mnoho použití například v automobilovém průmyslu, mobilních telefonech, videokonferencích či při rozpoznávání řeči. Při lokalizaci akustických signálů v poli mikrofonů je cílem získat informace o zdroji, a to jak v rovinné, tak i v prostorové oblasti nebo odstranění parazitních ruchů (šumů) tak, aby se takto degradovaný signál co nejvíce podobal signálu původnímu.

Tato práce se zabývá lokalizací statických zdrojů polem mikrofonů. Důraz je zvláště kladen na techniky (metody) lokalizace řečového signálu. Vybrané metody jsou v praktické části implementovány v prostředí MATLAB, kde je ověřena jejich praktická funkčnost v reálném i bezodrazovém prostředí.

První část je zaměřena na seznámení s problematikou sensorových polí, uspořádání jednotlivých polí, jejich strukturu, geometrii, rozdělíme typy prostředí a objasníme problematiku vzniku odrazů a ohybu zvuku. Dále je v této části popsána obecná problematika historie a použití sensorových polí.

Ve druhé části jsou popsány metody založené na tvarování přijímací charakteristiky, spektrálním odhadu vysokého rozlišení, odhadu časových zpoždění a metoda GCC PHAT. Dále je zde popsán postup lokalizace více zdrojů pomocí metody tvarování přijímací charakteristiky. V této části jsou také odvozeny teoretické postupy a předpoklady metod lokalizace akustických signálů.

Poslední část obsahuje dosažené výsledky testovaných metod, postup měření a popis měřicí aparatury.

1 ÚVOD DO PROBLEMATIKY

1.1 Historie a současný stav problematiky senzorových polí

Odborný termín pro lokalizaci zvuku je označován zkratkou SSL z anglického „sound source localization“. Lokalizaci můžeme rozdělit na techniky jednokanálové a vícekanálové.

Jednokanálové techniky se používaly v dřívějších dobách. Určení polohy zdroje signálu je nesmírně náročné, protože v tomto případě můžeme polohu zdroje jen hrubě odhadovat, například použitím směrových mikrofonů.

Při použití vícekanálových technik je akustický signál snímán více senzory. Nejjednodušší vícekanálovou technikou je technika dvoukanálová, která je založena na jednoduchém principu. Akustický signál zdroje je snímán směrovým mikrofonem a okolní hluk je v určité vzdálenosti od zdroje snímán všesměrovým mikrofonem. Pokud nyní tyto signály od sebe odečteme, získáme kvalitnější výsledek oproti použití pouze techniky jednokanálové.

Technika mikrofonových polí se začala rozvíjet v 80. letech v souvislosti s telekonferencemi (Telekonference - Flanagan 1985). Od této doby začal velký vývoj v oblasti lokalizace a začalo se přicházet na výhody použití mikrofonových polí. V dnešní době je velká většina videokonferenčních sálů vybavena buď headsetovými soupravami, nebo právě mikrofonovými poli. Výhodou použití mikrofonových polí v těchto místnostech je schopnost určit pozici mluvčího, tj. směr příchodu signálu, který je označován zkratkou DOA z anglického „Direction of Arrival“, a následně na tuto akci vyvolat určitý podnět. V našem případě by se jednalo například o nasměrování kamery na požadované místo, separování hlasu mluvčího, anebo pokud je videokonferenční místnost snímána více kamerami, jedná se například o spuštění záznamu určité kamery. Další využití jednoduchého pole je v automobilovém průmyslu. Nově vyrobené automobily mohou být vybaveny zakomponovanou handsfree sadou přímo v palubní desce. Pole je zde ve spoustě případů řešeno dvěma mikrofony. Jedním, který snímá hlas mluvčího, a druhým, který má za úkol snímat vnitřní a provozní hluk automobilu, jako je například hluk motoru, spínání cyklovačů a hluk ventilátorů [1].

Ve všech těchto případech se k lokalizaci používají tři základní techniky (metoda časových zpoždění, metoda tvarování přijímací charakteristiky pole a metoda, která je založena na spektrálních odhadech vysokého rozlišení). Některé tyto metody, nebo jejich modifikace, dokáží s použitím vhodného pole lokalizovat nejen směr šíření zvuku, ale také přesné umístění zdroje zvuku vzhledem k rozestavení senzorů

v prostoru.

Metody lokalizace ale nemusí být omezeny jen na akustické signály. Své uplatnění mají i při lokalizaci zářících zdrojů v radarech či sonarech a astronomii.

1.2 Uspořádání, struktura a geometrie pole

Pojem sensorové pole představuje skupinu snímačů určité fyzikální veličiny (ať již akustického tlaku, tepla, svítivosti nebo magnetického pole), které jsou vůči sobě v určitém postavení. Základní uspořádání můžeme rozčlenit do třech typů:

- lineární (senzory rozmístěné v řadě),
- rovinné (senzory rozmístěné na ploše),
- prostorové (senzory rozmístěné v prostoru).

Tyto skupiny dále můžeme dělit podle struktury:

- rovnoměrné (zkratka UA z anglického „uniform array“),
- nerovnoměrné.

Lineární uspořádání slouží k lokalizaci pouze v jedné polorovině. Tudíž azimut θ může nabývat hodnot 0° až 180° , respektive 180° až 360° . K objasnění vezměme v potaz skutečnost, že senzory v tomto uspořádání jsou umístěny na jedné přímce, tzn. výsledný azimut θ může mít obě znaménka, kladné i záporné. Proto u tohoto pole nejsme schopni přesně lokalizovat zdroj v celé rovině.

Druhým typem je uspořádání rovinné, kdy senzory leží v jedné rovině. Zde jsou senzory nejčastěji umístěny ve čtverci či obdélníku. Takto sestrojené pole pokryje celý 2-D (dvoudimenzionální) prostor, tudíž jsme schopni určit azimut θ v rozsahu hodnot 0° až 360° . Pokud je vzdálenost všech sousedních sensorů konstantní, můžeme toto pole označit za rovinné rovnoměrné. Pokud je vzdálenost odlišná, jedná se o pole nerovnoměrné. Tímto typem pole je možno lokalizovat zdroj i v 3-D prostoru, a tak určit úhel azimutu θ i elevace ϕ . V případě elevace ale nejsme jednoznačně schopni říci, ze kterého směru zdroj DOA přichází, viz část kap. 2.3.3.

Třetím a posledním typem je uspořádání prostorové. V tomto případě budeme chtít nalézt dva neznámé parametry. Směr zdroje signálu je vyjádřen úhlem azimutu θ a elevace ϕ . I v tomto typu uspořádání pole bývá ve většině případů použito pole s rozmístěním ve tvaru krychle či kvádrů. Tento typ již odstraňuje, oproti poli s uspořádáním rovinným, nejednoznačnost směru příchodu signálu.

Další geometrická rozmístění ale mohou být i kruhová, trojúhelníková a elipsoidní. Tvar pole má samozřejmě vliv na přesné určení výsledných úhlů. Nelze jednoznačně říci, která uspořádání je nejvhodnější použít. To je vždy potřeba zvážit pro daný problém. Při rozhodování je třeba vzít v úvahu, zda budeme převážně určovat zdroje zvuku vzdálené či blízké a kolik dimenzí budeme chtít do výsledku zahrnout.[1]

1.3 Odraz zvuku, šíření zvuku v bezodrazovém i v reálném prostředí

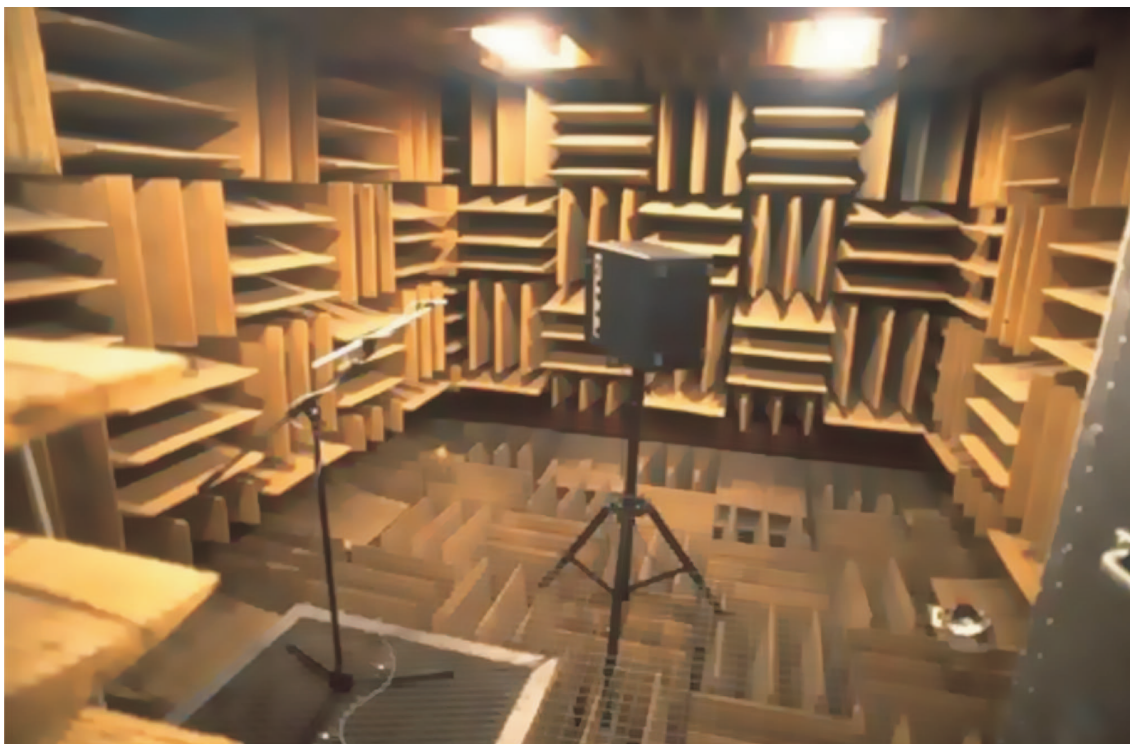
Převážná část metod, které jsou doposud popsány v technické literatuře a jsou určeny k lokalizaci akustických signálů, dokáží bez velkých problémů lokalizovat zdroje akustického signálu v bezodrazovém prostředí (v tzv. bezodrazové komoře). Většina videokonferenčních sálů je vybavena jen částečnými akustickými úpravami místnosti, tudíž většina prostředí, ve kterém máme být schopni lokalizovat směr příchodu signálu, je prostředí reálné. V prostředích bezodrazových většinou nebývá problém určit zdroj zvuku. Problém nastává pokud jsou metody použity v reálných podmínkách, kde se uplatní například faktory doba dozvuku, odraz zvuku, ohyb zvuku a šum. Pro následné další pochopení principu těchto metod se alespoň okrajově seznámíme se vznikem odrazů a některými parametry místnosti.

1.3.1 Bezodrazové prostředí

Bezodrazová komora

Výrazem bezodrazové prostředí definujeme takovou část prostoru, ve které nedochází k žádným, nebo jen minimálním odrazům. Takové prostředí je nazýváno mrtvá komora nebo také zvukotěsná komora s častým označením zkratkou ACH z anglického „Anechoic chamber“. Toto prostředí se využívá k mnohým měřením v oblasti akustiky nebo například k měřením frekvenčních charakteristik reproduktorů i mikrofónů, viz obr. 1.1.

Toto uzavřené prostředí je tvořeno zvukotěsnou místností, jejíž stěny neodrážejí žádné zvuky, přičemž jsou všechny zvuky pohlceny. Stěny místnosti, strop i podlaha by v ideálním případě měly být pokryty speciálním materiálem s velkým stupněm pohltivosti všech frekvencí. Tento materiál je dále tvarován (nejčastěji do tvaru jehlanu) a seskupován, aby nedocházelo ke zpětným odrazům a zvuk byl úplně pohlcen. Nad podlahu pokrytou takovýmto akustickým materiálem je poté instalována



Obr. 1.1: Bezodrazová komora s podlahou tvořenou kovovým roštem.

kovová či plastová mříž, která slouží jak k ukotvení stojanů a jiných pomůcek k samotnému měření, tak i k pohybu osob, které potřebují s těmito předměty manipulovat.

Tyto komory se vyznačují typickou hladinou hluku kolem 10 - 20 dB.

Semi-bezodrazové komory

Zatímco mrtvá komora pohlcuje zvuk ve všech směrech, semi-bezodrazové komory jsou tvořeny ze stejného materiálu jako mrtvé komory s tím rozdílem, že jejich podlaha není pokryta akustickým materiálem sloužícím k zabránění vzniku odrazů jako v předchozím případě (jehlany), ale nejčastěji je pokryta různými typy koberců či jiných materiálů s velkým koeficientem pohltivosti. Podlaha tudíž může sloužit jako pevná pracovní základna, na které mohou stát či být uchyceny těžší předměty, viz obr. 1.2. Tyto semi-bezodrazové komory bývají díky přijatelnější ceně než u ACH většinou součástí nahrávacích studií nebo různých ruchových studií. Takové typy komor se převážně vyznačují typickou hladinou hluku kolem 15 - 25 dB.

Volné prostředí

Jelikož se ve volném prostředí při lokalizaci směru příchodu signálu od akustického zdroje nenacházejí žádné překážky či odrazné plochy, nebo se vyskytují ve značné



Obr. 1.2: Semi-bezodrazová komora s pevnou podlahou.

vzdálenosti, nemusíme problematiku odrazů při lokalizaci vůbec uvažovat, jelikož dominantní vlnou, která dopadne na senzory, je vlna přímá.

1.3.2 Odraz zvuku

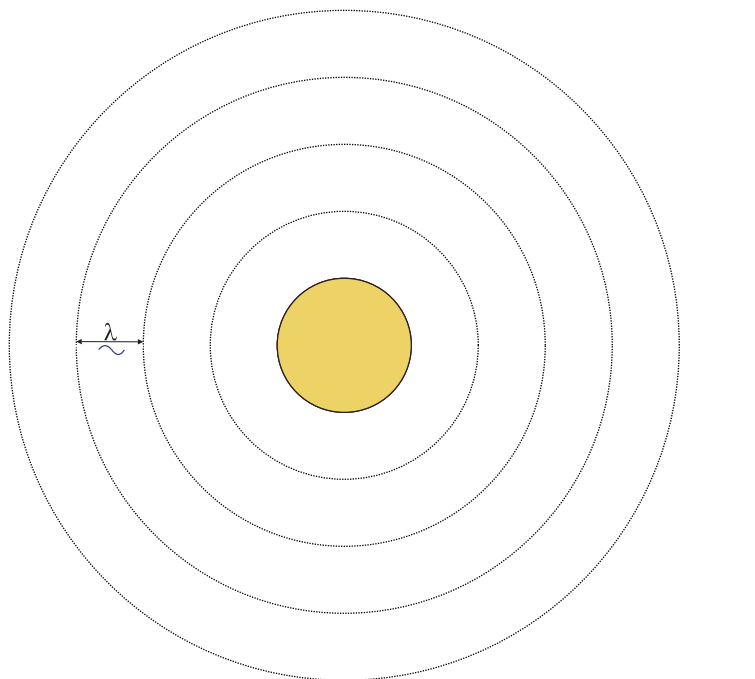
Charakteristickou veličinou zvukové vlny je tzv. vlnová délka označována λ . Vlnovou délku můžeme definovat jako:

$$\lambda = \frac{v_{\text{faz}}}{f}, \quad (1.1)$$

kde v_{faz} je fázová rychlost šířeného vlnění a f je frekvence vlnění. Tento parametr poté určuje vzdálenost mezi dvěma nejbližšími body se stejnou amplitudou i fází. Ve volném prostoru se zvuk šíří všemi směry v tzv. akustických vlnoplochách, viz obr. 1.3. Tyto vlnoplochy mají tvar koule. Při dostatečně velké vzdálenosti od zdroje je možno tyto kulové vlnoplochy nahradit částí roviny.

Vznik odrazu zvuku

Pokud zvuková vlna narazí ve svém směru šíření na překážku, při dopadu na tuto překážku dojde k jeho odrazu a část zvukové energie pronikne do prostředí druhého



Obr. 1.3: Šíření ve volném prostředí.

nebo se přemění v teplo. Pro odražený paprsek platí *zákon odrazu*:

$$\beta' = \beta, \text{ přičemž } \beta' \wedge \beta \in \chi, \quad (1.2)$$

kde úhel odrazu β' se rovná úhlu dopadu β , přičemž dopadající i odrážející vlnoplocha náleží polorovině dopadu. Pro intenzitu zvuku platí vztah:

$$I = \frac{E}{St}, \quad (1.3)$$

kde E je zvuková energie, S vyjadřuje obsah plochy, na kterou vlnoplocha dopadá a t je čas.

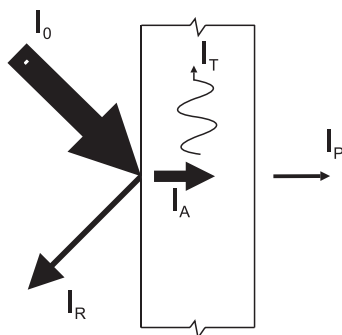
Pokud tedy budeme brát v úvahu značení z obr. 1.4, můžeme napsat závislosti mezi intenzitami:

$$I_A = I_T + I_P, \quad (1.4)$$

$$I_R = I_0 - (I_T + I_P), \quad (1.5)$$

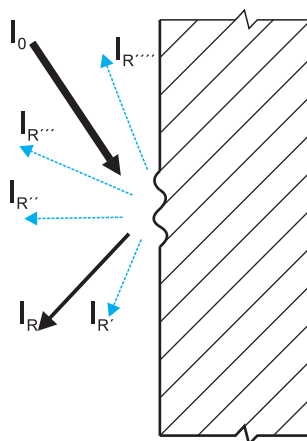
kde I_0 je intenzita zvuku dopadajícího na překážku, I_A je intenzita zvuku, která vstoupila do materiálu, I_T intenzita zvuku, která se přemění na teplo, I_P intenzita zvuku která vstoupí do druhého prostředí a I_R představuje intenzitu odraženého zvuku.

Dále je možné odvodit určité koeficientní podíly, které nám umožňují lepší pohled na materiál překážky:



Obr. 1.4: Odraz zvuku a rozložení intenzit.

- r - koeficient odrazivosti (reflexe),
- α - koeficient pohltivosti (absorpce),
- ζ - koeficient zvukové průzvučnosti (propustnosti).



Obr. 1.5: Odraz zvuku a rozptyl.

Koeficient odrazivosti r můžeme poté vyjádřit jako:

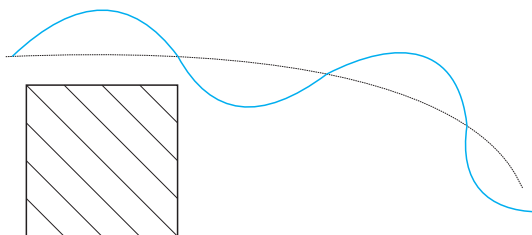
$$r = \frac{I_R}{I_0}. \quad (1.6)$$

Koeficient pohltivosti α je možné vyjádřit jako:

$$\alpha = \frac{I_A}{I_0}. \quad (1.7)$$

Koeficient zvukové průzvučnosti ζ je možné vyjádřit jako:

$$\zeta = \frac{I_P}{I_0}. \quad (1.8)$$

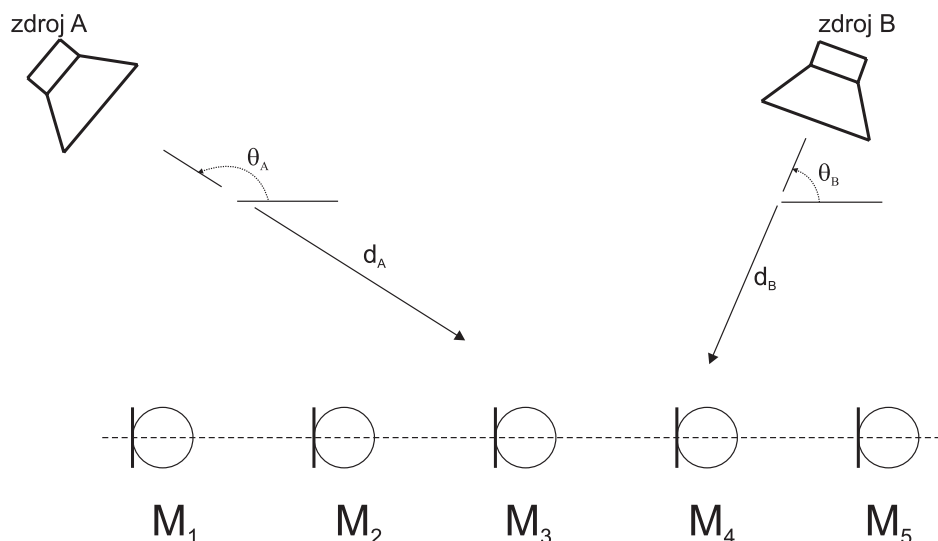


Obr. 1.6: Ohyb zvuku.

Pokud vlnová plocha dopadne na nerovný povrch, dojde k tzv. rozptylu zvuku do více směrů, viz obr. 1.5. Tento jev většinou nastává v reálných uzavřených prostředích, kde dominantní je opět vlna přímá, ale vznikají i vlny nepřímé, které vzniknou odražením od předmětů umístěných v prostoru a od stěn, které ohraničují daný prostor. Vlastnosti odrazu i koeficienty pohltivosti, zvukové průzvučnosti i odrazivosti jsou frekvenčně závislé [3] [2].

1.3.3 Ohyb zvuku

Ohyb zvuku je další jev, který ovlivňuje šíření zvuku v prostoru obr.(1.6). Pokud se v prostoru vyskytují předměty a objekty, které mají srovnatelné rozměry jako vlnové délky frekvencí ve slyšitelném pásmu, tj. vlnová délka λ dosahující velikosti v rozmezí od 1,7 cm do 17 m, dochází k ohybu zvukové vlny kolem tohoto předmětu. Díky tomuto jevu se zvuk dostane i za překážku. Tak jako odraz zvuku je i ohyb zvuku frekvenčně závislý.



Obr. 2.1: Lineární uspořádání pole.

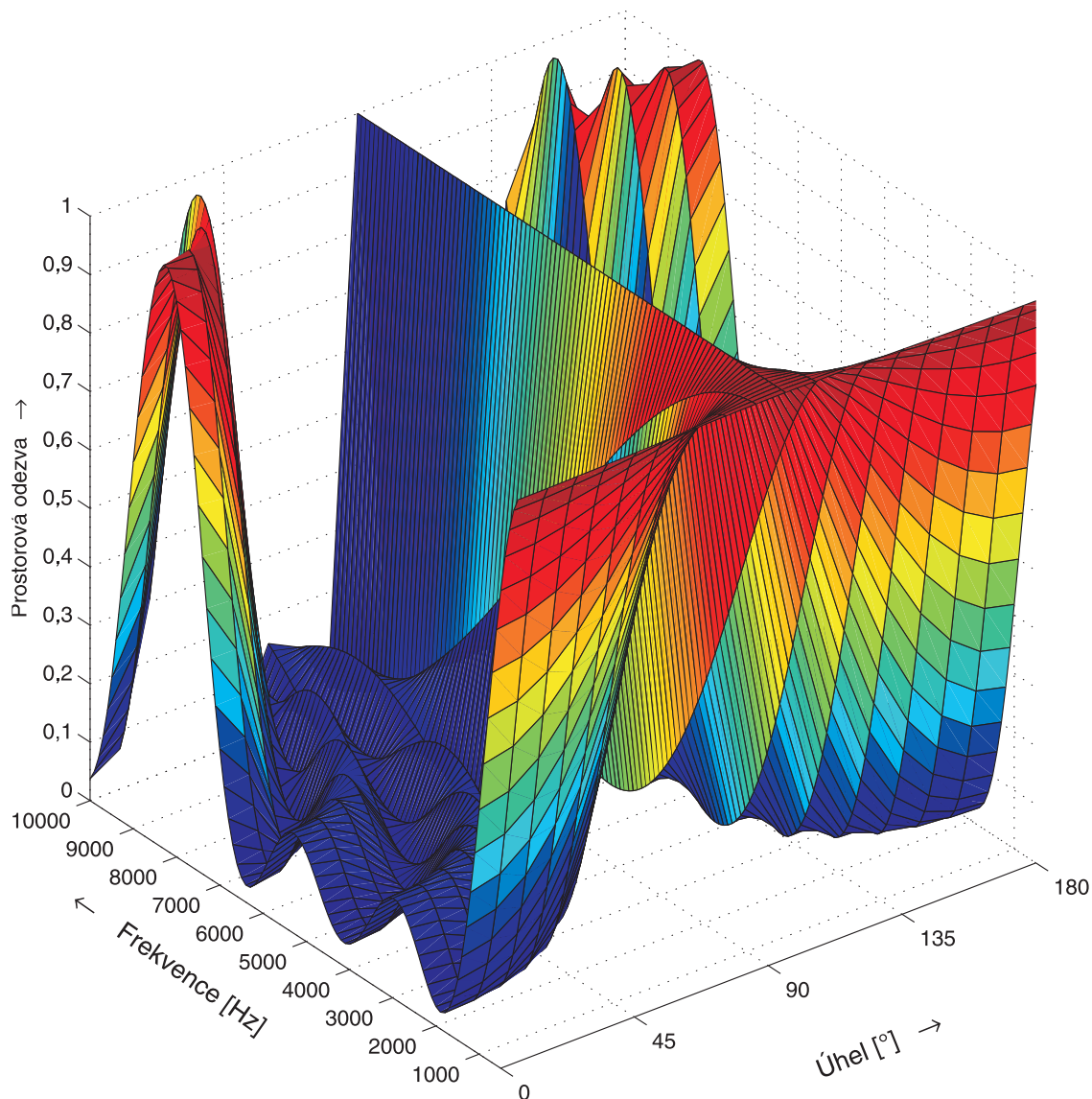
2 METODY LOKALIZACE DOA

Pro úspěšné lokalizování zdroje příchodu signálu DOA můžeme rozdělit metody do tří skupin. Metody založené na:

- tvarování přijímací charakteristiky,
- spektrálním odhadu vysokého rozlišení,
- odhadu časových zpoždění.

2.1 Metoda tvarování přijímací charakteristiky

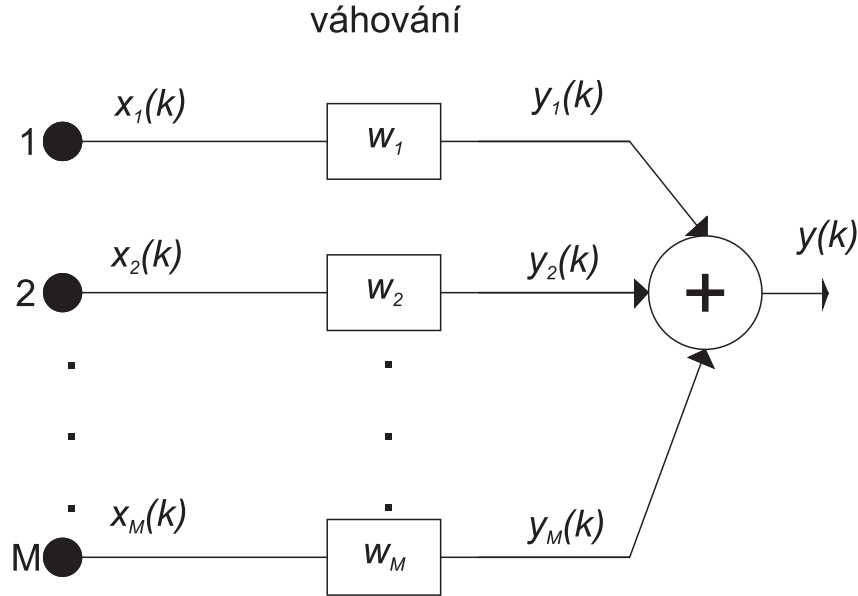
Metoda tvarování přijímací charakteristiky, kterou se zabývá tato část kapitoly, je označována také podle jejího anglického jména Beamforming. Tato metoda se využívá k separaci signálů (nutno použít velké množství mikrofonů), potlačení ruchů, ale i k samotné lokalizaci. Je nutno podotknout, že tato metoda dokáže určit více zdrojů, viz obr. 2.1. Čím více senzorů v této metodě použijeme, tím lépe jsme schopni potlačit nežádoucí šumy a ruchy. Nevýhodou této metody je ale její velká výpočetní náročnost. Pro každý úhel uvažovaného směru šíření je nutné znovu vyhodnotit odezvu pole (pro azimut θ v rozmezí 0° až 360° a elevaci ϕ -90° až 90°). Graf odezvy pole je zobrazen na obr. 2.2. Prostorová odezva pole byla počítána pro lineární pole pěti senzorů vzdálených mezi sebou $q = 0,04$ m a s vahami $w_i = 0,2$. Výpočetní nároky je samozřejmě možné částečně omezit krokem, pro který budeme pole vyhodnocovat. To má však negativní důsledek na přesnost vyhodnocení. Další možností, jak omezit



Obr. 2.2: Prostorová odezva tvarovače typu DAS.

výpočetní náročnost, je použití vícestupňového hledání. Vícestupňové hledání pracuje na principu prvotního počítání s větším krokem a následném zjemňování kroku a tím zpřesňování požadovaného parametru. Důsledek velké výpočetní náročnosti způsobuje nevhodnost použití této metody pro lokalizaci pohyblivých zdrojů.

Metoda je založena na tvarování přijímací charakteristiky pole, která na základě intenzit signálu a úhlu dopadu, ve kterém na senzory dopadnou signály, určí směr šíření vlny.



Obr. 2.3: Tvarovač (Beamformer) typu Delay And Sum.

2.1.1 Tvarovač typu Delay And Sum

Mezi nejjednodušší tvarovače, někdy označované také zkratkou BF z anglického "Beamformer", patří právě tvarovač typu Delay And Sum a je označován zkratkou DAS. Struktura tvarovače je naznačena na obr. 2.3. Signály, které budou dopadat kolmo na pole senzorů budou nezměněny. Pokud ale budou dopadat pod jiným úhlem než 90° , signály budou mít při dopadu na senzory mezi sebou odlišné zpoždění a intenzitu. Po natočení DAS je vždy sejmuta odezva pro úhel, o který je tvarovač pootočen a signály v jiném úhlu jsou potlačeny. Signály jsou dále upraveny pomocí váhování a následně jsou upravené signály jednotlivých kanálů sečteny. Váhy jsou voleny jako konstantní, většinou $w_i = 1$ nebo $w_i = \frac{1}{M}$. Směr DOA se rovná úhlu či úhlům, pro které DAS zjistí největší intenzity. Pro tvarovač typu DAS tedy platí:

$$y_i(f_k) = w_i e^{-j2\pi f c^{-1} q \cos \theta}, \quad (2.1)$$

kde w_i jsou váhy, f daná frekvence, c rychlost zvuku a q vzdálenost senzoru od vztažného bodu [4]. A

$$E_B(\theta) = \sum_{i=1}^M |y_i(k)|^2, \quad (2.2)$$

kde E_B je odezva pole, M je počet senzorů, $y_i(k)$ je výstup mikrofonomového pole v i -tém kroku.

2.1.2 Současná lokalizace více zdrojů

Metoda založená na tvarování přijímací charakteristiky má, na rozdíl od prosté metody odhadu časových zpoždění, schopnost lokalizovat více DOA (obr. 2.1). Metoda je schopna lokalizovat tolik zdrojů, kolik senzorů máme v poli k dispozici. Některé modifikace této metody jsou schopny lokalizace i více zdrojů než je vlastní počet senzorů v poli [1].

Tvarování přijímací charakteristiky je v tomto případě možné rozdělit do dvou kroků:

- zpoždění signálu pomocí celých vzorků - hrubé hledání směru příchodu (DOA),
- zpoždění signálu pomocí fázového posunu - jemné hledání směru příchodu (DOA).

Pokud jednotlivé signály, mimo vztažného, zpozdíme vůči sobě o určitý krok, jsme schopni vypočítat úhel natočení pole a odezvu pole. Musíme ale dodržet lineárnost zpoždění. Pokud tedy signál mající od vztažného senzoru vzdálenost velikosti q zpozdíme o časový interval odpovídající jednomu vzorku, signál o vzdálenosti $2q$ musíme zpozdít o časový interval odpovídající vzorkům dvěma. V tomto případě jsme schopni lokalizovat zdroj v určitém kroku daným zpožděním celého vzorku. Tento krok však nebude konstantní. Pro zpoždění více vzorků se krok bude zvětšovat, přičemž ve zpoždění odpovídajícímu malému počtu vzorků bude krok velice jemný a tím dostáváme schopnost rozlišit i desetinu stupně. Rozlišení metody je závislé na úhlu natočení γ , vzorkovací frekvenci a vzdálenosti senzorů q . Přepočít mezi počtem vzorků zpoždění a úhlem natočení je:

$$\gamma = \arcsin \left(\frac{k' \frac{1}{f_{vz}} c}{q} \right), \quad (2.3)$$

kde γ je natočení pole, k' představuje zpoždění ve vzorcích, f_{vz} je vzorkovací frekvence, c představuje rychlost zvuku a q je vzdálenost dvou senzorů.

Pro časový interval odpovídající většímu počtu vzorků je ale krok poměrně hrubý (3° až 10°). Pokud se tedy DOA bude nacházet v mezních hodnotách pole, je nutné výsledek zpřesnit. Toho můžeme dosáhnout zavedením fázového posunu k jednotlivým zpožděním pomocí celých vzorků. Fázového posunu jsme schopni docílit zavedením komplexního váhování w_i do výpočtu odezvy pole. Fázové posunutí je však možno aplikovat pouze na úzkopásmové signály. Proto je nutné zpracovávat jen úzkopásmové signály nebo širokopásmový signál vhodnými filtry rozdělit na frekvenčně omezené úseky [11].

2.2 Spektrální odhad vysokého rozlišení

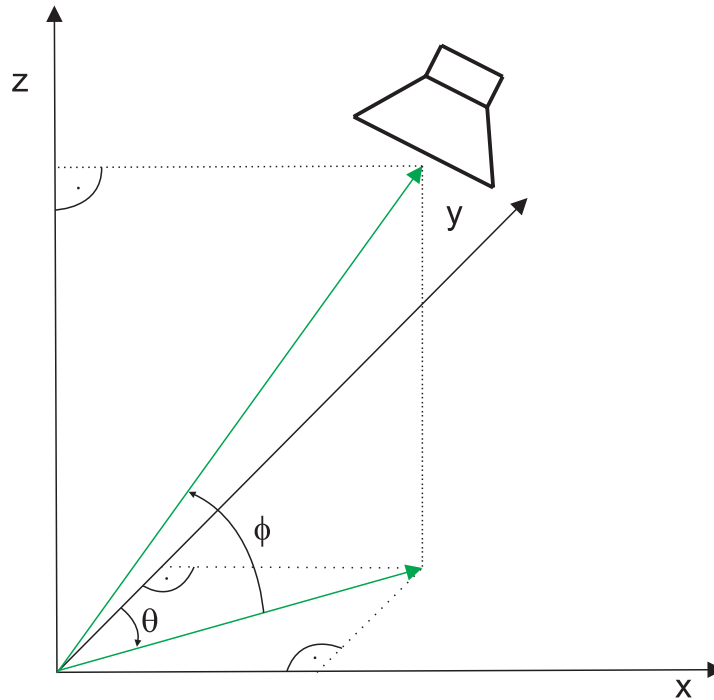
Metody spektrálního odhadu vysokého rozlišení využívají spektrální analýzu vysokého rozlišení, jako například spektrální odhad minimální odchytky (MV - z anglického „Minimum Variance“), autoregresivní modelování a jedna z nejpobulárnějších metod vícenásobné klasifikace signálu (MUSIC - z anglického „Multiple Signal Classification“). Algoritmy implementující metodu MV a MUSIC jsou vhodné pro typy pole jak prostorového i rovinného a dokáží lokalizovat i zdroje umístěné blízko sensorového pole. Naopak metoda využívající autoregresivní modelování je vhodná k lokalizaci pouze vzdáleného pole. Metody využívající spektrálního odhadu vysokého rozlišení nejsou většinou v praxi využívány k lokalizaci řečových signálů, a to z důvodu využití korelační matice, která vzniká průměrováním dat v intervalu, ve kterém jsou parametry odhadu zdroje a šumu neměnné [5].

Metody založené na spektrálním odhadu vysokého rozlišení jsou schopny lokalizovat více zdrojů signálu i v případě, že jsou zdroje velmi blízko sebe. V tomto případě metody vykazují znatelně lepší výsledky než metody založené na tvarování přijímací charakteristiky. Metody vykazují mnohem ostřejší špičky odezvy pole v místech, kde se vyskytuje zdroj signálu. Tyto metody jsou také znatelně méně výpočetně náročné než metody tvarování přijímací charakteristiky. Většina lokalizačních technik založená na odhadu vysokého rozlišení je vyvinuta pouze pro aplikaci na úzkopásmových signálech, a proto je metoda uplatňována například v radarové lokalizaci. Pokud by tato metoda byla aplikována na širokopásmový signál, její výpočetní nároky se razantně zvýší [5].

2.3 Metoda odhadu časových zpoždění

Tato metoda je aplikována v praxi pro nalezení směru příchodu zvukové vlny poměrně často, a to pro její nízkou výpočetní náročnost. Malá výpočetní náročnost nám pak umožňuje využít tuto metodu pro lokalizaci směru příchodu v reálném čase, například jako algoritmus v signálovém procesoru. Navíc je možno tuto metodu aplikovat na úzkopásmové i širokopásmové signály. Metoda je také označována zkratkou TDOA z anglického „Time Difference of Arrival“. Nevýhodou této metody je nemožnost lokalizování více než jednoho zdroje DOA. Metoda časových zpoždění vychází výhradně z geometrického modelu daného prostoru a nezávisí na fyzikálních charakteristikách uvažované soustavy. Naopak díky této metodě můžeme určité fyzikální veličiny, například rychlost zvuku, vypočítat [7].

Pro jednodušší výpočet této metody zavedeme několik zjednodušení, která musí být dodržena:



Obr. 2.4: Znázornění azimutu a elevace DOA.

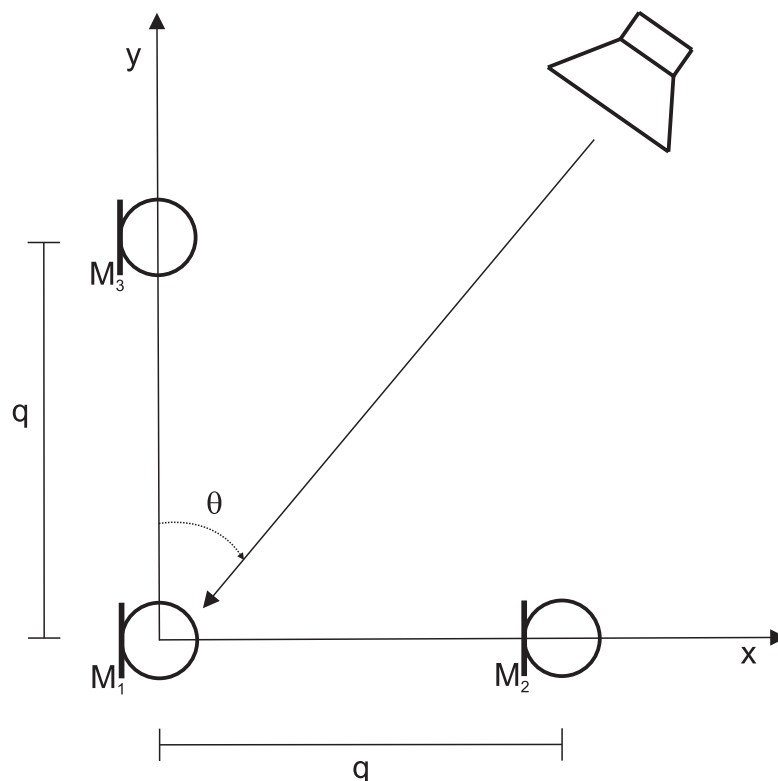
- převody signálů, které přicházejí ze senzorů, na A/D převodnicích musí být synchronní,
- zdroj zvuku i sensorové pole musí být umístěno z hlediska šíření signálu v jednom homogenním prostředí,
- vzdálenost sensorového pole od zdroje zvuku musí být mnohem větší než vzájemná vzdálenost mezi jednotlivými senzory v poli.

Při dodržení těchto podmínek můžeme v blízkosti sensorového pole nahradit akustickou vlnoplochu přímkou a zároveň je zajištěna synchronizace signálů a tím možnost správného určení časových zpoždění [1]. Metodu časových zpoždění můžeme využít pro lokalizaci zdroje zvuku v rovině i v prostoru. K pokrytí dvoudimenzionálního prostoru jsou postačující nejméně tři senzory, přičemž musí platit podmínky:

- alespoň jeden senzor bude ležet mimo přímkou tvořenou zbývajícími dvěma senzory,
- všechny senzory musí ležet v jedné rovině.

Pro prostorové pole platí:

- abychom mohli jednoznačně lokalizovat zdroj zvuku, prostorové pole musí být tvořeno minimálně čtyřmi senzory.



Obr. 2.5: Rozmístění senzorů v rovině.

Příčemž musí současně platit první podmínka z pole rovinného a podmínka z pole prostorového:

- alespoň jeden senzor musí ležet mimo rovinu tvořenou zbývajícími senzory.

Metodu nyní rozdělíme do dvou dílčích úseků - určení posuvů mezi senzory a výpočet zdroje DOA.

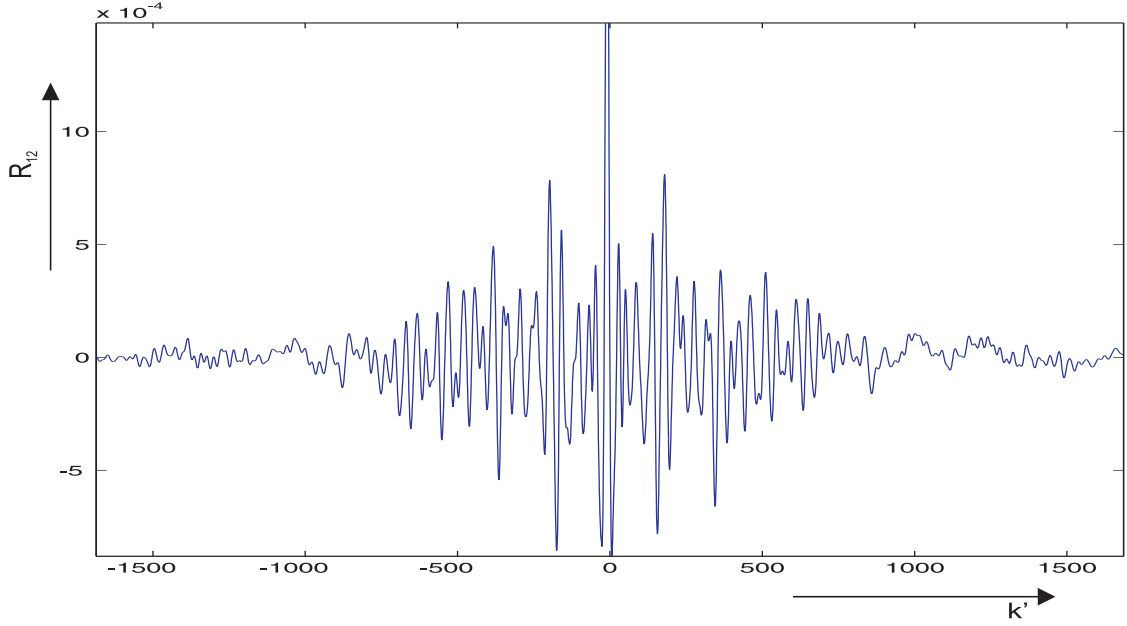
2.3.1 Určení posuvů signálu mezi senzory

Tato část si klade za cíl určit o kolik vzorků jsou signály ze senzorů posunuty. Toho lze dosáhnout například pomocí křížové korelační funkce. Pro jednoznačné určení zpoždění signálu musí platit vztah:

$$q_{\max} = \frac{\lambda_{\min}}{2}, \quad (2.4)$$

kde q je vzdálenost mezi sousedními senzory a λ_{\min} je vlnová délka maximální zpracovávané frekvence. Po dosazení vztahu $\lambda_{\min} = \frac{c}{f}$ bude platit:

$$f_{\max} < \frac{c}{q}, \quad (2.5)$$



Obr. 2.6: Křížová korelace.

kde c je rychlost zvuku. Jako vstupní veličiny vzájemné křížové korelační funkce je potřeba mít k dispozici stejně dlouhé vektory signálů o velikosti K , které byly současně naměřeny na senzorech. Křížová korelační funkce je pak definována vztahem:

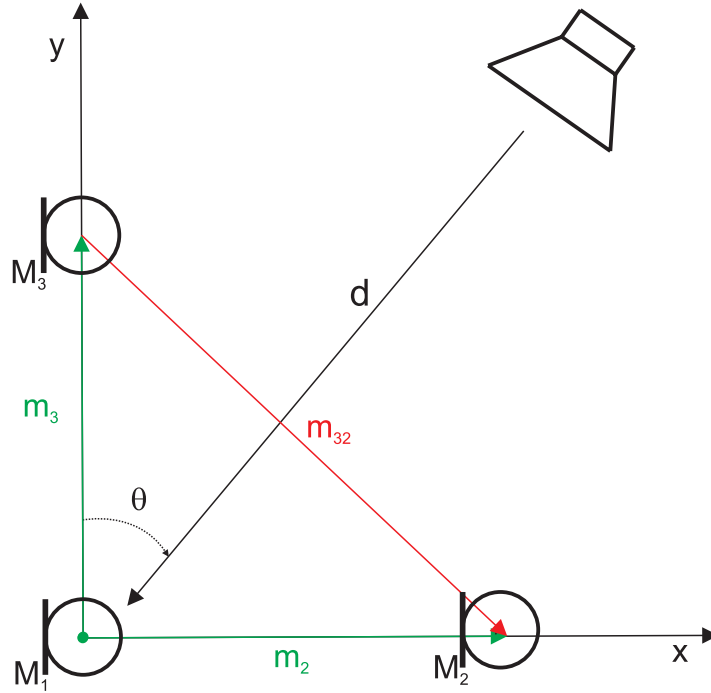
$$R_{12}(k') = \frac{1}{K} \sum_{k=1}^{K-|k'|} x_1(k)x_2(k+|k'|) \text{ pro } k' = -K-1, \dots, 0, \dots, K-1, \quad (2.6)$$

kde x_1 a x_2 jsou vektory vstupních signálů. R_{12} je tedy vektor o velikosti $2K-1$. Index posloupnosti, ve kterém se ve vektoru R_{21} vyskytne extrém, se rovná časovému zpoždění signálu mezi dvěma senzory ve vzorcích [1]. Na obr. 2.5 je naznačeno rozmístění senzorů (M_1, M_2, M_3) v prostoru a q (vzdálenost mezi senzory). Souřadnice každého senzoru v poli reprezentuje vektor $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_x$ (obr. 2.7). Z těchto vektorů nyní vytvoříme matici \mathbf{M} , kde každý sloupec reprezentuje polohu senzoru.

$$\mathbf{M} = \begin{pmatrix} m_{1x} & m_{2x} & m_{3x} \dots \\ m_{1y} & m_{2y} & m_{3y} \dots \\ m_{1z} & m_{2z} & m_{3z} \dots \end{pmatrix}$$

Nyní vytvoříme matici \mathbf{Q} o rozměrech $3 \times Q$, kde $Q = \frac{M!}{2!(M-2)!}$ a M je počet senzorů. Každý sloupec matice \mathbf{Q} bude představovat vektor dvojice senzorů a bude tvaru $m_{ij} = \mathbf{m}_i - \mathbf{m}_j$ pro $i, j = 1, 2, 3..M$ (obr. 2.7), přičemž $i > j$.

$$\mathbf{Q} = \begin{pmatrix} m_{2x} - m_{1x} & m_{3x} - m_{1x} \dots \\ m_{2y} - m_{1y} & m_{3y} - m_{1y} \dots \\ m_{2z} - m_{1z} & m_{3z} - m_{1z} \dots \end{pmatrix}$$



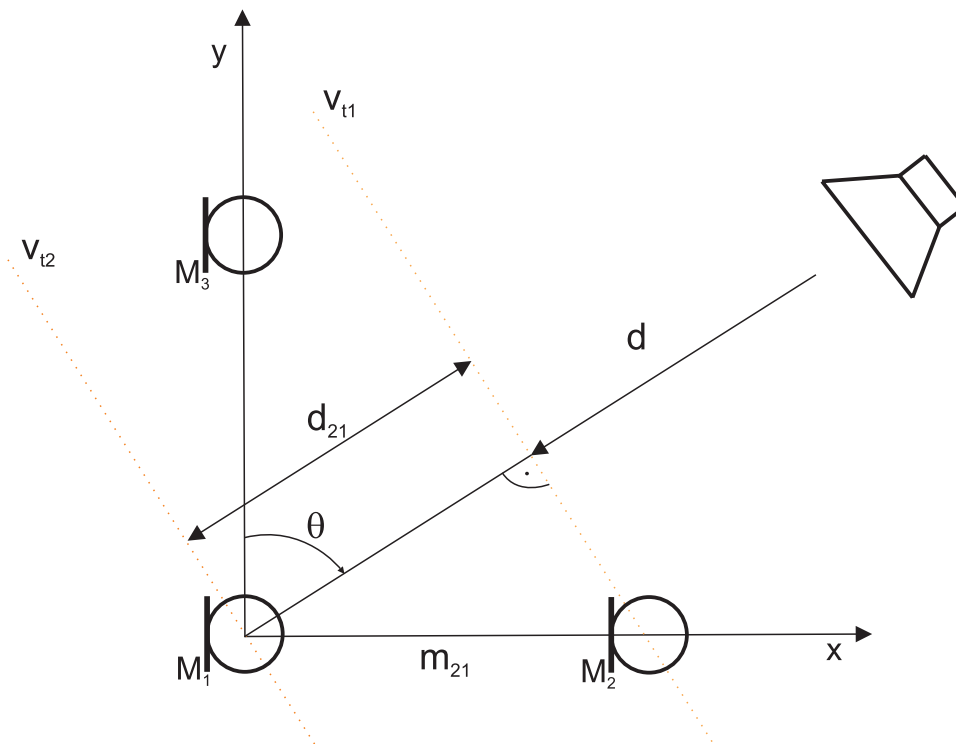
Obr. 2.7: Zobrazení vektorů.

Dále s využitím vzájemné křížové korelační funkce zjistíme zpoždění mezi vztažným senzorem a zbývajícími senzory. V našem případě je za vztažný senzor zvolen senzor M_1 . Vytvoříme matici \mathbf{R} , jejíž každý sloupec bude reprezentovat výsledný vektor vzájemné křížové korelace vztažného vektoru a zbývajících vektorů, přičemž nesmí být zaměněno pořadí. Velikost matice \mathbf{R} bude tedy $(2K - 1) \times (M - 1)$. V každém sloupci matice \mathbf{R} nalezneme extrémy a zjistíme příslušný index vzorku k' . Na obr. 2.6 je zobrazen průběh křížové korelační funkce a její extrém, který odpovídá $k' = 11$. Z obrázku je patrné, že průběh korelace je téměř symetrický, což je způsobeno zaznamenáváním signálů stejného zdroje, které se od sebe liší jen díky různé vzdálenosti šíření prostředím. Takto získaná zpoždění ve vzorcích jsme schopni snadno přepočítat, pokud známe rychlost zvuku a vzorkovací frekvenci zaznamenaných signálů, na časové zpoždění τ (2.7) i na zpoždění vzdálenostní d_{ij} (2.8):

$$\tau_{ij} = \frac{k'}{f_{vz}}, \quad (2.7)$$

$$d_{ij} = c\tau, \quad (2.8)$$

kde koeficienty ij představují indexy korelovaných signálů, k' je pak extrém daného sloupce matice \mathbf{R} , f_{vz} je vzorkovací frekvence signálů a c představuje rychlost světla. Právě toto zpoždění d_{ij} vyjadřuje rozdíl vzdáleností, které musí signál urazit od dopadu na senzor M_i až k dopadu na senzor M_j , viz obr. 2.8.



Obr. 2.8: Rozdíl vzdáleností signálu dopadajícího na dvojici senzorů.

2.3.2 Stanovení úhlu DOA

V tomto kroku již zbývá určit úhel DOA. Jednotkový vektor \mathbf{d} , který reprezentuje směr příchodu signálu (DOA), můžeme vyjádřit pro 3-D jako:

$$\mathbf{d} = [\cos \phi \sin \theta, \cos \phi \cos \theta, \sin \phi], \quad (2.9)$$

kde ϕ je elevace a θ azimut, viz obr. 2.4. Pro rovinu můžeme tento vzorec zjednodušit za předpokladu, že $\phi = 0$:

$$\mathbf{d} = [\cos 0 \sin \theta, \cos 0 \cos \theta, \sin 0] = [\sin \theta, \cos \theta]. \quad (2.10)$$

Vzájemný vztah mezi časovým zpožděním d_{ij} a vektorem m_{ij} spojující dva senzory (obr. 2.8) můžeme vyjádřit:

$$d_{ij} = (\mathbf{m}_i - \mathbf{m}_j)^T \mathbf{d}^T. \quad (2.11)$$

Po dosazení vzorce, který vyjadřuje DOA (2.10), můžeme napsat:

$$d_{ij} = m_{ij_x} \sin(\theta) + m_{ij_y} \cos(\theta), \quad (2.12)$$

kde m_{ij_x}, m_{ij_y} značí souřadnice vektoru spojující dva senzory. Tyto souřadnice jsou vyjádřeny v matici Q . Z tohoto vzorce jsme již nyní schopni spočítat azimut θ . Obdobné odvození je možné provést i pro 3-D a je identické s odvozením pro 2-D.

Pokud tedy zvolíme rozmístění pole z obr. 2.8 a souřadnice senzorů $\mathbf{m}_1 = [0, 0]$, $\mathbf{m}_2 = [0, 4]$, $\mathbf{m}_3 = [4, 0]$, můžeme napsat vztah mezi senzory M_1 a M_2 :

$$d_{21} = m_{21_x} \sin(\theta) + m_{21_y} \cos(\theta) = 4 \sin(\theta) + 0 \cos(\theta) = 4 \sin(\theta),$$

z čehož vyplývá:

$$\theta = \arcsin \frac{d_{21}}{m_{21_x}}. \quad (2.13)$$

Analogicky pro dvojici senzorů M_1 a M_3 :

$$d_{31} = m_{31_x} \sin(\theta) + m_{31_y} \cos(\theta) = 0 \sin(\theta) + 4 \cos(\theta) = 4 \cos(\theta),$$

z čehož vyplývá:

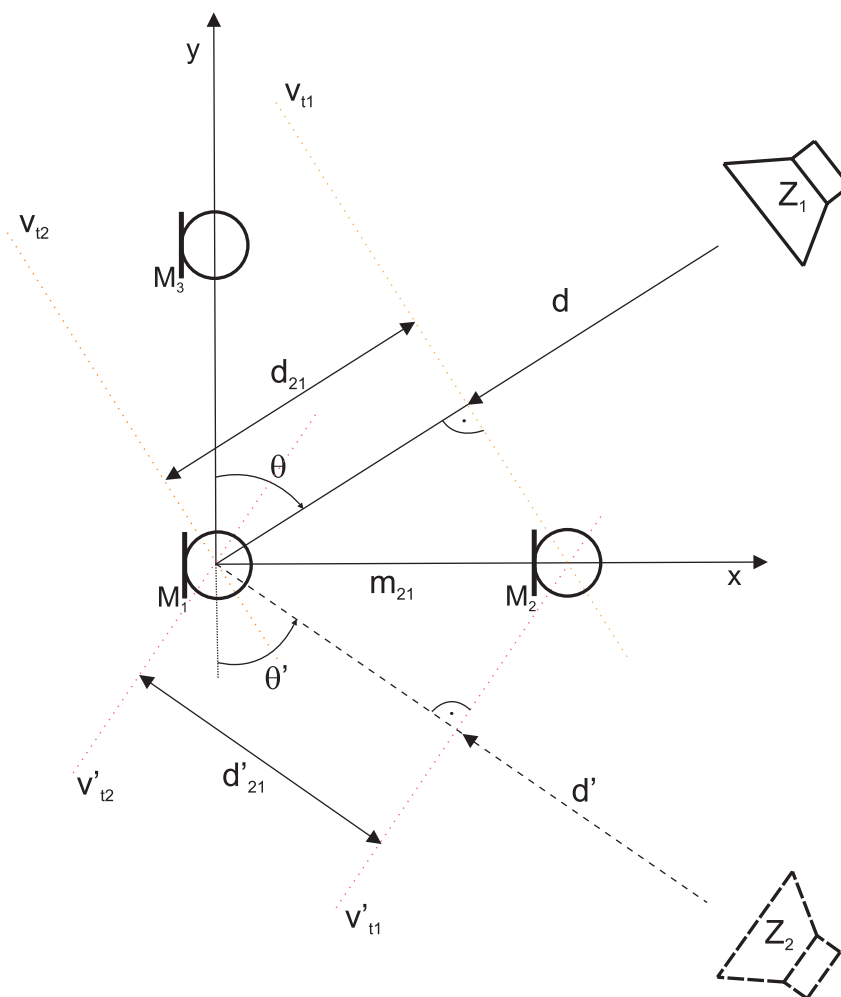
$$\theta = \arccos \frac{d_{31}}{m_{31_y}}. \quad (2.14)$$

2.3.3 Nejednoznačnost určení směru příchodu signálu

Pokud počítáme úhel právě ze dvou senzorů, nelze jednoznačně říci, zda tento úhel je právě směr příchodu signálu. Jak je naznačeno na obr. 2.9, kde \mathbf{d}' je alternativní směr DOA a v'_{ti} jsou akustické vlnoplochy v čase t , zdroj zvuku se může nacházet ve dvou polohách vzájemně zrcadlených k ose tvořené přímkou protínající příslušné senzory. Pokud vztáhneme určení úhlu vždy ke stejné ose, můžeme říci, že se úhly z příslušné dvojice senzorů budou vždy rovnat a budou mít pouze odlišná znaménka (tj. $+\theta$, $-\theta$). To znamená, že z každé dvojice senzorů jsme schopni určit právě dva směry příchodu signálu. Abychom mohli správně určit DOA, musíme úhly určit alespoň pro dvě dvojice senzorů a dopočítat k nim příslušné alternativní směry příchodu. Pokud nyní porovnáme všechny úhly DOA měřené ke stejné ose, měly by se nám právě dva úhly rovnat. Tento úhel je právě rovný směru příchodu signálu.

2.3.4 Metoda GCC PHAT

Tato metoda z velké části vychází z metody odhadu časových zpoždění TDOA, oproti této metodě však při jejím nasazení do reálných podmínek, kde se vyskytují odrazy, vykazuje přesnější výsledky. Metoda by měla být schopna lokalizovat zdroj i v prostředí s velkým výskytem šumu a je možno ji nasadit na úzkopásmové i širokopásmové signály. Nevýhodou této metody je nemožnost lokalizace více zdrojů zvukového signálu. Jako u metody odhadu časových zpoždění, metodu GCC PHAT je možné rozdělit do dvou kroků:



Obr. 2.9: Nejednoznačnost určení směru příchodu signálu.

1. určení zpoždění mezi senzory,
2. stanovení úhlu DOA.

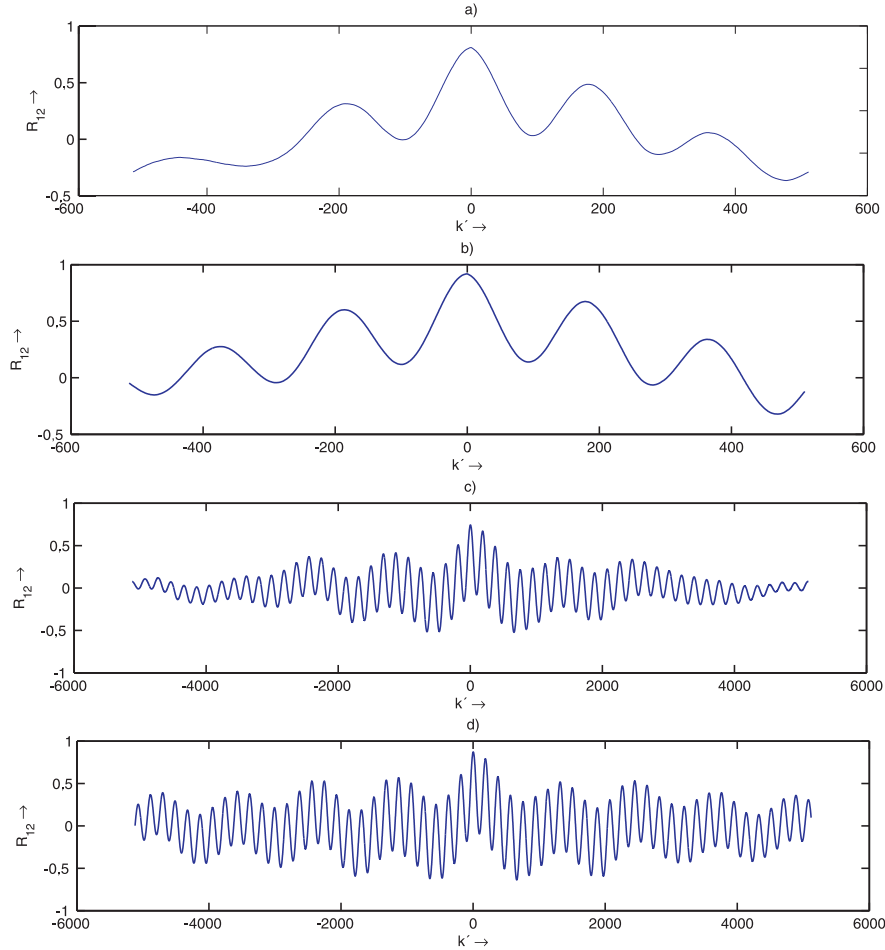
Při lokalizaci DOA touto metodou je možné použít lineární, rovinné i prostorové pole. Ve své první části se metoda při použití různých typů a struktur pole podstatně lišit nebude. Druhá část metody, tj. stanovení úhlu DOA, se bude lišit jen nepatrně a je zapotřebí vzít v úvahu geometrické postavení senzorů. Jelikož v našem případě bylo použito stejné rozmístění senzorů jako při metodě TDOA, stanovení úhlu DOA je totožné a bylo popsáno v části 2.3.2, proto již nebudeme znovu tuto problematiku popisovat.

Určení zpoždění mezi senzory

Jelikož metoda z velké části vychází právě z metody TDOA, musejí být dodrženy všechny předpoklady, které byly uvedeny v části 2.3 a podmínka (2.5) z části 2.3.1.

Korelaci GCC poté můžeme vyjádřit jako:

$$r_{12}(\tau) = \int_{-\infty}^{\infty} x_1(t + \tau)x_2(t)dt. \quad (2.15)$$



Obr. 2.10: Porovnání korelační funkce při zpracování metodou TDOA a GCC PHAT:

- a) *GCC PHAT (1 024 vzorků),*
- b) *křížová korelační funkce TDOA (1 024 vzorků),*
- c) *GCC PHAT (10 240 vzorků),*
- d) *křížová korelační funkce TDOA (10 240 vzorků).*

K ušetření výpočetního času je dobré veškeré operace se signály provádět ve frekvenční oblasti:

$$r_{12}(t) = \int_{-\infty}^{\infty} G_{X_1X_2}(f)e^{2\pi f}df = \int_{-\infty}^{\infty} X_1(f) \bullet (X_2^*(f))e^{2\pi f}df, \quad (2.16)$$

kde * značí komplexní sdružení, X_1 a X_2 jsou signály $x_1(t)$ a $x_2(t)$ vyjádřené přes Fourierovu transformaci (FFT). Toto vyjádření je velice podobné vyjádření křížové

korelační funkce v (2.6). Jelikož má tato metoda být odolná vůči odrazům, zavedeme pojem PHAT filtr, který řeší problematiku odrazů. PHAT filtr lze vyjádřit jako:

$$\psi(f) = |X_1(f) \bullet (X_2^*(f))|^{-1}. \quad (2.17)$$

Výsledný vztah poté můžeme zapsat jako:

$$R_{12}(t) = \int_{-\infty}^{\infty} \psi(f) G_{X_1 X_2}(f) e^{2\pi f t} df = \int_{-\infty}^{\infty} \psi(f) X_1(f) \bullet (X_2^*(f)) e^{2\pi f t} df, \quad (2.18)$$

tudíž:

$$R_{12}(t) = \frac{X_1(f) \bullet (X_2^*(f)) e^{2\pi f t}}{|X_1(f) \bullet (X_2^*(f))|}. \quad (2.19)$$

Pokud nyní z tohoto vztahu zjistíme absolutní extrém, tento extrém se rovná zpoždění ve vzorcích mezi dvěma senzory. Tento počet vzorků, o kolik je signál z dvojice jednoho senzoru zpožděn oproti druhému senzoru, lze snadno přepočítat na zpoždění v čase τ (2.7) nebo na zpoždění vzdálenostní d_{ij} (2.8). Po takto získaných zpoždění můžeme přistoupit k vypočtení úhlů azimutu θ a elevace ϕ směru příchodu signálu stejně jako v části kap. 2.3.2 [8], [9], [10], [11].

Porovnání křížové a zobecněné křížové korelace GCC PHAT je zobrazeno na obr. 2.10. Z něho je patrné, že u zobecněné křížové korelace GCC PHAT nevznikají tak výrazná okolní maxima jako u křížové korelační funkce, což znamená, že při hledání zpoždění je menší pravděpodobnost nalezení špatného maxima, které mohlo být způsobeno odrazem, než u metody TDOA.

3 DOSAŽENÉ VÝSLEDKY PŘI LOKALIZACI ZDROJŮ

3.1 Měřicí aparatura

Při měření byly použity mikrofony od značky Behringer s typovým označením ECM800. Mikrofony jsou zhotoveny jako kondenzátorové s elektretovou membránou a lineárním kmitočtovým rozsahem od 15 Hz do 20 kHz. Ke správné funkci musí být na mikrofon připojeno fantomové napájení v rozsahu +15 V až +48 V. Výstupní impedance mikrofonu je rovna 600 Ω , citlivost mikrofonu je -60 dB, odstupem signál - šum 78 dB a směrovou charakteristikou koule.



Obr. 3.1: Audio interface Event Layla, Mikrofon Behringer ECM800 a mikrofonní předzesilovač Behringer MIC2200.

Mikrofony byly uchyceny na konstrukci, která umožňovala sestavit 2-D i 3-D mikrofónové pole, a ta následně přichycena k točně obr. 3.2, která umožňovala natáčet mikrofony ke vztažnému bodu s rozsahem 0° až 360° s nastavitelným krokem, přičemž nejmenší krok je možné nastavit na hodnotu 5°.

Jako mikrofonní předzesilovače byly použity dvoukanalové elektronkové mikrofonní předzesilovače značky Behringer s typovým označením MIC2200. Předzesilovač obsahuje vypínatelné fantomové napájení +48 V, možnost otočení fáze, LED indikátory vybuzení, frekvenční rozsah 2 Hz až 200 kHz, regulátor zisku a parametrický ekvalizér. Předzesilovač má také galvanicky oddělené vstupy od výstupů.

Jako zdroje zvuku byly použity aktivní dvoupásmové reproduktorové soustavy.



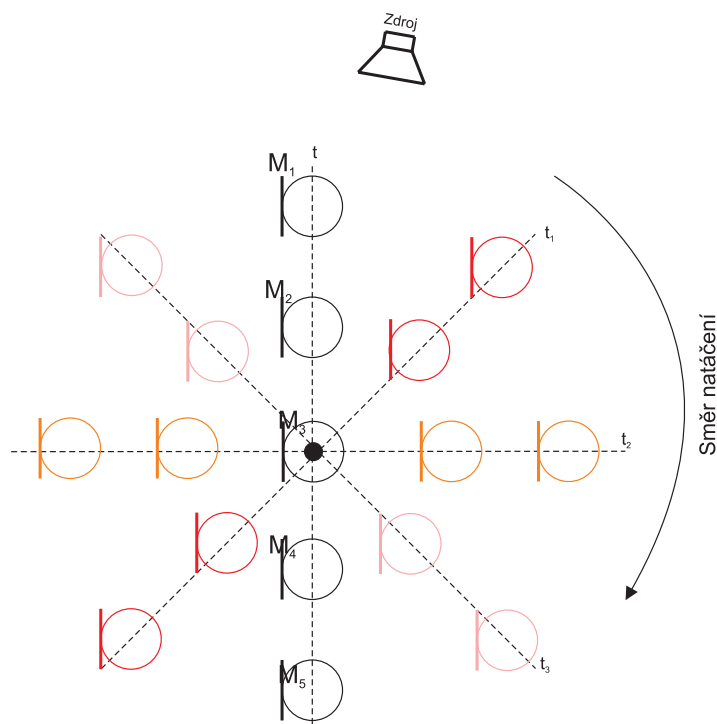
Obr. 3.2: Konstrukce pro uchycení mikrofonů na točně.

Pro záznam byl použit Multi-Track Recording System s A/D převodníky a interface k PC od Firmy Event/Echo s typovým označením Leyla Audio Digital Interface. Propojení převodníků a PC probíhá přes PCI sběrnici. Převodníky jsou osazeny nízkošumovými 24-bitovými obvody Motorola. Frekvenční rozsah interface je 10 Hz až 22 kHz s maximální vzorkovací frekvencí 50 kHz. Karta je schopna simultánně zaznamenat signály z osmi audio vstupů a přehrávat signály z deseti audio výstupů. Karta je také vybavena MIDI rozhraním, které jsme pro naše účely nevyužili. Jednotlivé signály byly synchronně zaznamenány do PC.

Jako softwarový nástroj byl použit muti-trackový editor Cubase 4 od firmy Steinberg, na kterém probíhal záznam, editace, analýza a následný export signálů. Zapojení měřicí aparatury je zobrazeno v příloze A.1.

3.1.1 Kalibrace měřicí aparatury

Některé metody lokalizace DOA, například metody založené na tvarování přijímací charakteristiky, vyžadují pro svou správnou funkčnost kalibraci mikrofonních zesilovačů. Správná kalibrace způsobí shodné zesílení signálu na všech výstupech z předzesilovačů. Kalibrace je prováděna v několika stupních pro co nejpřesnější výsledek. V první fázi musíme umístit mikrofony do jedné roviny se zdrojem zvuku a zaručit, aby všechny mikrofony byly od zdroje zvuku stejně vzdáleny ($\theta = 0$, $\phi = 0$). Dále je zapotřebí vybrat vhodný testovací signál. V běžné studiové praxi se využívají signály s průběhem sinus s frekvencí $f = 1$ kHz nebo bílý šum (white noise). Vybraný signál, v našem případě signál s průběhem sinus s frekvencí $f = 1$ kHz, bude přehrávat zdroj zvuku. Nyní k samotné kalibraci. V prvním kroku je nutno vyřadit, pokud je přítomen, equalizér na předzesilovači. To se provede nastavením zesílení všech pásem na hodnotu 0. Nyní nám led indikátory vybuzení ukazují přibližnou hodnotu vybuzení. Potenciometrem *Gain* umístěným u každého kanálu je možné

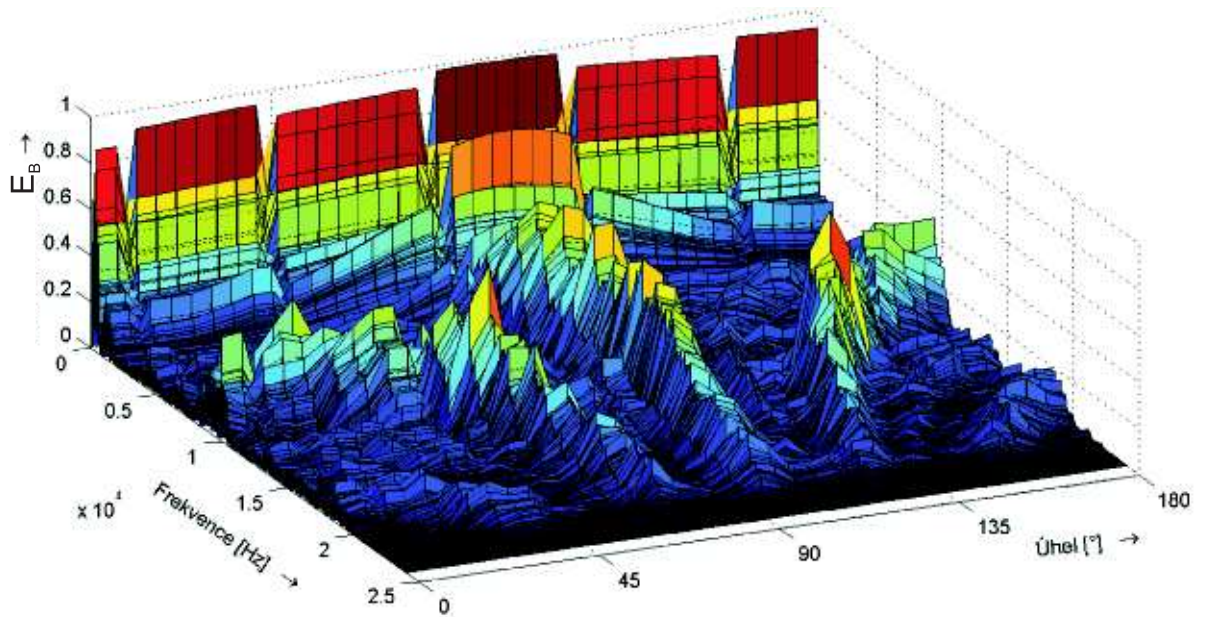


Obr. 3.3: Lineární uspořádání pole s naznačeným DOA.

toto vybuzení měnit. Nastavíme tedy potenciometry tak, aby indikátory na všech kanálech ukazovaly přibližně stejnou úroveň. Nyní naše hrubé nastavení zpřesníme v softwarovém editoru. Použitý software Cubase je doplněn také softwarovým mixážním pultem. Pokud ho zpřístupníme, máme před sebou zjednodušený mixážní pult, který splní naše požadavky. U každého tahového potenciometru se nachází softwarový indikátor vybuzení. Tento indikátor má šest režimů, které lze mezi sebou kombinovat. Každý indikátor nastavíme do režimu Fast Release a Meter Input. Tím způsobíme, že zobrazovaná úroveň signálu nebude ovlivněna žádnými softwarovými efekty, ekvalizéry ani tahovým potenciometrem. Pokud nyní znovu spustíme zdroj zvuku, indikátory by měly ukazovat stejnou úroveň číselně i graficky. Pokud tomu tak není, na příslušném kanálu znovu přenastavíme dle potřeby potenciometr *Gain*. V posledním kroku můžeme signály na vstupu převodníků zaznamenat a ve vhodném programu vypočítat jejich úroveň středního okamžitého výkonu dle vzorce:

$$P = \frac{1}{K} \sum_{s=0}^{K-1} x^2[k], \quad (3.1)$$

kde K je počet vzorků signálu a $x[k]$ posloupnost vzorků signálu. Úrovně středního okamžitého výkonu [6] každého signálu by se měly shodovat.



Obr. 3.4: Výsledný 3-D graf lokalizace metodou tvarování přijímací charakteristiky v bezodrazovém prostředí.

3.2 Výsledky metod v bezodrazovém prostředí

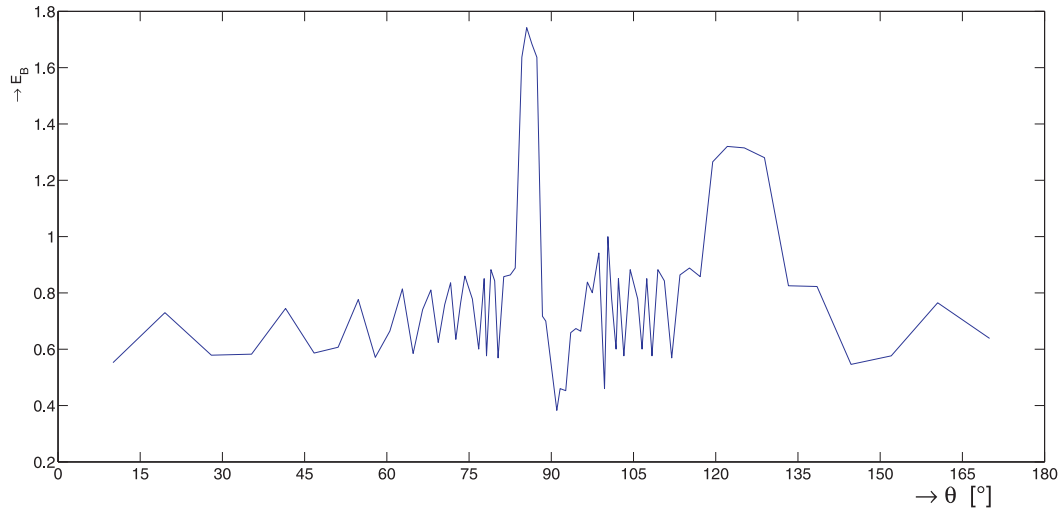
3.2.1 Metoda tvarování přijímací charakteristiky

Lokalizace jednoho zdroje zvuku

V této metodě bylo použito lineární pole s pěti mikrofony, viz obr. 3.3, vzdálených mezi sebou vždy o vzdálenost $q = 0,04$ m. Jako vztažný bod byl zvolen mikrofón M_3 , kolem kterého bylo mikrofónní pole natáčeno. Signál byl vzorkován frekvencí $f_{vz} = 48\,000$ Hz, váhové koeficienty nastaveny na $w_i = 0,20$, jako zdrojový signál byl použit bílý šum a pole s mikrofony bylo otáčeno s krokem 5° v úhlu $\theta = 0^\circ \div 180^\circ$. Bylo tedy zaznamenáno $36 \times 5 = 180$ signálů, které byly následně dosazeny do algoritmu tvořeným vztahy 2.1 a 2.2. Výsledek je zobrazen v trojrozměrném grafu (obr. 3.4), kde osy reprezentují azimut θ , frekvenci f a energii E_B . Správným úhlem směru příchodu signálu je úhel odpovídající umístění hlavního laloku přijetí energie. V našem případě bylo nejvíce energie E_B přijato v úhlu $\theta = 95^\circ$, což odpovídá umístění zdroje.

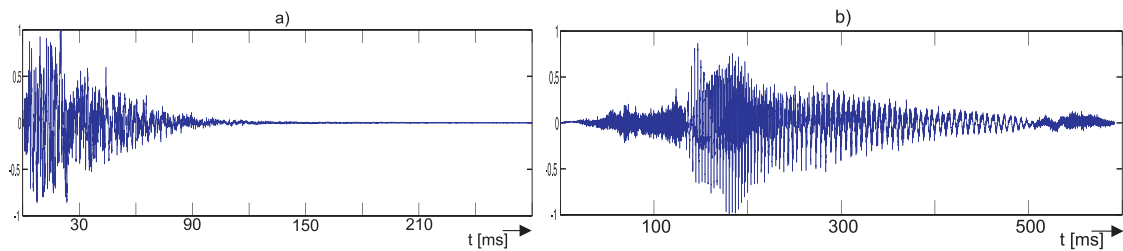
Lokalizace více směrů příchodu DOA

Při lokalizaci více zdrojů pomocí metody založené na tvarování přijímací charakteristiky bylo opět použito lineární pole s pěti mikrofony se vzdáleností $q = 0,04$ m.



Obr. 3.5: Výsledný 2-D graf lokalizace metodou tvarování přijímací charakteristiky v bezodrazovém prostředí pro frekvenci $f=1\,900$ Hz.

Vzorkovací frekvence byla nastavena na $f_{vz} = 48\,000$ Hz. Zdroje zvuku, které jsme lokalizovali, tvořily dva širokopásmové řečové signály. V prvním kroku lokalizace bylo počítáno s vahami $w_i = 0,20$. Pokud by přesnost úhlů v nalezených směrech DOA nebyla dostatečná, přistoupili bychom k zavedení fázových posuvů. Toto zpřesnění by ale vyžadovalo změnu vstupních signálů na úzkopásmové nebo zavedení vhodných pásmových propustí a signály by se musely zvláště fázově posouvat pro danou pásmovou propust.



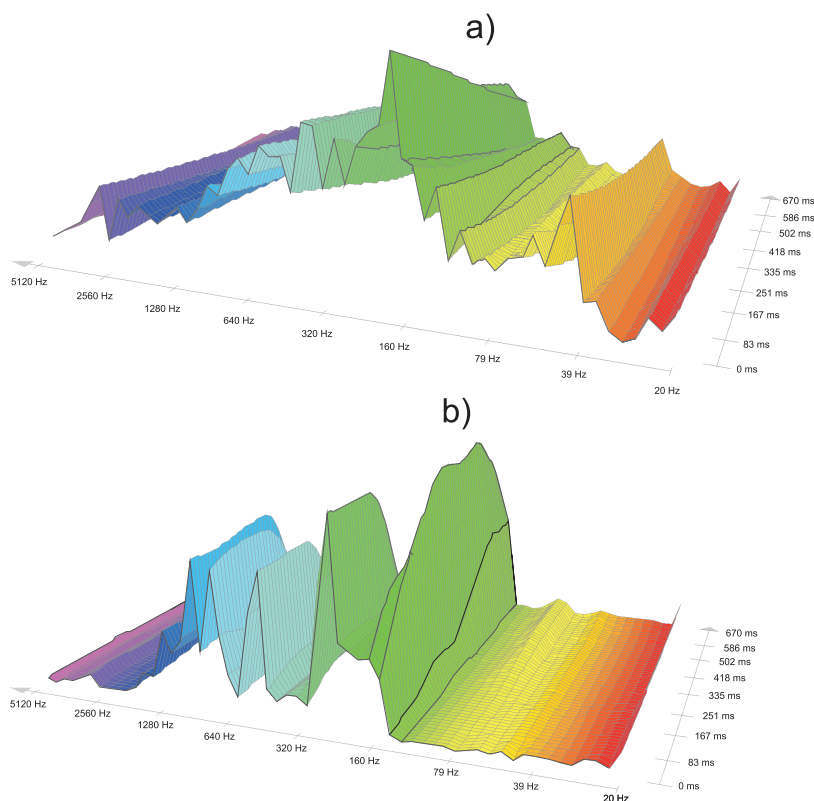
Obr. 3.6: Časový průběh testovacích signálů a) výstřel b) „five“

Na obr. 3.5 je zobrazen řez 3-D grafem odezvy pole pro frekvenci $f = 1\,900$ Hz. Maxima grafu odpovídají úhlům směru příchodu. Na obr. 3.5 se vyskytly maxima dvě a to maximum odpovídající $\theta_1 = 85^\circ$ a $\theta_2 = 128^\circ$. Tyto úhly odpovídají skutečným směřům DOA. Pro úhel směru DOA odpovídající θ_1 jsme schopni určit přesnost výsledku na $4,2^\circ$ a pro druhý směr příchodu DOA odpovídající úhlu θ_2 jsme výsledek schopni určit s přesností na $0,9^\circ$. Obě přesnosti jsou menší než je nejmenší možný krok točny, proto jsme do vztahu už nezaváděli zpoždění fázovým posunem, čímž

bychom mohli případně zpřesnit výsledek. Výpočetní algoritmus metody odhadu časových zpoždění byl implementován do prostředí MATLAB dle předešlého odvození. Výpočetní algoritmus byl založen v prvním kroku na získání zpoždění mezi jednotlivými mikrofony pomocí hledání maxima křížové korelační funkce (obr. 2.6) a poté byl proveden následný výpočet DOA postavený na vzorci 2.11. Cílem bylo vytvořit algoritmus, který by byl tvořen jen minimem cyklů, aby se přes svou malou výpočetní náročnost dal použít v reálném čase.

3.2.2 Metoda odhadu časových zpoždění

Abychom zamezili různým odrazům signálu, měření probíhalo v semi-bezodrazové komoře. Použito bylo rovinné pole tvaru pravoúhlého trojúhelníku se vzdáleností $q = 8$ cm a přeponou o velikosti 11,3 cm. Přesné rozmístění mikrofونů je zřetelné z obr. 2.5. Metoda byla měřena pro úhly v rozsahu $\theta_{\text{SKUT}} = -90^\circ$ až $\theta_{\text{SKUT}} = +90^\circ$



Obr. 3.7: Kmitočtové spektrum testovacích signálů a) výstřel b) „five“

s krokem rovným 15° . Signál byl vzorkován s frekvencí $f_{\text{vz}} = 48\,000$ Hz, kvantován na 16 bitů a rychlost zvuku byla nastavena na hodnotu $c = 340$ ms $^{-1}$. Jako vztahný mikrofون byl zvolen mikrofон M_1 , kolem kterého bylo celé pole natáčeno. Jako vstupní signály byly použity vzorky reprezentující slovo „five“ a ruch reprezentující *výstřel z revolveru ráže 9 mm* (obr. 3.6 a obr. 3.7). Úhel θ byl počítán pro

Tab. 3.1: Skutečný úhel a vypočítané úhly metodou TDOA.

	„five“		výstřel	
$\theta_{\text{SKUT}} [^\circ]$	$\theta_{m_{21}} [^\circ]$	$\theta_{m_{31}} [^\circ]$	$\theta_{m_{21}} [^\circ]$	$\theta_{m_{31}} [^\circ]$
-90	-76,89	-90	-62,2	-84,92
-75	-62,3	-74,59	-62,3	-74,59
-60	-52,83	-57,9	-52,83	-57,91
-45	-38,3	-44,9	-38,3	-44,9
-30	-26,27	-37,17	-26,27	-37,17
-15	-15	-27,69	-15,4	-27,69
0	0	-13,1	0	-27,69
15	15,4	27,69	15,4	27,69
30	26,27	37,17	26,26	37,16
45	45,09	51,69	45,099	51,69
60	62,3	69,25	52,83	69,25
75	76,89	79,8	76,89	79,8
90	76,89	95,08	76,89	95,08

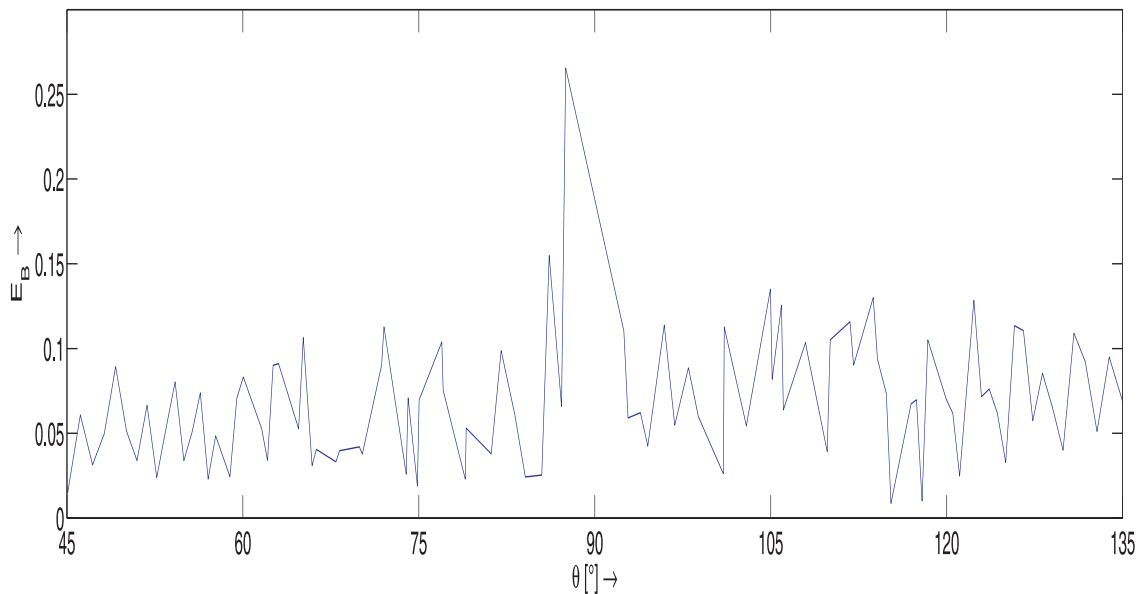
dvojice senzorů spojující vektor \mathbf{m}_{21} a \mathbf{m}_{31} . Přehled dosažených výsledků je uveden v tab. 3.1.

3.3 Výsledky metod v reálném prostředí

Veškerá měření metod v reálných podmínkách probíhaly v objektu ústavu Telekomunikací. Při tomto měření bylo nutno použít jiné HW zařízení pro převod A/D signálů. Jednalo se o zvukovou kartu Motu 8PRE vybavenou integrovanými předzesilovači a převodníky s maximální vzorkovací frekvencí 96 kHz s 24-bitovou hloubkou připojenou k PC přes rozhraní FireWire. Jako senzory byly opět použity mikrofony Behringer ECM800.

3.3.1 Metoda tvarování přijímací charakteristiky

S využitím pouze jednoho sensorového pole jsme touto metodou schopni lokalizovat pouze směr příchodu DOA a není možno přímo určit pozici mluvčího v prostoru. Pokud bychom chtěli získat přesnou polohu mluvčího, je nutné nasadit více sensorových polí. Při lokalizaci touto metodou bylo použito lineární pole o pěti mikrofonech se vzdáleností dvojice $q = 0,4$ m a vzorkovací frekvence A/D převodníků byla nastavena na 48 000 Hz. Jako testovací signál byl použit širokopásmový řečový vzorek



Obr. 3.8: Výsledný 2-D graf lokalizace metodou tvarování přijímací charakteristiky v reálném prostředí pro frekvenci $f = 508$ Hz.

mluvčího. Pole bylo opět jako v případě lokalizace více zdrojů metodou tvarování přijímací charakteristiky softwarově zpořádkováno, čímž jsme docílili natáčení pole a prohledání celé poloroviny. Jelikož prohledávací krok se většinou v oblastech pohyboval kolem úhlu $1,5^\circ$, nebylo do výpočtu zavedeno fázové zpoždění, čímž by mohlo být docíleno ještě větší přesnosti. Signál obsahoval největší spektrální hustotu na frekvenci $f = 508$ Hz. Na obr. 3.8 je zobrazena odezva pole E_B právě pro tuto frekvenci. Absolutní maximum grafu odpovídá vypočtenému úhlu $88,5^\circ$, přičemž zdroj se nacházel na úhlu $\theta_{SKUT} = 90^\circ$.

3.3.2 Metoda odhadu časových zpoždění

Při měření metodou TDOA byly použity typy rovinného pole ve tvaru rovnostranného trojúhelníku se vzdáleností $q = 0,2$ m a pole ve tvaru pravoúhlého trojúhelníku se vzdáleností $q = 0,2$ m a přeponou o velikosti $0,28$ m. Použitím těchto dvou odlišných polí zjistíme, zda měření v bezodrazovém prostředí bylo ovlivněno rozestavením mikrofónů. Vzorovací frekvence A/D převodníků byla nastavena na $44\,100$ Hz. Jako testovací signály byly zvoleny různé širokopásmové signály s odlišným rozložením spektrální hustoty v jednotlivých frekvenčních pásmech. Při výpočtu byly zpracovávány signály o velikosti $10\,240$, $1\,024$ a 512 vzorků.

Tab. 3.2: Závislost správného určení DOA na S/N pro testovací signál „Base“ pro geometrii pole do rovnoramenného trojúhelníku.

		TDOA		GCC PHAT	
θ [°]	S/N [dB]	$\theta_{m_{32}}$ [°]	$\theta_{m_{21}}$ [°]	$\theta_{m_{32}}$ [°]	$\theta_{m_{21}}$ [°]
0	-50	0	4,4217	1,1044	5,5303
	-45	0	4,4217	1,1044	5,5303
	-40	0	4,4217	1,1044	5,5303
	-35	0	4,4217	1,1044	5,5303
	-30	0	4,4217	1,1044	5,5303
	-25	0	4,4217	1,1044	5,5303
	-20	0	4,4217	1,1044	5,5303
	-15	0	4,4217	1,1044	5,5303
	-10	3,114	15,481	1,1044	10,582
	-5	7,514	15,481	1,1044	10,582
	0	8,211	18,451	1,1044	10,614

Rozložení do pravoúhlého trojúhelníku

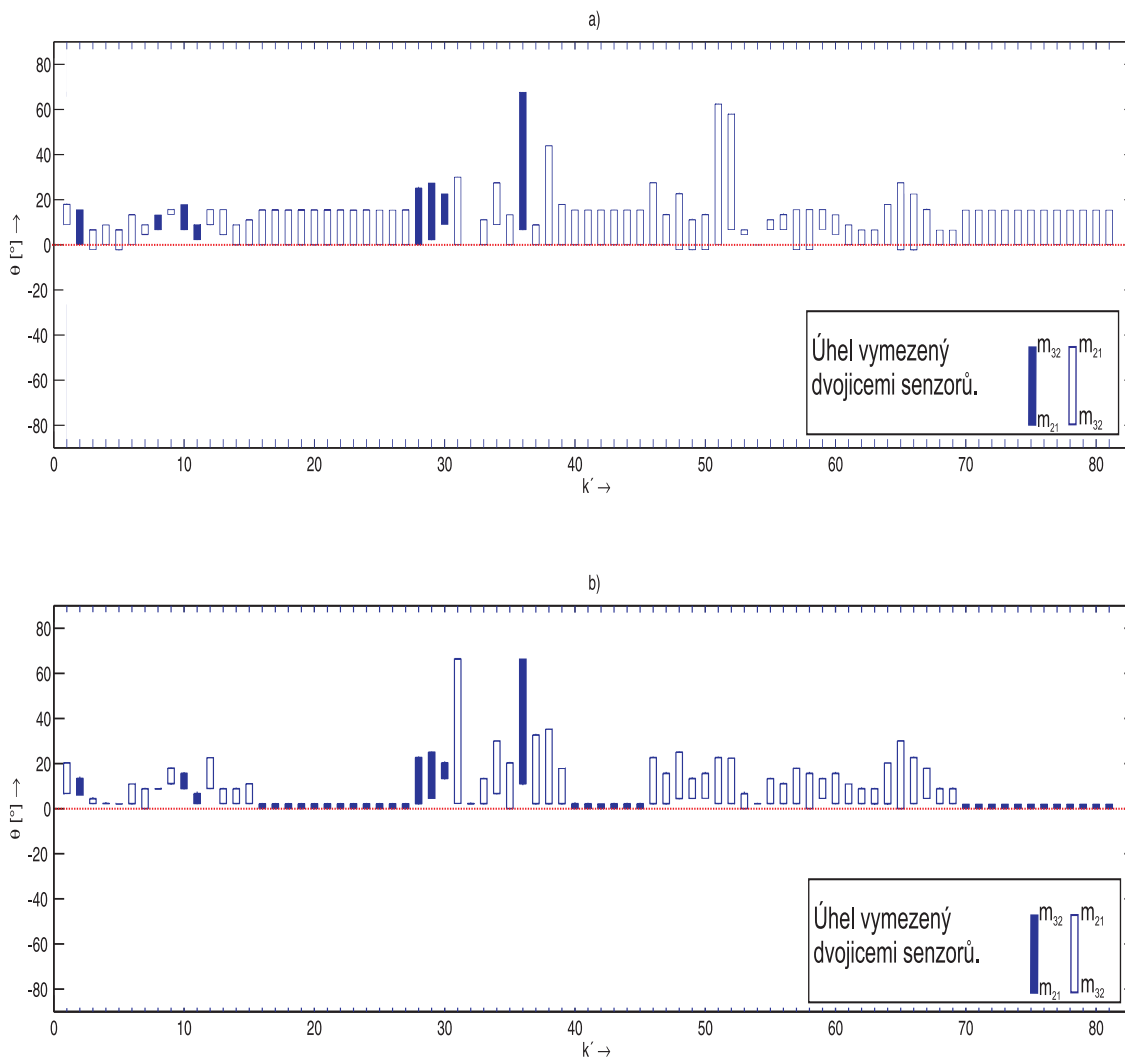
Rozmístění senzorů bylo zvoleno obdobně jako u měření v bezodrazovém prostředí dle obr. 2.5. Jako vztažný senzor byl zvolen senzor s označením M_1 . Funkčnost metody byla ověřena celkem na pěti různých DOA. V tab. B.1 jsou uvedeny výsledky měření úhlů DOA pro různé signály. Na obr. 3.9 jsou zobrazeny vypočtené úhly θ pro úhel směru příchodu DOA $\theta = 0^\circ$ se zpracovávaným počtem vzorků v bloku $k = 1024$ vzorků. Je zřejmé, že pro $k' = 36, 38, 51$ a 52 vzorků byl úhel určen špatně, což bylo způsobeno zpracováním bloku dat, jehož spektrální výkon v jednotlivých frekvenčních pásech byl zanedbatelný nebo byl srovnatelný s úrovní šumu.

Rozložení do rovnostranného trojúhelníku

Cílem změny geometrie pole bylo zjistit, jak velká je závislost přesností určení směru DOA při použití různého uspořádání pole. Jako v předchozím případě byl za vztažný senzor zvolen senzor s označením M_1 . V tab. B.2 jsou uvedeny výsledky této metody. Z tab. B.2 je patrné, že úhel θ , který je počítán z dvojice senzorů v zákrytu se zdrojem, je vypočten opět s odchylkou.

Prostředí s velkým šumovým pozadím

V některých reálných prostředích se může vyskytovat velké množství šumu či jiných akustických ruchů, které mohou narušit celý proces lokalizace DOA. Lokalizační



Obr. 3.9: Porovnání lokalizace metodou TDOA a GCC PHAT v reálných podmínkách se zpracovávaným blokem o velikosti $k = 1\ 024$.

metody jsou založeny na detekci nejhlasitějšího signálu. pokud by tedy v signálu začal převažovat hluk pozadí, metody by začaly lokalizovat právě zdroj hluku pozadí. Pokud ale i v tomto případě potřebujeme zdroj lokalizovat, je možné včlenit do lokalizačního algoritmu pásmovou propust, která musí potlačovat frekvenční pásma, které nejsou obsaženy ve zdroji signálu. Je však nutné upozornit, že je třeba znát charakter, respektive frekvenční rozsah hledaného zdroje signálu. Pokud tuto pásmovou propust zavedeme do hledacího algoritmu, metody lze použít i v těchto krajních případech. V tab. 3.2 a tab. 3.3 jsou zobrazeny výsledky lokalizace v prostředí se šumem. Jako hledaný signál bylo zvoleno slovo „Base“ a jako šum v pozadí byl zvolen bílý šum.

Tab. 3.3: Závislost správného určení DOA na S/N pro testovací signál „Base“ geometrii pole do rovnostranného trojúhelníku.

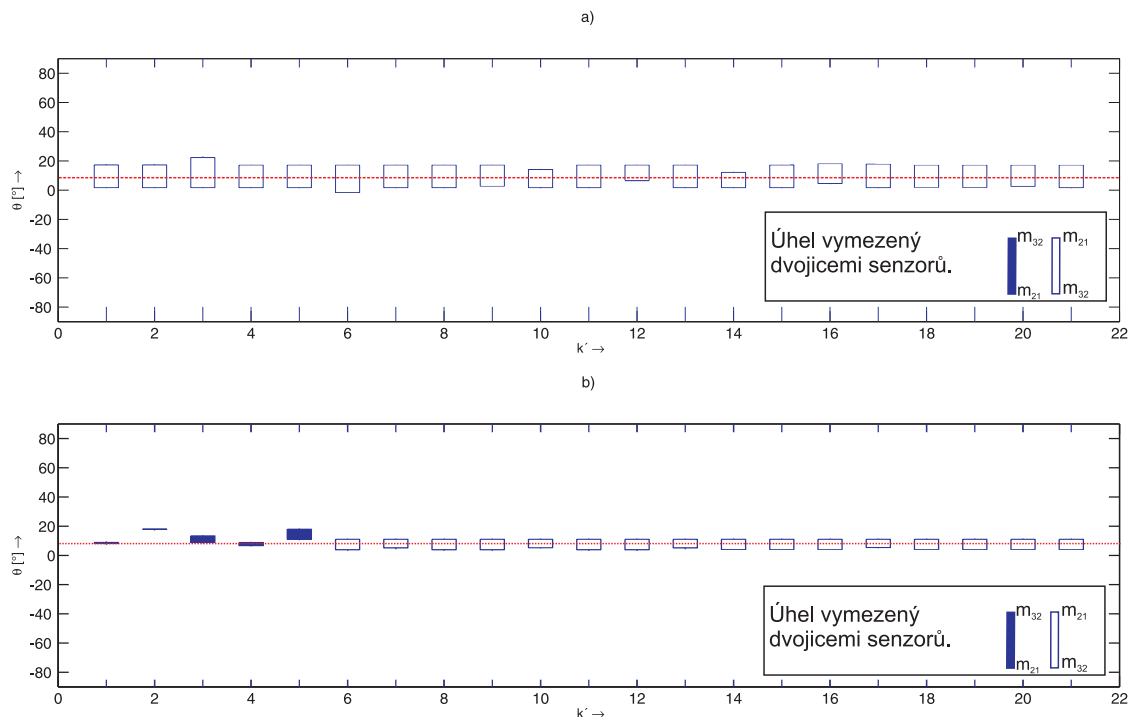
		TDOA		GCC PHAT	
θ [°]	S/N [dB]	$\theta_{m_{32}}$ [°]	$\theta_{m_{21}}$ [°]	$\theta_{m_{32}}$ [°]	$\theta_{m_{21}}$ [°]
-18	-50	-15,6548	-23,412	-16,805	-20,156
	-45	-15,6548	-23,412	-17,9622	-21,153
	-40	-15,6548	-23,412	-17,9622	-21,1416
	-35	-15,6548	-23,412	-16,805	-20,1562
	-30	-15,6548	-23,412	-16,805	-20,1566
	-25	-15,6548	-23,412	-16,805	-20,1567
	-20	-15,6548	-23,412	-16,805	-20,1563
	-15	-15,6548	-23,412	-16,805	-20,1561
	-10	-18,231	-27,541	-16,805	-20,1564
	-5	-20,158	-30,512	-16,805	-20,1561
0	-21,9748	-32,417	-16,805	-20,1563	

3.3.3 GCC PHAT

Při měření metody GCC PHAT bylo, pro lepší srovnání, použito pole ve tvaru pravoúhlého trojúhelníku (obr. 2.5) i rovnostranného trojúhelníku. Vypočtené úhly ve srovnání s metodou TDOA je zobrazeno v tab. B.1 a tab. B.2. Metoda vykazuje menší rozmezí hodnot, ve kterých by se DOA mělo pohybovat, viz obr. 3.9 a obr. 3.10. Je patrné, že tak jako u metody TDOA byl pro některé indexy k' úhel určen špatně (s velkým rozsahem), což bylo jako v předchozím případě způsobeno zpracováním bloku dat, jehož spektrální výkon v jednotlivých frekvenčních pásech byl zanedbatelný nebo byl srovnatelný s úrovní šumu.

Lokalizace přesného umístění

Jelikož v některých případech by bylo zapotřebí znát kromě směru příchodu signálu DOA ještě přesnou polohu zdroje signálu, byl proveden experiment, při kterém byla využita, pro své schopnosti odolávat odrazům, právě metoda GCC PHAT. Na stěny místnosti byly umístěny dvojice senzorů, které snímaly daný prostor. Vzdálenost dvojice senzorů byla nastavena na $q = 0,2$ m, přičemž k samotné lokalizaci je nutné znát přesné rozmístění mikrofونů. Jelikož dvojice senzorů jsou od sebe rozmístěny do značné vzdálenosti, výsledek je velmi přesný. Musíme vzít ale v úvahu fakt, že jedna dvojice senzorů nám nedokáže přesně určit směr příchodu signálu DOA z důvodu projevení nejednoznačnosti určení DOA (viz část 2.3.3) a nalezneme i směr



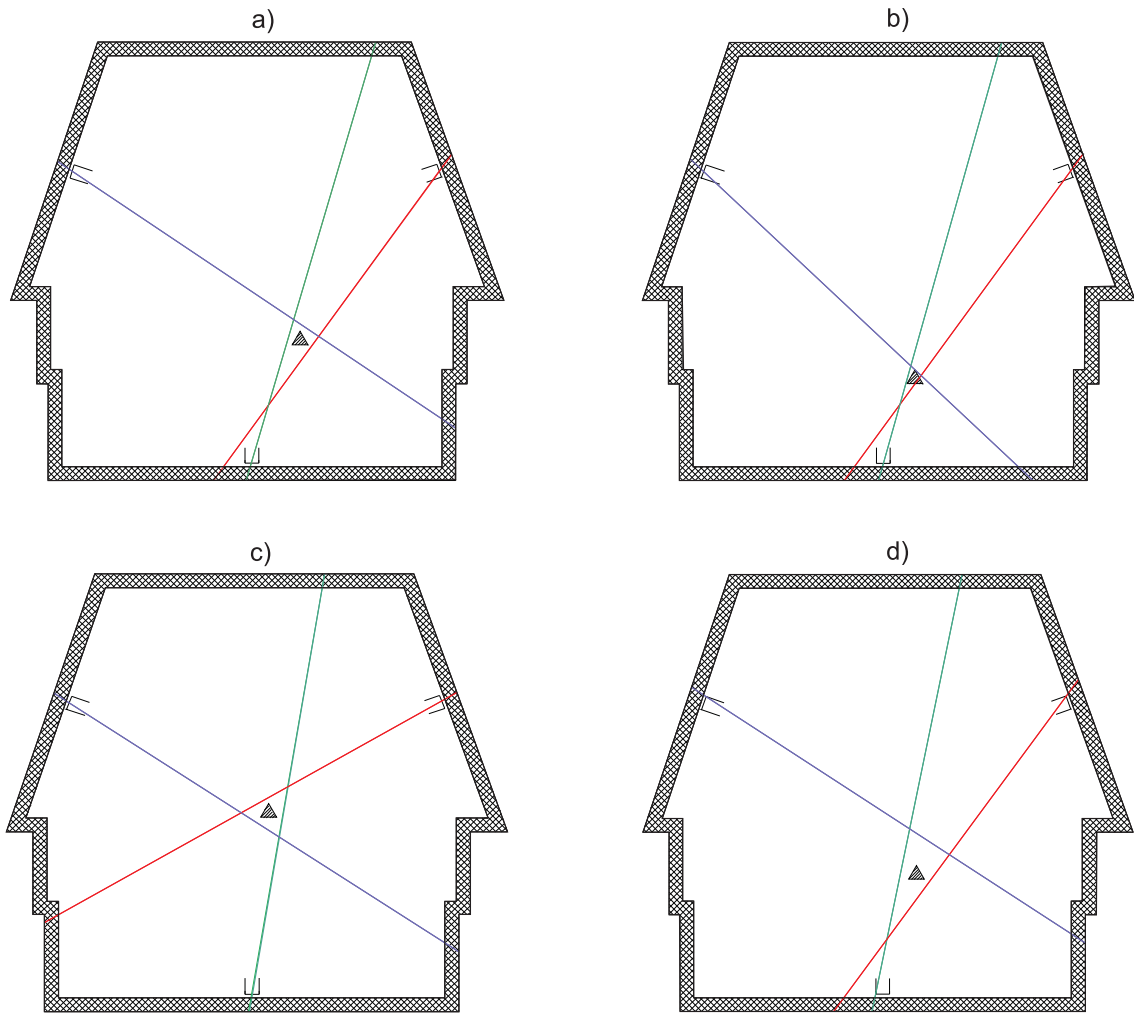
Obr. 3.10: Porovnání lokalizace metodou TDOA a GCC PHAT v reálných podmínkách se zpracovávaným blokem o velikosti $k = 512$.

alternativní. V porovnání s ostatními nalezenými DOA z dvojic senzorů jsme poté schopni určit výslednou polohu zdroje. Na obr. 3.11 jsou zobrazeny výsledky lokalizace přesného umístění pro testovací signál „jedna“ pro různé polohy mluvčího.

3.4 Shrnutí výsledků lokalizace DOA

Metoda tvarování přijímací charakteristiky

Při lokalizaci jednoho zdroje je z obr. 3.4 zřetelné, že hlavní lalok grafu odpovídá úhlu $\theta = 95^\circ$, což odpovídá směru příchodu signálu DOA. Jako testovací signál byl zvolen bílý šum, který se vyznačuje rovnoměrnou výkonovou spektrální hustotou. Odezva pole by tedy pro hlavní lalok grafu měla být konstantní. V našem případě však s rostoucí frekvencí klesá. Tento jev je zapříčiněn nerovností frekvenční charakteristiky zdrojů zvuku a vyšším pohlcováním vysokých kmitočtů ve vzduchu. Metoda byla výpočetně velmi náročná, tudíž by realizace této metody v reálném čase byla obtížná.



Obr. 3.11: Lokalizace přesného umístění metodou GCC PHAT.

Lokalizace více směrů příchodu DOA

Při lokalizaci více zdrojů se podařilo správně určit oba směry příchodu signálu. Oba úhly θ byly určeny s přesností lepší, než byl minimální krok točny, tudíž se do algoritmu nezavádělo fázové zpoždění. Výsledné směry jsou zobrazeny na obr. 3.5 jako absolutní maxima odpovídající úhlům $\theta_1 = 85^\circ$ a $\theta_2 = 128^\circ$. Pro lepší přehlednost a čitelnost úhlů odpovídajícím DOA byl graf proveden pouze pro jednu frekvenci $f = 1\,900$ Hz. Pokud v metodě tvarování přijímací charakteristiky budeme pracovat pouze s časovým zpožděním, je možné tuto metodu aplikovat na širokopásmové i úzkopásmové signály. Pokud do metody zavedeme zpoždění fázové, metoda je vhodná pouze pro zpracování signálů úzkopásmových. Pokud víme alespoň přibližně, jaké frekvenční spektrum bude signál obsahovat, algoritmus můžeme upravit právě na prohledávání pouze těchto frekvenčních pásem, a tím podstatně snížit výpočetní náročnost a dobu potřebnou ke zpracování.

V reálném prostředí byl pro lokalizaci pomocí metody tvarování přijímací charakteristiky použit právě algoritmus, který dokáže lokalizovat více zdrojů. Na obr. 3.8 je zobrazen výsledný směr příchodu signálu jako absolutní maximum. Výsledný úhel se od skutečného lišil přibližně o $1,5^\circ$. Pokud bychom do výpočtu zavedli ještě fázové zpoždění, výsledek by se mohl ještě zpřesnit, ale prodloužila by se výpočetní doba.

Metoda odhadu časových zpoždění

Z hlediska výpočetní náročnosti všech testovaných metod byla nejméně náročná právě metoda odhadu časových zpoždění a metoda GCC PHAT. V tab. 3.1 je uveden přehled dosažených výsledků z bezodrazového prostředí. Je patrné, že ve většině případů jsme byli schopni správně určit pouze úhel z jedné dvojice senzorů která není v zákrytu se zdrojem. Úhel, který byl vypočítán z druhé dvojice senzorů, obsahuje odchylku. Měření bylo prováděno i pro jiné testovací signály, abychom vyloučili spojitost mezi zdrojovým signálem a odchylkou. Dále bylo vyzkoušeno uchytit mikrofony v dané pozici do jiných směrů, čímž bychom alespoň nepatrně omezili odrazy od spodní desky točny. Nepřesný výsledek z jedné dvojice senzorů je patrně způsoben ohýbáním zvuku kolem senzorů. Bylo zjištěno, že při natáčení mikrofonového pole je nesprávně určeno zpoždění mezi jednou dvojicí senzorů, což má za následek nesprávně vypočítaný úhel. Při počítačových simulacích v teoretickém prostředí se tyto problémy nevyskytovaly. I když měření probíhalo v semi-bezodrazové komoře, odraz a ohyb zvuku mohl vzniknout mezi konstrukcí točny a mikrofonním polem.

V reálném prostředí byly použity typy pole s geometrickým uspořádáním do rovnostranného a pravoúhlého trojúhelníku. Cílem použití těchto polí bylo porovnání lokalizovaných směrů DOA a zjištění, zda problém, který nám způsoboval odchylku z jedné dvojice senzorů, je způsoben odrazem a ohybem zvuku. Z tabulek B.1 a B.2 vyplývá, že odchylka při použití různých polí není konstantní a mění se v závislosti na směru DOA. Tudíž by odchylka měla být způsobena právě těmito dvěma fyzikálními jevy. V tab. 3.2 a tab. 3.3 jsou uvedeny výsledky měření, kde jsme ověřovali schopnost využití metody i ve velmi zašumělém prostředí. Z naměřených výsledků je patrné, že vypočtené úhly z dvojic senzorů jsou až do $S/N = -15$ dB konstantní a nad tuto úroveň se vypočtené úhly θ začínají lišit.

GCC PHAT

Metoda GCC PHAT byla aplikována právě pro svou schopnost odolávat odrazům od stěn a předmětů v místnosti. Na obr. 2.10 je zobrazeno porovnání křížové korelační funkce použité v metodě TDOA a zobecněné křížové korelace GCC PHAT. V tab. B.1 a tab. B.2 jsou zaznamenány výsledné vypočtené úhly dvojic senzorů. Pokud porovnáme vypočtené úhly metodou TDOA a GCC PHAT na obr. 3.9 a obr. 3.10,

je patrné, že metoda GCC PHAT vždy určila přesnější rozmezí úhlu, ve kterém se směr příchodu signálu DOA pohybuje. Přesnějšího výsledku směru příchodu signálu DOA bylo také možné dosáhnout menším rámcem zpracovávaných dat. Na obr. 3.9 byly zpracovávány data v blocích o velikosti $k' = 1\,024$ vzorků, což při vzorkovací frekvenci $f_{vz} = 44\,100$ Hz odpovídá času $t = 23$ ms. I v tomto případě byl proveden test schopnosti nasazení metody v prostředí s velkým obsahem šumu. Výsledky a porovnání s metodou TDOA jsou uvedeny v tab. 3.2 a tab. 3.3, z čehož je patrné, že metoda GCC PHAT dokázala najít menší interval, ve kterém se směr příchodu DOA nachází, přičemž v tab. 3.3 pro úhel $\theta = 0^\circ$ metoda GCC PHAT našla stejný úhel až do hodnoty $S/N = 0$ dB.

4 ZÁVĚR

V této bakalářské práci byla zkoumána problematika lokalizace statických akustických zdrojů polem mikrofonů v bezodrazovém i reálném prostředí.

Vybrané metody a jejich modifikace byly implementovány v prostředí MATLAB, kde se ověřovala funkčnost odvozených teoretických postupů, vhodnost metod pro lokalizaci širokopásmových nebo úzkopásmových signálů, schopnost lokalizace více směrů příchodu signálu, přesnost a možnost nasazení v prostředích s velkým obsahem šumu. Dále byly v tomto prostředí navrženy funkce pro práci s vektory a maticemi, které budou přínosné i v další práci se signály.

Nejlepších výsledků lokalizace se dostavovaly u metod založených na tvarování přijímací charakteristiky. Podařilo se úspěšně lokalizovat jak jeden i směry příchodu dva. Jelikož ale bylo nutno výpočty odezvy pole E_B opakovat pro každý úhel, metoda tvarování přijímací charakteristiky je oproti metodě odhadu časových zpoždění více výpočetně náročná. Rozdíl, o který se maximálně může vypočtený úhel lišit od úhlu skutečného, je vždy roven prohledávacímu kroku metody v daném intervalu. Námi nalezený směr příchodu signálu DOA se od skutečného umístění lišil o $\theta = 1,5^\circ$.

U metody odhadu časových zpoždění TDOA se úhel podařilo vždy určit pouze v rozsahu vymezeném úhly dvojic senzorů. Dvojice senzorů, které byly s daným směrem DOA v zákrytu obsahovaly odchylku větší. Tento jev je s největší pravděpodobností způsoben odrazem a ohybem zvuku od jednotlivých senzorů či stěn. Metoda by dokázala lokalizovat zdroj signálu i v prostorách s větším obsahem aditivního bílého šumu až do hodnoty $S/N = -15$ dB. Nad tuto hodnotu dochází již k velkým nepřesnostem ve výpočtu úhlu. Nejlepších výsledků metoda dosahovala při zpracování rámců dat v rozmezí 512 až 1 024 vzorků, kde se výsledný rozsah výskytu směru příchodu signálu pohyboval kolem 15° , přičemž příčina toho rozsahu je popsána výše. Pokud bychom ale u této metody nasadili pole o větších rozměrech, měli bychom být schopni tyto odchylky z velké části eliminovat.

Metoda GCC PHAT v porovnání s metodou TDOA vykazovala mnohem lepší výsledky. Vypočtené úhly vymezovaly rozsah kolem 5° . Metodu je možné nasadit i do prostředí s velkým obsahem šumu, přičemž tato metoda zůstává přesnější než metoda TDOA. Touto metodou jsme také provedli experiment nalezení přesného umístění zdroje signálu, což by nám pomohlo získat další informace o zdroji. Vždy se nám podařilo se značně přesným výsledkem polohu zdroje nalézt.

LITERATURA

- [1] EKSLER, V. *Prostorová lokalizace a separace naslepo zdrojů akustických signálů polem mikrofónů : doktorská disertační práce*. Brno : Vysoké učení technické v Brně, 2006. 98 s.
- [2] VLACHÝ V. *Praxe zvukové techniky* Praha : Muzikus, 2008. 297 s. ISBN 80-86253-05-8.
- [3] KÁŇA, L. *Elektroakustika*. Brno : Vysoké učení technické, Fakulta Elektrotechniky a Komunikačních technologií, 2002. 91 s.
- [4] BRANDSTEIN, M., WARD, D. *Microphone Arrays: Signal Processing Techniques and Applications*. Berlín : Springer-Verlag, 2001. 398 s. ISBN 3-540-41953-5.
- [5] GRÉNAR, M. *Tvarování přijímací charakteristiky mikrofónového pole : bakalářská práce*. Brno : Vysoké učení technické v Brně, 2009. 42 s.
- [6] ZAPLATÍLEK, K., DOŇAR, B. *MATLAB: Začínáme se signály*. 1. vyd. Praha : BEN, 2006. 272 s. ISBN 80-7300-200-0.
- [7] RAJMÍČ, P. *Prostá metoda časových posunů a modifikovaná metoda časových posunů pro detekci směru přicházejícího zvuku*. [online]. ElektroRevue, 2002 [cit. 2009-11-07]. Dostupný z WWW: <<http://www.elektrorevue.cz/clanky/02066/index.html>>.
- [8] FERRAS FONT, M. *Multi-Microphone Signal Processing for Automatic Speech Recognition in Meeting Rooms : diplomová práce*. Berkeley, California : ICSI, 2005. 110 s.
- [9] VARMA, K. *Time-Delay-Estimate Based Direction-of-Arrival Estimation for Speech in Reverberant Environments : Diplomová práce*. Blacksburg, VA : The Bradley Department of Electrical and Computer Engineering Virginia Polytechnic Institute and State University in partial fulfillment of the requirements, 2002. 127 s.
- [10] QIN, B., et al. *Sound Source Localization in Domestic Environment*. [online]. IEEE Xplore, 2008, [cit. 2010-03-08]. Dostupný z WWW: <<http://www.ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4697676>>.
- [11] DiBiase J. H. *A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays : doktorská disertační*

práce. Providence, Rhode Island : Division of Engineering at Brown University,
2000. 112 s.

SEZNAM ZKRATEK

2-D	Dvojměrný prostor/pole – Two dimensional space/array
3-D	Trojměrný prostor/pole - Three dimensional space/array
BF	Tvarovač – Beamformer
DAS	Tvarovač typu Delay And Sum
DOA	Směr příchodu signálu – Direction of Arrival
MV	Spektrální odhad minimální odchylky - Minimum Variance
MUSIC	Metoda vícenásobné klasifikace signálu - Multiple Signal Classification
SSL	Lokalizace zvuku – Sound source localization
TDOA	Metoda odhadu časových zpoždění - Time Difference of Arrival
UA	Rovnoměrné pole - Uniform Array
ACH	Bezodrazová Komora - Anechoic chamber
GCC	Obecná křížová korelace - Generalized Cross Correlation
PHAT	Fázová transformace - Phase Transform
FFT	Fourierova transformace - Fourier Transform
A/D	Analogově digitální převodník - Analog to Digital Converter
S/N	Odstup signál-šum - Signal to Noise Ratio

SEZNAM SYMBOLŮ A VELIČIN

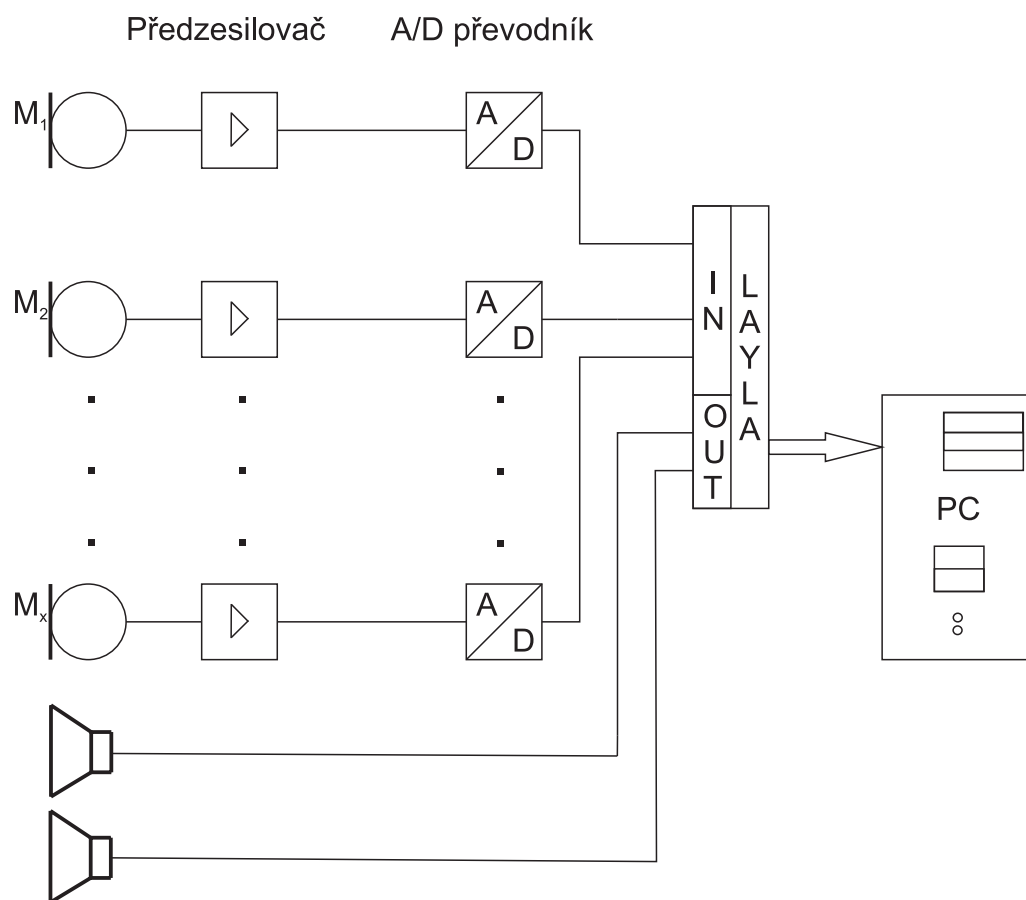
M	matice reprezentující rozmístění senzorů v poli ($3 \times M$)
Q	matice reprezentující vektory mezi dvojicemi senzorů ($3 \times Q$)
R	matice tvořená vektory R_{iy} ($Q \times K$)
α	koeficient pohltivosti
β	úhel dopadu zvuku
γ	úhel natočení pole
ς	koeficient zvukové průzvučnosti
ϕ	úhel elevace
λ_{\min}	vlnová délka maximální zpracovávané frekvence
θ	úhel azimutu
τ_{ij}	časové zpoždění signálu dvojice senzorů
E	zvuková energie
E_B	odezva pole pro všechny senzory
I	intenzita zvuku
K	délka signálu ve vzorcích
M	počet senzorů v poli
M_i	určení fyzického senzoru v poli
P	úroveň okamžitého středního výkonu
Q	počet vektorů spojující všechny dvojice senzorů v poli
R_{iy}	vzájemná korelační funkce dvou signálů
S	obsah
c	rychlost zvuku
\mathbf{d}	vektor reprezentující směr příchodu signálu
d_{ij}	rozdíl vzdáleností dopadu signálu mezi i -tým a j -tým senzorem

f	frekvence
f_{\max}	maximální zpracovávaná frekvence
f_{vz}	vzorkovací frekvence
k'	zpoždění ve vzorcích
k	počet vybraných vzorků
t	čas
\mathbf{m}_x	souřadnice x -tého senzoru
\mathbf{m}_{ij}	souřadnice vektoru mezi dvěma senzory
m_{ijx}	x -ová souřadnice vektoru mezi dvěma senzory
q	vzdálenost mezi senzory
q_{\max}	maximální vzdálenost dvou senzorů
r	koeficient odrazivosti
v_{faz}	fázová rychlost
w_i	váha i -tého mikrofonu
x_i	vektor vstupního signálu
$x[k]$	posloupnost vzorků signálu
y_i	odezva pole pro daný i -tý senzor

SEZNAM PŘÍLOH

A	Měřicí aparatura	60
B	Tabulky	61
C	Obsah přiloženého DVD	63

A MĚŘICÍ APARATURA



Obr. A.1: Zapojení měřicí aparatury.

B TABULKY

Tab. B.1: Skutečný úhel a vypočítané úhly metody TDOA s rozložením do pravoúhlého trojúhelníku v reálných podmínkách se zpracovávaným blokem vzorků o velikosti $k = 10\,240$.

		TDOA		GCC PHAT	
$\theta_{\text{SKUT}} [^\circ]$	SIGNAL	$\theta_{m_{32}} [^\circ]$	$\theta_{m_{21}} [^\circ]$	$\theta_{m_{32}} [^\circ]$	$\theta_{m_{21}} [^\circ]$
0	„base“	0	-4,4217	1,1044	5,5303
	„seconds“	0	-15,22	1,1044	10,582
	„Asia“	0	-17,9622	1,1044	19,1271
	„Amerika“	0	-36	1,1044	34,0245
	„jedna“	0	-39,5585	0	35,9506
-18	„base“	-15,6548	-22,306	-16,805	-19,1881
	„seconds“	-15,6548	-22,306	-16,805	-19,1881
	„Asia“	-15,6548	-22,674	-17,9622	-25,0897
	„Amerika“	-13,3732	-46,0622	-15,6548	-42,9098
	„jedna“	-15,6548	-22,306	16,805	19,1881
-65	„base“	-62,4517	-67,326	-67,694	-64,9103
	„seconds“	-62,4517	-67,326	-64,9442	-66,1238
	„Asia“	-58,0026	-67,326	60,1518	66,1238
	„Amerika“	-50,4415	-69,6998	52,2088	68,5177
	„jedna“	-43,9378	-72,0378	45,4919	70,8729
18	„base“	2,2092	22,67	1,1044	23,8762
	„seconds“	13,37	22,67	12,2406	23,8762
	„Asia“	11,113	22,674	8,8701	25,0897
	„Amerika“	6,6409	35,3262	4,4217	38,0814
	„jedna“	8,8701	35,3262	6,6409	38,0814
65	„base“	62,4517	67,326	60,1518	66,1238
	„seconds“	62,4517	67,326	58,0026	64,9103
	„Asia“	58,0026	67,326	55,9755	66,1238
	„Amerika“	58,0026	69,6998	54,0494	67,326
	„jedna“	47,0902	67,326	45,4919	66,1238

Tab. B.2: Skutečný úhel a vypočítané úhly metody TDOA s rozložením do rovnostranného trojúhelníku v reálných podmínkách se zpracovávaným blokem vzorků o velikosti $k = 10\,240$.

		TDOA		GCC PHAT	
$\theta_{\text{SKUT}} [^\circ]$	SIGNAL	$\theta_{m_{32}} [^\circ]$	$\theta_{m_{21}} [^\circ]$	$\theta_{m_{32}} [^\circ]$	$\theta_{m_{21}} [^\circ]$
0	„base“	0	11,113	13,3732	2,2092
	„seconds“	0	3,5483	7,694	2,2092
	„Asia“	0	7,6998	7,326	2,2092
	„Amerika“	0	-9,214	9,5585	2,2092
	„jedna“	0	9,5585	6,9421	2,2092
-18	„base“	-11,4214	-18,514	-15,4214	-18,514
	„seconds“	-11,4214	-19,214	-15,4214	-19,214
	„Asia“	-11,2451	-22,674	-15,2451	-22,674
	„Amerika“	-11,3732	-17,244	-13,3732	-17,244
	„jedna“	-10,8712	-18,452	-15,8712	-18,452
-65	„base“	-58,4517	-66,4356	-62,4517	-64,356
	„seconds“	-58,4517	-64,514	-62,4517	-64,214
	„Asia“	-67,4524	-65,4789	-61,1524	-65,215
	„Amerika“	-57,321	-67,415	-58,874	-67,415
	„jedna“	-58,4114	-64,1684	-60,414	-64,514
18	„base“	8,6092	18,541	10,6092	18,541
	„seconds“	13,742	17,541	13,987	17,541
	„Asia“	15,113	22,741	15,113	17,4741
	„Amerika“	16,1449	25,4162	16,1449	19,145
	„jedna“	18,8701	25,3262	14,8701	19,3262
65	„base“	62,4517	64,326	58,146	61,521
	„seconds“	62,4517	64,326	59,541	62,121
	„Asia“	58,0026	64,326	58,0026	60,894
	„Amerika“	58,0026	65,6998	58,0026	62,187
	„jedna“	49,0902	61,326	53,784	61,326

C OBSAH PŘILOŽENÉHO DVD

Součástí Bakalářské práce je i DVD, které obsahuje tyto adresáře:

- **DAS** - obsahuje zdrojové soubory navržených algoritmů pro lokalizaci jednoho i více zdrojů metodou tvarování přijímací charakteristiky popsané v části 2.1 pro prostředí MATLAB.
- **FCE** - obsahuje jednoduché funkce pro práci se signály.
- **GCC-PHAT** - obsahuje zdrojové soubory navrženého algoritmu pro lokalizaci metodou GCC-PHAT popsané v části 2.3.4 pro prostředí MATLAB.
- **PRESNE-UMISTENI** - obsahuje zdrojové soubory navrženého algoritmu pro lokalizaci metodou GCC-PHAT a TDOA pro lokalizaci přesného umístění zdroje pro prostředí MATLAB.
- **SIGNALS** - obsahuje zvukové vzorky testovaných zdrojových signálů ve formátu *.wav.
- **TDOA** - obsahuje zdrojové soubory navrženého algoritmu pro lokalizaci metodou TDOA ve 2-D i 3-D popsané v části 2.3 pro prostředí MATLAB.