

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

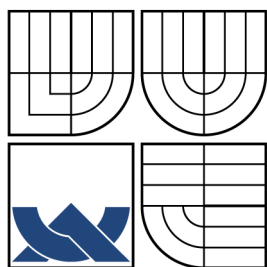
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

LOKALIZACE AKUSTICKÉHO ZDROJE V BLÍZKÉM POLI

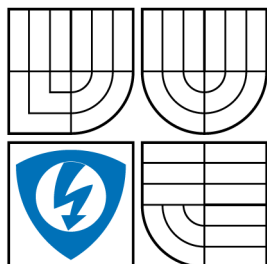
BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

DAVID ŠURÁNEK



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 AKUSTICKÉHO ZDROJE V BLÍZKÉM POLI LOCALIZATION OF AN ACOUSTIC SOURCE IN NEAR FIELD SITUATION

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

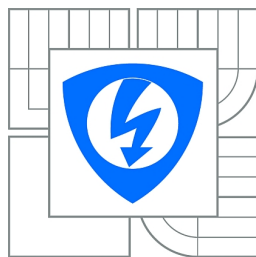
AUTOR PRÁCE
AUTHOR

DAVID ŠURÁNEK

VEDOUcí PRÁCE
SUPERVISOR

Ing. IVAN MÍČA

BRNO 2011



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Bakalářská práce

bakalářský studijní obor
Teleinformatika

Student: David Šuránek

ID: 115290

Ročník: 3

Akademický rok: 2010/2011

NÁZEV TÉMATU:

Lokalizace akustických zdrojů v blízkém poli

POKYNY PRO VYPRACOVÁNÍ:

Prostudujte metody lokalizace širokopásmových i úzkopásmových akustických zdrojů v blízkém poli mikrofónového pole. Zvolené lokalizační algoritmy implementujte v jazyce C nebo Matlab a srovnajte přesnost a spolehlivost lokalizace i výpočetní náročnost.

DOPORUČENÁ LITERATURA:

[1] BENESTY, J., CHEN, J., YITENG, H. Microphone Array Signal Processing. Berlin: Springer-Verlag, 2008. 250 s. ISBN 978-3-540-78612-2.

[2] EVEREST, F. A., The Master Handbook of Acoustics. McGraw/Hill, 2009. 528 s. ISBN 978-0071603324.

Termín zadání: 7.2.2011

Termín odevzdání: 2.6.2011

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.

ABSTRAKT

Tato práce se zabývá lokalizací akustického zdroje v blízkém poli. Úvod této práce se zabývá základním zkoumáním problematiky lokalizace akustických zdrojů. Následně jsou přiblíženy některé známé metody lokalizace akustických zdrojů. Další část se věnuje implementaci, praktickému ověření funkčnosti metod, srovnání jejich úspěšnosti a výsledkům lokalizace. V závěru jsou shrnuty dosažené výsledky.

KLÍČOVÁ SLOVA

zdroj zvuku, lokalizace, blízké pole, pole mikrofónů, křížová korelační funkce, GCC, PHAT, IPHAT, triangulace

ABSTRACT

KEYWORDS

acoustic source, localization, near field, microphone array, cross-correlation function, GCC, PHAT, IPHAT, triangulation

ŠURÁNEK, David *Lokalizace akustického zdroje v blízkém poli*: bakalářská práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2011. 48 s. Vedoucí práce byl Ing. Ivan Míča

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Lokalizace akustického zdroje v blízkém poli“ 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.

Brno

.....

(podpis autora)

OBSAH

1 Úvod	7
1.1 Historie a současný stav problematiky	7
2 Lokalizace akustických zdrojů a problematika s ní spojená	9
2.1 Pojmy „blízké“ a „vzdálené pole“	9
2.2 Senzorová pole	10
2.2.1 Prostorová konfigurace senzorového pole	11
2.3 Metody založené na spektrálním odhadu vysokého rozlišení	12
2.4 Metody založené na výpočtu zpoždění mezi senzory	12
2.4.1 Určení zpoždění signálu mezi jednotlivými senzory	13
2.4.2 Určování polohy zdroje	18
3 Implementace metod a praktické ověření	23
3.1 Experimentální ověření první	23
3.1.1 Úvod k měření a jeho popis	23
3.1.2 Vizualizace výsledků	25
3.2 Diskuze výsledků	28
3.3 Experimentální ověření druhé	28
3.3.1 Úvod k měření a jeho popis	28
3.3.2 Vizualizace výsledků	30
3.4 Diskuze výsledků	42
4 Závěr	43
Literatura	45
Seznam symbolů, veličin a zkratk	46
Seznam příloh	47
A Obsah přiloženého média	48

SEZNAM OBRÁZKŮ

2.1	Ilustrace směrů šíření	10
2.2	Vektor korelačních koeficientů $R_{12}(\tau)$	14
2.3	Detail vektoru korelačních koeficientů	15
3.1	Diagram rozložení pole a polohy zdroje při měření.	24
3.2	Vizualizace výsledků pro první měření (1/3).	25
3.3	Vizualizace výsledků pro první měření (2/3).	26
3.4	Vizualizace výsledků pro první měření (3/3).	27
3.5	Diagram rozložení pole a polohy zdroje při měření.	29
3.6	Diagram rozložení pole a polohy zdroje při měření.	29
3.7	Vizualizace výsledků pro druhé měření (1/12).	30
3.8	Vizualizace výsledků pro druhé měření (2/12).	31
3.9	Vizualizace výsledků pro druhé měření (3/12).	32
3.10	Vizualizace výsledků pro druhé měření (4/12).	33
3.11	Vizualizace výsledků pro druhé měření (5/12).	34
3.12	Vizualizace výsledků pro druhé měření (6/12).	35
3.13	Vizualizace výsledků pro druhé měření (7/12).	36
3.14	Vizualizace výsledků pro druhé měření (8/12).	37
3.15	Vizualizace výsledků pro druhé měření (9/12).	38
3.16	Vizualizace výsledků pro druhé měření (10/12).	39
3.17	Vizualizace výsledků pro druhé měření (11/12).	40
3.18	Vizualizace výsledků pro druhé měření (12/12).	41

1 ÚVOD

Lokalizace akustického zdroje je oblast zpracování signálů, která si klade za cíl poskytnutí informace o poloze zdroje v požadované úrovni kvality. Pro některé případy je tato úroveň dosažena již tím, že určíme „pouze“ směr příchodu signálu, tzv. DOA (direction of arrival – směr příchodu), jindy je naopak potřeba znát přímo polohu akustického zdroje. K těmto účelům, se používá vícero metod, z nichž každá má své výhody, či nevýhody, které jsou zohledňovány při výběru metody pro řešení zadaného problému.

Tato práce se zabývá problematikou lokalizace akustických zdrojů v blízkém poli produkujících širokopásmový signál.

V úvodu práce je velmi stručně zmíněna historie, vývoj a současný stav problematiky.

Poté následuje kapitola 2, ve které jsou popsány základní poznatky pro lokalizaci akustických zdrojů. Zmíněná kapitola obsahuje vysvětlení pojmů *blízké* a *vzdálené pole*. Vzápětí bude nastíněna problematika senzorových polí, a jejich konfigurací. Kapitola dále pokračuje popisem některých známých metod lokalizace, z nichž některé byly realizovány. Popis realizovaných metod je komplexnější. Nerealizované metody jsou spíše zmíněny a obsahují výčet vlastností.

Další kapitola 3 se věnuje implementaci zvolených metod, praktickému ověření funkčnosti. Obsahuje také vizualizace výsledků, které lépe demonstrují úspěšnost lokalizace. Zvolené metody jsou zde také porovnávány v různých ohledech.

V závěru jsou shrnuty výsledky práce a je diskutována možnost nasazení zkoušených metod v praxi.

1.1 Historie a současný stav problematiky

Metody lokalizace akustických zdrojů pomocí pole snímačů jsou známy mnoho desítek let. Jednou z nejstarších používaných metod byla metoda triangulace, která se používala již v první světové válce pro lokalizaci dělostřeleckých baterií. Obecně však platí, že metody pro lokalizaci lze aplikovat i na jiné zdroje, které produkují signály vlnového charakteru. Jednou z těchto technik je možnost využívání pole antén – radarů pro určení polohy objektu, která se také objevila již za první světové války, svého největšího rozmachu se však dočkala až za druhé světové války a její uplatnění trvá až do dnešní doby. Pole antén pro pozorování je v dnešní době používáno též v astronomii.

Systémy lokalizace pomocí polí snímačů jsou však používány také v mnoha oblastech moderní techniky, které je aplikují např. na telekonference, řečové rozpoznávání,

identifikaci mluvčího, záznam zvuků v zahlučených prostorách, nebo v prostorách s reverbancemi. Jinou možností aplikace je např. určení aktuální polohy objektu v prostoru, či určení jeho rychlosti.

K problematice lokalizace má velmi blízko problém separace signálů [2]. Při separaci signálů se též používá mikrofonních polí. Při separaci však není potřeba používat mnoho mikrofónů. Velká část algoritmů pro separaci si vystačí i se dvěma mikrofony [1]. Se zvyšujícím se počtem mikrofónů se však kvalita separace zvyšuje [2]. Pro příklad uveďme *noise cutting*, který ve své základní formě využívá dva mikrofony. První mikrofón snímá žádoucí signál, druhý mikrofón snímá hluky na pozadí. Signál z druhého mikrofónu je pak s opačnou fází přičten k signálu z prvního mikrofónu, čímž dojde k jeho alespoň částečné eliminaci. Tohoto nebo podobného algoritmu je v současné době hojně používáno například v mobilních telefonech, v „hands-free“ sadách, či dokonce i u sluchátek.

Ačkoliv jsou některé systémy polí navrženy tak, aby zaměřili ohnisko svého snímání na zdroj bez toho, aby jakkoliv měnili svou polohu, či orientaci, většina systémů sleduje polohu vybraného zdroje, či zdrojů, a fyzicky zaměřuje své snímací pole adekvátním směrem. Tyto automaticky se zaměřující pole mohou nahradit stávající techniky snímání, které zahrnují ruční ovládání zařízení, jako např. závěsné mikrofony či směrové mikrofony. Co je však ještě více zajímavé, systémy používající pole snímačů mají dostatek potenciálu k tomu, aby v některých situacích v budoucnu nahradili ruční, nebo náhlavní mikrofony. Tento potenciál je jim přisuzován zejména proto, že jak teoretická, tak praktická stránka polí snímačů je podrobována aktivnímu výzkumu[1].

2 LOKALIZACE AKUSTICKÝCH ZDROJŮ A PROBLEMATIKA S NÍ SPOJENÁ

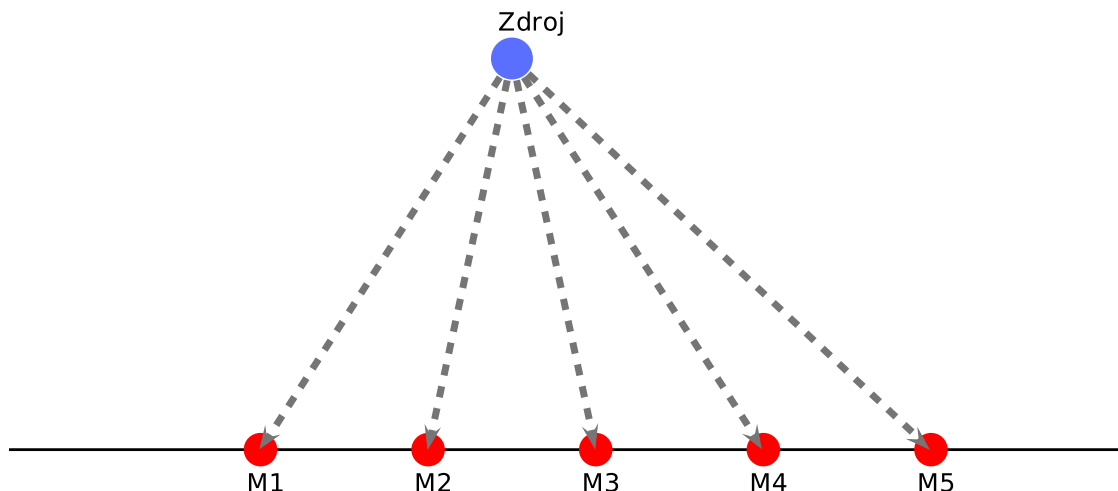
Lokalizace akustického zdroje ve své podstatě znamená odhad jeho souřadnic v kartézské soustavě. Dalšími možnými zjišťovanými parametry mohou být například směr, či rychlost pohybu, v případě že zdroj není nehybný[4]. Je také možno použít sférické souřadnice, v tomto případě budeme hledat velikosti rovinného a výškového úhlu (jinak řečeno azimutu a elevace), a také vzdálenost. Jak dále uvidíme, použití sférických souřadnic nám v některých případech může značně zjednodušit odvození pro výpočty výše uvedených parametrů, které jsou často souhrnně označovány jako DOA (direction of arrival – směr příchodu) [1].

Cílem této kapitoly je přiblížení způsobů, kterými různé metody určují polohu zdroje, tedy ho lokalizují v prostoru. Dále je uvažována jejich přesnost a možnost použití v blízkém poli. O rozdílech mezi tzv. blízkým a vzdáleným polem pojednává následující kapitola.

2.1 Pojmy „blízké“ a „vzdálené pole“

Uvažujeme-li obecné pole, složené z N mikrofونů, kde každý mikrofون má svou jedinečnou polohu, a dále zdroj, který má též jedinečné umístění, můžeme jednotlivé mikrofony se zdrojem proložit obecně N přímkami. Tyto přímky nám dávají tzv. přímou cestu, podél které se zvuk současně šíří od zdroje k mikrofونům. Udávají nám tedy tzv. směr šíření od zdroje k jednotlivým mikrofونům, viz obrázek 2.1. Pokud o výše zmíněných přímkách můžeme říci, že jsou téměř souběžné, v případě zjednodušení na 2-D (dvourozměrný prostor) rovnoběžné, mluvíme o situaci, kdy se akustický zdroj nachází ve vzdáleném poli sensorového pole (z angl. *far-field of the array*). Zmíněná situace nastává, pokud vzdálenost mezi zdrojem a sensorovým polem je mnohem větší, než je velikost, respektive šířka sensorového pole. Pro lepší představu uvažujme lineární konfiguraci sensorového pole. Sensory ležící na přímce nám svým prvním a posledním prvkem ohraničují úsečku. Pokud se bude zdroj nacházet ve vzdálenosti, která je mnohokrát větší, než je délka zmiňované úsečky, říkáme, že se zdroj nachází ve vzdáleném poli. Protože v této situaci se vůči velikosti sensorového pole jeví zakřivení kulové vlnoplochy šířícího se zvuku velmi malé, můžeme si tuto situaci zjednodušit tak, že šířící se vlnoplochu budeme uvažovat jako rovinnou.

Naopak, pokud bude vzdálenost zdroje od sensorového pole srovnatelná s velikostí tohoto pole, přímky budou tedy různoběžné, mluvíme o situaci, kdy se zdroj



Obr. 2.1: Ilustrace směrů šíření

nachází v blízkém poli sensorového pole (z angl. *near-field of the array*). Obrázek 2.1 tedy znázorňuje případ blízkého pole.

Zjednodušení v podobě vzdáleného pole s sebou však nese jistá omezení. Jelikož zakřivení vlnoplochy nám dává informaci o vzdálenosti, či také velikosti akustického zdroje, její rovinnou aproximací se o tyto informace ochuzujeme. Na druhou stranu se tímto zjednodušením velmi snižuje výpočetní náročnost, která bývá v některých situacích důležitější, než výsledná přesnost algoritmu [1]. Jako možný příklad můžeme uvést aplikace fungující v reálném čase.

2.2 Sensorová pole

Senzorovými poli je možno provádět zkoumání vlastností všech záření vlnového charakteru, proto jsou metody zkoumání vlnění známé i z jiných vědních oblastí, než je akustika. Jako příklad může být uvedena seizmologie, telekomunikace, či zdravotní diagnostika [3]. Signál, který chceme analyzovat, je zaznamenáván současně několika senzory a průběhy na jednotlivých senzorech jsou ukládány v patřičném formátu. V dnešní době probíhá uložení obvykle tak, že průběh signálů je zdigitalizován A/D převodníky a posloupnosti dat jsou uloženy do souboru obvykle na pevný disk počítače. Záznam průběhů z jednotlivých senzorů musí být paralelní.

V dalších kapitolách se budeme věnovat zejména metodám lokalizace akustického zdroje. Jak již však bylo několikrát zmíněno, tyto metody jsou stejně tak dobře aplikovatelné na jakýkoliv druh vlnění s přihlédnutím k tomu, že je nutno použít adekvátní snímače pro dané vlnění.

2.2.1 Prostorová konfigurace senzorového pole

Způsobů umístění senzorového pole v prostoru je v podstatě neomezený počet, s ohledem ke složitosti výpočtu se však používají obvykle rozložení vycházející ze základních geometrických objektů. Hlavním parametrem při volbě prostorové konfigurace je to, zda chceme lokalizovat zdroj v rovině (2-D), nebo v prostoru (3-D). Důležitou podmínkou, kterou je potřeba dodržet při rozmísťování senzorů, je vztah mezi vzdáleností senzorů a frekvencí snímaného vlnění

$$f_{\max} = \frac{v}{2d}, \quad (2.1)$$

kde $f_{\max}$ je maximální kmitočet zpracovatelný polem, v je vlnová rychlost zvuku a d je vzdálenost mezi senzory.

Pokud je pro nás výchozí 2-D (dvourozměrný prostor), můžeme si zvolit buď konfiguraci lineární, kde jsou senzory umístěny na přímce (v řadě), nebo rovinou, kde jsou senzory umístěny obvykle ve tvaru základních rovinných útvarů, jako např. čtverec nebo obdélník [2]. Tyto dvě zmíněné možnosti byly také dlouhou dobu nejčastěji používané [6]. V současné době se však obvykle využívá pole kruhového, které má jisté výhody, které vycházejí z matematických závislostí [1]. Tento matematický problém je však již mimo rozsah této práce a proto se jím nebudeme dále zabývat.

Pokud však hodláme lokalizovat zdroj v 3-D prostoru, musíme pole rozšířit o další dimenzi. Tato rozšířená pole bývají obvykle odvozena od rovinných polí do základních objektů, jako je krychle či kvádr či koule. Další zkoumanou možností je náhodné rozmístění pole, které může být pro některé aplikace vhodnější [1].

Samozřejmě lze použít i další možná rozložení. Výjimkou nejsou ani různé kombinace uvedených konfigurací. Rozmístění a tvar pole má vliv na přesnost lokalizace. Pro různé účely a situace jsou vhodnější různé konfigurace. Jelikož řešení různých úloh si vyžaduje různý přístup, nemůžeme žádné rozmístění označit jako obecně nejlepší pro všechny možné situace lokalizace.

2.3 Metody založené na spektrálním odhadu vysokého rozlišení

Jednou z možností lokalizace použitelných v blízkém poli může být využití některé z metod založených na spektrálním odhadu vysokého rozlišení [2]. Jako příklad uveďme třeba metodu MV (Minimum Variance – minimální odchylka), ale hlavně nejpopulárnější metodu MUSIC (Multiple Signal Classification – vícenásobná klasifikace signálu). Obě tyto metody dokáží lokalizovat zdroje v blízkém poli, navíc dokáží lokalizovat i více zdrojů současně a to i v případě, že se tyto zdroje nachází velmi blízko sebe. Tyto metody se však obvykle nepoužívají na širokopásmové signály (tedy ani na řeč). Většina jejich aplikací je vyvinuta s určením pro úzkopásmové signály, proto jsou uplatňovány například v radarové lokalizaci.

Metody založené na spektrálních odhadech vysokého rozlišení při lokalizaci více zdrojů dosahují dokonce lepších výsledků, než metody založené na tvarování přijímací charakteristiky, které jsou však určeny pro vzdálené pole. Proto se jimi tato práce nebude příliš zabývat.

V úzkopásmové oblasti nejsou metody spektrálního odhadu výpočetně příliš náročné. Toto se však mění, pokud jsou tyto metody aplikovány na širokopásmový signál. V tomto případě se jejich výpočetní nároky razantně zvyšují.

2.4 Metody založené na výpočtu zpoždění mezi senzory

Metody založené na určení časových zpoždění byly k lokalizaci zdrojů vlnění používány v minulosti hlavně kvůli své výpočetní nenáročnosti. Jejich obliba se však projevuje i v dnešní době, přestože máme k dispozici procesory, které jsou schopny zpracovávat i náročnější výpočty v relativně krátkém čase [1].

Postup lokalizace těmito metodami se dá rozdělit na dvě hlavní části:

- Určení zpoždění signálu mezi jednotlivými senzory.
- Provedení výpočetního algoritmu pro akustickou lokalizaci.

V obou těchto částech se používá vícero postupů, které je možné kombinovat a získat tak varianty vhodné pro různé použití. Nyní si představíme často používané metody z obou výše zmíněných částí.

2.4.1 Určení zpoždění signálu mezi jednotlivými senzory

Z pohledu určování zpoždění lze zaznamenávané signály rozdělit do dvou kategorií [6]. Jednak to mohou být tzv. signály typu *burst* (anglicky „výstřel“), které mají obvykle tvar sinusovky s klesající amplitudou, nebo signály se spojitou emisí. Pro určení zpoždění pro *burst* signály lze použít principu, při kterém se za dobu příchodu považuje čas, ve kterém dojde v záznamu k překročení jisté prahové úrovně (anglicky *threshold*). Toto je však metoda spíše historická a pro naše potřeby nepoužitelná, neboť nedosahuje potřebné úrovně univerzálnosti. My pro tuto úlohu použijeme modernější metody, které jsou použitelné i pro spojitě širokopásmové signály. Tyto metody jsou často odvozeny od korelační funkce. Korelační funkce a další často používané metody jsou popsány v následujících odstavcích.

Korelační funkce

Zjištění posunutí pomocí korelační funkce je velmi oblíbená metoda, je velmi rychlá a je možno ji použít i na poměrně krátké záznamy [3]. Možné nevýhody této metody ovšem spočívají v tom, že není příliš odolná proti aditivnímu šumu a zejména také to, že zpoždění mezi signály určí pouze s přesností na jeden vzorek [3].

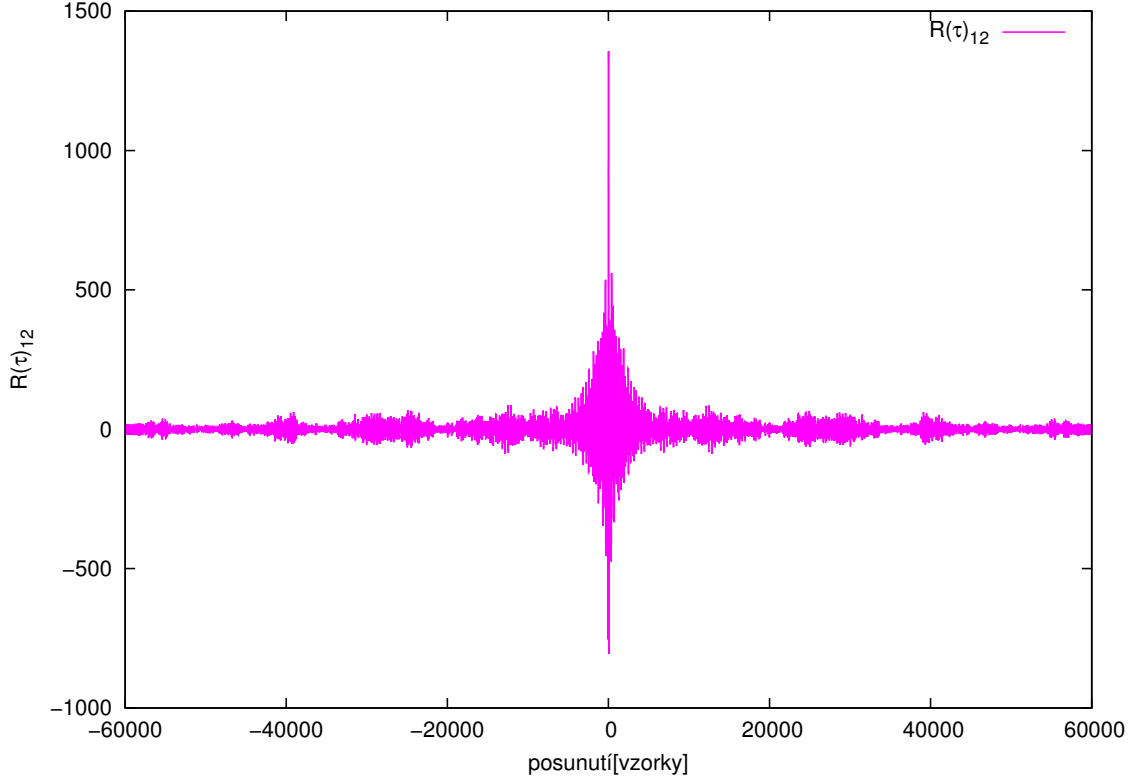
Korelační funkce v podstatě funguje tak, že porovnává podobnost dvou signálů pro všechna možná vzájemná posunutí podle vzorce (2.2). Provedeme-li tento výpočet pro všechna možná posunutí, získáme kompletní vektor korelačních koeficientů $R(\tau)$, ve kterém právě jednomu τ odpovídá největší hodnota korelačního koeficientu a právě pro tuto velikost posunutí o τ vzorků jsou si signály nejpodobnější.

Při výpočtu korelační funkce je vhodné vstupní signály co nejvíce sjednotit (normovat), například proto, že použité snímače nemusí mít stejné zesílení. Jako nejvhodnější se zde jeví metoda $signal = \frac{signal}{\max\{|signal|\}}$. Výsledkem potom bude signál s maximální absolutní hodnotou rovnou jedné. Nyní lze spočítat korelační funkci pomocí vzorce

$$R_{12}(\tau) = Q \cdot \sum_{k=-\infty}^{\infty} s_1(k + \tau) \cdot s_2^*(k), \quad \tau \in \mathbf{Z}, \quad (2.2)$$

kde $Q = 1$ pro standardní odhad korelační funkce, $Q = \frac{1}{N}$ pro vychýlený odhad a $Q = \frac{1}{N-\tau}$ pro nestranný odhad [3]. Vektor $R_{12}(\tau)$ vypočtený pro mikrofony M_1 a M_2 z praktické realizace je na obrázku 2.2 a jeho detail v místě globálního maxima na obrázku 2.3. Z obrázků je vidět, že vzájemná korelační funkce je téměř symetrická. To je dáno tím, že výpočet probíhá pro dva v podstatě stejné signály. Pokud zanedbáme šum, je jeden signál pouze zpožděn v čase. V případě, že bychom tento výpočet prováděli pro dva naprosto stejné signály, počítali bychom autokorelační funkci, která je symetrická podle osy y .

Po vypočtení koeficientů je nutno nalézt polohu největšího koeficientu a jeho index v poli. V korelační funkci se obvykle nachází několik lokálních maxim, proto je nutno nalézt maximum globální. Rozdíl mezi indexem nejvyššího koeficientu a středem celého vektoru nám udává posunutí ve vzorcích, pomocí něhož můžeme přes vzorkovací kmitočet f_{vz} získat časové zpoždění mezi signály.



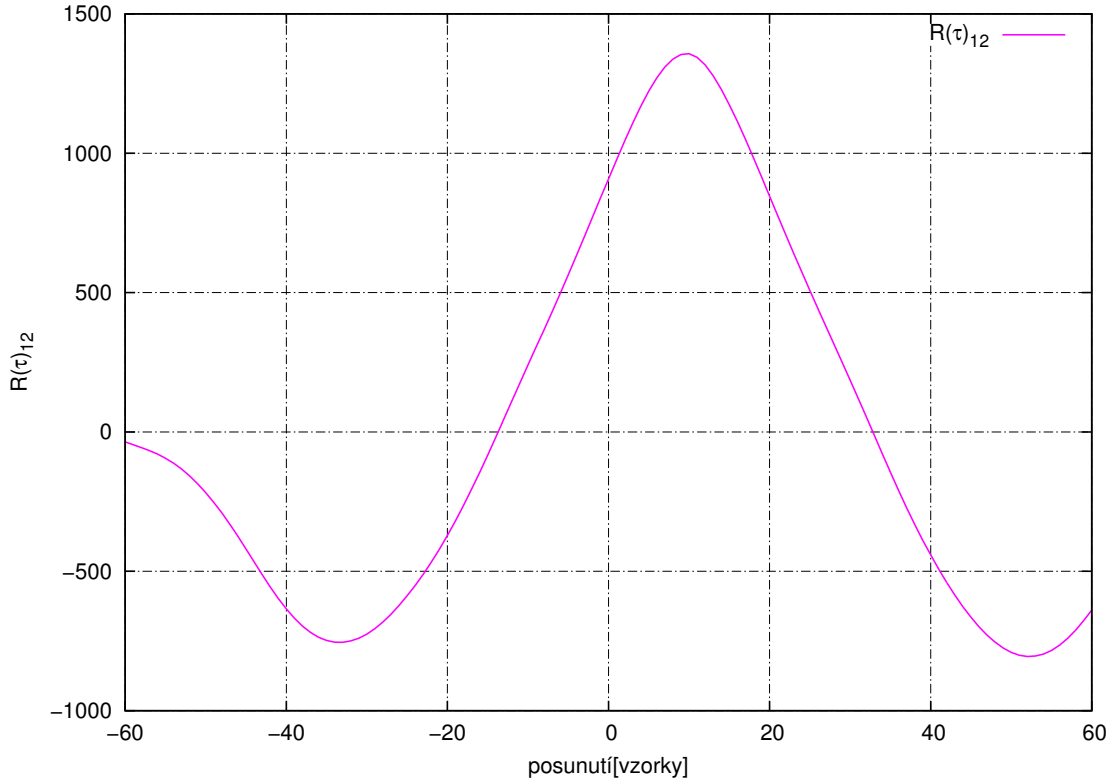
Obr. 2.2: Vektor korelačních koeficientů $R_{12}(\tau)$ mezi prvním a druhým mikrofonem s polohou zdroje [0 m; 0,6 m] pro řečový signál

GCC

GCC (Generalized Cross Correlation – zobecněná křížová korelace) je úspěšně používaná metoda pro určení časového rozdílu příchodu (TDOA) zvukových vln mezi dvojicí mikrofonů. Poskytuje navíc lepší vlastnosti v prostorách s reverbancemi či s hlukem na pozadí. Tyto vlastnosti lze navíc ještě zlepšit použitím některé z váhovacích funkcí.

Křížová korelace dvou signálů $x_1(t)$ a $x_2(t)$ je definována jako [1]:

$$c_{12}(\tau) \equiv \int_{-\infty}^{\infty} x_1(t)x_2(t + \tau)dt. \quad (2.3)$$



Obr. 2.3: Detail vektoru korelačních koeficientů z obrázku 2.2 v místě globálního maxima.

Fourierova transformace křížové korelační funkce je známá jako křížová spektrální hustota nebo také jako křížové spektrum [1] a je definována vztahem

$$C_{12}(\omega) = \int_{-\infty}^{\infty} c_{12}(\tau) e^{-j\omega\tau} d\tau. \quad (2.4)$$

Dosažením rovnice (2.3) do rovnice (2.4) a aplikováním konvolučních vlastností Fourierovy transformace [1] může být křížové spektrum vyjádřeno jako

$$C_{12}(\omega) = X_1(\omega) X_2'(\omega), \quad (2.5)$$

kde $X_1(\omega)$ je Fourierova transformace $x_1(t)$ a $X_2'(\omega)$ je komplexně sdružená Fourierova transformace $x_2(t)$. Inverzní Fourierovou transformací výrazu (2.5) dostaneme křížovou korelační funkci vyjádřenou pomocí $X_1(\omega)$ a $X_2'(\omega)$:

$$c_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_1(\omega) X_2'(\omega) e^{j\omega\tau} d\omega. \quad (2.6)$$

Funkce GCC, $R_{12}(\tau)$, je křížovou korelací dvou filtrovaných verzí $x_1(t)$ a $x_2(t)$. [1] S Pomocí Fourierovy transformace těchto filtrů označených jako $G_1(\omega)$ a $G_2(\omega)$ můžeme GCC zapsat jako

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (G_1(\omega) X_1(\omega)) (G_2(\omega) X_2(\omega))' e^{j\omega\tau} d\omega. \quad (2.7)$$

Matematickou úpravou rovnice (2.7) se lze dostat k výrazu

$$R_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_1(\omega) G_2'(\omega) X_1(\omega) X_2'(\omega) e^{j\omega\tau} d\omega. \quad (2.8)$$

Definováním frekvenčně závislé váhovací funkce $\Psi_{12}(\omega) \equiv G_1(\omega) G_2'(\omega)$ můžeme funkci GCC přepsat jako

$$R_{12}(\tau) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_{12}(\omega) X_1(\omega) X_2'(\omega) e^{j\omega\tau} d\omega. \quad (2.9)$$

Se správně zvolenou váhovací funkcí $\Psi_{12}(\omega)$ by se ve vektoru $R_{12}(\tau)$ mělo objevit globální maximum, které bude korespondovat s TDOA mezi mikrofony 1 a 2.

Některé základní a často používané váhovací funkce budou přiblíženy v dalších odstavcích.

Váhovací funkce fázové transformace PHAT

Velmi často používanou váhovací funkcí pro GCC je tzv. PHAT (Phase Transform – fázová transformace). Tato se sice v ideálních podmínkách jeví jako nepotřebná, avšak s jejím použitím dává GCC v reálných prostředích mnohem přesnější výsledky [1]. Z důvodu její zvýšené odolnosti proti ozvěnám a odrazům se stala velmi populární formou metody GCC označovanou jako GCC-PHAT. Tato váhovací funkce je definována vztahem

$$\Psi_{12}(\omega) \equiv \frac{1}{|X_1(\omega) X_2'(\omega)|}. \quad (2.10)$$

Váhovací funkce vylepšené fázové transformace IPHAT

IPHAT (Improved Phase Transform – vylepšená fázová transformace) je – jak sám název napovídá – vylepšenou verzí PHAT [7]. Ponechává si stejnou odolnost proti reverbancím jako PHAT a přidává ještě zvýšenou odolnost proti šumům a podává tak lepší výsledky v situacích kdy je menší odstup signálu od šumu [7]. IPHAT byla definována vztahem

$$\Psi_{12}(\omega) \equiv \frac{1}{|X_1(\omega) X_2'(\omega)|^\lambda}. \quad (2.11)$$

Kde hodnota činitele λ je odvozována z hodnoty SNR (Signal to Noise Ratio – odstup signálu od šumu) [7]. Experimentální výraz pro λ je dán jako

$$\lambda = \begin{cases} \lambda_0 & \sigma < \sigma_0, \\ \frac{\lambda_1 - \lambda_0}{\sigma_1 - \sigma_0} (\sigma - \sigma_0) + \lambda_0 & \sigma_0 < \sigma < \sigma_1, \\ \lambda_1 & \sigma > \sigma_1, \end{cases} \quad (2.12)$$

kde σ reprezentuje SNR.

Použitím této váhovací funkce samozřejmě vznikne více lokálních maxim ve vektoru $R_{12}(\tau)$ avšak globální maximum by se mělo nacházet vždy na pozici správného zpoždění i v případě velmi nízkého SNR, jak je dokázáno simulacemi v [7].

Váhovací funkce pásmové propusti

Jedná se o nejjednodušší možnou váhovací funkci. Tato funkce jednoduše potlačuje frekvence mimo zkoumané pásmo [1]. Nejlépe bude ukázat takovou funkci na jednoduchém příkladu. Chceme-li například lokalizovat řeč, jejíž pásmo je obvykle mezi 300 Hz a 6 kHz, váhovací funkce může mít tvar například

$$\Psi_{12}(\omega) = \begin{cases} 1 & 2\pi \cdot 300 \leq \omega \leq 2\pi \cdot 6000, \\ 0 & \text{jinak.} \end{cases} \quad (2.13)$$

Odstranění některých kmitočtů může být velmi přínosné. Například odstraněním dolních kmitočtů při lokalizaci řečníka můžeme eliminovat hluky, které mají většinu energie v tomto dolním pásmu. Typickým zástupcem takovýchto hluků na pozadí může být například hluk generovaný jednotkou klimatizace, či chlazením počítačů. Zdroje, které generují hluk o velkých vlnových délkách se navíc hůře lokalizují, což je způsobeno tím, že se špatně určuje směr šíření takových vlnění [1]. Váhovací funkce pásmové propusti se také často používá v kombinaci s dalšími váhovacími funkcemi, zejména PHAT, či IPHAT, kterým signály ořízne na požadované pásmo a tyto pak mohou podat ještě lepší výsledky.

Určování zpoždění s přesností na necelé vzorky

Pro co nejpřesnější určení polohy zdroje je nutné určit časové zpoždění mezi signály co nejpřesněji. Výše zmíněné metody však poskytují informaci o zpoždění pouze s přesností na celé vzorky. Z toho důvodu bylo nutné nalézt metodu, která dokáže zpoždění určit i na zlomky vzorků. Pro tento účel byl zvolen Lagrangeův interpolátor popsáný v [8].

Jak je známo, vzorkovaný signál lze rekonstruovat za předpokladu že je splněn vzorkovací teorém

$$f_{\text{vz}} > 2f_{\text{bw}}. \quad (2.14)$$

kde f_{vz} je vzorkovací kmitočet f_{bw} je šířka pásma vzorkovaného signálu.

Podle [7] pak můžeme psát

$$R_{12}(t) = \sum_{n=n_0-M}^{n_0+M} R_{12}(\tau) \text{sinc}(t - \tau), \quad (2.15)$$

kde $R_{12}(t)$ je vektor korelačních koeficientů v čase, $R_{12}(\tau)$ je vektor korelačních koeficientů ve vzorcích, n_0 je pozice, na které se nachází maximum v $R_{12}(\tau)$. $2M + 1$ vzorků bude použito pro restaurování $R_{12}(t)$. Interval t , na který se budeme soustředit je $t \in [n_0 - 1, n_0 + 1]$. Nalezením maximální hodnoty v $R_{12}(t)$ lze získat informaci o zpoždění mezi signály s přesností větší než jeden vzorek.

2.4.2 Určování polohy zdroje

V této části existují obecně dva přístupy. První možnost je ta, kdy chceme zjišťovat pouze DOA (direction of arrival – směr příchodu), což je základním cílem metody TDOA (Time Difference of Arrival – časový rozdíl příchodu) [2], která však uvažuje pouze vzdálené pole a proto se jí budeme věnovat jen okrajově. Další možnost je přímo určovat bod v prostoru, ve kterém se zdroj nachází. Tuto možnost nám bude reprezentovat metoda triangulace, která je pro tuto práci styčným bodem, neboť je v praktické části implementována a ověřena jednoduchým měřením.

Metoda TDOA

Tato metoda pro své správné fungování potřebuje splnění několika předpokladů pro výchozí podmínky [2]:

- Vzdálenost mikrofonního pole od zdroje musí být mnohem větší než je vzájemná vzdálenost mikrofonů v poli.
- Signály jsou centrovány a jejich A/D převod probíhá současně.

První předpoklad slouží pro zjednodušení. Při tomto rozložení pole a zdroje se zabýváme situací tzv. vzdáleného pole. Ve vzdáleném poli totiž můžeme uvažovat kulovou vlnoplochu, šířící se od zdroje, jako rovinou, a tak si výpočty podstatně zjednodušíme. Ochuzujeme se tím však o možnost určení vzdálenosti zdroje.

Pro pokrytí 2-D prostoru pak dále musí platit tyto podmínky:

- Je potřeba alespoň tři mikrofony.
- Alespoň jeden mikrofon musí ležet mimo přímku tvořenou zbývajících mikrofony.
- Všechny senzory musí ležet v jedné rovině se zdrojem.

Pro pokrytí 3-D prostoru musí platit tyto podmínky:

- Je potřeba alespoň čtyři mikrofony.
- Alespoň jeden senzor musí ležet mimo rovinu tvořenou zbývajících mikrofony.

Uvažujme nyní pole x , kde polohu každého mikrofonu reprezentuje vektor $m_1, m_2, \dots, m_N$. Z těchto vektorů pak vytvoříme matici

$$\mathbf{M} = \begin{pmatrix} m_{1x} & m_{2x} & \dots & m_{Nx} \\ m_{1y} & m_{2y} & \dots & m_{Ny} \\ m_{1z} & m_{2z} & \dots & m_{Nz} \end{pmatrix}, \quad (2.16)$$

kde každý sloupec představuje polohu jednoho mikrofonu.

Nyní vytvoříme matici $\mathbf{Q}$, o rozměrech $3 \times Q$, kde $Q = \frac{M!}{2!(M-2)!}$ a M je počet mikrofonů. Každý sloupec matice bude představovat vektor dvojice senzorů a bude

tvaru $m_{ij} = m_i - m_j$ pro $i, j = 1, 2, 3, \dots, M$, za podmínky, že $i > j$.

$$\mathbf{Q} = \begin{pmatrix} m_{2x} - m_{1x} & m_{3x} - m_{1x} & \dots & m_{Mx} - m_{(M-1)x} \\ m_{2y} - m_{1y} & m_{3y} - m_{1y} & \dots & m_{My} - m_{(M-1)y} \\ m_{2z} - m_{1z} & m_{3z} - m_{1z} & \dots & m_{Mz} - m_{(M-1)z} \end{pmatrix}. \quad (2.17)$$

Pomocí souřadnic z těchto matic a vypočteného zpoždění τ_{ij} mezi signály zaznamenanými mikrofony M_i a M_j , nyní můžeme spočítat vzdálenosti d_{ij} , kde i, j jsou koeficienty označující mikrofony, které musí vlnoplocha šířícího se signálu urazit, aby dorazila od mikrofonu M_i k mikrofonu M_j .

$$d_{ij} = v\tau_{ij}, \quad (2.18)$$

kde v je vlnová rychlost zvuku.

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

$$d = [\cos \phi \sin \theta, \cos \phi \cos \theta, \sin \phi]. \quad (2.19)$$

Vhodnou kombinací výše uvedených výrazů lze dospět ke vzorcům, které vyjadřují velikosti úhlů vyjadřující DOA:

$$\theta = \frac{d_{21}}{m_{21}}. \quad (2.20)$$

Obdobně potom pro určení ϕ .

Tato metoda sebou nese však jisté komplikace [5]. Pokud počítáme úhel právě ze dvou senzorů, nelze jednoznačně říci, zda tento úhel je právě směr příchodu signálu. 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. V případě, že nenajdeme dva naprosto shodné úhly, zvolíme dvojici úhlů, která se nachází ve stejném kvadrantu.

Metoda triangulace

Metoda triangulace je používána již od první světové války. Za své historie byla několikrát modifikována, avšak její princip zůstává stejný [6]. V současné době je používanější výše zmíněná metoda TDOA, zejména kvůli tomu, že určení DOA je v mnoha aplikacích dostačující.

V jednorozměrném případě, kdy dva mikrofony tvoří přímku a zdroj se tedy nachází na úsečce mezi nimi, je určení jeho polohy jednoznačně určeno rozdílem naměřených časů příchodů k jednotlivým mikrofonom.

V případě 2-D, tedy, když se zdroj nachází v rovině, může být rozdíl vzdáleností, které musí vlnoplocha urazit, aby dospěla k jednotlivým mikrofonom vypočten z časového rozdílu příchodů

$$\delta = \Delta t \cdot v, \quad (2.21)$$

kde δ je rozdíl drah, Δt je časový rozdíl příchodů a v je vlnová rychlost. Stejná hodnota rozdílu drah δ je získán, pokud zdroj leží na hyperbole, jejíž ohniska leží v souřadnicích mikrofonom. Použitím dalších snímačů můžeme polohu zdroje určit z průsečíku hyperbol. Obecně lze říct, že pokud chceme v N rozměrném prostoru určit polohu zdroje metodou triangulace, potřebujeme alespoň $N + 1$ mikrofonom. Pro 2-D prostor tedy potřebujeme alespoň tři mikrofony a pro 3-D alespoň čtyři mikrofony. Při použití minimálního počtu snímačů se můžou objevit dvojznačná řešení, kdy obě dvě varianty mají fyzikální smysl. Toto lze vyřešit přidáním dalšího senzoru, který nám poskytne další informaci, díky níž je možno určit, která z variant je správná. Kromě přímého matematického řešení, které je popsáno níže se někdy používá iterativních metod, či tabulkového řešení.

Malou modifikací při použití korelační funkce bylo omezení výpočtu pouze na oblast vzorků, ve které je možný výskyt maxima korelační funkce [3]. Toto bylo provedeno zejména kvůli rychlosti výpočtu. Velikost oblasti M , používaná pro výpočet korelační funkce, je dána maximální vzdáleností dvou mikrofonom v poli jako

$$M = \text{fix} \left(\frac{l \cdot f_{vz}}{v} \right), \quad (2.22)$$

kde l je maximální vzdálenost mikrofonom, f_{vz} je vzorkovací kmitočet, v je rychlost zvuku a $\text{fix}()$ je funkce vracející celou hodnotu čísla. Výpočet dle vzorce (2.2) pak probíhá od $-M$ do $+M$. V použitém algoritmu je navíc hodnota M zvětšena o jednotky vzorků, z důvodu možných nepřesností při instalaci pole mikrofonom. Jinak by se mohlo stát, že maximum se bude nalézat mimo zkoumaný interval.

Při řešení problému lokalizace akustického zdroje je v tomto algoritmu použita analytická metoda nalezení průsečíku resp. průsečíků dvou hyperbol, který je převzat z [6]. Hyperboly jsou určeny zpožděními mezi dvěma snímači. Souřadnice snímačů označíme $[x_1, y_1]$, $[x_2, y_2]$ a $[x_3, y_3]$. Při řešení vycházíme z parametrické rovnice hyperboly

$$r(\varphi) = \frac{p}{1 - \varepsilon \cos \varphi}, \quad (2.23)$$

kde $p = \frac{b^2}{a}$, $\varepsilon = \frac{e}{a}$ ($e > a$, $a \neq 0$) a $e^2 = a + b$. Z toho tedy

$$r(\varphi) = \frac{e^2 - a^2}{a - e \cos \varphi}. \quad (2.24)$$

Tato rovnice ovšem platí pouze pokud je druhé ohnisko v počátku souřadnic a obě ohniska se nachází na ose x . Pro naši potřebu je nutno použít rovnice obecnějšího tvaru, zde přímo pro jednotlivé složky souřadnic:

$$\begin{aligned}x &= x_f + \frac{e^2 - a^2}{a - e \cos(\varphi - \alpha)} \cos \varphi, \\y &= y_f + \frac{e^2 - a^2}{a - e \cos(\varphi - \alpha)} \sin \varphi,\end{aligned}\tag{2.25}$$

kde $[x_f, y_f]$ jsou souřadnice hlavního ohniska.

Nyní již máme rovnice, ze kterých můžeme vycházet při lokalizaci zdroje. Je však potřeba si uvědomit, že výsledné řešení spočívá v nalezení průsečíku 2 hyperbol, z nichž každá je určena rovnicemi (2.25).

Rovnice 1. hyperboly	Rovnice 2. hyperboly	
$x = x_1 + \frac{e_1^2 - a_1^2}{a_1 - e_1 \cos(\varphi - \alpha_1)} \cos \varphi$	$x = x_2 + \frac{e_2^2 - a_2^2}{a_2 - e_2 \cos(\varphi - \alpha_2)} \cos \varphi$	(2.26)
$y = y_1 + \frac{e_1^2 - a_1^2}{a_1 - e_1 \cos(\varphi - \alpha_1)} \sin \varphi$	$y = y_2 + \frac{e_2^2 - a_2^2}{a_2 - e_2 \cos(\varphi - \alpha_2)} \sin \varphi$	

kde body $[x_1, y_1]$ a $[x_2, y_2]$ jsou souřadnice ohnisek. Práci si podstatně zjednodušíme, pokud budeme při výpočtu uvažovat hyperboly se stejným ohniskem. V našem případě je toto zjednodušení možné realizovat tak, že jako ohnisko hyperbol zvolíme mikrofon, který je zdroji nejbližší, tzn. jeho relativní zpoždění je nejmenší. Budeme-li tedy uvažovat, že $x_1 = x_2$ z rovnic (2.26) pro souřadnice x , dostaneme

$$\frac{e_1^2 - a_1^2}{a_1 - e_1 \cos(\varphi - \alpha_1)} - \frac{e_2^2 - a_2^2}{a_2 - e_2 \cos(\varphi - \alpha_2)} = 0, \tag{2.27}$$

kterou lze vhodnými matematickými úpravami převést na tvar odpovídající

$$A \cos \varphi - B \sin \varphi = C, \tag{2.28}$$

kde

$$A = \frac{e_2}{e_2^2 - a_2^2} \cos \alpha_2 - \frac{e_1}{e_1^2 - a_1^2} \cos \alpha_1, \tag{2.29}$$

$$B = \frac{e_2}{e_2^2 - a_2^2} \sin \alpha_2 - \frac{e_1}{e_1^2 - a_1^2} \sin \alpha_1, \tag{2.30}$$

$$C = \frac{a_2}{e_2^2 - a_2^2} - \frac{a_1}{e_1^2 - a_1^2}. \tag{2.31}$$

Řešením rovnice (2.28) dostáváme dva kořeny φ_1 a φ_2 , pro které platí

$$\begin{aligned}\varphi_1 &= g + h, \\ \varphi_2 &= g - h,\end{aligned}\tag{2.32}$$

kde

$$g = \arctan\left(\frac{B}{A}\right), \quad (2.33)$$

$$h = \arccos\left(\frac{C}{A} \cos g\right). \quad (2.34)$$

Takto získané kořeny dosadíme do rovnic (2.26) pro první nebo druhou hyperbolu, čímž získáme dvě řešení, které otestujeme a vybereme to, které odpovídá vypočteným zpožděním. Může nastat situace, při které vyhovují obě řešení, v tom případě využijeme čtvrtého mikrofónu pro jednoznačné určení správného řešení.

3 IMPLEMENTACE METOD A PRAKTICKÉ OVĚŘENÍ

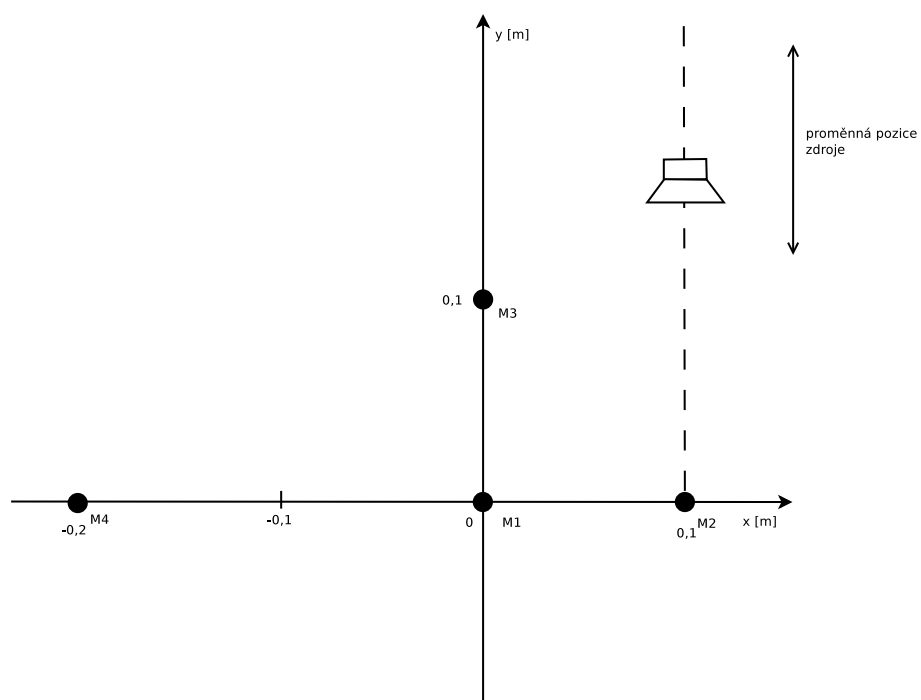
Tato kapitola se věnuje praktickému ověření funkčnosti zvolených metod při pokusech jak ve zjednodušených podmínkách semi-bezodrazové komory, tak i v reálných podmínkách. Všechny metody byly implementovány v prostředí GNU Octave a funkční soubory jsou přiloženy na přenosném disku. V následujících podkapitolách jsou jednotlivá měření popsána, jsou vyobrazeny výsledky lokalizace a tyto jsou poté diskutovány.

3.1 Experimentální ověření první

3.1.1 Úvod k měření a jeho popis

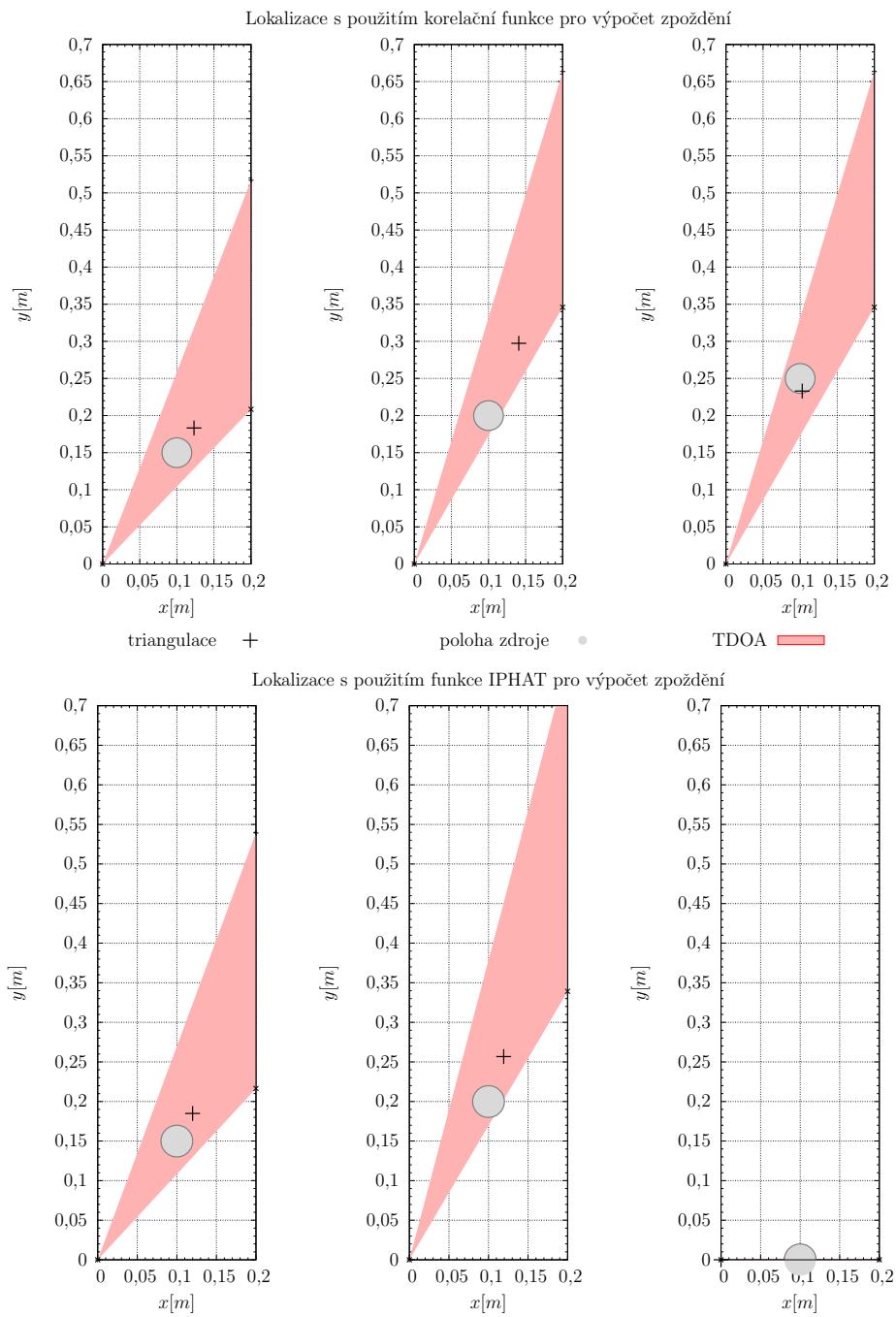
Toto experimentální ověření funkčnosti probíhalo v semi-bezodrazové komoře. Účel tohoto měření bylo prověření chování metody TDOA a metody triangulace v závislosti na vzdálenosti zdroje od pole. Předpoklad byl ten, že ve větších vzdálenostech zdroje od pole bude metoda TDOA přesnější výsledky lokalizace, než v menších vzdálenostech. Z rozdílu přesnosti by pak v ideálním případě mohla být stanovena přibližná hranice vzdáleného pole. Pro toto měření bylo zvoleno takové rozložení pole, aby na získané výsledky bylo možno aplikovat jak metodu triangulace, tak TDOA. Zvolené rozložení pole je na obrázku 3.1. Jako zdroj byl použit sluchátkový měnič, aby jsme se co nejlépe přiblížili ideálnímu bodovému zdroji. Zdroj jsme umisťovali do různých vzdáleností od mikrofonu číslo dvě po ose, která je na obrázku vyznačena přerušovanou čarou a je rovnoběžná s osou y . Zdrojovým signálem byl zvolen signál MLS.

Zaznamenané signály poté byly zpracovány metodami TDOA a triangulace za použití jak metody výpočtu zpoždění pomocí korelační funkce, tak i metodou IPHAT. Vypočtené výsledky lokalizace jsou vizualizovány na obrázcích 3.2, 3.3 a 3.4.

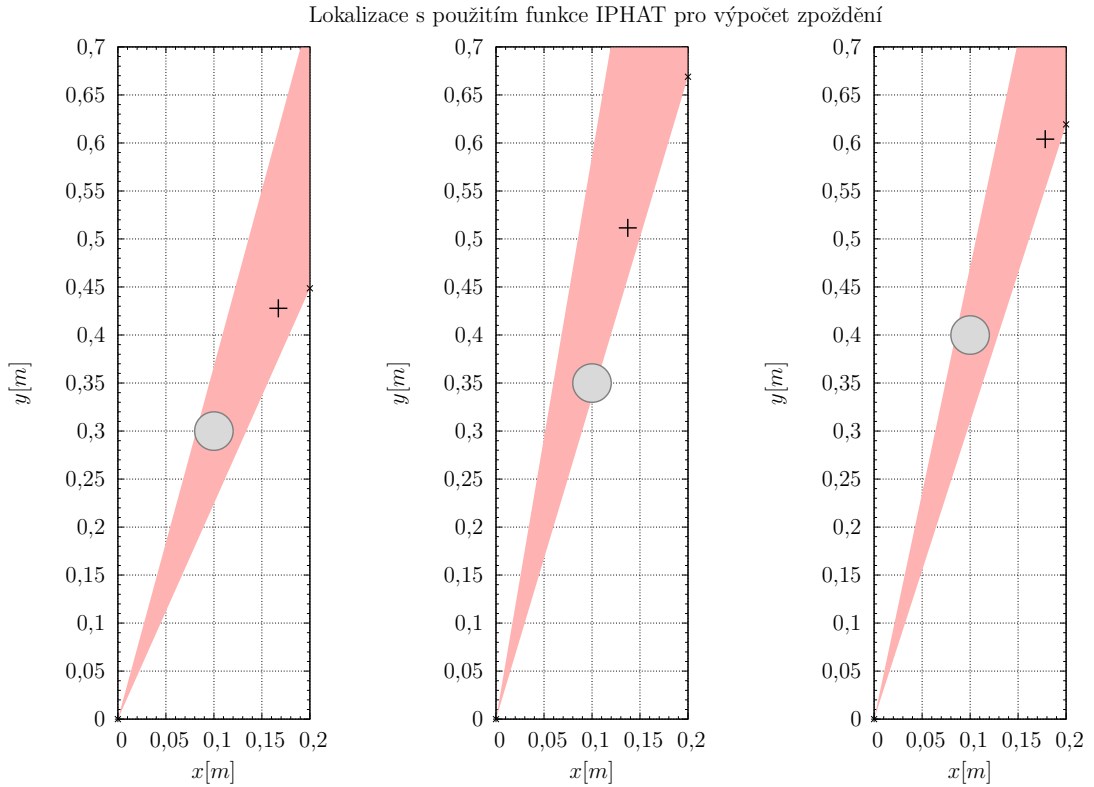
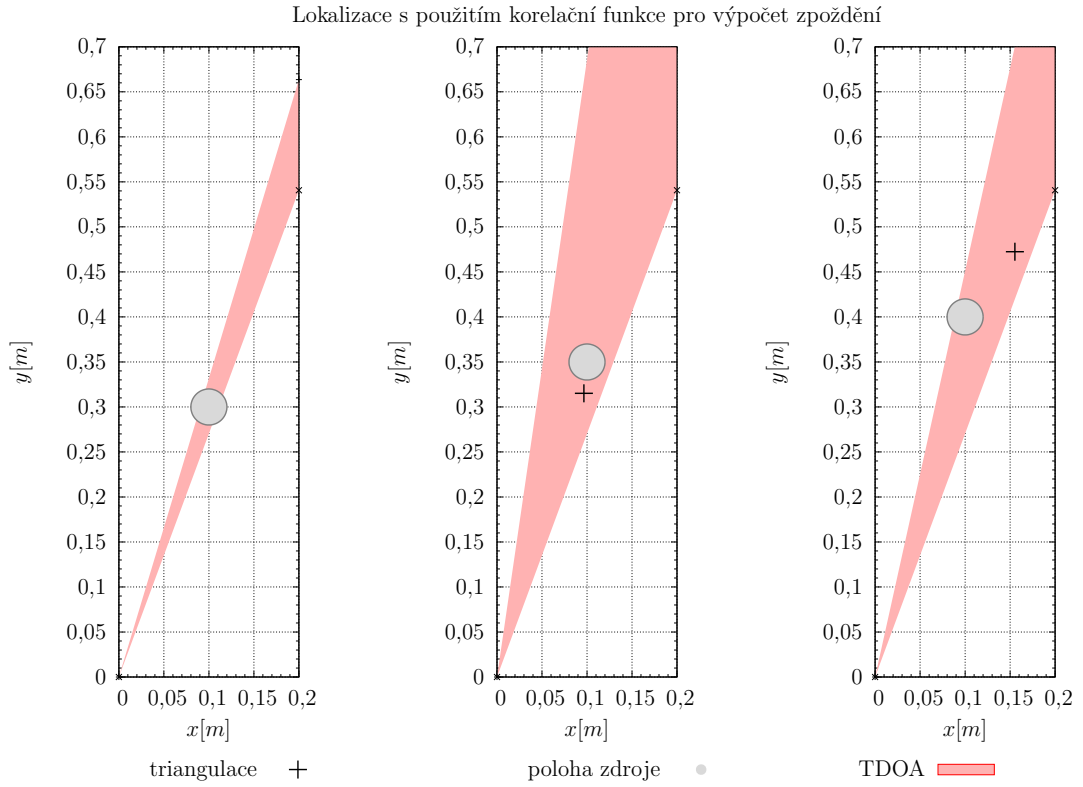


Obr. 3.1: Diagram rozložení pole a polohy zdroje při měření.

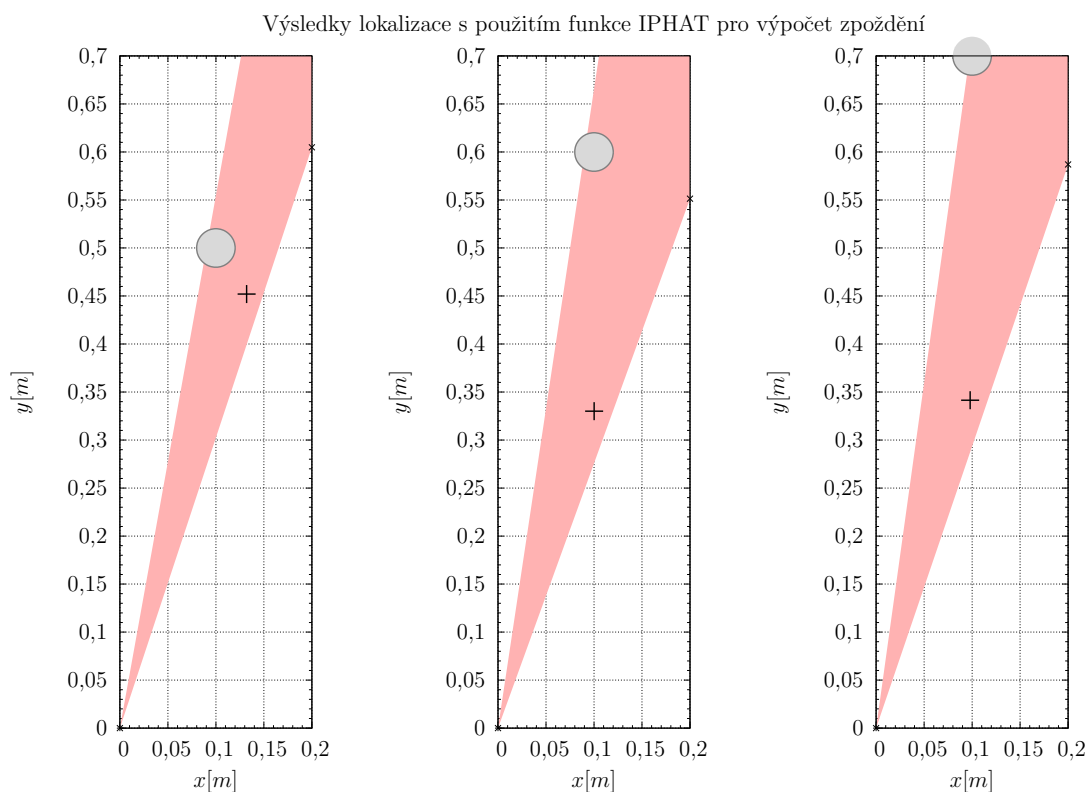
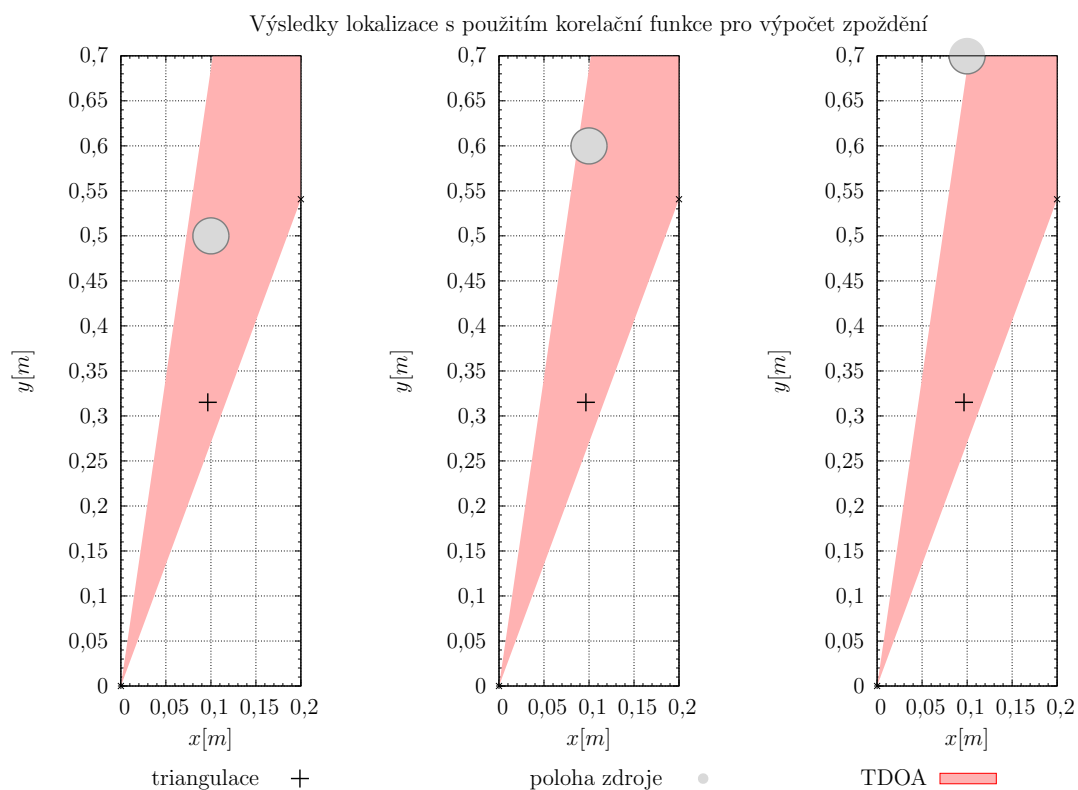
3.1.2 Vizualizace výsledků



Obr. 3.2: Vizualizace výsledků pro první měření (1/3). Triangulace zde vychází s chybou do 15 cm, TDQA nepodává příliš přesné výsledky. Metoda IPHAT podala lepší výsledek než korelační funkce při poloze zdroje $[0,1; 0,5]$. Pro poslední polohu tohoto grafu IPHAT nemohla být vypočtena – nahrané sekvence nebyly stejně dlouhé (pravděpodobně SW chyba).



Obr. 3.3: Vizualizace výsledků pro první měření (2/3). Lze postřehnout zpřesnění výsledků TDOA. V prvním případě KF triangulace určila zdroj, avšak mimo zobrazenou oblast. V dalších dvou případech byla přesnější než IPHAT. TDOA podává lepší výsledky při použití IPHAT. Naopak triangulace s IPHAT má zhoršující se trend úspěšnosti.



Obr. 3.4: Vizualizace výsledků pro první měření (3/3). KF pravděpodobně přestala fungovat z důvodu horší odolnosti proti šumu při nízkém SNR. IPHAT fungovala i v poloze zdroje $[0,1; 0,5]$, dále už však také nestačila.

3.2 Diskuze výsledků

Z vizualizací je patrné, že obě metody jsou funkční a jsou schopny zdroj lokalizovat. Bohužel sluchátkový měnič zřejmě nebyl dostatečně výkonný, aby mohl být dostatečně hlasitým zdrojem zvuku. Korelační funkce proto na obr. 3.4 lokalizuje zdroj pořád na stejném místě, respektive TDOA dává stejné úhly, lze se domnívat, že již nerozeznává zdroj a zpoždění je pravděpodobně vypočteno pro nějaké neměnné zpoždění šumu mezi jednotlivými kanály, toto však nebylo ověřeno. Na stejném obrázku se však ukazuje, že metoda IPHAT je opravdu o něco robustnější proti šumům a zvládne lokalizovat zdroj i za podmínek, kdy to korelační funkce již nezvládla. Tato robustnost se však nejvíce jeví být příliš výraznou.

Metoda triangulace nepodává příliš dobré výsledky. Lokalizace je dosti nepřesná, v některých případech nebyl zdroj dokonce lokalizován vůbec. Nezdá se ani, že by na úspěšnost měla metoda IPHAT pozitivní vliv, přestože v našem případě určuje zpoždění mezi signály i na necelé vzorky. Literatura však uvádí u této metody úspěšnost pouze kolem 2 % [6], není proto asi nutné se nad horšími výsledky příliš pozastavovat.

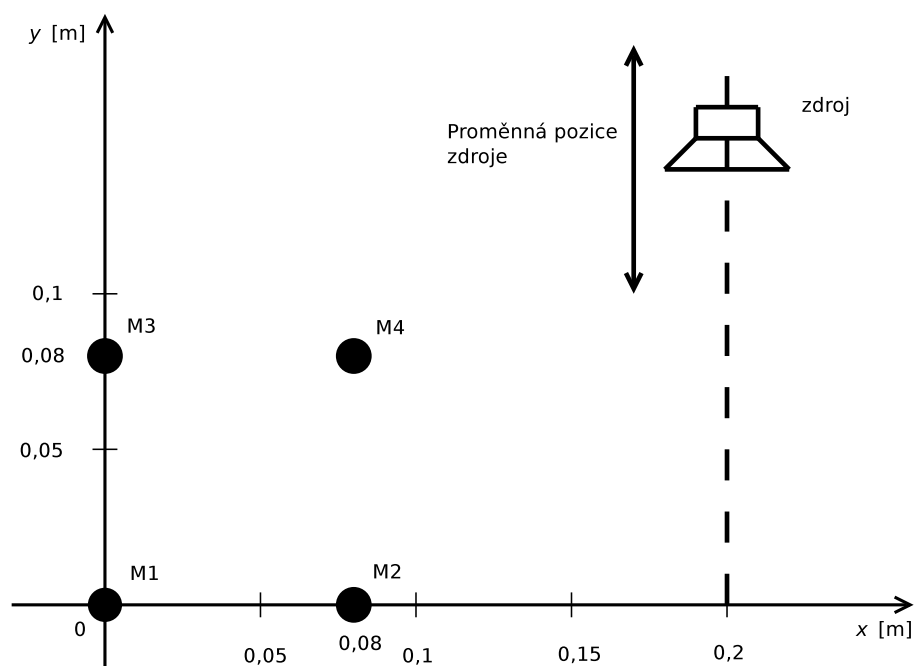
Použití metody IPHAT v tomto případě nepřineslo výsledky, které by měly výrazně přesnější. Tato metoda určení zpoždění však funguje spolehlivě. V jednom případě ale k výpočtu nedošlo (viz obrázek 3.3). Po krátkém hledání problému bylo zjištěno, že nahrané signály neměli stejnou délku, ačkoliv by měli být nahrávány synchronně. Domnívám se, že toto bylo způsobeno chybou softwaru Cubase. Při dodržení stejného postupu zaznamenávání se totiž tato chyba vyskytla vícekrát i při dalších měřeních.

3.3 Experimentální ověření druhé

3.3.1 Úvod k měření a jeho popis

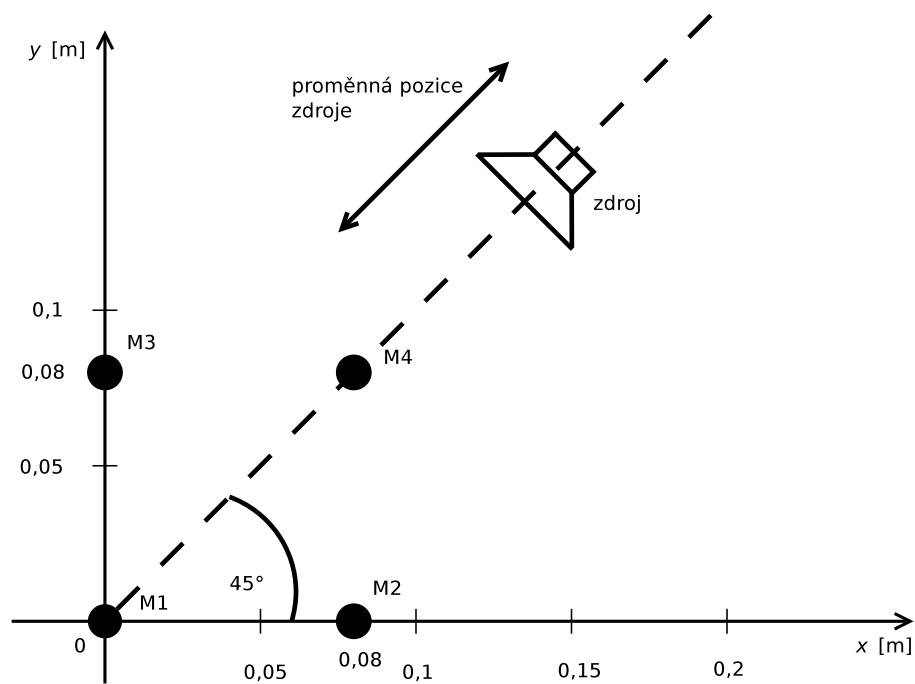
Záměr tohoto pokusu bylo ověřit funkčnost zvolených metod v reálných podmínkách a jejich schopnost lokalizovat zdroje jak úzkopásmové, tak širokopásmové. Měření tedy probíhalo v laboratoři PA-136, která poskytuje dobré zázemí pro podobná měření, jelikož je vybavena vhodným softwarem i hardwarem. Jako zástupce úzkopásmových signálů byl zvolen signál se sinusovým průběhem o frekvenci 1000 Hz. Za zástupce širokopásmových signálů byl zvolen růžový šum. Jako zdroj byl použit širokopásmový reproduktor značky Genelec.

Měření mělo dvě části. První probíhala dle obr. 3.5. Zdroj jsme umístili do bodu $[0, 2; 1]$ a postupně jsme ho posunovali rovnoběžně s osou y po 0,2 m až do vzdálenosti



Obr. 3.5: Diagram rozložení pole a polohy zdroje při měření.

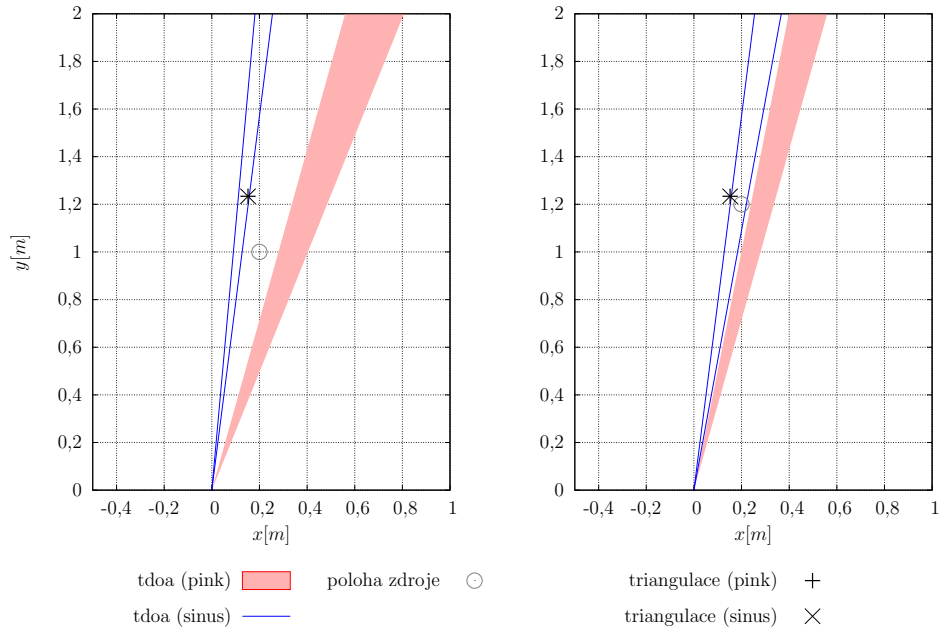
3m. V druhé části jsme zdroj umístili do vzdálenosti jednoho metru od počátku a v úhlu 45° , tak jak je znázorněno na obrázku 3.6.



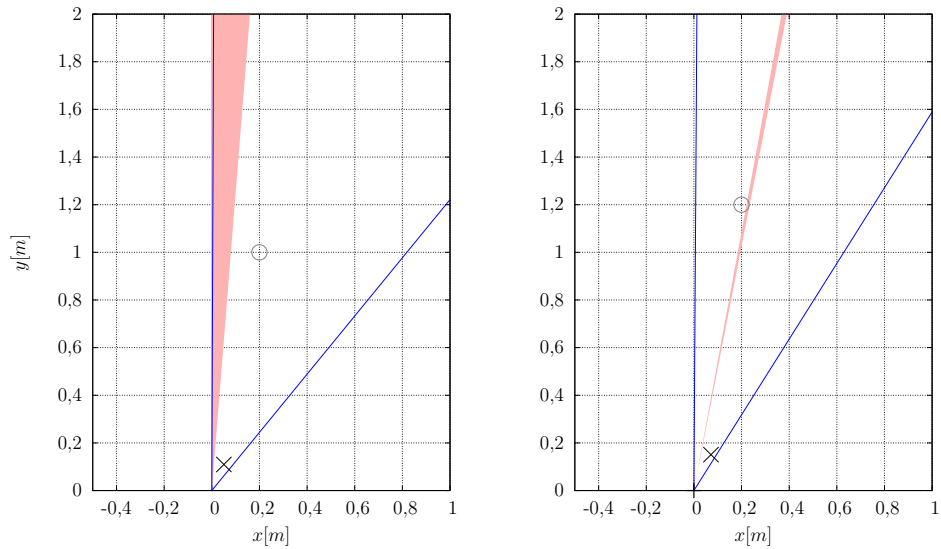
Obr. 3.6: Diagram rozložení pole a polohy zdroje při měření.

3.3.2 Vizualizace výsledků

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

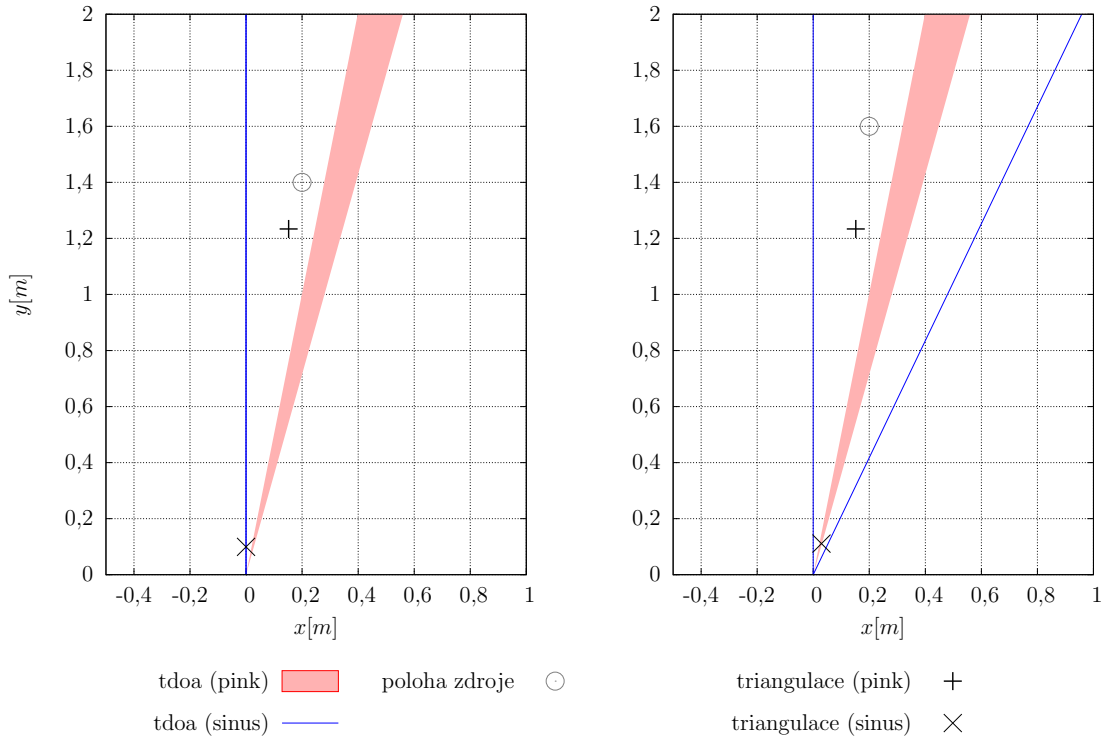


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

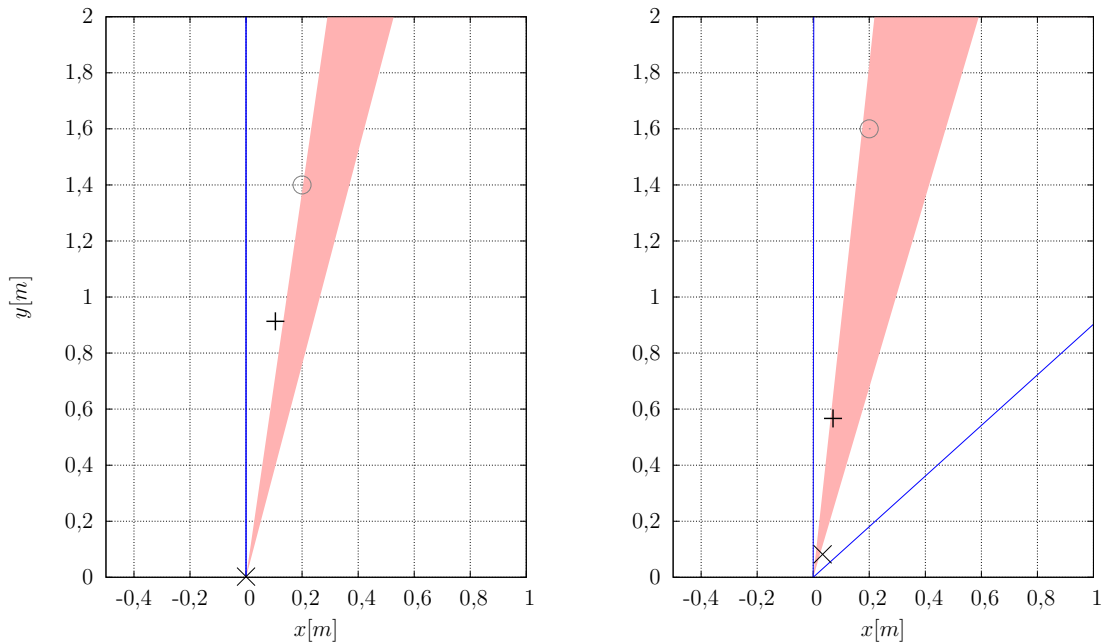


Obr. 3.7: Vizualizace výsledků pro druhé měření (1/12). Triangulace při použití KF zaměřila zdroj velmi přesně v poloze $[0,2; 1,2]$ pro oba druhy signálů. TDOA s KF velmi dobře lokalizovala sinusový signál. TDOA s IPHAT zde lokalizuje zdroj sinu hůře. Naopak u růžového šumu proběhla lokalizace velmi dobře ve vzdálenosti 1,2 m.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

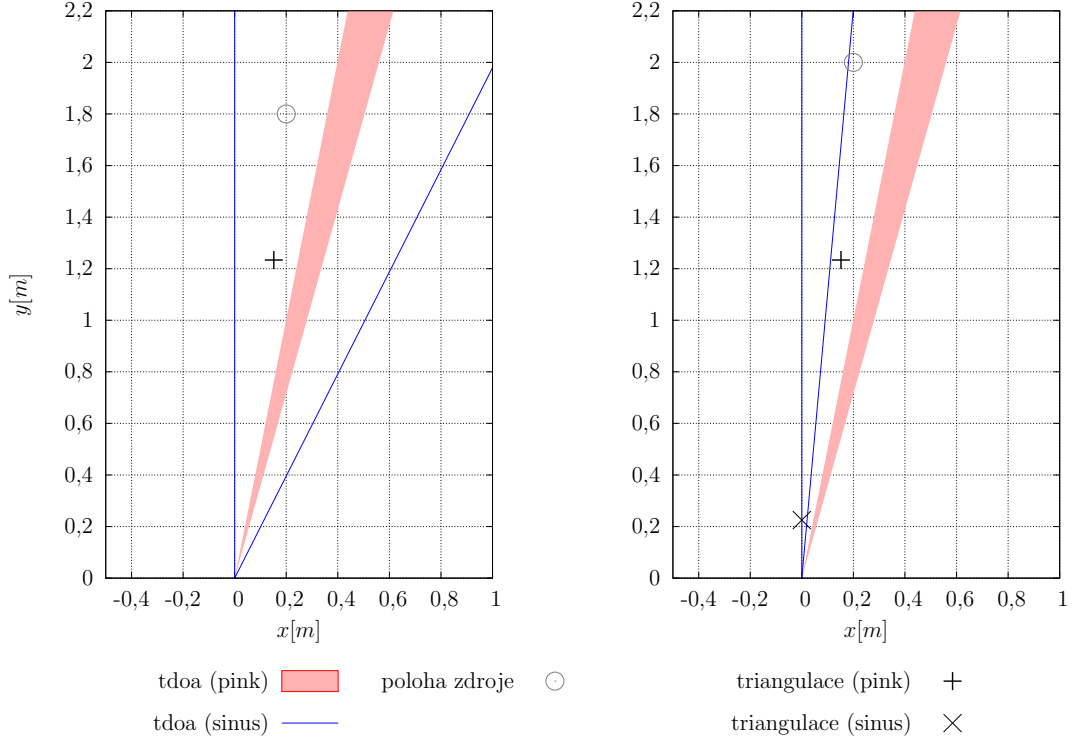


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

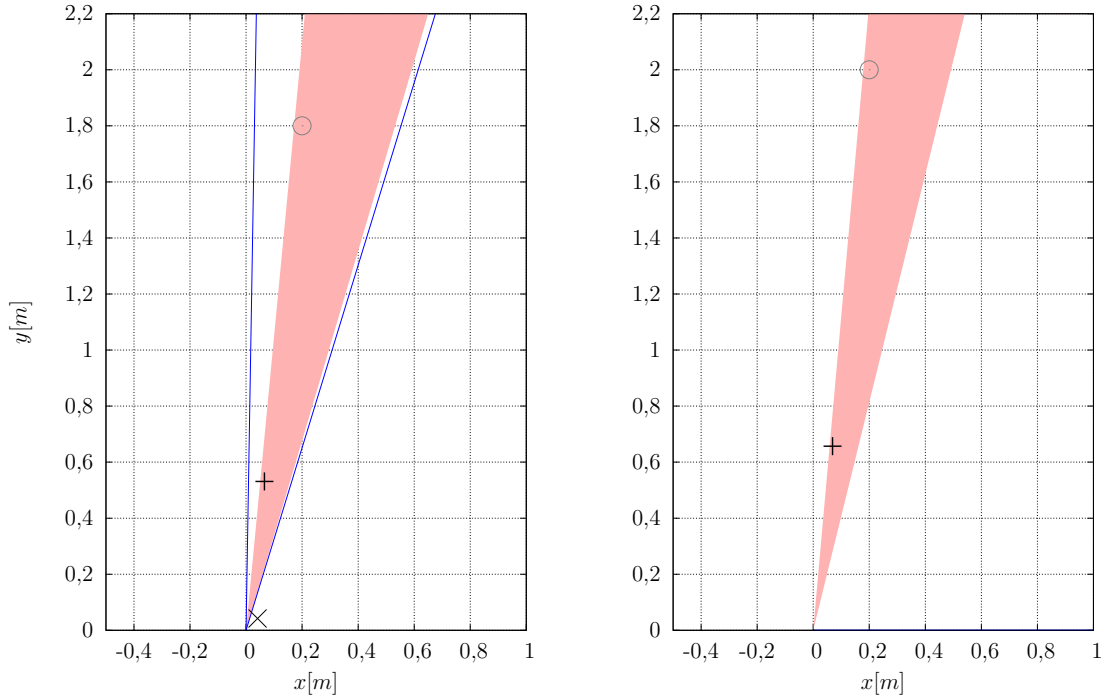


Obr. 3.8: Vizualizace výsledků pro druhé měření (2/12). KF + triangulace funguje jen pro růžový šum. KF + TDOA – pro sinus v prvním případě oba úhly shodně 90° . Celkově lepší výsledky pro růžový šum. IPHAT + TDOA dala pro sinus horší výsledky, růžový šum lokalizuje poměrně úspěšně. IPHAT + triangulace pro sinus nefunguje a pro šum postupně také ztrácí na přesnosti.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

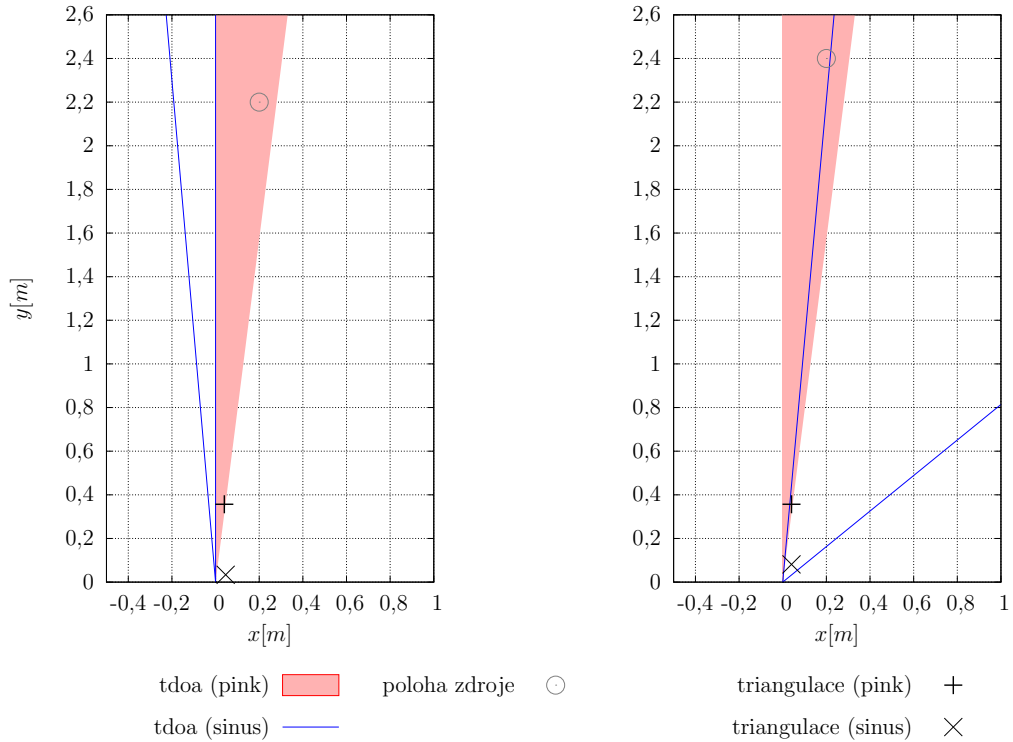


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

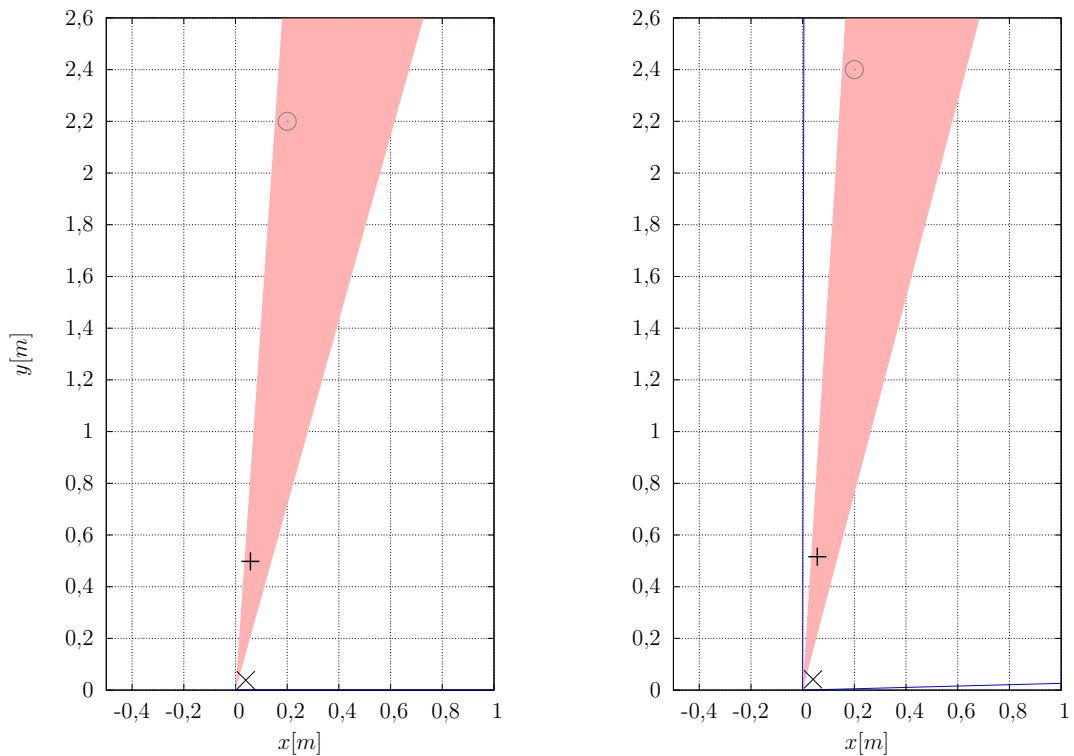


Obr. 3.9: Vizualizace výsledků pro druhé měření (3/12). KF + TDOA viditelně horší přesnost pro sinusový signál. Šum však také nebyl správně zaměřen. Triangulace při šumu zhoršuje schopnosti lokalizace, sinus nelokalizuje prakticky vůbec. IPHAT zde podává lepší výsledky, při vzdálenosti 2 m se však žádné dva úhly nenacházely ve společném kvadrantu. Triangulace tady už v podstatě nefunguje.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

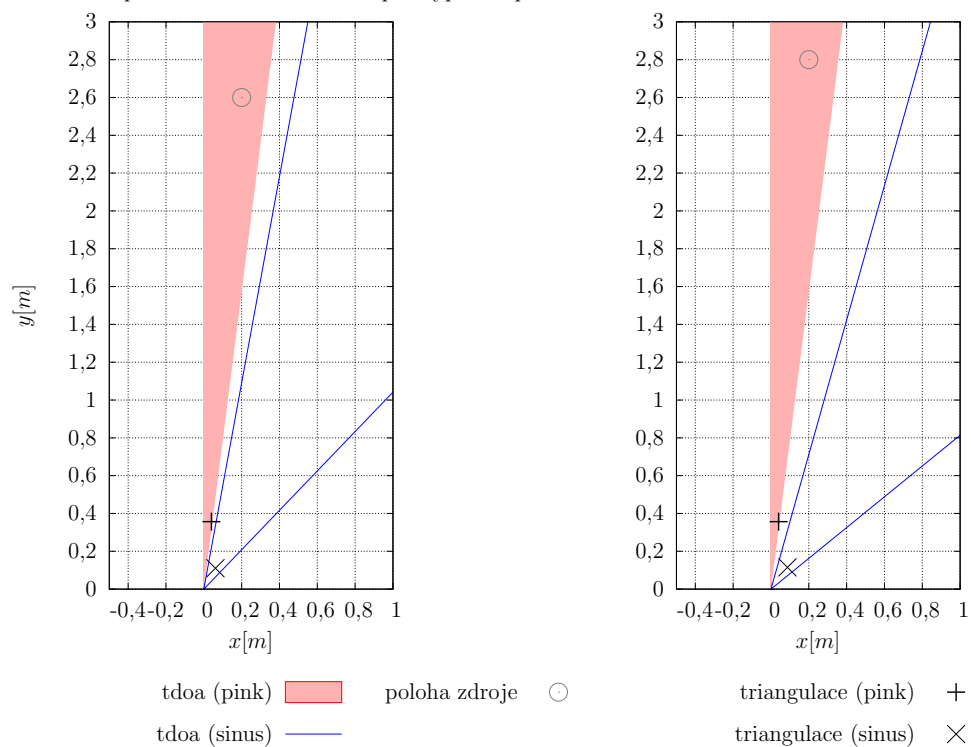


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

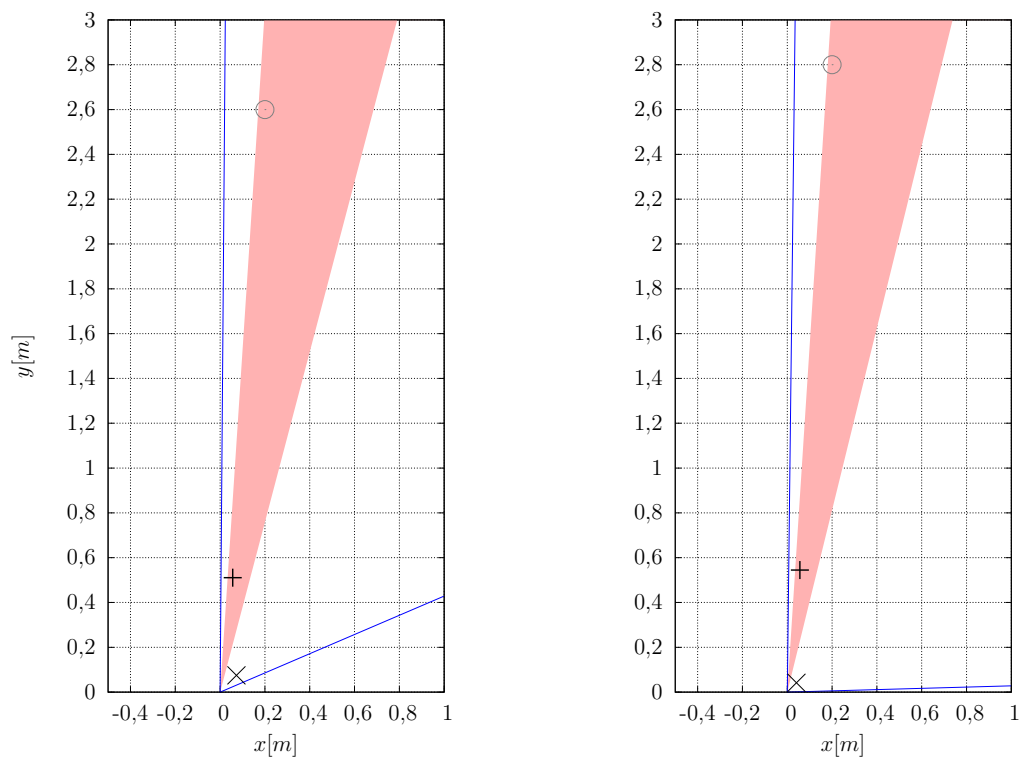


Obr. 3.10: Vizualizace výsledků pro druhé měření (4/12). Triangulace již přestává fungovat – nebudeme ji dále komentovat. KF + TDOA sinus lokalizovala s výrazně nižší přesností, zdroj šumu byl zaměřen spolehlivě. Při IPHAT + TDOA se opakuje situace z předešlého obrázku a úhly pro sinusový se znovu nepotkali. U vzdálenosti 2,4 m je lokalizace sinu velmi nepřesná. Šumový signál je lokalizován, avšak s menší přesností než u KF.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

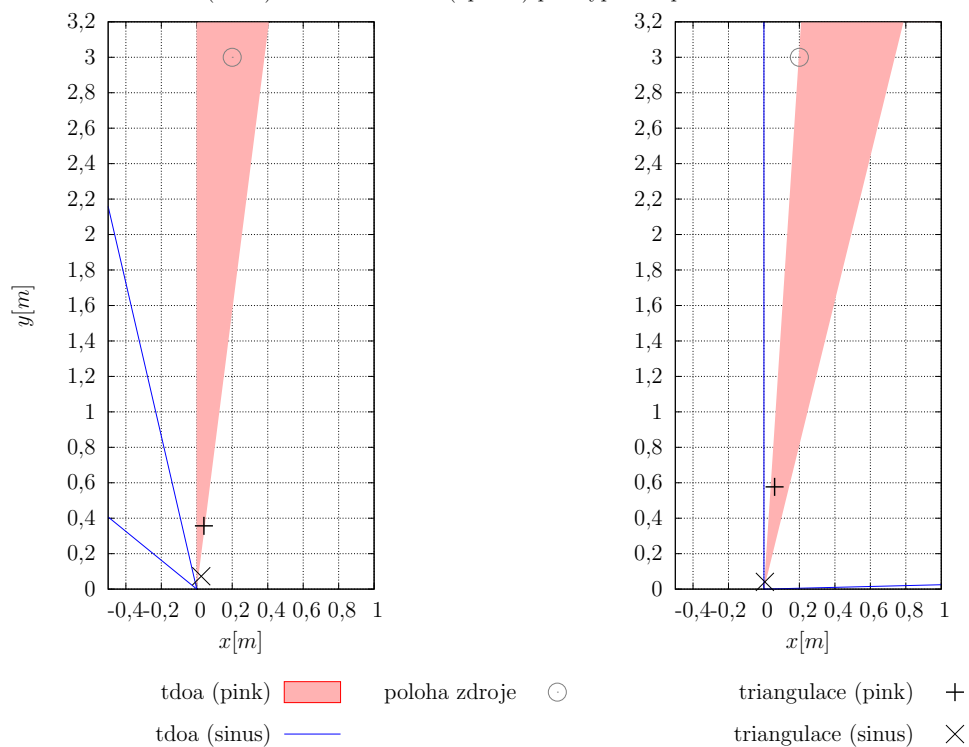


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění



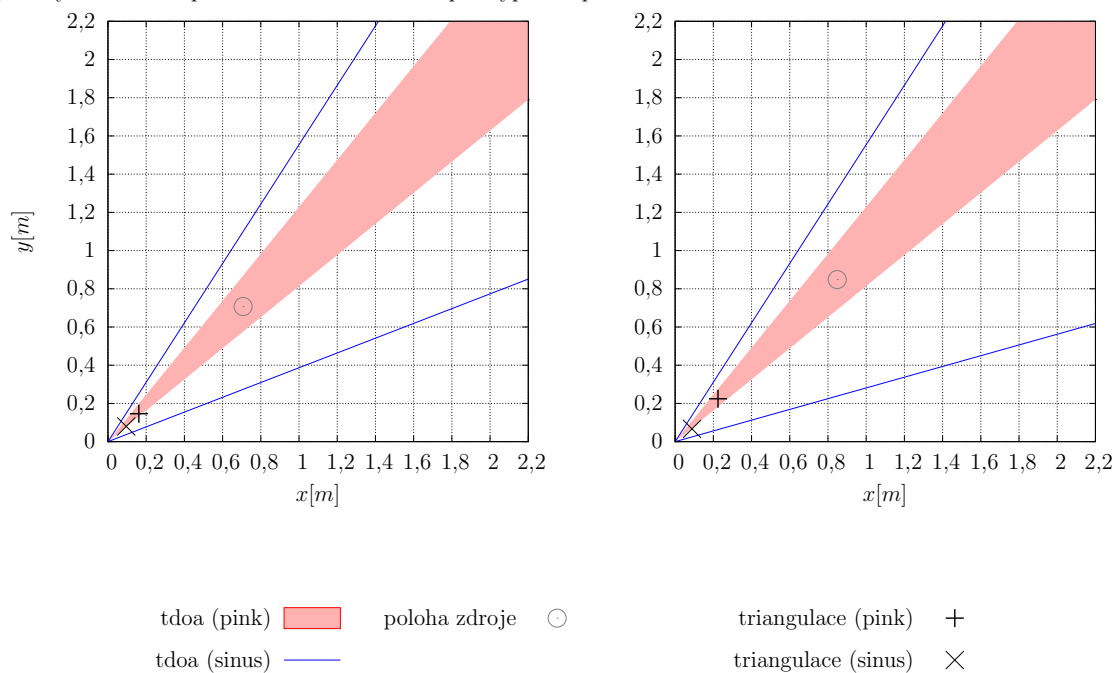
Obr. 3.11: Vizualizace výsledků pro druhé měření (5/12). Pro KF + TDOA má lokalizace sinusového signálu zhoršující se tendenci. IPHAT stále nepřesně lokalizuje sinus. Jeden pár mikrofónů však lokalizuje zdroj šumového signálu téměř přesně.

Výsledky lokalizace s použitím korelační funkce (vlevo) a funkce IPHAT (vpravo) pro výpočet zpoždění

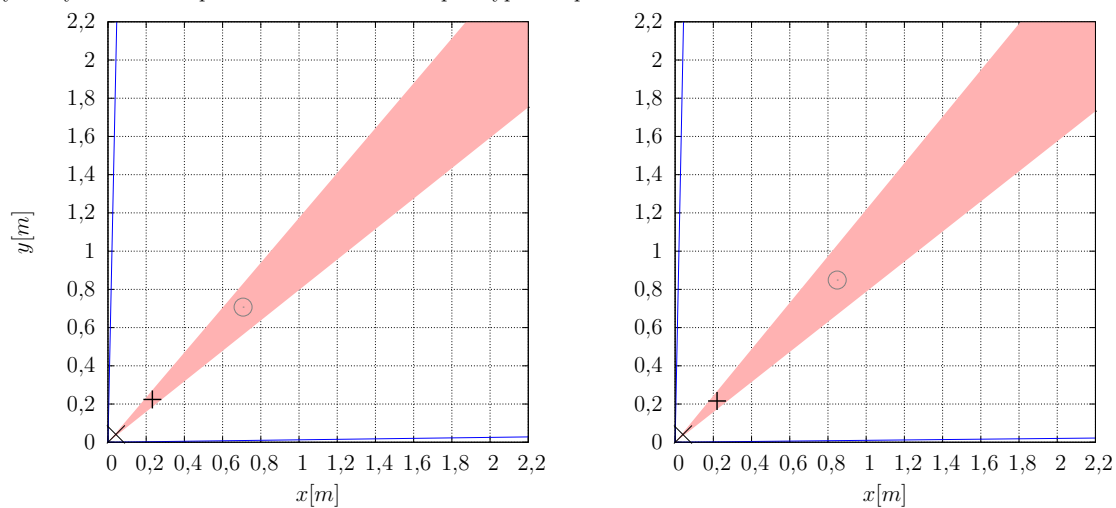


Obr. 3.12: Vizualizace výsledků pro druhé měření (6/12). Lokalizace sinusového signálu je v obou případech velmi nepřesná. Širokopásmový signál je naopak lokalizován metodami dobře.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

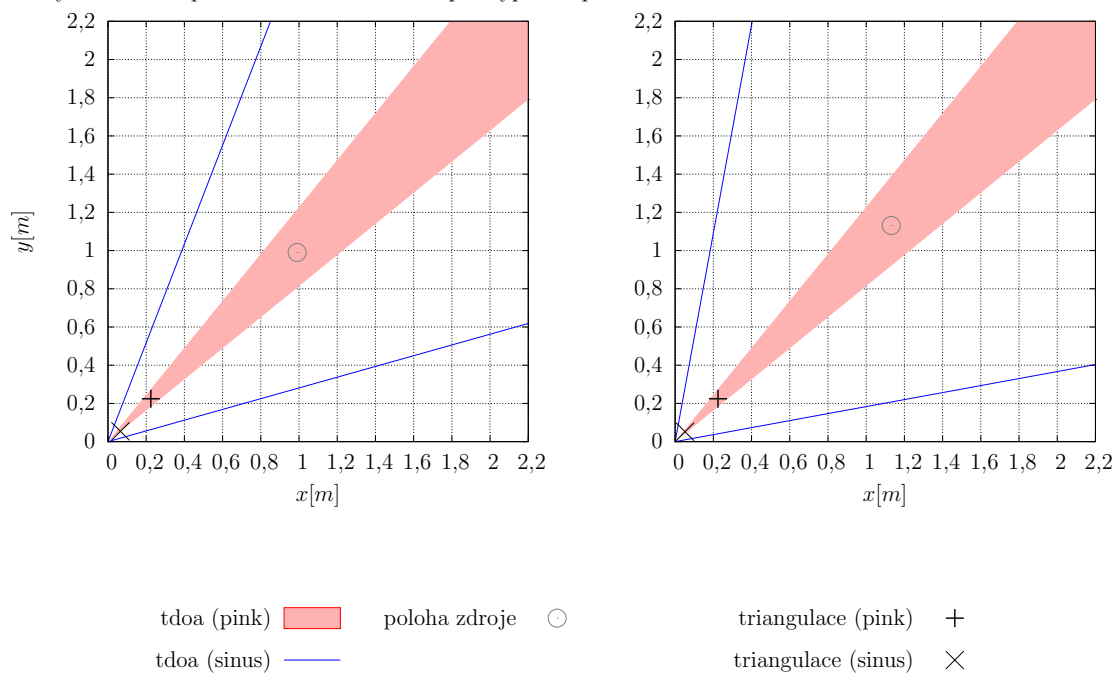


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

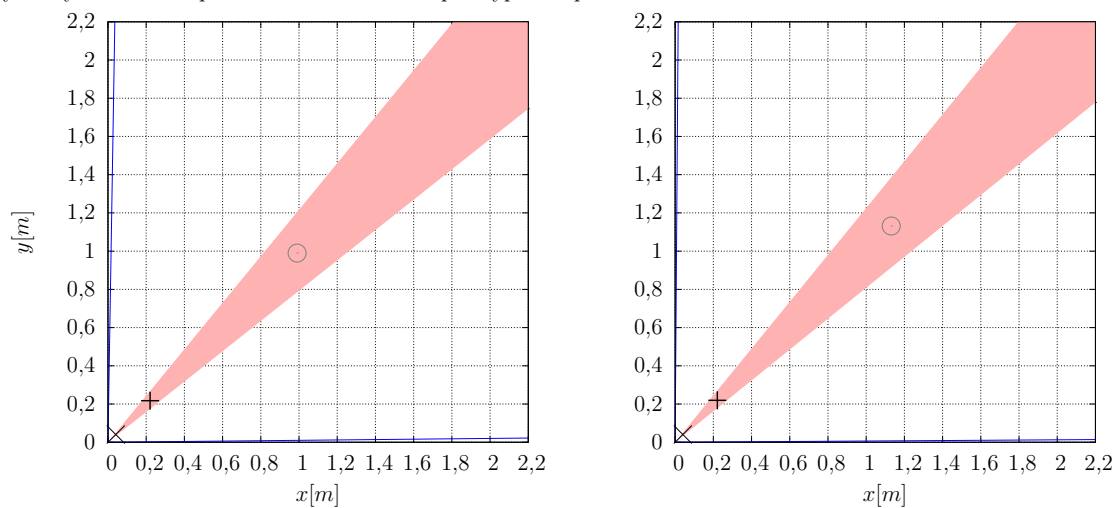


Obr. 3.13: Vizualizace výsledků pro druhé měření (7/12). Triangulace je velmi nepřesná, nebudeme ji tedy dále komentovat. KF + TDOA dává dobré výsledky pro šum, lokalizace zdroje sinu je však horší. IPHAT + TDOA podává nepatrně lepší výsledky v případě růžového šumu. Pro sinusový signál je však lokalizace velmi nepřesná.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

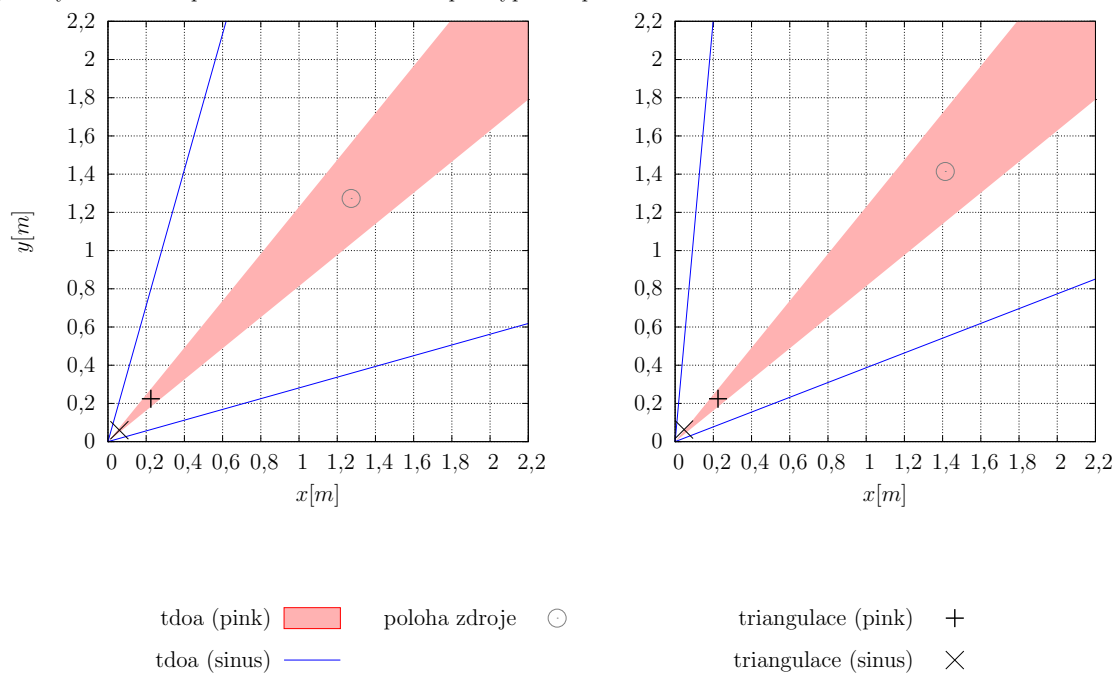


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

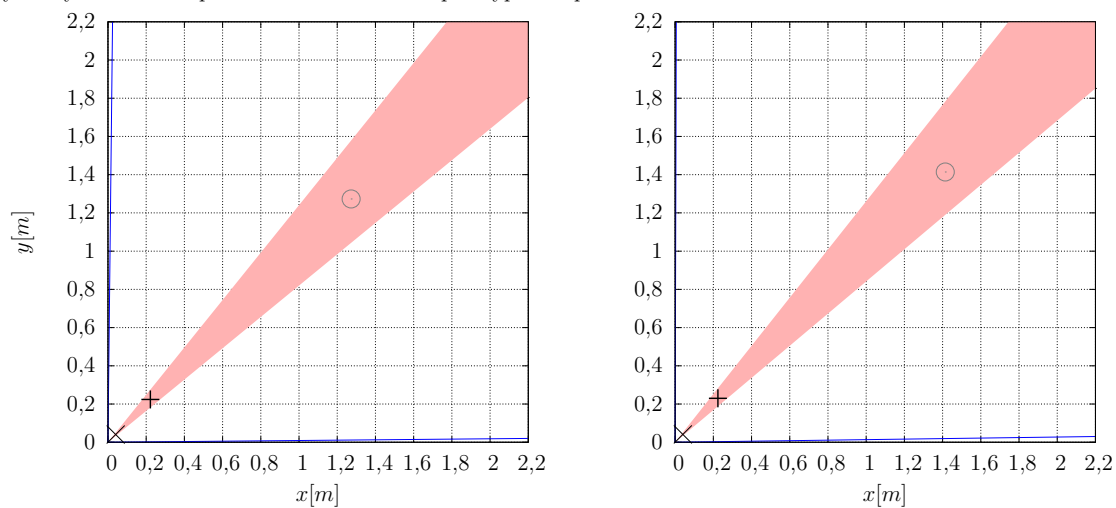


Obr. 3.14: Vizualizace výsledků pro druhé měření (8/12). Zde je situace velmi podobná obrázku 3.13. U KF navíc dochází k dalšímu zhoršení přesnosti pro sinusový signál.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

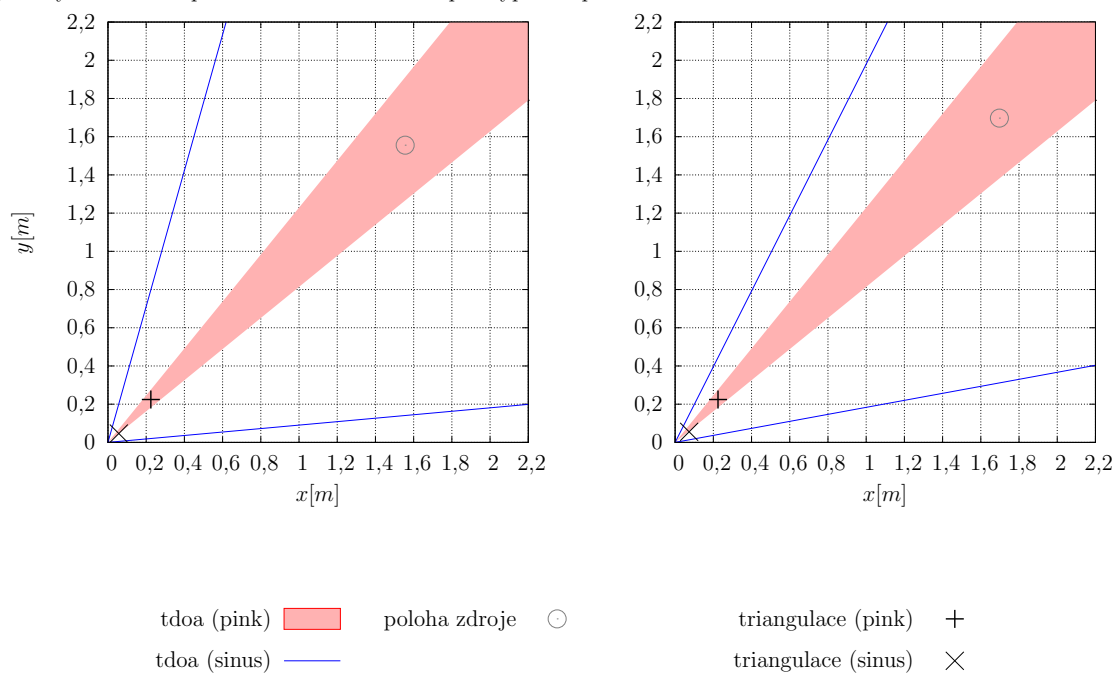


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

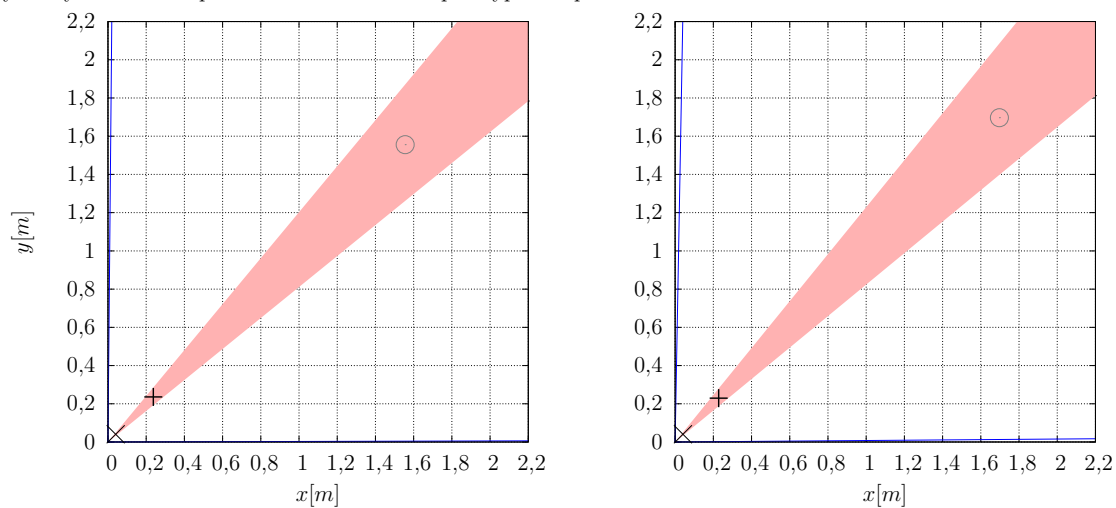


Obr. 3.15: Vizualizace výsledků pro druhé měření (9/12). U KF sledujeme stále se zhoršující přesnost pro sinusový signál. Pro šumový signál je přesnost lokalizace stabilní.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

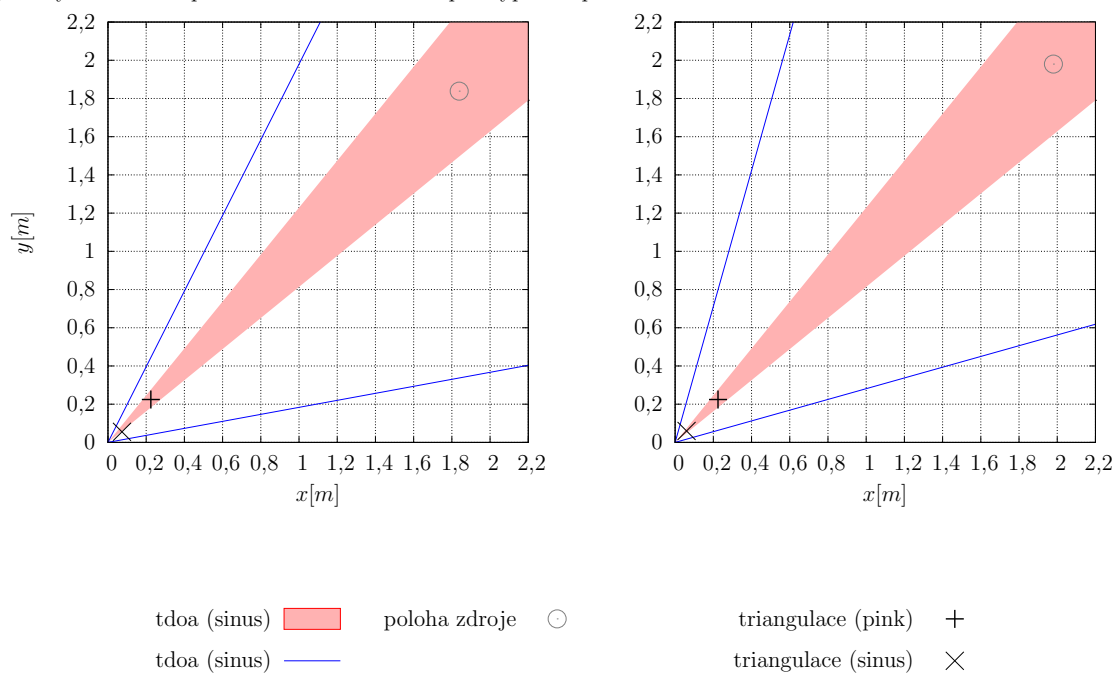


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění

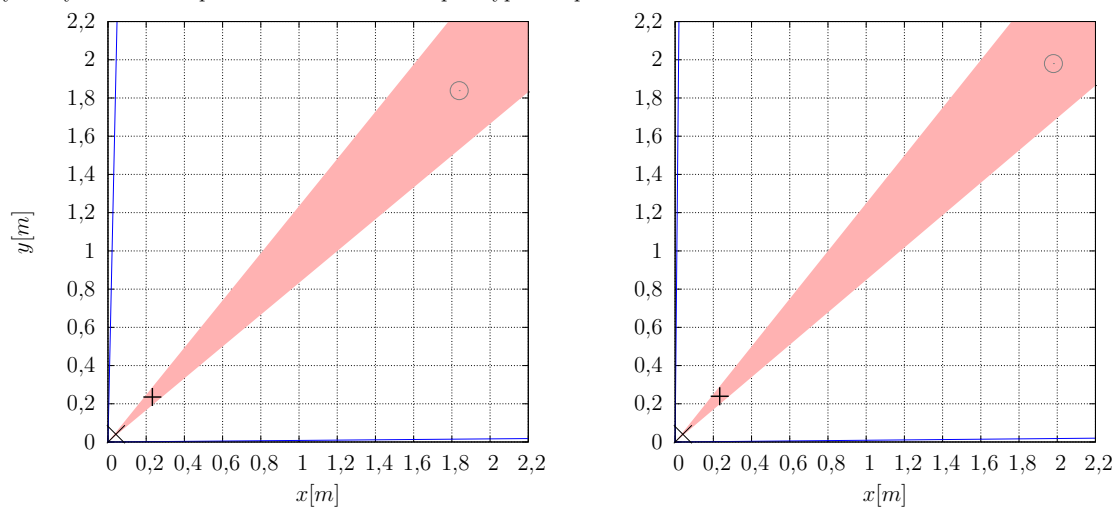


Obr. 3.16: Vizualizace výsledků pro druhé měření (10/12). Velmi podobná situace jako na obr. 3.16.

Výsledky lokalizace s použitím korelační funkce pro výpočet zpoždění

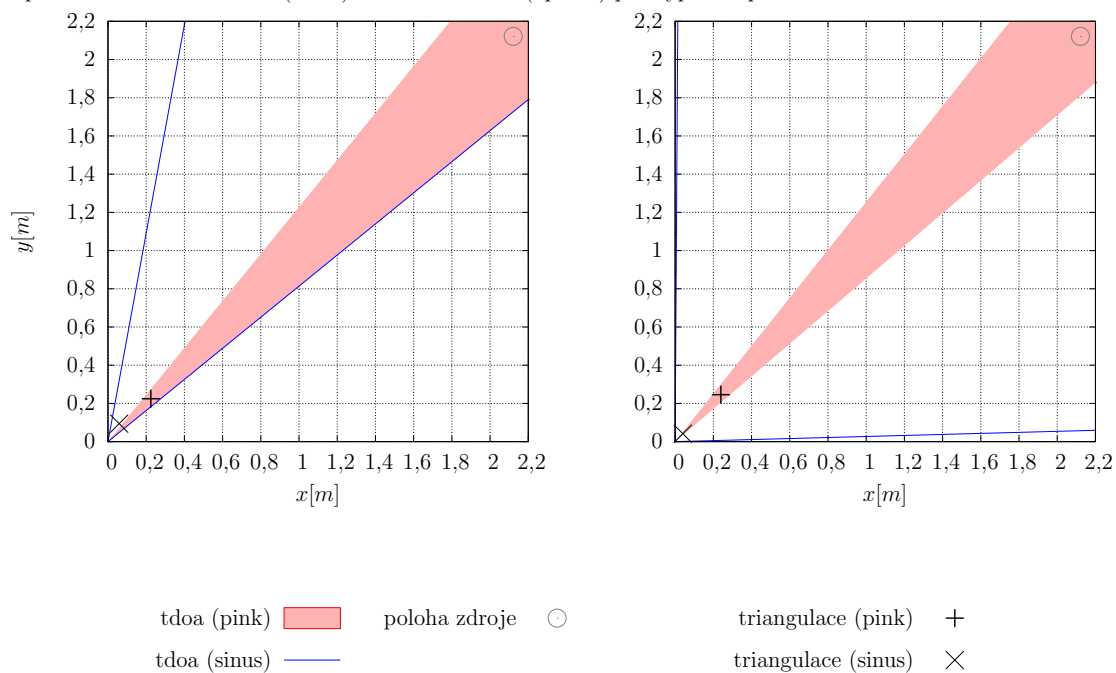


Výsledky lokalizace s použitím funkce IPHAT pro výpočet zpoždění



Obr. 3.17: Vizualizace výsledků pro druhé měření (11/12). i zde vidíme obdobný výsledek jako v předchozím případě.

Výsledky lokalizace s použitím korelační funkce (vlevo) a funkce IPHAT (vpravo) pro výpočet zpoždění



Obr. 3.18: Vizualizace výsledků pro druhé měření (12/12). KF + TDOA určili z jednoho páru mikrofónů stejný úhel pro oba druhy signálů. Jinak se však nezdá, že by došlo ke změně situace.

3.4 Diskuze výsledků

Při diskuzi výsledků se budeme nejprve věnovat první části experimentu, která probíhala dle obrázku 3.5. Prvně však podotkneme několik faktů. Pokud na vizualizaci není vidět poloha zdroje vypočtená triangulací, pak funkce nevrátila výsledek, což znamená, že nedošlo k protnutí hyperbol, nebo se výsledek nachází mimo zobrazenou oblast. Obě tyto možnosti však znamenají stejný výsledek, že lokalizace byla neúspěšná. Také by mělo být připomenuto, že výsledky metody TDOA by měly zaměřovat přímo zdroj. Zdroj by tedy neměl ležet mezi přímkami, ale přímkou by měly zdrojem procházet a být tzv. „přes sebe“. Pokud není zobrazen výsledek pro metodu TDOA, znamená to, že se ve výsledku nenacházely žádné dva úhly ve stejném kvadrantu.

Obecně můžeme říci, že se pro širokopásmové signály u TDOA jako přesnější jeví metoda IPHAT, jistě i pro to, že zpoždění určí i na zlomky vzorků. Avšak použijeme-li tuto metodu na úzkopásmový, v našem případě harmonický signál, zdá se, že metoda naprosto přestává fungovat a korelační funkce si v tomto případě vede lépe. Domnívám se, že je to způsobeno tím, že IPHAT si odruší šumovou složku signálu a vzhledem k velké periodicitě výsledného signálu pak již není schopen určit přesné zpoždění.

Metoda triangulace se při těchto pokusech jeví jako naprosto nevyhovující, její výsledky se zdají být, zejména ve větších vzdálenostech, spíše náhodné.

Hranice vzdáleného pole se nezdá být ani zde příliš patrná. Při použití IPHAT-TDOA na MLS signál se jeden z úhlů již od vzdálenosti 1,2 m trefuje velmi blízko skutečné polohy zdroje. Dá se tedy předpokládat, že při dalším zvětšování vzdálenosti bychom se dostali do situace, kdy oba úhly budou velmi blízko správnému úhlu.

Nyní se podíváme na další část pokusu, která probíhala dle obrázku 3.6. Začneme metodou TDOA. Zde se nezdá, že by korelační funkce podávala horší výsledky než IPHAT. Při lokalizaci sinusového signálu se spíše zdá, že tradiční korelační funkce si vede lépe než IPHAT. To že se zdroj nachází vždy téměř přesně na ose úhlu, který svírají přímkou výsledky je způsobeno pravděpodobně zakřivením šířící se vlnoplochy.

U metody triangulace se bohužel nedá říci, že by v tomto případě poskytla nějaký relevantní výsledek lokalizace. Můžeme si sice všimnout, že propojením počátku a výsledkem lokalizace bychom získali úhel, který ve většině případů směřuje ke skutečnému zdroji zvuku. Toto však není požadovaný výsledek. Vzhledem k výpočetní náročnosti a nespolehlivosti se metoda triangulace jeví jako nevhodná.

4 ZÁVĚR

V této bakalářské práci se zabývala některými vybranými metodami lokalizace akustických zdrojů. Implementovány byly metody ve své podstatě základní, jmenovitě metoda triangulace a metoda TDOA. Vybrané metody byly implementovány a byly prakticky testovány jak v semi-bezodrazové komoře, tak v prostředí velmi blízkém reálným podmínkám. Vypočtené výsledky byly vizualizovány a podrobně popsány v příslušných kapitolách.

Z výsledků plyne, že algoritmus metody triangulace je funkční a v ideálních podmínkách je schopen zdroj zvuku opravdu lokalizovat. V praxi se však ukazuje, že spolehlivost tohoto algoritmu je velmi nízká, téměř až nulová. Pokusili jsme se zvýšit úspěšnost lokalizace touto metodou použitím funkce IPHAT, která byla navíc rozšířena o schopnost určení zpoždění mezi signály na zlomky vzorků. Avšak ani toto posílení nepřineslo žádné viditelné zvýšení přesnosti metody. Funkce IPHAT navíc razantně zvýšila již tak vysoké výpočetní nároky tohoto algoritmu. Toto budou s největší pravděpodobností také důvody, proč není tato metoda při lokalizaci akustických zdrojů příliš používána. Triangulace je navíc schopna lokalizovat pouze jeden zdroj zvuku, což může být v moderních aplikacích také omezující faktor.

Na druhou stranu metoda TDOA se zdá být poměrně schopným algoritmem. Její výpočetní náročnost není zvláště vysoká a její schopnost lokalizace zdrojů může být v některých jednoduchých aplikacích plně dostačující. Jelikož se nám při našich měřeních nepovedlo určit hranici vzdáleného pole, a jelikož vypočtené úhly sobě nebyly příliš blízké ani v největších měřeních vzdálenostech, předpokládám, že jsme se se zdrojem neocitli ve vzdáleném poli mikrofonního pole. To, že zdroj signálu je např. na obr. 3.11 lokalizován jednou dvojicí mikrofónů téměř přesně je pravděpodobně způsobeno tím, že se vůči této dvojici nachází ve velmi ostrém úhlu a zakřivení vlnoplochy zde není příliš patrné, nemůže tedy zkreslovat výsledek. Z toho se dá vyvodit, že přesnost metody TDOA je částečně závislá i na poloze zdroje. I přesto si myslím, že po jednoduchých modifikacích (například zprůměrováním vypočtených úhlů) je metoda poměrně dobře nasaditelná i v situaci blízkého pole.

Jelikož jsou obě metody založeny na výpočtu časových zpoždění, je velmi nutné věnovat velkou pečlivost přípravě mikrofonního pole, zejména potom přesnému umístění jednotlivých měřicích mikrofónů. Z důvodu velkých výkyvů v přesnosti metody triangulace se nabízí otázka, do jaké míry je přesnost ovlivněna tvarem pole, případně polohou zdroje. Můj předpoklad je ten, že tyto faktory budou hrát obrovskou roli ve schopnosti lokalizovat. Pro její nespolehlivost a vysokou výpočetní náročnost bych však tuto metodu nedoporučoval ani pro další studium, ani pro žádnou reálnou aplikaci.

Jako další oblast zkoumání bych navrhoval modernější metody, které jsou založené

například na tvarování přijímací charakteristiky. Jako příklad uveďme základní metodu beamformingu – *delay and sum*. Tato metoda je však také určena pro vzdálené pole. Další velmi zajímavou alternativou by mohli být metody založené na spektrálním odhadu vysokého rozlišení, jako například velmi populární metoda MUSIC (Multiple Signal Classification – vícenásobná klasifikace signálu). MUSIC je schopna lokalizovat zdroje v blízkém poli a dokonce je schopna lokalizovat více zdrojů najednou. Nevýhodou je však rapidně rostoucí výpočetní náročnost při zvyšování šířky pásma lokalizovaného signálu. Přesto si myslím, že pro přesnou lokalizaci bude spolu s příbuznými metodami i přes své nedostatky mnohem zajímavější oblastí výzkumu, zejména pro možnost nasazení v reálných podmínkách.

LITERATURA

- [1] 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.
- [2] EKSLER, V. *Prostorová lokalizace a separace naslepo zdrojů akustických signálů polem mikrofónů*. Doktorská disertační práce. Vysoké učení technické v Brně, 2006. 102 s.
- [3] DRBAL, O. *Měření zpoždění mezi signály EEG* [online]. Diplomová práce. České vysoké učení technické, 2004. 67 s. Dostupné z URL: <<http://ondrej.drbal.cz/school/dp/dp.pdf>>.
- [4] BEZDÍČEK, M. *Lokalizace pohyblivých akustických zdrojů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2010. 56 s. Vedoucí diplomové práce Ing. I. Míča.
- [5] MIKEŠ, P. *Lokalizace akustických zdrojů polem mikrofónů*. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2010. 63 s., 4 s. příloh. Bakalářská práce. Vedoucí práce Ing. I. Míča.
- [6] HORA, P. *Lokalizace zdrojů akustické emise* [online]. Vědecká zpráva. Západočeská univerzita, Institut technologie a spolehlivosti. Plzeň, 1994. 70 s. Dostupné z URL: <http://www.cdm.cas.cz/publications/hora/ph_rep1994_local.pdf>.
- [7] Qin, B; Zhang, H; Fu, Q; Yan, Y. *Subsample Time Delay Estimation via Improved GCC PHAT Algorithm*. Odborný vědecký článek. ThinkIT Speech Lab., institute of Acoustics, CAS, 2008. 4 s.
- [8] Välimäki, V. *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. Disertační práce. Helsinki University of Technology, Faculty of Electrical Engineering, Laboratory of Acoustics and Audio Signal Processing, 1995. 194 s.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

DSP Digital Signal Processing – číslicové zpracování signálů

f_{vz} vzorkovací kmitočet

f_{bw} šířka pásma vzorkovaného signálu

DOA direction of arrival – směr příchodu

2-D dvourozměrný prostor

3-D trojrozměrný prostor

$R(\tau)$ vektor korelačních koeficientů

DFT Discrete Fourier Transform – diskrétní Fourierova transformace

FFT Fast Fourier Transform – rychlá Fourierova transformace

GCC Generalized Cross Correlation – zobecněná křížová korelace

TDOA Time Difference of Arrival – časový rozdíl příchodu

MV Minimum Variance – minimální odchylka

MUSIC Multiple Signal Classification – vícenásobná klasifikace signálu

f_{max} maximální kmitočet zpracovatelný polem

v vlnová rychlost zvuku

d vzdálenost mezi senzory

Δd chyba umístění senzoru

PHAT Phase Transform – fázová transformace

IPHAT Improved Phase Transform – vylepšená fázová transformace

SNR Signal to Noise Ratio – odstup signálu od šumu

KF Korelační funkce

SEZNAM PŘÍLOH

A Obsah přiloženého média

48

A OBSAH PŘILOŽENÉHO MÉDIA

- Elektronická verze semestrálního projektu ve formátu `pdf`.
- Naměřené signály ve formátu `wav` i `m`.
- Složka naprogramovaných souborů v GNU Octave.