

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

دانشگاه یزد
دانشکده مهندسی برق و کامپیوتر
گروه مخابرات

پایان نامه
برای دریافت درجه کارشناسی ارشد
مهندسی برق مخابرات گرایش سیستم

مکان یابی منبع صوت با استفاده از شبکه های حسگر صوتی Ad-hoc

استاد راهنما: دکتر حمیدرضا ابوطالبی
پژوهش و نگارش: حامد قناعت پیشه

مهر ۱۳۹۲

تقدیم به پدر بزرگوار و مادر مهربانم

آن دو فرشته‌ای که از خواسته‌هایشان گذشتند، سختی‌ها را به جان خریدند و خود را سپر بالای مشکلات و ناملایمات کردند تا من به جایگاهی که اکنون در آن ایستاده‌ام برسم.

تقدیر و تشکر

منتّ خدای را عزّوجلّ که طاعتش موجب قربت است و به شکر اندرش مزید نعمت.
هر نفسی که فرو می‌رود ممد حیات است و چون بر می‌آید مفرح ذات. پس در هر نفسی دو
نعمت موجود است و بر هر نعمت شکری واجب.

از استاد گرامی جناب آقای دکتر ابوطالبی برای راهنمایی اینجانب در به انجام
رسانیدن این پایان‌نامه تشکر و قدردانی می‌کنم. همچنین از آقای مهندس سید حمید یزدانی
که من را در تهیه این پایان‌نامه یاری دادند، سپاس‌گزاری می‌کنم.

چکیده

با پیشرفت تکنولوژی در محاسبه‌گرهای همراه و وسایل مخابراتی (همچون گوشی تلفن همراه و PDA (Personal Digital Assistance)) می‌توان از این تجهیزات به عنوان یک آرایه میکروفونی با چیدمان Ad-hoc استفاده نمود. مکان‌یابی مکان‌گوینده یکی از مسایل مهم و مطرح در پردازش گفتار است. استفاده از آرایه میکروفونی Ad-hoc برای مبحث مکان‌یابی گوینده در جاهایی که میکروفون در دسترس نیست و یا زیر ساخت مناسبی فراهم نیست، می‌تواند مفید واقع شود. با این وجود استفاده از آرایه میکروفونی Ad-hoc به دلیل ماهیت Ad-hoc شبکه سبب ایجاد چالش‌های جدید در بحث مکان‌یابی گوینده می‌شود.

در این پژوهش روندی برای محاسبه مکان گوینده نسبت به میکروفون‌ها در چیدمان میکروفونی Ad-hoc با استفاده از ماتریس فاصله بین میکروفون‌ها و فاصله بین گوینده تا میکروفون‌ها در شرایط محیط واقعی پیشنهاد شده است. برای بدست آوردن فاصله بین میکروفونی از دو روش براساس انرژی و مدل همدوسی نویز پخشنده (Diffuse Noise) استفاده شده است. همچنین برای تخمین فاصله بین گوینده تا میکروفون‌ها از دو روش فرم بسته و حل عددی بهره برده شده است. به دلیل مکان‌یابی در شرایط نویزی و پژواکی نیاز بوده تا از روش‌های مقاوم برای تخمین TDOA (Time Difference of Arrival) استفاده شود. روش GCC-PHAT (Generalized Cross-Correlation Phase Transform) با پهنای پنجره بالا (نیم ثانیه) و هم‌چنین روش GCC-PHAT به همراه تفریق طیفی (Spectral Subtraction) برای تخمین TDOA پیشنهاد شده است. شبیه‌سازی‌ها نشان داده است که استفاده از دو روش نامبرده برای تخمین TDOA بار محاسباتی پایینی نسبت به روش GCC-MPHAT دارد و همچنین از دقت قابل قبولی برخوردار است. همچنین از شیوه درون‌یابی برای افزایش وضوح نمودار در روش حل عددی استفاده شده است. شبیه‌سازی‌ها نشان داده است که این شیوه بار محاسباتی را کاهش می‌دهد.

کلمات کلیدی:

مکان‌یابی منبع گفتار، آرایه میکروفونی Ad-hoc، نویز پخشنده، MDS، TDOA، GCC-

PHAT

فهرست مطالب

عنوان	صفحه
فصل اول - پیشگفتار.....	۱
۱-۱- معرفی انواع آرایه‌های میکروفونی	۱
۲-۱- آشنایی با انواع محیط‌های پردازشی	۳
۱-۲-۱- پردازش میدان دور	۳
۲-۲-۱- پردازش میدان نزدیک	۴
۳-۱- انواع روش‌ها در مکان‌یابی منبع صوتی	۵
۱-۳-۱- مکان‌یابی با استفاده از اختلاف زمان رسیدن سیگنال به میکروفون‌ها (TDOA)	۵
۱-۱-۳-۱- روش‌های تخمین TDOA	۶
۲-۳-۱- مکان‌یابی با استفاده از انرژی (SRP-based)	۷
۳-۳-۱- مکان‌یابی با استفاده از روش همبستگی بین سیگنال‌ها	۷
۴-۱- شبکه Ad-hoc	۸
۱-۴-۱- کاربردهای شبکه Ad-hoc	۸
۲-۴-۱- کاربرد شبکه Ad-hoc در مکان‌یابی	۹
۳-۴-۱- انگیزه و چالش کاربرد ساختار Ad-hoc در مکان‌یابی منبع گفتار	۹
۵-۱- خطا در شبکه‌های میکروفون بیسیم با چیدمان Ad-hoc در کنفرانس‌های صوتی	۱۰
۶-۱- طرح کلی پایان‌نامه	۱۲
فصل دوم - کالیبراسیون آرایه میکروفونی.....	۱۴
۱-۲- مکان‌یابی منبع صوتی براساس انرژی دریافتی حسگرهای صوتی	۱۵
۲-۲- انواع روش‌های کالیبراسیون آرایه میکروفونی	۱۷
۳-۲- کالیبراسیون آرایه میکروفونی براساس انرژی	۱۹
۱-۳-۲- الگوریتم اول	۲۰
۱-۱-۳-۲- شبیه‌سازی سیگنال گفتار در محیط واقعی	۲۳
۲-۱-۳-۲- شبیه‌سازی الگوریتم اول	۲۴
۲-۳-۲- الگوریتم دوم	۲۵
۱-۲-۳-۲- آزمایش الگوریتم دوم [۲۳]	۲۸
۳-۳-۲- ارزیابی مقاومت الگوریتم‌ها در برابر خطای سنکرون [۲۳]	۳۱

۳۲	۴-۲- کالیبراسیون آرایه میکروفونی Ad-hoc با استفاده از مدل همدوسی نویز پخشنده
۳۳	۴-۲-۱- روابط کالیبراسیون آرایه میکروفونی از طریق مدل همدوسی نویز پخشنده
۳۷	۴-۲-۵- مقایسه شبیه‌سازی الگوریتم‌های کالیبراسیون از نظر زمان اجرا و خطا در شرایط مختلف محیطی
۳۸	۴-۲-۶- نتیجه‌گیری
۴۰	فصل سوم- تعیین فاصله بین منبع صدا و میکروفون‌ها در آرایه با چیدمان Ad-hoc
۴۲	۳-۱- الگوریتم اول (فرم بسته)
۴۴	۳-۲- الگوریتم دوم (حل عددی)
۴۷	۳-۳- نتایج شبیه‌سازی
۵۱	۳-۴- خلاصه و نتیجه‌گیری
۵۲	فصل چهارم- مکان‌یابی گوینده در آرایه میکروفونی با چیدمان Ad-hoc
۵۳	۴-۱- تعیین TDOA با روش GCC-PHAT + Spectral subtraction
۵۶	۴-۲- تأثیر پهنای پنجره بر روی تخمین TDOA
۵۹	۴-۳- افزایش وضوح نمودار با شیوه درون‌یابی
۶۰	۴-۴- تجمیع اجزای مختلف پیشنهادی برای سیستم مکان‌یابی گوینده در چیدمان میکروفونی Ad-hoc
۶۳	۴-۵- مکان‌یابی منبع گفتار آرایه میکروفونی با چیدمان Ad-hoc با MDS
۶۴	۴-۵-۱- شبیه‌سازی برای تعیین مکان گوینده با روش MDS
۶۶	۴-۶- خلاصه و نتیجه‌گیری
۶۸	فصل پنجم- نتیجه‌گیری و پیشنهادها
۶۹	۵-۱- جمع‌بندی و نتیجه‌گیری
۷۰	۵-۲- فعالیت‌های انجام شده و دستاوردهای این پژوهش
۷۱	۵-۳- پیشنهادهایی برای ادامه کار
۷۳	پیوست‌ها:
۷۳	پیوست آ- معرفی روش MDS
۷۴	پیوست ب- روش‌های دسته‌بندی داده
۷۸	منابع و مراجع

فهرست شکل‌ها

عنوان	صفحه
شکل ۱-۱: نمایش نمونه‌ای از یک آرایه‌ی میکروفونی منظم و دوبعدی (گرفته شده از [۲])	۲
شکل ۱-۲: نمایش یک شبکه Ad-hoc با استفاده از موبایل و لپ‌تاپ (گرفته شده از [۳])	۳
شکل ۱-۳: نمایش ساختار موقعیت منبع گفتار و آرایه میکروفونی در پردازش میدان نزدیک (گرفته شده از [۱])	۴
شکل ۲-۱: نمایش نمودار تضعیف براساس فاصله. محور افقی فاصله برحسب متر و محور عمودی تضعیف برحسب $1/m$ است (گرفته شده از [۶])	۲۱
شکل ۲-۲: نمایش پاسخ ضربه اتاق	۲۴
شکل ۲-۳: نمایش موقعیت صحیح میکروفون و موقعیت میکروفون متناظر تخمین زده شده	۲۵
شکل ۲-۴: نمایش موقعیت صحیح میکروفون و موقعیت میکروفون متناظر تخمین زده شده در حالت سنکرون (واحدها برحسب متر است) (گرفته شده از [۲۳])	۲۹
شکل ۲-۵: نمایش موقعیت واقعی میکروفون و موقعیت تخمین زده شده به همراه موقعیت گوینده متناظر با استفاده از الگوریتم اول (واحدها برحسب متر است) (گرفته شده از [۲۳])	۲۹
شکل ۲-۶: نمایش موقعیت صحیح لپ‌تاپ و موقعیت تخمین زده شده لپ‌تاپ‌های متناظر به همراه موقعیت تخمین زده شده گوینده در الگوریتم دوم و آزمایش دوم (واحدها برحسب سانتی‌متر است) (گرفته شده از [۲۳])	۳۱
شکل ۲-۷: نمایش ارزیابی مقاومت خطای سنکرون روی کانال‌های صوتی. محور افقی میزان شیفت سنکرون و محور عمودی خطای تخمین را نشان می‌دهد (گرفته شده از [۲۳])	۳۲
شکل ۲-۸: تطبیق دادن تابع Sinc روی همدوسی یک فریم از دو سیگنال میکروفونی ضبط شده و در میدان نویز پخشنده؛ فاصله درست ۲۰ cm است و فاصله تخمین زده شده ۱۹.۳ cm است (گرفته شده از [۳۱]).	۳۴
شکل ۲-۹: تطبیق دادن تابع Sinc روی میانگین همدوسی در ۱۰۰ فریم از دو سیگنال ضبط شده در میدان نویز پخشنده؛ فاصله درست ۲۰ cm است و فاصله تخمین زده شده ۱۹.۸ cm است (گرفته شده از [۳۱]).	۳۵
شکل ۲-۱۰: نمایش مکان میکروفون‌های بدست آمده از روش کالیبراسیون به همراه مکان دقیق میکروفون‌ها.	۳۶

- شکل ۱۱-۲: تغییرات خطای کالیبراسیون برحسب تعداد منبع مورد استفاده در مدل‌سازی همدوسی نویز پخشنده ۳۶
- شکل ۱۲-۲: تغییرات خطای کالیبراسیون روش مبتنی بر انرژی برحسب پژواک در سه سطح مختلف نویز سفید ۳۸
- شکل ۱-۳: تغییرات $\frac{G_{0,1}}{G_{0,2}}$ برحسب فاصله (r_0) . (گرفته شده از [۳۲]) ۴۶
- شکل ۲-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۱ در پژواک‌های مختلف و در نویز محیطی ۲۵ dB برای دو روش فرم بسته و حل عددی ۴۸
- شکل ۳-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۲ در پژواک‌های مختلف و در نویز محیطی ۲۵ dB برای دو روش فرم بسته و حل عددی ۴۸
- شکل ۴-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۳ در پژواک از منبع گفتار شماره‌های مختلف و در نویز محیطی ۲۵ dB برای دو روش فرم بسته و حل عددی ۴۹
- شکل ۵-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۱ در سطح نویز مختلف و در پژواک محیطی ۱۰۰ ms برای دو روش فرم بسته و حل عددی ۴۹
- شکل ۶-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۲ در سطح نویز مختلف و در پژواک محیطی ۱۰۰ ms برای دو روش فرم بسته و حل عددی ۵۰
- شکل ۷-۳: نمایش خطای فاصله بدست آمده از منبع گفتار شماره ۳ در سطح نویز مختلف و در پژواک محیطی ۱۰۰ ms برای دو روش فرم بسته و حل عددی ۵۰
- شکل ۱-۴: مقایسه تغییرات دو الگوریتم GCC-MPHAT و GCC-PHAT + SPECTRAL SUB. برحسب SNR در پژواک ۲۵۰ ms ۵۵
- شکل ۲-۴: مقایسه تغییرات دو الگوریتم GCC-MPHAT و GCC-PHAT + Spectral Subtraction. برحسب پژواک در SNR 15 dB ۵۵
- شکل ۳-۴: اثر پهنای پنجره بر روی تخمین TDOA با دو روش GCC-PHAT و GCC-MPHAT ۵۷
- شکل ۴-۴: نمایش خطای بدست آمده از روش GCC-PHAT و GCC-MPHAT در پهنای پنجره نیم و یک ثانیه ۵۷
- شکل ۵-۴: بلوک دیاگرام بدست آوردن مکان گوینده در چیدمان میکروفونی Ad-hoc ۶۲
- شکل ۶-۴: تغییرات خطای مکان‌یابی گوینده برحسب پژواک در SNR 15 dB ۶۵
- شکل ۷-۴: تغییرات خطای مکان‌یابی گوینده برحسب نویز در پژواک 200 ms ۶۶
- شکل آ-۱: نمایش چیدمان میکروفون‌ها و منبع گفتار به همراه مکان واقعی آنها ۷۴

شکل ب-۱: دسته‌بندی داده با روش دسته‌بندی GMM؛ تعداد داده ۵۰ تا است و جواب صحیح ۱
است. ۷۶

شکل ب-۲: دسته‌بندی داده با روش دسته‌بندی K-means؛ تعداد داده ۵۰ تا است و جواب صحیح ۱
است. ۷۷

فهرست جدول‌ها

عنوان	صفحه
جدول ۱-۲ : مقایسه زمان اجرای الگوریتم کالیبراسیون بر پایه انرژی و مدل همدوسی نویز پخشنده	۳۷
جدول ۱-۳: مختصات میکروفون و مختصات منبع گفتار برای اتاق شبیه‌سازی شده (مقادیر برحسب متر هستند).....	۴۷
جدول ۲-۳: فاصله منابع تا نزدیکترین میکروفون به خود (مقادیر برحسب متر هستند).....	۴۷
جدول ۳-۳: زمان اجرا برای الگوریتم فرم بسته و حل عددی.....	۵۱
جدول ۱-۴: زمان اجرای الگوریتم پیشنهادی GCC-PHAT+Spectral subtraction و روش GCC-MPHAT	۵۴
جدول ۲-۴: نمایش خطای بدست آمده از روش GCC-PHAT با پهناى پنجره نیم ثانیه.....	۵۸
جدول ۳-۴: نمایش خطای بدست آمده از روش GCC-PHAT با پهناى پنجره یک ثانیه.....	۵۸
جدول ۴-۴: زمان محاسبه دو روش GCC-MPHAT و GCC-MPHAT.....	۵۹
جدول ۵-۴: فاصله بدست آمده و خطای حاصله در استفاده از شیوه درون‌یابی.....	۶۰
جدول ۶-۴: زمان اجرای الگوریتم حل عددی با استفاده از شیوه درون‌یابی.....	۶۰
جدول ۷-۴: نمایش خطای بدست آمده از تخمین مکان منبع گفتار با روش فرم بسته.....	۶۴
جدول ۸-۴: نمایش خطای بدست آمده از تخمین مکان منبع گفتار با روش حل عددی.....	۶۴
جدول ۹-۴: نمایش تغییرات ماتریس فاصله در تخمین مکان منبع گفتار توسط شیوه MDS و فرم بسته.....	۶۵

حروف اختصاری

علامت اختصاری

عنوان

AWGN: Additive White Gaussian Noise	نویز سفید گوسی جمع پذیر
CC: Cross-Correlation.....	همبستگی متقابل
dB: Decibel.....	واحدی که نسبت بین دو مقدار الکتریسیته یا صوت را بیان میکند
GCC: Generalized Cross-Correlation.....	همبستگی متقابل تعمیم یافته
GMM: Gaussian Mixture Models	مدل ترکیبی گوسی
MDS: Multi Dimensional Scaling	مقیاس چند بعدی
Mic: Microphone.....	میکروفون
ML: Maximum Likelihood.....	بیشترین شانس
MPHAT: Modified PHAse Transform.....	تبدیل فاز تعمیم یافته
ms: mili second.....	میلی ثانیه
m: meter.....	متر
PDA: Personal Digital Assistance.....	کمکیار دیجیتال شخصی
PHAT: PHAse Transform.....	تبدیل فاز
RT60: Reverbration time By 60 dB attenuation.....	زمان پژواک با تضعیف ۶۰ دی بی
Sec: Second	ثانیه
SNR: Signal to Noise Ratio.....	نسبت سیگنال به نویز
TDE: Time Delay Estimation.....	تخمین تأخیر زمان
TDOA: Time Difference of Arrival	اختلاف زمان دریافت

فصل اول

پیشگفتار

هدف از این فصل آشنایی با انواع آرایه‌های میکروفونی^۱، محیط‌های پردازشی و انواع روش‌های مکان‌یابی منبع صوت است؛ همچنین به معرفی آرایه Ad-hoc پرداخته شده و از مسایل و مشکلات کار با این آرایه صحبت می‌نمائیم. روش‌های استفاده شده در این پایان‌نامه نیز بطور خلاصه بیان شده است. در انتهای این فصل، ساختار کلی پایان‌نامه بیان شده است.

۱-۱- معرفی انواع آرایه‌های میکروفونی

آرایه حسگر^۲ به مجموعه‌ای از چند حسگر گفته می‌شود که در مکان‌های مختلف قرار گرفته و برای اندازه‌گیری و بدست آوردن سیگنال دلخواه (که بسته به کاربرد، نوع سیگنال می‌تواند متفاوت باشد) مورد استفاده قرار می‌گیرند. حسگرها به عنوان مبدل^۳، انرژی سیگنال دریافتی را به انرژی الکتریکی تبدیل می‌نمایند، تا بتوان بر روی آنها پردازش‌های دلخواه را انجام داد. خروجی هر حسگر صوتی با میدان موج^۴ در مکان حسگر^۵ متناسب است. این ویژگی این امکان را فراهم می‌آورد که بتوان با استفاده از آرایه به فیلتر نمودن فضای^۶ پرداخت.

آرایه‌های میکروفونی را از نظر نحوه چیدمان می‌توان به دو دسته کلی زیر تقسیم کرد:

^۱ Microphone Array

^۲ Sensor Array

^۳ Converter

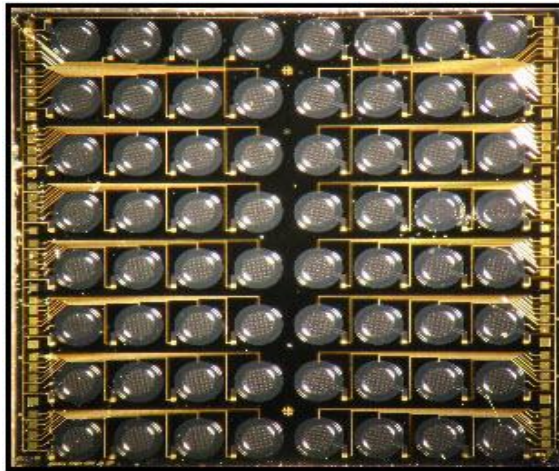
^۴ Wave Field

^۵ Sensor Position

^۶ Spatial Filter

- آرایه میکروفونی منظم و مشخص
- آرایه میکروفونی نامشخص (Ad-hoc)

در آرایه‌های میکروفونی منظم [۱]، مکان میکروفون‌ها مشخص و معلوم است و رابطه مشخص و خاصی بین فواصل آنها وجود دارد. همچنین می‌توان چیدمان مختلفی از میکروفون‌ها را برای کاربردهای گوناگون بکار برد. از چیدمان‌های ممکن می‌توان: چیدمان خطی (تک بعدی)^۱، دو بعدی و دایروی را اشاره کرد.

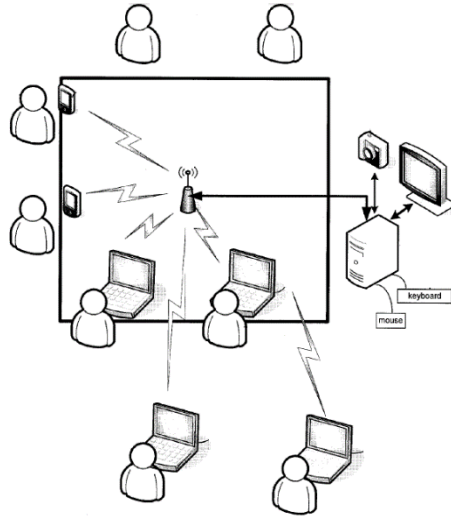


Microscope image of the MEMS microphone array chip (1cm x 1cm)

شکل ۱-۱: نمایش نمونه‌ای از یک آرایه‌ی میکروفونی منظم و دوبعدی (گرفته شده از [۲])

در آرایه‌های میکروفونی نامنظم^۲، برعکس آرایه‌های میکروفونی منظم، مکان حسگرها نامشخص است و رابطه خاصی با یکدیگر ندارند.

^۱ Linear Arrays
^۲ Ad-hoc



شکل ۱-۲: نمایش یک شبکه Ad-hoc با استفاده از موبایل و لپ تاپ (گرفته شده از [۳])

۱-۲-۱- آشنایی با انواع محیط‌های پردازشی

در حالت کلی دو تکنیک مرسوم پردازش میدان دور^۱ و پردازش میدان نزدیک^۲ [۱] در پردازش آرایه‌ای استفاده می‌شود که در زیر به معرفی آن دو می‌پردازیم:

۱-۲-۱-۱- پردازش میدان دور

در این پردازش، فرض بر فاصله زیاد بین آرایه و منبع گفتار است. اگر منابع صوتی به اندازه‌ی کافی از آرایه‌های میکروفونی دور باشد می‌توان این چنین تصور نمود که امواج دریافتی از این منابع، صفحه‌ای (تخت)^۳ است. شرط استفاده از این نوع پردازش در آرایه‌های خطی به صورت زیر است:

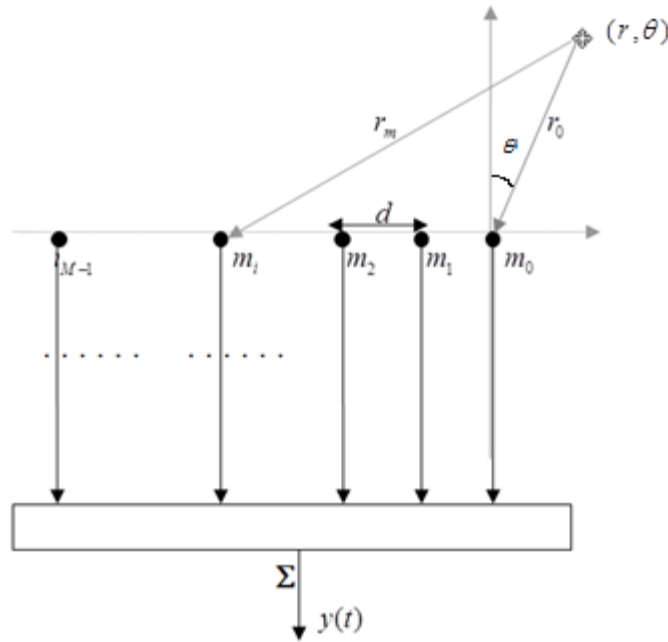
$$r \geq \frac{2D^2}{\lambda_{\min}} \quad (1-1)$$

که در آن r ، فاصله‌ی منبع گفتار تا مرکز آرایه، D طول آرایه و λ_{\min} حداقل طول موج کاری است. به حداقل فاصله مجاز r ، اصطلاحاً فاصله Fraunhofer می‌گویند.

^۱ Far-Field Processing
^۲ Near-Field Processing
^۳ Planet Wave

۱-۲-۲- پردازش میدان نزدیک

در این پردازش که از دقت بالاتری نسبت به پردازش میدان دور برخوردار است، امواج رسیده از منابع صوتی، کروی (غیر تخت) در نظر گرفته می‌شوند. با فرض منبع گفتار در موقعیت (r, θ) باشد، ساختار زیر را می‌توان در نظر گرفت:



شکل ۱-۳: نمایش ساختار موقعیت منبع گفتار و آرایه میکروفونی در پردازش میدان نزدیک (گرفته

شده از [۱])

با دو دیدگاه می‌توان به قضیه نگاه کرد:

چنانچه $x(t)$ سیگنال دریافتی توسط میکروفون مرجع (m_0) و v سرعت صوت در فضای

آزاد باشد، برای سیگنال ضبط شده توسط m امین میکروفون داریم:

$$x_m(t) = \frac{r}{r_m} x\left(t - \left(\frac{r_m - r}{v}\right)\right) \quad (۲-۱)$$

$$r_m = |\bar{r} + m\bar{d}| = \sqrt{(m\bar{d})^2 + r^2 - 2rmd \cos(\theta)} \quad (۳-۱)$$

$$y(t) = \sum_{m=0}^{M-1} \frac{r}{r_m} x\left(t - \left(\frac{r_m - r}{v}\right)\right) \quad (۴-۱)$$

$$\Rightarrow H(f, \theta) = \sum_{m=0}^{M-1} \frac{r}{r_m} \exp(-j2\pi f \left(\frac{r_m - r}{v}\right)) \quad (۵-۱)$$

همانطور که ملاحظه می‌شود برخلاف پردازش میدان دور، در اینجا دامنه سیگنال‌های دریافتی در حسگرهای مختلف، متفاوت می‌باشد.

در دیدگاه دوم با یک دید کلی‌تر و برای هر آرایش دلخواهی از میکروفون‌ها به مسئله نگاه می‌کنیم. فرض کنید میکروفون‌ها در موقعیت مختصاتی (x_m, y_m, z_m) و منبع صوتی مورد نظر در موقعیت مختصاتی (x_s, y_s, z_s) قرار داشته باشند، در این صورت:

$$x_m(t) = \frac{1}{r_m} s\left(t - \frac{r_m}{v}\right) \quad (6-1)$$

$$r_m = \sqrt{(x_s - x_m)^2 + (y_s - y_m)^2 + (z_s - z_m)^2} \quad (7-1)$$

در عبارت فوق فرض بر آن است که دامنه سیگنال، متناسب با معکوس فاصله تضعیف شود.

۱-۳- انواع روش‌ها در مکان‌یابی منبع صوتی

در تعیین مکان منبع گفتار روش‌های متفاوتی وجود دارد که هر یک بسته به شرایط، نیاز و دقت لازم بکار گرفته می‌شود. یک دسته‌بندی مهم و کلاسیک از این روش‌ها به صورت زیر است [۱]:

- مکان‌یابی با استفاده از اختلاف زمان رسیدن سیگنال به میکروفون‌ها (TDOA)^۱
- مکان‌یابی با استفاده از انرژی سیگنال‌ها (SRP)^۲
- مکان‌یابی با استفاده از روش ماتریس همبستگی سیگنال‌ها^۳

۱-۳-۱- مکان‌یابی با استفاده از اختلاف زمان رسیدن سیگنال به میکروفون‌ها

(TDOA)

مکان‌یابی با استفاده از روش اختلاف زمان رسیدن سیگنال به میکروفون‌ها (TDOA) یک شیوه دو مرحله‌ای است: اولین قدم تعیین مقدار TDOA جفت میکروفون‌ها، و قدم دوم پیدا کردن

^۱ Time Difference of Arrival

^۲ Steered Response Power

^۳ Signal Correlation Matrix

تخمین ML^1 برای مکان منبع گفتار از روی اطلاعات TDOA و موقعیت میکروفون‌ها است. هنگامی که TDOA یک جفت میکروفون بدست آید، انتظار می‌رود که منبع گفتار روی یک هذلولی و به کانون موقعیت دو میکروفون، مکان‌یابی شود.

۱-۱-۳-۱ روش‌های تخمین TDOA

TDOA یکی از پارامترهای مهم در مبحث مکان‌یابی منبع گفتار است. TDOA به دلیل اختلاف فاصله میکروفون‌ها تا گوینده ایجاد می‌شود. تخمین TDOA در واقع بدست آوردن تأخیر بین سیگنال‌های دریافتی است.

یکی از روش‌های معمول در بدست آوردن میزان تأخیر بین دو سیگنال، استفاده از همبستگی متقابل دو سیگنال (CC^2) است [۴]. در نسخه تعمیم‌یافته این روش (GCC^3)، این کار در حوزه فرکانس^۴ انجام می‌گیرد تا باعث کاهش بار محاسباتی در آن شود. همچنین در روند محاسبه از وزن‌هایی استفاده می‌شود تا از اثر عوامل مخرب محیطی کاسته شود. این وزن‌ها، بسته به نوع محیط ضبط سیگنال اعمال می‌شود.

چندین مورد از توابع وزنی مورد استفاده به صورت زیر قابل بیان است:

• $GCC-PHAT^5$

• $GCC-ML$

• $GCC-MPHAT^6$ [۴]

زمانی که سیگنال با پژواک^۷ محیط همراه است از تابع وزنی $GCC-PHAT$ برای تخمین TDOA استفاده می‌شود. این تابع وزنی به شدت به نویز همراه با سیگنال اثر منفی نشان داده و در شرایط نویزی جواب خوبی را از آن نمی‌توان بدست آورد.

^۱ Maximum Likelihood

^۲ Cross-Correlation

^۳ Generalized Cross-Correlation

^۴ Frequency Domain

^۵ GCC- PHAse Transform

^۶ GCC-Modified PHAse Transform

^۷ Reverbration

تابع وزنی GCC-ML، هنگامی که سیگنال به نویز محیط مخلوط شده باشد مورد استفاده قرار می‌گیرد. این تابع وزنی نیز در شرایط پژواکی اثری نامطلوب بر روی جواب دارد. از تابع وزنی GCC-MPHAT [۴] در شرایط نویزی-پژواکی قابل استفاده است. با این وجود، بار محاسباتی این تابع وزنی در مقایسه با دو مورد قبل بیشتر می‌باشد.

۱-۳-۲- مکان‌یابی با استفاده از انرژی (SRP-based)

در روش مکان‌یابی با استفاده از انرژی به جستجوی در تمام نقاط فضا و یافتن نقاط با انرژی بیشتر پرداخته می‌شود. این روند تا رسیدن به بیشترین انرژی ادامه می‌یابد. مکان این نقطه به عنوان منبع گفتار در نظر گرفته می‌شود. این کار در عمل، معادل بیشینه سازی توان خروجی شکل‌دهنده پرتوها می‌باشد. این روش به دلیل نیاز به جستجوی زیاد، بار محاسباتی سنگینی را به همراه دارد.

۱-۳-۳- مکان‌یابی با استفاده از روش همبستگی بین سیگنال‌ها

در این روش با استفاده از تجزیه و تحلیل روی ماتریس همبستگی^۱ و مقادیر ویژه^۲ بین سیگنال‌ها انجام می‌گیرد و تخمین طیف با تفکیک‌کنندگی بالا^۳ در مکان‌یابی بدست می‌آید. این روش‌ها بار محاسباتی به مراتب کمتری نسبت به روش انرژی دارد. همچنین این روش‌ها تخمین بهتری از مکان منبع را بدست می‌دهد ولی استفاده از این روش‌ها برای پردازش سیگنال گفتار در محیط اکوستیکی با محدودیت‌هایی همچون ایستادن نبودن سیگنال (گفتار و نویز) و شرط مستقل نبودن سیگنال از نویز محیط مواجه هست [۵].

^۱ Correlation Matrix

^۲ Eigenvalues

^۳ Maximum Resolution

۴-۱- شبکه Ad-hoc

یک شبکه Ad-hoc، نوع خاصی از شبکه‌های بدون زیر ساخت است. لغت Ad-hoc در اصطلاح، کلمه‌ای با ریشه لاتین و به معنای "به این منظور" بوده که مفهوم آن موقتی بودن است [۳]. علت این نامگذاری آن است که این شبکه‌ها بطور موقتی ایجاد می‌شوند؛ بخصوص برای مکان‌هایی که زیرساخت ثابتی وجود ندارد و یا نیاز به ایجاد سریع شبکه است (مثل میدان جنگ). در اینگونه شبکه‌ها، همبندی شبکه دائماً در حال تغییر بوده و علت این تغییر، سیار بودن گره‌های شبکه است. علاوه بر این، گره‌های جدیدی ممکن است در هر لحظه به شبکه اضافه و یا گره‌هایی از شبکه حذف شوند و یا این که بعضی از گره‌ها خود را به حالت خاموش درآورند. هیچ کنترل مرکزی ثابتی برای پشتیبانی پیکربندی شبکه و یا باز پیکربندی شبکه وجود ندارد. شبکه‌های Ad-hoc که برای اولین بار توسط وزارت دفاع آمریکا در سیستم‌های نظامی و عملیاتی مورد استفاده قرار گرفته بود، از سال ۱۹۷۰ بطور عمومی مورد بهره برداری واقع شد.

۴-۱-۱- کاربردهای شبکه Ad-hoc

به طور کلی زمانی که زیرساخت ثابتی نداریم و ایجاد و نصب زیرساخت غیر عملی بوده و یا مقرون به صرفه نباشد، استفاده از شبکه Ad-hoc مفید می‌نماید. از جمله این کاربردها می‌توان به موارد ذیل اشاره نمود:

شبکه‌های شخصی

شامل تلفن‌های سلولی، کامپیوترهای کیفی و ساعت‌های مچی.

محیط‌های نظامی

این شبکه‌ها به دلیل قابلیت گسترش سریع و خودسازمان بودن در میدان جنگ کاربرد وسیعی دارد. همچنین در هنگام جنگ برای دسترسی سربازان به یکدیگر یا دسترسی ادوات نظامی به یکدیگر، در مناطقی که هیچ زیرساختی وجود ندارد، مورد استفاده دارد و از این طریق می‌توان سربازان در حال حرکت را پشتیبانی کرد.