بسم الله الرحمن الرحيم

دانشگاه یزد
دانشکده مهندسی برق وکامپیوتر
گروه مهندسی مخابرات

پایان‌نامه
برای دریافت درجه کارشناسی ارشد
مهندسی برق- مخابرات

مکان‌یابی چند گوینده همزمان

استاد راهنما:دکتر حمیدرضاابوطالبی
استاد مشاور: دکتر وحید ابوطالبی
پژوهش و نگارش: هدیه هلی

مهرماه ۱۳۹۰

تقدیم به

پدر و مادر عزیزم

والدینی که بودنشان تاج افتخاریست بر سرم و نامشان دلیلی‌است بر بودنم،

چرا که این دو وجود پس از پروردگار مایه هستی‌ام بوده‌اند

این پایان‌نامه با حمایت‌های

مالی مرکز تحقیقات مخابرات

ایران به انجام رسیده‌است.

تقدیر و تشکر

نفس می نیارم زد از شکر دوست که شکری ندانم که در خورد اوست

آنچه در پیش رویتان گشوده می‌شود، حاصل راهنمایی و کمک اساتید و عزیزانی است که بدون یاریشان این مهم ممکن نمی‌گشت.

از جناب آقای دکتر حمیدرضا ابوطالبی به جهت قبول راهنمایی این پژوهش و زحمات بی‌دریغ و پر شمار ایشان و جناب آقای دکتر وحید ابوطالبی به جهت راهنمایی‌ها و کمک‌های ارزشمندشان کمال تشکر را می‌نمایم. در نهایت از حمایت‌های مادی و معنوی "مرکز تحقیقات مخابرات ایران" در راستای انجام این پژوهش قدردانی می‌نمایم.

چکیده

مکان‌یابی منبع گفتار در بسیاری از سیستم‌های پردازش گفتار از قبیل سیستم‌های تله‌کنفرانس، سیستم‌های بازشناسی صوت، اتاق‌های کنفرانس هوشمند و سیستم‌های کمک شنوایی کاربرد دارد. امروزه مکان‌یابی منبع صوت با کمک آرایه‌های میکروفونی از موضوعات مورد توجه می‌باشد. این آرایه‌ها در نهایت سیگنال دارای کیفیت مطلوب را از منبع صوتی تهیه می‌کنند و منابع صوتی غیر دلخواه را حذف می‌کنند. اما یکی از چالش‌های پیش رو در کاربردهای ذکر شده، بحث مکان‌یابی چند گوینده همزمان است. در این پایان‌نامه به بررسی یکی از روش‌های مکان‌یابی چندین گوینده همزمان که بر اساس الگوریتم‌های جداسازی کور منابع (BSS) است، می‌پردازیم. BSS را می‌توان به صورت مجموعه‌ای از شکل‌دهنده‌های پرتو صفر و تطبیقی کور درنظر گرفت. در حقیقت در حالی‌که برای یافتن مکان منابع، به اطلاعاتی در مورد مکان منابع نیاز داریم تا شکل‌دهنده پرتو را به سمت آن‌ها هدایت کنیم، BSS بدون داشتن این اطلاعات اولیه، امکان بازیابی سیگنال‌ها را از سیگنال‌های ضبط‌شده در میکروفون‌ها -که ترکیبی از سیگنال‌های همزمان هستند- فراهم می‌کند. این به آن معناست که فیلترهای جداکننده BSS حاوی اطلاعات مهمی در مورد مکان منابع هستند. در نهایت با توجه به فیلترهای جداکننده حاصل از الگوریتم BSS، می‌توان الگوی تشعشعی بدست آورد که با یافتن مینیمم‌ها در این الگو، قادر به تخمین مکان منابع هستیم. در این پایان‌نامه به‌منظور بهبود عملکرد الگوریتم مکان‌یابی در محیط‌های انعکاسی، روشی پیشنهاد کردیم که در آن تنها بزرگترین ضریب فیلترهای جداکننده که متناظر با مسیر مستقیم است را درنظر گرفته و الگوی تشعشعی را محاسبه می‌کنیم. همچنین برای کاهش بار محاسباتی، محدوده فرکانسی موردنظر برای فرآیند متوسط‌گیری را کاهش می‌دهیم. نتایج شبیه‌سازی گویای دقت بالاتر الگوریتم پیشنهادی در تخمین مکان منابع همزمان نسبت به روش‌های مشابه می‌باشد.

فهرست مطالب

فصل اول : مقدمه ۱
۱-۱- مکان‌یابی منابع صوتی به کمک آرایه‌های میکروفونی ۱
۱-۲- اتاق کنفرانس هوشمند ۲
۱-۳- سابقه تحقیق ۴
۱-۴- ساختار پایان‌نامه ۶
فصل دوم : معرفی و جداسازی کور منابع ۷
۲-۱- مقدمه ۷
۲-۲- مدل بکار رفته برای مساله جداسازی کور منابع ۷
۲-۲-۱- مدل ترکیب لحظه‌ای ۸
۲-۲-۲- مدل ترکیب کانولوتیو ۹
۲-۳- کاربردهای جداسازی کور منابع ۱۱
۲-۳-۱- کاربردهای جداسازی کور منابع در مهندسی پزشکی ۱۲
۲-۳-۲- کاربردهای مخابراتی ۱۳
۲-۳-۳- کاربردهای جداسازی کور منابع برای سیگنال‌های صوتی و موزیک ۱۳
۲-۴- موارد ابهام در جداسازی کور منابع به صورت لحظه‌ای و کانولوتیو ۱۴
۲-۵- حل BSS ایده‌آل ۱۵
فصل سوم : روش‌های مکان‌یابی چندین گوینده همزمان بر اساس جداسازی کور منابع ..۱۷
۳-۱- مقدمه ۱۷
۳-۲- تخمین TDOA با استفاده از ضرایب فیلترهای BSS ۱۸
۳-۲-۱- مرحله اول : تخمین TDOA ۱۸
۳-۲-۲- مرحله دوم : تخمین مکان منابع ۲۲
۳-۳- رابطه بین BSS حوزه فرکانس با شکل‌دهی وفقی پرتو در حوزه فرکانس برای مخلوط کانولوتیو ۲۳
۳-۳-۱- BSS حوزه فرکانس برای مخلوط کانولوتیو ۲۴
۳-۳-۲- ABF حوزه فرکانس ۲۶
۳-۳-۲-۱- ABF برای سیگنال هدف S_1 سیگنال مزاحم S_2 ..... ۲۷
۳-۳-۳- ABF برای سیگنال هدف S2 سیگنال مزاحم S1 ..... ۲۹
۳-۴- مکان‌یابی با استفاده از BSS-ADF ۳۴
۳-۴-۱- الگوهای تشعشعی جهت‌دهی‌شده BSS ۳۴
۳-۴-۲- محاسبه الگوی جهت‌دهی‌شده متوسط ۳۵
۳-۴-۳- مدل‌های میدان- نزدیک و میدان-دور ۳۶

۳-۴-۳-۱- مدل انتشار میدان- نزدیک ۳۷
۳-۴-۳-۲- فرض میدان- دور ۳۷
فصل چهارم : پیاده سازی و ارزیابی روش مکان‌یابی چند گوینده مبتنی بر الگوریتم BSS و ارایه راهکار جهت بهبود این روش ۳۹
۴-۱- مقدمه ۳۹
۴-۲- تخمین TDOA با استفاده از روش‌های مبتنی بر BSS..... ۳۹
۴-۳- تخمین مکان منابع به کمک الگوی تشعشعی هدایت شده در محیط بدون انعکاس۴۳
۴-۴- تخمین مکان منابع به کمک الگوی تشعشعی هدایت شده در محیط انعکاسی ۴۹
۴-۵- روش پیشنهادی جهت بهبود دقت مکان‌یابی در محیط‌های انعکاسی ۵۷
۴-۵-۱- شرح تئوری روش پیشنهادی..... ۵۷
۴-۵-۲- نتایج شبیه‌سازی روش پیشنهادی ۵۸
۴-۶- تغییر پیشنهادی جهت کاهش حجم محاسبات ۶۴
فصل پنجم: جمع بندی و نتیجه‌گیری ۶۶
۵-۱- کارهای انجام شده در این پایان نامه ۶۶
۵-۲- نوآوری‌ها و دستاوردهای این تحقیق ۶۶
۵-۳- پیشنهادهایی برای ادامه کار ۶۷
مراجع ۶۸

فهرست جداول

جدول ۴-۱- نتایج حاصل از تخمین TDOA برای حالت دو منبع و دو میکروفون ۴۳
جدول ۴-۲- نتایج حاصل از مکان‌یابی دو منبع همزمان فعال ۵۴
جدول ۴-۳- نتایج حاصل از مکان‌یابی سه منبع همزمان فعال ۵۷
جدول ۴-۴- خطای ناشی از مکان‌یابی توسط دو روش BSS-ADP-DP برای دو منبع همزمان فعال ۶۱
جدول ۴-۵- میانگین خطای حاصل از دو روش BSS-ADP-DP برای اتاقی با ابعاد ۶*۹*۵ متر ۶۱
جدول ۴-۶- میانگین خطای حاصل از دو روش BSS-ADP-DP برای اتاقی با ابعاد ۵*۴*۳ متر ۶۴
جدول ۴-۷- میانگین خطای حاصل از روش BSS-ADP-DP برای f_{max} های متفاوت ۶۵

فهرست اشکال

شکل ۲-۱- مدل BSS ۱۰
شکل ۳-۱- مدل SIMO ۲۰
شکل ۳-۲- مدل MIMO ۲۰
شکل ۳-۳- دو مجموعه از ساختار ABF ۲۷
شکل ۳-۴- مسیرهای معادله ۳۰
شکل ۳-۵- موقعیت هندسی منبع و میکروفون نسبت به مرکز آرایه ۳۵
شکل ۴-۱- شمای قرارگیری منابع و میکروفون‌ها ۴۰
شکل ۴-۲- بردارهای جدا کننده حاصل از الگوریتم BSS ایده‌آل برای زمان انعکاس ۴۱۵۰۰ms
شکل ۴-۳- الگوی تشعشعی هر یک از خروجی‌ها در سه فرکانس ۴۵
شکل ۴-۴- الگوی تشعشعی متوسط‌گیری شده برای دو منبع مورد نظر واقع در ۲۰- درجه و۳۵ درجه ۴۶
شکل ۴-۵- الگوی تشعشعی خروجی ۴۸
شکل ۴-۶- الگوی تشعشعی متوسط‌گیری شده برای ۶ منبع همزمان فعال ۴۹
شکل۴-۷- بردارهای جدا کننده حاصل از الگوریتم BSS ایده‌آل برای زمان انعکاس ۱۵۰ms ... ۵۱
شکل ۴-۸- الگوی تشعشعی متوسط‌گیری شده برای زمان انعکاس ۲۰۰ms ۵۲
شکل ۴-۹- الگوی تشعشعی متوسط‌گیری شده برای دو منبع همزمان فعال ۵۳
شکل ۴-۱۰- الگوی تشعشعی متوسط‌گیری شده برای سه منبع همزمان فعال ۵۶
شکل ۴-۱۱- الگوی تشعشعی متوسط‌گیری شده برای دو منبع همزمان فعال ۶۰
شکل ۴-۱۲- الگوی تشعشعی متوسط‌گیری شده برای دو منبع همزمان فعال ۶۳

فصل اول

مقدمه

۱-۱- مکان‌یابی منابع صوتی به کمک آرایه‌های میکروفونی

مکان‌یابی در حقیقت مشخص کردن مکان گوینده‌های فعال مختلف در محیط مورد نظر می‌باشد. مکان‌یابی منبع گفتار در کاربردهای متعددی از قبیل اتاق‌های کنفرانس هوشمند، سیستم‌های بازشناسی صوت و سیستم‌های کمک شنوایی استفاده می‌شود. در زمینه مکان‌یابی تک‌گوینده به کمک آرایه‌های میکروفونی تاکنون تحقیقات وسیعی صورت گرفته و روش‌های مختلفی برای آن ارائه شده است. این روش‌ها را می‌توان به سه دسته کلی تقسیم‌بندی کرد:

- در روش‌های دسته اول توان خروجی آرایه ماکزیمم می‌شود.
- روش‌های دسته دوم با پردازش ماتریس همبستگی سیگنال‌های دریافتی توسط آرایه مکان منبع را بدست می‌آورد.
- در روش‌های دسته سوم با استفاده از اطلاعات تفاوت زمان رسیدن سیگنال[1] (TDOA) به هر میکروفون مکان منبع تخمین زده می‌شود.

اصول حاکم بر این روش‌ها در مرجع [۱] بررسی شده است.

در برخی کاربردها از جمله اتاق کنفرانس هوشمند و سیستم‌های کمک شنوایی، یکی از چالش‌های پیش‌رو بحث مکان‌یابی چندین گوینده همزمان است. به عنوان مثال در حین مکالمه در یک اتاق کنفرانس، ممکن است دو یا چند نفر همزمان شروع به صحبت کردن کنند. از آنجا که هدف ما ضبط و ثبت مکالمات و صحبت‌های تمامی اعضای حاضر در نشست است، بایستی این سیگنال‌های همزمان با حداکثر کیفیت ضبط شوند تا سیستم‌های بهسازی و بازشناسی گفتار قادر

---

[1]- Time Difference Of Arrival

به انجام سایر پردازش‌ها با کمترین خطا بوده و در نهایت بتوانیم بهترین خروجی را از سیستم دریافت کنیم.

۱-۲- اتاق کنفرانس هوشمند[1]

همانطور که در بخش قبل ذکر شد، یکی از کاربردهای مکان‌یابی گوینده در بحث اتاق کنفرانس هوشمند است. با بهبود سیستم‌های مخابراتی و کوچکتر شدن و سریعترشدن آن‌ها، هدف ایجاد اتاق‌های کنفرانس هوشمند به واقعیت نزدیک‌تر می‌شود. هدف از اتاق کنفرانس هوشمند، مهیا کردن سرویس‌هایی است که به ارتباط مستقیم انسان و کامپیوتر نیازی ندارد. بلکه با استفاده از آرایه‌های میکروفونی و دوربین‌ها، اطلاعات اتاق دریافت شده، آنالیزهای لازم انجام می‌شود و در نهایت به افراد این امکان داده می‌شود که فقط بر روی اهدافشان متمرکز شوند.

در اتاق کنفرانس هوشمند به چند سوال باید پاسخ داده شود.

۱- چه کسی سخن می‌گوید؟

۲- چه می‌گوید؟

۳- چگونه و با چه حالتی سخن می‌گوید؟ شاد، ناراحت، عصبی و...

۴- گوینده کجا قرار دارد و یا چه کسی در حال سخن گفتن است؟ (شامل مکان‌یابی و دنبال کردن گوینده)

برای رسیدن به جواب سؤالات فوق و همچنین دستیابی به هدف نهایی که تهیه صورت‌جلسه مکتوب در این اتاق‌هاست، سه مرحله اساسی طی می‌شود:

- مکان‌یابی[2] و دنبال کردن[3] گوینده
- بهسازی گفتار[4] که در این مرحله عملیات حذف نویز و پژواک برای بهبود کیفیت صوت ضبط شده، صورت می‌گیرد و سیگنال برای ورود به مرحله نهایی آماده می‌شود.

---

[1]- Smart Meeting Room
[2]- Localization
[3]- Tracking
[4]- Speech enhancement

- بازشناسی گفتار[1] که در این قسمت سیگنال صوتی ضبط شده و ارتقا یافته در مراحل قبل، به متن تبدیل می‌شود.

مکان‌یابی و دنبال‌کردن گوینده نقش مهمی را در پردازش‌های اتاق کنفرانس ایفا می‌کنند. مکان گوینده برای انتخاب و جهت‌دهی دوربین برای فیلم‌برداری و همچنین برای جهت‌دهی بیم یک شکل‌دهنده پرتو[2] آرایه میکروفونی [۱] برای بهبود کیفیت گفتار ورودی به یک سیستم بازشناسی گفتار نقش تعیین کننده دارد. همچنین از اطلاعات مکان گوینده می‌توان در جهت شناسایی یا تشخیص هویت گوینده (با استفاده از اطلاعات صوتی و تصویری) استفاده کرد.

مکان‌یابی می‌تواند در حالت میدان نزدیک[3] [۲] انجام شود که در این حالت به دنبال یافتن مکان دقیق منبع هستیم و یا در حالت میدان دور[4] [۲] انجام شود که در این صورت مکان دقیق منبع مورد نظر نبوده، بلکه کافی است تا جهت رسیدن سیگنال‌ها[5] (DOA) تخمین زده شود.

همانطور که پیش از این هم گفته شد، برای دستیابی به بهترین سیگنال ضبط‌شده، قدم اول مکان‌یابی گوینده است. بحث مکان‌یابی چند گوینده همزمان، موضوعی است که در چند سال اخیر به آن توجه شده است. برای حل این مسأله دو رویکرد اساسی وجود دارد:

- در رویکرد اول از یک شکل دهنده پرتو برای یافتن چندین قله یک تابع انرژی مانند توان پاسخ هدایت‌شده[6] استفاده می‌شود [۳و۴].
- در رویکرد دوم ابتدا TDOA به وسیله آرایه میکروفونی تخمین زده شده، سپس مکان منابع با استفاده از تکنیک‌های دسته‌بندی مناسب تخمین زده می‌شود [۵و۶]. در حالتی که TDOA تخمین زده می‌شود، تأخیرهای نسبی سیگنال‌ها برای جفت میکروفون‌های مختلف تخمین زده می‌شود. تعداد TDOAهای تخمین زده شده به تعداد منابع و تعداد

---

[1]- Speech recognition
[2]- Beamformer
[3]- near field
[4]- far field
[5]- Direction of arrival
[6]- Steered Response Power

بعد مورد نظر بستگی دارد. پس از انجام مرحله اول، TDOAهای تخمین زده شده برای محاسبه مکان منابع در فضای سه بعدی و یا دو بعدی مورد استفاده قرار می‌گیرند.

در شرایطی که در محیط نویز و انعکاس وجود داشته باشد، تخمین TDOA تحت تأثیر این عوامل مخرب قرار گرفته و در نتیجه در تخمین مکان منابع دچار خطا می شویم. اما روش‌های دسته اول نسبت به وجود نویز و انعکاس مقاوم‌ترند. در مقابل حجم محاسباتی روش‌های دسته اول در مقایسه با دسته دوم بیشتر است. در ادامه، به چند روش از رویکردهای فوق اشاره خواهیم کرد.

۱-۳- سابقه تحقیق

عمده کارهای اساسی در زمینه مکان‌یابی تک گوینده، حاصل تحقیقات چندین مرکز فعال در این حوزه است. از جمله این مراکز می‌توان به سه گروه آزمایشگاه تحقیقاتی LEMS دانشگاه Brown [۷] ، مرکز تحقیقاتی WATRI و مرکز تحقیقاتی APL [۸] اشاره کرد. اما از آنجاکه مکان‌یابی چندین گوینده همزمان، موضوعی است که در چند سال اخیر به آن توجه شده، تحقیقات متمرکز و منسجمی بر روی آن صورت نگرفته است. اما به طور پراکنده روش‌های مختلفی برای این منظور توسط افراد مختلف ارائه شده که به چند مورد از آن‌ها اشاره می‌کنیم.

مکان‌یابی و جداسازی چندین منبع سیگنال با استفاده از روش Histogram Mapping در سال ۱۹۹۵ معرفی شد [۹] و در سال ۲۰۰۹ مقاله‌ای منتشر شد که در آن این روش بهبود داده شد [۱۰] . در این روش سیگنال ورودی به هر میکروفون ابتدا با استفاده از تبدیل فوریه کوتاه‌مدت (STFT)[1] به اجزاء فرکانسی تجزیه می‌شود. وقتی چندین گوینده همزمان درحال صحبت کردن هستند، هر میکروفون مخلوطی از چندین سیگنال صحبت را دریافت می‌کند. به دلیل وجود تفاوت‌های طیفی بین صدای افراد مختلف، می‌توان این فرض را درنظر گرفت که یک سلول فرکانسی از سیگنال ورودی در هر فریم زمانی، مؤلفه طیفی تنها یک سیگنال گفتار را شامل می‌شود. در واقع سیگنال‌های صحبت متفاوت در حوزه زمان-فرکانس، به ندرت با هم همپوشانی دارند. پس از آن TDOA بین هر جفت از میکروفون‌ها و به تبع آن DOA هر یک از منابع، محاسبه شده و برای هر مؤلفه فرکانسی، هیستوگرامی رسم می‌شود که در آن محور افقی شامل DOA و

[1]- Short Time Fourier Transform

محور عمودی فراوانی آنها را نشان می‌دهد. با میانگین‌گیری از هیستوگرام‌های بدست آمده، نهایتاً به نموداری دست خواهیم یافت که وجود هر قلّه در آن به منزله وجود یک منبع بوده و DOA متناظر با آن قلّه، DOA مربوط به یکی از منابع است. از محاسن این روش بار محاسباتی پایین آن است، اما از اشکالات این روش محدود بودن نوع آرایه میکروفونی، عدم درنظر گرفتن پژواک محیط و دقت پایین آن است.

از جمله الگوریتم‌های دیگر می‌توان به آنالیز دو مرحله‌ای نقشه‌های صوتی اشاره کرد. پیش از این از نقشه‌های صوتی برای مکان‌یابی تک‌گوینده استفاده می‌شد [۱۱] . اما در سال ۲۰۰۸ این روش به یافتن مکان دو گوینده همزمان تعمیم داده شد [۱۲] . در این دسته از روش‌ها از معیارهایی مانند همبستگی متقابل تعمیم یافته GCC[1] که از محبوبترین روش‌های تخمین TDOA است، استفاده می‌شود. اما از آنجاکه این معیارها بر اساس مدل انتشاری ایده‌آل هستند، وجود پژواک در محیط بر عملکرد این دسته از روش‌ها تاثیر منفی داشته و دقت آن را تا حدود زیادی کاهش می‌دهد.

از روش‌های دیگری که برای این موضوع بکار رفته است، می‌توان به استفاده از تکنیک‌های زیرفضای سیگنال و ESPRIT[2] [۱۳]، بکارگیری الگوریتم MUSIC[3] برای تخمین TDOA در مرحله اول و سپس استفاده از تکنیک‌های دسته‌بندی [۵] و ... اشاره کرد. باید توجه داشت که در این دسته از روش‌ها نیز از مدل انتشاری تک مسیری استفاده شده و بنابراین برای محیط‌های واقعی چندان مناسب به نظر نمی‌رسد. ضمن اینکه حداکثر تعداد منابعی هم که می‌توان آن‌ها را آشکارسازی کرد، به تعداد میکروفون‌ها بستگی دارد.

روشی که در این پایان نامه قصد داریم آن را بررسی کنیم، روشی مبتنی بر جداسازی کور منابع (BSS) می‌باشد که در آن اطلاعات مربوط به مکان، از فیلترهای جداکننده[4] حاصل از الگوریتم جداسازی کور منابع استخراج می‌شود. مقالات مربوط به این روش اکثراً توسط A.

---

[1]- Generalized Cross Correlation
[2]- Estimation of Signal Parameters via Rotational Invariance Techniques
[3]- Multiple Signal Classification
[4]- demixing filters

Lombard و W. Kellermann و H. Buchner [۱۴-۱۵و۱۶] ارائه شده‌اند. اما مقالاتی هم توسط M.Ikram در سال ۲۰۰۲ [۱۷] و H. Sawada در سال ۲۰۰۵ [۱۸] ارائه شده که در آن‌ها به حل مشکل ابهام جایگشت توجه شده است و اطلاعات مکان در هر زیر طیف فرکانسی[1] بطور جداگانه محاسبه می‌شود. در سال ۲۰۰۹ با بهبود روش BSS با استفاده از متوسط‌گیری الگوهای تشعشعی جهت‌دهی شده[2] (BSS-ADP) روی همه زیر طیف‌های فرکانسی، این روش برای هر آرایه دلخواه مورد استفاده قرار گرفت [۱۹]. از جمله مزیت‌های این روش نسبت به سایر روش‌ها این است که قابلیت آشکارسازی تا بیش از ۶ گوینده فعال را دارا بوده و در محیط‌های با نویز و انعکاس بالا هم عملکرد بسیار مناسبی دارد. علاوه بر آن می‌توان این روش را برای هر آرایه میکروفونی دلخواه بکار گرفت.

۱-۴- ساختار پایان نامه

در این پایان نامه به دلیل مزایایی که برای روش مکان‌یابی بر اساس BSS ذکر شد، این روش را مورد مطالعه قرار داده و به بررسی و تحلیل آن خواهیم پرداخت.

در فصل دوم، مقدمه‌ای بر روش‌های جداسازی کور منابع ذکرخواهد شد.

در فصل سوم، از نتایج حاصل از جداسازی کور منابع برای تخمین مکان منابع استفاده شده و دو روش موجود برای تخمین مکان منابع را معرفی می‌نماییم. ابتدا مکان‌یابی بر اساس تخمین TDOA بررسی شده و در ادامه با بیان رابطه بین BSS و شکل‌دهنده پرتو، روش مکان‌یابی با استفاده از الگوی تشعشعی جهت‌دهی شده[3] را معرفی خواهیم کرد.

در فصل چهارم، به بررسی و ارزیابی نتایج حاصل از شبیه‌سازی‌ها پرداخته و روشی را به‌منظور بهبود دقت مکان‌یابی ارایه خواهیم داد.

در نهایت در فصل پنجم، به جمع‌بندی مطالب و موضوعات مورد بحث در این پایان نامه و برشماری دستاوردهای آن پرداخته و پیشنهادهایی را برای ادامه کار مطرح خواهیم کرد.

---

[1]- frequency bin
[2]- BSS Averaged Directivity Pattern
[3] - Directivity pattern

فصل دوم

معرفی جداسازی کور منابع

۲-۱- مقدمه

در سال‌های اخیر، کارهای بسیاری در زمینه جداسازی منابع (به خصوص با الگوریتم‌های کور) انجام شده‌است. در جداسازی کور منابع، هم منابع و هم فرآیند ترکیب آنها ناشناخته است و تنها ترکیبی از منابع در دسترس است. در اکثر موارد، هدف الگوریتم جداسازی کور منابع، بازیابی منابع از روی ترکیب ضبط شده آن‌ها یا حداقل شناسایی سیستم ترکیب‌کننده[۱] است. مدل به کار رفته در الگوریتم‌های جداسازی، مدل ترکیب لحظه‌ای[۲] و یا مدل کانولوتیو[۳] است. در کارهای اخیر در مقالات بیشتر از مدل کانولوتیو استفاده شده است. چراکه این مدل با حالات واقعی که انعکاس در محیط وجود دارد، انطباق بیشتری دارد. در ادامه به برخی از کاربردهای جداسازی کور منابع اشاره می‌شود. همچنین مدل ترکیب لحظه‌ای و کانولوتیو مورد استفاده در جداسازی کور منابع و موارد ابهامی در مسئله جداسازی کور منابع بررسی خواهد شد.

۲-۲- مدل بکار رفته برای مسأله جداسازی کور منابع

در بخش‌های زیر ابتدا مدل‌های مختلف سیستم ترکیب که در کاربردهای مختلف جداسازی کور منابع داریم، معرفی می‌شوند.

---

[1]- mixing system
[2]- instantaneous mixture
[3]- convolutive mixture

۲-۲-۱- مدل ترکیب لحظه‌ای

ساده‌ترین نوع مسأله جداسازی کور منابع، مربوط به حالتی است که مدل لحظه‌ای را برای سیستم ترکیب داریم که در آن هیچ مؤلفه تأخیر یافته‌ای از سیگنال‌های منبع وجود ندارد. این مدل را می‌توان به صورت زیر بیان کرد:

$$x_p(n) = \sum_{q=1}^{N} h_{qp} s_q(n) + n_p(n) \qquad (۳-۲)$$

که در آن $x_p(n)$ برای $p=1,\ldots,m$، سیگنال‌های دریافت‌شده توسط میکروفون‌ها است که از ترکیب خطی مولفه‌های منبع مستقل $s_q(n)$ برای $q=1,\ldots,N$ ایجاد شده‌اند. در رابطه فوق، h_{qp} بیانگر ضرایب وزنی اسکالر از منبع qام به میکروفون pام است. $n_p(n)$ مولفه نویز در هر میکروفون و n اندیس زمانی گسسته است. با توجه به مدل لحظه‌ای بیان شده در رابطه (۲-۳)، در مسأله جداسازی کور منابع به دنبال یافتن ضرایب مناسبی برای سیستم جداساز هستیم که با اعمال آن به سیگنال‌های مشاهده شده در میکروفون‌ها، سیگنال‌های خروجی حاصل، تخمین مناسبی از سیگنال‌های منابع اصلی را بدست بدهد:

$$y_p(n) = \sum_{q=1}^{m} w_{qp} x_q(n) \qquad (۴-۲)$$

که در آن $y_p(n)$ سیگنال جداشده در خروجی و w_{qp} ضرایب وزنی از qامین میکروفون به pامین خروجی است.

در کاربردهایی از جداسازی کور منابع مانند کاربرد آن در شناسایی مولفه‌های فعالیت مغز از روی سیگنال‌های EEG -که در آن از جداسازی کور منابع برای شناسایی فاکتورهای پنهان از روی سری‌های زمانی موازی استفاده می‌شود- از مدل لحظه‌ای استفاده شده است. از دیگر زمینه‌های کاربردی جداسازی کور منابع که در آن از مدل لحظه‌ای استفاده می‌شود، می‌توان به استخراج ویژگی‌های مستقل در پردازش تصویر و بهبود کیفیت تصویر اشاره کرد. در این پایان‌نامه با محیط‌های آکوستیک سروکار داریم. از آنجا که در مدل لحظه‌ای، مولفه‌های تأخیریافته سیگنال‌های منبع اصلی درنظر گرفته نمی‌شود، این مدل برای محیط‌های آکوستیکی مدل مناسبی نیست. از این رو در بخش بعدی، مدل کانولوتیو معرفی می‌شود.

۲-۲-۲- مدل ترکیب کانولوتیو

با تعمیم مدل ترکیب لحظه‌ای و درنظر گرفتن مولفه‌های تاخیر یافته سیگنال‌های منبع $s_q(n)$، مدل کانولوتیو را برای سیستم ترکیبی خواهیم داشت که این مدل، شامل پاسخ ضربه‌های محدود در زمان (FIR) به جای ضرایب اسکالر است. به این ترتیب با درنظر گرفتن فیلترهای FIR با M نمونه زمانی، مدل کانولوتیو را برای سیستم ترکیبی خواهیم داشت:

$$x_p(n) = \sum_{q=1}^{N} \sum_{k=0}^{M-1} h_{qp}(k) s_q(n-k) + n_p(n) \qquad (۵-۲)$$

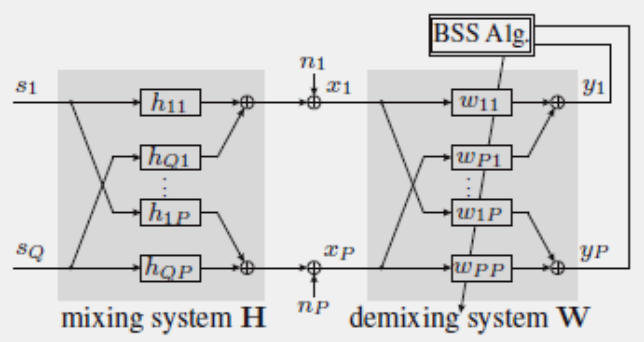
که در آن $h_{qp}(k)$ بیانگر ضرایب فیلتر FIR است که مسیر بین منبع qام تا میکروفون pام را مدل می‌کند. به این نکته باید توجه کرد که در این مدل فرض شده که سیگنال‌های منبع نقطه‌ای هستند و بنابراین می‌توان مسیر از منبع q به میکروفون p را توسط فیلترهای FIR مدل کرد. مشابه مدل لحظه‌ای، نویز جمع شونده در میکروفون‌ها هم توسط مولفه جمع شونده $n_p(n)$، در رابطه مدل لحاظ شده است. با فرض خطی بودن انتشار موج آکوستیکی، می‌توان مسیر آکوستیکی بین دو نقطه (منبع نقطه‌ای و میکروفون) را توسط یک تابع تبدیل خطی مدل کرد. با وجود اینکه امکان وقوع برخی پدیده‌های جانبی مانند جذب غیرخطی در محیط‌های واقعی وجود دارد، این مدل در اغلب موارد به خوبی عمل می‌کند. از طرفی به دلیل اینکه در حالت کلی محل منابع ثابت نیست، مسیرهای آکوستیکی عموماً متغیر با زمان است. در شبیه‌سازی‌ها، از فیلترهای FIR برای مدل کردن پاسخ‌های ضربه آکوستیکی استفاده می‌شود. یک روش بسیار مناسب و پرکاربرد برای مدلسازی پاسخ ضربه آکوستیکی، روش تصویر است [۲۰].

هدف در مسأله جداسازی کور منابع، یافتن ضرایب فیلترهای جداساز[1] FIR است به گونه‌ای که سیگنال‌های خروجی $y_q(n)$ $(q=1,\ldots,N)$ حاصل از رابطه زیر، تخمین خوبی از سیگنال‌های منبع اصلی را بدست می‌دهند:

[1] - demixing

$$y_q(n) = \sum_{p=1}^{m} \sum_{k=0}^{L-1} w_{qp}(k) x_p(n-k) \qquad (۶-۲)$$

پارامتر L، طول فیلترهای جداساز w_{qp} است. مدل ترکیب کانولوتیو به همراه سیستم جداساز در بلوک دیاگرام شکل (۲-۱) نمایش داده شده است. با توجه به این شکل، جداسازی کور منابع را می‌توان به عنوان یک تکنیک کور چند ورودی-چند خروجی درنظر گرفت.



شکل(۲-۱) : مدل BSS [۱۹].

در بیشتر الگوریتم‌های جداسازی کور منابع، فرض می‌شود که تعداد منابع و میکروفون‌ها با هم برابر است. با استفاده از روش‌های مناسبی می‌توان موارد کلی‌تر با تعداد دلخواه منابع و سنسورها را به فرم مدل جداسازی کور منابع استاندارد درآورد. حالتی که در آن تعداد حسگرها از تعداد منابع بیشتر است، به جداسازی کور منابع فرامعین[۱] معروف است. روش معروف برای حل این مسأله، استفاده از آنالیز مولفه‌های اصلی (PCA) بمنظور کاهش بعد مسأله و درنظر گرفتن تنها N مولفه اول و در نتیجه کاهش مساله فرامعین به مسأله استاندارد جداسازی کور منابع است. حالتی که در مسأله جداسازی کور منابع، تعداد میکروفون‌ها از تعداد منابع کمتر باشد را اصطلاحاً جداسازی کور منابع فرو معین[۲] می‌نامند. در تعدادی از الگوریتم‌ها از خاصیت تنک بودن منابع در حوزه زمان-فرکانس برای بدست آوردن بخش‌هایی که مربوط به سیگنال منبع است، استفاده شده است. بسیاری از محققان، روش‌هایی را برای تخمین تنکی در حوزه زمان-فرکانس بر مبنای مدل

[1]- over determined
[2] - under determined