بسمه تعالی

دانشکده مهندسی برق و کامپیوتر

گروه مهندسی مخابرات

پایان‌نامه

برای دریافت درجه کارشناسی ارشد

مهندسی برق – مخابرات

شکل‌دهی پرتو در شبکه‌های حسگر صوتی AD-HOC

استاد راهنما: دکتر حمیدرضا ابوطالبی

پژوهش و نگارش: سیدحمید یزدانی

شهریور ۱۳۹۲

بسم الله الرحمن الرحيم

تقدیم به پدر بزرگوار و مادر مهربانم

آن دو فرشته ای که از خواسته هایشان گذشتند، سختی ها را به جان خریدند و خود را سپر بلای مشکلات و ناملایمات کردند تا من به جایگاهی که اکنون در آن ایستاده ام برسم

خداوندا به ما توفیق تلاش در شکست، صبر در نومیدی، رفتن بی همراه، جهاد بی سلاح، کار بی پاداش، فداکاری در سکوت، دین بی دنیا، مذهب بی عوام، عظمت بی نام، خدمت بی نان، ایمان بی ریا، خوبی بی نمود، گستاخی بی خامی، مناعت بی غرور، عشق بی هوس، تنهایی در انبوه جمعیت و دوست داشتن بی آنکه دوست بداند، را عنایت فرما

یارب دل ما را تو به رحمت جان ده

درد همه را به صابری درمان ده

این بنده چه داند که چه می باید جست

داننده تویی هر آنچه دانی آن ده

در اینجا می‌خواهم از کسی که همواره مشوق و راهنمای من در دوران این پایان‌نامه بود تشکری صمیمانه داشته باشم. از استاد پرمایه و عزیزم جناب آقای دکتر حمیدرضا ابوطالبی که به حق، استادی کردند و شاگرد کوچک خود را همواره با دلگرمی‌های خود و با دانش خود به سمت درهای گشوده علم و پیشرفت پیش بردند کمال تشکر و احترام را قائل هستم.

سیدحمید یزدانی

چکیده:

بهسازی گفتار یکی از شاخه‌های مهم بحث پردازش گفتار است. امروزه با تنوع در کاربردهای پردازش گفتاری و وجود عوامل بسیار در تخریب سیگنال، امر بهسازی و بهبود کیفیت سیگنال گفتار از اهمیت زیادی برخوردار است. شکل‌دهی پرتو یکی از روش‌های بهسازی گفتار بوده و در آرایه‌های میکروفونی استفاده می‌شود. توجه به نوع چینش و محدودیت‌های موجود در چینش میکروفون‌ها پنجره‌ای را به سوی آرایه‌میکروفونی با چینش Ad-hoc با مزیت عدم محدودیت در نوع چینش گشوده است.

در این پایان‌نامه، هدف اصلی بررسی راهکارهایی برای شکل‌دهی پرتو در آرایه‌های میکروفونی Ad-hoc است. با توجه به عدم شناخت نسبت به مکان و چینش میکروفون‌ها و همچنین پراکنده بودن میکروفون‌ها در محیط ایده خوشه‌بندی میکروفون‌ها در این پایان‌نامه مورد توجه قرار گرفته است. برای میدان‌های نویزی متفاوت روش‌های متفاوتی برای خوشه بندی پیشنهاد شده است. توجه به تابع همدوسی میدان نویزی پخشنده و همچنین انرژی سیگنال‌های دریافتی از منابع مختلف نکته اساسی برای خوشه‌بندی در این پایان‌نامه به حساب می‌آید. بر اساس خوشه‌ها ساختار جدیدی برای شکل‌دهنده پرتو GSC پیشنهاد شده است.

در این پایان‌نامه نگاه جدیدی به عوامل مخرب سیگنال گفتار و میدان‌های نویزی شده است. این نگاه وابسته به نوع آرایه‌های میکروفونی است که در این پایان‌نامه مورد توجه قرار گرفته‌اند. سعی شده است تمامی روش‌های پیشنهادی در عین سادگی، بسیار کاربردی و عملی باشند. همچنین روش خوشه‌بندی در شرایط نویز جهتی سعی شده است که برای تمامی میدان‌های نویزی قابل به‌کارگیری باشد.

به طور کلی می‌توان هدف دومی نیز از این پایان‌نامه بدست آورد. هدف دوم کاهش پیچیدگی محاسبات و افزایش کیفیت خروجی بطور همزمان است که این امر با استفاده از آرایه‌های میکروفونی Ad-hoc بدست می‌آید.

کلمات کلیدی: بهسازی گفتار، شکل‌دهی پرتو، آرایه‌های میکروفونی Ad-hoc، خوشه‌بندی و شکل‌دهنده پرتو GSC.

فهرست مطالب:

فصل اول ۱

پیشگفتار ۱

۱-۱-عوامل مؤثر در تخریب سیگنال ۳

۱-۱-۱-میدان‌های نویزی ۳

۱-۱-۲-انعکاس ۴

۱-۱-۳-تداخل ۷

۱-۲-ساختار پایان‌نامه: ۷

فصل دوم ۹

پردازش هدفمند آرایه میکروفونی و شکل دهی پرتو ۹

۲-۱- آرایه میکروفونی کلاسیک و پردازش آرایه‌ای ۹

۲-۱-۱-شکل‌دهی پرتو کلاسیک ۱۰

۲-۱-۲-شکل‌دهی پرتوِ تأخیر و جمع (DSB) ۱۱

۲-۱-۳-شکل‌دهنده‌های وفقی پرتو GSC وTF_GSC ۱۱

۲-۲- آرایه میکروفونی Ad-hoc، تعریف و بیان ویژگی‌ها ۱۴

۲-۲-۱-هم‌پوشانی فضایی ۱۶

۲-۳- شکل‌دهی پرتو در آرایه‌های میکروفونی Ad-hoc ۱۸

۲-۴- خلاصه ۲۱

فصل سوم ۲۳

شکل‌دهی پرتو کور در آرایه‌های AD-HOC، در شرایط میدان نویزی پخشنده ۲۳

۳-۱- بررسی جامع‌تر آرایه‌های میکروفونی Ad-hoc ۲۳
عامل مخرب نویز: ۲۷
عامل مخرب پژواک: ۲۷

۳-۲- بررسی و شبیه‌سازی روش خوشه‌بندی ۲۸
۳-۲-۱- معیاری برای تعیین میزان نزدیکی بین میکروفون ها و رتبه‌بندی خوشه‌ها ۲۹
۳-۲-۲- شبیه‌سازی روش خوشه‌بندی پیشنهادشده در [] ۳۴

۳-۳- توسعه روش خوشه‌بندی برای شرایط محیطی مختلف ۴۰
۳-۳-۱- توسعه روش خوشه‌بندی ۴۱
۳-۳-۲- توسعه نحوه رتبه‌بندی ۴۵

۳-۴-شبیه‌سازی روش توسعه یافته و ارزیابی آن ۴۶

۳-۵- خلاصه ۵۱

فصل چهارم ۵۲

شکل‌دهی پرتو کور در آرایه‌های AD-HOC، در شرایط نویز جهتی ۵۲

۴-۱- بررسی شرایط نویز جهتی و تداخل با نگاه خوشه‌بندی میکروفون‌ها ۵۳

۴-۲- پیشنهاد روش خوشه‌بندی برای شرایط نویز جهتی و تداخل ۵۷

۴-۳- پیاده‌سازی روش پیشنهادی برای خوشه‌بندی ۵۸
۴-۳-۱- فریم‌بندی سیگنال میکروفون‌ها و محاسبه میانگین انرژی ۶۰
۴-۳-۲- تعیین خوشه‌های نهایی و رتبه‌دهی آن‌ها ۶۲
۴-۳-۳- پیشنهاد معیار دیگری برای خوشه‌بندی برای شرایط خاص‌تر ۶۶

۴-۴- پیشنهاد ساختار جدیدی برای شکل‌دهنده پرتو وفقی GSC ۶۷

۴-۵- مقایسه نتایج برای روش GSC متمرکز و GSC خوشه‌بندی شده ۷۰

۴-۶- خلاصه ۷۳

فصل پنجم ۷۵

جمع‌بندی و نتیجه‌گیری ۷۵

۵-۱-خلاصه‌ای از کارهای انجام‌گرفته ۷۵

۵-۲- دستاوردها و نوآوری‌های این پایان‌نامه ۷۷

۵-۳-پیشنهادهایی برای ادامه کار ۷۸

مراجع و منابع ۸۰

فهرست شکل‌ها

شکل(۱-۴): نمایش اندازه پاسخ ضربه یک محیط صوتی دارای بازتاب ۶
شکل ۱-۵- نمایش یک محیط نمونه برای پردازش گفتار ۷
شکل(۲-۱): دیاگرام بلوکی GSC ۱۲
شکل ۳-۱- مدل سیگنال دریافتی توسط آرایه میکروفونی ۲۴
شکل ۳-۲- چگالی انرژی مسیر مستقیم و بازتاب‌ها بر حسب فاصله برگرفته از [] ۲۸
شکل (۳-۳) - اندازه مربع همدوسی برای سه فاصله بین میکروفونی متفاوت ۲۹
شکل(۳-۴)- نشان دادن میزان خطای محاسبه اختلاف زمانی سیگنال دریافتی برای فاصله‌های مختلف بین میکروفونی ۳۳
شکل (۳-۵)- محیط سناریوی اول مورد استفاده در فصل سوم ۳۵
شکل (۳-۶)- محیط سناریوی دوم مورد استفاده در فصل چهارم ۳۶
شکل ۳-۷- نحوه تغییرات PESQ خروجی برای گوینده‌های اول (شکل بالا) و دوم (شکل وسط) و سوم (شکل پایین) در سناریوی اول بر حسب تغییرات A و خروجی نزدیک‌ترین خوشه (رنگ سبز). ۳۹
شکل ۳-۸- تغییرات نسبت سیگنال به نویز قطعه‌ای بر حسب A برای گوینده اول (شکل بالا) و گوینده دوم (شکل پایین) برای ترکیب خوشه‌ها، نزدیک‌ترین خوشه و تمام میکروفون‌ها ۴۰
شکل ۳-۹- نمونه‌ای از پاسخ ضربه محیط برای بررسی ماهیت سیگنال‌های رسیده به میکروفون‌ها از دید بازتاب و انتشار ۴۶
شکل ۳-۱۰- نتایج پیاده‌سازی روش‌های خوشه‌بندی بر اساس قاعده (ستون دوم) و روش پیشنهادی (ستون سوم) به همراه خوشه‌های درست بر اساس مکان دقیق میکروفون‌ها (ستون اول) در چهار چینش میکروفونی مختلف ۴۷
شکل ۳-۱۱- نتیجه اعمال فیلتر کالمن و تخمین فاصله‌ها و اعمال تکنیک MDS و کالیبراسیون میکروفون‌ها در چهار چینش متفاوت برای بهتر نشان دادن روش پیشنهادی برای خوشه‌بندی. ۴۸
شکل ۳-۱۲- بلوک دیاگرام خلاصه‌ای از کارهای انجام گرفته در فصل سوم ۵۱
شکل ۴-۱- نمایی از یک محیط و آرایه‌ی میکروفونی AD-HOC با انواع نویز و تداخل و چندین گوینده ۵۳
شکل ۴-۲- چگالی انرژی مسیر مستقیم و بازتاب‌ها بر حسب فاصله برگرفته از [] ۵۶
شکل ۴-۳- یک سناریوی نمونه برای نشان دادن ایده اصلی در این فصل برای خوشه‌بندی ۵۶
شکل ۴-۴- بلوک دیاگرام روش پیشنهادی برای خوشه‌بندی ۵۸
شکل ۴-۵- سناریوی اول مورد استفاده در فصل سوم برای پیاده‌سازی الگوریتم‌ها ۵۹
شکل ۴-۶- سناریوی دوم مورد استفاده در فصل سوم برای پیاده‌سازی الگوریتم‌ها ۵۹
شکل ۴-۷- تغییرات میانگین انرژی در تمام میکروفون‌ها در لحظه شروع گفتار در سناریوی دوم و گوینده اول. ۶۱
شکل ۴-۸-تغییرات میانگین انرژی در تمام میکروفون‌ها در لحظه شروع گفتار در سناریوی دوم و گوینده دوم.. ۶۱

شکل ۴-۹-تغییرات میانگین انرژی در تمام میکروفون‌ها در لحظه شروع گفتار در سناریوی دوم و گوینده سوم. ۶۲
شکل ۴-۱۰- نمودار تغییرات انرژی برای گوینده اول (بالا) و گوینده دوم (وسط) و گوینده سوم (پایین) در سناریوی دوم بر حسب میکروفون‌ها ۶۵
شکل ۴-۱۱- نمودار تغییرات آنتروپی برای گوینده اول (بالا) و گوینده دوم (وسط) و گوینده سوم (پایین) در سناریوی دوم بر حسب میکروفون‌ها ۶۵
شکل ۴-۱۲- ساختار کلی شکل‌دهنده پرتو GSC ۶۸
شکل۴-۱۳- نمای کلی از یک فیلتر وفقی برای کاهش نویز مکانی ۶۸
شکل ۴-۱۴- نمای کلی شکل‌دهنده پرتو پیشنهادی ۶۹
شکل ۴-۱۵- نمودار تغییرات معیار کیفی PESQ بر حسب SNR برای دو روش GSC متمرکز و خوشه‌بندی پیشنهادشده برای گوینده اول در سناریوی دوم ۷۱
شکل ۴-۱۶-نمودار تغییرات معیار کیفی SEG.SNR بر حسب SNR برای دو روش GSC متمرکز و خوشه‌بندی پیشنهادشده برای گوینده اول در سناریوی دوم ۷۱
شکل ۴-۱۷-نمودار تغییرات معیار کیفی LLR بر حسب SNR برای دو روش GSC متمرکز و خوشه‌بندی پیشنهادشده برای گوینده اول در سناریوی دوم ۷۲
شکل ۴-۱۸-نمودار تغییرات معیار کیفی WSS بر حسب SNR برای دو روش GSC متمرکز و خوشه‌بندی پیشنهادشده برای گوینده اول در سناریوی دوم ۷۲
شکل ۴-۱۹- مقایسه خروجی دو حالت برای GSC خوشه‌بندی شده. در حالت اول روش پیشنهادی اعمال شده ولی در حالت دوم در شاخه پایینی همان خوشه شاخه بالا استفاده شده است. ۷۳
شکل۴-۲۰- بلوک دیاگرام خلاصه کارهای انجام شده در فصل چهارم ۷۴

فهرست جداول

جدول (۳-۱) - نتایج شبیه‌سازی روش خوشه‌بندی هیماوان در سناریوی اول ۳۶
جدول (۳-۲) - نتایج شبیه‌سازی روش خوشه‌بندی هیماوان در سناریوی دوم ۳۷
جدول (۳-۳)- میزان PESQ خروجی شکل‌دهنده پرتو تأخیر و جمع برای حالت‌ها و سناریوهای مختلف ۳۸
جدول(۳-۴)- پارمترهای معادله گوس-مارکف و روابط فیلتر کالمن برای تخمین فاصله بین میکروفونی ۴۳
جدول ۳-۵- میانگین شماره فریم لحظه آغاز سیگنال برای حالت‌های مختلف محیطی برای گوینده اول در سناریوی دوم ۴۹
جدول ۳-۶-میانگین شماره فریم لحظه آغاز سیگنال برای حالت‌های مختلف محیطی برای گوینده دوم در سناریوی دوم ۴۹
جدول ۳-۷-میانگین شماره فریم لحظه آغاز سیگنال برای حالت‌های مختلف محیطی برای گوینده سوم در سناریوی دوم ۵۰
جدول ۴-۱- نتیجه اعمال روش خوشه‌بندی جدید در سناریوی اول برای تمام گوینده‌ها ۶۳
جدول ۴-۲- نتیجه اعمال روش خوشه‌بندی جدید در سناریوی دوم برای تمام گوینده‌ها ۶۳
جدول ۴-۳- رتبه گفتاری و رتبه نویزی بدست آمده از روش پیشنهادی برای سناریوی اول و تمام گوینده‌ها ۶۶
جدول ۴-۴-رتبه گفتاری و رتبه نویزی بدست آمده از روش پیشنهادی برای سناریوی دوم و تمام گوینده‌ها ۶۶

حروف اختصاری

فصل اول

پیشگفتار

در چند دهه اخیر، ارتباط بین انسان و ماشین‌ها یکی از مهم‌ترین زمینه‌های تحقیقاتی در حوزه پردازش گفتار بوده است. همچنین ارتباط گفتاری بهترین و قدرتمندترین رابط کاربری با کامپیوترها به حساب می‌آید. با توجه به رشد و گستردگی تکنولوژی ، پیشرفت در پردازش گفتار و زمینه‌های دیگر مانند پردازش تصویر، اهداف جدیدی در حوزه ارتباطات شکل گرفته است. با توجه به این پیشرفت‌ها، نسل جدید کامپیوترها باید به نیازهای بشر وفق داده شوند و به سمت هوشمند شدن تمام سیستم‌ها متمایل شوند. بنابراین باید سیستم‌ها بتوانند در شرایط متفاوت عملکرد قابل قبولی داشته باشند.

یک مثال جالب در این زمینه، تلاش‌های صورت گرفته برای توسعه اتاق‌های هوشمند** است. یک اتاق هوشمند، یک فضای بسته تشکیل شده از چندین میکروفون و دوربین است که برای کمک و تکمیل کردن فعالیت‌های انسان‌ها طراحی شده‌اند. هدف از این اتاق، دادن خدماتی همچون ضبط هوشمند صدا و تصویر، ارائه خلاصه‌ای از جلسات به طور هوشمند و ... است.

در حوزه صدا و گفتار، اهداف مختلفی از قبیل آشکارسازی حضور صدا، بازشناسی هوشمند گفتار، شناسایی هویت گوینده و مکان گوینده** و ... تعریف شده است. در اکثر این اهداف، زمانی که سناریوهای کنترل شده (میکروفون‌های دستی) وجود داشته باشند، تقریباً سیستم‌های موجود عملکرد قابل قبولی دارند. با پیشرفت کاربردهای گفتاری هوشمند (بدون استفاده از دست) که در آنها استفاده از میکروفون‌های دستی امکان‌پذیر نیست، در عملکرد سیستم‌ها مشکلاتی به وجود می‌آید.

در شرایط جدید، سیگنال‌های صوتی توسط میکروفون‌هایی که از گوینده فاصله نسبتاً زیادی دارند ضبط می‌شوند. به دلیل فاصله زیاد گوینده و میکروفون‌ها، سیگنال گفتار مطلوب به علت وجود عوامل مخربی همچون نویز، پژواک، جهت و حرکت چندین گوینده همزمان و تداخل و ... توسط میکروفون‌ها با کیفیت پایین ضبط می‌شود. کاهش کیفیت سیگنال میکروفون‌ها باعث افت عملکرد سیستم‌های پردازش گفتار می‌شود.

با توجه به پاراگراف قبل، باید به بحث بهبود کیفیت سیگنال میکروفون‌ها توجه ویژه‌ای شود. در چند دهه اخیر به این موضوع توجه بسیاری شده و روش‌های مختلفی نیز برای آن ارائه شده است []. تخریب سیگنال‌ها به اندازه‌ای در سیستم‌های پردازش گفتاری با اهمیت است که امروزه شاخه بزرگی با عنوان بهسازی سیگنال گفتار در حوزه پردازش گفتار شکل گرفته است. به طور کلی بهسازی سیگنال گفتار، به معنای بالا بردن کیفیت و قابلیت فهم سیگنال گفتار و بهبود عملکرد سیستم‌های پردازش گفتار هوشمند در حضور شرایط محیطی مخرب است.

همانطور که گفته شد، در سیستم‌های ارتباط گفتاری، امکان دور بودن منبع سیگنال مطلوب از میکروفون‌ها وجود دارد. منبع مورد نظر، موج صوتی را تولید می‌کند. بخشی از این امواج، به طور مستقیم به میکروفون‌ها می‌رسند. بخشی دیگر شامل بازتاب‌ها[1] است که با تأخیر به میکروفون می‌رسند. در کنار این موارد، نویز پس‌زمینه نیز وارد میکروفون می‌شود. برای از بین بردن تأثیرات

[1] - Reflections

بازتاب‌ها، نویز پس‌زمینه و دیگر تداخل‌ها باید سیستم بهسازی گفتار با کارایی بالا طراحی شود. سیستم‌های بهسازی گفتار به دو دسته تک میکروفونه و چندمیکروفونه (پردازش آرایه‌ای) افراز می‌شوند. در فصل دوم، درباره این سیستم‌ها بحث خواهد شد.

به دلیل این که نمی‌توان همیشه میکروفون را در نزدیکی منبع مطلوب قرار داد، سیگنال توسط ۱- بازتاب‌ها (که نتیجه انتشار چند جهته هستند) و ۲- نویز (که به دلیل وجود منابع مزاحم است)، تخریب می‌شود. یکی از تفاوت‌های اساسی بین نویز و بازتاب این است که بازتاب‌ها وابسته به سیگنال مطلوب هستند در حالی که نویز معمولاً مستقل (یا ناهمبسته) از سیگنال مطلوب است. به همین علت خیلی از سیستم‌های بهسازی گفتار که برای حذف نویز طراحی شده‌اند، کارایی لازم در حضور بازتاب را ندارند.

میزان تخریب سیگنال توسط بازتاب با افزایش فاصله میکروفون تا منبع، افزایش می‌یابد. این تخریب باعث کاهش قابلیت فهم سیگنال گفتار، کاهش کارایی سیستم‌های کنترلی گفتار و کاهش کارایی الگوریتم‌های کدگذاری در سیستم‌های تلفنی، می‌شوند. در بخش‌های بعدی، برخی از عوامل مخرب سیگنال گفتار بررسی خواهند شد.

۱-۱-عوامل مؤثر در تخریب سیگنال

در این بخش، به بررسی اجمالی مهمترین عوامل مخرب که نویز، تداخل و انعکاس هستند پرداخته می‌شود:

۱-۱-۱-میدان‌های نویزی

به طور کلی به میدان‌های صوتی در غیاب سیگنال گفتار مورد مطلوب، میدان نویزی گفته می‌شود. در حقیقت، نویز مجموعه‌ای از سیگنال‌های مزاحم است که با قرار گرفتن در میدان صوتی گفتار مطلوب باعث افت کیفیت آن می‌شود. در محیط‌های واقعی، این عامل مخرب از منابع

متنوعی سرچشمه می‌گیرد. از این رو میدان‌های نویزی از تنوع وسیعی برخوردار هستند و می‌توان از دیدگاه‌های مختلفی آن‌ها را تقسیم‌بندی کرد.

در دیدگاه اول، بر اساس میزان همبستگی سیگنال‌های نویز در نقاط مختلف از فضا به دو دسته نویزهای همبسته و نویزهای ناهمبسته افراز می‌شوند؛ در یک میدان نویزی همبسته، امواج بدون هیچ انعکاس و یا پراکندگی به میکروفون‌ها می‌رسند به نحوی که این سیگنال‌ها در نقاط مختلف فضا از جمله روی میکروفون‌ها کاملاً به یکدیگر همبسته می‌باشند. پس می‌توان چنین نتیجه گرفت که میدان نویزی همبسته، از منبعی در فضای آزاد و بدون مانع عمده در پخش صدا حاصل می‌شود. این در حالی است که در میدان‌های نویزی ناهمبسته، سیگنال‌های نویز در نقاط مختلف فضا از همبستگی بسیار پایینی برخودار هستند. از نمونه‌های این دسته از میدان‌های نویزی می‌توان به نویز الکتریکی نیکروفون‌ها اشاره کرد.

در دیدگاه دوم، بر اساس میزان پویایی نویز می‌توان آن را به دو دسته ایستان و غیر ایستان افراز نمود. در دیدگاه سوم []، نویزهای محیط در قالب دو دسته نویزهای جهت‌دار و همه‌جهته (یا diffuse) تقسیم شده‌اند. همان طور که از نام آنها نیز مشخص است، در دسته اول، سیگنال‌های مزاحم از یک جهت مشخصی از فضا دریافت می‌شوند. در دسته دوم، این سیگنال‌ها از جهت‌های مختلفی از محیط به میکروفون می‌رسند. میدان نویزی diffuse، برآیند سیگنال‌های ناخواسته‌ای است که با احتمال مساوی و انرژی یکنواخت در همه جهت‌های محیط منتشر شده و همگی دارای دامنه برابر و فاز تصادفی هستند[]؛ این مدل که تلفیقی از اجزای همبسته و ناهمبسته است، به دلیل کاربرد زیادی که در شرایط واقعی (همچون محیط‌های کاری و یا محیط داخل خودرو) دارد، از اهمیت بسیار بالایی برخوردار است. در این پژوهش، دیدگاه سوم مورد توجه قرار گرفته است.

۱-۱-۲-انعکاس