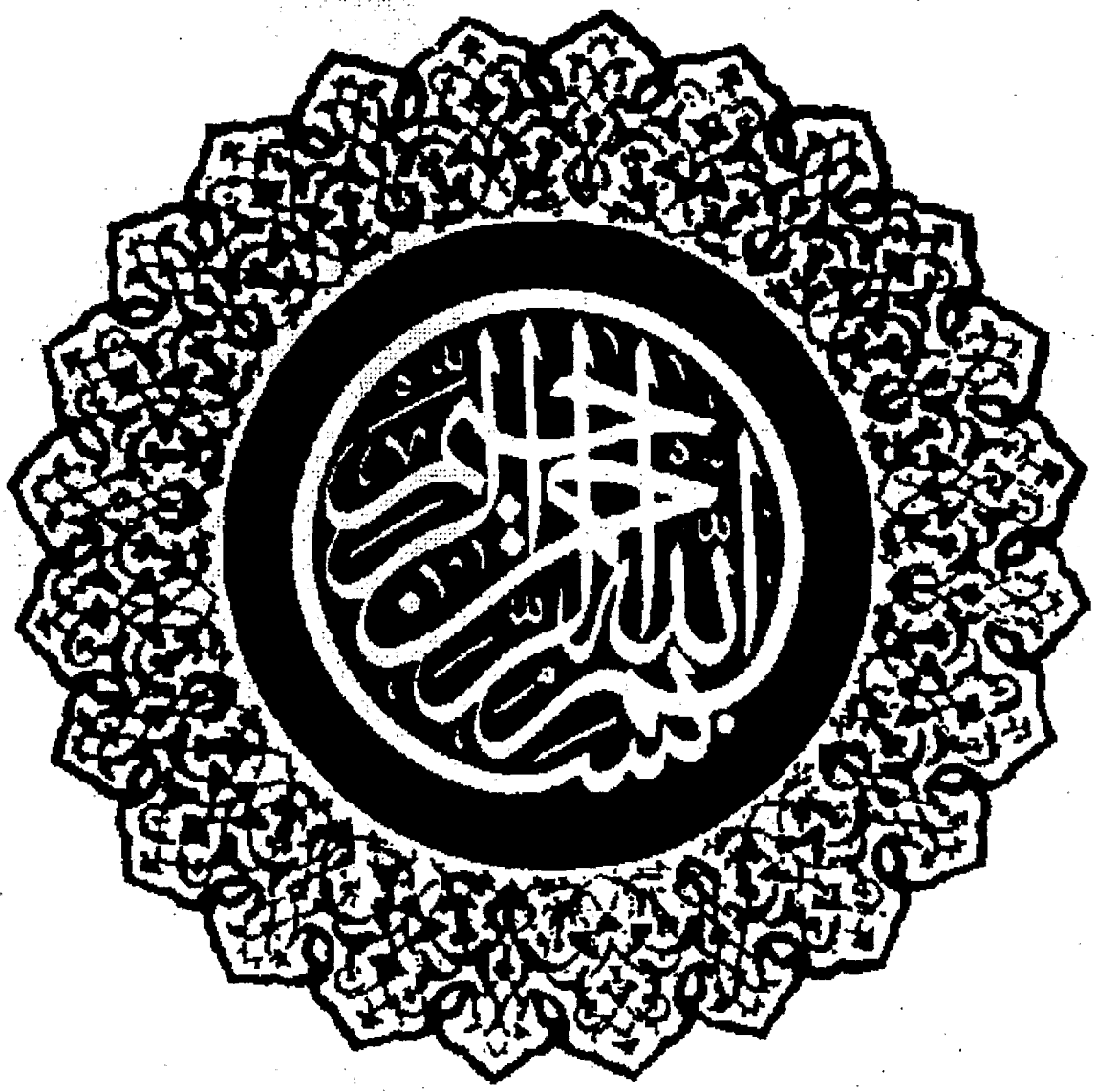

٩٤٢٩٨

مجتمع فنی و مهندسی

دانشکده مهندسی برق

پایان نامه کارشناسی ارشد مهندسی برق ـ مخابرات

عنوان:

بهبود عملکرد روشهای چندکانالهٔ بهسازی گفتار در شرایط واقعی

استاد راهنما:

دکتر حمیدرضا ابوطالبی

استاد مشاور:

دکتر مسعودرضا آقابزرگی صحاف

پژوهش و نگارش:

مجید پوراحمدی

۱۳۸۶/۱۱/۲۸

اسفند ۱۳۸۴

۹۲۲۹۶

تقدیم به دو اسوهٔ صبر و عطوفت

پدر و مادر عزیزم

که همواره غمخوار و پشتیبانم در زندگی بوده اند

تقدیر و تشکّر

اکنون که به یاری خداوند متعال موفق به ارائهٔ این پایان نامه گردیده ام بر خود لازم می دانم تـا از زحمات و تلاشهای بی شائبهٔ استاد راهنمای گرامی ام، جناب آقای دکتر ابوطالبی کـه از آغـاز دورهٔ کارشناسی ارشد همواره از راهنمایی های علمی و مشـورتهای برادرانـهٔ ایشـان اسـتفاده نمـوده ام، کمال تشکر و قدردانی را داشته باشم.

همچنین رعایت ادب لازم می دارد تا از جناب آقای دکتر آقابزرگی که زحمت مشاوره و راهنمـایی مرا در تکمیل این پایان نامه بر عهده داشته اند خاضعانه و خالصانه تشکر و قدردانی نمایم.

از دبیرخانهٔ محترم شورای عالی اطلاع رسانی که حمایت مالی این پایان نامه را عهده دار بـوده انـد نیز سپاسگزاری می نمایم. مسلماً حمایتهای آن شورای محترم زمینـه سـاز رشـد و ترقـی جامعـهٔ علمی کشور عزیزمان ایران خواهد بود.

در پایان سپاس بیکران خود را تقدیم کلیهٔ اساتید محترم دانشـکدهٔ بـرق کـه افتخـار شـاگردی آن عزیزان نصیب من گردیده، می نمایم و امیدوارم تا در تمام مراحـل زنـدگی بـا حمایتهـای خداونـد متعال موفق و مؤید باشند.

بسمه تعالی

مدیریت تحصیلات تکمیلی

صورتجلسه دفاعیه پایان نامه دوره کارشناسی ارشد

جلسه دفاعیه پایان نامه تحصیلی خانم /آقای **مجید پوراحمدی**

دانشجوی کارشناسی ارشد **مجتمع فنی و مهندسی** دانشگاه یزد،دررشته/ گرایش. **مهندسی برق - مخابرات**

تحت عنوان: **بهبود و عملکرد روشهای چند کاناله بهسازی گفتار در شرایط واقعی**

وتعداد واحد: ٦ واحد درتاریخ ۱۰ / ۱۲ /۱۳۸٤

باحضور اعضای هیات داوران متشکل از :	نام ونام خانوادگی	امضاء
۱-استاد راهنما	دکتر حمیدرضا ابوطالبی	
۲- استاد مشاور	دکتر مسعود رضا آقابزرگی	
۳-داورخارج از گروه	دکتر فرشاد الماس گنج	
۴-داور داخل گروه	دکتر قاسم میرجلیلی	

تشکیل گردید و پس ازارزیابی پایان نامه توسط هیات داوران، بادرجه عالی و نمره: به عدد ۱۹/۷۵ به حروف نوزده و هفتاد و پنج مورد تصویب قرارگرفت.

نماینده تحصیلات تکمیلی دانشگاه (ناظر)

نام ونام خانوادگی: دکتر ولی کلانتر

امضاء

فهرست مطالب

فصل اول : پیش گفتار ۱

۱-۱) پیش گفتار ۱

۱-۲) ساختار این پایان نامه ۲

فصل دوم : حذف کنندهٔ وفقی نویز ۵

۲-۱) حذف کنندهٔ وفقی نویز ۵

۲-۲) تابع همدوسی فضایی ۹

۲-۲-۱) میزان کاهش نویز ۹

۲-۲-۲) میدان نویزی diffuse ۱۱

۲-۳) نتیجه گیری ۱۲

فصل سوم : مشکل نویزهای ناهمبسته در کانالها ۱۴

۳-۱) شکل دهنده های موج ۱۴

۳-۲) فضای دیتا ۱۹

۳-۲-۱) شکل دهنده موج با محدودیت خطی و واریانس مینیمم ۱۹

۳-۳) وفقی سازی در فضای بیم ۲۳

۳-۳-۱) عملکرد شکل دهنده موج در حذف نویز ۲۵

۳-۴) ساختار GSC ۲۷

۳-۴-۱) عملکرد GSC در حذف نویز ۳۰

۳-۵) نتیجه گیری ۳۲

فصل چهارم : جلوگیری از گفتار نشتی ۳۳
۴-۱) مقابله با گفتار نشتی در کانال مرجع ۳۴
۴-۱-۱) استفاده از شکل دهنده های موج برای کاهش گفتار نشتی در ANC ۳۴
۴-۱-۲) استفاده از ساختار فیلترهای وفقی ضربدری برای حذف نویز در حضور گفتار نشتی ۳۵
۴-۱-۲-۱) از حذف نویز تا جداسازی سیگنالها ۳۶
۴-۲) روشهای پیاده سازی CTRANC ۴۰
۴-۲-۱) مزیّت اطلاعات اضافی برای جداسازی سیگنال ۵۰
۴-۲-۲) علّی بودن ساختار CTRANC ۵۱
۴-۲-۳) پایداری فیلترها ۵۳
۴-۳) ساختار ACTRANC ۵۵
۴-۴) نتیجه گیری ۵۹

فصل پنجم : جداسازی کور منابع ۶۱
۵-۱) معرفی کلی مسألهٔ جداسازی کورمنابع (BSS) ۶۲
۵-۱-۱) سیستم ترکیب کننده ۶۳
۵-۱-۲) فرضهای خاص آماری برای منابع ۶۵
۵-۲) نحوهٔ پیاده سازی الگوریتم جداکننده ۶۷
۵-۳) جداسازی کور منابع صوتی ۷۰
۵-۳-۱) الگوریتم MAD ۷۳
۵-۳-۱-۱) ترکیب کننده های کانوالوکننده ۷۴
۵-۳-۲) حذف نویز زمینه ۷۹
۵-۴) نتیجه گیری ۸۴

فصل ششم : پیاده سازی، نتیجه گیری و جمع بندی ۸۵
۶-۱) شبیه سازی ۸۵
۶-۱-۱) سیستم ANC ۸۵
۶-۱-۲) CTRANC و ACTRANC ۹۰

۶-۱-۳) سیستم ترکیبی ACTRANC و BSS ۱۰۱
۶-۲) نتیجه گیری ۱۰۳
۶-۳) نوآوری ۱۰۵
۶-۴) پیشنهادات ۱۰۶

مراجع **۱۰۸**

ضمیمهٔ ۱ **۱۱۴**
ضمیمهٔ ۲ **۱۱۵**

فهرست شکلها

شکل (۲ـ۱): ساختار حذف کننده وفقی نویز ۶
شکل (۲-۲): میزان کاهش نویز برحسب مربع تابع همدوسی ۱۰
شکل (۳-۱): یک آنتن نوعی و شکل بیم قلمی آن ۱۶
شکل (۳-۲): ساختار یک fixed Beamformer ۱۷
شکل (۳-۳): اولین ساختار Beamformer وفقی ۱۸
شکل (۳-۴): ساختار Beamformer وفقی ۱۹
شکل (۳-۵): تابش موج به دو سنسورمجاور در آرایه شامل M سنسور ۲۰
شکل (۳-۶): پترن آرایه برای چهار مقدار مختلف α ۲۵
شکل (۳-۷): ساختار GSC ۲۹
شکل (۳-۸): اجرای ساختار GSC به صورت یک مسأله بهینگی استاندارد ۳۰
شکل (۳-۹): نمایش کلی یک GSC ۳۱
شکل (۴-۱): مدل ساده شده کانال بین منابع و سنسورها ۳۳
شکل (۴-۲): ساختار سیستم ANC ۳۷
شکل (۴-۳): ساختار CTRANC ۳۹
شکل (۴-۴): مدل کانال ۴۱
شکل (۴-۵): اجرای feed forward ۴۲
شکل (۴-۶): اجرای feed back ۴۲
شکل (۴-۷): سیستم تولید Z و U ۴۸
شکل (۴-۸): جهت تابش منابع ۵۵
شکل (۴-۹): سیستم CTRANC فقط در حضور سیگنال ۵۷

شکل (۴-۱۰): ساختار ACTRANC ۵۸
شکل (۴-۱۱): نمایش حرکت منبع نویز بین دو میکروفون ۵۸
شکل (۴-۱۲): مدل تولید سیگنال و ساختار CTRANC ۵۹
شکل (۵-۱): مدل سادهٔ یک سیستم جداسازی کور منابع ۶۹
شکل (۵-۲): BSS در حوزهٔ فرکانس ۷۵
شکل (۵-۳): الف) سیستم ترکیبی BSS و فیلتربانک. ب) بلوک دیاگرام تبدیل موجک مورد استفاده در فیلتر بانک ۸۲
شکل (۵-۴): سیستم پیشنهادی برای حذف همزمان نویزهای نقطه ای و زمینه ۸۳
شکل (۶-۱): a) گفتار تمیز (بدون نویز) b) گفتار نویزی کانال primary با نویز پایین گذر c) خروجی ANC d) خروجی با نویز همهمه e) خروجی با وجود نشت گفتار در کانال مرجع ۸۹
شکل (۶-۲): خروجی سیستم CTRANC برای حالتی که a) منبع نویز نزدیک میکروفون مرجع باشد b) منبع نویز روی عمود منصّف محور میکروفونها باشد c) منبع نویز نزدیک به میکروفون primary باشد ۹۴
شکل (۶-۳): خروجی سیستم ACTRANC برای حالتی که a) منبع نویز نزدیک میکروفون مرجع باشد b) منبع نویز روی عمود منصّف محور میکروفونها باشد c) منبع نویز نزدیک به میکروفون primary باشد ۹۵
شکل (۶-۴): بهبود SNR برای مقادیر مختلف تأخیر انتشار ۹۶
شکل (۶-۵): LAR-distance برای گفتار نویزی، خروجی CTRANC و ACTRANC ۹۸
شکل (۶-۶): خروجی روش a) CTRANC و b) BSS در محیطی با انعکاس بالا ۱۰۰
شکل (۶-۷): مقایسهٔ روشهای مختلف بهسازی گفتار a) گفتار تمیز b) خروجی BSS c) BSS+ACTRANC d) فیلتر وینر+BSS e) BSS+MMSE f) روش تفریق طیفی+BSS ۱۰۳

فهرست جداول

جدول (۶-۱): SNR خروجی سیستم ANC در دو حالت الف) بدون نشت گفتار و مقادیر مختلف SNR کانال primary ب) با وجود نشت گفتار و با $SNR_{Pri} = 5\,dB$ ۸۷

جدول (۶-۲): تأثیر درجهٔ APA روی SNR خروجی ۹۹

جدول (۶-۳): SNR خروجی ACTRANC برای تأخیرهای مختلف ۹۹

جدول (۶-۴): بهبود SNR در CTRANC و ACTRANC ۱۰۱

جدول (۶-۴): بهبود SNR در روشهای مختلف با تغییر SNR ورودی BSS ۱۰۲

چکیده

استفادهٔ گسترده از سیستم های مخابراتی مستلزم کارکرد مناسب این سیستم ها در شرایط نویزی می باشد. در محیطهای واقعی، نویزهای محیط عملکرد سیستم را تا حد زیادی خدشه دار ساخته به طوری که وجود یک الگوریتم کاهش دهندهٔ نویز در این سیستمها امری ضروری به نظر می رسد. گسترده ترین سیستم های دو کانالهٔ حذف نویز، حذف کننده های وفقی نویز می باشد که در آن یک الگوریتم وفقی با الگوریتم LMS (Least Mean Square) برای حذف نویز مورد استفاده قرار می گیرد. اجرای مؤثر این سیستم مستلزم آن است که نویز مرجع که با کانال مرجع ضبط می شود همبستگی زیادی با نویز موجود در کانال دیگر داشته باشد. همچنین باید از نفوذ سیگنال مطلوب در کانال مرجع جلوگیری کرد. تأثیر این سیگنال نشتی، کاهش عملکرد سیستم می باشد، زیرا سیگنال موجود در کانال مرجع باعث حذف سیگنال مطلوب همراه با نویز توسط سیستم می شود.

برای غلبه بر مشکل نویزهای ناهمبسته تاکنون روشهای مختلفی پیشنهاد شده که استفاده از شکل دهنده های موج و روش GSC (Generalized Sidelobe Canceller) از معروفترین آنهاست. اما عملکرد این سیستم ها در فرکانسهای مختلف متفاوت بوده و در مجموع نمی توان کارکرد قابل ملاحظه ای را از آنها انتظار داشت. نکتهٔ قابل توجه این است که امروزه به علّت کوچک شدن مجموعه سیستمهای مورد استفاده و نزدیکتر شدن فاصلهٔ میکروفونها، مشکل سیگنال نشتی به مراتب شدیدتر از نویز ناهمبسته شده و بیشتر تحقیقات در این راستا در حال اجرا می باشد. برای غلبه بر سیگنال نشتی اوّلین ساختار قابل استفاده، CTRANC (CrossTalk Resistant Adaptive Noise Canceller) ها بود که بعداً به علت اشکال آنها در مواجهه با نویزهای جهت دار، ساختار کاملتر آن به نام ACTRANC (Asymmetric CTRANC) ارائه شد. در این ساختار با استفاده از ویژگی شکل دهی موج و با تأخیردهی مناسب به سیگنال دریافتی از میکروفون اصلی می توان عملکردی به مراتب بهتر نسبت به سیستم متداول CTRANC داشت.

در اینصورت با حرکت منبع نویز و تغییر موقعیت آن نسبت به میکروفونها، خدشه ای در عملکرد سیستم در حذف نویز وارد نخواهد شد. در این پایان نامه با بررسیهای مختلف و مقایسهٔ نتایج حاصل، بهترین مقدار تأخیر روی کانال اصلی را برای روش ACTRANC بدست آورده و علاوه برآن با استفاده از الگوریتم Affine Projection، الگوریتم جدیدی را برای وفقی سازی همزمان فیلترهای موجود در سیستم معرفی نموده ایم. امّا شبیه سازیها و همچنین کارکرد سیستم در محیط های واقعی نشان می دهد که در محیط های با انعکاس بالا اینگونه سیستم ها کارایی مناسبی نخواهد داشت. به همین دلیل توجه اکثر محققین به سوی سیستم های BSS (Blind Source Separation) جذب شده که در شرایط واقعی با انعکاس نسبتاً بالا، عملکرد قابل قبولی از خود نشان می دهند.

با این وجود اگر علاوه بر نویزهای نقطه ای ذکر شده، نویز زمینه هم در محیط وجود داشته باشد سیستم BSS از حذف آن ناتوان خواهد بود. در این پایان نامه یک روش جدید بهسازی گفتار در محیطهای شامل دو منبع صوتی و همراه با منابع نویز زمینه ارائه خواهد شد. روش ارئه شده، شامل ترکیب یک واحد جداسازی کور منابع (BSS) به همراه ساختار ACTRANC می باشد. واحد اول منابع نقطه ای مثل سیگنالهای صوتی دو گوینده نزدیک به هم را جدا کرده درحالیکه واحد دوم نویز زمینه را از این سیگنالها حذف خواهد کرد. شبیه سازیهای مختلف، کارکرد مناسب این سیستم را با استفاده از سیگنالهای گفتار آغشته به نویز همهمه (babble) نشان می دهد.

فصل اوّل:

پیش گفتار

۱-۱) پیش گفتار

در طول چند دهۀ اخیر و با پیشرفت جنبه های مختلف پردازش سیگنال گفتار، روشهای وسیعی برای پاکسازی سیگنال گفتار از نویز ارائه گردیده است. از یک منظر می توان کلیۀ این روشها را بر اساس تعداد کانال (میکروفون) ورودی به دو دسته تقسیم کرد:

۱) روشهای تک کاناله

۲) روشهای چند کاناله

در روشهای تک کاناله تنها یک میکروفون ورودی در دسترس بوده و اندازه گیری مشخصات نویز می باید در بخشهای غیرگفتار (سکوت) همان کانال انجام گیرد. اساس کار این روشها بر فرض ایستان بودن موضعی نویز می باشد. پس عملکرد این روشها در مواجهه با نویزهای با دینامیک زمانی زیاد دچار خدشه خواهد شد.

در روشهای چند کاناله، دو یا چند گیرنده در ورودی سیستم بهسازی گفتار مورد استفاده قرار می گیرد. مسلّماً افزایش تعداد میکروفون ورودی بر توانایی سیستم در حذف نویز خواهد افزود، ولی در مقابل بر هزینه و پیچیدگی سیستم هم اضافه خواهد شد.

تاکنون روشهای چند کانالۀ زیادی در بحث بهسازی گفتار مطرح شده که یکی از مهمترین این روشها، روش دو کاناله فیلتر وفقی می باشد. در این روش، یک ورودی به عنوان نویز مرجع و دیگری به عنوان سیگنال آغشته به نویز مورد استفاده قرار می گیرد. همانطور که انتظار می رود کار فیلتر وفقی در این سیستم، تخمین نویز همراه با سیگنال از روی نویز مرجع می باشد. با کم کردن خروجی فیلتر وفقی از سیگنال اصلی به گفتار تمیز می رسیم.

وجود ورودی مرجع، امکان مواجهه با نا ایستانی نویز را فراهم می آورد. به علاوه اینکه وجود کانال مرجع، نیاز به دانستن اطلاعات اولیۀ موجود در مورد سیگنالهای گفتار و نویز را مرتفع می سازد. از اینرو در محیطهای با نویز بسیار زیاد که اطلاعات اولیۀ موجود در مورد سیگنالهای

گفتار و نویز چندان معتبر نیست، استفاده از روش فیلتر وفقی به نتایج کاملاً برتر منجر می شود. عملکرد این سیستم وفقی تا وقتی که ورودی نویز مرجع فقط نویز موجود در محیط باشد بسیار خوب بوده و میزان SNR خروجی این سیستم بسیار مطلوب می باشد.

همانگونه که اشاره شد فیلتر وفقی با استفاده از همبستگی میان نویز مرجع و نویز همراه سیگنال، تخمینی از نویز سیگنال ارائه می دهد. قاعدتاً برای اینکه این دو نویز همبسته باشند باید فاصلهٔ میان میکروفون مرجع و میکروفون سیگنال کم باشد. تحقّق این موضوع موجب پدیدار شدن مشکل دیگری شده که از آن به نام گفتار نشتی[1] یاد می شود زیرا کم کردن فاصلهٔ میان دو میکروفون باعث نشت قسمتی از گفتار در کانال مرجع می شود و با اعمال پردازش وفقی روی سیگنال، این گفتار نشتی از گفتار اصلی کم خواهد شد.

پس در اینجا با دو مشکل که بر طرف کردن هریک موجب بروز دیگری می شود روبرو هستیم. یکی مسألهٔ نویزهای ناهمبسته و دیگری وجود گفتار نشتی.

۱-۲) ساختار این پایان نامه

برای مواجهه با نویزهای ناهمبسته روشهای مختلفی پیشنهاد شده که از جملهٔ آنها استفاده از شکل دهنده های موج و روش GSC[2] را می توان نام برد. در فصل دوم در این مورد به تفصیل سخن خواهیم گفت.

برای جلوگیری از بروز مسألهٔ گفتار نشتی هم تاکنون راه حلهای مختلفی ارائه گردیده که از میان آنها می توان موارد زیر را ذکر کرد:

۱) روش شکل دهی موج[1] به عنوان مرحله پیش پردازش فیلتروفقی

1- Speech Leakage
2- Generalized Side-lob Canceller

۲) استفاده از دو فیلتر وفقی با ساختار ضربدری (CTRANC)[۲]

۳) جداسازی کور منابع (BSS)[۳]

در روش شکل دهی موج - همانگونه که در فصل سوّم بیان می شود- سعی بر آن است تا با اعمال تأخیرهای مناسب به سیگنال دریافتی تخمین درستی از منابع نویز و سیگنال و جهت تابش آنها به دست آید. در ساختار فیلترهای وفقی ضربدری - که در مورد آن در فصل چهارم بطور کامل بحث شده- سعی می شود یک فیلتر وفقی با استفاده از سیگنال دریافتی تخمین دقیقی از نویز مرجع برای فیلتر دوّم ایجاد کند و فیلتر دوّم هم از سیگنال دریافتی تخمینی از سیگنال گفتار به عنوان ورودی فیلتر اوّل فراهم نماید. در پایان کار، خروجی فیلتر اول نویز مرجع و خروجی فیلتر دوّم سیگنال بهسازی شدهٔ گفتار خواهد بود.

مطابق آنچه در فصل پنجم بیان می شود، در روش BSS هدف تخمین سیگنالهای منابع اصلی است حتی اگر آنها کاملاً مستقل از هم نباشند. در اینجا لفظ کور بودن به این دلیل استفاده می شود که هیچ اطلاع قبلی در مورد طبیعت منابع یا کانال در اختیار نیست.

در حالت کلّی روش BSS را به دو صورت می توان اجرا نمود؛ اگر منابع اوّلیه را مستقل در نظر بگیریم، هدف از اعمال BSS پیدا کردن تبدیلاتی است که اطمینان می دهد سیگنالهای خروجی تا حدّ ممکن مستقل هستند. امّا اگر منبع اوّلیه ناهمبسته باشند، BSS سعی دارد تا خروجیهایی را پیدا کند که ناهمبسته باشند. در فصل پنجم یکی از الگوریتمهای BSS بر پایهٔ ناهمبسته بودن خروجیها را که از آن بنام الگوریتم MAD[۴] نام برده می شود معرفی خواهیم کرد. علاوه براین با استفاده از این الگوریتم و همچنین استفاده از ساختار فیلترهای وفقی ضربدری، سیستمی را معرفی می کنیم که هدف آن حذف نویزهای زمینه می باشد.

1- Beam forming
2-Cross Talk Resistant Adaptive Noise Canceler
3- Blind Source Separation
4- Multple Adaptive Decorrelation

در ادامه پس از معرفی روش ANC در فصل دوم، به ترتیب در فصول بعدی هریک از راه حلهای بیان شده را بررسی خواهیم نمود.

فصل دوّم:

حذف کنندهٔ وفقی نویز

۲-۱) حذف کنندهٔ وفقی نویز

حذف کننده های وفقی نویز (ANC[1]) – همانگونه که از نام آنها پیداست- سعی دارند تا با استفاده از یک عملکرد وفقی، نویز را از سیگنال دریافتی به گونه ای کم کنند تا نسبت سیگنال به نویز خروجی سیستم بهبود یابد. البته کم کردن مستقیم نویز از گفتار بدون آنکه هیچگونه پردازشی روی آن انجام گیرد، کمی غیر منطقی به نظر می رسد، حتی ممکن است این کار سبب افزایش توان متوسط نویز خروجی هم شود. اما هنگامی که این عمل تفریق، توسط یک فرآیند وفقی کنترل شود به دست آوردن یک سیستم حذف نویز، معقول خواهد بود [۴۸].

اولین فعالیت تحقیقاتی انجام شده درزمینه حذف نویز با استفاده از فیلتر وفقی به کارهای Widrow و همکارانش در دانشگاه استنفورد و نیز به کارهای Kelly در آزمایشگاه بل بر می گردد. Widrow در مقالهٔ معروف خود در سال ۱۹۷۵ [۴۸] بکارگیری فیلتر وفقی برای حذف نویز ۶۰ هرتز برق شهر از روی خروجی یک سیستم اندازه گیری و ثبت سیگنال قلب را گزارش نمود.

به طور عمده یک حذف کننده وفقی نویز یک سیستم دو ورودی با فیدبک حلقه بسته وفقی مطابق شکل (۲-۱) می باشد.

دو ورودی سیستم از یک جفت سنسور به نامهای سنسور اصلی[2] و سنسور مرجع[3] بدست می آیند که در حالت کلی بصورت زیر توصیف می شوند:

- الف) سنسور اصلی سیگنال اطلاعات $s(n)$ را که همراه با نویز جمع شوندهٔ $v_0(n)$ است به صورت زیر دریافت می کند:

1- Adaptive Noise Canceler
2- primary
3-reference