دانشکده فنی و مهندسی

پایان‌نامه دوره کارشناسی ارشد مهندسی برق–الکترونیک

کاهش شنوایی نویز محیط‌های صنعتی

حمزه معظمی گودرزی

استاد راهنما:

دکتر سعید سیدطبایی

استاد مشاور:

دکتر حامد ساجدی

تابستان ۱۳۸۸

دانشکده فنی و مهندسی

صورت جلسه هیئت داوران رساله کارشناسی ارشد

جلسه دفاعیه پروژه کارشناسی ارشد مربوط به آقای حمزه معظمی گودرزی به شماره دانشجویی ۸۵۷۵۱۴۵۰۲ در رشته الکترونیک در روز ۸۸/۶/۳۱ در دانشکده فنی و مهندسی با حضور افراد ذیل تشکیل شد، نتیجه به قرار زیر است :

☑ پروژه نامبرده با نمره ۱۹٫۲۵ قابل قبول می‌باشد.

☐ پروژه نامبرده مردود می‌باشد.

☐ پروژه نامبرده به شرط انجام اصلاحات جزیی قابل قبول می‌باشد. نمرا دانشجو متعاقباً اعلام می‌شود.

☐ نام استاد راهنمای اول		دانشگاه :	امضاء
☐ نام استاد راهنمای دوم		دانشگاه :	امضاء
☐ نام استاد مشاور اول		دانشگاه :	امضاء
☐ نام استاد مشاور دوم		دانشگاه :	امضاء
☐ نام داور اول		دانشگاه :	امضاء
☐ نام داور دوم		دانشگاه :	امضاء
☐ نام داور سوم		دانشگاه :	امضاء
☐ نام داور چهارم		دانشگاه :	امضاء
☐ نام نماینده محترم پژوهشی			امضاء

تقدیم به

پدر و مادرم،

که همواره به عنوان دو شمع سوزان، روشن کنندهٔ راه زندگی و

پیشرفت من هستند.

و به همه کسانی که به من آموختند.

تشکر و قدردانی

بدینوسیله مراتب احترام و تشکر خود را از تمامی عزیزانی که مرا در تهیه این پایان‌نامه یاری کرده‌اند، به‌ویژه جناب آقای **دکتر سعیدسیدطبایی**، استاد راهنما و جناب آقای **دکتر حامد ساجدی**، استاد مشاور اعلام میدارم.

همچنین از اساتید بزرگوار آقایان **دکتر حمید شیخ‌زاده نجار** و **دکتر منصور ولی** که داوری این رساله را تقبل فرمودند سپاسگزارم و از خداوند متعال توفیق روز افزون آنها را خواستارم.

حمزه معظمی گودرزی

تابستان ۱۳۸۸

چکیده

یکی از موضوعات مهم پردازش سیگنال، کاهش و حذف نویز ناخواسته از سیگنال اصلی و ارتقاء آن است. زندگی پر از صداهاست و ما همیشه طالب شنیدن صداهای خوش و حیاتی هستیم و از صداهای نامطبوع و خطرناک گریزانیم. اغلب در محیط‌های صنعتی، کارگران در معرض سر و صدای شدید دستگاههایی از قبیل، دستگاههای پرس، برش، فرز، رنده و ... قرار می‌گیرند که در دراز مدت آسیب‌های جدی بر سلامتی آنها وارد می‌سازد و همچنین باعث افت کارایی و کیفیت عملکرد آنها در ساعات کاری می‌شود. تلاشها برای برطرف کردن این مشکل در حال انجام است. یکی از روش‌ها، استفاده از هدفون‌های الکترونیکی کاهش دهنده نویز و ارتقاء گفتار می‌باشد. تکنیک‌های ارتقاء می‌توانند به دو گروه تک‌کاناله و چندکاناله تقسیم شوند. سیستم‌های تک‌کاناله، متداولترین نوع الگوریتم‌های بلادرنگ می‌باشد چراکه به سادگی قابل پیاده‌سازی هستند و به طور نسبی هزینه آنها کمتر از سیستمهایی با چند کانال ورودی می‌باشد. در میان روش‌های تک‌کاناله حذف نویز، روش تفریق طیفی بدلایل متعددی از جمله توانایی بالا در حذف نویز و سرعت بالای الگوریتم از اهمیت زیادی برخوردار هستند. مهمترین چالش در روش تفریق طیفی، ایجاد اعوجاج و ایجاد نویز موسیقی در سیگنال پردازش‌شده می‌باشد که بدلیل تخمین نادقیق طیف نویز به‌وجود می‌آید. مهمترین وظیفه الگوریتم‌های ارتقاء گفتار تخمین طیف نویز می‌باشد.

در این پایان‌نامه از روش تفریق طیفی برای کاهش و حذف شنوایی نویز بشدت غیرایستای محیط‌های صنعتی از روی سیگنال گفتار در محیط کارخانه استفاده شد. یک راهکار بهینه مبتنی بر کمترین آمارگان‌های وزن‌دار شده چندباندی برای تخمین این نویزها پیشنهاد شد و عملکرد آن مورد آزمایش قرار گرفت. مشخص شد که الگوریتم پیشنهادی قادر است به سرعت تغیرات طیف نویز را دنبال کند و سیگنال پردازش شده دارای حداقل اعوجاج و نویز موسیقی می‌باشد. این گفته‌ها با هر دو آزمایشات علمی(روابط ریاضی) و ادراکی(نظر شخص انسانی) تایید می‌شوند.

کلید واژه: ارتقاء گفتار، نویز غیرایستای صنعتی، تفریق طیفی، تخمین طیف نویز، مینیمم آمارگانها.

فهرست مطالب

عنوان **صفحه**

فهرست جدول‌ها ..... ج
فهرست شکل‌ها ..... د

فصل ۱- مقدمه ۱
۱-۱- تعریف مسئله ۱
۱-۲- راهکارهای موجود ۳
۱-۳- خلاصه کارهای انجام گرفته ۴
۱-۴- هدف از این مطالعه ۶
۱-۵- فصل بندی این پایان‌نامه ۶

فصل ۲- روشهای مختلف ارتقاء گفتار و تخمین طیف نویز ۸
۲-۱- پیش زمینه‌های تئوری سیگنال صحبت ۸
۲-۱-۱- دستگاه شنوایی انسان ۸
۲-۱-۲- ویژگیهای امواج صوتی ۱۲
۲-۲- دسته بندی روشهای ارتقاء گفتار ۱۴
۲-۲-۱- کاهش نویز با استفاده از خاصیت پریودیک بودن گفتار یا نویز ۱۷
۲-۲-۲- ارتقاء گفتار مبتنی بر مدل‌سازی ۱۸
۲-۲-۳- ارتقاء گفتار مبتنی بر ضوابط و معیارهای ادراکی ۱۹
۲-۲-۴- ارتقاء گفتار مبتنی بر تخمین کوتاه-مدت دامنه طیفی ۲۰
۲-۳- روش تفریق طیفی ۲۱
۲-۳-۱- اصول و روابط تکنیک تفریق طیفی ۲۱
۲-۳-۲- روش تفریق طیفی به عنوان یک فیلتر: ارتباط آن با فیلترینگ وینر ۲۳
۲-۳-۳- محدودیت‌های تفریق طیفی ۲۴
۲-۳-۴- اصلاح تفریق طیفی و روشهای مربوطه دیگر ۲۶
۲-۳-۴-۱- تفریق طیفی با استفاده از فاکتور فوق تفریق و فاکتور کف طیفی ۲۷
۲-۳-۴-۲- تفریق طیفی غیرخطی ۲۸
۲-۳-۴-۳- تفریق طیفی توسعه یافته ۲۹
۲-۴- الگوریتم‌های تخمین طیف نویز ۳۱

فصل ۳- الگوریتم تخمین نویز پیشنهادی ۳۸
۳-۱- مقدمه ۳۸
۳-۲- آنالیز طیفی و شکل موج نویز فرز ۳۹
۳-۳- تشریح الگوریتم تخمین نویز پیشنهادی ۴۱

۳-۴- مروری بر سیستم پیشنهادی ۴۴
فصل ۴- پیاده سازی و ارزیابی کارایی ۴۶
۴-۱- جزئیات پیاده‌سازی ۴۶
۴-۲- روشهای ارزیابی ریاضی ۵۰
۴-۲-۱- نسبت سیگنال به نویز(SNR) ۵۱
۴-۲-۲- اندازه‌گیری ایتاکورا-سایتو(IS) ۵۱
۴-۲-۳- شیب طیفی وزندار(WSS) ۵۲
۴-۲-۴- اندازه‌گیری اعوجاج طیفی Bark اصلاح‌شده(MBSD) ۵۲
۴-۳- ارزیابی کارایی ۵۳
۴-۴- نتیجه‌گیری و ارائه پیشنهادات ۶۰
۴-۴-۱- نتیجه‌گیری ۶۰
۴-۴-۲- پیشنهادات ۶۰
فهرست مراجع ۶۲
واژه نامه فارسی به انگلیسی ۶۷
واژه نامه انگلیسی به فارسی ۶۸

فهرست جدول‌ها

عنوان **صفحه**

جدول ۴-۱: نتایج حاصل شده از میانگین‌گیری مقادیر بدست آمده از معیارهای SNR، BSD،IS و WSS. ۵۸
جدول ۴-۲: نحوه ارزیابی ادراکی ۵۸
جدول ۴-۳: نتایج ارزیابی ادراکی الگوریتم پیشنهادی ۵۹

فهرست شکل‌ها

عنوان **صفحه**

شکل ۱-۱: نویز جمع‌شونده پشت‌زمینه. ۲
شکل ۱-۲: یک هدفون کاهش نویز الکتریکی. ۲
شکل ۱-۳: طرح اولیه از یک سیستم حذف نویز الکتریکی. ۳
شکل ۱-۴: تداخل ویرانگر. ۴
شکل ۲-۱: نمودار کارکردی گوش انسان[۹]. ۱۰
شکل ۲-۲: واحدهای شدت صدا [۹]. ۱۰
شکل ۲-۳: تشخیص فاز توسط گوش انسان[۹]. ۱۳
شکل ۲-۴: موج صوتی ویولن. ویولن موج دندانه‌اره‌ای ایجاد می‌کند(شکل a)، صدای دریافت شده شامل فرکانس اساسی و همسازهای آن است(شکل b)[۹]. ۱۳
شکل ۲-۵: یک سیستم حذف کننده وفقی نویز با یک میکروفن مرجع ۱۶
شکل ۲-۶: بلوک دیاگرام کلی یک سیستم تفریق طیفی. ۲۰
شکل ۲-۷: خصوصیات نویز موسیقی در حوزه‌ی فرکانس برای ۳۰ فریم[۲۴]. ۲۵
شکل ۲-۸: ساختار اولیه تفریق طیفی توسعه یافته. ۲۹
شکل ۲-۹:نمایش بخش گفتاری و بخش غیرگفتاری یک نمونه شکل موج صوتی[۷]. ۳۱
شکل ۲-۱۰: بلوک دیاگرام یک سیستم تفریق طیفی با استفاده از تکنیک VAD برای تخمین طیف نویز. ۳۲
شکل ۳-۱: شکل موج نویز فرز در حالت درگیری. ۴۰
شکل ۳-۲: طیف فرکانسی نویز فرز در حالت درگیری. ۴۰
شکل ۳-۳: طیف توان هموارسازی‌شده و کف نویز تخمین‌زده شده[۵۲]. ۴۱
شکل ۳-۴: انرژی نویز در هر زیرباند[۶۳]. ۴۳
شکل ۴-۱: شکل موج و اسپکتوگرام یک نمونه صدای گفتاری که توسط یک خانم ادا شده. ۴۷
شکل ۴-۲: جمله خراب‌شده در SNR ۵dB ۴۷
شکل ۴-۳: جمله خراب‌شده در SNR ۰dB ۴۸
شکل ۴-۴: به ترتیب شکل موج یک نمونه گفتار خراب‌شده با نویز سفید گوسین با SNR ۵ dB ، و سیگنال ارتقاء یافته توسط الگوریتم‌های Beruti، Kamath، Martin و الگوریتم پیشنهادی. ۵۵
شکل ۴-۵: به ترتیب اسپکتوگرام یک نمونه گفتار خراب‌شده با نویز سفید گوسین با SNR ۵ dB ، و سیگنال ارتقاء یافته توسط الگوریتم‌های Beruti، Kamath، Martin و الگوریتم پیشنهادی. ۵۵

شکل ۴-۶: به ترتیب شکل موج یک نمونه گفتار خراب‌شده با نویز فرز در حالت غیردرگیری با SNR=۵dB و سیگنال ارتقاء یافته توسط الگوریتم‌های Beruti، Kamath، Martin و الگوریتم پیشنهادی. ۵۶
شکل ۴-۷: به ترتیب اسپکتوگرام یک نمونه گفتار خراب‌شده با نویز فرز در حالت غیردرگیری با SNR=۵dB و سیگنال ارتقاء یافته توسط الگوریتم‌های Beruti، Kamath، Martin و الگوریتم پیشنهادی. ۵۶
شکل ۴-۸: به ترتیب شکل موج حاصل از سیگنال تمییز، سیگنال نویزی، سیگنال ارتقاءیافته توسط الگوریتم Martin و سیگنال ارتقاءیافته توسط الگوریتم پیشنهادی. ۵۷
شکل ۴-۹: به ترتیب اسپکتوگرام حاصل از سیگنال تمییز، سیگنال نویزی، سیگنال ارتقاءیافته توسط الگوریتم Martin و سیگنال ارتقاءیافته توسط الگوریتم پیشنهادی. ۵۷
شکل ۴-۱۰: مقایسه ادراکی کیفیت سیگنال صحبت حذف نویزشده توسط الگوریتم Martin و پیشنهادی. ۵۹

فصل ۱- مقدمه

۱-۱- تعریف مسئله

نویز صوتی از معضلات دنیای امروزی است. گسترش تکنولوژی و صنعت باعث ایجاد صداهایی با شدت زیاد در اطراف ما شده است. اگر شدت صدا از حد مجاز خود بالاتر برود، به افرادی که در مواجهه با آن هستند آسیب وارد می شود. البته شدت این آسیب با توجه به شدت صدا و مدت زمان مواجهه با آن متغیر است. صدای بلند، چه موسیقی باشد و چه صدایی آزاردهنده در صنعت، می تواند بر شنوایی ما تأثیرات نامطلوبی داشته باشد.

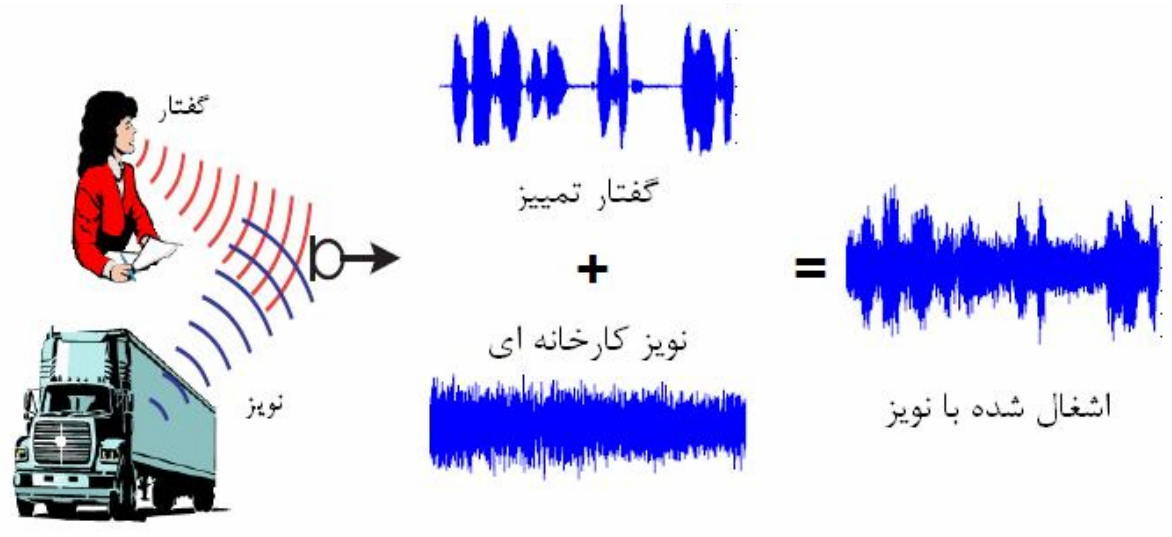
شدت صدا را با دسی بل می‌سنجند، به عنوان مثال، شدت صدای ناشی از پرواز یک هواپیمای جت از فاصله ۱۰۰ متری برابر با ۱۳۰ دسی بل است. شدت صدا در یک اتاق ساکت و خلوت ۲۰ دسی‌بل است.

بر اساس استانداردهای کمیته فنی بهداشت حرفه‌ای کشور مقدار حد مجاز تماس شغلی صدا برای ۸ ساعت کار روزانه(یک شیفت کار)، برابر با ۸۵ دسی بل است که به ازای هر ۳ دسی بل افزایش صدا، زمان کار فرد در آن محیط باید به نصف، کاهش یابد تا از آثار سوء صدا بر سلامتی فرد جلوگیری شود. (به ازای هر ۳ دسی‌بل افزایش شدت صوت، میزان فشاری که به پرده صماخ گوش وارد می‌شود، دو برابر می‌شود). مثلا اگر کارگری در محیط کار خود در معرض تماس با صدایی معادل ۸۸ دسی‌بل باشد، زمان مجاز کار او ۴ ساعت است و اگر میزان صدا به ۹۱ دسی‌بل برسد فقط مجاز به ۲ ساعت مواجهه با صدا خواهد بود.

آلودگی صوتی علاوه براینکه به طور مستقیم تاثیرات غیرقابل برگشتی بر روی سیستم شنوایی انسان وارد می‌کند، می‌تواند بر روی بسیاری از اندام‌های دیگر بدن نظیر سیستم قلب و عروق، دستگاه گوارش، اعصاب و روان، ایجاد عوارضی به صورت سردرد، پرخاشگری، اضطراب، تغییرات در خلق و خو، افزایش فشار خون و ... شود. سر و صدا و نویز بر سلامتی ما تاثیر می‌گذارد، در ارتباطات ما دخالت دارد و بر لذت بردن ما از زندگی اثر می‌گذارد(شکل ۱-۱).

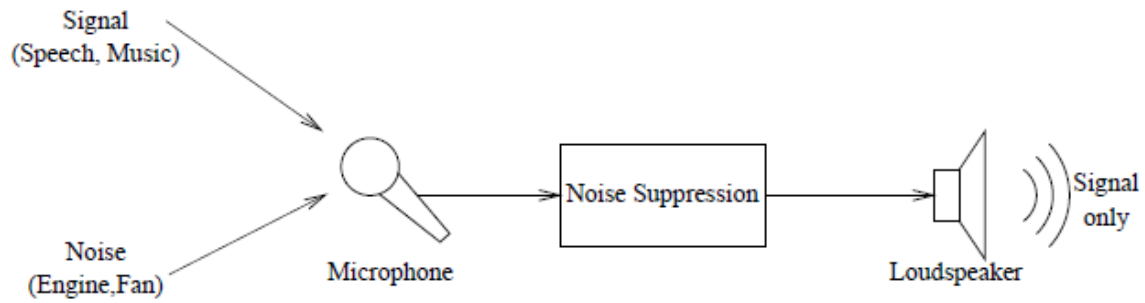
اغلب در محیط‌های صنعتی، کارگران در معرض سر و صدای شدید دستگاههایی از قبیل، دستگاههای پرس، برش، فرز، رنده و ... قرار می‌گیرند که در دراز مدت آسیب‌های جدی بر سلامتی آنها وارد می‌سازد و همچنین باعث افت کارایی و کیفیت عملکرد آنها در ساعات کاری می‌شود. در بعضی از محیط‌های کاری استفاده از گوشی حفاظتی مناسب نیست چراکه اولا این گوشیها فقط در حذف نویزهای فرکانس بالا موثر بوده و حال آنکه اکثر نویزهای صنعتی، فرکانس هارمونیک‌های اصلی زیر ۸۰۰ Hz دارند. دوما، این گوشی‌ها هر دو/ سیگنال مهم و غیر مهم را حذف می‌کنند درحالیکه بعضی صداهای هشداردهنده، صداهای گفتاری و ... برای ما اهمیت دارند و نباید از شنیدن آنها جلوگیری شود.

برای حل این مشکل، استفاده از هدفون‌های الکترونیکی کاهش نویز پیشنهاد می‌شود(شکل ۱-۲) به این صورت که در قسمت بیرونی گوشی‌ها، میکروفنی تعبیه شده که سیگنال اکوستیکی آلوده به نویز را به یک سیگنال الکتریکی تبدیل کرده و با استفاده از پردازشگر داخل گوشی، عمل جداسازی نویز از سیگنال، انجام شده و سپس سیگنال توسط یک بلندگو در داخل گوشی پخش شود(شکل ۱-۳).



شکل ۱-۱: نویز جمع‌شونده پشت‌زمینه.

شکل ۱-۲: یک هدفون کاهش نویز الکتریکی.



شکل ۱-۳: طرح اولیه از یک سیستم حذف نویز الکتریکی.

۱-۲- راهکارهای موجود

عمل حذف نویز و یا کاهش آن از سیگنال گفتاری آلوده به نویز به دو صورت مورد توجه است[۱۷]:

۱- روشهای کنترل فعال نویز (ANC)[۱]

۲- روشهای ارتقاء گفتار (SE)[۲]

کنترل فعال نویز سعی در ایجاد سکوت در نقاط مشخص می‌کند. روش‌های متداول حذف نویز اکوستیکی همانند استفاده از مواد جذب کننده صدا، عموما در فرکانس‌های پایین قابل استفاده نیستند، چون در این فرکانس‌ها طول موج صدا بزرگ بوده و به تناسب آن، ابعاد و ضخامت مواد جذب کننده و خفه کننده صدا باید بزرگ انتخاب شوند. کنترل فعال نویز(ANC) شامل یک سیستم الکتریکی-آوایی یا الکتریکی-مکانیکی است، که نویز ناخواسته(اولیه) را بر مبنای اصل جمع آثار به طور فیزیکی حذف می‌کند. مطابق با شکل ۱-۴، یک موج ضد نویز (نویز ثانویه) با دامنه برابر و فاز مخالف تولید شده و پس از پخش توسط یک بلندگو، به‌طور اکوستیکی با نویز اولیه ترکیب می‌شود و منجر به حذف هر دو نویز می‌شود(تداخل ویرانگر) که درنتیجه انرژی صوت به گرما تبدیل خواهد شود. تمام پروسه بستگی به سرعت انتشار نسبتا کند موج صدا، در مقایسه با پردازش سریع سیگنال الکتریکی دارد. چون در این صورت است که قبل از اینکه نویز ناخواسته در میکروفن به بلندگوی حذف کننده برسد، شکل موج معکوس می‌تواند ایجاد شود. از آنجاییکه ویژگیهای منبع نویز آوایی و محیط، محتویات فرکانسی، دامنه، فاز و تندی صدا(بر حسب زمان) متغیر با زمان است، لذا نویز نامطلوب غیر ایستا است. یک سیستم ANC باید به طور وفقی باشد تا با این تغییرات سازگار شود. فیلترهای وفقی، ضرایبشان را تنظیم می‌کنند تا سیگنال خطا را مینیمم کنند و می‌توانند به عنوان فیلترهای پاسخ ضربه محدود(FIR)[۳] و پاسخ ضربه نامحدود(IIR)[۴] استفاده شوند سیستم‌های ANC به دو گروه کنترل به روش پیشخور[۵] (جلوسو) و

---

[۱] Active Noise Control

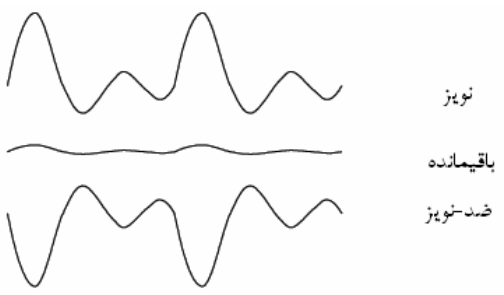
[۲] Speech Enhancement

[۳] Finite Impulse Response

[۴] Infinite Impulse Response

[۵] Feed forward

پسخور[1](فیدبک) تقسیم می‌شوند. عیب اصلی این روش این است که عمل حذف نویز تنها در یک ناحیه خاص از فضا انجام می‌گیرد و عملکرد آن را در کاربردهای خاص محدود می‌کند ضمن اینکه تاکنون هیچ کس ادعای حل کامل آن را ندارد.



شکل ۱-۴: تداخل ویرانگر.

هدف از ارتقاء گفتار(Speech Enhancement) حذف و یا کاهش نویز اکوستیکی از ترکیب سیگنال-به اضافه-نویز و رسیدن به سیگنال مطلوب می‌باشد. چندین راه برای دسته‌بندی انواع روشهای ارتقاء گفتار وجود دارد. بعضی از تکنیک‌ها بر مدل ریاضی سیستم تولیدگفتار تکیه دارد، و به عنوان مدل‌های پارامتریک نام برده می‌شوند. دسته دیگری از روش‌ها، روش‌های غیر پارامتریک هستند که هیچ فرضی درباره‌ی سیستم تولید گفتار انجام نمی‌دهند، اما بر تخمین نویز مداخله‌گر تکیه دارند. روش‌های ارتقاء گفتار می‌تواند به راهکار تک‌کاناله و چندکاناله، آماری و غیر آماری، در حوزه زمان یا فرکانس، وفقی و غیروفقی نیز دسته بندی شوند.

یک سیستم ارتقاء گفتار معمولا مبتنی بر فرضیات و محدودیت‌های خاص وابسته به کاربرد و محیط، استوار است. به طور کلی عملکرد الگوریتم‌های ارتقاء گفتار می‌تواند توسط عوامل ذکر شده در زیر محدود شود: محدودیت در اختیار داشتن اطلاعاتی از ویژگی‌های منابع نویز که می‌تواند اطلاعات پیشینی در مورد سیگنال آلوده‌شده به نویز را برای ما میسر کند، محدودیتی از قبیل سرعت تغییرات زمانی(غیرایستایی) موجود در سیگنال آلوده‌شده با نویز، محدودیتی مبتنی بر مدل‌سازی سیگنال و یا نویز مانند محدودیت الگوریتم برای مدل‌سازی نویزهای ناهمبسته[2]، محدودیتی مبتنی بر در اختیار داشتن سیگنال آلوده به نویز با یک نسبت سیگنال به نویز پایین و غیره.

ما در این پایان نامه سعی در کاهش و حذف شنوایی نویز محیط‌های صنعتی از صدای گفتاری و سیگنال‌های مهم(ارتقاء گفتار) داریم.

۱-۳- خلاصه کارهای انجام گرفته

بازیابی سیگنال گفتاری مطلوب از ترکیب گفتار و نویز پس‌زمینه، از جمله قدیمی‌ترین اهداف محققان پردازش سیگنال و سیستم‌های ارتباطی می‌باشد که تا به امروز ادامه داشته است.

---

[1] Feed back

[2] Uncorrelated

هدف اصلی ارتقاء گفتار کاهش دادن اثرات نویز، بالابردن کیفیت و قابل فهم بودن سیگنال گفتاری می‌باشد. این در کاربردهای مختلفی از قبیل در محیط‌هایی با نویزهای مداخله‌گر پس‌زمینه(برای مثال در دفاتر کار، خیابان‌ها، اتومبیل‌ها، کارخانه‌ها و غیره) و در سیستم‌های بازشناسی گفتار، محیط‌های هندفری برای ماشین‌ها، وسایل کمک‌شنیداری و غیره از اهمیت بالایی برخوردار می‌باشد.
در طول این سالها، محققان و مهندسان چندین روش برای برطرف کردن این مشکل پیشنهاد کرده‌اند ولی تاکنون به‌خاطر پیچیدگی سیگنال گفتاری، این ناحیه از تحقیقات با مشکلاتی همراه بوده‌است. به طور معمول کاهش نویز بدون ایجاد اعوجاج در سیگنال گفتار به سختی امکان‌پذیر است و بنابراین عملکرد سیستم‌های ارتقاء گفتار به ارتباطی میان اعوجاج گفتاری و کاهش نویز محدود می‌شود.

با توجه به [۱] ، [۲] ، [۳] ، [۴] و [۳۳]، سیستم‌های ارتقاء گفتار می‌توانند بر مبنای تعداد کانال‌های ورودی(یک/دو/چند)، حوزه پردازشی(زمانی/فرکانسی)، نوع الگوریتم(وفقی/غیر وفقی) تقسیم‌بندی شوند. به طور نمونه در بیشتر متون ارتقاء گفتاری، استراتژی‌های مختلف پردازش گفتاری را به دو گروه وسیع تکنیک‌های تک کاناله و چند کاناله تقسیم‌بندی کرده‌اند. به طور معمول در بیشتر موقعیت‌ها نویز و گفتار در یک کانال حضور دارند(سیستم‌های تک کاناله) و این یکی از مشکل‌ترین موقعیت‌ها برای ارتقاء گفتار می‌باشد. پیچیدگی و سادگی در به‌کارگیری هر یک از این طرح‌ها فاکتور مهم دیگری است چراکه اکثر الگوریتم‌های ارتقاء گفتار و کاهش نویز در کاربردهای بلادرنگ[۱] و قابل حمل[۲] از جمله تلفن‌های سلولی، وسایل کمک شنیداری، کیت‌های هندفری، گوشی‌های حذف نویز و ... به کار گرفته می‌شوند. همچنین هزینه‌ی کم برای به کارگیری این روشها از جمله عوامل انتخاب آنها برای کاربردهای مورد نظر می‌باشد.

روش تفریق طیفی، یکی از شناخته‌ترین تکنیک‌های تک کاناله برای کاهش نویز بوده‌است [۵]، [۶] و[۱]. تفریق طیفی، طیف توان گفتار تمییز را از تفاضل طیف توان نویز از طیف توان گفتار نویزی، تخمین می‌زند. این روش به خاطر پیچیدگی کم و به کارگیری نسبتا ساده، توجه بسیار زیادی را در طی سالهای گذشته به خود اختصاص داده است. این روش به طور کلی یک نویز باقی‌مانده که معمولا نویز موسیقی نامیده می‌شود، تولید می‌کند. این مشکلات به دلیل تخمین نادقیق نویز در فریم موردنظر و خطای ناشی از اختلاف سیگنال تخمین زده شده و سیگنال اصلی می‌باشد.
تخمین طیف نویز یکی از چالش‌برانگیزترین وظایف الگوریتم‌های ارتقاء گفتار تک کاناله می‌باشد چراکه ما تنها به سیگنال گفتار نویزی دسترسی داریم و هیچ سیگنال مرجعی برای نویز در دسترس نمی‌باشد و گفتار تمییز نمی‌تواند قبل از اینکه با نویز آمیخته شود، پیش‌پردازش شود. برای تخمین نویزهای ایستا، اغلب میانگین‌گیری از طیف سیگنال نویزی در طی اولین پریود سکوت(غیرگفتاری) کافی می‌باشد[۷]. برای تخمین نویزهای غیرایستا، نیاز است که طیف نویز به طور پیوسته تخمین‌زده شود و به‌هنگام شود. این کار بسیار سخت است چراکه ما هیچ اطلاعات پیشین از نویز در اختیار نداریم. برای این کار نیاز به یک آشکارساز فعال گفتار/ سکوت(VAD)[۳] می‌باشد تا وقفه‌های گفتاری را پیدا کرده و تخمین نویز را

[۱] Real-Time
[۲] Portabale
[۳] Voice Activity Detector

به‌هنگام کند[۸]. از آنجاییکه برای محیط‌های نویزی بشدت غیرایستا و یا محیط‌هایی با SNR پایین، تشخیص مطمئن بخشهای غیر گفتاری خیلی مشکل می‌باشد، تکنیکهای موثر دیگری که نیاز به یک الگوریتم جداسازی گفتار / سکوت واضح ندارند و حتی در بخش‌های گفتاری قادر به به‌هنگام تخمین نویز هستند، پیشنهاد شده‌اند [۶۲-۵۲].

۱-۴- هدف از این مطالعه

هدف از این پایان‌نامه، کاهش و حذف شنوایی نویز محیط‌های صنعتی از سیگنال گفتار در محیط کارخانه با استفاده از روش تفریق طیفی می‌باشد. این نویزها اغلب شامل یک بخش باند پهن و تعداد محدودی از سیگنال‌های سینوسی غالب با بیشترین انرژی نویز در فرکانس‌های پایین می‌باشد که به شدت کیفیت سیگنال گفتاری را پایین می‌آورد.

ما یک راهکار بهینه مبتنی بر کمترین آمارگانهای چندباندی برای تخمین نویزهای به‌شدت غیرایستای صنعتی پیشنهاد می‌کنیم که به خوبی نویز باند پهن و هارمونیک‌های اصلی فرکانس پایین غیر ایستای آنها را حذف می‌کند و سیگنال پردازش شده دارای حداقل اعوجاج و نویز موسیقی خواهد شد. روش پیشنهادی نیازی به آشکارساز گفتار/ سکوت نداشته و تخمین نویز را حتی در بخش‌های گفتاری به‌هنگام می‌کند. این روش در مقایسه با سایر روش‌های مرسوم تخمین طیف نویز از کارایی بسیار بالاتری برخوردار است. برای ارزیابی عملکرد این الگوریتم در نسبت‌های سیگنال-به- نویز (SNR)[۱] مختلف، ما به طور دیجیتالی نویز نوعی فرز را با سیگنال‌های گفتاری با استفاده از نرم‌افزارMatlab، جمع کرده و روش پیشنهاد شده را بر روی آن اعمال می‌کنیم و عملکرد آن را با ارزیابی‌های ریاضی[۲] و انسانی[۳] نشان می‌دهیم.

۱-۵- فصل بندی این پایان‌نامه

در فصل ۲ خلاصه‌ای از روش‌های مختلف ارتقاء گفتار با تاکید بر روش تک کاناله تفریق طیفی بیان می‌شود. همچنین محدودیت‌ها و روشهای اصلاح تکنیک تفریق طیفی آورده شده است و در آخر روشهای مختلف تخمین طیف توان نویز به عنوان مهمترین وظیفه الگوریتم‌های تک کاناله ارتقاء گفتار مورد بحث قرار می‌گیرد. در فصل ۳ الگوریتم پیشنهادی ارائه می‌شود. در فصل ۴ پیاده‌سازی الگوریتم و نتایج آزمایشگاهی آورده می شود و نهایتا نتیجه گیری خواهیم کرد.

---

[۱] Signal-to-Noise Ratio

[۲] Objective

[۳] Subjective

فصل دوم

روشهای مختلف ارتقاء گفتار و تخمین طیف نویز

فصل ۲- روشهای مختلف ارتقاء گفتار و تخمین طیف نویز

۲-۱- پیش زمینه‌های تئوری سیگنال صحبت

۲-۱-۱- دستگاه شنوایی انسان

گوش انسان یک عضو به گونه‌ای فزاینده پیچیده است[۹]. قضیه وقتی پیچیده‌تر می‌شود که اطلاعات ارسالی از دو گوش در یک شبکهٔ پیچیدهٔ گیج کننده که همانا مغز انسان باشد با هم ترکیب می‌شوند. به یاد داشته باشیم که بیان فوق یک گذر کلی بر قضیه است و تعداد زیادی از پدیده‌ها و آثار دقیق مرتبط با گوش انسان هنوز به درستی درک نشده‌اند.

شکل ۲-۱ قسمت اعظم ساختارها و پردازشهایی را که گوش انسان در بر دارد به تصویر می‌کشد. گوش خارجی از دو بخش تشکیل شده است: نرمی پوست قابل مشاهده و غضروف متصل به کنار سر و کانال گوش که لوله‌ایست به قطر تقریبی ۰،۵ سانتیمتر و تا حدود ۳ سانتیمتر در داخل سر فرو می‌رود. این ساختارها صداهای محیط را به بخشهای حساس گوش میانی و گوش داخلی که در درون استخوانهای جمجمه محافظت می‌شود راهبری می‌کنند. در انتهای کانال گوش یک ورقهٔ نازک از نسوج که پردهٔ صماخ[۱] یا طبل گوش نامیده می‌شود کشیده شده است. امواج صدا با برخورد به پردهٔ صماخ باعث لرزش آن می‌شوند. گوش میانی مجموعه‌ای از استخوانهای کوچک است که لرزش مزبور را به حلزون گوش[۲] (گوش داخلی) انتقال می‌دهند و در آنجا این لرزشها تبدیل به ضربه‌های عصبی می‌گردند. حلزون گوش یک لولهٔ پر از مایع است که به زحمت قطر آن به ۲ میلیمتر و طول آن به ۳ سانتیمتر می‌رسد. اگر چه حلزون گوش در شکل ۲-۱ به صورت یک لولهٔ مستقیم نشان داده شده اما در واقع به دور خودش همانند صدف حلزون پیچ خورده است و وجه تسمیهٔ آن که ریشه در کلمه‌ای یونانی به معنای حلزون دارد نیز این واقعیت است.

وقتی یک موج صوتی سعی دارد از هوا وارد مایع شود تنها کسر کوچکی از آن از بین دو محیط عبور می‌کند و باقیماندهٔ انرژی آن بازتابیده می‌شود. دلیل این امر مقاومت مکانیکی پایین هوا (ناشی از پایین بودن میزان فشار صوتی و سرعت بالای ذرات هوا که به نوبهٔ خود از چگالی پایین و تراکم‌پذیری بالای آنها نشأت می‌گیرد) در برابر مقاومت مکانیکی بالای مایع است. به عبارت ساده‌تر دلیل این امر مشابه

[۱] tympanic membrane

[۲] cochlea

دلیل این موضوع است که برای ایجاد موج با دست در درون آب به تلاش بیشتری به نسبت انجام این کار در هوا نیازمندیم. تفاوت موجود باعث بازتابش قسمت اعظم صوت در مرز هوا/مایع می‌گردد.

گوش میانی یک شبکهٔ تطبیق مقاومت[۱] است که کسر انرژی صوتی وارد شده به مایع گوش داخلی را زیاد می‌کند. برای نمونه ماهی پردهٔ صماخ یا گوش میانی ندارد چرا که نیازی به شنیدن در هوا ندارد. تغییر شدت، بیشتر ناشی از تفاوت مساحت پردهٔ صماخ (که صدا را از هوا دریافت می‌کند) و دریچه بیضوی (که مطابق شکل ۲-۱ صدا را به داخل مایع انتقال می‌دهد) می‌باشد. مساحت پردهٔ صماخ حدوداً ۶۰ میلیمتر مربع است حال آن که دریچهٔ بیضوی[۲] حدوداً ۴ میلیمتر مربع مساحت دارد. از آنجا که فشار برابر است با نسبت نیرو به مساحت، این تفاوت مساحت، فشار موج صدا را حدوداً ۱۵ برابر افزایش می‌دهد.

در داخل حلزون گوش پردهٔ اصلی[۳] قرار دارد که ساختاری را برای ۱۲۰۰۰ سلول حسی که شکل‌دهندهٔ عصب حلزونی است ایجاد می‌کند. پردهٔ اصلی در نزدیکی دریچهٔ بیضوی بسیار سفت است و در انتهای دیگر انعطاف‌پذیرتر است که این امر به این عضو کمک می‌کند تا به عنوان تحلیلگر طیف فرکانسی عمل کند. وقتی پردهٔ اصلی در معرض یک سیگنال با فرکانس بالا قرار می‌گیرد در قسمت سفت‌تر طنین می‌اندازد که سبب تحریک سلولهای عصبی نزدیک به دریچهٔ بیضوی می‌گردد. به همین ترتیب فرکانسهای پایین موجب تحریک انتهای دورتر پردهٔ اصلی می‌شوند. این امر موجب پاسخگویی رشته‌های خاص عصب حلزونی در برابر فرکانسهای خاص می‌گردد. این سازوکار اصل مکان[۴] نامیده می‌شود و در سراسر مسیر به سمت مغز حفظ می‌شود.

طرح کدگذاری اطلاعات دیگری نیز در شنوایی انسان به کار می‌رود که اصل رگبار[۵] نامیده می‌شود. سلولهای عصبی اطلاعات را با تولید پالسهای الکتریکی کوچکی که پتانسیل کنش[۶] نامیده می‌شوند انتقال می‌دهد. یک سلول عصبی واقع بر پردهٔ پایینی می‌تواند اطلاعات صوتی را با تولید یک پتانسیل کنش در پاسخ هر سیکل لرزش کدگذاری کند. برای نمونه یک موج صدای ۲۰۰ هرتزی می‌تواند توسط یک نورون ایجاد کنندهٔ ۲۰۰ پتانسیل کنش در ثانیه نشان داده شود. در هر صورت این روش تنها در فرکانسهای زیر حدوداً ۵۰۰ هرتز – بالاترین سرعت ممکن تولید پتانسیل کنش در نورونها – به کار می‌آید. گوش انسان برای غلبه بر این مشکل به نورونها اجازه می‌دهد که برای انجام این کار دسته‌جمعی عمل کنند. برای نمونه یک صدای ۳۰۰۰ هرتزی می‌تواند توسط ده سلول عصبی که هر کدام ۳۰۰ ضربه در ثانیه علامت می‌دهند نشان داده شود. این پدیده بازهٔ کارایی اصل رگبار را تا ۴ کیلوهرتز گسترش می‌دهد که بالاتر از بازهٔ عملیاتی اصل مکان می‌باشد.

---

[۱] impedance matching

[۲] oval windows

[۳] basilar membrane

[۴] place principle

[۵] volley principle

[۶] action potential