بسم الله الرحمن الرحيم

٩٣٩٥٤

بسمه تعالی

دانشگاه مازندران

مجتمع آموزش عالی فنی و مهندسی نوشیروانی

موضوع:

بهینه سازی Voice Active Detector (VAD) بر مبنای آستانه گذاری موجک

Improvement of VAD with Wavelet Tresholding

جهت اخذ درجه کارشناسی ارشد

رشته مهندسی برق – گرایش الکترونیک

استاد راهنما:

دکتر محمد رضا کرمی ملائی

استاد مشاور:

دکتر حسین میار نعیمی

نگارش:

محدثه اسحاقی

زمستان ۱۳۸۶

۱۳۸۷ /۳/ ۲۳

۹۳۹۵۴

باسمه تعالی

دانشگاه مازندران

معاونت آموزشی

تحصیلات تکمیلی

ارزشیابی پایان نامه در جلسه دفاعیه

مجتمع آموزش عالی فنی مهندسی نوشیروانی

نام و نام خانوادگی دانشجو: محدثه اسحاقی نیموری — شماره دانشجوئی: ۸٤٥۱۳٦۹۰۰٦

رشته تحصیلی: مهندسی برق – الکترونیک — مقطع: کارشناسی ارشد

سال تحصیلی: نیمسال دوم ۸۷ – ۱۳۸۶

عنوان پایان نامه :

«بهینه سازی VAD(Voie Active Detector) بر مبنای آستانه گذاری موجک»

تاریخ دفاع: ۱۳۸٦/۱۲/۲۰

نمره پایان نامه (به عدد): ۱۹/۵

نمره پایان نامه (به حروف): نوزده و نیم

۱۳۸۷ /۳/ ۲۳

هیات داوران :

استاد راهنما: دکتر محمد رضا کرمی ملائی — امضا

استاد مشاور: دکتر حسین میار نعیمی — امضا

استاد مدعو: دکتر غلامرضا اردشیر — امضا

استاد مدعو: دکتر عطاء ا... ابراهیم زاده — امضا

نماینده کمیته تحصیلات تکمیلی: دکتر سعید لسان — امضا

با تشکر

از استاد گرانقدر و ارجمندم

جناب آقای دکتر کرمی ملائی

تقدیم به:

پدر

مادر

و

معلمانم

به من آموختند : تلاش، حرکت در مسیر زیستن است.

به من آموختند : اندیشه، پرواز در آسمان زندگی است.

به من آموختند : برای رسیدن به قلّه باید درست اندیشید و بسیار تلاش کرد.

چکیده:

در سال های اخیر همزمان با رشد و توسعه ی سیستم های پردازش گفتار، کاربرد آنها در زمینه های متعددی از قبیل مخابراتی، چند رسانه ای، کامپیوتر های واسط بین انسان و ماشین و غیره گسترش یافته است. امروزه، آشکارسازگفتار[1]، مؤلفه ی اصلی سیستم های ارتباطی مختلف از جمله کدگذاری گفتار، بازشناسی گفتار، Hands-free telephony و حذف اکو محسوب می گردد. در سیستم های رادیویی سلولی مانند GSM و CDMA (که در مد انتقال گسسته می باشند) تشخیص زمان های گفتار و سکوت سبب کاهش تداخل و کاهش توان مصرفی وسایل قابل حمل شده و ظرفیت سیستم را بهبود می بخشد.

در این پایان نامه الگوریتم های مختلف آشکارسازگفتار بررسی گردیده و سپس با استفاده از تبدیل موجک روش های جدیدی ارائه می شود که ایده اصلی این روش های بر استفاده از تبدیل موجک استوار است.

با بررسی روش های موجود دو نکته مشخص شد: اول آنکه این الگوریتم ها دقت لازم را در SNR پایین ندارند بنابراین نتایج حاصل از آنها قابل اطمینان نیستند و دوم، این روشها بطور معمول برای محیط های نویزی ویژه ای طراحی شده اند و نمی توان از آنها در محیط های نویزی متفاوت استفاده نمود. پس از مطالعات بیشتر معلوم شد که در بیشتر روش های فوق از پنجره گذاری ثابت با دقت زمان – فرکانس یکنواخت استفاده می شود در حالی که می دانیم سیگنال گفتاری غیر ایستا و دارای مؤلفه های ناپایدار می باشد و در سیگنال های غیر ایستا تبدیل فوریه با دقت زمان– فرکانس یکنواخت به خوبی پاسخ نمی دهد. این مشکل با استفاده از آنالیز چند دقتی موجک[2] حل می شود.

در ابتدا با استفاده از تبدیل موجک و پارامتر انرژی روشی ساده برای جداسازی قسمت های گفتاری و نویزی ارائه کردیم که نتایج خوبی در پی داشت ولی توانایی جداسازی قسمت های صدادار و بی صدای یک سیگنال را در نویزهای مختلف با SNRهای دلخواه نداشت. به همین دلیل به اصلاح این روش پرداختیم و

[1] -voice active detector.
[2] -wavelet transform.

الگوریتم دیگر بر مبنای تبدیل بسته موجک و پارامتر انرژی و آستانه گذاری پیشنهاد نمودیم. در این روش از تفاوتهای طیفی موجود بین گفتار[1] و نویز استفاده گردید.

در ادامه ی این روند الگوریتم جدید دیگری را ارائه کردیم که مبتنی براستخراج مشخصه ی جدیدی (زاویه ی هر نمونه در تبدیل موجک) می باشد.

در پایان، مقایسه ی روشهای پیشنهادی با روشهای موجود، نشان می دهد که روشهای پیشنهادی نتنها در SNRهای پایین به خوبی پاسخ می دهند بلکه در نویزهای مختلف نیز دقت لازم و کافی را دارا می باشند.

[1] -speech.

فهرست مطالب

عنوان	صفحه
✓ **بخش اول: مفاهیم**	۱
فصل اول: مقدمه	۲
۱.۱- آشنایی	۲
۲.۱- نویز آسان	۵
۳.۱- سکوت	۵
۴.۱- کاربردهای آشکارسازگفتار	۶
۵.۱- ساختار این پایان نامه	۶
فصل دوم: فرمول ها و اطلاعات	۹
۱.۲- مدل سیگنال	۹
۲.۲- اندازه ی نسبت سیگنال به نویز	۱۰
- قسمت بندی SNR	۱۱
۳.۲- پارامترهای مهم آشکارسازگفتار	۱۱
۴.۲- اطلاعات	۱۲
۱.۴.۲- گفتار	۱۲
۲.۴.۲- مشخصات گفتار	۱۲
— حروف صدادار وحروف بی صدا	۱۳
— مدولاسیون فرکانسی	۱۳
— نسبت هارمونی	۱۴
- قسمت های بی صدای گفتار	۱۴

فهرست مطالب

عنوان	صفحه
۳.۴.۲- منابع گفتاری	۱۴
- منبع گفتار تمیز TIMIT	۱۴
۴.۴.۲- انواع نویز	۱۵
- نویز سفید	۱۵
- نویز همهمه	۱۵
- نویز رنگی	۱۶
۵.۴.۲- ترکیب گفتار و نویز	۱۷
✓ **بخش دوم: آشکارسازگفتار**	۱۸
فصل سوم: آشکارسازگفتار	۱۹
۱.۳- مشخصه ها و کلاس ها	۱۹
۲.۳- انتخاب مشخصه ها	۲۰
۳.۳- استخراج مشخصه ها	۲۱
۴.۳- الگوریتم	۲۱
۵.۳- الگوریتم های آشکارسازی گفتار	۲۳
۱.۵.۳- الگوریتم های حوزه زمان	۲۴
۲.۵.۳- الگوریتم های حوزه فرکانس	۲۵
فصل چهارم: تبدیل موجک	۲۹
۱.۴- تاریخچه موجک	۲۹
۲.۴- تبدیل فوریه	۳۰
۳.۴- آنالیز فوریه در بازه زمانی کوتاه	۳۱

فهرست مطالب

عنوان	صفحه
۴.۴- آنالیز موجک	۳۲
۵.۴- آنالیز موجک چه کاری می تواند انجام دهد ؟	۳۲
۶.۴- آنالیز موجک چیست؟	۳۴
۷.۴- تعداد ابعاد	۳۵
۸.۴- تبدیل موجک پیوسته (CWT)	۳۵
۹.۴- مقیاس گذاری	۳۷
۱۰.۴- تغییر مکان دهی	۳۸
۱۱.۴- پنج گام آسان تبدیل پیوسته موجک (CWT)	۳۹
۱۲.۴- مقیاس و فرکانس	۴۱
۱۳.۴- چه چیزی در تبدیل پیوسته موجک، پیوسته است ؟	۴۲
۱۴.۴- دقتهای زمانی و فرکانسی	۴۳
۱۵.۴- روابط مربوط به CWT	۴۳
۱۶.۴- ویژگیهای موجک (admissibility and regularity conditions)	۴۴
۱۷.۴- چند نمونه از موجک های کلاسیک	۴۵
۱۸.۴- تبدیل گسسته موجک	۴۶
— فیلترگذاری تک مرحله ای : تقریبات و جزئیات	۴۷
— آنالیز بسته موجک	۴۹
- تعداد سطوح	۵۰
۱۹.۴- بازسازی موجک	۵۰

فهرست مطالب

عنوان	صفحه
۲۰.۴- تجزیه و بازسازی چند مرحله ای	۵۱
۲۱.۴- کاربردهای موجک	۵۲
- آشکارسازی گفتار با استفاده از تبدیل موجک و تابع خود همبستگی	۵۳
فصل پنجم: سیستمهای آشکارسازگفتار و ارزیابی آنها	۵۷
G.729B – ۱.۵	۵۷
GSM FR/HR/EFR -۲.۵	۵۸
۳.۵- آشکارسازگفتار فازی	۶۰
GSM AMR1/2 -۴.۵	۶۰
۵.۵- آشکارسازگفتار ساده	۶۱
۶.۵- پارامترهای ارزیابی آشکارسازهای گفتار	۶۱
۷.۵- یک مثال ارزیابی	۶۳
۸.۵- نتایج ارزیابی	۶۴
✓ **بخش سوم روش های پیشنهادی و جمع بندی**	۶۶
فصل ششم: روش های پیشنهادی	۶۷
۱.۶- الگوریتم ساده آشکارسازگفتار با استفاده از تبدیل موجک	۶۷
– الگوریتم	۶۷
– الگوریتم انتخاب آستانه	۶۹
۲.۶- الگوریتم آشکارسازگفتار با استفاده از تبدیل بسته موجک و آستانه گذاری	۷۱
– الگوریتم با استفاده از تبدیل بسته ای موجک	۷۱
- الگوریتم انتخاب آستانه	۷۲

فهرست مطالب

عنوان	صفحه
۳.۶- آشکارسازگفتار با استفاده از تبدیل بسته موجک و Teager Energy Operator	۷۴
- تبدیل موجک و Teager Energy Operator	۷۵
- ترکیب دستی بردارهای سیگنال	۷۶
- سیگنال مرجع	۷۸
- استخراج مشخصه ها	۷۹
- آستانه گذاری	۸۰
فصل هفتم: نتایج حاصل از پیاده سازی الگوریتم های پیشنهادی	۸۱
۱.۷- الگوریتم ساده آشکارسازگفتار با استفاده از تبدیل موجک	۸۱
۲.۷- الگوریتم آشکارسازگفتار با استفاده از تبدیل بسته موجک و آستانه گذاری	۸۵
۳.۷- آشکارسازگفتار با استفاده از تبدیل بسته موجک و Teager Energy Operator	۹۳
فصل هشتم: ارزیابی نتایج	۱۰۲
۱.۸- ارزیابی نتایج	۱۰۲
- الگوریتم جدید آشکارسازگفتار با استفاده از تبدیل بسته موجک و آستانه گذاری	۱۰۲
- آشکارسازگفتار با استفاده از تبدیل بسته موجک و Teager Energy Operator	۱۰۶
جمع بندی	۱۱۰
منابع	۱۱۲

فهرست شکل ها

عنوان	صفحه
شکل۱.۱- یک مدل ساده از آشکارسازگفتار	۴
شکل ۲.۱- یک جمله گفتاری همراه با نویز	۵
شکل ۳.۱- خروجی آشکارسازگفتار ایده ال	۵
شکل ۱.۲- spectrogram جمله ی بدون نویز" She had your dark suit and greasy wash-water all year"	۱۳
شکل ۲.۲- spectrogram جمله ی " wash-water all year She had your dark suit and greasy" با نویز سفید	۱۶
شکل ۳.۲- spectrogram جمله ی " wash-water all year She had your dark suit and greasy" با نویزهمهمه	۱۶
شکل ۱.۳- نمودار اندازه یک سیگنال گفتاری بر حسب زمان	۲۴
شکل ۲.۳- بلوک دیاگرام آنالیز کپسترال	۲۷
شکل ۳.۳ - میزان خطای خروجی فیلتر معکوسLPC به ازای حروف صدا دار و بی صدا	۲۷
شکل ۱.۴- تبدیل فوریه	۳۰
شکل ۲.۴ - آنالیز فوریه زمان کوتاه	۳۱
شکل ۳.۴- تبدیل موجک	۳۲
شکل ۴.۴- مقایسه بین سیگنال زمانی، فرکانسی و STFT و موجک	۳۳
شکل ۵.۴- سیگنال سینوسی با یک ناپیوستگی کوچک	۳۳
شکل ۶.۴- تبدیل فوریه و موجک سیگنال مورد نظر	۳۴
شکل ۷.۴- سیگنال سینوسی و شکل موجک	۳۵

فهرست شکل ها

عنوان	صفحه
شکل ۸.۴- تبدیل فوریه	۳۶
شکل ۹.۴- تبدیل موجک	۳۶
شکل ۱۰.۴- اثر فاکتور مقیاس در سیگنال سینوسی	۳۷
شکل ۱۱.۴- اثر فاکتور مقیاس در موجک	۳۸
شکل ۱۲.۴- شیفت دادن یک موجک	۳۸
شکل ۱۳.۴- محاسبه عدد c و انتخاب شکل موجک	۳۹
شکل ۱۴.۴- محاسبه عدد c و انتخاب شکل موجک	۳۹
شکل ۱۵.۴- محاسبه عدد c و انتخاب شکل موجک	۴۰
شکل ۱۶.۴- نمودار زمان – مقیاس	۴۰
شکل ۱۷.۴- نمایش ضرایب تبدیل پیوسته موجک	۴۱
شکل ۱۸.۴- تطابق بین فرکانس و مقیاس	۴۱
شکل ۱۹.۴- موجک Morlet	۴۵
شکل ۲۰.۴- موجک hat Mexican	۴۵
شکل ۲۱.۴- موجک Meyer	۴۶
شکل ۲۲.۴- موجک های Symlets	۴۶
شکل ۲۳.۴- سیگنال اصلیS از دو فیلتر عبور کرده و دو سیگنال به عنوان خروجی ایجاد می شود	۴۷

فهرست شکل ها

عنوان	صفحه
شکل ۲۴.۴- در فرایند سمت راست با کاهش نمونه ضرایب تبدیل موجک تولید می شود	۴۸
شکل ۲۵.۴- درخت تجزیه ی موجک	۴۸
شکل ۲۶.۴- تجزیه ی سیگنال با روش آنالیز موجک	۴۹
شکل ۲۷.۴- درخت تجزیه ی آنالیز بسته موجک	۵۰
شکل ۲۸.۴- روند بازسازی سیگنال در تبدیل موجک	۵۱
شکل ۲۹.۴- فرایند Up sampling	۵۱
شکل ۳۰.۴- فرایند تجزیه و بازسازی	۵۲
شکل ۳۱.۴- روندنمای روش پیشنهادی	۵۳
شکل ۳۲.۴- نتایج حاصل از VAD پیشنهادی	۵۶
شکل ۱.۵- روندنمای آشکارسازگفتار G.729B	۵۹
شکل ۲.۵- بعضی از پارامترهای کمّی	۶۲
شکل۱.۶- بلوک دیاگرام الگوریتم پیشنهادی اول	۷۰
شکل۲.۶- بلوک دیاگرام الگوریتم پیشنهادی دوم	۷۴
شکل ۳.۶— مرز تصمیم برای جداسازی قسمت های صدادار سیگنال	۷۴
شکل ۴.۶— بلوک دیاگرام الگوریتم پیشنهادی سوم	۷۵
شکل ۵.۶- زاویه دسی بل مقادیر بردارنمونه ها	۷۹

فهرست شکل ها

عنوان	صفحه
شکل ۱.۱.۷-الف. سیگنال بدون نویز	۸۲
شکل ۱.۱.۷-ب. سیگنال نویزی با سیگنال به نویز ۶	۸۲
شکل ۱.۱.۷-ج. خروجی آشکارسازگفتار	۸۲
شکل ۱.۱.۷-د. جداسازی حروف صدادار و بی صدا	۸۳
شکل ۲.۱.۷-الف. سیگنال بدون نویز	۸۳
شکل ۲.۱.۷-ب. سیگنال نویزی	۸۳
شکل ۲.۱.۷-ج. خروجی آشکارسازگفتار	۸۴
شکل ۳.۱.۷-الف. سیگنال بدون نویز	۸۴
شکل ۳.۱.۷-ب. سیگنال نویزی	۸۵
شکل ۳.۱.۷-ج. خروجی آشکارسازگفتار	۸۵
شکل ۱.۲.۷-الف. سیگنال بدون نویز	۸۶
شکل ۱.۲.۷-ب. سیگنال نویزی(نویز سفید با SNR=0 اضافه شده)	۸۷
شکل ۱.۲.۷-ج. روندنمای سیگنال بدون نویز	۸۷
شکل ۱.۲.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۸۷
شکل ۱.۲.۷-ر. خروجی آشکارسازگفتار پیشنهادی	۸۸
شکل ۲.۲.۷-الف. سیگنال بدون نویز	۸۸

فهرست شکل ها

عنوان	صفحه
شکل ۲.۲.۷-ب. سیگنال نویزی (نویزرنگی با SNR=۵ اضافه شده)	۸۹
شکل ۲.۲.۷-ج. روندنمای سیگنال بدون نویز	۸۹
شکل ۲.۲.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۸۹
شکل ۲.۲.۷-ر. خروجی آشکارسازگفتار پیشنهادی	۹۰
شکل ۳.۲.۷-الف. سیگنال بدون نویز	۹۰
شکل ۳.۲.۷-ب. سیگنال نویزی (نویز همهمه با SNR=۱۰ اضافه شده)	۹۰
شکل ۳.۲.۷-ج. روندنمای سیگنال بدون نویز	۹۱
شکل ۳.۲.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۹۱
شکل ۳.۲.۷-ر. خروجی آشکارسازگفتار پیشنهادی	۹۱
شکل ۴.۲.۷-الف. سیگنال بدون نویز فارسی	۹۲
شکل ۴.۲.۷-ب. سیگنال نویزی (نویز سفید با SNR=۵ اضافه شده)	۹۲
شکل ۴.۲.۷-ج. روندنمای سیگنال بدون نویز	۹۳
شکل ۴.۲.۷-ر. خروجی آشکارسازگفتار پیشنهادی	۹۳
شکل ۱.۳.۷-الف. سیگنال بدون نویز	۹۴
شکل ۱.۳.۷-ب. سیگنال نویزی (نویز سفید با SNR=0 اضافه شده)	۹۵
شکل ۱.۳.۷-ج. روند نمای سیگنال بدون نویز	۹۵
شکل ۱.۳.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۹۵

فهرست شکل ها

عنوان	صفحه
شکل ۱.۳.۷-ر. فرم گفتاری صدادار سیگنال	۹۶
شکل ۱.۳.۷-ز. خروجی آشکارسازگفتار	۹۶
شکل ۲.۳.۷-الف. سیگنال بدون نویز	۹۷
شکل ۲.۳.۷-ب. سیگنال نویزی (نویز رنگی با SNR=۵ اضافه شده)	۹۷
شکل ۲.۳.۷-ج. روندنمای سیگنال بدون نویز	۹۷
شکل ۲.۳.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۹۸
شکل ۲.۳.۷-ر. فرم گفتاری صدادار سیگنال	۹۸
شکل ۲.۳.۷-ز. خروجی آشکارسازگفتار	۹۸
شکل ۳.۳.۷-الف. سیگنال بدون نویز	۹۹
شکل ۳.۳.۷-ب. سیگنال نویزی (نویز همهمه با SNR=۱۰ اضافه شده)	۹۹
شکل ۳.۳.۷-ج. روندنمای سیگنال بدون نویز	۱۰۰
شکل ۳.۳.۷-د. مرز جداسازی قسمتهای صدادار گفتاری	۱۰۰
شکل ۳.۳.۷-ر. خروجی آشکارسازگفتار	۱۰۰

فهرست نمودارها

عنوان	صفحه
نمودار ۱.۵- درصد خطای کمّی هریک از آشکارسازها	۶۵
نمودار ۲.۵- نتایج مرحله کیفی	۶۵
نمودار ۱.۱.۸- الف. مقایسه پارامتر CS حاصل از روش پیشنهادی در نویزهای مختلف با SNR های متفاوت	۱۰۴
نمودار ۱.۱.۸- ب. مقایسه پارامتر CL حاصل از روش پیشنهادی در نویزهای مختلف با SNR های متفاوت	۱۰۵
نمودار ۱.۱.۸- ج. مقایسه دقت پارامتر CS روش پیشنهادی با دو روش دیگر	۱۰۵
نمودار ۲.۱.۸- الف. مقایسه پارامتر CS حاصل از روش پیشنهادی در نویزهای مختلف با SNR های متفاوت	۱۰۸
نمودار ۲.۱.۸- ب. مقایسه پارامتر CL حاصل از روش پیشنهادی در نویزهای مختلف با SNR های متفاوت	۱۰۸
نمودار ۲.۱.۸- ج. مقایسه دقت پارامترCS روش پیشنهادی با روش دیگر	۱۰۹

فهرست جداول

عنوان	صفحه
جدول ۱.۳- مراحل و مباحث مهم در طراحی سیستم آشکارسازگفتار	۲۳
جدول ۱.۱.۸-الف. نتایج حاصل از پیاده سازی الگوریتم پیشنهادی دوم	۱۰۳
جدول ۱.۱.۸-ب. مقایسه بین دقت الگوریتم پیشنهادی با دقت الگوریتم آشکارسازگفتار استاندارد G.729B و الگوریتم CHEN	۱۰۳
جدول ۲.۱.۸-الف. نتایج حاصل از پیاده سازی الگوریتم پیشنهادی سوم	۱۰۷
جدول ۲.۱.۸-ب. مقایسه دقت پارامتر CS الگوریتم پیشنهادی با دقت الگوریتم آشکارسازگفتار استاندارد G.729B	۱۰۷