دانشگاه تبریز
دانشکده مهندسی برق و کامپیوتر
گروه مخابرات

پایان نامه

برای دریافت درجه کارشناسی ارشد در رشته مهندسی برق

گرایش مخابرات-سیستم

عنوان

جداسازی تک‌گوشی سیگنال گفتار بر اساس ردیابی Pitch

استاد راهنما

دکتر مسعود گراوانچی‌زاده

استاد مشاور

دکتر میرهادی سیدعربی

پژوهشگر

صنم ایمانی شاملو

بهمن ۱۳۹۰

بسم الله الرحمن الرحيم

به پاس تعبیر عظیم و انسانی شان از کلمه ایثار و از خودگذشتگی

به پاس عاطفه سرشار و گرمای امید بخش وجودشان که در این سردترین روزگاران بهترین پشتیبان است

به پاس قلب های بزرگشان که فریادرس است و سرگردانی و ترس در پناهشان به شجاعت می گراید

و به پاس محبت های بی دریغشان که هرگز فروکش نمی کند

این مجموعه را به پدر و مادر عزیزم و همسر مهربانم تقدیم می کنم

تقدیر و تشکر

شکر و سپاس پروردگار عالم را که موفقم گردانید تا در عرصه ی علم و دانش پله های سعادت و تعالی را طی کنم و در این راه خطیر و پر نشیب و فراز بزرگوارانی خالصانه و دلسوزانه تأثیرگذار بودند که قدردانی از آنان را بر خود لازم می دانم.

بهترین تشکراتم را تقدیم استاد گرانقدر جناب آقای دکتر مسعود گراوانچی زاده می نمایم که با زحمات بی دریغ و راهنمایی های ارزشمندشان همواره پشتیبان و راه گشایم بودند.

هم چنین از استادمشاور محترم جناب آقای دکتر میرهادی سید عربی به خاطر راهنمایی های مفید و ارزنده شان تشکر و قدردانی می نمایم.

در نهایت از تمام دوستانی که در این راه صمیمانه یاریم نمودند، تشکر نموده و از خداوند متعال برایشان آرزوی توفیق روزافزون دارم.

این پایان نامه طبق قرارداد شماره ی ۱۵۷۰۵/۵۰۰/ت از تاریخ ۹۰/۱۰/۲۸ تحت حمایت های مالی و معنوی مرکز تحقیقات

مخابرات ایران قرار گرفته است.

نام خانوادگی: ایمانی شاملو **نام:** صنم

عنوان پایان نامه: جداسازی تک‌گوشی سیگنال گفتار بر اساس ردیابی Pitch

استاد راهنما : دکتر مسعود گراوانچی‌زاده

استاد مشاور: دکتر میرهادی سیدعربی

مقطع تحصیلی: کارشناسی ارشد **رشته:** مهندسی برق **گرایش:** مخابرات-سیستم

دانشگاه: دانشگاه تبریز **دانشکده:** مهندسی برق و کامپیوتر

تاریخ فارغ‌التحصیلی: ۱۳۹۰/۱۱/۱۷ **تعداد صفحه:** ۱۳۰

کلید واژه‌ها: جداسازی تک‌گوشی سیگنال گفتار، آنالیز ترکیب شنیداری محاسباتی، جداسازی گفتار صدادار، ردیابی گام، کلاس‌بندی، تابع خود-هم‌بستگی پوش بهبود یافته.

چکیده

حضور نویز در محیط‌های طبیعی غیرقابل اجتناب است. استخراج گفتار مورد نظر (هدف) از نویز پس‌زمینه دارای حوزه‌ی کاربرد وسیعی از قبیل تشخیص خودکار گفتار، وسایل کمک‌شنوایی و سیستم‌های مخابراتی راه دور می‌باشد. یک شنونده انسانی توانایی قابل توجهی در جداسازی ترکیب صوتی و توجه به یک صوت هدف دارد. این فرآیند ادراکی، آنالیز ترکیب شنیداری (Auditory Scene Analysis) نامیده می‌شود. هدف آنالیز ترکیب شنیداری محاسباتی (Computational Auditory Scene Analysis) شبیه‌سازی عمل پردازش صوت توسط انسان می‌باشد. بسیاری از کاربردها، نظیر مخابرات راه دور و بازیابی اطلاعات صوتی، نیازمند راه حل تک‌گوشی هستند. برای سیگنال‌های تک‌گوشی (تک میکروفونه)، باید ویژگی‌های ذاتی گفتار یا تداخل در نظر گرفته شوند. در این پایان‌نامه، اساس فرآیند جداسازی منحنی گام است که ابتدا، از گفتار جداشده، بر اساس گام غالب، تخمین زده شده و سپس، با توجه به شرط‌های روانی-صوتی (Psychoacoustic) تصحیح می‌شود. مشکل اصلی در سیستم‌های پیشین CASA، جداسازی نامناسب گفتار صدادار در ناحیه‌ی فرکانس بالا می‌باشد. در این پایان‌نامه، دو الگوریتم جدید برای بهبود فرآیند جداسازی گفتار صدادار، به ویژه در ناحیه‌ی فرکانس بالا، پیشنهاد شده است. الگوریتم اول از تابع خود-هم‌بستگی پوش بهبود یافته (Enhanced Envelope Autocorrelation Function) و هم‌بستگی پوش پاسخ بین کانالی برای برچسب-گذاری واحدهای زمان-فرکانسی استفاده می‌کند. الگوریتم دوم مبتنی بر کلاس‌بندی واحدهای زمان-فرکانسی به گفتار صدادار و نویز غیرگفتار، در ناحیه‌ی فرکانس بالا، می‌باشد. نتایج مقایسه روش‌های پیشنهادی با مدل Hu, Wang نشان می‌دهد که بطور کلی، فرآیند جداسازی گفتار صدادار به طور قابل

توجهی در ناحیه‌ی فرکانس بالا بهبود یافته و در عین حال ردیابی گام بهتری صورت گرفته است. همچنین، الگوریتم پیشنهادی دوم نتایج بهتری را نسبت به الگوریتم اول، در امر جداسازی گفتار صدادار از نویزهای غیرگفتار، ارائه می‌دهد.

پیشگفتار ۱
فصل ۱: پیشینه‌ی پژوهش **۲**
۱.۱ مقدمه ۳
۲.۱ سیستم شنیداری انسان ۳
۳.۱ درآمدی بر CASA ۵
۴.۱ ساختار کلی سیستم‌های CASA ۶
۵.۱ هدف محاسباتی CASA ۷
۶.۱ کاربردهای CASA ۸
۷.۱ روش‌های جداسازی گفتار بر اساس CASA ۱۰
۱.۷.۱ جداسازی گفتار بر اساس ویژگی‌ها ۱۰
۱.۱.۷.۱ جداسازی گفتار به صورت تک‌گوشی ۱۱
۲.۱.۷.۱ جداسازی گفتار به صورت دوگوشی ۱۳
۲.۷.۱ آنالیز ترکیب بر اساس مدل ۱۴
۳.۷.۱ جداسازی گفتار با استفاده از شبکه‌های عصبی ۱۵
۸.۱ ارزیابی سیستم‌های CASA ۱۵
۹.۱ چالش‌های CASA ۱۷
فصل ۲: بررسی روش‌های جداسازی گفتار صدادار بر اساس آنالیز ترکیب شنیداری محاسباتی **۱۹**
۱.۲ مقدمه ۲۰
۲.۲ جداسازی گفتار به صورت تک‌گوشی ۲۰
۳.۲ محیط شنیداری ۲۳
۴.۲ استخراج ویژگی‌ها ۲۵
۱.۴.۲ استخراج پوش ۲۶
۲.۴.۲ تخمین گام ۲۷
۱.۲.۴.۲ تناوب و فرکانس پایه ۲۷
۲.۲.۴.۲ فرکانس پایه و گام ۲۸
۳.۲.۴.۲ تخمین $F0$ یک منبع ۲۹
۴.۲.۴.۲ تخمین $F0$ چند منبع ۳۴
۳.۴.۲ کوریلوگرام و هم‌بستگی بین کانالی ۳۵
۵.۲ بازسازی شکل موج خروجی ۳۸
۶.۲ جداسازی گفتار صدادار با استفاده از الگوریتم تکرارشونده ۳۸
۱.۶.۲ تخمین ماسک باینری ایده‌ال با استفاده از گام سیگنال مورد نظر ۳۹
۲.۶.۲ تخمین گام با استفاده از ماسک باینری ایده‌ال ۴۰

۳.۶.۲ ارزیابی ۴۰
۷.۲ جداسازی گفتار صدادار بر اساس مجموعی از ویژگی‌ها و توزیع انرژی ۴۱
۸.۲ جداسازی گفتار صدادار بر اساس راهبرد دقیق گروه‌بندی هارمونیک‌ها ۴۲
۹.۲ جداسازی گفتار صدادار و ارزیابی آن توسط آزمون‌های قابلیت فهم مقدماتی ۴۳
۱۰.۲ جداسازی گفتار صدادار با استفاده از زمینه‌های اتفاقی جداکننده ۴۵
۱۱.۲ جداسازی گفتار صدادار بر اساس هم‌بستگی نوسانی ۴۵
۱۲.۲ جداسازی گفتار صدادار و پس‌زمینه بر اساس دسته‌بندی هارمونیکی-زمانی ۴۶
۱۳.۲ جداسازی گفتار صدادار بر اساس ردیابی گام و مدولاسیون دامنه ۴۶
۱.۱۳.۲ بلوک دیاگرام ۴۶
۲.۱۳.۲ پیاده‌سازی الگوریتم ۴۶
۳.۱۳.۲ نتایج شبیه‌سازی الگوریتم ۴۷
فصل ۳: ارائه‌ی روش‌های جدید جداسازی گفتار صدادار به صورت تک‌گوشی ۵۸
۱.۳ مقدمه ۵۹
۲.۳ الگوریتم پیشنهادی مبتنی بر EEACF و هم‌بستگی پوش پاسخ بین کانالی ۶۰
۱.۲.۳ آنالیز اولیه و استخراج ویژگی‌ها ۶۲
۱.۱.۲.۳ تجزیه‌ی سیگنال ۶۲
۲.۱.۲.۳ استخراج ویژگی‌های شنوایی ۶۳
۲.۲.۳ جداسازی اولیه ۶۹
۱.۲.۲.۳ بخش‌بندی اولیه ۶۹
۲.۲.۲.۳ گروه‌بندی اولیه ۷۰
۳.۲.۲.۳ فلوچارت مرحله‌ی جداسازی اولیه ۷۲
۳.۲.۳ ردیابی گام و برچسب‌گذاری واحدها ۷۳
۱.۳.۲.۳ ردیابی گام ۷۳
۲.۳.۲.۳ برچسب‌گذاری واحدها ۷۶
۳.۳.۲.۳ فلوچارت مرحله‌ی ردیابی گام و برچسب‌گذاری واحدهای زمان-فرکانسی ۷۷
۴.۲.۳ جداسازی نهایی ۷۸
۱.۴.۲.۳ فلوچارت مرحله‌ی جداسازی نهایی ۷۹
۳.۳ الگوریتم پیشنهادی مبتنی بر کلاس‌بندی واحدهای زمان-فرکانسی ۸۰
فصل ۴: نتایج و پیشنهادات ۸۵
۱.۴ مقدمه ۸۶
۲.۴ شرایط شبیه‌سازی ۸۶
۱.۲.۴ دادگان ۸۶

۲.۲.۴ معیارهای ارزیابی ۸۷
۱.۲.۲.۴ معیارهای ارزیابی subjective ۸۷
۲.۲.۲.۴ معیارهای ارزیابی objective ۸۸
۳.۴ نتایج شبیه‌سازی الگوریتم پیشنهادی مبتنی بر EEACF و هم‌بستگی پوش پاسخ بین کانالی ۸۹
۴.۴ نتایج شبیه‌سازی الگوریتم پیشنهادی مبتنی بر کلاس‌بندی ۱۰۱
۵.۴ مقایسه‌ی عملکرد الگوریتم‌های پیشنهادی و الگوریتم Hu, Wang ۱۱۱
۶.۴ پیشنهادات ۱۲۵
مراجع ۱۲۶

فهرست اختصارات

ACF	Autocorrelation Function
AMDF	Average Magnitude Difference Function
ASA	Auditory Scene Analysis
ASR	Automatic Speech Recognition
CASA	Computational Auditory Scene Analysis
CRF	Conditional Random Fields
DRF	Discriminative Random Fields
EM	Expectation-Maximization
ERB	Equivalent Rectangular Bandwidth
GF	Gammatone Feature
GFCC	Gammatone Frequency Cepstral Coefficients
GMM	Guassian Mixture Models
HMM	Hidden Markov Models
HTC	Harmonic-Temporal Clustering
IBM	Ideal Binary Mask
IID	Interaural Intensity Difference
ITD	Interaural Time Difference
MAP	Maximum Autocorrelation Peak
ML	Maximum Likelihood
MLP	Multi-Layer Perceptron
SACF	Summary Autocorrelation Function
SDF	Squared Difference Function
WALE	Windowed Autocorrelation Lag Energy

فهرست جداول

جدول (۱.۲): میانگین نسبت سیگنال به نویز (SNR) (نسبت به سیگنال تمیز) برای سیگنال‌های جداسازی شده و سیگنال‌های به دست آمده از ماسک باینری ایده‌ال، مربوط به ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۴۸

جدول (۲.۲): میانگین درصد اتلاف انرژی (P_EL) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۴۹

جدول (۳.۲): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی dB -5 ..... ۴۹

جدول (۴.۲): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی dB 0 ..... ۵۰

جدول (۵.۲): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی dB 5 ..... ۵۰

جدول (۶.۲): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی dB 10 ..... ۵۰

جدول (۷.۲): میانگین نسبت سیگنال به نویز (SNR) سیگنال‌های جداسازی شده نسبت به سیگنال‌های به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) برای ترکیب جملات صدادار با انواع نویز در SNR-های مختلف ورودی..... ۵۱

جدول (۱.۳): مقایسه ویژگی‌های به کار رفته در مدل Hu & Wang [3, 25] و الگوریتم پیشنهادی اول..... ۷۸

جدول (۲.۳):مقایسه ویژگی‌های به کار رفته در مدل Hu & Wang [3, 25] و الگوریتم پیشنهادی دوم..... ۸۴

جدول (۱.۴): شرایط آزمایشگاهی برای شبیه‌سازی الگوریتم پیشنهادی مبتنی برEEACF و هم‌بستگی پوش پاسخ بین کانالی..... ۸۹

جدول (۲.۴): میانگین نسبت سیگنال به نویز (SNR) (نسبت به سیگنال تمیز) برای سیگنال‌های جداسازی شده و سیگنال‌های به دست آمده از ماسک باینری ایده‌ال، مربوط به ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۹۰

جدول (۳.۴): میانگین درصد اتلاف انرژی (P_EL) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۹۰

جدول (۴.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی dB -5 ..... ۹۱

جدول (۵.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 0 dB ۹۱

جدول (۶.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 5 dB ۹۱

جدول (۷.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 10 dB ۹۲

جدول (۸.۴): میانگین نسبت سیگنال به نویز (SNR) سیگنال‌های جداسازی شده نسبت به سیگنال‌های به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) برای ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۹۶

جدول (۹.۴): شرایط آزمایشگاهی برای شبیه‌سازی الگوریتم پیشنهادی مبتنی بر کلاس‌بندی..... ۱۰۱

جدول (۱۰.۴): میانگین نسبت سیگنال به نویز (SNR) (نسبت به سیگنال تمیز) برای سیگنال‌های جداسازی شده و سیگنال‌های به دست آمده از ماسک باینری ایده‌ال، مربوط به ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۱۰۲

جدول (۱۱.۴): میانگین درصد اتلاف انرژی (P_EL) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۱۰۲

جدول (۱۲.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی -5 dB ۱۰۳

جدول (۱۳.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 0 dB ۱۰۳

جدول (۱۴.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 5 dB ۱۰۳

جدول (۱۵.۴): میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز در SNR ورودی 10 dB ۱۰۴

جدول (۱۶.۴): میانگین نسبت سیگنال به نویز (SNR) سیگنال‌های جداسازی شده نسبت به سیگنال‌های به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) برای ترکیب جملات صدادار با انواع نویز در SNR-های ورودی مختلف..... ۱۰۴

جدول (۱۷.۴): مقایسه‌ی میانگین نسبت سیگنال به نویز (SNR) (نسبت به سیگنال تمیز) سیگنال‌های جداسازی شده مربوط به ترکیب جملات صدادار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی -5 dB..... ۱۱۲

جدول (۱۸.۴): مقایسه‌ی میانگین درصد اتلاف انرژی (P_EL) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی -5 dB..... ۱۱۲

جدول (۱۹.۴): مقایسه‌ی میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدا-دار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی -5 dB..... ۱۱۳

جدول (۲۰.۴): مقایسه‌ی میانگین نسبت سیگنال به نویز (SNR) سیگنال‌های جداسازی شده توسط الگوریتم‌های مختلف نسبت به سیگنال‌های به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) برای ترکیب جملات صدادار با انواع نویز برای SNR ورودی 5- dB و انواع نویز..... ۱۱۳

جدول (۲۱.۴):مقایسه‌ی میانگین نسبت سیگنال به نویز (SNR) (نسبت به سیگنال تمیز) سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی 5 dB..... ۱۱۴

جدول (۲۲.۴): مقایسه‌ی میانگین درصد اتلاف انرژی (P_EL) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی 5 dB..... ۱۱۴

جدول (۲۳.۴): مقایسه‌ی میانگین درصد نویز باقی‌مانده (P_NR) در سیگنال‌های جداسازی شده از ترکیب جملات صدادار با انواع نویز توسط الگوریتم‌های مختلف برای SNR ورودی 5 dB..... ۱۱۵

جدول (۲۴.۴): مقایسه‌ی میانگین نسبت سیگنال به نویز (SNR) سیگنال‌های جداسازی شده توسط الگوریتم‌های مختلف نسبت به سیگنال‌های به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) برای ترکیب جملات صدادار با انواع نویز برای SNR ورودی 5 dB و انواع نویز..... ۱۱۵

فهرست شکل‌ها

شکل(۱.۱): ساختار گوش انسان.....۴

شکل (۲.۱): ساختار سیستم نمونه‌ی CASA.....۶

شکل (۱.۲): طرح کلی سیستم نمونه‌ی CASA برای جداسازی گفتار به صورت تک‌گوشی.....۲۱

شکل (۲.۲): (a) تجزیه‌ی زمان-فرکانسی جمله‌ای با گوینده‌ی مرد، "Why were you all weary?" (b) شکل موج جمله. (c) تجزیه‌ی زمان-فرکانسی جمله‌ی ترکیب شده با نویز مهمانی در SNR = 3 dB. (d) شکل موج ترکیب. (e) رشته‌ی هدف شکل یافته از تمام واحدهای زمان-فرکانسی (نواحی سفید) مربوط به سیگنال موردنظر (ماسک باینری ایده‌ال). (f) شکل موج بازسازی شده از رشته‌ی هدف.....۲۲

شکل (۳.۲): پاسخ فرکانسی فیلتر گاماتون. برای وضوح بیشتر، 32 کانال نمایش داده شده است.....۲۴

شکل (۴.۲): مدولاسیون دامنه‌ی پاسخ فیلتر. (a) پاسخ فیلتری با فرکانس مرکزی 3.2 kHz به جمله‌ی نمایش داده شده در شکل (۲.۲) (b) با گوینده‌ی مرد. (b) پوش پاسخ مربوطه.....۲۷

شکل (۵.۲): بلوک دیاگرام تخمین $F0$ یک منبع در حوزه‌ی زمان-فرکانس.....۳۴

شکل (۶.۲): کوریلوگرام برای جمله‌ی شکل (۲.۲) (b) با گوینده‌ی مرد. هم‌بستگی بین کانالی در شکل سمت راست و کوریلوگرام مجموع در شکل پایین نمایش داده شده است.....۳۶

شکل (۷.۲): بلوک دیاگرام سیستم CASA بر اساس استفاده از ویژگی توزیع انرژی.....۴۱

شکل (۸.۲): بلوک دیاگرام سیستم چند-مرحله‌ای CASA بر اساس ردیابی گام و مدولاسیون دامنه.....۴۷

شکل (۹.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده.....۵۲

شکل (۱۰.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۹.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی.....۵۲

شکل (۱۱.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N3 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده.....۵۳

شکل (۱۲.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۱.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی.....۵۳

شکل (۱۳.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N6 با نسبت سیگنال به نویز 5 dB-. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۵۴

شکل (۱۴.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۳.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۵۴

شکل (۱۵.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۵۵

شکل (۱۶.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۵.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۵۵

شکل (۱۷.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N3 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۵۶

شکل (۱۸.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۷.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۵۶

شکل (۱۹.۲): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N6 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۵۷

شکل (۲۰.۲): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۹.۲) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۵۷

شکل (۱.۳): بلوک دیاگرام الگوریتم پیشنهادی مبتنی بر EEACF و هم‌بستگی پوش پاسخ بین کانالی..... ۶۰

شکل (۲.۳): (a) طیف‌نگار جمله‌ی ترکیب شکل (۲.۲) (d). (b) طیف‌نگار خروجی فیلتربانک گاماتون جمله‌ی مربوطه. (c) طیف‌نگار خروجی مدل سلول‌های مویی. (d) کاکلیگرام..... ۶۳

شکل (۳.۳): (a) تابع خود-هم‌بستگی پوش پاسخ جمله‌ی ترکیب شکل (۲.۲) (d) در فریم 70 و کانال 90. (b) شکل موج تابع خود-هم‌بستگی پوش بهبود یافته‌ی مربوطه..... ۶۶

شکل (۴.۳): (a) طیف تابع خود-هم‌بستگی پوش پاسخ در فریم 70 برای جمله‌ی ترکیب در شکل (۲,۲) (d) با گوینده-ی مرد. (b) طیف تابع خود-هم‌بستگی پوش بهبود یافته در فریم 70. (c) شکل موج تابع خود-هم‌بستگی پوش بهبود یافته در تمام کانال‌ها در فریم 70..... ۶۷

شکل (۵.۳): ویژگی‌های شنیداری. (a) کوریلوگرام در فریم 70 برای جمله‌ی ترکیب در شکل (۲.۲) (d) با گوینده‌ی مرد. هم‌بستگی بین کانالی در شکل سمت راست و کوریلوگرام مجموع در شکل پایین نمایش داده شده است. (b) کوریلوگرام در فریم 70 برای جمله‌ی تمیز در شکل (۲.۲) (b). هم‌بستگی بین کانالی در شکل سمت راست و کوریلوگرام مجموع در شکل پایین نمایش داده شده است. (c) کوریلوگرام پوش جمله‌ی ترکیب. هم‌بستگی بین کانالی در شکل سمت راست نمایش داده شده است. (d) کوریلوگرام پوش جمله‌ی تمیز. هم‌بستگی بین کانالی در شکل سمت راست نمایش داده شده است..... ۶۸

شکل (۶.۳): نتیجه‌ی بخش‌بندی اولیه برای جمله‌ی ترکیب شکل (۲.۲) (d). نواحی خاکستری مربوط به بخش‌های فرکانس پایین و نواحی سفید مربوط به بخش‌های فرکانس بالا هستند..... ۷۰

شکل (۷.۳): فلوچارت مرحله‌ی جداسازی اولیه..... ۷۲

شکل (۸.۳): نتایج ردیابی گام برای جمله‌ی نشان داده شده در شکل (۲.۲) (b). (a) تناوب‌های گام غالب با ’’x’’ نمایش داده شده‌اند. (b) منحنی گام تخمین زده شده برای گفتار مورد نظر با ’’x’’ نمایش داده شده است. خط پر منحنی گام به‌دست آمده از گفتار تمیز پیش از ترکیب است که با نرم‌افزار Praat [51] به‌دست آمده است..... ۷۸

شکل (۹.۳): فلوچارت مرحله‌ی ردیابی گام و برچسب‌گذاری واحدهای زمان-فرکانسی..... ۷۷

شکل (۱۰.۳): فلوچارت مرحله‌ی جداسازی نهایی..... ۷۹

شکل (۱۱.۳): بلوک دیاگرام الگوریتم پیشنهادی مبتنی بر کلاس‌بندی..... ۸۰

شکل (۱.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۳

شکل (۲.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۳

شکل (۳.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۴

شکل (۴.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۳.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۴

شکل (۵.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N6 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۵

شکل (۶.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۵.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۵

شکل (۷.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۶

شکل (۸.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۷.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۶

شکل (۹.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۷

شکل (۱۰.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۹.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۷

شکل (۱۱.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N3 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۸

شکل (۱۲.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۱.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۸

شکل (۱۳.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N6 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۹۹

شکل (۱۴.۴): ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۳.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی اول. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۹۹

شکل (۱۵.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۰۵

شکل (۱۶.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۵.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۰۵

شکل (۱۷.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۰۶

شکل (۱۸.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۷.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۰۶

شکل (۱۹.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N6 با نسبت سیگنال به نویز 5- dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۰۷

شکل (۲۰.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۱۹.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۰۷

شکل (۲۱.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۰۸

شکل (۲۲.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۲۱.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۰۸

شکل (۲۳.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N2 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۰۹

شکل (۲۴.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۲۳.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۰۹

شکل (۲۵.۴): طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N3 با نسبت سیگنال به نویز 5 dB. (b) شکل موج جمله‌ی ترکیب. (c) طیف‌نگار جمله‌ی صدادار تمیز. (d) شکل موج جمله‌ی تمیز. (e) طیف‌نگار جمله‌ی جداسازی شده. (f) شکل موج جمله‌ی جداسازی شده..... ۱۱۰

شکل (۲۶.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی ترکیب شکل (۲۵.۴) (b). (b) شکل موج به‌دست آمده از ماسک باینری ایده‌ال. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی دوم. (d) شکل موج به‌دست آمده از ماسک تخمینی خروجی..... ۱۱۰

شکل (۲۷.۴): (a) طیف‌نگار جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز dB 5- .(b) طیف‌نگار جمله‌ی صدادار تمیز. (c) طیف‌نگار جمله‌ی جداسازی شده توسط الگوریتم Hu, Wang. (d) طیف‌نگار جمله‌ی جداسازی شده توسط الگوریتم پیشنهادی بر اساس EEACF و هم‌بستگی پوش پاسخ بین کانالی. (e) طیف‌نگار جمله‌ی جداسازی شده توسط الگوریتم پیشنهادی بر اساس کلاس‌بندی..... ۱۱۶

شکل (۲۸.۴): (a) شکل موج جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز dB 5- .(b) شکل موج جمله‌ی صدادار تمیز. (c) شکل موج جمله‌ی به دست آمده از ماسک باینری ایده‌آل. (d) شکل موج جمله‌ی جدا-سازی شده توسط الگوریتم Hu, Wang. (e) شکل موج جمله‌ی جداسازی شده توسط الگوریتم پیشنهادی بر اساس EEACF و هم‌بستگی پوش پاسخ بین کانالی. (f) شکل موج جمله‌ی جداسازی شده توسط الگوریتم پیشنهادی بر اساس کلاس‌بندی..... ۱۱۷

شکل (۲۹.۴): (a) ماسک باینری ایده‌ال برای جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز 5- dB. (b) ماسک باینری ایده‌ال تخمینی توسط الگوریتم Hu, Wang. (c) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی بر اساس EEACF و هم‌بستگی پوش پاسخ بین کانالی. (d) ماسک باینری ایده‌ال تخمینی توسط الگوریتم پیشنهادی بر اساس کلاس‌بندی..... ۱۱۸

شکل (۳۰.۴): (a) میانگین نتایج SNR نسبت به سیگنال خروجی به دست آمده از ماسک باینری ایده‌ال (رابطه‌ی (۱.۴)) در الگوریتم Hu, Wang، الگوریتم پیشنهادی اول و الگوریتم پیشنهادی دوم برای جملات صدادار ترکیب شده با انواع نویز با نسبت سیگنال به نویز dB 5-..... ۱۱۸

شکل (۳۱.۴):(a) میانگین نتایج SNR (نسبت به سیگنال تمیز) سیگنال‌های خروجی الگوریتم Hu, Wang، الگوریتم پیشنهادی اول، الگوریتم پیشنهادی دوم و ماسک باینری ایده‌ال برای جملات صدادار ترکیب شده با انواع نویز با نسبت سیگنال به نویز dB 5-..... ۱۱۹

شکل (۳۲.۴): (a) نتیجه‌ی ردیابی گام برای جمله‌ی صدادار v3 ترکیب شده با نویز N0 با نسبت سیگنال به نویز dB 5- توسط الگوریتم Hu, Wang. (b) نتیجه‌ی ردیابی گام توسط الگوریتم پیشنهادی بر اساس EEACF و هم‌بستگی پوش پاسخ بین کانالی. (c) نتیجه‌ی ردیابی گام توسط الگوریتم پیشنهادی بر اساس کلاس‌بندی. منحنی گام تخمین زده شده با "x" نمایش داده شده است. خط پر، منحنی گام به‌دست آمده از گفتار تمیز پیش از ترکیب است که با نرم‌افزار Praat به‌دست آمده است..... ۱۱۹

شکل (۳۳.۴):شکل‌ها به‌ترتیب از بالا. شکل اول، مقایسه‌ی میانگین نتایج SNR (نسبت به سیگنال تمیز) سیگنال خروجی به دست آمده از الگوریتم Hu, Wang، الگوریتم پیشنهادی اول، الگوریتم پیشنهادی دوم و ماسک باینری ایده‌ال برای تمام جملات و تمام نویزها. شکل دوم، مقایسه‌ی میانگین نتایج SNR (نسبت به سیگنال تمیز) سیگنال خروجی به دست آمده از الگوریتم Hu, Wang، الگوریتم پیشنهادی اول، الگوریتم پیشنهادی دوم و ماسک باینری ایده‌ال برای تمام جملات و نویزهای غیرگفتار. شکل سوم، مقایسه‌ی میانگین نتایج SNR (نسبت به سیگنال تمیز) سیگنال خروجی به دست آمده از الگوریتم Hu, Wang، الگوریتم پیشنهادی اول و ماسک باینری ایده‌ال برای تمام جملات و نویزهای گفتار و گفتارگونه..... ۱۲۱