**دانشگاه صنعتی امیر کبیر**
دانشکده مهندسی برق

رساله دکتری مهندسی برق

عنوان

# بهبود عملکرد سیستم های جداسازی سیگنال گفتار دو گوینده در حالت تک میکروفونه

نگارش
**محمد حسین رادفر**

استاد راهنما
**دکتر ابواقاسم صیادیان**

۱۳۸٦

بسمه تعالي

تاريخ :
..۸۶/۷/.۱۰.....
شماره : .....

**فرم اطلاعات پايان نامه**
**كارشناسي ارشد و دكترا**

**دانشگاه صنعتي اميركبير**
( پلي تكنيک تهران )

**معاونت پژوهشي**
فرم پروژه تحصيلات تكميلي ۷

**مشخصات دانشجو**

نام و نام خانوادگي : محمد حسين رادفر دانشجوي آزاد

شماره دانشجويي: ۸۱۱۲۳۹۱۹ دانشكده : برق رشته تحصيلي: مخابرات

---

**نام و نام خانودگي استاد راهنما : ابوالقاسم صيادیان**

---

**عنوان به فارسي: بهبود عملكرد سيستم های جداسازی سيگنال در حالت تک ميكروفونه**

**عنوان به انگليسي:**
**Performance improvement of single channel speech separation techniques**

**نوع پروژه: دكترا** **كاربردي**

---

**تاريخ شروع:** ۱۳۸۱ **تاريخ خاتمه :** ۱۳۸۶ **تعداد واحد:** ۲۴

**سازمان تامين كننده اعتبار :**

---

**واژه هاي كليدي به فارسي : جداسازی سيگنال گفتار - ماسک دودويی – تحليل محاسباتی محيط شنوايی -**
**واژه هاي كليدي به انگليسي :**
**Speech separation –binary mask – computational auditory scene analysis**

---

**نظرها و پيشنهادها به منظور بهبود فعاليت هاي پژوهشي دانشگاه:**

**استاد راهنما:**

**دانشجو:**

**امضاء استاد راهنما :** **تاريخ:** ۱۳۸۶/۷/۱۰

---

**نسخه ۱:** معاونت پژوهشي
**نسخه ۲ :** كتابخانه و به انضمام دو جلد پايان نامه به منظور تسويه حساب با كتابخانه و مركز اسنادومدارک علمی

**سپاسگزاری :**

در اینجا لازم می دانم از آقای دکتر صیادیان که در تمامی مراحل این رساله، مشوق و راهنمای من بودند و همواره مرا از نظریات و راهنمایی های ارزشمندشان بهره‌مند نموده‌اند تشکر و قدردانی کنم. اگرچه هیچ کلامی نمی تواند گویای نقش ایشان در انجام این رساله و جبران محبت ایشان باشد.

همچنین از آقای دکتر ریچارد دانسرو بخاطر زحمات بی دریغ و راهنمایی های ارزنده‌شان، کمال تشکر را دارم.

ضمنا از اساتید محترمی که قبول زحمت نموده، در جلسه دفاع از رساله و همچنین پیشنهاد رساله شرکت کردند و با پیشنهادهایی راهگشا امکان انجام این پروژه را فراهم نمودند ، کمال تشکر را دارم.

**چکیده:**

در این رساله روش جداسازی سیگنال دو گوینده، زمانی که فقط به سیگنال مخلوط دریافت شده از یک میکروفون دسترسی داریم مورد بررسی قرار می‌گیرد. بطور کلی روش های جداسازی سیگنال در حالت تک میکروفونه به دو گروه تقسیم می شود: گروه اول روش هایی را شامل می شوند که اصول حاکم بر سیستم شنوایی را برای جداسازی دنبال می کنند و گروه دوم روش هایی که با توجه به اطلاعات قبلی از دو گوینده، سیگنال مخلوط را مدل می کنند. هدف اصلی این رساله پوشش نقاط ضعف روش های موجود است. برای این منظور چهار راه حل ارائه می شود: (۱) ترکیب روش های گروه اول و دوم، این روش علاوه براینکه مزایای هر دو روش را حفظ می کند، مستقل از گوینده می باشد. (۲) الگوریتم جدیدی برای تخمین فرکانس های اصلی دو گوینده زمانی که فقط به سیگنال مخلوط دسترسی داریم ارائه می شود. فرکانس های اصلی دو گوینده به عنوان ویژگی متمایز کننده دو گوینده در روش های گروه اول کاربرد دارند. (۳) روش تعمیم یافته ای برای روش های مدل سازی وابسته به گوینده بر اساس شناسایی گوینده ها از سیگنال مخلوط ارائه می شود. به این ترتیب می توان روش های وابسته به گوینده را به تعداد نامحدودی از گوینده ها تعمیم داد. (۴) بر اساس روش تخمین کمینه کردن متوسط مربعات خطا[a] (MMSE)، فیلتر جدیدی بنام ماسک نرم برای جداسازی سیگنال دو گوینده ارائه می شود. علاوه بر ارائه روش های جدید جداسازی ، در این رساله دو تقریب بنام های تقریب ماکزیمم- مخلوط و ماسک دودویی که در روش های جداسازی کاربرد فراوانی دارد، موردآنالیز قرار گرفته و با استفاده از روش هایی ریاضی ، اثبات می شوند. اهمیت این آنالیز ها از آن جهت است که دیدگاه جدیدی با مستدلات ریاضی برای ابزارهای که بمدت ۲۰ سال بصورت تجربی مورد استفاده قرار می گرفته، ارائه می شود.

[a] Minimum mean square error

# فهرست مندرجات


**۱ مقدمه** **۱**

**۲ مروری بر روش های جداسازی سیگنال در حالت تک کاناله** **۱۱**

۲-۱ روش های جداسازی مبتنی بر تحلیل محاسباتی محیط شنوایی . . . . . ۱۱

۲-۱-۱ روش های CASA تصریحی و تلویحی . . . . . ۱۱

۲-۱-۲ انتقال سیگنال مخلوط به فضای ویژگی مناسب (تجزیه) . . . . . ۱۲

۲-۱-۳ استخراج نشان های مناسب برای جداسازی . . . . . ۱۳

۲-۱-۴ تخمین سیگنال های گفتار( ترکیب) . . . . . ۱۵

۲-۲ روش های جداسازی تک کاناله بر اساس مدل سازی گویندها . . . . . ۱۸

۲-۲-۱ ویژگی های اساسی . . . . . ۱۸

۲-۲-۲ بررسی روش های جداسازی مبتنی بر روش های استنتاج آماری . . . . . ۲۱

**۳ جداسازی سیگنال های گفتار مبتنی بر ترکیب روش های CASA و مدل سازی** **۲۷**

۳–۱ مقدمه . . . . . ۲۷

۳–۲ رابطه بین بهره و نسبت توان سیگنال–به–سیگنال . . . . . ۲۸

۳–۳ تخمینگر بیشینه–مخلوط . . . . . ۳۰

۳–۴ ترکیب روش های آنالیز عملکرد محیط شنوایی با روش های مبتنی بر مدل سازی . . ۳۴

۳–۴–۱ تجزیه سیگنال گفتار به سیگنال تحریک و پوش طیف . . . . . ۳۸

۳–۴–۲ تخمین فرکانس اصلی سیگنال گفتار دو گوینده از سیگنال مخلوط . . . . . ۳۹

۳–۵ مدل سازی توابع توزیع منابع . . . . . ۳۹

۳–۶ تابع توزیع سیگنال مخلوط . . . . . ۴۰

۳–۷ تخمین منابع با استفاده از روش درستنمایی بیشینه . . . . . ۴۲

۳–۸ ارزیابی عملکرد مدل و مقایسه . . . . . ۴۳

۳–۸–۱ استخراج بردارهای ویژگی و مدل کردن توابع توزیع مخلوط گوسی . . . . ۴۳

۳–۸–۲ معرفی تکنیک های مورد استفاده برای مقایسه . . . . . ۴۴

۳–۸–۳ نتایج . . . . . ۴۶

**۴ تخمین فرکانس اصلی سیگنال های گفتار از سیگنال مخلوط ۵۳**

۴–۱ مقدمه . . . . . ۵۳

۴-۲ معرفی مختصر مدل سینوسی و هارمونیکی سیگنال گفتار . . . . . ۵۶

۴-۳ تخمین فرکانس های اصلی . . . . . ۵۷

۴-۴ گروه بندی، جداسازی و درون یابی . . . . . ۶۱
۴-۴-۱ گروه بندی . . . . . ۶۱
۴-۴-۲ جداسازی . . . . . ۶۳
۴-۴-۳ درون یابی . . . . . ۶۴

۴-۵ مقایسه و نتایج آزمون ها . . . . . ۶۵

**۵ تعمیم روش های جدا سازی وابسته به گوینده‌ها بااستفاده از روش تشخیص گوینده و ماسک نرم ۷۷**

۵-۱ مقدمه . . . . . ۷۷

۵-۲ مروری کلی بر روش پیشنهادی . . . . . ۷۸

۵-۳ تشخیص هویت گوینده‌ها از سیگنال مخلوط . . . . . ۷۹

۵-۴ ارزیابی مدل تعمیم یافته . . . . . ۸۲
۵-۴-۱ ارزیابی مدل پیشنهادی تشخیص گوینده ها از سیگنال مخلوط . . . . . ۸۲

۵-۵ جداسازی سیگنال های گفتار با استفاده از روش ماسک نرم . . . . . ۸۳

۵-۶ استخراج ماسک نرم با استفاده از روش MMSE . . . . . ۸۶
۵-۶-۱ فرمول بندی مسئله . . . . . ۸۶
۵-۶-۲ استخراج ماسک نرم (محاسبه عامل اول) . . . . . ۸۷
۵-۶-۳ محاسبه احتمال زیر منابع (عامل دوم) . . . . . ۸۹
۵-۶-۴ تخمینگرمنتجه . . . . . ۹۰

۵-۷ کاهش حجم محاسبات با استفاده از تقریب معادلات . . . . . ۹۱
۵-۷-۱ تقریب تابع توزیع سیگنال مخلوط . . . . . ۹۱
۵-۷-۲ انتخاب فیلتر نرم با بیشترین احتمال وقوع . . . . . ۹۲

۵-۸ مقایسه ماسک نرم با فیلتر وینر . . . . . ۹۳

۵-۹ آزمون ها . . . . . ۹۴

**۶ نتیجه گیری و پیشنهادات ۱۰۰**

**۷ ضمیمه ۱ ۱۰۳**

**۸ ضمیمه ۲ ۱۰۶**

# لیست اشکال


۱–۱ (a)روش استاندارد جداسازی کور منابع که در آن با استفاده از دومیکروفون سیگنال گفتار دو گوینده با استفاده از معکوس ماتریس مخلوط ($W$) از یکدیگر جدا می گردند. (b) روش تک کاناله (تک میکروفونه) که در آن با استفاده از روش های جداسازی کور منابع نمی توان سیگنال ها را جدا کرد. . . . . ۴

۱–۲ دیاگرام روش های مختلف جداسازی سیگنال گفتار. . . . . ۵

۱–۳ راهکارهای پیشنهادی در این رساله برای بهبود سیستم های جداسازی گفتار در حالت تک کاناله. . . . . ۹

۲–۱ بلوک دیاگرام روش های جداسازی مبتنی بر تحلیل محاسباتی محیط شنوایی (CASA) . ۱۲

۲–۲ نمایش زمان–فرکانس با استفاده از: (a) سپکتروگرام و (b) ککلیوگرام برای یک سیگنال گفتار. . . . . ۱۳

۲-۳ نشان هایی که در مدل های CASA برای جداسازی مورد استفاده قرار می گیرند: پیوستگی فرکانسی ، طبیعت تناوبی اصوات واکدار ، و شروع و پایان آواها. . . . . ۱۴

۲-۴ (a) مدل عمومی بکار رفته برای جداسازی با استفاده از فیلتر شانه‌ای. (b)روش تولید فیلتر شانه‌ای. . . . . ۱۷

۲-۵ جداسازی سیگنال با استفاده از روش مدل سازی. . . . . ۲۰

۲-۶ مدل های کلی برای جداسازی مبتنی بر مدل سازی : (a) روش چندی کردن برداری ، (b)روش مدل مخلوط گوسی (c) روش مدل مارکوف پنهان. . . . . ۲۵

۲-۷ چگونگی استفاده از ماسک دودیی برای جداسازی سیگنال های گفتار. در این شکل، (a) سپکتروگرام سیگنال مخلوط و (b) و (c) بترتیب ماسک های دودویی گوینده اول و دوم را نشان می دهد (برای حالت ایده‌ال) . حاصل ضرب این ماسک های در سپکتروگرام سیگنال مخلوط در شکل های (d) و (e) نشان داده شده است. برای مقایسه، سپکتروگرام های اصلی سیگنال های گفتار دو گوینده نیز در شکل های (f) و (g) آورده شده است. . . . . ۲۶

۳-۱ مقایسه بین $g_y^2$ و $g_1^2 + g_2^2$. برای مشاهده اختلاف جزئی دو منحنی قسمت بزرگ شده‌ای در مستطیل کوچکتر نشان داده شده است. . . . . ۲۹

۳-۲ (a) سیگنال مخلوط اصلی ،(b) سیگنال بدست آمده با استفاده تخمینگر ماکزیمم-مخلوط، (c) خطای تخمین. . . . . ۳۳

۳-۳ تابع تجربی توزیع احتمال فاز تبدیل تبدیل گسسته فوریه سیگنال گفتار. . . . . ۳۴

۳-۴ ساختار کلی سیستم پیشنهادی که در آن مشخصات لوله صوتی به عنوان یک عامل متمایز کننده جدید به سیستم CASA که فقط از مشخصه فرکانس اصلی برای جداسازی استفاده می کنند اضافه شده است. . . . . ۳۵

۳-۵ بلوک دیاگرام مدل ترکیبی پیشنهادی. . . . . ۳۷

۳-۶ بلوک دیاگرام روش جدا سازی با استفاده از روش بیشینه بردارهای چندی شده. . . . ۴۶

۳-۷ میانگین نتایج SNR به ازاء مقادیر SSR برای ۱۶ سیگنال جدا شده‌از ۸ سیگنال مخلوط زن+زن. خط مربع شکل ،□ ، خط لوزی شکل ، ◇ ،و خط دایره‌ای شکل ، ○ ، نتایج حاصله را بترتیب برای مدل پیشنهادی (Proposed)، سیگنال تحریک (Excitation)، و لگاریتم طیف (Spectrum) نشان می دهد. توجه شود محور $x$ در مقیاس لگاریتم رسم شده است. . . . ۴۹

۳-۸ میانگین نتایج SNR به ازاء مقادیر SSR برای ۱۶ سیگنال جدا شده‌از ۸ سیگنال مخلوط مرد+مرد. خط مربع شکل ،□ ، خط لوزی شکل ،◇ ، و خط دایره‌ای شکل ، ○ ، نتایج حاصله را بترتیب برای مدل پیشنهادی Proposed)، سیگنال تحریک (Exciation)، و لگاریتم طیف (Spectrum) نشان می دهد. توجه شود محور $x$ در مقیاس لگاریتم رسم شده است. . . . ۵۰

۳-۹ میانگین نتایج SNR به ازاء مقادیر SSR برای ۱۶ سیگنال جدا شده‌از ۸ سیگنال مخلوط مرد+زن. خط مربع شکل ،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ، ○ ، نتایج حاصله را بترتیب برای مدل پیشنهادی Proposed)، سیگنال تحریک (Exciation)، و لگاریتم طیف (Spectrum) نشان می دهد. توجه شود محور $x$ در مقیاس لگاریتم رسم شده است. . . . ۵۱

۳–۱۰ مثالی از سیگنال های جدا شده با استفاده از روش پیشنهادی. (a) سیگنال گفتار گوینده اول ، (b) سیگنال گفتار گوینده دوم ، (c) سیگنال مخلوط ، (d) سیگنال تخمین گفتار گوینده اول ، و (e)سیگنال تخمین گفتار گوینده دوم. . . . . ۵۲

۴–۱ بلوک دیاگرام روش پیشنهادی برای آشکار سازی منحنی های فرکانس اصلی دو گوینده. ۵۵

۴–۲ لگاریتم طیفی سیگنال مخلوط. خط با فلش دو جهته نمایشگر بازه‌ای است که در آن فرکانس های اصلی رخ می دهند. همچنین قله های طیفی متناظر با فرکانس اصلی دو گوینده ($f_o^x = ۲۲۰\text{Hz}$ و $f_o^y = ۱۸۴\text{Hz}$ ) با علامت های $o$ و $+$ مشخص شده اند. . . . . ۶۰

۴–۳ فرایند تعیین قله های طیفی متناظر با هارمونیک های فرکانس اصلی. قله های متناظر با فرکانس اصلی ۲۲۶ هرتز با علامت+ مشخص شده اند. خطوط موازی عمودی هارمونیک های وابسته به این فرکانس اصلی را نشان می دهد. فرکانس اصلی گوینده دیگر برابر است با ۹۳ هرتز و با علامت $*$ مشخص شده است. . . . . ۶۱

۴–۴ الگوریتم پیشنهادی برای تخمین فرکانس های اصلی دو گوینده در سطح فریم. . . . ۶۲

۴–۵ فرایند گروه‌بندی فرکانس اصلی بدست آمده از مرحله آشکارسازی. . . . . ۶۳

۴–۶ منحنی های فرکانس های اصلی بعد از اعمال فرایند جداسازی و درون یابی. . . . . ۶۵

۴–۷ میانگین نتایج GER بر حسب TIR برای روش های P ، W و C برای ۴۹ مخلوط مرد—مرد. . . . . ۶۹

۴-۸ میانگین نتایج GER بر حسب TIR برای روش های P ، W و C برای ۴۹ مخلوط زن--زن. . . . . ۷۰

۴-۹ میانگین نتایج GER بر حسب TIR برای روش های P ، W و C برای ۹۸ مخلوط مرد--زن. . . . . ۷۰

۴-۱۰ میانگین نتایج FER بر حسب TIR برای روش های P ، W و C برای ۴۹ مخلوط مرد—مرد. . . . . ۷۱

۴-۱۱ میانگین نتایج FER بر حسب TIR برای روش های P ، W و C برای ۴۹ مخلوط زن--زن. . . . . ۷۱

۴-۱۲ میانگین نتایج FER بر حسب TIR برای روش های P ، W و C برای ۹۸ مخلوط مرد--زن. . . . . ۷۲

۴-۱۳ میانگین خطای V→U بر حسب TIR برای روش های P ، W برای ۴۹ مخلوط مرد—مرد. . . . . ۷۲

۴-۱۴ میانگین خطای V→U بر حسب TIR برای روش های P و W برای ۴۹ مخلوط زن--زن. . . . . ۷۳

۴-۱۵ میانگین خطای V→U بر حسب TIR برای روش های P و W برای ۹۸ مخلوط مرد--زن. . . . . ۷۳

۴-۱۶ میانگین خطای U→V بر حسب TIR برای روش های P و W برای ۴۹ مخلوط مرد—مرد. . . . . ۷۴

۴-۱۷ میانگین خطای U→V بر حسب TIR برای روش های P و W برای ۴۹ مخلوط زن—زن. . . . . ۷۴

۴-۱۸ میانگین خطای U→V بر حسب TIR برای روش های P ، W و C برای ۹۸ مخلوط مرد—زن. . . . . ۷۵

۴-۱۹ درصد خطای جداسازی بر حسب TIR برای روش های P ، W برای ۱۹۶ مخلوط . . . ۷۵

۴-۲۰ منحنی های فرکانس تخمین زده شده سیگنال گفتار یک گوینده از سیگنال مخلوط با استفاده از روش های پیشنهادی P (a) ، W (b)و C (c). . . . . ۷۶

۵-۱ بلوک دیاگرام روش پیشنهادی. . . . . ۷۹

۵-۲ الگوریتم تشخیص گوینده‌ها در هنگامی که فقط سیگنال مخلوط در دسترس باشد. . . ۸۱

۵-۳ درصد تشخیص گوینده از سیگنال مخلوط بر حسب مقادیر TIR برای گوینده هدف. خطوط دایره شکل نتایج حاصل از فریم های UV-V و خطوط لوزی شکل برای تمامی فریم ها. ۸۳

۵-۴ درصد تشخیص گوینده از سیگنال مخلوط بر حسب مقادیر TIR برای گوینده تداخلی. خطوط دایره شکل نتایج حاصل از فریم های UV-V و خطوط ستاره شکل برای تمامی فریم ها. . . . . ۸۴

۵–۵ نمایش CSM برای منابع. هر CSM داخل مستطیل های خط چین قرار دارد. . . . . ۸۵

۵–۶ بلوک دیاگرام تخمینگر MMSE برای تخمین $x_{1d}$. هر مستطیل نشان دهنده یک ماسک نرم می باشد. . . . . ۹۱

۵–۷ تقریب تابع توزیع سیگنال مخلوط. . . . . ۹۲

۵–۸ میانگین نتایج SSNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+مرد. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ،، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۶

۵–۹ میانگین نتایج SSNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط زن+زن. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ،، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۶

۵–۱۰ میانگین نتایج SSNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+زن. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ،، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۷

۵–۱۱ میانگین نتایج SNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+مرد. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ، ، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۷

۵–۱۲ میانگین نتایج SNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط زن+زن. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۸

۵–۱۳ میانگین نتایج SNR به ازاء مقادیر زیر منابع (CSM) برای ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+زن. خط مربع شکل،□ ، خط لوزی شکل ،◇ ،و خط دایره‌ای شکل ،○ ، نتایج حاصله را بترتیب بترتیب برای برای روش ماسک نرم (Soft)، ماسک دودیی (Hard)، وینر (Wiener) نشان می دهد. توجه شود محور $x$ در مقایس لگاریتم رسم شده است. . . . . ۹۸

۵–۱۴ مقایسه میانگین نتایج SSNR به ازاء مقادیر زیر منابع (CSM) ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+مرد (خط مربع شکل،□ ) ، ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط زن+زن (خط لوزی شکل ،◇ ) و ۲۰ سیگنال جدا شده‌از ۱۰ سیگنال مخلوط مرد+زن (خط دایره‌ای شکل ،○ )، با استفاده از روش ماسک نرم (Soft). توجه شود تعداد زیر منابع(CSM) برای این آزمون برابر ۶۴ قرار داده شده است. . . . . ۹۹

۷–۱ الگوریتم پیشنهادی توسط پاول برای تخمین پوش طیف سیگنال گفتار. . . . . ۱۰۴

۷–۲ روش بازسازی طیف یک فریم واکدار سیگنال گفتار.(a) طیف اصلی بهمراه پوش طیف.(b) طیف بازسازی شده‌بهمراه پوش طیف. . . . . ۱۰۵

۳-۷ روش بازسازی طیف یک فریم بی واک سیگنال گفتار.(a) طیف اصلی بهمراه پوش طیف.(b) طیف بازسازی شده‌بهمراه پوش طیف. . . . . ۱۰۵

# لیست جداول


۳-۱ نسبت سیگنال به نویز (SNR) برای سیگنال های جدا شده از مخلوط های زن+زن برای : (۱) مدل پیشنهادی (Proposed) (سطر سوم) ، (۲) بااستفاده از سیگنال تحریک (Exciation (سطر چهارم) ، و (۳) با استفاده از لگاریتم طیف (Spectrum) (سطر پنجم). سطر اول شماره سیگنال مخلوط و سطر دوم شماره سیگنال جدا شده را نشان می دهد. . . . . ۴۸

۳-۲ نسبت سیگنال به نویز (SNR) برای سیگنال های جدا شده از مخلوط های مرد+مرد برای : (۱) مدل پیشنهادی (Proposed) (سطر سوم) ، (۲) بااستفاده از سیگنال تحریک (Excitation) (سطر چهارم) ، و (۳) با استفاده از لگاریتم طیف (Spectrum)(سطر پنجم). سطر اول شماره سیگنال مخلوط و سطر دوم شماره سیگنال جدا شده را نشان می دهد. . . . . ۴۸

۳-۳ نسبت سیگنال به نویز (SNR) برای سیگنال های جدا شده از مخلوط های زن+مرد برای : (۱) مدل پیشنهادی (Proposed) (سطر سوم) ، (۲) بااستفاده از سیگنال تحریک (Exciation) (سطر چهارم) ، و (۳) با استفاده از لگاریتم طیف (Spectrum) (سطر پنجم). سطر اول شماره سیگنال مخلوط و سطر دوم شماره سیگنال جدا شده را نشان می دهد. . . . . ۴۸

**فهرست نمادهای اختصاری :**

| نماد | عبارت کامل | معنی |
|---|---|---|
| CASA | Computaional Auditory Scene Analysis | آنالیز عملکرد محیط شنوایی |
| BSS | Blind Source Separation | جداسازی کور منابع |
| $c_i$ | Codebook of speker i | کتاب کد گوینده $i$ |
| DCT | Disrete Cosine Transform | تبدیل کسینوسی گسسته |
| D | FFT order | رتبه تبدیل فوریه |
| $e$ | Excitation Signal | سیگنال تحریک |
| $E\{\cdot\}$ | Expectation Operation | عملگر امید ریاضی |
| FER | Fine Error Rate | نرخ خطای ناچیز |
| $f_o$ | Fundamental Frequency | فرکانس اصلی |
| $f(\cdot)$ | Probability Density Function | تابع چگالی احتمال |
| $\mathcal{F}_D(\cdot)$ | Discrete Fourier Transform | تبدیل فوریه گسسته |
| GMM | Gaussian Mixture Model | مدل مخلوط گوسی |
| GER | Gross Error Rate | نرخ خطای قابل توجه |
| HMM | Hidden Morkov Model | مدل مارکوف پنهان |
| $h$ | Vocal Tract-related Filter | فیلتر لوله صوتی |
| HMC | Harmonic Matching Classifer | کلاسه بندی کننده تطبیقی هارمونیکی |
| ICA | Independent Component Analysis | آنالیز مولفه های مستقل |
| $K$ | Number of Gaussian Mixture | تعداد مولفه های گوسی |
| $\lambda$ | Scale factor | عامل مقایسه |
| LP | Linaer Prediction | پیش بینی خطی |
| $\mathcal{L}$ | Laplas Distribution | توزیع لاپلاس |
| MIXMAX | Mixture-Maximaization | ماکزیمم مخلوط |
| MMSE | Minimum Mean Square Error | کمینه متوسط مربع خطا |
| MAP | Maximum a posteriori | بیشنه توابع پیشین |
| ML | Maximum Likelihood | بیشنه درستنمای |

**فهرست نمادهای اختصاری :**

| نماد | عبارت کامل | معنی |
|---|---|---|
| MLT | Modulated Lapped Transform | تبدیل همپوشانی ادغام شده |
| $N$ | Window Length | طول پنجره |
| $\mathcal{N}$ | Normal Distribution | توزیع نرمال |
| PDF | Probability Density Function | تابع چگالی احتمال |
| $\phi_y$ | Phase of Mixed signal | فاز سیگنال مخلوط |
| SCSS | Single Channel Speech separation | جداسازی سیگنال در حالت تک کاناله |
| SD | Spectral Distortion | اعوجاج طیفی |
| SIR | Speaker Identification Rtae | نرخ شناسای گوینده |
| $\sigma$ | Variance | واریانس |
| TIR | Target to Interference Ratio | نسبت سیگنال هدف به سیگنال تداخلی |
| $\mu$ | Mean Vector | بردار میانگین |
| U | Unvoiced Frame | فریم بی واک |
| V | Voiced Frame | فریم واکدار |
| USD | Usable Speech Detection | آشکار سازی گفتار قابل استفاده |
| VQ | Vector Quantization | چندی کردن برداری |
| $W$ | Applied Window | پنجره مورد استعمال |
| $y$ | Mixed Signal | سیگنال مخلوط |