

طرح دو فیلتر جدید برای بهبود کیفیت گفتار مبتنی بر توزیع احتمال پسین برای ضرایب موجک

بهزاد زمانی دهکردی^۱ احمد اکبری^۱ بابک ناصرشریف^{۱و۲}

^۱ دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران
^۲ دانشکده فنی و مهندسی، دانشگاه گیلان، گیلان، ایران

چکیده

روش‌های حذف نویز مبتنی بر موجک عموماً بر اساس آستانه‌گذاری استوارند که این امر نیز ریشه در فرض تمرکز انرژی سیگنال در تعداد کمی از ضرایب تبدیل موجک سیگنال دارد. عیب عمده روش‌های آستانه‌گذاری در سیگنال گفتار، تخریب گفتار در فرکانس‌های بالا و عملکرد ضعیف آنها در حضور نویزهای رنگی است. از این‌رو دو روش فیلتر کردن ضرایب تبدیل موجک به عنوان جایگزینی برای روش‌های آستانه‌گذاری مطرح می‌گردد. در مقاله حاضر ابتدا با انگیزه نشان دادن برتری روش‌های فیلتر کردن بر آستانه‌گذاری، دو روش حذف نویز آستانه‌گذاری نرم با حد آستانه SURE و فیلتر وینر با هم مقایسه شده‌اند. علاوه بر این دو فیلتر جدید بر مبنای توزیع احتمال پسین ضرایب تبدیل موجک پیشنهاد شده است. نتایج کار حاضر نشان می‌دهند که فیلتر وینر نسبت به روش آستانه‌گذاری، دارای عملکرد بهتر و کیفیت شنیداری بالاتری می‌باشد. به علاوه، مقایسه فیلترهای پیشنهادی با فیلتر وینر در حوزه موجک نشان می‌دهد که این فیلترها دارای کارایی بالاتری نسبت به فیلتر وینر هستند، همچنین فیلترهای پیشنهادی در حوزه موجک نسبت به همتایشان در حوزه فرکانس، نسبت سیگنال به نویز را بیشتر افزایش داده و اعوجاج کمتری بر روی سیگنال تمیز ایجاد می‌نمایند.

کلمات کلیدی: تبدیل موجک، فیلتر وینر، درخت موجک بسته‌ای، آستانه‌گذاری، MAP.

۱- مقدمه

محیط کار یا نویز مربوط به وزش باد در محیط آزاد نیز از جمله عوامل افت کیفیت صوت در سیستم‌های هدست می‌باشد. یکی از کاربردهای دیگر بهبود گفتار سیستم‌های بازشناسی گفتار می‌باشند. معمولاً سیستم‌های بازشناسی گفتار با سیگنال‌های صوتی ضبط شده در محیط آزمایشگاهی آموزش داده می‌شوند، در حالیکه در عمل باید در محیط‌های واقعی که در آن نویز وجود دارد مورد استفاده قرار گیرند. در این حالت برای جلوگیری از پایین آمدن دقت بازشناسی، نیاز به حذف نویز از گفتار احساس می‌شود. نمونه دیگری از کاربرد حذف نویز کدکننده‌هایی هستند که برای کد کردن گفتار بکار می‌روند. این کدکننده‌ها معمولاً برای کار با سیگنال تمیز طراحی می‌گردند حال آنکه در عمل برای شرایط واقعی مورد استفاده قرار می‌گیرند که سیگنال گفتار آغشته به نویز محیط می‌باشد.

استفاده از بهبود گفتار، در بسیاری از کاربردهایی که صدا در یک محیط نویزی دریافت می‌گردد، ضروری است. از جمله این موارد به دریافت صدا از طریق میکروفن تلفن، بدون استفاده از گوشی^۱ و نیز میکروفن‌های رومیزی همه‌جهته در یک کنفرانس صوتی می‌توان اشاره کرد. هدست‌هایی که روی گوش قرار می‌گیرند و با تکنولوژی بلوتوث^۲ با تلفن همراه ارتباط برقرار می‌کنند نیز برای حفظ کیفیت صوت به بهبود صوت نیاز دارند. در این دستگاه‌ها به علت فاصله چند سانتیمتری میکروفن با دهان گوینده صدای مربوط به نویز محیط اطراف نیز دریافت می‌گردد و موجب افت کیفیت صوت می‌شود. سر و صدای اضافی مانند همهمه افراد در

می‌شود و حد آستانه در باندهای بالایی برای فریم واکنش افزایش و برای فریم بی‌واک کاهش می‌یابد [۹]. در کل روش‌های آستانه در حذف نویزهای سفید و ایستا خوب عمل می‌کنند. در حالیکه عملکرد خوبی در نویزهای رنگی ندارند. برای حل این مشکل در سال ۲۰۰۳، Wang و Lu روش آستانه‌گذاری در باندهای بحرانی درخت تجزیه موجک را ارائه دادند. در این تکنیک ابتدا آستانه ماسک نویز، برای هر فریم تخمین زده شده و سپس این آستانه برای تنظیم آستانه نهایی بکار گرفته می‌شود. اندازه آستانه بطور اتوماتیک در باندهای بالا در طول یک سگمنت واکنش و در باندهای پائین در طول یک سگمنت بی‌واک افزایش می‌یابد [۱۰].

روش‌های آستانه‌گذاری از نظر نتایج شنوائی [۱۱، ۱۲] دارای عملکرد ضعیفی هستند [۱۳، ۱۴، ۳۴]. برای حل مشکل روش‌های حد آستانه، استفاده از فیلتر مطرح شده است که در آن پارامترهای فیلتر وابسته به شرایط آماری ضرایب تبدیل موجک تنظیم می‌گردد. فیلتر وینر و فیلتر تفاضل طیف و فیلتر کالمن نمونه‌هایی از فیلترهای بکار رفته در حوزه تبدیل موجک هستند [۱۵، ۱۶، ۱۷]. بطور عمده روش‌های طراحی فیلتر بهبود گفتار در حوزه موجک با الهام از اصول روش‌های طراحی فیلتر بهبود گفتار در حوزه فرکانس انجام می‌شوند. در مقاله حاضر، طراحی فیلتر در حوزه موجک و بر اساس بیشینه کردن احتمال پسین برای ضرایب موجک سیگنال تمیز با الهام از مرجع [۱۸، ۱۹] در حوزه تبدیل فوری مورد توجه قرار گرفته است. از آنجا که روش کمینه کردن خطای میانگین مربعات (MMSE) و میزان نویز غیر رنگی باقیمانده در مرجع [۱۸]، نیازمند محاسبات نمائی و توابع بسل است، از همین رو استفاده از فیلترهایی که با محاسبات کمتر عملکردی نزدیک به روش کمینه کردن خطای میانگین مربعات داشته باشند، در مرجع [۱۹] مد نظر قرار گرفته است. در مقاله حاضر بر همین روال عمل شده است. بر این اساس دو فیلتر در حوزه موجک در این مقاله پیشنهاد شده‌اند که در مقایسه با فیلتر وینر نتایج بهتری داشته‌اند. همچنین در مقایسه با فیلترهای مشابهشان در حوزه تبدیل فوری، دارای نتایج بهتری بوده‌اند.

ساختار ادامه مقاله به این صورت است که بخش دوم مقاله، به مروری کلی بر تبدیل موجک و معرفی انواع درخت تجزیه موجک اختصاص دارد. بخش سوم به بررسی ویژگیهای موجک و دو گرایش کلی حذف نویز در حوزه موجک از قبیل فیلتر کردن وینر و آستانه‌گذاری نرم پرداخته می‌شود. در ادامه در بخش چهارم با استفاده از دو معیار MAP-G و MAP-SG دو فیلتر در حوزه تبدیل موجک پیشنهاد شده است. بخش پنجم نیز به مقایسه دو روش فیلتر کردن وینر و روش آستانه‌گذاری نرم اختصاص دارد. سپس در بخش ششم نتایج بدست آمده از پیاده‌سازی دو فیلتر پیشنهادی در برابر فیلتر وینر در دو حوزه موجک و فرکانس آورده شده است. در آخر نیز به جمع‌بندی کلی اختصاص دارد.

۲- تبدیل موجک

تبدیل موجک پیوسته (که انتگرال موجک نیز نامیده می‌شود) توسط Grossman و Morlet پایه گذاری شده است [۲۰]. تئوری موجک بر مبنای یک مجموعه از فیلترهایی است که توسط انبساط و انقباض موجک مولد ایجاد می‌شوند. همه موجک‌ها از مقیاس موجک مادر بدست می‌آیند. بنابراین فقط کافی است که یک فیلتر طراحی شده و بقیه فیلترها با تغییر مقیاس در زمان و فرکانس روی این فیلتر بدست می‌آیند.

شکل ۱ ساختار ساده تبدیل موجک را نشان می‌دهد. در این تبدیل سیگنال ورودی به دو زیرباند بالا و پائین تفکیک می‌شود. در ادامه برای بازسازی از دو فیلتر دیگر استفاده می‌شود. آنچه در شکل دیده می‌شود درخت موجکی با یک سطح می‌باشد. $h_a(k)$ فیلتر پائین‌گذر و $h_d(k)$ فیلتر بالاگذر می‌باشند. پاسخ فرکانسی آنها وابسته به نوع تابع پایه موجک می‌باشد و برای تجزیه سیگنال به کار

در این حالت نیز بهبود گفتار و حذف نویز می‌تواند صحت عملکرد کدکننده را افزایش دهد.

برای حذف نویز از سیگنال گفتار، می‌توان طیف فرکانسی سیگنال نویزی را به گونه‌ای اصلاح نمود که پس از بازیابی سیگنال در حوزه زمان، صداهای پس زمینه گفتار حذف گردد و یا کاسته شود. به این منظور ابتدا باید سیگنال گفتار با تبدیلی نظیر تبدیل فوری به حوزه فرکانس انتقال یابد و سپس اصلاحات لازم بر روی طیف فرکانسی به شیوه‌ای صورت گیرد تا نویز کاهش یابد. از جمله‌ی شناخته شده‌ترین این روشها، روش تفاضل طیف توان و نیز فیلتر وینر^۲ اعمال شده بر طیف می‌باشد [۱]. گرچه این روش‌ها در عمل موفق بوده‌اند و پیاده سازی آنها نیز بسیار ساده می‌باشد. با این حال مشکلات و کاستی‌هایی نیز دارند. یکی از معایب این روشها اعوجاجی است که روی سیگنال مفید ایجاد می‌نمایند. مشکل دیگر ایجاد یک نوع نویز پس زمینه مصنوعی در این روشهاست که به نویز موسیقی مشهور شده است. علاوه بر این بنا به ماهیت تبدیل فوری و نایستانت بودن سیگنال گفتار، عملیات تفاضل طیف و یا فیلتر کردن باید بر روی طیف یک پنجره محدود از سیگنال صورت گیرد که کوچک بودن پنجره عموماً موجب محدودیت در رزولوشن فرکانسی طیف سیگنال صوتی می‌گردد و کارائی اصلاح طیف و یا فیلتر کردن پائین می‌آید.

از دیگر روش‌های بهبود گفتار استفاده از تبدیل موجک است. ویژگی اصلی تبدیل موجک بهره جستن از پنجره‌های زمانی با طول متفاوت برای باندهای فرکانسی مختلف است. به این ترتیب با استفاده از تبدیل موجک می‌توان به دقت فرکانسی بالا در باندهای فرکانسی پایین در عین حفظ دقت در رزولوشن زمانی دست یافت. به این ترتیب تبدیل موجک محدودیت‌های ناشی از کار با پنجره محدود در تبدیل فوری را ندارد. مهمترین روش برای بهبود گفتار با استفاده از تبدیل موجک، روش آستانه‌گذاری ضرایب تبدیل موجک است. این روش بر این اساس عمل می‌کند که سیگنال گفتار همانند بسیاری از سیگنالهای دیگر، انرژی سیگنال غالباً در تعداد کمی از ضرایب تبدیل موجک متمرکز است. این ضرایب نسبت به ضرایب دیگر تبدیل موجک خود سیگنال، یا ضرایب تبدیل موجک هر سیگنال دیگری (بویژه نویز) که انرژی آن بر روی تعداد زیادی از ضرایب گسترش می‌یابد، بزرگتر هستند. بنابراین می‌توان با صفر کردن ضرایب کوچکتر، علاوه بر حفظ اطلاعات مهم سیگنال اصلی، نویز را نیز محدود کرد. مطابق با این ویژگی، ضرایب تبدیل موجک سیگنال با یک حد آستانه^۳ مقایسه می‌شوند و مقدار ضرایب کوچکتر از این حد آستانه برابر صفر قرار داده می‌شوند. این عمل را می‌توان به نوعی، فیلتر کردن ضرایب تبدیل موجک دانست. روش‌های حد آستانه علاوه بر حذف نویز مستقیم از ضرایب موجک، برای تخمین طیف فرکانسی سیگنال گفتار نیز بکار گرفته شده‌اند [۲، ۳، ۴].

در آستانه‌گذاری ضرایب موجک دو مسئله مطرح می‌شود: شیوه تعیین حد آستانه و نحوه اعمال حد آستانه بر ضرایب. مسئله دوم با تعیین یک تابع موسوم به تابع آستانه‌گذاری حل می‌شود. توابع آستانه‌گذاری متفاوتی همچون توابع آستانه‌گذاری نرم، سخت، نیمه‌نرم و گاروت و گاروت پله‌ای^۵ در کارهای تحقیقاتی متفاوت تعریف شده‌اند [۵، ۳۴]. برای تعیین حد آستانه روش‌های متفاوتی وجود دارد که حد آستانه را مستقل یا وابسته به تابع آستانه‌گذاری تعیین می‌نمایند. روش‌هایی چون کمینه کردن نرخ یابش نادرست^۶ (MFDR) [۵] و روش تاپ^۷ [۶] حد آستانه را مستقل از نوع تابع اعمال حد آستانه محاسبه می‌نمایند. روش کمینه کردن تخمین مخاطره نا ارب استین^۸ (SURE) نیز روش دیگری برای تعیین حد آستانه است که به نوع تابع اعمال حد آستانه وابسته است.

رایج‌ترین روش تعیین حد آستانه، روشی است که Johnston و Donoho برای حذف نویز سفید گوسی پیشنهاد داده‌اند. این روش برای نویزهای رنگی به خوبی عمل نمی‌کند [۷، ۸] شیخ‌زاده و ابوطالبی در سال ۲۰۰۱ روشی بهبود یافته پیشنهاد داده‌اند که در آن سیگنال به فریم‌های واکنش و بی‌واک تقسیم‌بندی

شکل ۴ یک درخت موجک بسته ای که دارای ۴ زیرباند یکنواخت می باشد را نشان می دهد. در درخت موجک بسته ای همه زیرباندها و نودهای داخلی درخت به یک میزان رشد می کنند، در نتیجه تحلیل فرکانسی در تمام محدوده فرکانسی یکسان خواهد بود.

۳- معرفی دو گرایش مختلف برای حذف نویز از ضرایب موجک

روش های حذف نویز از سیگنال گفتار در حوزه فرکانس، از این واقعیت استفاده می نمایند که طیف نویز جمع شونده ایستادن، در حوزه فرکانس قابل تخمین بوده و تکنیک پذیری نویز و سیگنال گفتار در این حوزه بهتر انجام می گیرد. در این روش ها فیلتر حذف نویز با استفاده از تخمین طیف نویز محاسبه می گردد و بر سیگنال نویزی اعمال می شود. برای تخمین پارامترهای فیلتر بایستی سیگنال فریم بندی شود. از طرفی در تبدیل فوریه نسبت تعداد نمونه ها در حوزه زمان با تعداد زیرباندها برابر می باشد. بنابراین برای داشتن جزئیات فرکانسی بیشتر باید پنجره آنالیزی بزرگتری در نظر بگیریم. انتخاب پنجره آنالیزی بزرگ از یک طرف موجب می گردد جزئیات زمانی سیگنال در نظر گرفته نشود و قسمتهای گذرای سیگنال در طی پردازش فرکانسی آسیب ببینند و از طرف دیگر به علت غیر ایستادن بودن سیگنال در پنجره های بزرگ دقت آنالیز زمان فرکانس کم شود. به علاوه با میانگین گیری از نمونه های دور گذشته در نتایج شنوائی اثر آکو نیز ظاهر می گردد. کوچک کردن پنجره آنالیز نیز موجب کم شدن جزئیات فرکانسی سیگنال شده و موجب اعوجاج در سیگنال بهبود یافته می گردد. تبدیل موجک با فراهم کردن مقیاس زمان - فرکانس دلخواه این عیب را برطرف نموده است.

با استفاده از تبدیل موجک می توان زیرباندها را بطور دلخواه روی بازه فرکانس توزیع کرد و توزیع زیرباندها را به سیستم شنوائی انسان نزدیک نمود [۲۱، ۲۲]. بنابراین حوزه موجک برای پردازش گفتار مناسب می باشد [۲۳]. در ضمن می دانیم که انواع نویزها توزیع مختلفی روی زیرباندهای فرکانسی دارند، از همین رو می توان توزیع زیرباندها را نیز بر اساس نوع نویز تعیین کرد. تبدیل موجک می تواند از توابع پایه مختلف بهره بگیرد و تابعی را که دارای عملکرد بهتری است، بکار برد.

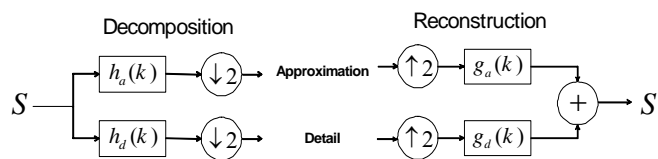
از مزایای دیگر موجک نسبت به فوریه بار محاسباتی کمتر آن می باشد [۷، ۲۴]. می دانیم که در سیگنال گفتار فاز از اهمیت خاصی برخوردار است، در تبدیل موجک فاز سیگنال جدا نمی گردد، بنابراین فاز نیز در پروسه بهبود سیگنال گفتار بکار برده می شود. در صورتی که در اغلب روش های حذف نویز در حوزه فرکانس فرض می کنند فاز سیگنال بهبود یافته با فاز سیگنال نویزی برابر است. با توجه به این ویژگی ها در این مقاله به روش های حذف نویز مبتنی بر موجک پرداخته شده است. بطور کلی می توان گفت، حذف نویز از ضرایب موجک به دو روش آستانه گذاری یا اعمال فیلتر نظیر فیلتر وینر در زیرباندها انجام می گیرد که هر یک از دو روش در ادامه شرح داده شده است.

۳-۱- گرایش فیلتر کردن ضرایب موجک - فیلتر وینر

برای محاسبه فیلتر وینر در حوزه موجک به صورت زیر عمل می شود. فرض کنید سیگنال گفتار تمیز S با نویز جمع شونده n آغشته شده باشد. در این صورت سیگنال نویزی x در حوزه زمان از رابطه (۱) تبعیت می کند.

$$x(i, m) = s(i, m) + n(i, m), \quad i = 0, 1, \dots, L-1 \quad (1)$$

می روند. خروجی فیلتر پائین گذر ضرایب تقریب و خروجی فیلتر بالا گذر ضرایب تفصیل را تشکیل می دهند.

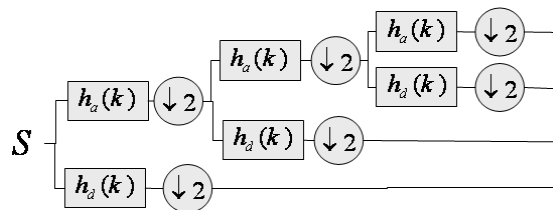


شکل ۱- ساختار تبدیل و عکس تبدیل موجک

مجموعه موجکها به عنوان یک بانک فیلتر برای آنالیز سیگنال می تواند در نظر گرفته شود. در این راستا، بانک فیلتر موجک را به صورت درخت نیز نمایش می دهند. می توان تبدیل موجک را بفرم درخت موجک گسسته دوتائی و درخت موجک بسته ای^۱ در نظر گرفت.

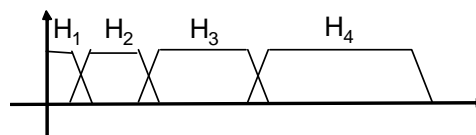
۲-۱- درخت موجک گسسته دوتائی و موجک بسته ای

شکل ۲ ساختار یک درخت بانک فیلتر را نشان می دهد که هم ارز با تبدیل موجک گسسته دوتائی روی یک دنباله می باشد. در این نوع درخت گسترش درخت از طرف زیرباندهای پائین یعنی باندهای فرکانس پائین می باشد. بنابراین تحلیل فرکانسی دقیق تری در این زیرباندها داریم.

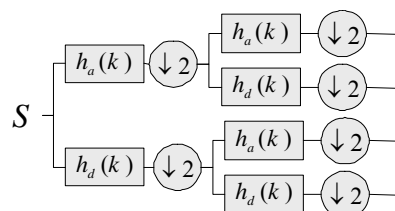


شکل ۲- درخت موجک گسسته دوتائی با ۴ زیرباند نامساوی

در مرحله اول طیف سیگنال ورودی به دو زیرباند بالائی و زیرباند پائینی تفکیک شده، سپس زیرباند پائینی دوباره به دو زیرباند پائینی و بالائی تفکیک می شود و این عمل برای زیرباند پائینی حاصل تکرار شده تا در مجموع ۴ زیرباند نهائی ایجاد شوند. شکل ۳ مشخصه طیفی فیلتربانک حاصل را نشان می دهد.



شکل ۳- مشخصه طیفی فیلتربانک موجک نشان داده شده در شکل ۲



شکل ۴- درخت موجک بسته ای با ۴ زیرباند یکنواخت

دیدگاه دیگر برای بدست آوردن فیلتر این است که برای هر نمونه از یک زیرباند از یک فریم، یک فیلتر محاسبه شود. که آنرا می‌توان در رابطه (۸) نشان داد:

$$\hat{S}_j(i, m) = H_j(i, m) \cdot X_j(i, m) \quad (8)$$

که با توجه به این امر فیلتر وینر را مجدداً بفرم رابطه (۹) تعریف می‌کنیم:

$$H_j(i, m) = \frac{R_{pr,j}(i, m)}{R_{pr,j}(i, m) + 1} \quad (9)$$

$R_{pr,j}(i, m)$ را بفرم رابطه (۱۰) که یک رابطه بازگشتی می‌باشد، تعریف می‌کنیم. $R_{pr,j}(i, m)$ در واقع نسبت واریانس سیگنال تمیز به واریانس سیگنال نویز در طی فریم‌های متوالی می‌باشد.

$$R_{pr,j}(i, m) = \alpha R_{pr,j}(i, m-1) + (1-\alpha) \cdot \max\left\{\frac{\hat{S}_j^2(i, m)}{\hat{N}_j^2(i, m)}, 0\right\} \quad (10)$$

برای تخمین $S_j^2(i, m)$ از تفاضل سیگنال نویزی و سیگنال نویز تخمین زده شده استفاده می‌شود. در این رابطه α فاکتور فراموشی می‌باشد که بر اساس روش واریانس بالانس شده مطرح شده در [۲۵] بدست می‌آید.

۳-۲- گرایش آستانه‌گذاری روی ضرایب موجک - آستانه‌گذاری نرم

ایده اصلی آستانه‌گذاری روی ضرایب تبدیل موجک بر اساس این واقعیت است که در سیگنال گفتار، انرژی سیگنال غالباً در تعداد کمی از ضرایب تبدیل موجک متمرکز است. این ضرایب نسبت به ضرایب دیگر تبدیل موجک خود سیگنال، یا ضرایب تبدیل موجک هر سیگنال دیگری از جمله نویز که انرژی آن بر روی تعداد زیادی از ضرایب گسترش می‌یابد، بزرگتر هستند. بنابراین می‌توان با صفر کردن ضرایب کوچکتر، علاوه بر حفظ اطلاعات مهم سیگنال اصلی، نویز را نیز محدود کرد. مطابق با این ویژگی، ضرایب تبدیل موجک سیگنال با یک حد آستانه مقایسه می‌شوند و مقدار ضرایب کوچکتر از این حد آستانه برابر صفر قرار داده می‌شوند. از روش‌های آستانه‌گذاری می‌توان آستانه‌گذاری نرم [۸، ۲۶]، آستانه‌گذاری سخت [۸]، آستانه‌گذاری Quantile، آستانه‌گذاری نیمه نرم [۲۷]، آستانه‌گذاری با روش وفقی BayesShrink [۲۸، ۲۹] و آستانه‌گذاری سخت اصلاح شده [۲۹] را نام برد. در زیر به بررسی روش آستانه‌گذاری نرم می‌پردازیم که با رابطه (۱۱) نشان داده شده است.

$$\hat{S}_j(i, m) = \begin{cases} 0 & \text{if } |X_j(i, m)| < T_j(m) \\ \text{sgn}(X_j(i, m)) (|X_j(i, m)| - T_j(m)) & \text{otherwise} \end{cases} \quad (11)$$

که در آن $X_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال گفتار نویزی در زیرباند \hat{m} از فریم \hat{m} و $\hat{S}_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال بهبودیافته در زیرباند \hat{m} از فریم \hat{m} است. همچنین $T_j(m)$ حد آستانه زیرباند \hat{m} از فریم \hat{m} می‌باشد، در ادامه جزئیات نحوه محاسبه انواع آستانه آمده است.

که در آن i اندیس زمان، L طول فریم و m اندیس فریم می‌باشد. در این صورت اگر بر رابطه (۱) تبدیل موجک را اعمال نمائیم، خواهیم داشت:

$$X_j(i, m) = S_j(i, m) + N_j(i, m), \quad i = 0, 1, \dots, L_j - 1 \quad (2)$$

که در آن $X_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال گفتار نویزی در زیرباند \hat{m} از فریم \hat{m} ، و $S_j(i, m)$ و $N_j(i, m)$ نمونه متناسب برای سیگنال تمیز و سیگنال نویز می‌باشند. و L_j طول زیرباند \hat{m} می‌باشد. می‌توان مطابق رابطه (۳) برای بدست آوردن سیگنال بهبود یافته سیگنال نویزی را فیلتر نمود، در این رابطه برای هر زیرباند از یک فریم یک فیلتر بدست می‌آید:

$$\hat{S}_j(i, m) = H_j(m) \cdot X_j(i, m) \quad (3)$$

در صورتی که از فیلتر وینر برای پیاده‌سازی رابطه بالا استفاده کنیم تابع تبدیل آن بصورت رابطه (۴) تعریف می‌شود:

$$H_j(m) = \frac{E\{S_j^2(i, m)\}}{E\{S_j^2(i, m)\} + \text{est}\{N_j^2(m)\}} \quad (4)$$

که در آن $H_j(m)$ اندازه فیلتر برای زیرباند \hat{m} از فریم \hat{m} می‌باشد که در رابطه (۳) بکار می‌رود. $E\{S_j^2(i, m)\}$ انرژی زیرباند \hat{m} از فریم \hat{m} سیگنال تمیز و $\text{est}\{N_j^2(m)\}$ تخمین انرژی نویز در زیرباند \hat{m} از فریم \hat{m} می‌باشد که از رابطه (۵) بدست می‌آید.

$$\text{est}\{N_j^2(m)\} = \alpha \cdot \text{est}\{N_j^2(m-1)\} + (1-\alpha) \cdot E\{\hat{N}_j^2(i, m)\} \quad (5)$$

که $E\{\hat{N}_j^2(i, m)\}$ انرژی زیرباند \hat{m} از فریم \hat{m} نویز تخمین زده شده، $\text{est}\{N_j^2(m-1)\}$ تخمین انرژی نویز در زیرباند \hat{m} از فریم $(m-1)$ و α فاکتور فراموشی می‌باشد که بر اساس روش واریانس بالانس شده مطرح شده در [۲۵] بدست می‌آید. در این روش اندازه ضریب فراموشی به طول فریم بستگی دارد به این صورت که α برای فریم‌های بزرگ به یک نزدیک شده و برای فریم‌های کوچک به صفر نزدیک می‌شود. رابطه محاسبه α به در رابطه (۶) آمده است:

$$\alpha = \frac{T_{sm} \times f_s - L/2}{T_{sm} \times f_s + L/2} \quad (6)$$

که L طول فریم، f_s فرکانس نمونه‌برداری سیگنال ورودی و T_{sm} زمان تثبیت شدن تخمین توان سیگنال می‌باشد که معمولاً برابر 0.2 ثانیه در نظر گرفته می‌شود [۲۵]. $E\{S_j^2(i, m)\}$ انرژی سیگنال تمیز در زیرباند \hat{m} از فریم \hat{m} از رابطه (۷) بدست می‌آید.

در حالت ایده‌آل برای تخمین‌ها از سیگنال‌های نویز و تمیزی که در دسترس هستند استفاده می‌شود. ولی در شرایط واقعی که فقط سیگنال نویزی در اختیار است، از مناطقی از سیگنال که فقط نویز وجود دارد برای به روز کردن تخمین انرژی نویز استفاده می‌شود و تا یافتن بخش دیگری از سیگنال که فقط نویز در آن حاضر باشد، از تخمین قبلی نویز استفاده می‌شود.

$$E\{S_j^2(i, m)\} = E\{X_j^2(i, m)\} - \text{est}\{N_j^2(m)\} \quad (7)$$

۳-۲-۱- انتخاب آستانه

روش‌های مختلفی را می‌توان برای انتخاب حد آستانه در نظر گرفت، اما در اغلب کاربردها، روش پیشنهاد شده توسط Donoho [۸] مورد استفاده قرار می‌گیرد که به فرم رابطه (۱۲) محاسبه می‌شود.

$$T = \hat{\sigma}_n \sqrt{2 \log(N)} \quad (12)$$

که N ، طول فریم جاری و $\hat{\sigma}_n$ تخمینی از انحراف معیار کلی نویز است. در برخی از منابع این روش تعیین حد آستانه، روش VisuShrink نامیده می‌شود. از آنجا که مقدار واقعی واریانس نویز σ_n^2 در دسترس نیست. در مورد نویز سفید گوسی می‌توان انحراف معیار $\hat{\sigma}_n$ را با توجه به رابطه (۱۳) از روی ضرایب موجک تخمین زد.

$$\hat{\sigma}_n = \frac{MAD_1}{0.6745} = \frac{\text{median}(|W_1^D| - \text{median}(W_1^D))}{0.6745} \quad (13)$$

در این رابطه MAD بیانگر Median Absolute Difference ضرایب موجک تفصیل مربوط به سطح یک است. اگر از تبدیل بسته موجک برای بدست آوردن ضرایب موجک استفاده کرده باشیم، مقدار حد آستانه از رابطه (۱۴) محاسبه می‌شود.

$$T = \hat{\sigma}_n \sqrt{2 \log(N \log_2 N)} \quad (14)$$

روش آستانه‌بندی پیشنهاد شده توسط Donoho را می‌توان توسعه داد و حد آستانه را با توجه به رابطه (۱۵) به طور مجزا برای هر سطح در تبدیل بسته موجک محاسبه کرد [۳۰].

$$T_j = \hat{\sigma}_{n_j} \sqrt{2 \log(N)} \quad (15)$$

انحراف معیار نویز برای هر سطح نیز از رابطه (۱۶) تخمین زده می‌شود.

$$\hat{\sigma}_{n_j} = \frac{MAD_j}{0.6745} \quad (16)$$

که در آن مقدار MAD برای هر سطح از روی ضرایب موجک تفصیل همان سطح محاسبه می‌شود.

$$MAD_j = \text{median}(|W_j^D| - \text{median}(W_j^D)) \quad (17)$$

این روش آستانه‌بندی در منابع مختلف با عنوان روش SureShrink شناخته می‌شود.

در این بخش دو دیدگاه کلی برای حذف نویز در حوزه موجک بررسی شدند. برای هر دیدگاه روش‌های متعددی در مقالات ارائه شده است که در اینجا دو روش مرسوم تر در حوزه موجک مطرح شدند، که در بخش پنجم مقاله به مقایسه این دو روش می‌پردازیم.

۴- معرفی فیلترهای مبتنی بر توزیع احتمال پسین ضرائب موجک

روش حذف نویز ارائه شده توسط Ephraim در حوزه فرکانس [۱۸]، علیرغم بهینه بودن از نقطه نظر کمینه کردن خطای میانگین مربعات (MMSE) روی دامنه سیگنال در حوزه فرکانس، نیازمند محاسبات نمایی و توابع بسل می‌باشد. از همین رو در سال‌های اخیر فیلترهای متعددی که دارای عملکرد مشابه و محاسبات کمتر بودند ارائه شده‌اند از جمله می‌توان به فیلترهای Joint-MAP، MAP-SAE⁺ و MMSE اشاره کرد که همگی در حوزه فرکانس می‌باشند [۱۹]. با توجه به ویژگی‌های مطرح شده برای تبدیل موجک در بخش سوم مقاله که نشان‌دهنده کارآمد بودن موجک برای پردازش گفتار می‌باشد، بر آن شدیم که از این تبدیل بجای تبدیل فوری استفاده کنیم. در اینجا با الهام از مرجع [۱۸] و نیز دو معیار تخمین MAP-G و MAP-SG دو فیلتر حذف نویز در حوزه موجک پیشنهاد می‌نمائیم. طریقه محاسبه فیلترهای پیشنهادی در ادامه توضیح داده می‌شوند. به علاوه در بخش‌های بعدی این فیلترها پیاده‌سازی و ارزیابی خواهند شد.

۴-۱- پیشنهاد معیار MAP-G برای محاسبه فیلتر

معیار MAP-G توزیع احتمال پسین سیگنال تمیز را بشرط داشتن سیگنال نویزی ماکزیمم می‌کند. اگر سیگنال گفتار و نویز دارای توزیع تصادفی گوسی با میانگین صفر و واریانس σ_s^2 و σ_n^2 باشند، می‌توان احتمال سیگنال تمیز را بشرط داشتن سیگنال نویزی بفرم رابطه (۱۸) نوشت. در [۳۱] با استفاده از معیار Kurtosis نشان داده شده است که ضرائب موجک در زیرباند‌های درخت تجزیه موجک بسته‌ای دارای توزیع گوسی با واریانس کم می‌باشند.

$$\begin{aligned} p(S_j(i, m) | X_j(i, m)) & \quad (18) \\ & \propto p(X_j(i, m) | S_j(i, m)) p(S_j(i, m)) \\ & \propto \frac{S_j(i, m)}{\pi^2 \sigma_{S_j(i)}^2 \sigma_{N_j(i)}^2} \exp\left(-\frac{|X_j(i, m) - S_j(i, m)|^2}{\sigma_{N_j(i)}^2} - \frac{S_j^2(i, m)}{\sigma_{S_j(i)}^2}\right) \end{aligned}$$

که $\sigma_{S_j(i)}^2 = E[S_j^2(i)]$ و $\sigma_{N_j(i)}^2 = E[\hat{N}_j^2(i)]$ به ترتیب واریانس سیگنال نویز و واریانس سیگنال تمیز برای نمونه i ام از زیرباند j ام در طول فریم‌های متوالی می‌باشند. از رابطه فوق لگاریتم طبیعی می‌گیریم، و آنرا J می‌نامیم، بنابراین:

$$J = -\frac{|X_j(i, m) - S_j(i, m)|^2}{\sigma_{N_j(i)}^2} - \frac{S_j^2(i, m)}{\sigma_{S_j(i)}^2} + \ln(S_j(i, m)) + cte \quad (19)$$

که در آن cte یک مقدار ثابت می‌باشد. اگر از رابطه فوق نسبت به $S_j(i, m)$ مشتق می‌گیریم، خواهیم داشت:

$$\frac{\partial}{\partial S_j(i, m)} J = -\frac{2(S_j(i, m) - X_j(i, m))}{\sigma_{N_j(i)}^2} - \frac{2S_j(i, m)}{\sigma_{S_j(i)}^2} + \frac{1}{S_j(i, m)} \quad (20)$$

با مساوی صفر قرار دادن رابطه (۲۰) خواهیم داشت:

که σ_S^2 و σ_N^2 به ترتیب واریانس سیگنال نویز و واریانس سیگنال تمیز برای نمونه نام از زیرباند نام در طول فریم‌های متوالی می‌باشند. ضرایب موجک سیگنال تمیز دارای توزیع super Gaussian بوده که $p(S_j(i, m))$ بفرم زیر خواهد بود:

$$p(S_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^\nu(i, m)}{\sigma_{S_j(i)}^{\nu+1}} \exp\left(-\mu \frac{S_j(i, m)}{\sigma_{S_j(i)}}\right) \quad (27)$$

که Γ تابع گاما و نیز μ و ν پارامترهای توزیع هستند. با جایگذاری روابط (26) و (27) در رابطه (25) داریم:

$$p(S_j(i, m)|X_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^\nu(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2 \sigma_{S_j(i)}^{\nu+1}} \times \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2} - \mu \frac{S_j(i, m)}{\sigma_{S_j(i)}}\right) \times I_0\left(\frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) \quad (28)$$

از سوئی رابطه بسل را می‌توان بفرم رابطه (29) نیز حساب کرد:

$$I_0(|x|) \approx \frac{1}{\sqrt{2\pi}|x|} \exp(|x|) \quad (29)$$

اگر رابطه (29) را در رابطه (28) جایگزین کنیم، داریم:

$$p(S_j(i, m)|X_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^{\nu-\frac{1}{2}}(i, m) X_j^{\frac{1}{2}}(i, m)}{2\sqrt{\pi} \sigma_{N_j(i)} \sigma_{S_j(i)}^{\nu+1}} \times \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2} - \mu \frac{S_j(i, m)}{\sigma_{S_j(i)}} + \frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{X_j^{\frac{1}{2}}(i, m)}{2\sqrt{\pi} \sigma_{N_j(i)} \sigma_{S_j(i)}^{\nu+1}} \exp\left(-\frac{X_j^2(i, m)}{\sigma_{N_j(i)}^2}\right) S_j^{\nu-\frac{1}{2}}(i, m) \times \exp\left(-\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} - S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right)\right) \approx S_j^{\nu-\frac{1}{2}}(i, m) \exp\left(-\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} - S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right)\right) \quad (30)$$

از رابطه فوق لگاریتم طبیعی می‌گیریم، و آنرا J می‌نامیم.

$$J = -\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} \quad (31)$$

$$S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right) + \left(\nu - \frac{1}{2}\right) \ln(S_j(i, m))$$

از رابطه بالا نسبت به $S_j(i, m)$ مشتق می‌گیریم، خواهیم داشت:

$$-2\left(\sigma_{S_j(i)}^2 + \sigma_{N_j(i)}^2\right) S_j^2(i, m) + 2\sigma_{S_j(i)}^2 X_j(i, m) S_j(i, m) + \sigma_{S_j(i)}^2 \cdot \sigma_{N_j(i)}^2 = 0 \quad (21)$$

معادله فوق را بر اساس متغیر $S_j(i, m)$ حل می‌کنیم، برای سادگی روابط بدست آمده، دو مقدار $R_{pr,j}(i, m)$ تعریف شده در رابطه (10) و $R_{po,j}(i, m)$ بفرم رابطه (22) را تعریف می‌کنیم. $R_{po,j}(i, m)$ در واقع نسبت سیگنال به نویز در فریم جاری می‌باشد.

$$R_{po,j}(i, m) = \frac{\hat{S}_j^2(i, m)}{\hat{N}_j^2(i, m)} \quad (22)$$

با این فرض و با حل معادله درجه دوم (21) خواهیم داشت:

$$S_j(i, m) = \left\{ \frac{R_{pr,j}(i, m)}{2(1 + R_{pr,j}(i, m))} + \sqrt{\frac{R_{pr,j}^2(i, m) + 2(1 + R_{pr,j}(i, m)) R_{pr,j}(i, m)}{(R_{po,j}(i, m) + 1) 2(1 + R_{pr,j}(i, m))}} \right\} X_j(i, m) \quad (23)$$

بنابراین فیلتر بفرم رابطه (24) در خواهد آمد:

$$H_j(i, m) = \frac{R_{pr,j}(i, m)}{2(1 + R_{pr,j}(i, m))} + \sqrt{\frac{R_{pr,j}^2(i, m) + 2(1 + R_{pr,j}(i, m)) R_{pr,j}(i, m)}{(R_{po,j}(i, m) + 1) 2(1 + R_{pr,j}(i, m))}} \quad (24)$$

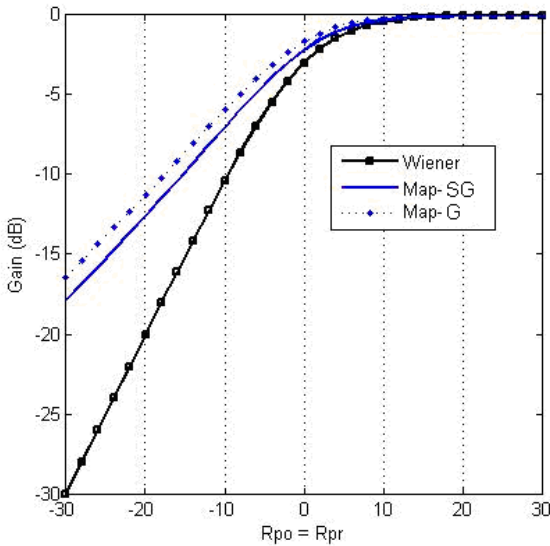
۴-۲- پیشنهاد معیار MAP-SG برای محاسبه فیلتر

در این قسمت نیز با معیار MAP-SG سعی می‌شود تا احتمال تخمین دامنه سیگنال بهبودیافته را با توجه به داشتن دامنه سیگنال نویزی ماکزیمم کنیم که مشابه با معیار MAP-G عمل می‌شود. در این حالت فرض می‌کنیم که ضرایب موجک سیگنال تمیز دارای توزیع super Gaussian بوده و احتمال ضرایب موجک سیگنال نویزی به شرط ضرایب موجک سیگنال تمیز دارای توزیع Rice باشد. با این فرض احتمال $p(S_j(i, m)|X_j(i, m))$ بفرم زیر تعریف می‌شود:

$$p(S_j(i, m)|X_j(i, m)) \propto p(X_j(i, m)|S_j(i, m)) p(S_j(i, m)) \quad (25)$$

ضرایب موجک سیگنال نویزی به شرط ضرایب موجک سیگنال تمیز دارای توزیع Rice می‌باشد که می‌توان در حوزه موجک $p(X_j(i, m)|S_j(i, m))$ بفرم رابطه (26) نوشت: [35]

$$p(X_j(i, m)|S_j(i, m)) = \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2} \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2}\right) I_0\left(\frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) \quad (26)$$

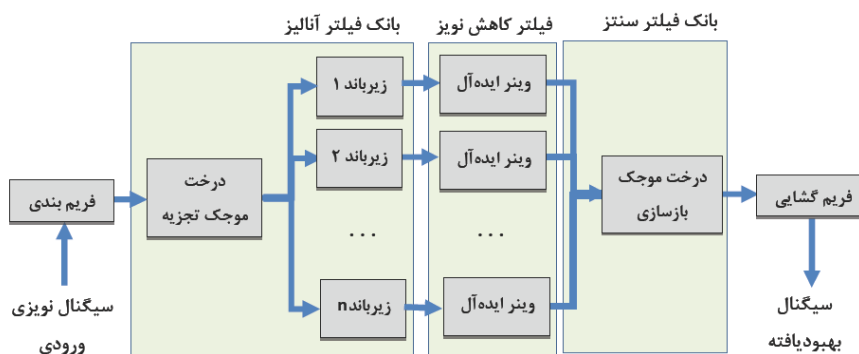


شکل ۵- اندازه سه فیلتر وینر، MAP-G و MAP-SG در حالت بهینه برای مقادیر مختلف R_{pr}

شکل ۵ اندازه سه فیلتر وینر، MAP-G و MAP-SG در حالت بهینه برای مقادیر مختلف R_{pr} را نشان می‌دهد. در این نمودار محور افقی اندازه R_{pr} را بر حسب دسی بل نشان می‌دهد که از -30dB تا +30dB تغییر می‌کند. محور عمودی نیز اندازه فیلتر را بر حسب دسی بل نشان می‌دهد. چنانچه در شکل ۵ دیده می‌شود، در سیگنال به نویزهای پائین، فیلتر وینر نویز بیشتری را نسبت به دو فیلتر دیگر حذف می‌کند. بنابراین انتظار می‌رود که تخریب بیشتری نیز صورت بگیرد. ولی برای نرخ سیگنال به نویز بیشتر از صفر اندازه این سه فیلتر بهم نزدیک می‌شود که انتظار عملکرد یکسانی از سه فیلتر می‌رود. در قسمت ششم مقاله به بررسی این سه فیلتر در دو حوزه فرکانس و موجک پرداخته می‌شود.

۵- مقایسه گرایش‌های حذف نویز از ضرائب موجک معرفی شده در بخش ۳

بمنظور ارزیابی و مقایسه دو روش آستانه‌گذاری و فیلترینگ، روش‌های آستانه‌گذاری نرم و فیلتر وینر با یکدیگر مقایسه شده‌اند. برای فیلتر وینر رابطه (۴) و برای روش آستانه‌گذاری نرم رابطه (۱۲) و تعیین حد آستانه SureShrink یعنی رابطه (۱۶) در شرایط ایده‌آل بکار رفته‌اند، به این صورت که برای بدست آوردن انرژی و انحراف معیار نویز یک زیرباند از خود سیگنال نویز استفاده شده است.



شکل ۶- بلوک دیاگرام سیستم کاهش نویز

$$\frac{\partial}{\partial S_j(i, m)} J = -\frac{2S_j(i, m)}{\sigma_{N_j(i)}^2} - \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2} \right) + \left(\nu - \frac{1}{2} \right) \frac{1}{S_j(i, m)} \quad (32)$$

با مساوی صفر قرار دادن معادله بالا و حل این معادله درجه دوم بر حسب متغیر $S_j(i, m)$ رابطه (۳۳) حاصل می‌شود.

$$S_j(i, m) = \frac{\mu\sigma_{N_j(i)}^2 - X_j(i, m)}{4\sigma_{S_j(i)}} + \sqrt{\left(-\frac{\mu\sigma_{N_j(i)}^2 - X_j(i, m)}{4\sigma_{S_j(i)}} + \frac{X_j(i, m)}{2} \right)^2 - \frac{\sigma_{N_j(i)}^2}{2} \left(\nu - \frac{1}{2} \right)} \quad (33)$$

با ساده‌سازی به رابطه زیر می‌رسیم:

$$S_j(i, m) = \left\{ \frac{\mu\sigma_{N_j(i)}^2}{4\sigma_{S_j(i)}X_j(i, m)} - \frac{1}{2} + \sqrt{\left(-\frac{\mu\sigma_{N_j(i)}^2}{4\sigma_{S_j(i)}X_j(i, m)} + \frac{1}{2} \right)^2 - \frac{\sigma_{N_j(i)}^2}{2X_j^2(i, m)} \left(\nu - \frac{1}{2} \right)} \right\} X_j(i, m) \quad (34)$$

برای سادگی روابط بدست آمده، دو مقدار $R_{pr,j}(i, m)$ و $R_{po,j}(i, m)$ تعریف شده در روابط (۱۰) و (۲۲) بکار می‌بریم. بنابراین فیلتر بفرم زیر در خواهد آمد:

$$H_j(i, m) = \frac{\mu}{4R_{pr,j}(i, m)R_{po,j}(i, m)} - \frac{1}{2} + \sqrt{\left(-\frac{\mu}{4R_{pr,j}(i, m)R_{po,j}(i, m)} + \frac{1}{2} \right)^2 - \frac{1}{2R_{po,j}^2(i, m)} \left(\nu - \frac{1}{2} \right)} \quad (35)$$

در این مقاله مقادیر دو پارامتر μ و ν ، با توجه به [36] بترتیب ۱.۵ و ۰.۰۱ در نظر گرفته شده‌اند. دو فیلتر پیشنهادی با معیارهای MAP-G و MAP-SG به میزان تخمین‌های R_{pr} و R_{po} وابسته می‌باشند. در حالت بهینه دو تخمین R_{pr} و R_{po} برابرند. شکل ۵ اندازه سه فیلتر وینر با رابطه (۹)، MAP-G با رابطه (۲۴) و MAP-SG با رابطه (۳۵) را در حالت بهینه برای مقادیر مختلف R_{pr} نشان می‌دهد.

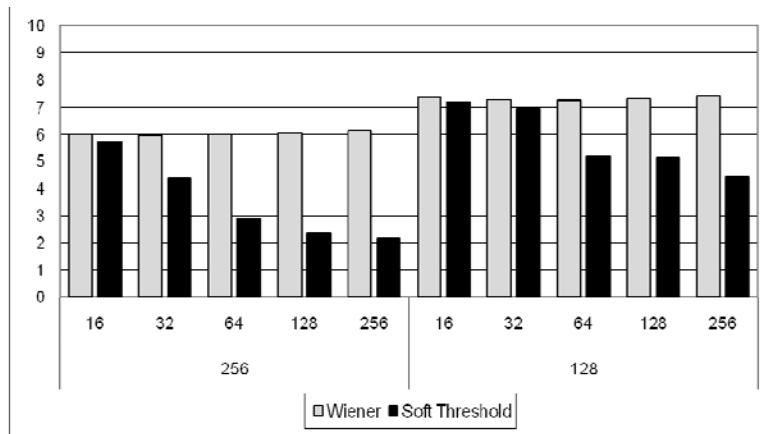
۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 0dB و 20dB نشان می‌دهد. نتایج نظری نشان می‌دهد که مقدار نسبت سیگنال به نویز بهبودیافته برای روش فیلتر کردن وینر بیشتر از روش آستانه‌گذاری نرم می‌باشد. تفاوت سطح نسبت سیگنال به نویز بهبودیافته دو روش برای نرخ سیگنال به نویز ورودی 0dB بیشتر از تفاوت سطح نسبت سیگنال به نویز بهبودیافته برای نرخ سیگنال به نویز ورودی 10dB می‌باشد، چراکه در نویز بیشتر عملکرد فیلتر وینر بهتر بوده و همه نویز را حذف می‌کند ولی در روش آستانه‌گذاری نویز زیادی باقی مانده و از طرف دیگر اعوجاج بیشتری در سیگنال مفید بوجود آمده است.

چنانچه در شکل ۷ دیده می‌شود، با افزایش تعداد زیرباند مقدار نسبت سیگنال به نویز برای روش آستانه‌گذاری کاهش می‌یابد، در واقع آستانه‌گذاری در تعداد زیرباندهای کمتر بهتر عمل می‌کند. ولی با افزایش تعداد زیرباند مقدار نسبت سیگنال به نویز برای روش وینر افزایش می‌یابد، زیرا جزئیات فرکانسی افزایش می‌یابد، بنابراین تخمین دقیقتری از فیلتر داشته و انتظار عملکرد بهتری برای وینر در زیرباندهای بیشتر داریم. از سویی با افزایش طول فریم از ۱۲۸ نمونه به ۲۵۶ نمونه مقدار نسبت سیگنال به نویز بهبودیافته برای روش وینر اندکی کاهش می‌یابد، چرا که با افزایش طول فریم رزلوشن زمانی سیگنال کاهش یافته است در حالیکه تعداد زیرباندها تغییری نکرده است و لذا دقت تخمین پارامترهای فیلتر وینر کم می‌شود.

بلوک دیاگرام سیستم کاهش نویز فیلتر کردن وینر مورد استفاده در این ارزیابی در شکل ۶ نشان داده شده است. همانطور که در شکل مشاهده می‌کنید، ابتدا سیگنال ورودی به فریم‌هایی با طول مشخص تقسیم می‌شود. در ادامه فریم توسط فیلتربانک موجک به N زیرباند تقسیم می‌گردد و روی هر زیرباند یک فیلتر وینر ایده آل اعمال می‌شود. پس از آن با استفاده از عکس تبدیل موجک فریم بهبود یافته بدست می‌آید و در نهایت فریم‌ها در کنار هم قرار می‌گیرند تا سیگنال بهبود یافته نهایی بدست آید.

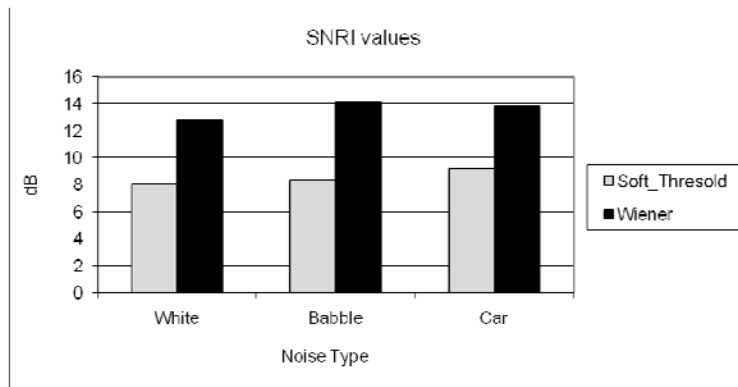
تخمین مؤلفه‌های فیلتر وینر ایده‌آل از روی دو سیگنال نویز و سیگنال گفتار انجام می‌گیرد. مقادیر برای پارامترهای سیستم با توجه به نتایج گزارش شده در مقاله [۳۷] انتخاب شده‌اند. که با توجه به مرجع [۳۸، ۳۷] از موجک بسته‌ای برای تجزیه سیگنال به زیرباندها استفاده شد، که در آن زیرباندها یکنواخت می‌باشند. همچنین طول فریم از ۱۲۸ تا ۱۰۵۴ و تعداد زیرباندها از ۱۶ تا ۲۵۶ در نظر گرفته شده است. تابع پایه موجک بکار رفته نیز $daub10^{11}$ می‌باشد. در پیاده‌سازی آستانه‌گذاری نرم برای آنکه تخمین بهتری از آستانه داشته باشیم، تخمین آستانه را روی فریم بزرگتری انجام می‌دهیم. این عمل به این خاطر انجام می‌گیرد که در واقع با فرض اولیه آستانه‌گذاری یعنی بینهایت بودن فریم تطابق بیشتری داشته باشد.

شکل ۷ مقادیر نسبت سیگنال به نویز بهبودیافته را برای دو روش آستانه‌گذاری نرم و وینر با طول فریم ۱۲۸ و ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴،



نرخ سیگنال به نویز ورودی 20dB

شکل ۷- مقادیر نسبت سیگنال به نویز بهبودیافته دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۱۲۸ و ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 0dB و 20dB



شکل ۸- مقادیر نسبت سیگنال به نویز بهبودیافته برای دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویزهای گوسی سفید، نویز همهمه و نویز ماشین با نرخ سیگنال به نویز ورودی 10dB

شکل ۱۰ نتایج تست شنوائی را برای دو روش آستانه‌گذاری نرم و فیلتر وینر زمانیکه طول فریم ۲۵۶ و تعداد زیرباند ۱۶ و ۲۵۶ نشان می‌دهد. نتایج شنوائی نشان می‌دهد که میزان تخریب گفتار و نیز میزان نویز باقیمانده در روش آستانه‌گذاری بسیار بیشتر از روش فیلتر وینر می‌باشد از طرفی روش فیلتر کردن دارای مقدار کمی نویز موزیکال می‌باشد ولی در مجموع تست شنوائی شنوندگان سیگنال‌های بهبودیافته بروش فیلتر کردن را به سیگنال‌های بهبودیافته بروش آستانه‌گذاری ترجیح داده‌اند. برای طول فریم‌های دیگر و تعداد زیرباندهای دیگر نیز نتایج مشابهی بدست آمده است که برای جلوگیری از طولانی شدن مقاله به ذکر همین نتایج بسنده می‌کنیم.

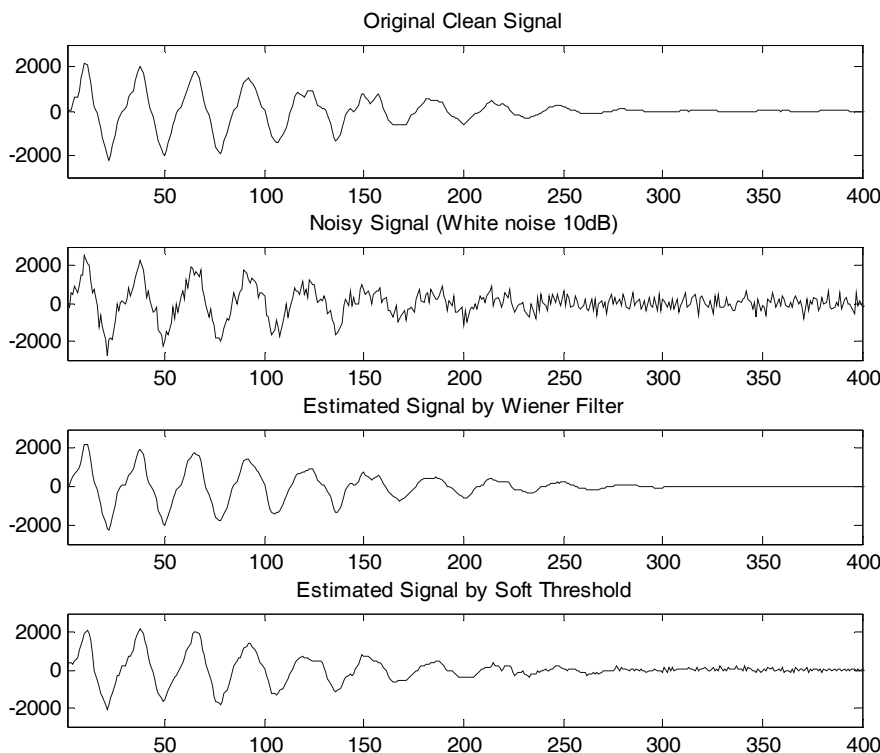
۶- مقایسه و ارزیابی فیلترهای پیشنهادی در بخش ۴

در این بخش به بررسی فیلترهای ارائه شده می‌پردازیم، سیستم حذف نویز مطابق با بلوک دیاگرام شکل ۶ در عمل پیاده‌سازی شده است. در این پیاده‌سازی بجای فیلتر وینر ایده‌آل سه فیلتر وینر غیر ایده‌آل، MAP-G و MAP-SG که روابط آنها بترتیب در (۱۱)، (۲۶) و (۳۷) آمده است، استفاده شده است. در واقع تخمین نویز از روی قسمتهایی از سیگنال نویزی که تنها شامل نویز می‌باشد صورت می‌گیرد. در این ارزیابی از مویک بسته‌ای برای تجزیه سیگنال به زیرباندها استفاده شد. همچنین طول فریم ۲۵۶ نمونه و تعداد زیرباندها از ۱۶ تا ۲۵۶، تابع پایه مویک نیز db10 در نظر گرفته شده است. جدول ۱ مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG با طول فریم ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد.

شکل ۸ مقادیر نسبت سیگنال به نویز بهبودیافته برای دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویزهای گوسی سفید، نویز همپه و نویز ماشین با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد. چنانکه در شکل هم دیده می‌شود مقدار نسبت سیگنال به نویز بهبودیافته در حضور نویزهای مختلف برای روش وینر بیشتر از روش آستانه‌گذاری نرم می‌باشد. و روش فیلتر کردن نتایج بهتری نسبت به روش آستانه‌گذاری در حضور نویزهای مختلف دارد. این تفاوت برای نویز همپه که تا حد زیادی نالیستان است بیشتر مشخص می‌باشد. این امر کاملاً قابل پیش بینی بوده چون در شرایطی مانند همپه که نویز زمینه مشخصاتی شبیه به سیگنال اصلی دارد، توزیع سیگنال و نویز روی زیرباندهای مختلف شبیه به همدیگر خواهد بود و آستانه‌گذاری کارایی خود را تا حدی از دست می‌دهد، حال آنکه فیلتر بهینه با تخمین دقیق تر از نویز، کارا تر است.

در شکل ۹ قسمتی از سیگنال نویزی، سیگنال بهبود یافته بروش فیلتر کردن وینر و سیگنال بهبود یافته بروش آستانه‌گذاری نرم را نشان می‌دهد. سیگنال نویزی شامل نویز و سیگنال گفتار می‌باشد. چنانچه در شکل هم دیده می‌شود، نویز بیشتری در روش فیلتر کردن وینر قسمت اعظم نویز حذف می‌شود ولی در روش آستانه‌گذاری بدلیل غیر بهینه بودن روش آستانه‌گذاری مقداری از نویز باقی می‌ماند. همچنین روش فیلتر کردن در حذف نویز بهتر عمل کرده و سیگنال بهبودیافته به سیگنال تمیز شباهت بیشتری دارد.

تست شنوائی به این صورت انجام شده است که شنوندگان به سیگنال بهبود یافته گوش داده و به ۲ جنبه وجود نویز زمینه^{۱۲} و تخریب صدا^{۱۳} امتیازی بین ۱ تا ۳ داده‌اند. در این تست به سیگنال بهبود یافته ای که کیفیت خوبی دارند امتیاز یک، به سیگنالهای بهبودیافته ای که کیفیت پایین تری دارند ولی از نظر شنونده قابل قبول هستند، امتیاز دو و به سیگنالهایی که کیفیت نازلی دارند و از دید شنونده غیرقابل قبول هستند، امتیاز سه داده می‌شود.



شکل ۹- شکل موج ۴۰۰ نمونه از سیگنال نویزی، سیگنال بهبود یافته بروش فیلتر کردن وینر و سیگنال بهبود یافته بروش آستانه‌گذاری نرم با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

شنوایی نشان داد که کیفیت سیگنال بهبود یافته در تعداد زیرباندهای بیشتر، بهتر از کیفیت سیگنال بهبود یافته در تعداد زیرباندهای کمتر می‌باشد و میزان تخریب گفتار در آنها کمتر می‌باشد.

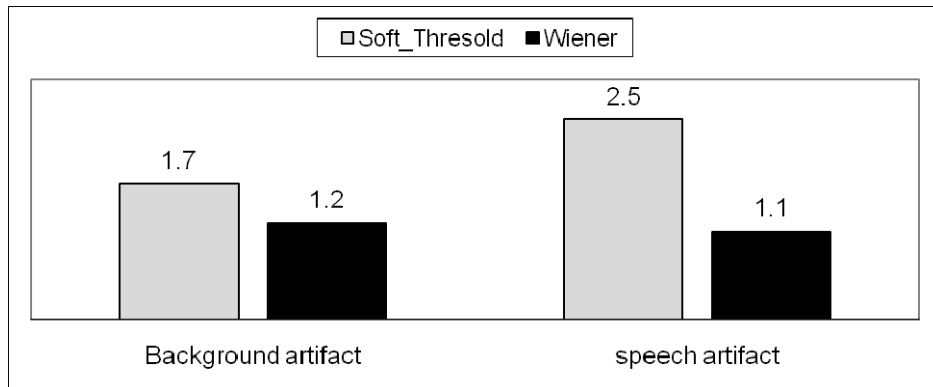
شکل ۱۱ اندازه سه فیلتر وینر، MAP-G و MAP-SG را برای یک فریم با ۶۴ زیرباند نشان می‌دهد. در این نمودار محور افقی فرکانس و محور عمودی اندازه فیلتر در مقیاس دسی بل را نشان می‌دهد. چنانچه که در نمودار دیده می‌شود اندازه فیلتر MAP-G بالاتر از دو فیلتر دیگر و اندازه فیلتر وینر پائینتر از دو فیلتر دیگر می‌باشد. از همینرو در جدول (۱) برای تعداد زیرباند ۶۴، مقدار نسبت سیگنال به نویز بهبودیافته برای روش فیلتر وینر بیشتر و برای فیلتر MAP-G کمتر می‌باشد.

در این مقاله علاوه بر تست شنوایی از مقدار MOS تخمینی با ابزار PESQ^۴ برای ارزیابی نتایج نیز استفاده شده است. قابل ذکر است این روش از طرف ITU به عنوان استاندارد ارزیابی کیفیت گفتار پیشنهاد شده است. این روش برای استفاده بر روی یک پهنه زیاد از شرایط شبکه شامل اتصالات آنالوگ، کدکننده‌ها، بسته‌های گم شده و پارامتر تأخیر تهیه شد. نویز زمینه و نویز پردازش نیز می‌تواند توسط معیار PESQ با استفاده از سیگنال تمیز بدست آید. روش PESQ به فایل بهبود یافته نمره‌ای بین صفر و ۴.۵ می‌دهد. که نمره صفر به سیگنال نویزی و تخریب شده و نمره ۴.۵ به سیگنال تمیز داده می‌شود. در این روش هر چه عدد بزرگتر باشد به معنی کیفیت بالاتر می‌باشد. [۳۲، ۳۳]

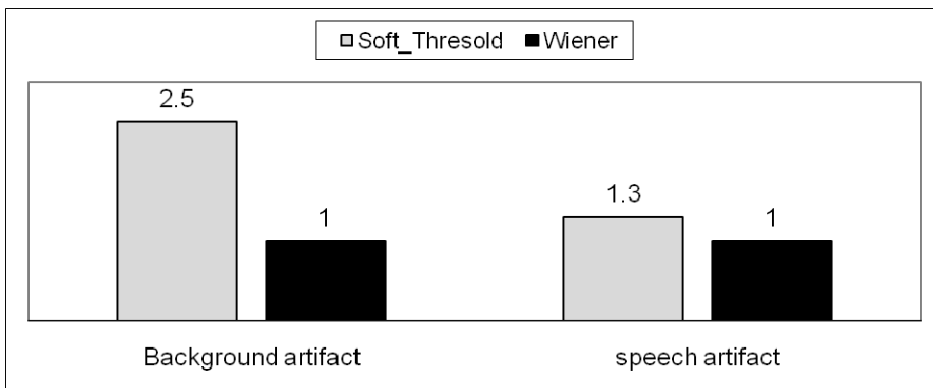
جدول ۱- مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG با طول فریم ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

Number of subbands	Methods		
	MAP-SG	MAP-G	Wiener
16	10.05	9.9	9.77
32	10.28	10.12	10.27
64	10.24	10.09	10.44
128	10.05	9.88	10.46
256	9.58	9.43	10.25

چنانچه در جدول دیده می‌شود مقادیر نسبت سیگنال به نویز بهبودیافته برای سه روش فوق برای یک زیرباند خاص دارای مقادیر نزدیک بهم می‌باشد و این اختلاف در حدود 1dB می‌باشد که می‌توان از آن صرفه نظر کرد. از سویی نیز مقدار نسبت سیگنال به نویز برای زیرباندهای مختلف نیز تقریباً یکسان می‌باشد. در واقع از لحاظ تست نظری تفاوت چندانی بین تعداد زیرباندها وجود ندارد، ولی تست

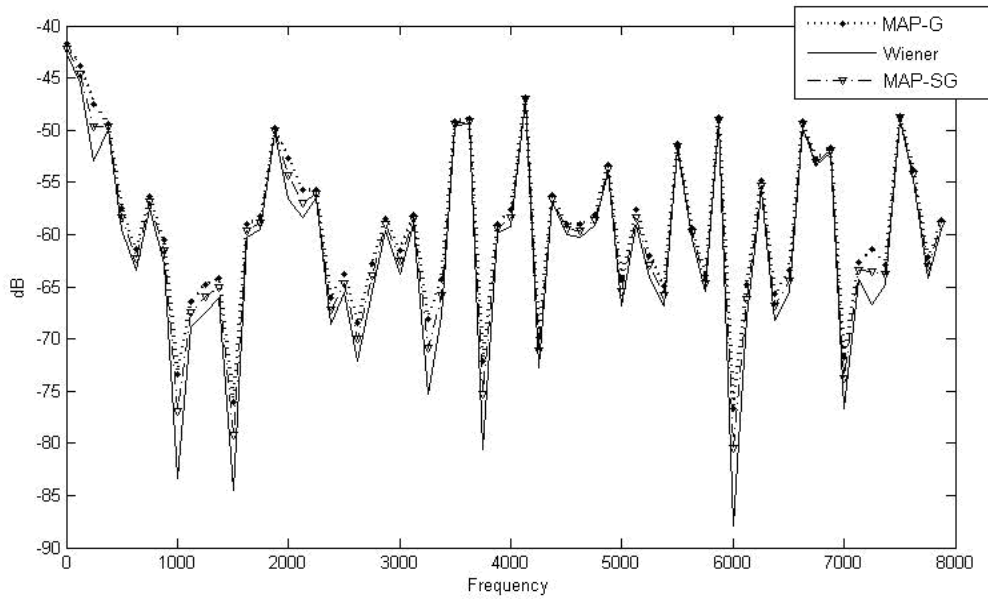


تعداد زیرباند = ۱۶

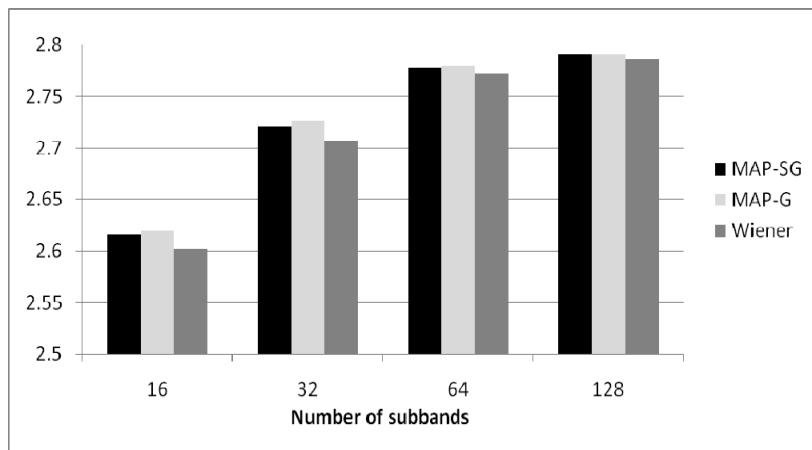


تعداد زیرباند = ۲۵۶

شکل ۱۰- نتایج تست شنوایی برای دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB



شکل ۱۱- اندازه سه فیلتر وینر، MAP-G و MAP-SG را برای یک فریم با طول ۲۵۶ نمونه و تعداد زیرباند ۶۴ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB



شکل ۱۲- مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای فریم ۲۵۶ نمونه‌ای و تعداد زیرباند ۱۶، ۳۲، ۶۴ و ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

به نویز در حوزه فرکانس می‌باشد، که این امر به دلیل خواص موجک می‌باشد. به این معنی که می‌توان در حوزه موجک مقیاس زمان- فرکانس مختلف داشته باشیم که این خاصیت موجک را ابزاری قدرتمند برای تحلیل سیگنال گفتار می‌کند.

جدول ۲- مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG در دو حوزه موجک و فوریه با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

	Wavelet	FFT
MAP-SG	10.05	4.35
MAP-G	9.88	3.4
Wiener	10.46	6.44

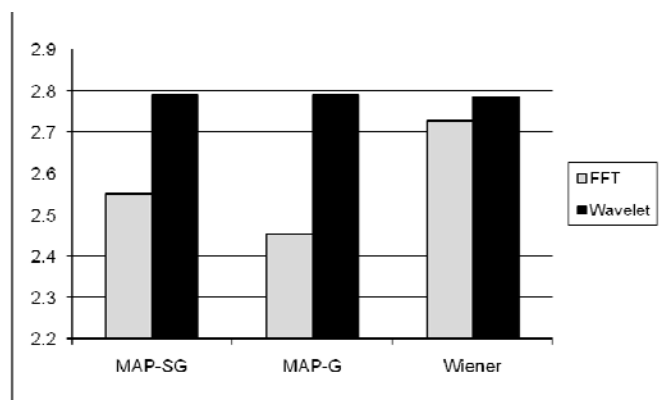
شکل ۱۲ مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای فریم ۲۵۶ نمونه‌ای و تعداد زیرباند ۱۶، ۳۲، ۶۴ و ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد. چنانچه دیده می‌شود روش MAP-G دارای مقدار بیشتری نسبت به دو روش دیگر در طول فریم‌های مختلف دارد. همچنین نتایج بهتری برای تعداد زیرباند بیشتر نسبت به تعداد زیرباند کمتر بدست آمد که این امر بدلیل تحلیل فرکانسی دقیقتر در زیرباندهای بیشتر می‌باشد.

نتایج بدست آمده از آزمایشات با نتایج تئوری که در شکل ۵ نشان داده شده بود مطابقت دارد و دیده می‌شود که فیلتر وینر دارای کمی تخریب گفتار بیشتر نسبت به دو فیلتر پیشنهادی است. در ادامه سه فیلتر فوق را در حوزه فرکانس پیاده سازی شده و نتایج آنها با معادله‌شان در حوزه موجک مقایسه گردید. جدول ۲ نتایج تست نظری یعنی مقدار نسبت سیگنال به نویز بهبودیافته را برای این سه فیلتر در دو حوزه نشان می‌دهد. همانطور که در جدول دیده می‌شود مقدار نسبت سیگنال به نویز برای هر سه فیلتر در حوزه موجک بیشتر از مقدار نسبت سیگنال

شکل ۱۳ طیف نگار قسمتی از سیگنال بهبودیافته را که توسط روش‌های مختلف حذف نویز در دو حوزه موجک و فرکانس بدست آمده است را نشان می‌دهد. مطابق با شکل در تمامی طیفها فرکانس‌های اصلی سیگنال حفظ شده‌اند ولی در روش‌های حذف نویز در حوزه فرکانس، نویز زمینه بیشتری باقی مانده است و همین امر باعث کاهش کیفیت شنیداری آنها می‌شود. چنانچه در طیف سیگنالهای بهبودیافته‌ی بدست آمده از فیلترهای پیشنهادی در حوزه موجک نیز دیده می‌شود نویز بیشتری حذف شده است بدون اینکه مؤلفه‌های اصلی فرکانس مخدوش شوند. این مطلب برای هر سه فیلتر وینر، MAP-G و MAP-SG صادق می‌باشد.

شکل ۱۴ مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای دو حوزه فرکانس و موجک با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB را نشان می‌دهد. چنانچه دیده می‌شود مقادیر MOS تخمینی با PESQ برای هر سه فیلتر در حوزه موجک بیشتر از سه فیلتر فوق در حوزه فرکانس می‌باشد و بهبود حاصل شده توسط روش موجک در مقایسه با FFT می‌تواند به تجزیه مناسب تر سیگنال توسط موجک نسبت داده شود.

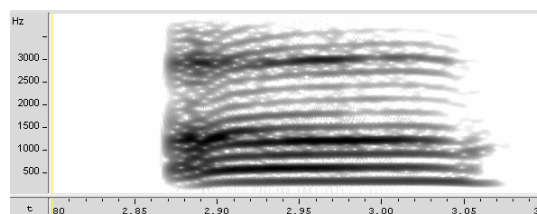
در مجموع می‌توان با استفاده از نتایج ذکر شده ادعا نمود دو فیلتر پیشنهادی و فیلتر وینر دارای کیفیت بهتری هم از نظر تست نظری و هم از نظر تست عملی و شنوائی نسبت به معادلشان در حوزه فرکانس می‌باشند.



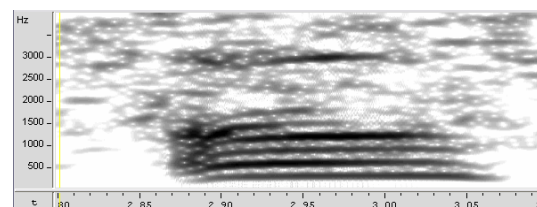
شکل ۱۴- مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای دو حوزه فرکانس و موجک با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

۷- نتیجه گیری

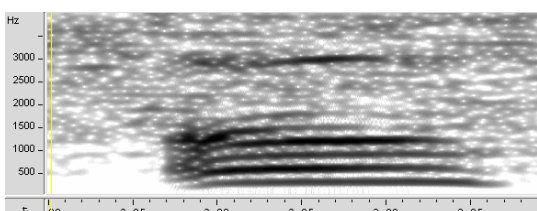
در کار حاضر، ابتدا نشان داده شد که روش فیلتر کردن ضرایب تبدیل موجک بر آستانه‌گذاری این ضرایب، چه از نظر عملی و چه از دیدگاه نظری برتری دارد. دلیل این امر می‌تواند در تخمین هوشمندانه نویز در روش‌های فیلترینگ باشد. در ادامه دو فیلتر جدید در حوزه موجک با الهام از روش‌های بهینه فیلتر کردن در حوزه فرکانس، پیشنهاد شده‌اند. این فیلترها، احتمال پسین ضرایب تبدیل موجک سیگنال بهبود یافته را بیشینه می‌کنند. فیلترهای پیشنهادی در حوزه موجک، از لحاظ ارزیابی نظری عملکردی همپا، و از لحاظ ارزیابی با معیار PESQ عملکردی بهتر نسبت به فیلتر وینر در حوزه موجک دارند. به عبارت دیگر اختلاف مقدار نسبت سیگنال به نویز بهبودیافته برای هر سه نوع فیلتر در حد 1dB می‌باشد که قابل صرف نظر کردن است. از طرفی نیز در ارزیابی PESQ فیلتر MAP-G عملکرد بهتری نسبت به دو فیلتر دیگر دارد.



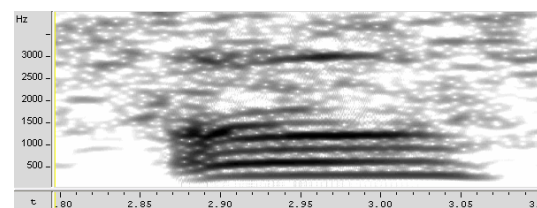
(الف) سیگنال تمیز



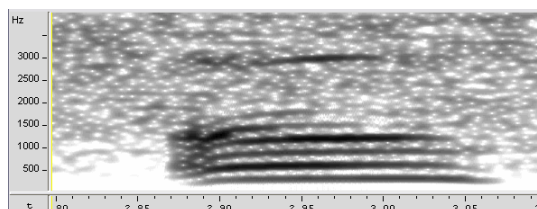
(ب) حذف نویز با فیلتر وینر در حوزه موجک



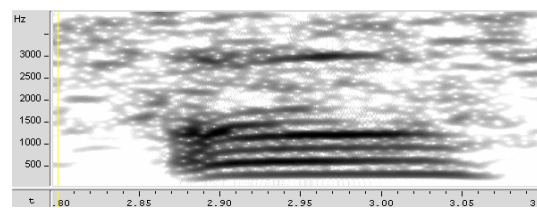
(ج) حذف نویز با فیلتر وینر در حوزه فرکانس



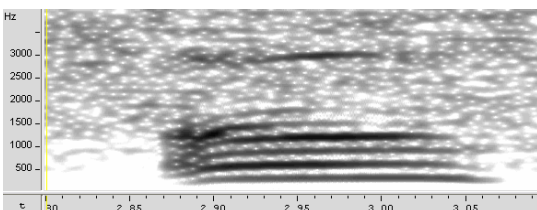
(د) حذف نویز با فیلتر MAP-G در حوزه موجک



(ه) حذف نویز با فیلتر MAP-G در حوزه فرکانس



(و) حذف نویز با فیلتر MAP-SG در حوزه موجک



(ز) حذف نویز با فیلتر MAP-SG در حوزه فرکانس

شکل ۱۳- طیف نگار قسمتی از سیگنال بهبودیافته بدست آمده از روش‌های مختلف حذف نویز در دو حوزه موجک و فرکانس

transform," *Speech Communication*, Vol. 41, pp. 409-427, 2003.

[11] E. Zwicker, and H. Fastl, *Psychoacoustics: Facts and Models*, New York, Springer-Verlag, 1990.

[12] N. Virag, "Single channel speech enhancement based on masking properties of the human auditory system", *IEEE Trans. Speech and Audio Processing*, Vol. 7, No. 2, pp. 126-137, 1999.

[13] Y. Shao, and C. H. Chang, "A versatile speech enhancement system based on perceptual wavelet denoising," *Proc, IEEE Int'l Symp. on Circuits and Systems*, pp. 864-867, 2005.

[14] S. H. Chen, and J. F. Wang, "Speech Enhancement Using Perceptual Wavelet Packet Decomposition and Teager Energy Operator", *Journal of VLSI Signal Processing*, Vol. 36, No. 2-3, pp. 125 - 139, 2004.

[15] G. Ju, and L. Lee, "Speech Enhancement and Improved Recognition Accuracy by Integrating Wavelet Transform and Spectral Subtraction Algorithm," *Proc, European Conf. on Speech Communication and Technology*, pp. 1370-1380, 2003.

[16] N. Ma, M. Bouchard, and R. A. Goubran, "A Wavelet Kalman Filter With Perceptual Masking For Speech Enhancement In Colored Noise," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal processing*, Vol. 1, pp. 149-152, 2005.

[17] I. Y. Soon, S. N. Kah, and C. K. Yeo, "Wavelet for speech denoising," *Proc, IEEE TENCON*, Vol. 2, pp. 479-482, 1997.

[18] Y. Ephraim, and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. 32, No. 6, pp. 1109-1121, 1984.

[19] P. J. Wolfe, and S. J. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," *Proc, 11th IEEE Workshop Statistical Signal Processing*, pp. 496-499, 2001.

[20] A. Grossman, R. Martinet, and J. Morlet, "Reading and understanding continuous wavelet transform," *Proc, Int'l Conf. Wavelets*, pp. 2-20, 1989.

[21] Y. H. Huang, and T. D. Chiueh, "A new audio coding scheme using a forward masking model and perceptually weighted vector quantization," *IEEE Trans. Speech and Audio Processing*, Vol. 10, No. 5, pp. 325-335, 2002.

[22] Y. Shao, and C. H. Chang, "A Generalized Perceptual Time-Frequency Subtraction Method for Speech Enhancement," *Proc, IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2537-2540, 2006.

[23] C. Tantibundhit, J. R. Boston, C. C. Li, J. D. Durrant, S. Shaiman, K. Kovacyk, and A. El-Jaroudi, "Speech

در انتها نیز نتایج بیانگر آنند که فیلترهای پیشنهادی در حوزه موجک، از دیدگاه ارزیابی نظری دارای عملکرد برجسته‌ای نسبت به معادلشان در حوزه فرکانس هستند، این اختلاف در مقدار نسبت سیگنال به نویز بهبودیافته برای دو روش پیشنهادی در حوزه موجک در قیاس با مقدار نسبت سیگنال به نویز بهبودیافته برای فیلترهای معادلشان در حوزه فوریه حدود 6dB می‌باشد و این اختلاف برای فیلتر وینر حدوداً 4dB است. علاوه بر این معیار PESQ نیز مؤید عملکرد مناسب این فیلترها در حوزه موجک است. از نظر پیچیدگی محاسباتی، محاسبه فیلتر در روش MAP-G حدود ۵ برابر و در روش MAP-SG حدود ۷ برابر محاسبات لازم برای فیلتر وینر در هر زیربانند می‌باشد. پیچیدگی محاسباتی اعمال فیلترها یکسان می‌باشد، چراکه طول فیلترها یکسان است. قابل ذکر است در روشهای مورد استفاده محاسبات ناشی از تبدیل موجک بیشتر از محاسبات مربوط به فیلترهاست از اینرو افزایش محاسبات فوق‌الذکر در روشهای پیشنهادی اثر قابل توجهی در زمان اجرای کل فرایند حذف نویز ندارند.

مراجع

[1] A. Huang, A. Acero, and H. Hon, *Spoken Language processing*, Prentice Hall, 2001.

[2] Y. Hu, and P. C. Loizou, "Incorporating a psychoacoustical model in frequency domain speech enhancement," *Proc, IEEE Signal Processing Letters*, Vol. 11, pp. 270-273, 2004.

[3] A. C. Cristan, and A. T. Walden, "Multitaper power spectrum estimation and thresholding: Wavelet packets versus wavelets," *IEEE Trans. Signal Processing*, Vol. 50, pp. 2976 - 2986, 2002.

[4] Y. Hu, and P. C. Loizou, "Speech enhancement based on wavelet thresholding the multitaper spectrum," *IEEE Trans. Speech and Audio Processing*, Vol. 12, No. 1, pp. 59-67, 2004.

[5] K. P. Soman, and K. I. Ramachandran, *Insight into wavelets: From Theory to Practice*, India, Prentice-Hall, 2005.

[6] S. Jung, Y. Kwon, and S. Yang, "Speech Enhancement By Wavelet Packet Transform With Best Fitting Regression Line In Various Noise Environments," *IEEE Int'l Conf. on Acoustics, Speech, and Signal processing*, Vol. 1, pp. 469-472, 2006.

[7] D. L. Donoho, and I. M. Johnston, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* 81, pp. 425-455, 1994.

[8] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, pp. 613-627, 1995.

[9] H. Sheikhzadeh, and H. R. Abutalebi, "An improved wavelet based speech enhancement system," *Proc, European Conf. on Speech Communication and Technology*, pp. 1855-1858, 2001.

[10] C. T. Lu, and H. C. Wang, "Enhancement of single channel speech based on masking property and wavelet

Speech Model," *EURASIP Journal on Applied Signal Processing*, No. 7, pp. 1110-1126, 2005.

[۳۷] ب. زمانی، ب. ناصرشریف، و ا. اکبری، "بررسی اثر باندهای بحرانی و پارامترهای بکارگیری تبدیل موجک در بهبود گفتار"، *مجموعه مقالات یازدهمین کنفرانس بین‌المللی کامپیوتر ایران*، ۱۳۸۴.

[۳۸] ب. زمانی، س. رسول‌زاده، ب. ناصرشریف، و ا. اکبری، "بکارگیری بانک‌فیلتر WOLA در بهبود گفتار بر مبنای تبدیل موجک به منظور کاهش پیچیدگی محاسباتی"، *مجموعه مقالات چهاردهمین کنفرانس بین‌المللی برق ایران*، ۱۳۸۵.



بهزاد زمانی دانشجوی دوره دکترای هوش مصنوعی در دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت می‌باشد. او همچنین درجه کارشناسی‌ارشد هوش مصنوعی را از دانشگاه علم و صنعت در سال ۱۳۸۴ و کارشناسی سخت‌افزار را از دانشگاه صنعتی اصفهان در سال ۱۳۸۱ اخذ نموده است. در حال حاضر وی در آزمایشگاه پردازش صوت و گفتار در زمینه تشخیص گفتار و روش‌های متمایزسازی بر مبنای توابع هسته مشغول به تحقیق است. زمینه تحقیق وی شامل بهبود گفتار، بازشناسی گفتار، تبدیل موجک، تبدیلات متمایزساز ویژگی و توابع هسته می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

bzamani@iust.ac.ir



احمد اکبری درجه کارشناسی را در مهندسی الکترونیک و درجه کارشناسی‌ارشد را در مهندسی مخابرات از دانشگاه صنعتی اصفهان در سالهای ۱۳۶۶ و ۱۳۶۹ دریافت کرد. وی درجه دکترای را در پردازش سیگنال و مخابرات راه دور از دانشگاه رن فرانسه در سال ۱۳۷۴ اخذ نمود و در همان سال به عنوان عضو هیات علمی به دانشگاه علم و صنعت پیوست و در حال حاضر با عنوان دانشیار دانشکده مهندسی کامپیوتر مشغول به فعالیت است. وی سرپرست مرکز تحقیقات فناوری اطلاعات و آزمایشگاه پردازش صوت و گفتار در این دانشگاه است. در ضمن ایشان سمت سردبیری مجله علمی پژوهشی موسسه پردازش هوشمند علام را بر عهده دارند و از سال ۱۹۹۹ عضو انجمن کامپیوتر ایران می‌باشند. زمینه تحقیق وی شامل بهبود گفتار، تشخیص گفتار، مقاوم‌سازی بازشناسی گفتار، انتقال صوت از طریق شبکه و پروتکل‌های مربوطه می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

akbari@iust.ac.ir



بابک ناصرشریف متولد ۱۳۵۳ در تهران است. وی درجه کارشناسی را در رشته مهندسی کامپیوتر گرایش سخت‌افزار از دانشگاه صنعتی امیرکبیر در سال ۱۳۷۶ دریافت نمود. وی موفق به اخذ درجه کارشناسی‌ارشد و دکترای در رشته مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه علم و صنعت ایران در سالهای ۱۳۷۹ و ۱۳۸۶ گردید. وی از سال ۱۳۸۶ تا کنون عضو هیئت علمی گروه مهندسی کامپیوتر در دانشکده فنی دانشگاه گیلان است. زمینه تحقیق وی شامل بهبود گفتار، بازشناسی گفتار و مقاوم‌سازی آن، تبدیل موجک، مدل‌سازی و بازشناسی الگو و واژه‌یابی گفتاری می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

nasser_s@iust.ac.ir

enhancement using transient speech components," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing*, pp. 833-836, 2006.

[24] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelet and Wavelet Transforms*, Prentice Hall Inc., 1998.

[25] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal smoothing and Minimum Statistic," *IEEE Trans. on Speech and Audio Processing*, Vol. 9, No. 5, 2001.

[26] J. Weaver, D. M. Healy, and L. D. Cromwell, "Filtering noise from images with wavelet transforms," *Magnetic Resonance in Medicine*, Vol. 21, No. 2, pp. 288-95, 1991.

[27] J. W. Seok, and K. S. Bae, "Speech enhancement with reduction of noise components in the wavelet domain," *Proc, IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing*, pp. 1323-1326, 1997.

[28] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive Wavelet Thresholding for Image Denoising and Compression," *IEEE Trans. Image Processing*, Vol. 9, No. 9, pp. 1532-1546, 2000.

[29] I. J. Kim, S. I. Yang, and Y. Kwon, "Speech enhancement using adaptive wavelet shrinkage," *Proc, IEEE Int'l Symp. on Industrial Electronics*, Vol. 1, pp. 501-504, 2001.

[30] I. M. Johnston, and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *Journal of the Royal Statistical Society, Series B*, Vol. 59, pp. 319-351, 1997.

[31] R. Moussaoui, J. Rouat, and R. Lefebvre, "Wavelet based independent component analysis for multi-channel source separation," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing*, pp. 645-648, 2006.

[32] "Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end speech quality assessment of narrow band telephone network and speech codec," ITU-T Recommend, pp. 862, Feb 2001.

[33] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," *Proc, IEEE ICASSP*, Vol. 2, pp. 749-752, 2001.

[34] S. Ayat, M. Manzuri, and R. Dianat, "An Improved Wavelet-Based Speech Enhancement by Using Speech Signal features," *International Journal of Computers & Electrical Engineering*, Vol. 32, No. 6, pp. 411-425, 2006.

[35] R. McAulay, and M. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 28, no. 2, pp. 137-145, 1980.

[36] T. Lotter, and P. Vary, "Speech Enhancement by MAP Spectral Amplitude Estimation Using a Super-Gaussian

اطلاعات بررسی مقاله:

تاریخ ارسال:

تاریخ اصلاح:

تاریخ قبول شدن:

نویسنده مرتبط:

-
- ¹ Hands Free
 - ² Bluetooth
 - ³ Wiener Filter
 - ⁴ Threshold
 - ⁵ Step Garrote
 - ⁶ Minimizing the False Discovery Rate
 - ⁷ Top
 - ⁸ Minimizing Stein's Unbiased Risk Estimate
 - ⁹ Wavelet Packet
 - ¹⁰ Maximum A Posteriori Spectral Amplitude Estimator
 - ¹¹ Daubechy Order 10
 - ¹² Background Artifact
 - ¹³ Speech Artifact
 - ¹⁴ Perceptual Evaluation of Speech Quality