

معرفی فیلترهایی مناسب برای بهسازی گفتار در حوزه موجک

هدی کاویانی شهرضایی شیما طیبیان احمد اکبری

دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران

چکیده

تاکنون روش‌های متفاوتی برای کاهش نویز در سیستم‌های پردازش گفتار، استفاده شده‌اند که برخی از آنها در حوزه تبدیل موجک صورت گرفته‌اند. روش‌های بهسازی گفتار در این حوزه، از طریق آستانه‌بندی و یا با فیلتر کردن ضرایب نویزی، محقق می‌شود. در این مقاله، ابتدا مجموعه فیلترهای مناسبی که در روش فیلتر کردن در حوزه تبدیل موجک قابل استفاده هستند، بر اساس روابط ریاضی معرفی شده‌اند. سپس با استفاده از فیلترهای این مجموعه، دو فیلتر جدید تعریف شده است که فیلتر اول از ترکیب فیلترهای مینیمم و ماکزیمم حاصل شده است و فیلتر دوم که بر پایه فیلتر مینیمم بنا شده است، توانی متغیر با مقادیر کوچکتر از یک را به این فیلتر نسبت می‌دهد. اساس تعریف این فیلترها بر این اصل استوار است که فیلتر پیشنهادی در فریم‌های سکوت بیشترین نویز را از سیگنال حذف کرده و در فریم‌های گفتار همزمان با کاهش نویز، کمترین اعوجاج را بر روی سیگنال گفتار ایجاد کند.

کلمات کلیدی: بهسازی گفتار، تبدیل موجک، فیلتر کردن، نویز باقیمانده، انرژی سیگنال.

۱- مقدمه

آستانه‌بندی [۶-۱۰] و یا با فیلتر کردن ضرایب نویزی [۱۱-۱۳]، محقق می‌شود. در روش‌های بهبود گفتار بر پایه فیلتر کردن در حوزه موجک، سیگنال ورودی پس از پنجره‌بندی، به تعدادی زیرباند تجزیه می‌گردد سپس با اعمال فیلترهای بهبود گفتار بر روی هر یک از زیرباندها، از تأثیر نویز کاسته می‌شود. در مرحله بعد، با ترکیب زیرباندهای بهبود، سیگنال بهسازی شده نهایی حاصل می‌شود. در این مقاله ابتدا با کمک روابط ریاضی، رابطه میان انرژی سیگنال‌های ورودی و خروجی با انرژی توزیع شده در زیرباندهای درخت موجک، بررسی شده است. سپس با استفاده از همین خاصیت، مجموعه‌ای از فیلترهای قابل استفاده در این حوزه معرفی شده‌اند که از میان این فیلترها، کوچکترین آنها بیشترین نویز را از سیگنال حذف می‌کند ولی بزرگترین فیلتر کمترین اعوجاج را ایجاد می‌کند. با توجه به این اثرات فیلتر پیشنهادی اول، با استفاده از ترکیب این دو فیلتر ایجاد شده است به طوری که در فریم‌های گفتار فیلتر ماکزیمم و در فریم‌های سکوت، فیلتر مینیمم بیشترین تأثیر را داشته باشد. فیلتر پیشنهادی دوم با توجه به خاصیت حذف حداکثری نویز توسط فیلتر مینیمم تعریف شده است، به طوری که توانی متغیر به این فیلتر نسبت داده شده است که در فریم‌های گفتار مقدار فیلتر را اندکی کاهش می‌دهد تا از بروز اعوجاج در سیگنال جلوگیری شود و در

با رشد سریع سیستم‌های مخابرات صوت، بهسازی گفتار نویزی تبدیل به یک موضوع مهم و فعال تحقیقاتی شده است. هدف اصلی در حذف نویز این است که سیگنال بازسازی شده تا حد امکان به سیگنال اصلی نزدیک بوده و در عین حال خواص اصلی سیگنال همچون میزان همواری آن نیز حفظ شود. روش‌های بهسازی گفتار در دو حوزه فرکانس و زمان تدوین شده‌اند. از تکنیک‌های حوزه زمان می‌توان به مدل مخفی مارکف، فیلتر وینر و فیلتر کالمن اشاره کرد [۱-۳]. روش‌های مبتنی بر تبدیل موجک و تبدیل فوریه نیز مثال‌هایی از تکنیک‌های بهسازی گفتار در حوزه فرکانس هستند [۴-۱۳]. هدف از استفاده تبدیلات، انتقال سیگنال نویزی از حوزه زمان به حوزه دیگری است که در حوزه جدید امکان استخراج آسان‌تر و با کیفیت بهتر سیگنال اصلی از سیگنال نویزی وجود دارد. در این میان تبدیل موجک با داشتن مزایایی نظیر دارا بودن مقیاس‌های زمان-فرکانس مختلف و پیچیدگی محاسباتی پایین‌تر نسبت به تبدیل فوریه، این هدف را بهتر برآورده می‌سازد. بهسازی گفتار بر مبنای تبدیل موجک، از طریق

در بعضی از موارد، شنوندگان سیگنال نویزی اولیه را به سیگنال بهسازی شده بوسیله سیستم مبتنی بر تبدیل موجکی که در آن از آستانه‌گذاری استفاده شده است، ترجیح داده‌اند.

۲-۲- روش‌های بر مبنای فیلتر کردن

در روش‌های بر مبنای فیلتر کردن، پس از پنجره‌بندی و محاسبه ضرایب تبدیل موجک، فیلتر مورد نظر بر روی ضرایب زیرباندها اعمال می‌شود. رابطه (۱) چگونگی اعمال فیلتر را بر روی زیر باندها نشان می‌دهد:

$$\bar{S}(m, f, j) = H(m, f) \times X(m, f, j) \quad (1)$$

که در این رابطه $X(m, f, j)$ مؤلفه j ام از زیرباندها f از فریم m از سیگنال گفتار نویزی را نشان می‌دهد و $\bar{S}(m, f, j)$ مؤلفه j ام از زیرباندها f از فریم m از سیگنال بهبودیافته را نشان می‌دهد و $H(m, f)$ ، فیلتر بهبود گفتار مربوط به زیرباندها f از فریم m می‌باشد.

استفاده از روش فیلتر کردن برای بهسازی گفتار از سال‌های دورتر همواره مورد توجه بوده است ولی استفاده از آن بیشتر در حوزه تبدیل فوریه بوده است. از جمله شناخته‌شده‌ترین روش‌های این زمینه می‌توان به روش تفاضل طیف توان و استفاده از فیلترهای طیفی نظیر فیلتر وینر اشاره نمود [۴]. از جمله اثرات نامطلوب این شیوه بهبود گفتار میتوان به ایجاد اعوجاج در سیگنال و ایجاد نویز موسیقی اشاره نمود. علاوه بر این، بنا به ماهیت تبدیل فوریه و نایستانت بودن سیگنال گفتار، عملیات باید بر روی طیف دوره کوتاه سیگنال صورت گیرد که محدودیت‌هایی در تنظیم رزولوشن زمانی و فرکانسی ایجاد می‌نماید. از این رو با استفاده از فیلترها در حوزه تبدیل موجک به جای تبدیل فوریه می‌توان این اثرات نامطلوب را تا حد زیادی کاهش داد.

۳- تبدیل موجک

تبدیل موجک یکی از تبدیل‌هایی است که در سال‌های اخیر در کاربردهای مختلفی از پردازش سیگنال استفاده شده است. در این تبدیل، بر خلاف تبدیل فوریه، توابع پایه، توابع سینوسی و کسینوسی نیستند، بلکه هر مجموعه از توابع با دارا بودن شرایط خاصی می‌توانند به عنوان توابع پایه مورد استفاده قرار گیرد.

در هر تبدیل موجک، یک تابع تغییر مقیاس پایه $\phi(t)$ وجود دارد که مجموعه توابع تغییر مقیاس از انتقال این تابع پایه بدست می‌آیند:

$$\phi_k(t) = \phi(t - k); k \in Z \quad (2)$$

اگر فضای پوشانده شده توسط $\phi_k(t)$ را V_0 بنامیم، آنگاه هر تابع $f(t)$ در این فضا را می‌توانیم بصورت ترکیب خطی از توابع $\phi_k(t)$ بنویسیم:

$$f(t) = \sum_k a_k \phi_k(t); f(t) \in V_0 \quad (3)$$

این رابطه بسط $f(t)$ را در مقیاس صفر نشان می‌دهد. می‌توان آن را در مقیاس‌های دیگر نیز بسط داد. برای این کار از مجموعه‌ای از توابع که از تغییر مقیاس و انتقال تابع اصلی $\phi(t)$ بدست می‌آید، استفاده می‌شود:

فریم‌های سکوت مقدار فیلتر تقریباً بدون تغییر باقی می‌ماند تا بیشترین نویز از سیگنال حذف شود.

ساختار این مقاله به صورت ذیل است؛ در بخش بعد مرور مختصری بر روی بهسازی گفتار مبتنی بر تبدیل موجک، انجام گرفته است. در بخش سوم، به تئوری تبدیل موجک پرداخته و در ادامه در بخش چهارم، رابطه بین انرژی سیگنال‌های ورودی و خروجی با انرژی توزیع شده در زیرباندهای درخت موجک، بررسی شده است. سپس، در بخش پنجم مجموعه‌ای از فیلترهای قابل استفاده در بهسازی گفتار بر مبنای تبدیل موجک، ارائه شده‌اند. در بخش ششم، فیلترهای پیشنهادی در حوزه موجک معرفی شده است. در بخش هفتم مراحل انجام عملیات بهسازی گفتار با استفاده از فیلترهای پیشنهادی بیان شده است و در بخش هشتم نتایج پیاده‌سازی روش‌های پیشنهادی مورد ارزیابی قرار گرفته است. در انتها، در بخش پایانی، نتیجه‌گیری مطالب ارائه شده است.

۲- بهسازی گفتار مبتنی بر تبدیل موجک

روش‌های بهسازی گفتار با کمک تبدیل موجک را می‌توان به دو دسته عمده طبقه بندی کرد:

- ۱) روش‌های بر مبنای آستانه‌گذاری
- ۲) روش‌های بر مبنای فیلتر کردن

۲-۱- روش‌های بر مبنای آستانه‌گذاری

اساس کار آستانه گذاری موجک بر این امر استوار است که، برای اکثر سیگنال‌های موجود در زندگی حقیقی تعداد محدودی از ضرایب موجک در باندهای پایین، برای بازسازی تخمین خوبی از سیگنال اصلی کافی هستند. معمولاً این ضرایب در مقایسه با دیگر ضرایب مقدار بالاتری دارند و در نتیجه بوسیله حذف کردن ضرایبی که کمتر از یک مقدار حد آستانه هستند، می‌توانیم بدون از دست دادن اطلاعات مهم سیگنال، نویز را حذف کنیم.

الگوریتم این روش نویز زدایی به این ترتیب است که ابتدا سیگنال نویزی را پنجره‌بندی می‌کنیم. در ادامه، تبدیل موجک پنجره مورد نظر را محاسبه می‌کنیم. در مرحله بعد، بر مبنای یک الگوریتم آستانه‌گذار و یک مقدار آستانه، تعدادی از ضرایب موجک را حذف می‌کنیم. سپس، تبدیل معکوس موجک را بر ضرایب بدست آمده اعمال می‌کنیم. در آخر، با جمع و همپوشانی پنجره‌های بدست آمده سیگنال تمیز را محاسبه می‌کنیم.

روش آستانه گذاری موجک، ضرایب جزئیات را حذف می‌کند زیرا این ضرایب مؤلفه‌های فرکانسی بالای سیگنال را آشکار می‌کنند، در حالیکه، فرض بر این است که اطلاعات مهم سیگنال در فرکانس‌های پایین قرار گرفته‌اند و این فرض برای گروه بزرگی از سیگنال‌ها درست است. پس این بدان معنی است که در فرکانس‌های بالا، نویز تأثیر بیشتری از سیگنال اصلی دارد و به عبارت دیگر در فرکانس‌های بالا، نویز قسمت بیشتری از ضرایب را، نسبت به فرکانس‌های پایین اشغال می‌کند.

روش آستانه‌گذاری موجک، معایبی را هم برای استفاده در سیگنال‌های گفتار داراست. از آن جمله این است که قسمت‌هایی از گفتار مانند نواحی صامت که اطلاعات مهمی از سیگنال را در فرکانس‌های بالا حمل می‌کنند، در قسمت جزئیات ضرایب قرار می‌گیرند. از دیگر معایب این روش آن است که، فقط در تست‌های عینی موفق عمل میکند و این برای کاربردی نظیر گفتار که هر دو جنبه عینی و شنیداری در آن مهم هستند، کافی به نظر نمی‌رسد. زیرا به عنوان مثال،

تبدیل موجک گسسته را می‌توان هم به صورت دوتایی و هم به صورت موجک بسته‌ای در نظر گرفت که در درخت موجک بسته‌ای هم ضرایب جزئیات و هم ضرایب تخمین در سطح درخت گسترش می‌یابند [۱۳]. در کار حاضر از یک درخت بسته موجک در سه سطح و با ۸ زیرباند استفاده شده است.

۴- رابطه بین انرژی زیرباندهای تبدیل موجک با انرژی سیگنال‌های ورودی و خروجی

می‌توان ثابت کرد که انرژی سیگنال ورودی با میانگین انرژی زیرباندها در حوزه تبدیل موجک برابر است. فرض کنیم $h_a = [a_0, \dots, a_L]$ ضرایب فیلتر پایین‌گذر و $h_d = [d_0, \dots, d_L]$ ضرایب فیلتر بالاگذر تبدیل موجک باشند. از آنجا که پایه‌های موجکی که معمولاً برای بهسازی گفتار استفاده می‌شوند دارای فیلترهایی با تعداد ضرایب زوج هستند، می‌توانیم فرض کنیم که L عددی زوج باشد. همچنین فرض کنیم $x = [x_1, \dots, x_N]$ سیگنال ورودی باشد. در اینصورت داریم:

$$x * h_a(j) = \sum_{k=0}^j a_k x_{j-k} \quad (12)$$

حال انرژی سیگنال در زیر باند اول به صورت زیر بدست می‌آید:

$$|x * h_a|^2 = \sum_{j=0}^{N+L-1} \left(\sum_{k=0}^j a_k x_{j-k} \right)^2 \quad (13)$$

با کمک اتحاد مربع چندجمله‌ای می‌توان نوشت:

$$|x * h_a|^2 = \sum_{j=0}^{N+L-1} \sum_{k=0}^j a_k^2 x_{j-k}^2 + 2 \sum_{j=0}^{N+L-1} \sum_{k=0}^j \sum_{n=1}^{L-k} a_k a_{k+n} x_{j-k} x_{j-k-n} \quad (14)$$

می‌توان ثابت کرد که مقدار نهایی رابطه (۱۳) به صورت رابطه (۱۵) می‌باشد:

$$|x * h_a|^2 = \left(\sum_{k=0}^L a_k^2 \right) \left(\sum_{j=0}^N x_j^2 \right) + 2 \sum_{i=1}^{L/2} \left\{ \sum_{k=0}^{L-2i} a_k a_{k+2i} \right\} \left(\sum_{j=0}^{N-2i} x_j x_{j+2i} \right) + 2 \sum_{i=1}^{L/2} \left\{ \sum_{k=0}^{L-2i+1} a_k a_{k+2i-1} \right\} \left(\sum_{j=0}^{N-2i+1} x_j x_{j+2i-1} \right) \quad (15)$$

با استفاده از رابطه (۱۰) داریم:

$$|x * h_a|^2 = \left(\sum_{j=0}^N x_j^2 \right) + 2 \sum_{i=1}^{L/2} \left\{ \sum_{k=0}^{L-2i+1} a_k a_{k+2i-1} \right\} \left(\sum_{j=0}^{N-2i+1} x_j x_{j+2i-1} \right) \quad (16)$$

با توجه به رابطه (۱۱)، انرژی سیگنال در سطح دوم به صورت زیر بدست

می‌آید:

$$|x * h_d|^2 = \left(\sum_{j=0}^N x_j^2 \right) - 2 \sum_{i=1}^{L/2} \left\{ \sum_{k=0}^{L-2i+1} a_k a_{k+2i-1} \right\} \left(\sum_{j=0}^{N-2i+1} x_j x_{j+2i-1} \right) \quad (17)$$

$$\phi_{k,j}(t) = 2^{-j/2} \phi(2^{-j}t - k); j, k \in Z \quad (4)$$

اگر فضای پوشانده شده توسط $\phi_{k,j}(t)$ را V_j بنامیم، آنگاه خواهیم داشت:

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots \quad (5)$$

ملاحظه می‌شود که بسط $f(t)$ در مقیاس‌های بالاتر (زهای بزرگتر)، در واقع یک تقریب درشت‌نمایی از آن را به ما می‌دهد. یعنی در مقیاس‌های بالاتر جزئیات را از دست خواهیم داد. برای رفع این مشکل مجموعه توابع دیگری تعریف می‌شود که تفاوت بین فضاهای V_j را بپوشاند. این مجموعه توابع، موجک‌ها می‌باشند که آنها را با $\psi_{j,k}(t)$ نشان می‌دهیم. بنابراین اگر فضای پوشانده شده با توابع $\psi_{j,k}(t)$ را W_j بنامیم، آنگاه خواهیم داشت:

$$V_j = V_{j+1} \oplus W_j \quad (6)$$

توابع $\psi_{j,k}(t)$ از تغییر مقیاس و انتقال تابع موجک مادر $\psi(t)$ به صورت زیر بدست می‌آیند:

$$\psi_{s,\tau}(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x-\tau}{s}\right) \quad (7)$$

که در آن s و τ به ترتیب پارامترهای مقیاس و تبدیل می‌باشند. در تکنیک‌های بهسازی گفتار مبتنی بر تبدیل موجک معمولاً از تبدیلات موجک گسسته استفاده می‌شود. در این تبدیلات از فیلترهای بالاگذر و پایین‌گذر مرتبط با توابع موجک مادر و مقیاس استفاده می‌شود که روابط آن صورت زیر است:

$$a_{j+1}(k) = \sum_m h_a(m-2k) a_j(m) \quad (8)$$

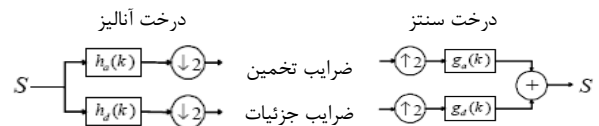
$$d_{j+1}(k) = \sum_m h_d(m-2k) d_j(m) \quad (9)$$

که h_a و h_d فیلترهای بالاگذر و پایین‌گذر و a_i و d_i به ترتیب ضرایب تخمین و جزئیات تبدیل موجک می‌باشند. روابط زیر میان ضرایب تخمین و جزئیات برقرار است.

$$\sum a_k a_{k+2i} = \delta(i) \quad (10)$$

$$a_k = (-1)^k d_{L-k} \quad (11)$$

شکل (۱) ساختاری از یک تبدیل موجک ساده با دو زیرباند را نمایش می‌دهد:



شکل ۱- ساختار درخت آنالیز و سنتز موجک

حال با میانگین گرفتن از انرژی این دو سطح داریم:

می‌توان ثابت کرد انرژی سیگنال خروجی به گونه‌ای که شرایط فوق را برآورده نماید، باید به صورت زیر باشد (جزئیات اثبات در ضمیمه آورده شده است):

$$|\hat{s}|^2 = (|x| - |n|)(|x| + a|n|); \quad -1 \leq a \leq 1 \quad (26)$$

$$\frac{|x * h_a|^2 + |x * h_d|^2}{2} = \left(\sum_{j=0}^N x_j^2 \right) \quad (18)$$

حال مجموعه فیلترهای مناسب برای بهسازی گفتار در حوزه تبدیل موجک را به صورت رابطه (۲۷) تعریف می‌کنیم:

بنابر رابطه (۱۸)، انرژی سیگنال ورودی و خروجی با میانگین انرژی در زیرباندهای مختلف برابر است که این خاصیت به سطح‌های بالاتر درخت تجزیه موجک هم قابل تعمیم است. حال قصد داریم به کمک این خاصیت فیلترهای مناسب برای استفاده در حوزه موجک را به دست بیاوریم.

$$h_i = \frac{\sqrt{(|x_i| - |n_i|)(|x_i| + a|n_i|)}}{x_i}; \quad -1 \leq a \leq 1 \quad (27)$$

۵- مجموعه فیلترهای مناسب برای بهسازی گفتار در حوزه تبدیل موجک

که h_i فیلتر زیرباند i ام و n تعداد زیرباندها است. همچنین منظور از x_i و n_i به ترتیب ضرایب سیگنال نویزی و سیگنال نویز در زیرباند i ام از درخت موجک می‌باشد. با انتخاب فیلتر به صورت رابطه (۲۷) و با توجه به رابطه:

فرض کنیم x سیگنال نویزی ورودی باشد. در این صورت داریم:

$$|\hat{s}|^2 = \frac{1}{n} \sum_n |x_i h_i|^2 \quad (28)$$

$$x = s + n \quad (19)$$

تمام فیلترهای تعریف شده توسط رابطه (۲۷)، برای بهسازی گفتار در حوزه موجک قابل استفاده هستند.

که s معرف سیگنال تمیز و n سیگنال نویز می‌باشد. هدف یافتن سیگنال خروجی \hat{s} می‌باشد به طوری که این سیگنال به حد کافی به s نزدیک باشد. این کار معمولاً با استفاده از مینیمم کردن معیار مجموع مربعات خطا صورت می‌گیرد که رابطه آن به صورت زیر است:

$$E = (s - \hat{s})^2 \quad (20)$$

۶- ارائه فیلترهایی مناسب در حوزه تبدیل موجک

۶-۱- فیلتر پیشنهادی اول

در این بخش قصد داریم فیلتری را معرفی کنیم که از ترکیب دو فیلتر از مجموعه فیلترهای رابطه (۲۷) بدست آمده است. برای بدست آوردن یک فیلتر ترکیبی مناسب، ابتدا باید دو فیلتر با کارایی‌های مطلوب و مختلف انتخاب کنیم و سپس ضریبی مناسب این دو فیلتر را با هم ترکیب کنیم.

ولی از آنجا که بعد از عملیات فیلتر کردن در حوزه موجک، بدست آوردن تخمینی از سیگنال خروجی در حوزه زمان کار بسیار دشواری است و از سوی دیگر با میانگین‌گیری از انرژی زیرباندهای درخت موجک می‌توانیم تقریبی از انرژی سیگنال خروجی بدست آوریم، از معیار زیر استفاده می‌کنیم:

$$E = |s|^2 - |\hat{s}|^2 \quad (21)$$

۶-۱-۱- انتخاب فیلترها برای ترکیب شدن

از میان فیلترهای بیان شده در رابطه (۲۷)، به ازای $a = -1$ فیلتری با مقدار مینیمم و به ازای $a = 1$ فیلتر ماکزیمم بدست خواهد آمد. از آنجا که هر چه مقدار فیلتر کمتر باشد، ضرایب سیگنال را بیشتر کاهش می‌دهد، قادر است مقدار نویز بیشتری را از گفتار حذف کند. این در حالی است که با اعمال فیلتری با مقدار بیشتر، به سبب تغییر کمتری که بر روی ضرایب ایجاد می‌شود، اعوجاج کمتری را شاهد خواهیم بود. اگر این فیلترها را h_{\min} و h_{\max} بنامیم، داریم:

واضح است که با مینیمم کردن این رابطه، رابطه (۲۰) نیز مینیمم خواهد شد. بنابراین فیلتر پیشنهادی باید به گونه‌ای عمل کند که انرژی سیگنال خروجی هر چه بیشتر به $|s|^2$ نزدیک باشد. می‌دانیم:

$$|s|^2 = |x|^2 + |n|^2 - 2xn \quad (22)$$

بنابراین برای مینیمم شدن رابطه (۲۱)، مقدار $|\hat{s}|^2$ باید تابعی درجه دوم از $|x|$ و $|n|$ باشد. همچنین با توجه به خاصیت نویز جمع شونده n شرایط زیر باید برقرار باشند:

$$h_{\min_i} = \frac{|x_i| - |n_i|}{x_i} \quad (29)$$

$$1) |x| = |n| \Rightarrow |\hat{s}|^2 = 0 \quad (23)$$

$$h_{\max_i} = \frac{\sqrt{|x_i|^2 - |n_i|^2}}{x_i} \quad (30)$$

$$2) |n| = 0 \Rightarrow |\hat{s}|^2 = |x|^2 \quad (24)$$

که h_{\min_i} و h_{\max_i} مقادیر فیلترها در زیرباند i ام از درخت موجک می‌باشند. با توجه به خواصی که از این دو فیلتر ذکر شد، باید آنها را طوری ترکیب

$$3) 0 \leq |\hat{s}|^2 \quad (25)$$

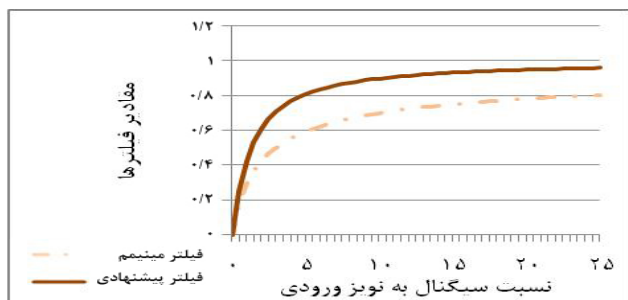
با توجه به این خواص، ضریبی که به این صورت تعریف می‌شود برای استفاده در رابطه (۳۳) مناسب می‌باشد.

۶-۲- فیلتر پیشنهاد دوم

با توجه به آنچه در بخش‌های قبل ذکر شد، از میان فیلترهای بیان شده در رابطه (۲۷)، به ازای $a = -1$ فیلتری با مقدار مینیمم بدست خواهد آمد. از آنجا که هر چه مقدار فیلتر کمتر باشد، ضرایب سیگنال را بیشتر کاهش می‌دهد، قادر است مقدار نویز بیشتری را از گفتار حذف کند. ولی عیب استفاده از این فیلتر، ایجاد اندکی اعوجاج در فریم‌های گفتار می‌باشد. برای حل این مشکل باید مقدار فیلتر را در این فریم‌ها اندکی افزایش دهیم. برای این منظور از فیلتری تواندار به صورت رابطه (۳۵) استفاده می‌کنیم که مشابه این فیلتر قبلاً در حوزه تبدیل فوریه مطرح شده و نتایج خوبی را به دنبال داشته است [۵].

$$h = h_{\min}^{1-h_{\min}} \quad (35)$$

شکل (۲) نمودار مقادیر دو فیلتر h و h_{\min} را نشان می‌دهد:



شکل ۲- نمودار مقادیر فیلتر مینیمم و فیلتر پیشنهادی دوم بر حسب میزان نسبت سیگنال به نویز ورودی

مقایسه این دو فیلتر نشان می‌دهد که در نسبت‌های سیگنال به نویز پایین تفاوت محسوسی بین دو فیلتر مشاهده نمی‌شود و هر چه مقدار نسبت سیگنال به نویز ورودی بیشتر باشد تفاوت بین این دو فیلتر بیشتر خواهد شد. بنابراین این فیلتر جدید در فریم‌های سکوت که نسبت سیگنال به نویز پایین‌تر است مقدار نویز را نسبت به فیلتر قبلی افزایش نمی‌دهد، در حالیکه در فریم‌های گفتار با افزایش مقدار فیلتر از ایجاد اعوجاج در سیگنال جلوگیری می‌کند.

۷- بهسازی گفتار با استفاده از فیلترهای پیشنهادی

مراحل انجام عملیات بهسازی گفتار بوسیله فیلتر پیشنهادی اول در شکل (۳) نشان داده شده است:

برای انجام عملیات بهسازی گفتار توسط هر کدام از فیلترهای پیشنهادی، ابتدا سیگنال ورودی را پنجره‌بندی کرده و سپس ضرایب تبدیل موجک آن را محاسبه می‌کنیم. در مرحله بعد مقدار انرژی نویز، به‌روز رسانی شده تا از آن برای بدست آوردن مقدار فیلتر استفاده شود. به‌روز رسانی مقدار انرژی نویز از طریق روش تخمین نویز پیوسته صورت می‌گیرد. روش کار به این صورت است که ابتدا یک تخمین اولیه از نویز در نظر می‌گیریم و سپس در هر مرحله با اعمال فیلتر بدست آمده از فریم قبلی، تخمین جدیدی از نویز بدست آورده و با ضریبی ثابت، آن را با تخمین انرژی نویز اولیه ترکیب می‌کنیم [۷]. سپس با داشتن انرژی نویز

کنیم که در فریم‌های سکوت سیگنال h_{\min} بیشترین تأثیر را داشته باشد، چون در این بخش هدف تنها حذف نویز از سیگنال است. از سوی دیگر در این فریم‌ها دوبار استفاده از این فیلتر، نویز بیشتری حذف خواهد شد و در نتیجه بهتری را شاهد خواهیم بود. می‌توان ثابت کرد که این کار معادل با استفاده از دوم این فیلتر است. زیرا از آنجا که با یکبار استفاده از این فیلتر در قسمت‌های سکوت، میزان انرژی سیگنال و نویز به یک نسبت کاهش می‌یابد، اگر بخواهیم بار دیگر از این فیلتر و با مقادیر جدید استفاده کنیم داریم:

$$\text{new_h_min}_i = \frac{|k||x_i| - |k||n_i|}{|k||x_i|} \quad (31)$$

حال با فاکتورگیری از $|k|$ و ساده کردن آن از صورت و مخرج داریم:

$$\text{new_h_min}_i = \frac{|x_i| - |n_i|}{|x_i|} = h_{\min}_i \quad (32)$$

بنابراین مقدار فیلتر با پارامترهای جدید تغییر نمی‌کند و می‌توانیم دوباره همان مقدار قبلی را استفاده کنیم. بنابراین دو بار استفاده از فیلتر معادل با استفاده از h_{\min}^2 خواهد بود. پس فیلتر پیشنهادی را به صورت ترکیب خطی h_{\max} و h_{\min}^2 تعریف می‌کنیم.

$$h_i = \alpha_i h_{\max}_i + (1 - \alpha_i) h_{\min}_i^2 \quad (33)$$

در این رابطه منظور از α_i ، ضریب فیلتر در زیرباند نام از درخت موجک می‌باشد.

۶-۱-۲- یافتن ضریب مناسب

با توجه به خواص دو فیلتر h_{\max} و h_{\min} ، ضریب α را باید به گونه‌ای انتخاب کنیم که در فریم‌های سکوت سیگنال، این مقدار به صفر میل کند با این هدف که فیلتر h_{\min}^2 بیشترین تأثیر را داشته باشد. از سوی دیگر این انتخاب باید به گونه‌ای باشد که در فریم‌های گفتار ضریب α به یک نزدیک شود تا تأثیر h_{\max} را افزایش دهد. با این توضیحات α را به صورت رابطه زیر تعریف می‌کنیم:

$$\alpha_i = \frac{(|x_i|^2 - |n_i|^2)}{|x_i|^2} \quad (34)$$

انتخاب α بصورت رابطه (۳۴)، دارای خواص زیر است:

- (۱) از آنجا که مقدار تفاضل انرژی سیگنال و نویز در هر فریم یکی از عوامل مشخص‌کننده سکوت یا گفتار بودن فریم است، انتخاب α به صورت تابعی از این مقدار، تا حدی نشان‌دهنده احتمال سکوت یا گفتار بودن فریم موردنظر خواهد بود.
- (۲) مقدار α بین ۰ و ۱ تغییر می‌کند.
- (۳) با انتخاب توان دوم، α در فریم‌های سکوت تا حد زیادی به صفر نزدیک می‌شود و تأثیر h_{\max} در این فریم‌ها به حداقل می‌رسد. در حالیکه در فریم‌های گفتار که کاهش نویز با کمترین اعوجاج هدف ماست، هر دو فیلتر اثرگذار خواهند بود.

در سیگنال بهبودیافته، دخیل هستند. این معیار، عددی بین ۰ و ۴/۵ به سیگنال ورودی تخصیص می‌دهد. هر چه این عدد به ۴/۵ نزدیکتر باشد، کیفیت سیگنال بهبودیافته به سیگنال گفتار تمیز نزدیکتر خواهد بود [۱۴].

جداول (۱)، (۲)، (۳)، (۴) و (۵)، میزان بهبود نسبت سیگنال به نویز ورودی در حضور نویز سفید، صورتی، ماشین، کارخانه و همه‌ها، با نرخ‌های سیگنال به نویز متفاوت را به کمک فیلتر مینیمم، ماکزیمم، وینر و فیلترهای پیشنهادی، نشان می‌دهد.

جدول ۱- نتایج ارزیابی میزان بهبود نسبت سیگنال به نویز ورودی برای نویز سفید

فیلتر پیشنهادی دوم	فیلتر پیشنهادی اول	فیلتر وینر	فیلتر ماکزیمم	فیلتر مینیمم	SNR ورودی
۱۰/۵۰۹ db	۱۰/۱۷ db	۹/۹۴۳ db	۷/۶۹۱ db	۹/۳۴۳ db	۰ db
۹/۳۸۳ db	۹/۳۶۷ db	۸/۹۹۱ db	۷/۱۱۲ db	۷/۲۳۲ db	۵ db
۸/۳۸۹ db	۸/۱۷۲ db	۸/۰۷۱ db	۶/۵۵ db	۶/۳۹۶ db	۱۰ db
۷/۵۶۸ db	۷/۲۴۲ db	۷/۳۳۵ db	۶/۰۳ db	۵/۰۷۱ db	۱۵ db
۶/۹۰۷ db	۶/۶۰۲ db	۶/۹۰۷ db	۵/۵۹۳ db	۴/۰۶۱ db	۲۰ db
۶/۳۹۷ db	۶/۱۰۳ db	۶/۲۱ db	۵/۲۴۲ db	۳/۲۶۹ db	۲۵ db

جدول ۲- نتایج ارزیابی میزان بهبود نسبت سیگنال به نویز برای نویز صورتی

فیلتر پیشنهادی دوم	فیلتر پیشنهادی اول	فیلتر وینر	فیلتر ماکزیمم	فیلتر مینیمم	SNR ورودی
۸/۴۰۲ db	۸/۵۶۵ db	۸/۴۰۳ db	۶/۵۷۸ db	۷/۸۱۷ db	۰ db
۷/۸۶۶ db	۷/۸۶۹ db	۷/۸۶۶ db	۶/۱۰۹ db	۶/۵۱۶ db	۵ db
۷/۴۲۲ db	۷/۳۰۵ db	۷/۴۲۲ db	۵/۷۴۱ db	۵/۵۵ db	۱۰ db
۶/۸۹۷ db	۵/۸۴۷ db	۶/۸۹۷ db	۵/۴۳ db	۴/۸۱۹ db	۱۵ db
۶/۳۳۱ db	۶/۴۵۷ db	۶/۳۳۱ db	۵/۱۷۹ db	۴/۳۴۱ db	۲۰ db
۵/۸۷ db	۶/۱۰۳ db	۵/۸۷ db	۴/۹۲۶ db	۳/۷۵۸ db	۲۵ db

جدول ۳- نتایج ارزیابی میزان بهبود نسبت سیگنال به نویز ورودی برای نویز ماشین

فیلتر پیشنهادی دوم	فیلتر پیشنهادی اول	فیلتر وینر	فیلتر ماکزیمم	فیلتر مینیمم	SNR ورودی
۷/۶۴۲ db	۷/۸۸۶ db	۷/۵۴۸ db	۶/۶۸۲ db	۶/۳۰۷ db	۰ db
۶/۸۵۲ db	۶/۲۹۵ db	۶/۷۸۷ db	۶/۱۱۴ db	۴/۷۰۶ db	۵ db
۶/۳۳۴ db	۶/۷۸۵ db	۶/۲۷۷ db	۵/۷۱۱ db	۳/۴۹۹ db	۱۰ db
۵/۹۹۷ db	۶/۳۰۶ db	۵/۹۴۴ db	۵/۴۳۷ db	۲/۴۹۴ db	۱۵ db
۵/۷۵۷ db	۵/۸۵۹ db	۵/۶۹۹ db	۵/۲۱ db	۲/۷۰۵ db	۲۰ db
۵/۵۲۲ db	۵/۴۱۶ db	۵/۴۴۷ db	۴/۹۵۴ db	۲/۰۸ db	۲۵ db

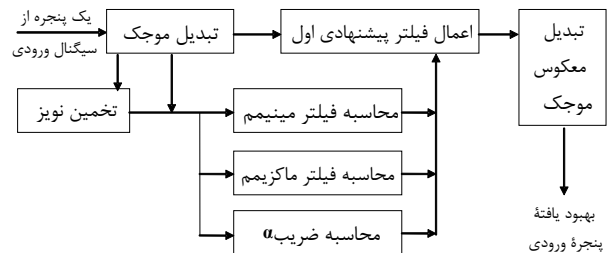
جدول ۴- نتایج ارزیابی میزان بهبود نسبت سیگنال به نویز ورودی برای نویز کارخانه

فیلتر پیشنهادی دوم	فیلتر پیشنهادی اول	فیلتر وینر	فیلتر ماکزیمم	فیلتر مینیمم	SNR ورودی
۶/۹ db	۷/۰۵۸ db	۶/۰۱۹ db	۴/۳۳۳ db	۷/۰۵ db	۰ db
۶/۴۷۳ db	۶/۲۷۵ db	۵/۶۱ db	۴/۱ db	۶/۰۴۹ db	۵ db
۶/۱۲۲ db	۵/۸۰۱ db	۵/۲۸۸ db	۳/۹۱۴ db	۵/۴۶۸ db	۱۰ db
۵/۸۳۳ db	۵/۴۱۱ db	۵/۰۲۵ db	۳/۷۵۴ db	۴/۸۵۸ db	۱۵ db
۵/۶۱۲ db	۵/۰۸۷ db	۴/۷۹۲ db	۲/۶۰۹ db	۴/۳۶۱ db	۲۰ db
۵/۳۸۳ db	۴/۷۹۳ db	۴/۵۶۳ db	۴/۵۶۳ db	۳/۹۲۲ db	۲۵ db

جدول ۵- نتایج ارزیابی میزان بهبود نسبت سیگنال به نویز ورودی برای نویز همه‌ها

فیلتر پیشنهادی دوم	فیلتر پیشنهادی اول	فیلتر وینر	فیلتر ماکزیمم	فیلتر مینیمم	SNR ورودی
۵/۸۶۱ db	۷/۱۷۷ db	۴/۹۸۷ db	۳/۲۰۵ db	۶/۸۶۶ db	۰ db
۵/۶۵۱ db	۷/۳۹۸ db	۴/۷۷۲ db	۳/۰۸۱ db	۵/۹۸۵ db	۵ db
۵/۵۱۱ db	۷/۱۱۷ db	۴/۶۰۷ db	۲/۹۸۴ db	۵/۲۵۴ db	۱۰ db
۵/۳۸۸ db	۶/۷۹۵ db	۴/۴۶ db	۲/۸۹۸ db	۴/۶۱۹ db	۱۵ db
۵/۲۶۲ db	۶/۴۲۱ db	۴/۳۲۶ db	۲/۸۱۹ db	۴/۰۶۶ db	۲۰ db
۵/۱۱۲ db	۶/۰۴۵ db	۴/۱۸۶ db	۲/۷۲۴ db	۳/۶۰۲ db	۲۵ db

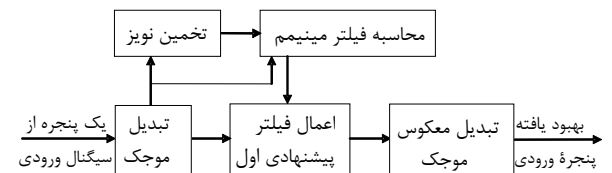
مقدار فیلترهای پایه و در مرحله بعد، مقدار فیلتر نهایی را در هر زیرباند محاسبه کرده و آن را بر ضرایب موجک اعمال می‌کنیم. سپس با کمک تبدیل معکوس موجک، فریم‌های بهبودیافته حاصل می‌شوند. با همپوشانی و جمع این فریم‌ها سیگنال بهسازی شده نهایی حاصل می‌شود.



شکل ۳- سیستم کاهش نویز با استفاده از فیلتر پیشنهادی اول

شکل (۴) مراحل انجام این عملیات را برای فیلتر پیشنهادی دوم نشان

می‌دهد:



شکل ۴- سیستم کاهش نویز با استفاده از فیلتر پیشنهادی دوم

۸- نتایج پیاده‌سازی

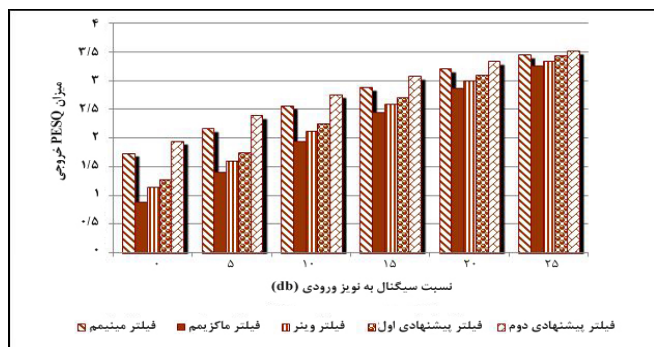
در این قسمت به مقایسه کارایی فیلترهای پیشنهادی با فیلترهای مینیمم، ماکزیمم و فیلتر وینر، پرداخته شده است.

تعداد سطوح درخت آنالیز، برای هر دو فیلتر بانک مورد آزمایش، سه سطح می‌باشد. مقایسات بر روی شش سیگنال گفتار تمیز مختلف (مربوط به سه زن و سه مرد) از پایگاه داده TIMIT انجام شده‌اند. برای ساخت سیگنال نویزی از پنج نوع نویز سفید، صورتی، همه‌ها، کارخانه و ماشین با نسبت‌های سیگنال به نویز ۰، ۵، ۱۰، ۱۵، ۲۰ و ۲۵ دسیبل، استفاده شده است و برای بدست آوردن نتایج بین ۶ جمله متوسط گیری شده است. برای طول قاب از اندازه ۲۵۶ نمونه استفاده شده است، همپوشانی میان قاب‌ها ۵۰٪ می‌باشد.

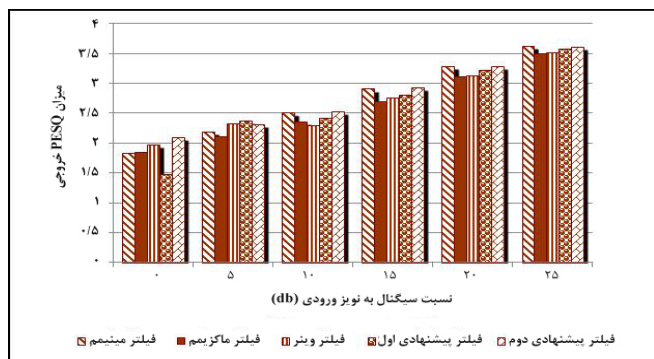
به منظور انجام مقایسات از معیارهای عینی نسبت سیگنال به نویز (SNR) و معیار ارزیابی شنیداری کیفیت گفتار (PESQ) استفاده شده است. برای محاسبه معیار عینی نسبت سیگنال به نویز از رابطه زیر استفاده شده است:

$$SNR = \frac{\sum_{i=1}^{SignalLength} S(i)^2}{\sum_{i=1}^{SignalLength} [S(i) - \hat{S}(i)]^2} \quad (36)$$

در رابطه بالا S و \hat{S} به ترتیب سیگنال تمیز و سیگنال بهبود یافته می‌باشند. معیار PESQ، یکی از معیارهای ارزیابی شنیداری کیفیت گفتار است که در محاسبه آن پارامترهای مختلفی از قبیل میزان اعوجاج، اکو، خفگی و نویز باقیمانده



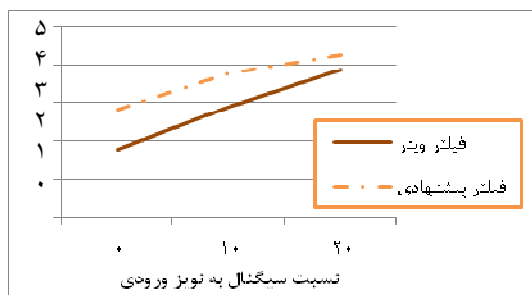
شکل ۸- مقایسه فیلترها از لحاظ میزان PESQ خروجی در حضور نویز جمع‌شونده کارخانه



شکل ۹- مقایسه فیلترها از لحاظ میزان PESQ خروجی در حضور نویز جمع‌شونده همه‌مه

گرچه معیار PESQ در برخی موارد برتری فیلتر مینیمم را نشان می‌دهد، ولی با توجه به اینکه این فیلتر با ارزیابی صورت گرفته توسط معیار نسبت سیگنال به نویز برتر قابل توجهی را نشان نمی‌دهد، نمی‌توان آن را برتر دانست. با این توضیح می‌توان گفت که همانند نتایج معیار نسبت سیگنال به نویز، فیلتر پیشنهادی اول بهترین نتیجه را داشته و به‌طور متوسط حدود ۰/۵ واحد برتری را نشان می‌دهد. بعد از این فیلتر، فیلتر پیشنهادی دوم به خصوص در نسبت‌های سیگنال به نویز بالاتر بهترین نتیجه را داشته و به‌طور متوسط ۰/۲ واحد برتری را نشان می‌دهد.

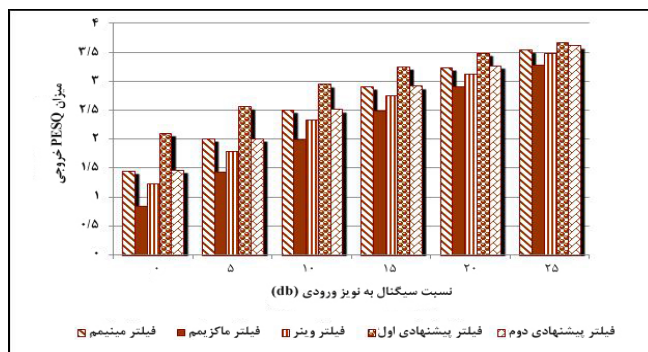
با توجه به برتری فیلتر پیشنهادی اول، تست شنوایی را بر روی فیلتر پیشنهادی اول و در مقایسه با فیلتر وینر انجام داده‌ایم. آزمایش‌ها بر روی دو سیگنال (۱ مرد و ۱ زن) از پایگاه داده TIMIT انجام شده که با سه نویز سفید، ماشین و همه‌مه تخریب شده‌اند. شکل‌های (۱۰)، (۱۱) و (۱۲) نتایج مربوط به این آزمایش را نشان می‌دهد. بطوریکه هر چه مقدار نویز باقیمانده کمتر بوده، عدد بالاتری به آن نسبت داده شده است.



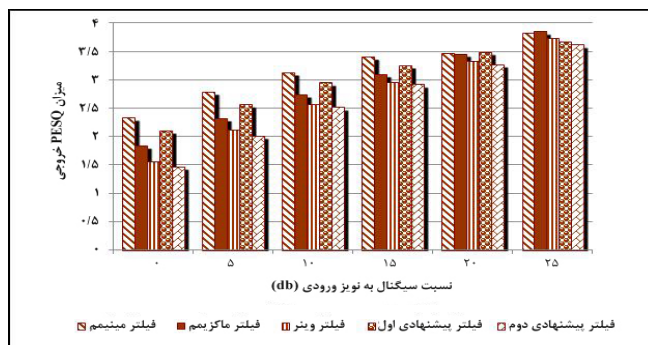
شکل ۱۰- مقایسه دو فیلتر وینر و فیلتر پیشنهادی اول، از لحاظ مقدار نویز باقیمانده کمتر در حضور نویز جمع‌شونده سفید

نتایج حاکی از آن است که فیلتر پیشنهادی اول، تقریباً تمام موارد بر دیگر فیلترها برتری داشته است بطوریکه، مثلاً در مورد نویز همه‌مه این تفاوت بین ۱ تا ۴ دسیبل بوده است. به‌طور متوسط می‌توان گفت که فیلتر پیشنهادی دوم حدود ۲ دسیبل از دیگر فیلترها برتر بوده است. همچنین، می‌توان دید که این برتری در نسبت‌های سیگنال به نویز پایین، محسوس‌تر بوده و کارایی فیلتر بالاتر است. در مورد فیلتر پیشنهادی دوم هم می‌توان گفت که نسبت به دیگر فیلترها، غیر از فیلتر پیشنهادی دوم، برتر بوده است و به‌طور میانگین حدود ۱ دسیبل برتری داشته است. این فیلتر بر خلاف فیلتر پیشنهادی اول، در نسبت‌های سیگنال به نویز بالاتر بهتر عمل می‌کند.

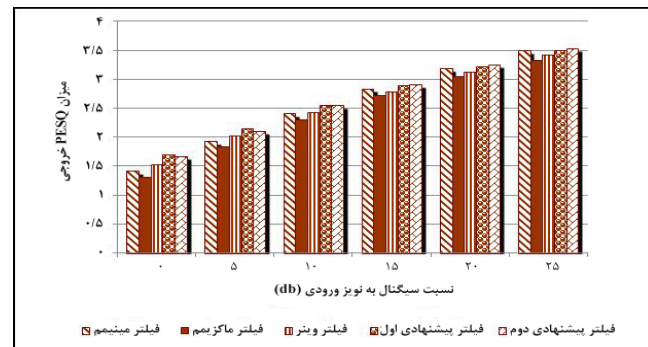
در ادامه به بررسی فیلترهای پیشنهادی با معیار PESQ خواهیم پرداخت که نتایج آن در شکل‌های (۵)، (۶)، (۷)، (۸) و (۹) نمایش داده شده است.



شکل ۵- مقایسه فیلترها از لحاظ میزان PESQ خروجی در حضور نویز جمع‌شونده سفید



شکل ۶- مقایسه فیلترها از لحاظ میزان PESQ خروجی در حضور نویز جمع‌شونده صوتی



شکل ۷- مقایسه فیلترها از لحاظ میزان PESQ خروجی در حضور نویز جمع‌شونده ماشین

کارایی فیلترهای پیشنهادی را بیش از این بهبود بخشید که این کار در حیطه فعالیت‌های پژوهشی آتی قرار خواهد گرفت.

مراجع

[1] H. Sameti, and L. Deng, "Nonstationary-State Hidden Markov Model Representation of Speech Signals for Speech Enhancement," *Signal Processing*, vol. 82, no. 2, pp. 205-227, 2002.

[2] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New Insights Into the Noise Reduction Wiener Filter," *IEEE Transaction on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1218-1234, 2006.

[3] T. K. Kumar, "Optimum Cascaded Design for Speech Enhancement Using Kalman Filter," *Proc. World Academy of Science, Engineering and Technology*, 2008.

[4] T. Lotter, P. Vary, "Noise Reduction By Maximum A Posteriori Spectral Amplitude Estimation With Supergaussian Speech Modeling," *Proc. of the International Workshop on Acoustic Echo and Noise Control*, pp. 83-86, 2003.

[5] B. Ayad, "A Noise Suppressor," <http://www.freepatentsonline.com/EP1242992>.

[6] S. Tabibian, and A. Akbari, "Noise Reduction from Speech Signal based on Wavelet Transform and Kullback-Leibler Divergence," *Proc. of the International Symposium on Telecommunications*, pp. 787-791, 2008.

[7] M. Johnson, X. Yuan, and Y. Ren, "Speech Signal Enhancement through Adaptive Wavelet Thresholding," *Speech Communication*, vol. 49, no. 2, pp. 123-133, 2007.

[8] Q. Fu, and E. Wan, "Perceptual Wavelet Adaptive Denoising of Speech," *Proc. of the Euro Conference on Speech Communication and Technology*, pp. 577-580, 2003.

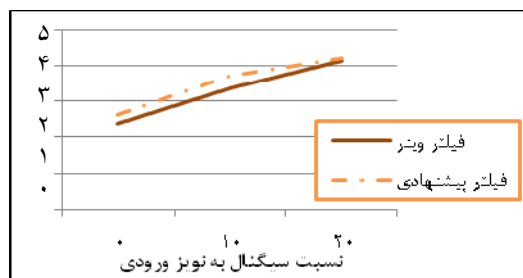
[9] C. A. Medina, and R. Alcaim, "Wavelet Denoising Of Speech Using Neural Networks For Threshold Selection," *Electronics Letters*, vol. 39, no. 25, pp. 1869-1871, 2003.

[10] M. Bahoura, and J. Rouat, "A New Approach for Wavelet Speech Enhancement," *Proc. of the Euro Conference on Speech Communication and Technology*, pp. 1937-1940, 2001.

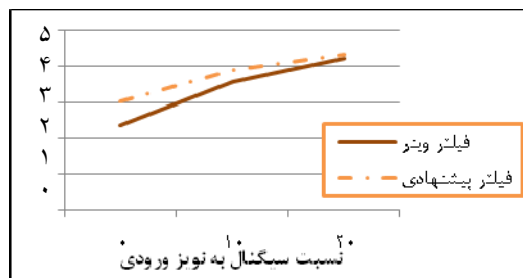
[11] Y. Shao, and C. Chang, "A Generalized Time-Frequency Subtraction Method For Robust Speech Enhancement Based On Wavelet Filter Banks Modeling Of Human Auditory System," *IEEE Transactions On Systems, Man, And Cybernetics*, vol. 37, no. 4, pp. 877-889, 2007.

[12] M. Rahmani, M. Mohammadi, and A. Akbari, "Background Noise Control for Speech Enhancement," *Proc. of the Conference on Computer Society of Iran*, 2006.

[13] F. Ykhlef, A. Guessoum, and D. Berkani, "Combined Spectral Subtraction and Wiener Filter Methods in Wavelet



شکل ۱۱- مقایسه دو فیلتر وینر و فیلتر پیشنهادی اول، از لحاظ مقدار نویز باقیمانده کمتر در حضور نویز جمع‌شونده ماشین



شکل ۱۲- مقایسه دو فیلتر وینر و فیلتر پیشنهادی اول، از لحاظ مقدار نویز باقیمانده کمتر در حضور نویز جمع‌شونده مهمه

مشاهده می‌شود که در هر سه نویز فیلتر پیشنهادی نسبت به فیلتر وینر به‌خصوص در مورد نسبت‌های سیگنال به نویز پایین، نتیجه بهتری را داشته است، بطوریکه در نسبت سیگنال به نویز ۰، حداقل ۰/۲ و حداکثر ۱/۱ واحد افزایش را شاهد هستیم.

مقایسه این دو فیلتر از لحاظ قابلیت فهم، تفاوت قابل ملاحظه‌ای را نشان نمی‌دهد و اکثر شنوندگان به هر دو فیلتر، مقادیر مساوی نسبت داده‌اند.

همچنین، در تست‌های شنوایی انجام شده، ۸۲٪ از شنوندگان سیگنال بهسازی شده با کمک فیلتر پیشنهادی را ترجیح داده‌اند و فقط در ۴٪ موارد، فیلتر وینر برتری داده شده است.

با توجه به این نتایج می‌توان گفت که فیلتر پیشنهادی اول، بهترین نتیجه را نسبت به دیگر فیلترها، به‌خصوص در نسبت‌های سیگنال به نویز پایین‌تر داراست.

۹- نتیجه‌گیری

در این مقاله مجموعه فیلترهای قابل استفاده در حوزه تبدیل موجک، با کمک روابط ریاضی، معرفی شدند و با کمک فیلترهای این مجموعه، دو فیلتر جدید در این حوزه تعریف شد که فیلتر اول از ترکیب دو فیلتر مینیمم و ماکزیمم و دیگری از تواندار کردن فیلتر مینیمم حاصل شده است. ارزیابی‌ها، برتری فیلتر پیشنهادی اول را به خصوص در مورد نسبت‌های سیگنال به نویز پایین تأیید می‌کند. بطوریکه افزایش حدود ۲ دسیبل، در میزان بهبود نسبت سیگنال به نویز، مشاهده می‌شود. میزان بهبود معیار PESQ نیز با بکارگیری فیلتر پیشنهادی اول، در حدود ۰/۵ واحد افزایش داشته است.

در نسبت‌های سیگنال بالاتر، کارایی خوب فیلتر پیشنهادی دوم محسوس است، به‌طوریکه افزایش حدود ۲ دسیبلی در نسبت سیگنال به نویز و حدود ۰/۲ واحد افزایش در مورد میزان PESQ خروجی، مشاهده می‌شود.

به نظر می‌رسد با توجه به نتایج خوب این فیلترها، بتوان با انتخاب α بهینه‌تر، در مورد فیلتر پیشنهادی اول و انتخاب مناسب‌تر توان فیلتر، در مورد فیلتر پیشنهاد دوم،



هدی کاویانی دارای مدرک کارشناسی ریاضی محض و فارغ‌التحصیل مقطع کارشناسی ارشد رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک از دانشگاه علم و صنعت ایران می‌باشد که تحقیقات او در زمینه بهسازی گفتار است.

آدرس پست‌الکترونیکی ایشان عبارت است از:

hoda.kaviani@gmail.com



شیمای طبیبیان دانشجوی مقطع دکترای رشته هوش مصنوعی و رباتیک در دانشگاه علم و صنعت ایران می‌باشد. تحقیقات او در زمینه پردازش گفتار است که بر روی موضوعات بهسازی گفتار و واژه‌یابی گفتار تمرکز بیشتری یافته است.

آدرس پست‌الکترونیکی ایشان عبارت است از:

shimatabibian@iust.ac.ir



احمد اکبری دانشیار دانشگاه علم و صنعت در رشته مهندسی کامپیوتر می‌باشد. او مدرک دکترای خود را در سال ۱۹۹۵ از دانشگاه رن فرانسه اخذ کرده است. زمینه کاری او بر روی تحلیل سیگنال گفتار، بازشناسی گفتار گسسته، بهسازی گفتار، سیستم‌های گفتگو، واسط ارتباط انسان و کامپیوتر مبتنی بر گفتار، کاربرد صدا روی وب، شبکه‌های کامپیوتری و امنیت شبکه متمرکز شده است.

آدرس پست‌الکترونیکی ایشان عبارت است از:

akbari@iust.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۸/۳/۲۵

تاریخ اصلاح: ۸۹/۵/۲۶

تاریخ قبول شدن: ۸۹/۷/۶

نویسنده مرتبط: دکتر احمد اکبری، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

Domain for Noise Reduction," *Proc. of the International Symposium on Control, Communications, and Signal Processing*, 2006.

[14] A. Graps, "An Introduction to Wavelet", *IEEE Computational Science and Engineering*, vol. 2, no. 2, pp. 57-61, 1995.

[15] "Objective Measurement of Active Speech Level," *ITU-T, Geneva*, P. 56, March 1993.

ضمیمه الف

در اینجا قصد داریم انرژی سیگنال خروجی را به گونه‌ای تعریف کنیم که سه شرط مورد نظر برقرار باشد:

$$1) |x| = |n| \Rightarrow |\hat{s}|^2 = 0 \quad (\text{الف-1})$$

$$2) |n| = 0 \Rightarrow |\hat{s}|^2 = |x|^2 \quad (\text{الف-2})$$

$$3) 0 \leq |\hat{s}|^2 \quad (\text{الف-3})$$

از رابطه (الف-۱) می‌توان نتیجه گرفت:

$$|\hat{s}|^2 = (|x| - |n|) \times P \quad (\text{الف-4})$$

که منظور از P تابعی درجه اول از |x| و |n| است. حال برای برآورده شدن شرط دوم، P باید به صورت زیر تعریف شود:

$$P = (|x| + a|n|) \quad (\text{الف-3})$$

بنابراین:

$$|\hat{s}|^2 = (|x| - |n|) \times (|x| + a|n|) \quad (\text{الف-4})$$

در ادامه برای بدست آوردن مقدار a، از شرط سوم کمک می‌گیریم که نتیجه می‌شود:

$$-1 \leq a \leq 1 \quad (\text{الف-5})$$

بنابراین انرژی سیگنال خروجی باید به صورت زیر باشد:

$$|\hat{s}|^2 = (|x| - |n|)(|x| + a|n|); -1 \leq a \leq 1 \quad (\text{الف-6})$$