

بکارگیری بانک فیلتر WOLA در بهبود گفتار بر مبنای تبدیل ویولت به منظور کاهش پیچیدگی محاسباتی

سیامک رسولزاده، بهزاد زمانی دهکردی، بابک ناصرشریف، احمد اکبری

دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران

s_rasoulzadeh, bzamani@comp.iust.ac.ir, nasser_s, akbari@iust.ac.ir

چکیده - استفاده از تبدیل ویولت بجای تبدیل فوریه برای بهبود گفتار در سالهای اخیر بسیار مورد توجه واقع شده است. تبدیل ویولت با استفاده از پنجره های زمانی با طول متفاوت برای زیرباندهای فرکانسی مختلف، رزولوشن زمان-فرکانسی مناسبی برای مدل کردن سیگنالهای نا ایستانی همچون سیگنال گفتار فراهم می آورد. همچنین استفاده از بانک فیلتر WOLA موجب کاهش بار محاسباتی می شود. در این مقاله بانک فیلتر WOLA در یک سیستم کاهش نویز بکار گرفته شده است. نتایج نشان می دهد که بار محاسباتی با استفاده از بانک فیلتر WOLA نسبت به روش معمول بدون از دست دادن کیفیت ۱۷,۵٪ کاهش یافته است.

کلید واژه- تبدیل ویولت، بانک فیلتر WOLA، فیلتر وینر، درخت ویولت بسته ای.

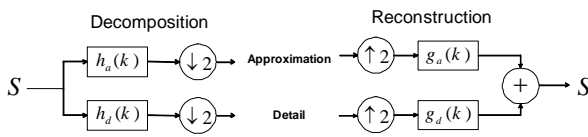
۱- مقدمه

عنوان نویز موسیقی اشاره نمود. علاوه بر این بنا به ماهیت تبدیل فوریه و نایستار بودن سیگنال گفتار، عملیات باید بر روی طیف دوره کوتاه سیگنال صورت گیرد که محدودیتهایی در تنظیم رزولوشن زمانی و فرکانسی ایجاد می نماید.

از دیگر روشهای بهبود گفتار استفاده از تبدیل ویولت است. ویژگی اصلی تبدیل ویولت بهره جستن از پنجره های زمانی با طول متفاوت برای باندهای فرکانسی مختلف است. به این ترتیب با استفاده از تبدیل ویولت می توان به دقت فرکانسی بالا در باندهای فرکانسی پایین و دقت فرکانسی پایین در باندهای فرکانسی بالا دست یافت. بنا به این خصوصیت، تبدیل ویولت ابزار قدرتمندی برای مدل کردن سیگنالهای نایستانی همانند سیگنال گفتار است که دارای تغییرات آرام در فرکانسهای پایین و تغییرات ناگهانی در فرکانسهای بالا می باشد به این ترتیب تبدیل ویولت محدودیت تبدیل فوریه را برای کار بر روی طیف دوره کوتاه دارا نیست.

استفاده از بهبود گفتار، در بسیاری از کاربردهایی که صدا در یک محیط نویزی دریافت می گردد، ضروری است. در سالهای اخیر با گسترش کاربرد هدستها در موبایل، لزوم استفاده از الگوریتمهای کارا برای ارتقای کیفیت صوت بطور جدی مطرح می باشد چرا که در سیستمهای کاهش نویز در شرایط واقعی مسائلی از قبیل حافظه مورد نیاز، تاخیر ایجاد شده و توان مصرفی دارای اهمیت خاصی می باشند. لذا نیاز به کاهش بار محاسباتی احساس می شود.

برای حذف نویز از سیگنال گفتار، می توان طیف فرکانسی سیگنال نویزی را به گونه ای اصلاح نمود که پس از بازیابی سیگنال در حوزه زمان، صداهای پس زمینه گفتار کاهش یابد. از جمله روشهای این زمینه، می توان به روش تفاضل طیف توان و استفاده از فیلترهای طیفی نظیر فیلتر وینر اشاره نمود [1]. از جمله اثرات نامطلوب این شیوه می توان به ایجاد اعوجاج در سیگنال و ایجاد مولفه های جدید نویز با



شکل ۱ - ساختار تبدیل و عکس تبدیل ویولت

۲- تبدیل ویولت

تبدیل ویولت توسط Grossman و Morlet پایه گذاری شده است [4].

شکل (۱) ساختار ساده تبدیل ویولت را نشان می دهد. در این تبدیل سیگنال ورودی به دو زیرباند بالا و پائین تفکیک می شود. و در ادامه برای بازسازی از دو فیلتر دیگر استفاده می شود. آنچه در شکل دیده می شود درخت ویولتی با یک سطح می باشد. $h_a(k)$ فیلتر پائین گذر و $h_d(k)$ فیلتر بالا گذر می باشند که پاسخ فرکانسی آنها وابسته به نوع تابع پایه ویولت می باشد و برای تجزیه سیگنال به کار می روند. چون سیگنال خروجی هر یک از فیلترهای پائین گذر و بالا گذر نصف فرکانس سیگنال ورودی است بنابر تئوری نایکوئیست می توان نیمی از نمونه ها را دور ریخت.

تبدیل ویولت را می توان به صورت درخت ویولت گسسته دوتائی و ویولت بسته ای (Wavelet Packet) در نظر گرفت. [4,5]

۳- حذف نویز از ضرایب ویولت

حذف نویز از ضرایب ویولت با اعمال فیلتر وینر در زیرباندها انجام می گیرد که این روش در زیر شرح داده شده است:

۳-۱- فیلتر وینر

رابطه (۲) فرمول فیلتر وینر را در حوزه فرکانس نشان می دهد. که در آن P_n طیف توان نویز و P_s طیف توان سیگنال تمیز می باشد.

$$H = \frac{P_s}{P_s + P_n} \quad (2)$$

برای محاسبه فیلتر وینر در حوزه ویولت طبق زیر عمل می شود. فرض کنید سیگنال گفتار تمیز s با نویز جمع شوند n آلوده شده باشد. آنگاه می توان گفت سیگنال

تبدیل فوریه دارای پیچیدگی محاسباتی $O(N \log N)$ بوده در صورتیکه تبدیل ویولت دارای پیچیدگی محاسباتی $O(N)$ می باشد. بنابراین بکار بردن تبدیل ویولت می تواند موجب کاهش بار محاسباتی گردد.

در الگوریتم های کاهش نویز به روش اصلاح طیفی طول فریم آنالیز و سنتز نقش مهمی در کیفیت ایفا می نماید. از طرف دیگر مرتبه محاسبات یک فریم متناسب با طول آن می باشد. بانک فیلتر WOLA با شکستن یک فریم به چند زیرفریم و اعمال تبدیل روی هر زیرفریم موجب صرفه جویی در میزان محاسبات می شود. در عین حالی که طول فریم مورد پردازش تغییر نمی کند.

در مرجع [2] احدی از WOLA بعنوان بانک فیلتر در پیاده سازی یک روش تشخیص گفتار استفاده شده است. در این مقاله تنها از قسمت آنالیز بانک فیلتر استفاده گردیده است و اشاره شده است که بانک فیلتر WOLA را می توان بصورت کامل در یک سیستم بهبود گفتار نیز بکار برد. در مرجع [3]، D. Hermann از بانک فیلتر WOLA در پیاده سازی کدکننده زیرباندهای Bluetooth (SBC) در سیستم صوتی بی سیم استفاده کرده است. در این مقاله یک کمک پردازنده WOLA برای انجام موازی بخش های مختلف الگوریتم SBC بکار رفته است.

در کار حاضر از تبدیل ویولت به همراه فیلتر وینر بجای تبدیل فوریه استفاده شده است. همچنین استفاده از WOLA در یک سیستم مبتنی بر ویولت به منظور کاهش محاسبات در این مقاله پیشنهاد شده است. نتایج نشان می دهد که در سیستم پیشنهادی با کاهش ناچیز در کیفیت شنیداری خروجی، حدود ۲۰٪ کاهش محاسبات بدست می آید.

ساختار ادامه مقاله به صورت زیر است. بخش دوم مقاله، به مروری کلی بر تبدیل ویولت اختصاص دارد. بخش سوم به نحوه محاسبه فیلتر وینر و تخمین نویز در حوزه تبدیل ویولت می پردازد. در بخش چهارم به بررسی بانک فیلتر WOLA پرداخته شده است. در ادامه بخش پنجم به ارزیابی اثر بانک فیلتر WOLA در سیستم کاهش نویز تخصیص داده شده است. در آخر، بخش ششم نیز حاوی یک جمع بندی کلی است.

یک نزدیک شده و برای فریمهای کوچک به صفر نزدیک می‌شود.

$E\{S_j(i, m)\}$ انرژی سیگنال تمیز در زیرباند λ ام فریم λ ام از رابطه (۸) بدست می‌آید.

$$E\{S_j(i, m)\} = E\{D_j(i, m)\} - est\{N_j(i)\} \quad (\lambda)$$

در حالت ایده آل برای تخمین ها از سیگنالهای نویز و تمیزی که در دسترس هستند استفاده می‌شود. ولی در شرایط واقعی که فقط سیگنال نویزی در اختیار است، قسمتی از سیگنال که فقط دارای نویز می‌باشد برای به روز کردن تخمین انرژی نویز استفاده می‌شود و تا یافتن بخش دیگری از سیگنال که فقط حاوی نویز باشد، از تخمین قبلی نویز استفاده می‌گردد. درکار حاضر، حالت ایده آل در نظر گرفته شده است.

۴- بانک فیلتر WOLA

سیستم هایی که قادر به بازسازی کامل سیگنال هستند محدودیت های خاصی را بر روی نحوه طراحی بانک فیلترهای آنالیز و سنتز اعمال می‌کنند. یکی از این سیستمها بانک فیلتر WOLA است. این ساختار در ابتدا برای پیاده سازی بانک فیلتر DFT طراحی شد. [7] مدل دیگری که برای این بانک فیلتر طراحی شد بانک فیلتر چند فازی است. این دو ساختار از لحاظ محاسباتی یکسان بوده و تنها در نحوه تحلیل داده‌ها متفاوت می‌باشند.

شکل (۳) بلوک دیاگرام بانک فیلتر آنالیز و سنتز WOLA را نشان می‌دهد. [8] که L طول فریم اولیه، N طول فریم ثانویه و R میزان همپوشانی دو فریم متوالی را نشان می‌دهند. شکل (۳-الف) بانک فیلتر آنالیز و شکل (۳-ب) بانک فیلتر سنتز را برای WOLA نشان می‌دهد.

در عمل صف ورودی هر بار با دریافت R نمونه جدید از سیگنال گفتار ورودی به همین میزان شیفت پیدا می‌کند. سپس صف ورودی در پنجره آنالیز ضرب می‌شود که در حقیقت یک فیلتر پایین گذر بطول L است (پنجره آنالیز با $W_A(1:L)$ نمایش داده شده است). در ادامه بردار حاصل به قسمتهایی به طول N تقسیم شده و این قسمتها باهم جمع می‌شوند تا یک بردار بطول N حاصل شود (این بردار را با $Z(1:N)$ نشان داده شده است). حال تبدیل اصلی روی این

نویزی x در حوزه زمان از رابطه (۳) تبعیت می‌کند.

$$x(i, m) = s(i, m) + n(i, m), \quad (3)$$

$$m = 0, 1, \dots, L - 1$$

که در آن m اندیس زمان، i اندیس فریم و L طول فریم می‌باشد. رابطه (۴) معادل حوزه ویولت معادله (۳) را نشان می‌دهد.

$$X_j(i, m) = S_j(i, m) + N_j(i, m), \quad (4)$$

$$m = 0, 1, \dots, L_j - 1$$

که $X_j(i, m)$ نمونه λ ام از تبدیل ویولت سیگنال گفتار نویزی در زیرباند λ ام از فریم λ ام، $S_j(i, m)$ نمونه λ ام از تبدیل ویولت سیگنال تمیز در زیرباند λ ام از فریم λ ام و $N_j(i, m)$ نمونه λ ام از تبدیل ویولت سیگنال نویز در زیرباند λ ام از فریم λ ام بوده و L_j طول زیرباند λ ام می‌باشد. می‌توان رابطه (۵) را برای بدست آوردن سیگنال بهبود یافته بیان کرد:

$$\hat{S}_j(i, m) = H_j(i).X_j(i, m) \quad (5)$$

فیلتر وینر مورد استفاده در رابطه بالا صورت رابطه (۶) تعریف می‌شود:

$$H_j(i) = \frac{E\{S_j(i, m)\}}{E\{S_j(i, m)\} + est\{N_j(i)\}} \quad (6)$$

در رابطه فوق $H_j(i)$ اندازه فیلتر برای زیرباند λ ام فریم λ ام می‌باشد که در رابطه (۸) بکار می‌رود. $E\{S_j(i, m)\}$ انرژی زیرباند λ ام فریم λ ام سیگنال تمیز و $est\{N_j(i)\}$ تخمین انرژی نویز در زیرباند λ ام فریم λ ام می‌باشد که از رابطه (۷) بدست می‌آید.

$$est\{N_j(i)\} = \alpha.est\{N_j(i-1)\} + (1-\alpha).E\{N_j(i, m)\} \quad (7)$$

در این رابطه $E\{N_j(i, m)\}$ انرژی زیرباند λ ام فریم λ ام نویز تخمین زده شده، $est\{N_j(i-1)\}$ تخمین انرژی نویز در زیرباند λ ام فریم $(i-1)$ ام و α فاکتور فراموشی می‌باشد که بر اساس روش واریانس بالانس شده مطرح شده در [6] بدست می‌آید. در این روش اندازه ضریب فراموشی به طول فریم بستگی دارد به این صورت که α برای فریم های بزرگ به

critically sampled خواهد بود، که در آن فرکانس قطع پنجره آنالیز برابر $2\pi/N$ خواهد شد.

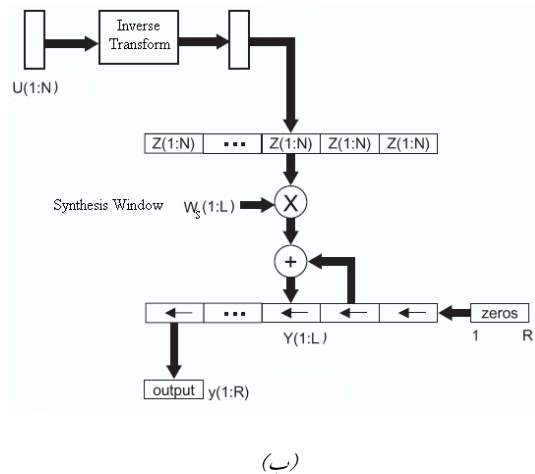
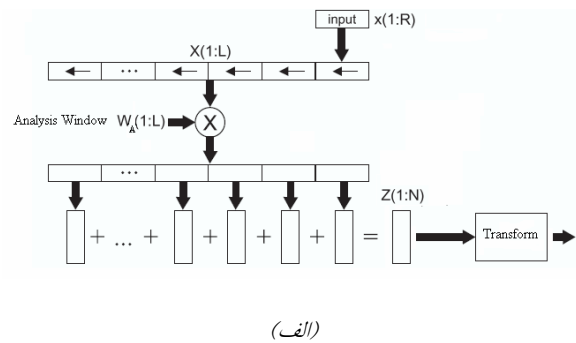
در سیستم‌هایی که از بانک‌فیلترهای آنالیز و سنتز برای تحلیل و بازسازی کامل سیگنال استفاده می‌کنند، بیشترین بار محاسباتی مربوط به بانک‌فیلترهای آنالیز و سنتز می‌باشد. [9] پیچیدگی محاسباتی در روشهای معمول در بانک‌فیلترها، از درجه L می‌باشد، که L طول یک فریم از سیگنال ورودی است. اما در بانک‌فیلتر WOLA این تبدیل می‌تواند به نسبت L/N ، بار محاسباتی را کمتری کند. [2]

۵- آزمایشات و نتایج اثر بانک‌فیلتر WOLA

به منظور مقایسه اثر فیلتربانک WOLA یکبار عمل فریم‌بندی را با استفاده از روش معمولی انجام داده و یکبار با استفاده از WOLA این عمل را انجام می‌دهیم.

شکل (۴) بلوک دیاگرام مربوط به حذف نویز بر مبنای ویولت با استفاده از فریم‌بند معمولی را نشان می‌دهد. همانطور که در شکل مشاهده می‌کنید، ابتدا سیگنال ورودی به فریم‌هایی با طول مشخص با در نظر گرفتن همپوشانی تقسیم می‌شود. سپس روی هر فریم پنجره ای اعمال می‌شود. در ادامه فریم توسط بانک‌فیلتر ویولت به N زیرباندها تقسیم می‌گردد و روی هر زیرباند یک فیلتر وینر ایده آل اعمال می‌شود. پس از آن با استفاده از عکس تبدیل ویولت فریم بهبود یافته بدست می‌آید و در نهایت با در نظر گرفتن همپوشانی فریم‌ها در کنار هم قرار می‌گیرند تا سیگنال بهبود یافته نهائی بدست آید.

شکل (۵) بلوک دیاگرام مربوط به حذف نویز بر مبنای ویولت با استفاده از بانک‌فیلتر WOLA را نشان می‌دهد. همانطور که در شکل مشاهده می‌کنید، ابتدا سیگنال ورودی به فریم‌هایی با طول مشخص با در نظر گرفتن همپوشانی تقسیم می‌شود. سپس روی هر فریم پنجره ای اعمال می‌شود. در ادامه هر فریم به چند قسمت مساوی تقسیم شده و این قسمت‌ها با هم جمع شده و یک فریم با طول کوچکتر را ایجاد می‌کنند. سپس فریم بدست آمده توسط بانک‌فیلتر ویولت به N زیرباندها تقسیم می‌گردد و روی هر زیرباند یک فیلتر وینر ایده آل اعمال می‌شود. پس از آن با استفاده از عکس تبدیل ویولت فریم بهبود یافته بدست می‌آید و در نهایت با در نظر گرفتن روش سنتز WOLA

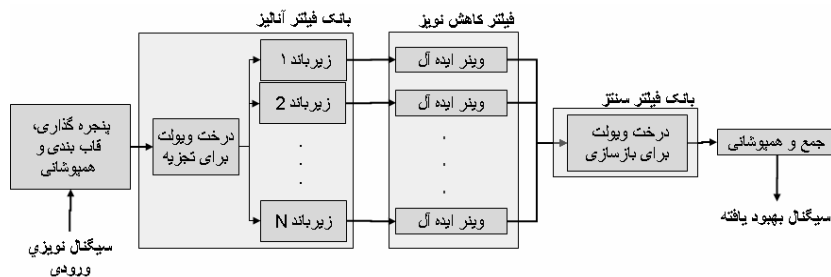


شکل ۳ - بانک‌فیلتر آنالیز و سنتز WOLA

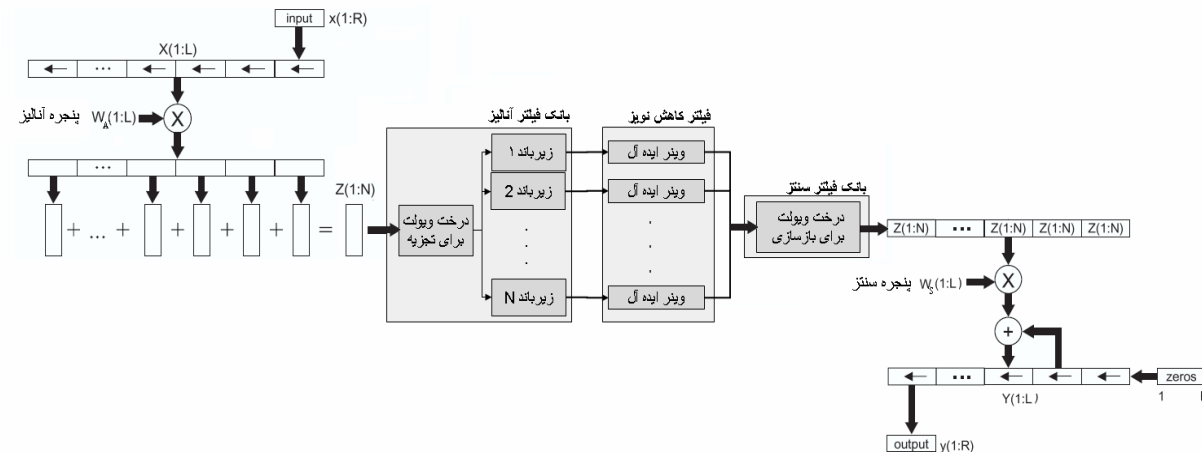
بردار اعمال می‌گردد. این تبدیل می‌تواند تبدیل فوریه و یا ویولت و یا هر تبدیل دیگری باشد مثلاً در بانک‌فیلتر DFT یک تبدیل فوریه سریع بر روی بردار N تایی حاصل انجام می‌شود. سپس در حوزه فرکانس بر روی بردار حاصل پردازش بهبود کیفیت انجام می‌گیرد.

در ادامه برای انجام بازسازی کامل ابتدا تبدیل معکوس قسمت آنالیز اعمال می‌گردد که مثلاً در بانک‌فیلتر DFT از نوع IFFT خواهد بود. سپس بردار N تایی حاصل بصورت متناوب تکرار می‌شود تا به طول پنجره سنتز برسد. پنجره سنتز با $W_S(1:L)$ نمایش داده شده است. در نهایت پنجره سنتز در بردار حاصل ضرب خواهد شد و طی یک عمل همپوشانی و جمع با بردار خروجی قبلی و یک عمل شیفت R تایی نمونه‌های خروجی هر مرحله حاصل می‌شوند و با کنار هم قرار دادن این نمونه‌ها سیگنال بهبود یافته نهایی حاصل می‌شود.

اگر مقادیر پارامترهای R و N برابر باشند بانک‌فیلتر از نوع



شکل ۴ - بلوک دیاگرام سیستم کاهش نویز بر مبنای ویولت با فریم بندی ساده



شکل ۵ - بلوک دیاگرام سیستم کاهش نویز بر مبنای ویولت با استفاده از بانک فیلتر WOLA

مساوی از فریم اولیه بدست می آید. در کاربردهای بهبود گفتار اندازه همپوشانی معمولاً یکی از این دو مقدار است.

برای بررسی تأثیر استفاده از روش WOLA بعنوان بانک فیلترهای آنالیز و سنتز در روش wavelet به مقایسه پیچیدگی محاسباتی دو روش پرداخته شده است. ابتدا پیچیدگی محاسباتی روش wavelet و سپس پیچیدگی محاسباتی تلفیق دو روش wavelet و WOLA بررسی می شوند. برای معتبر بودن مقایسه باید طول فریم اولیه و تعداد زیرباندهای حاصل در هر دو روش یکسان در نظر گرفته شود.

پارامترهای روش wavelet شامل طول فریم اولیه (N) و طول تابع پایه (m) و تعداد زیرباندها (S) خواهد بود. در روش wavelet ابتدا یک فریم از سیگنال ورودی انتخاب می شود و در پنجره آنالیز بطول همان فریم ضرب می شود. سپس این فریم طی مراحل S به S زیرباند تقسیم می شود. برای داشتن S زیرباند باید $\text{Log}_2 S$ مرحله تکرار شود و در

فریمهای نهائی ساخته شده و با کنار هم قرار دادن این فریمها، سیگنال بهبود یافته نهائی بدست می آید.

۵-۱- اثر بانک فیلتر WOLA بر پیچیدگی محاسباتی

از آنجائیکه بیشترین بار محاسباتی هر سیستم بهبود گفتار در بانک فیلترهای آنالیز و سنتز رخ می دهد، تمام تلاشها بمنظور کاهش محاسبات در این دو بانک فیلتر صورت می گیرد. استفاده از روش WOLA در بانک فیلترهای آنالیز و سنتز با کاهش طول فریم اولیه و بازسازی کامل آن باعث کاهش محاسبات خواهد شد. کاهش طول فریم اولیه با شکستن آن به دو یا چهار تکه و جمع آنها و تولید یک فریم ثانویه صورت می گیرد. تعداد شکستن فریمها به میزان همپوشانی فریمهای متوالی بستگی دارد. اگر همپوشانی فریمهای متوالی 50% باشد فریم ثانویه از مجموع دو نیمه فریم اولیه بدست خواهد آمد و اگر همپوشانی به 75% برسد فریم ثانویه از مجموع چهار تکه

روش فریم اولیه به طول N انتخاب شده و در پنجره آنالیز ضرب می‌شود. فریم ورودی wavelet دارای طول کمتر یعنی N/a خواهد بود. پارامتر a به میزان همپوشانی فریم‌های متوالی بستگی دارد. در صورتیکه همپوشانی 50% باشد $a=2$ و اگر همپوشانی 75% باشد $a=4$ خواهد بود. مراحل آنالیز و سنتز wavelet روی این فریم انجام می‌شوند و فریمی بطول N/a از روش wavelet حاصل می‌شود. سپس با بکارگیری مرحله سنتز WOLA که شامل N عمل ضرب و جمع است به فریم نهایی بعنوان خروجی دست پیدا می‌کنیم. تعداد محاسبات که شامل عملیات‌های ضرب و جمع است بصورت رابطه (۱۱) خواهد بود.

$$2N + 4m \frac{N}{a} L + 4m(m-1)(2^L - 1) \quad (11)$$

به عنوان مثال در جدول (۱) تعداد محاسبات به ازای هر خروجی در دو روش برای طول پنجره، تعداد زیرباندها و همپوشانی 50% نشان داده شده است.

جدول ۱ - تعداد محاسبات به ازای هر خروجی در دو روش

N	m	L	تعداد ضرب و تعداد جمع به ازای هر خروجی		میزان کاهش محاسبات
			Wavelet with WOLA	Wavelet without WOLA	
256	40	6	4035.25	4993.25	19.19%
256	40	5	2315.25	3113.25	25.63%

۵-۲- ارزیابی اثر بانک فیلتر WOLA

تخمین مؤلفه های فیلتر وینر ایده آل از روی دو سیگنال نویز و سیگنال گفتار انجام می‌گیرد. در این ارزیابی از ویولت بسته‌ای برای تجزیه سیگنال به زیرباندها استفاده شد، که در آن زیرباندها یکنواخت می‌باشند. همچنین طول فریم ۱۲۸ و ۲۵۶، تعداد زیرباندها از ۳۲ تا ۱۲۸ و اندازه همپوشانی ۵۰٪ در نظر گرفته شده است. آزمایشها روی ۶ فایل شامل صدای زن و مرد و در حضور نویز سفید گوسی با نرخ سیگنال به نویز ورودی 10dB انجام گرفته است. فایل‌های ورودی با فرکانس 8kHz نمونه برداری می‌شود و هر نمونه 16 بیت می‌باشد. شکل (۶) مقادیر SNRI را برای دو روش کاهش نویز با استفاده از بانک فیلتر WOLA و بدون استفاده از آن زمانی که طول فریم ۱۲۸

هر مرحله کانولوشن یک فریم با دو فیلتر بدست آید. این دو فیلتر از همان تابع پایه بدست می‌آیند و شامل دو فیلتر پایین‌گذر و بالاگذر هستند که برای بدست آوردن فریم‌های approximation و detail استفاده می‌شوند. نتیجه هر مرحله شامل دو فریم جدید خواهد بود که طول هر یک نصف طول فریم اولیه است. اینکار برای مرحله بعدی نیز صورت می‌گیرد تا به تعداد زیرباندهای موردنظر دست پیدا کنیم. با داشتن S زیرباند، مرحله پردازش روی سیگنال صوتی صورت می‌گیرد و سپس همان مراحل آنالیز بصورت عکس برای بدست آوردن سیگنال اولیه تکرار خواهد شد.

برای بدست آوردن میزان محاسبات آنالیز باید در نظر داشته باشیم که در هر مرحله دو کانولوشن بین فریم ورودی آن مرحله با دو فیلتر بالا و پایین‌گذر صورت می‌گیرد. تعداد ضرب‌ها و جمع‌های مرحله m به طول فریم ورودی آن مرحله $(N/2^{n-1})$ و طول دو فیلتر بالا و پایین‌گذر (m) بستگی دارد. هر عمل کانولوشن بین دو تابع بطول‌های N و m به تعداد $m(N+m-1)$ عمل ضرب و جمع نیاز دارد. تعداد محاسبات در مرحله آنالیز wavelet برابر خواهد بود با:

$$\sum_{n=1}^L \left(\frac{N}{2^{n-1}} + m - 1 \right) \cdot 2^n \cdot m = 2mNL + 2m(m-1)(2^L - 1) \quad (9)$$

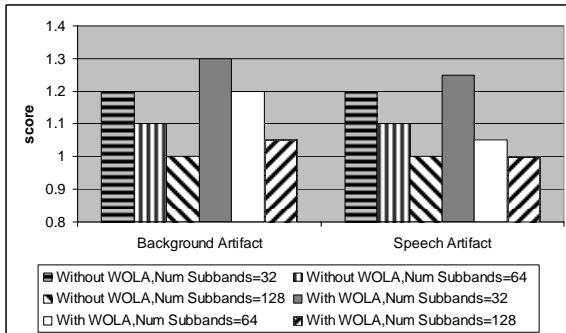
که در آن N طول فریم اولیه، m طول تابع پایه ویولت و L تعداد سطوح مورد نیاز برای رسیدن به S زیرباند است. ($L = \log_2 S$)

همین محاسبات در مرحله سنتز نیز تکرار خواهد شد. سپس برای بازسازی کامل یک عمل همپوشانی-جمع که شامل N عمل جمع است صورت می‌گیرد. بنابراین تعداد ضرب‌ها و جمع‌ها در روش wavelet بصورت رابطه (۱۰) خواهد بود.

$$N + 4mNL + 4m(m-1)(2^L - 1) \quad (10)$$

با بکارگیری روش WOLA بعنوان بانک فیلترهای آنالیز و سنتز در روش wavelet طول فریم اولیه بعنوان فریم ورودی اولین مرحله wavelet کاهش خواهد یافت. در این

قابل تحمل نیستند، امتیاز سه داده می‌شود.



شکل ۷ - نتایج تست شنوائی برای دو روش کاهش نویز با استفاده از بانک‌فیلتر WOLA و بدون استفاده از آن با طول فریم ۱۲۸ و تعداد زیرباند ۳۲، ۶۴ و ۱۲۸ در حضور نویز سفید با نرخ سیگنال به نویز 10dB

شکل (۷) نتایج تست شنوائی را برای دو روش کاهش نویز نشان داده است. بر اساس نتایج تست شنوائی، هر دو روش در طول فریمها و تعداد زیرباندها دارای عملکرد تقریباً مشابهی می‌باشند. به این معنی که کیفیت بدست آمده در شرایط برابر از هر دو الگوریتم به هم نزدیک می‌باشد.

زمان محاسبه برای هر الگوریتم در شرایط مختلف اندازه‌گیری شد. سپس میانگین زمانهای بدست آمده، محاسبه شد. طبق نتایج عملی، الگوریتم کاهش نویز بر مبنای ویولت با استفاده از بانک‌فیلتر WOLA 17.5% نسبت به الگوریتم کاهش نویز بر مبنای ویولت بدون استفاده از بانک‌فیلتر WOLA کاهش زمان محاسبه را به همراه داشت.

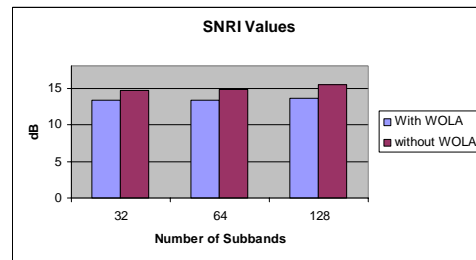
۶- نتیجه‌گیری

با توجه به نتایج بدست آمده از ارزیابی‌های نظری و شنوائی مشخص شد که بهبود کیفیت گفتار با استفاده از بانک‌فیلتر WOLA در یک سیستم کاهش نویز مبتنی بر ویولت دارای کیفیتی در حد قابل قبول می‌باشد. از طرفی سیستم کاهش نویز ارائه شده با بانک‌فیلتر WOLA موجب کاهش ۱۷,۵ درصدی در بار محاسباتی می‌شود که این کاهش محاسبات در کاربردهای عملی حائز اهمیت می‌باشد.

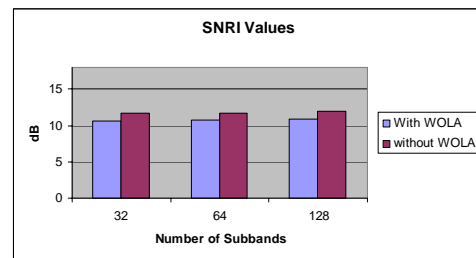
مراجع

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

و ۲۵۶ و تعداد زیرباند ۳۲، ۶۴ و ۱۲۸ می‌باشد، در حضور نویز گوسی سفید با نرخ سیگنال به نویز 10dB نشان می‌دهد. چنانچه در شکل دیده می‌شود، در هر طول فریمی و هر تعداد زیرباندی مقادیر SNRI بدست آمده برای هر دو روش تقریباً برابر می‌باشد و تفاوت آنها در حدود 1dB می‌باشد که قابل صرفنظر کردن می‌باشد. نتایج مشابه برای نویزهای همهمه و نویز ماشین با نرخ سیگنال به نویز مختلف نیز بدست آمد.



طول فریم = ۱۲۸



طول فریم = ۲۵۶

شکل ۶ - مقادیر SNRI برای دو روش کاهش نویز مبتنی بر ویولت با استفاده از بانک‌فیلتر WOLA و بدون استفاده از آن با طول فریم ۱۲۸ و ۲۵۶، تعداد زیرباند ۳۲، ۶۴ و ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز 10dB

همانطور که در شکل (۶) دیده می‌شود، در هر دو روش حذف نویز طول فریم مهمترین پارامتر در سیستم بهبود کیفیت گفتار می‌باشد.

تست شنوائی به این صورت انجام شده است که شنوندگان به سیگنال بهبود یافته گوش داده و به ۲ جنبه نویز زمینه (Background Artifact) و تخریب صدا (Speech Artifact) امتیازی بین ۱ تا ۳ داده‌اند. در این تست به سیگنالهای بهبود یافته که کیفیت خوبی دارند امتیاز یک، به سیگنالهای بهبود یافته ای که کیفیت پایین تری دارند ولی از نظر شنونده قابل تحمل هستند، امتیاز دو و به سیگنالهایی که کیفیت نازلی دارند و از دید شنونده



- [6] Martin R., Noise Power Spectral Density Estimation Based on Optimal smoothing and Minimum Statistic. IEEE Transaction on Speech and Audio Processing, Vol.9, No.5, July 2001.
- [7] R. E. Crochiere, "A Weighted Overlap-Add Method of Short-Time Fourier Analysis / Synthesis", IEEE Transaction on Acoustics, Speech, Signal Processing, vol. ASSP-28, pp 98-102, Feb 1980
- [8] R.E. Crochiere and L.R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, 1983
- [9] de Bont, F., Groenewegen, M., and Oomen, W., "A High Quality Audio-Coding System at 128 kb/s", 98 th AES Convention, Feb. 25-28, 1995.
- [2] S.M. ahadi, H. Sheikhzadeh, R. L. Brennan, G. H. Freeman, "An Efficient Front-End For Automatic Speech Recognition", 2003 international conference on electronics, circuits and systems (ICECS2003)
- [3] D. Herman, R. L. Brennan, H. sheikhzadeh, E. Cornu, "Low Power Implementation of The Bluetooth Subband Audio Codec", Proceeding of the 2004 international Conference on Acoustics Speech and Signal Processing (ICASSP 2004)
- [4] Grossman A., Martinet R., Morlet J., *Reading and understanding continuous wavelet transform*. Pringer-Verlag, pp. 2-20, 1989.
- [5] Cohen I., *Enhancement of speech using bark-scaled wavelet packet decomposition*. European Conf. on Speech Communication and Technology, 2001.