

کاهش بار محاسباتی سیستمهای حذف نویز مبتنی بر موجک با استفاده از WOLA

بهزاد زمانی دهکردی^{*}، سیامک رسولزاده[†]، بابک ناصر شریف[‡]، احمد اکبری[§]

چکیده

سیستمهای بهبود گفتار مبتنی بر موجک موفقیت خود را در عمل نشان داده اند، با استفاده از این سیستم ها سیگنال صوتی در هر یک از دو حوزه زمان و فرکانس با رزولوشن دلخواه تجزیه می گردد. سپس با استفاده از آستانه گذاری یا روشهای فیلتر کردن اجزا سیگنال بهبود یافته بازسازی می گردد. در مقاله حاضر یک چارچوب مناسب برای حذف نویز با استفاده از تبدیل موجک پیشنهاد شده است که در آن پیچیدگی محاسباتی با استفاده از روش WOLA محدود شده است. نتایج بهبود گفتار و ارزیابی تعداد محاسبات، نشان می دهند که بکارگیری بانک فیلتر WOLA نه تنها سبب کاهش محاسبات سیستم حذف نویز مبتنی بر موجک می گردد، بلکه کیفیت سیگنال بهبود یافته را نیز به خوبی حفظ می کند. درصد کاهش محاسبات با توجه به طول فریم ورودی سیستم کاهش نویز مابین ۲۶% تا ۴۲% می باشد.

کلمات کلیدی

تبدیل موجک، فیلتر بانک WOLA، فیلتر وینر، درخت موجک بسته ای.

Reducing Computational Complexity of Wavelet based Noise Reduction Algorithms Using WOLA

Behzad Zamani, Siamak Rasulzadeh, Babak Nassersharif, ahmad akbari

Abstract

Speech enhancement systems using wavelet transform recently show a good performance in practice. In these systems, first, speech signal is divided into sub-bands. Then, thresholding or filtering methods are applied into each sub-band. Finally, enhanced speech signal is reconstructed using modified sub-bands. In this paper, we propose a method for reducing computational complexity of wavelet transform using WOLA. Speech enhancement and computational cost evaluation results show that using WOLA, not only reduce computational cost in wavelet noise reduction systems, but also save quality of speech signal. The amount of this reduction in computation cost is between 26% and 42% that is dependent to length of input frame.

Keywords

Wavelet transform, WOLA filter bank, Wiener filter, Wavelet packet tree.

^{*} دانشجوی دوره دکترا، دانشکده کامپیوتر، دانشگاه علم و صنعت، bzamani@iust.ac.ir

[†] دانشجوی دوره کارشناسی ارشد، دانشکده کامپیوتر، دانشگاه علم و صنعت، s_rasulzadeh@comp.iust.ac.ir

[‡] دانشجوی دوره دکترا، دانشکده کامپیوتر، دانشگاه علم و صنعت، s_nasser@iust.ac.ir

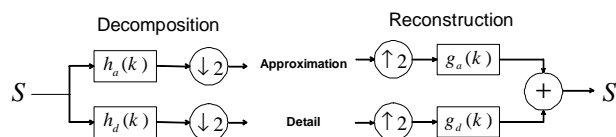
[§] عضو هیئت علمی دانشکده کامپیوتر، دانشگاه علم و صنعت، akbari@iust.ac.ir

۱- مقدمه

محاسباتی درخت موجک تشریح شده است. در بخش چهارم به بررسی بانک فیلتر WOLA و بکارگیری آن در سیستم کاهش نویز به منظور کاهش پیچیدگی محاسباتی پرداخته شده است. سپس در بخش پنجم ارزیابی نتایج حاصل از اثر روش پیشنهادی یعنی بکارگیری بانک فیلتر WOLA در سیستم حذف نویز گزارش شده است و در انتها نتیجه‌گیری آورده شده است.

۲- تبدیل موجک

تبدیل موجک پیوسته (که انتگرال موجک نیز نامیده می‌شود) توسط Grossman و Morlet پایه‌گذاری شده است [۶]. شکل ۱ ساختار ساده تبدیل موجک را نشان می‌دهد. در این تبدیل سیگنال ورودی به دو زیرباند بالا و پائین تفکیک می‌شود. و 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 که پاسخ فرکانسی آنها وابسته به نوع تابع پایه موجک می‌باشد و برای تجزیه سیگنال به کار می‌روند. چون سیگنال خروجی هر یک از فیلترهای پائین‌گذر و بالاگذر نصف فرکانس سیگنال ورودی است بنابراین تئوری ناپکوئیست می‌توان نیمی از نمونه‌ها را حذف نمود.



شکل (۱): ساختار تبدیل و عکس تبدیل موجک

مجموعه موجکها همانند یک بانک فیلتر برای آنالیز سیگنال می‌باشند. می‌توان تبدیل موجک را به صورت دو درخت موجک گسسته دوتائی و موجک بسته ای** در نظر گرفت. [۶، ۷]

در درخت بازسازی از فیلترهای جداگانه سنتز که $g_a(k)$ و $g_d(k)$ نامیده می‌شوند برای بازسازی استفاده می‌شود. البته همانگونه که هنگام تجزیه درخت موجک عمل نمونه‌کاهی انجام شد، هنگام بازسازی نیز عمل افزایشی انجام می‌شود.

۳- سیستم کاهش نویز بر مبنای موجک

۳-۱- دیاگرام سیستم کاهش نویز مبتنی بر موجک

روشهای حذف نویز در حوزه موجک را می‌توان به دو دسته تقسیم نمود، روشهای بر مبنای آستانه‌گذاری و روشهای فیلترینگ. در سیگنال گفتار، غالباً قسمت عمده انرژی در تعداد کمی از ضرایب تبدیل موجک متمرکز است. این ضرایب نسبت به ضرایب سیگنال نویز که انرژی آن بر روی تعداد زیادی از ضرایب گسترش می‌یابند، بزرگتر هستند.

حذف نویز از سیگنال گفتار در بسیاری از کاربردها از جمله تلفنهای بدون گوشی که صدا در یک محیط نویزی دریافت می‌گردد، یک نیاز ضروری است. روشهای حذف نویز از سیگنال گفتار را می‌توان به دو گروه کلی تقسیم نمود: روشهای اصلاح طیف فرکانسی سیگنال نویزی مبتنی بر تبدیل فوریه و روشهای حذف نویز مبتنی بر تبدیل موجک. را به گونه ای اصلاح نمود که پس از بازیابی سیگنال در حوزه زمان، صداهای پس زمینه گفتار کاهش یابد. در روشهای مبتنی بر تبدیل فوریه، بنا به ماهیت تبدیل فوریه و نایبستان بودن سیگنال گفتار، عملیات باید بر روی طیف دوره کوتاه سیگنال صورت گیرد که محدودیتهایی در تنظیم رزولوشن زمانی و فرکانسی ایجاد می‌نماید و علاوه بر این اثرات نامطلوبی همچون اعوجاج در سیگنال گفتار و نویز موسیقی را به همراه دارد. از جمله شناخته‌ترین روشهای مبتنی بر تبدیل فوریه می‌توان به روش تفاضل طیف توان و استفاده از فیلترهای طیفی نظیر فیلتر وینر اشاره نمود [۱، ۲].

در روشهای مبتنی بر تبدیل موجک، از پنجره‌های زمانی با طول متفاوت برای باندهای فرکانسی مختلف استفاده می‌گردد برای حذف نویز استفاده از آستانه‌گذاری ضرایب تبدیل موجک، متداولترین روش است [۱۱، ۱۲، ۱۳]. به علت عملکرد ضعیف روشهای آستانه در نویزهای رنگی، روش آستانه‌گیری در باندهای بحرانی درخت تجزیه موجک ارائه گردیده است [۵]. به این ترتیب که ابتدا آستانه ماسک نویز برای یک فریم تخمین زده شده و سپس این آستانه برای تنظیم آستانه نهائی بکار گرفته می‌شود که اندازه آستانه بطور اتوماتیک در باندهای بالا در طول یک قطعه واکندار و در باندهای پائین در طول یک قطعه بی‌واک افزایش می‌یابد [۵]. برای حل مشکل روشهای حد آستانه، استفاده از فیلترهایی نظیر وینر در حوزه تبدیل موجک برای یک قاب پیشنهاد شده است که پارامترهای فیلتر وابسته به شرایط آماری ضرایب تبدیل موجک تنظیم می‌گردد [۱، ۲]. نگارندگان این مقاله نیز در کار پیشین خود [۱]، فیلتر وینر را در درخت هرس شده تبدیل موجک بسته‌ای مبتنی بر مقیاسهای فرکانسی Mel بکار گرفته‌اند که نتایج شنوایی و نظری مناسبی داشته است. در کار حاضر با استفاده از بانک فیلتر WOLA، روشی برای کاهش محاسبات سیستمهای حذف نویز موجک ارائه گردیده است که کارایی سیستم حذف نویز را نیز کم نمی‌کند. بانک فیلتر WOLA پیش از این برای کاهش محاسبات در تبدیل فوریه مورد استفاده قرار گرفته و پیاده‌سازی نیز گردیده است [۳، ۴، ۱۰]. این بانک فیلتر با تقسیم یک قاب به چندین زیرقاب و بکارگیری تبدیل فوریه بر روی هر زیرقاب، زمینه کاهش محاسبات را فراهم می‌آورد بدون اینکه طول قاب را تغییر دهد [۳، ۴، ۱۰].

در ادامه بخش دوم به معرفی تبدیل موجک اختصاص دارد. در بخش سوم سیستم کاهش نویز در حوزه موجک تعریف شده است و به نحوه محاسبه فیلتر وینر در حوزه موجک می‌پردازد و در ادامه پیچیدگی

** Wavelet Packet Tree

جدول (۱): تعداد جمع و ضرب حقیقی در بانک فیلتر موجک به ازای R خروجی

عملیات	جمع حقیقی	ضرب حقیقی
windowing	-----	N
Decomposition	$L(m-1)N + (m-1)^2.(2^{L+1} - 2 - L)$	$L.m.N + m(m-1).(2^{L+1} - 2 - L)$
Reconstruction	$L(m-1)N + (m-1)^2.(2^{L+1} - 2 - L)$	$L.m.N + m(m-1).(2^{L+1} - 2 - L)$
Overlap-Add	N	-----
مجموع	$N + 2[L(m-1)N + (m-1)^2.(2^{L+1} - 2 - L)]$	$N + 2[L.m.N + m(m-1).(2^{L+1} - 2 - L)]$

نویز استفاده می شود. درخت برابر L از رابطه ۷ بدست می آید. باید توجه داشت که طول فیلترهای پایین‌گذر و بالاگذر در همه سطوح ثابت است.

$$\begin{aligned}
 MUL - computations &= \sum_{i=1}^L 2^i . m . \left[\frac{N}{2^i} + \sum_{k=1}^i \frac{m-1}{2^k} \right] \\
 &= L.m.N + m(m-1)(2^{L+1} - 2 - L) \\
 ADD - computations &= \sum_{i=1}^L 2^i . (m-1) . \left[\frac{N}{2^i} + \sum_{k=1}^i \frac{m-1}{2^k} \right] \\
 &= L.(m-1).N + (m-1)^2(2^{L+1} - 2 - L)
 \end{aligned} \quad (7)$$

برای برآورد میزان محاسبات در بانک فیلتر موجک باید پارامترهایی از این بانک فیلتر که در میزان محاسبات موثرند را مورد نظر قرار دهیم. از این پارامترها می توان به طول فریم ورودی، تعداد سطوح آنالیز، مرتبه تابع پایه، همپوشانی پنجره های متوالی و ساختار درخت موجک اشاره کرد.

در این بانک فیلتر همانگونه که در شکل ۲ دیده شد برای تجزیه سیگنال ورودی به تعدادی زیربانند از یک درخت موجک استفاده می شود. درخت موجک می تواند دارای ساختار متقارن یا نامتقارن باشد. نحوه تولید زیرباندها در محاسبه میزان عملیات های ضرب و جمع در درخت موجک نقش عمده ای دارد، از اینرو در ادامه به ارزیابی محاسبات انجام شده در بانک فیلتر آنالیز پرداخته می شود.

۳-۳- پیچیدگی محاسباتی موجک

۱-۳-۳- پیچیدگی محاسباتی درخت موجک بسته ای

چنانچه در قسمت دوم مقاله دیده شد، محاسباتی که طی تشکیل درخت موجک و بازسازی درخت موجک انجام می شود به عملیات کانولوشن بستگی دارد، زیرا تولید هر سطح از درخت موجک به گذر سیگنال ورودی آن سطح از فیلترهای بالاگذر و پایین گذر مربوط می شود.

می دانیم که کانولوشن دو بردار بطول های N و m باعث تولید بردار خروجی بطول (N + m - 1) می شود. از آنجاییکه برای تولید هر خروجی در عمل کانولوشن به m عمل ضرب و m عمل جمع نیاز داریم (با فرض m < N) لذا میزان محاسبات لازم برای یک عمل کانولوشن برابر m(N + m - 1) عمل ضرب و جمع می شود.

در هر مرحله از ساختن درخت موجک تعداد عملیات کانولوشن دو برابر می شود. همچنین تعداد خروجی های هر کانولوشن برابر (N + m - 1)/2 خواهد بود. بنابراین در هر سطح میزان محاسبات به طول فریم ورودی آن سطح و تعداد زیرباندهای آن سطح بستگی دارد. یعنی در سطح mام تعداد زیرباندها برابر 2^m خواهد بود. بنابراین می توان گفت که در سطح mام تعداد 2^m عمل کانولوشن انجام می شود. تعداد محاسبات در تشکیل درخت موجک را با فرض طول فریم اولیه برابر N و طول فیلترهای بالاگذر و پایین گذر برابر m و تعداد سطوح تجزیه

از آنجاییکه برای بازسازی درخت موجک در بانک فیلتر سنتز از همین روش استفاده می شود بنابراین میزان محاسبات تبدیل موجک معکوس نیز به همین میزان خواهد بود.

در جدول ۱ میزان محاسبات انجام شده در بانک فیلتر موجک بدون در نظر گرفتن فیلتر وینر بعنوان فیلتر حذف نویز برای یک فریم N نمونه ای نشان داده شده است.

میزان همپوشانی پنجره های متوالی بر تعداد محاسباتی که به ازای هر خروجی انجام می شود تاثیر گذار است. اگر همپوشانی پنجره ها برابر P باشد میزان پیشروی پنجره در حوزه زمان برابر R خواهد بود که برابر $(1 - P) * N$ خواهد بود. اگر همپوشانی پنجره ها برابر 50% باشد محاسبات انجام شده در جدول ۱ به ازای هر N/2 خروجی انجام می گیرد و اگر میزان همپوشانی به 75% برسد این محاسبات برای N/4 خروجی انجام می شود که نشان دهنده دو برابر شدن محاسبات به ازای هر خروجی می باشد.

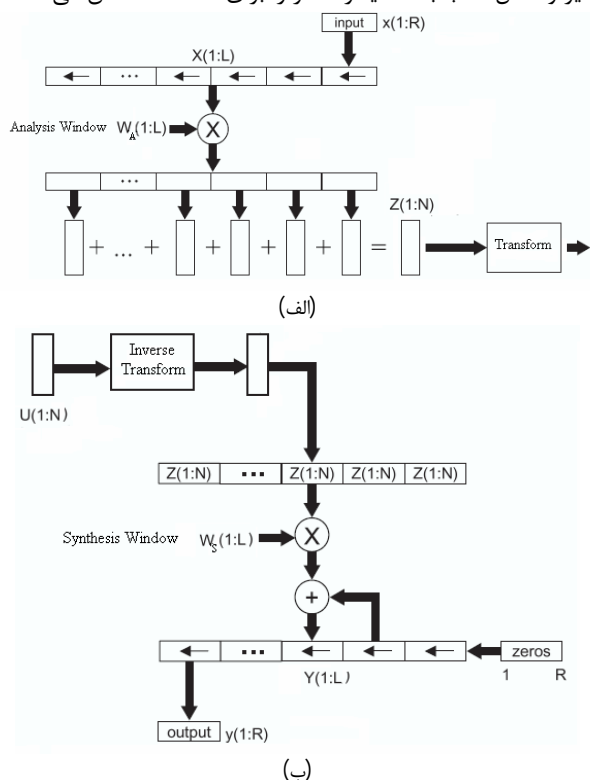
۲-۳-۳- پیچیدگی محاسباتی بانک فیلتر Mel

فرض کنید که $E \subset \{(l, n): 0 \leq l \leq L, 0 \leq n \leq 2^l\}$ نشان دهنده مجموعه گره های درخت موجک باشد، که l سطح گره و n شماره گره در آن سطح می باشد. برای گره های انتهائی درخت، پهنای باند و فرکانس مرکزی هر باند به صورت رابطه ۸ می باشد:

$$\begin{aligned}
 f_{l,n} &= 2^{-l} [G_c^{-1}(n) + 0.5] . F_s / 2 \\
 \Delta_{l,n} &= 2^{-l} . F_s / 2
 \end{aligned} \quad (8)$$

که در آن GC انحراف جایگشت کد gray و Fs فرکانس نمونه برداری می باشد. [۷]

همپوشانی دو فریم متوالی را نشان می‌دهند. شکل ۴-الف بانک فیلتر آنالیز و شکل ۴-ب بانک فیلتر سنتز را برای WOLA نشان می‌دهد.

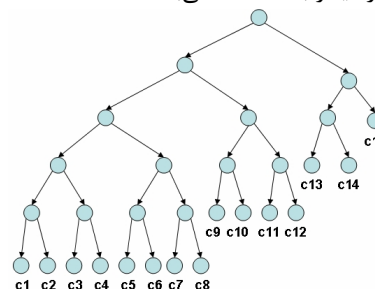


شکل (۴): بانک فیلتر آنالیز و سنتز WOLA

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 پایین‌گذر 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 با $Z(1:N)$ نشان داده شده است). حال تبدیل اصلی روی این بردار اعمال می‌گردد. این تبدیل می‌تواند تبدیل فوریه و یا موجک و یا هر تبدیل دیگری باشد مثلاً در بانک فیلتر 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 بانک فیلتر 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. برای اینکه در این پیاده‌سازی حافظه کمتری مصرف شود تنها پنجره آنالیز ذخیره می‌شود. پنجره سنتز نیز از روی پنجره آنالیز بدست می‌آید و در حقیقت نسخه کوچک شده از پنجره آنالیز می‌باشد.

مطابق شکل ۳ ساختار درخت از ریشه شروع به گسترش می‌کند تا درخت موجک مبتنی بر بانک فیلتر Mel ساخته شود. درخت مبتنی بر بانک فیلتر Mel پیشنهادی دارای ۱۵ زیربانده می‌باشد که درخت موجک آن حداکثر تا عمق ۵ گسترش یافته است. پهنای فرکانسی هر بانده این درخت در جدول ۲ نشان داده شده است. از آنجا که در گسترش درخت موجک از تجزیه دوتایی استفاده شده است لذا درخت حاصل تقریبی از فیلتر بانک Mel می‌باشد.



شکل (۳): درخت موجک بر اساس فیلتر بانک Mel (۱۵ زیربانده)

جدول (۲): گستره باندهای بحرانی در درخت موجک مبتنی بر فیلتر

بانک Mel			
محدوده فرکانسی (Hz)	بانده بحرانی	محدوده فرکانسی (Hz)	بانده بحرانی
1000~1250	C9	0~125	C1
1250~1500	C10	125~250	C2
1500~1750	C11	250~375	C3
1750~2000	C12	375~500	C4
2000~2500	C13	500~625	C5
2500~3000	C14	625~750	C6
3000~4000	C15	750~875	C7
		875~1000	C8

رابطه ۹ تعداد محاسبات جمع و ضرب را برای بانک فیلتر موجک مبتنی بر Mel نشان می‌دهد.

$$\text{تعداد ضرب} = 4.25(m)(N) + 24.5(m)(m-1) \quad (9)$$

$$\text{تعداد جمع} = 4.25(m-1)(N) + 24.5(m-1)^2$$

۴- پیشنهاد بکارگیری WOLA در سیستم کاهش نویز برای کاهش پیچیدگی محاسباتی

۴-۱- بانک فیلتر WOLA

سیستم‌هایی که قادر به بازسازی کامل سیگنال هستند محدودیت‌های 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 پیاده‌سازی بانک فیلتر DFT طراحی شد. [۹] مدل دیگری که برای این بانک فیلتر طراحی شد بانک فیلتر چند مرحله‌ای است. این دو ساختار از لحاظ محاسباتی یکسان بوده و تنها در نحوه تحلیل داده‌ها متفاوت می‌باشند.

شکل ۴ بلوک دیگرام بانک فیلتر آنالیز و سنتز WOLA را نشان می‌دهد. [۱۰] که L طول فریم اولیه، N طول فریم ثانویه و R میزان

قسمت بندی (Segmentation) وابستگی دارند و به صورت زیر تعریف می شوند.

$$OS - Factor = \frac{M}{R} \quad (10)$$

دو پارامتر فوق بر کیفیت سیگنال خروجی و میزان محاسبات تأثیر می‌گذارند. با افزایش Seg-factor کیفیت سیگنال حاصل کاهش می‌یابد و با افزایش OS-factor میزان محاسبات افزایش می‌یابد. باید یادآوری کنیم که در پیاده‌سازی این بانک فیلتر همپوشانی پنجره‌های متوالی حداقل باید ۷۵٪ باشد. این مقدار به این دلیل انتخاب می‌شود که بتوان سیگنال گفتار را مورد بازسازی کامل قرار داد.

۴-۳- میزان محاسبات با بکارگیری WOLA در بانک فیلتر موجک

با استفاده از ایده WOLA در بانک فیلتر موجک طول فریم ورودی به قسمت آنالیز کاهش می‌یابد. به این ترتیب اگر طول فریم اولیه برابر N و طول فریم ثانویه برابر M و میزان پیشروی پنجره نمونه‌برداری برابر R باشد برای اینکه شرایط بازسازی کامل با وجود WOLA فراهم گردد، مقدار پارامتر R حداکثر می‌تواند نصف M باشد. به این ترتیب در بانک فیلتری که از ایده WOLA استفاده می‌کند میزان همپوشانی پنجره‌های متوالی حداقل باید ۷۵٪ باشد. بنابراین می‌توان گفت تعداد محاسبات جمع برابر با رابطه ۱۱ و تعداد محاسبات ضرب برابر با رابطه ۱۲ می‌باشد.

$$2N + 2[L(m-1)M + (m-1)^2 \cdot (2^{L+1} - 2 - L)] \quad (11)$$

$$2N + 2[L \cdot m \cdot M + (m-1)^2 \cdot (2^{L+1} - 2 - L)] \quad (12)$$

۵- پیاده‌سازی و آزمایشات

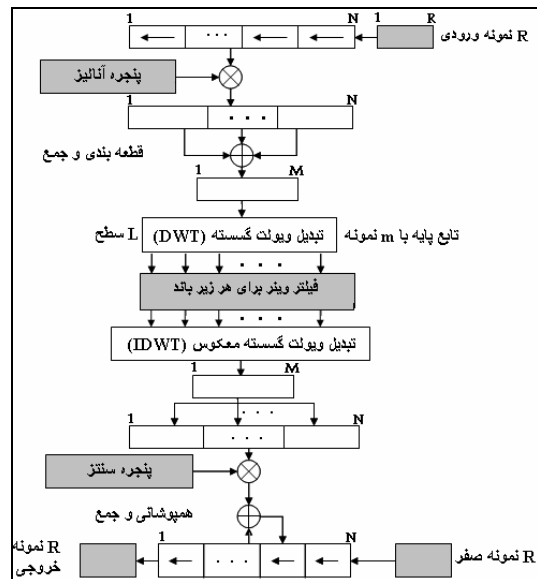
در بررسی دو پارامتر OS-factor و Seg-factor در سیستم پیشنهادی بهبود کیفیت مبتنی بر موجک با WOLA با در نظر گرفتن نتایج ارزیابی نظری و عملی، مقدار ۲ برای دو پارامتر فوق بدست آمد. در ادامه به مقایسه سیستم بهبود کیفیت مبتنی بر موجک بدون WOLA و روش پیشنهادی بهبود کیفیت مبتنی بر موجک با WOLA می‌پردازیم، هر دو سیستم دارای درخت تجزیه موجک بر مبنای Mel می‌باشند. میزان محاسبات (تعداد عمل جمع و ضرب) بانک فیلتر موجک با ساختار درخت Mel با استفاده از رابطه ۹ بدست آید. در حالیکه میزان محاسبات همین بانک فیلتر در صورت بکارگیری WOLA از رابطه ۱۱ و ۱۲ بدست می‌آید. به عنوان مثال اگر طول فریم ۲۵۶ نمونه باشد و نوع تابع پایه db5 باشد، در تشکیل درخت موجک بر مبنای Mel

اگر مقادیر پارامترهای R و N برابر باشند بانک فیلتر از نوع نمونه برداری بحرانی خواهد بود، که در آن فرکانس قطع پنجره آنالیز برابر $2\pi/N$ خواهد شد.

در سیستم‌هایی که از بانک فیلترهای آنالیز و سنتز برای تحلیل و بازسازی کامل سیگنال استفاده می‌کنند، بیشترین بار محاسباتی مربوط به بانک فیلترهای آنالیز و سنتز می‌باشد. [۱۱] پیچیدگی محاسباتی در روشهای معمول در بانک فیلترها، از درجه L می‌باشد، که L طول یک فریم از سیگنال ورودی است. اما در بانک فیلتر WOLA این تبدیل حداکثر L/N مرتبه، بار محاسباتی را کمتر می‌کند. [۳]

۴-۲- سیستم بهبود گفتار با بکارگیری WOLA

در شکل ۵ یک سیستم بهبود گفتار که از WOLA برای کاهش طول سیگنال ورودی به بانک فیلتر آنالیز استفاده می‌شود نشان داده شده است.



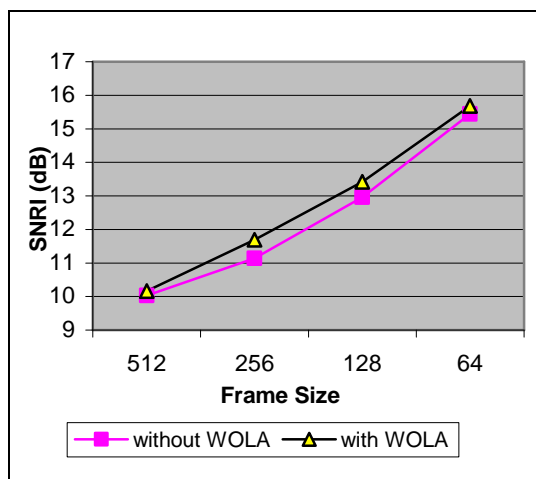
شکل (۵): سیستم بهبود گفتار بر اساس بانک فیلتر موجک

با بکارگیری WOLA

در این سیستم ابتدا سیگنال ورودی در قسمت آنالیز به چند قسمت تقسیم می‌شود و با جمع این قسمت‌ها فریمی حاصل می‌شود که توسط درخت موجک به زیرباندهایی تقسیم می‌شود. پس از بکارگیری فیلتر وینر توسط قسمت سنتز نمونه‌های بهبود یافته سیگنال گفتار بازیابی می‌شوند.

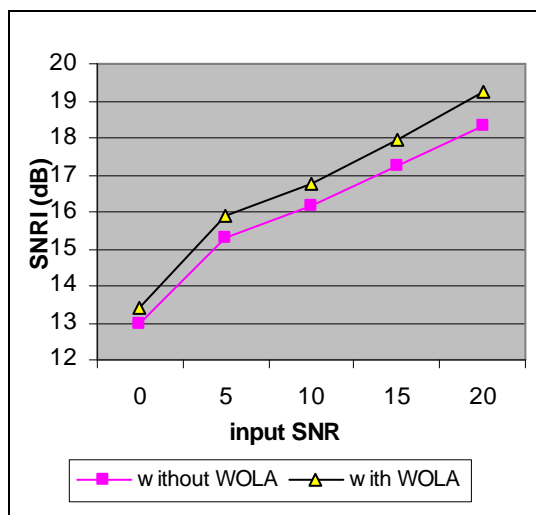
بانک فیلتر WOLA همواره از نوع فرامونه‌برداری^{††} شده خواهد بود. پارامتر فرامونه‌برداری برابر نسبت M/R است که M طول تبدیل زمان-فرکانس و R فاکتور کاهش نرخ یا میزان پیشروی پنجره‌های متوالی می‌باشد. در این بانک فیلتر نیز استفاده از ایده WOLA با دو پارامتر اساسی همراه است. این دو پارامتر OS-factor و Seg-factor می‌باشند که به ترتیب به فرامونه برداری (OverSampling) و

^{††} Oversampled filterbank



شکل (۷): نتایج SNRI برای دو روش حذف نویز در حضور نویز سفید 0dB

شکل ۸ نتایج SNRI را برای دو روش فوق با طول فریم ۱۲۸ در حضور نویز سفید با SNR ورودی مختلف نشان می‌دهد. در این نمودار محور افقی SNR ورودی و محور عمودی اندازه SNRI را برحسب dB نشان می‌دهد. چنانچه دیده می‌شود اندازه SNRI برای روش پیشنهادی نسبت به روش معمولی در اینجا نیز حدود 1dB بیشتر می‌باشد. نتایج مشابهی نیز برای طول فریم‌های دیگر نیز بدست آمد.

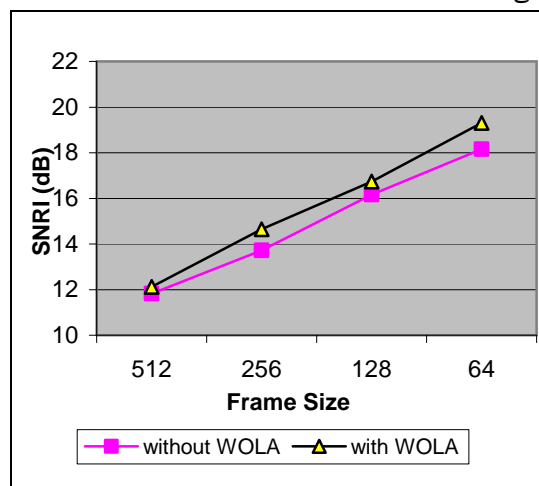


شکل (۸): نتایج SNRI برای دو روش حذف نویز با طول فریم ۱۲۸ در حضور نویز سفید با SNR ورودی مختلف

همچنین در ارزیابی شنوایی مشخص شد که تخریب صدا در روش پیشنهادی کمتر شده است و این روش دارای اعوجاج کمتری است. و کیفیت سیگنال‌های بهبود یافته بروش حذف نویز پیشنهادی یعنی بکارگیری WOLA در حد کیفیت سیگنال‌های بهبود یافته بروش حذف نویز معمولی می‌باشند.

بدون بکار بردن WOLA با توجه به رابطه ۹ تعداد محاسبات برای یک فریم ۳۹۶ عمل و در حالت بکارگیری WOLA با توجه به رابطه ۱۱ و ۱۲ تعداد محاسبات برای ۲۴۲ عمل خواهد بود، که میزان محاسبات به اندازه ۳۸٫۷٪ کاهش یافته است. با توجه به روابط فوق می‌توان گفت که پیچیدگی محاسباتی روش پیشنهادی کمتر از روش معمولی است و اندازه این کاهش به طول فریم بستگی دارد، یعنی در فریم بزرگ کاهش محاسبات بیشتری مشاهده می‌شود. که این کاهش محاسبات جمع و ضرب برای دو حالت از بکارگیری درخت Mel یکی حالت بدون WOLA با همپوشانی 75٪ نسبت به حالت بکارگیری WOLA با پارامترهای $Os-factor = seg-factor = 2$ بین ۲۶٪ تا ۴۲٪ می‌باشد.

با توجه به مرجع [۱] مقدار طول فریم برابر ۶۴، ۱۲۸، ۲۵۶ و ۵۱۲، همپوشانی را ۷۵٪ در نظر می‌گیریم. برای تجزیه موجک از درخت موجک مبتنی بر Mel استفاده می‌کنیم. برای ارزیابی سیستم حذف نویز از عبارتهای بیان شده توسط گویندگان مرد و زن که با نویزهای سفید، ماشین و همهمه با نرخ سیگنال به نویزهای 0dB، 10dB و 20dB آلوده شده اند، استفاده می‌نمائیم. در ارزیابی از معیار بهبود SNR یا SNRI استفاده شده است، این معیار در مرجع [۱۶] بطور کامل تشریح شده است. شکل ۶ و ۷ نتایج SNRI را برای دو روش فوق در حضور نویز سفید 10dB نشان می‌دهند. در این نمودارها محور افقی طول فریم و محور عمودی اندازه SNRI را برحسب dB نشان می‌دهد.



شکل (۶): نتایج SNRI برای دو روش حذف نویز در حضور نویز سفید 10dB

چنانچه دیده می‌شود اندازه SNRI برای روش پیشنهادی نسبت به روش معمولی حدود ۱ تا 1.5dB بیشتر می‌باشد. نتایج مشابهی نیز برای نویزهای دیگر با نرخ سیگنال به نویز مختلف بدست آمد.

- [10] Crochiere R.E. and Rabiner L.R., *Multirate Digital Signal Processing*, Prentice-Hall, 1983.
- [11] Bont F., Groenewegen M., and Oomen W., *A High Quality Audio-Coding System at 128 kb/s*, Audio Engineering Society Conference, Feb 1995.
- [12] Donoho, D.L., *De-noising by soft-thresholding*. IEEE Trans. Inform. Theory 41 (3), 613–627, 1995.
- [13] Ing Y.S., Soo N.K., Chai K.Y., *Wavelet for speech denoising*. IEEE. TENCON '97 vol.2, pp. 479 – 482, 1997.
- [14] Sheikhzadeh, H., Abutalebi, H.R., *An improved wavelet based speech enhancement system*, In: Proc. European Conf. on Speech Communication and Technology, pp. 1855–1858, 2001.
- [15] Mallat S., *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
- [16] Paajanen E., Ayad B., Mattila V., *New Objective Measures for Characterization of Noise Suppression Algorithms*, IEEE Workshop on Speech Coding, pp. 23-25, 2000.

۶- نتیجه گیری

در این مقاله روشی برای حذف نویز با استفاده از فیلتر وینر و تبدیل موجک پیشنهاد گردید. استفاده از فیلتر بانک Mel می تواند در تعداد زیرباند کمتر، به کیفیت بالاتری در سیگنالهای صوتی بهبودیافته نسبت به توزیع زیرباندها بطور یکنواخت دست یافت. در این مقاله به تشریح کامل پیچیدگی محاسباتی آن پرداخته شد و برای کاهش این پیچیدگی روشی پیشنهاد شد که در آن از WOLA در سیستم کاهش نویز بر مبنای موجک استفاده شد. همانطور که دیده شد فیلتر بانک WOLA دارای دو پارامتر OS-factor و Seg-factor می باشد که با افزایش Seg-factor کیفیت سیگنال حاصل کاهش می یابد و با افزایش فاکتور OS-factor میزان محاسبات افزایش می یابد. از نتایج حاصل از بکارگیری WOLA در سیستم کاهش نویز مبتنی بر موجک چنین نتیجه گیری می شود که بهترین مقدار برای این دو پارامتر ۲ می باشد. همچنین بکارگیری WOLA در سیستم کاهش نویز بر اساس درخت Mel ضمن کاهش حجم محاسبات انجام شده کیفیت سیگنال گفتار را نیز در حد خوبی حفظ کند. درصد کاهش محاسبات با توجه به طول فریم ورودی سیستم کاهش نویز حداقل ۲۶٪ و حداکثر ۴۲٪ می باشد.

مراجع

- [۱] بهزاد زمانی، احمد اکبری، بابک ناصر شریف، بررسی اثر باندهای بحرانی و پارامترهای بکارگیری تبدیل موجک در بهبود گفتار، یازدهمین کنفرانس بین‌المللی کامپیوتر ایران، ۱۳۸۴.
- [2] Huang A., Acero A., Hon H., *Spoken language processing*. Prentice Hall, 2001.
- [3] Ahadi S.M., Sheikhzadeh H., Brennan R.L., Freeman G.H., *An Efficient Front-End For Automatic Speech Recognition*, international conference on electronics, circuits and systems (ICECS2003).
- [4] Herman D., Brennan R.L., Sheikhzadeh H., Cornu E., *Low Power Implementation of The Bluetooth Subband Audio Codec*, Proceeding of the 2004 international Conference on Acoustics Speech and Signal Processing (ICASSP 2004).
- [5] Lu C.T., Wang H.C., *Enhancement of single channel speech based on masking property and wavelet transform*. *Speech Communication*, Volume 41, Issues 2-3, pp. 409-427, 2003.
- [6] Grossman A., Martinet R., Morlet J., *Reading and understanding continuous wavelet transform*. Springer-Verlag, pp. 2-20, 1989.
- [7] Cohen I., *Enhancement of speech using bark-scaled wavelet packet decomposition*. European Conf. on Speech Communication and Technology, 2001.
- [8] 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.
- [9] Crochiere R. E., *A Weighted Overlapp-Add Method of Short-Time Fourier Analysis/Synthesis*, IEEE Transaction on Acoustics, Speech, Signal Processing, vol. ASSP-28, pp. 98-102, Feb 1980.