

# آستانه گذاری وفقی ضرایب موجک برای پاک سازی سیگنال گفتار نویزی

فاطمه شیخ علیشاهی، حمیدرضا ابوطالبی و محمدرضا تابان

وقفه<sup>4</sup> با استفاده از محاسبه آنتروپی هیستوگرام<sup>5</sup> ضرایب موجک با SNR کم ارائه دادند. Lu و Wang [6] تا [7] با بهره گیری از خاصیت آستانه ماسک شونده نویز<sup>6</sup> (NMT) در سیگنال گفتار الگوریتم Donoho را اصلاح کردند. Johnson و همکارانش [11] نیز با در نظر گرفتن مدل جدیدی برای ساختار شنوایی انسان، کارایی الگوریتم آستانه گذاری وفقی ضرایب موجک در SNR کم گزارش داده اند.

در تحقیق حاضر، برخلاف کارهای قبلی که از یک تابع آستانه گذاری سخت<sup>7</sup> یا نرم<sup>8</sup> استفاده می شد، تابعی بینابینی برای آستانه گذاری ضرایب بررسی شده است؛ در واقع حد آستانه و تابع آستانه گذاری به گونه ای تنظیم می شود که زیرباندهای موجکی که دارای انرژی گفتاری با محتوای بسیار زیاد هستند (مانند مؤلفه های فرکانس پایین واکه ها<sup>9</sup>)، از حد آستانه کوچک تر و تابع آستانه گذاری سخت پیوسته استفاده کنند و زیرباندهای با محتوای ناچیز از انرژی گفتار (مانند نواحی سکوت<sup>10</sup> و مؤلفه های فرکانس بالای واکه ها) از حد آستانه بزرگ تر و تابع آستانه گذاری نرم استفاده کنند. در نواحی با وضعیت بینابین دو حالت فوق، تابع آستانه گذاری به صورت وفقی و متناسب با توان سیگنال و نویز در آن زیرباند، تابعی مابین دو وضعیت حدی آستانه گذاری سخت و آستانه گذاری نرم می باشد.

در راستای انجام این تحقیق، برای بررسی تأثیر نحوه رشد درخت موجک (یعنی موجک معمولی، بسته موجک<sup>11</sup> کامل و یا بسته موجک متناسب با باندهای بحرانی<sup>12</sup>) و نیز بررسی تأثیر نوع موجک در عملکرد سیستم بهسازی گفتار، آزمایش های متعددی انجام و نتایج حاصل از آن ارائه و در مراحل بعدی کار مورد استفاده قرار گرفته است.

ساختار مقاله به شرح زیر می باشد: بخش 2 به مروری بر تبدیل موجک معمولی و بسته موجک می پردازد. اصول اولیه حذف نیز در حوزه موجک در بخش 3 معرفی می شود. شرح جزئیات روش بهبود یافته در بخش های 4 و 5 بیان می شود. در بخش 6 نتایج شبیه سازی الگوریتم پیشنهادی در مقایسه با دو روش بهسازی گفتار (ارائه شده در [7] و [9]) نشان داده می شوند. در پایان، جمع بندی نتایج این تحقیق در بخش 7 ارائه می شود.

## 2- تبدیل موجک و بسته موجک

ایجاد تبدیل موجک گسسته با تکرار بانک فیلتر بر روی شاخه تابع مقیاس بندی پایین گذر امکان پذیر است [12]. به این روش، الگوریتم

چکیده: این مقاله به مبحث بهسازی گفتار در حوزه موجک می پردازد. در روش پیشنهادی، بعد از تجزیه سیگنال نویزی به باندهای موجک تابع آستانه گذاری وفقی روی ضرایب موجک اعمال می شود. در زیرباندهایی که دارای انرژی گفتار با محتوای بسیار زیاد هستند، از حد آستانه کوچک تر و تابع آستانه گذاری سخت استفاده می شود و برعکس، در زیرباندهای با محتوای ناچیز از انرژی گفتار، حد آستانه بزرگ تر و تابع آستانه گذاری نرم مورد استفاده واقع می شود. در نواحی با وضعیت بینابین دو حالت فوق، تابع آستانه گذاری به صورت وفقی و مابین دو وضعیت حدی آستانه گذاری سخت و آستانه گذاری نرم تعیین می شود. پارامتری که تابع آستانه گذاری و حد آستانه را در هر زیرباند موجک تعیین می کند با نسبت توان گفتار و نویز در هر زیرباند رابطه دارد. آزمایش های انجام شده در مقایسه با روش های قبلی نشان می دهد که با اعمال این تکنیک، نویز به نحو مطلوبی حذف شده و میزان اعوجاج در گفتار خروجی کاهش می یابد. علاوه بر این، نتایج شبیه سازی حکایت از آن دارد که افزایش رشد درخت موجک در بهبود خروجی سیستم بهسازی تأثیر داشته و نوع موجک مناسب، وابسته به نوع نویز موجود در محیط می باشد.

کلید واژه: بهسازی گفتار، تبدیل موجک، آستانه گذاری وفقی.

## 1- مقدمه

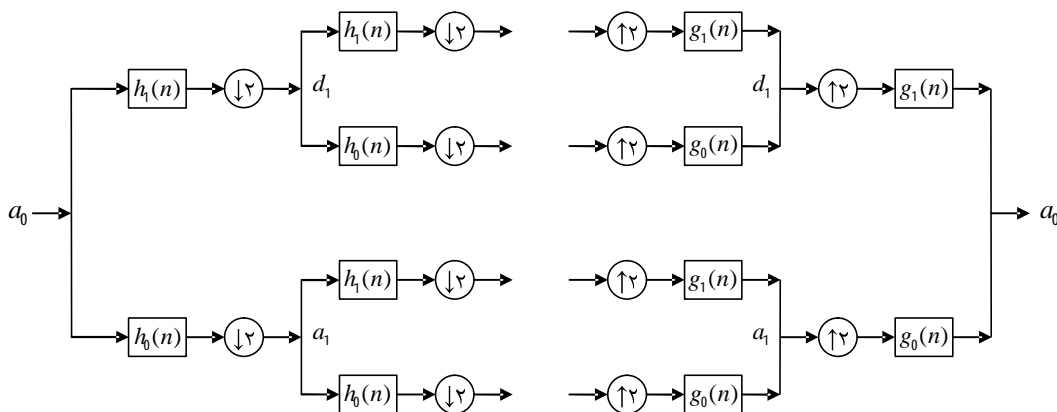
با رشد سریع سیستم های مخابرات صوت، بهسازی گفتار<sup>1</sup> نویزی یک موضوع مهم و فعال تحقیقاتی شده است. در طی دهه های گذشته، روش های تک کاناله و چندکاناله متنوعی برای بهسازی گفتار ارائه شده اند. از میان تکنیک های تک کاناله، آستانه گذاری<sup>2</sup> ضرایب موجک گفتار نویزی یکی از روش های مطرح در این زمینه است. این روش بر پایه تئوری حذف نویز در حوزه موجک که در سال 1995 توسط Donoho [1] ارائه گردید، بنا نهاده شد. چندین محقق از الگوریتم تئوری Donoho در حذف نویز گفتار بهره گرفته [2] تا [11] و با استفاده از ویژگی های گفتار، گفتار نویزی را بهسازی نمودند. Rouat و Bahoura [2] با بهره گیری از عملگر انرژی تیگر<sup>3</sup> در زیرباندهای موجک، حد آستانه وفقی متغیر با زمان و فرکانس برای بهبود کیفیت سیگنال گفتار نویزی ارائه کردند. Chang و همکارانش [5] الگوریتم مؤثری برای تخمین نویز (بدون تشخیص نواحی

این مقاله در تاریخ 6 شهریور ماه 1386 دریافت و در تاریخ 14 تیر ماه 1387 بازنگری شد.

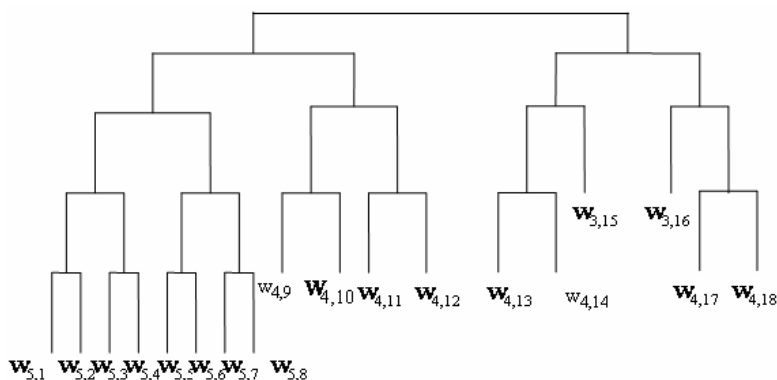
فاطمه شیخ علیشاهی، آزمایشگاه تحقیقاتی پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، یزد (email: fatemalishahi@yahoo.com).  
حمیدرضا ابوطالبی، آزمایشگاه تحقیقاتی پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، یزد (email: habutalebi@yazduni.ac.ir).  
محمدرضا تابان، آزمایشگاه تحقیقاتی پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، یزد (email: mrtaban@yazduni.ac.ir).

1. Speech Enhancement
2. Thresholding
3. Teager

4. Pause
5. Entropy of Histogram
6. Noise Masking Threshold
7. Hard
8. Soft
9. Vowels
10. Silence
11. Wavelet Packet
12. Critical Bands



شکل 1: درخت باینری کامل برای تبدیل کامل بسته موجک دومقیاسه.



شکل 2: درخت تبدیل بسته موجک بر اساس باندهای بحرانی (PWPDP).

یک محیط نویزی را دارند. پاسخ فرکانسی ساختار شنوایی انسان بر اساس میزان درک مؤلفه‌های فرکانسی صوت به باندهای بحرانی تقسیم می‌شود. در بسیاری از پدیده‌های شنیداری از قبیل درک صدا، گام صدا و ... باندهای بحرانی از اهمیت زیادی برخوردار هستند [13]. یک معیار برای تعیین باندهای بحرانی، بارک<sup>2</sup> است. رابطه میان هر بارک  $z$  برحسب فرکانس  $f$  به‌طور تقریبی برابر است با [13]

$$z(f) = 13 \arctan(7,6 \times 10^{-4} f) + 3,5 \arctan(1,33 \times 10^{-4} f)^2 \text{ [Bark]} \quad (3)$$

که  $f$  فرکانس برحسب هرتز است. پهنای باند بحرانی متناظر با فرکانس مرکزی فیلترها را می‌توان با رابطه زیر بیان نمود [11]

$$CBW(f_c) = 25 + 75 \times (1 + 1,4 \times 10^{-6} f_c^2)^{0,69} \text{ [Hz]} \quad (4)$$

که  $f_c$  فرکانس مرکزی بر حسب هرتز است. از دید تئوری، محدوده فرکانس شنوایی انسان در 20 تا 20000 هرتز قرار دارد و تقریباً 25 بارک را در برمی‌گیرد. چون مقیاس بارک تابع غیر خطی از فرکانس  $f$  می‌باشد، اولین قدم برای ساختن تبدیل بسته موجک تعیین پهنای باند سیگنال گفتار - به‌منظور تنظیم تعداد بارک معتبر - است. اگر پهنای باند سیگنال گفتار 4 kHz باشد، محدوده فرکانسی گفتار با 18 باند بحرانی تقریب زده می‌شود. ساختار درختی PWPDP<sup>3</sup> متناظر با این باندهای بحرانی به‌فرم شکل 2 درمی‌آید. پهنای باندهای متناظر با درخت PWPDP نیز در شکل 3 نشان داده شده است. با استفاده از این PWPDP، سیگنال گفتار به 18 باند بحرانی تجزیه می‌شود.

Mallat گفته می‌شود. این الگوریتم با یک دقت فرکانسی لگاریتمی برای پردازش اکثر سیگنال‌ها (نه همه آنها) عملکرد خوبی دارد. اما پهنای باند زیرباندهای متناظر با این نوع تبدیل موجک نمی‌تواند به‌خوبی ساختار شنوایی انسان را مدل کند. به‌منظور غلبه بر عیب فوق، می‌توان شاخه موجک بالاگذر درخت الگوریتم Mallat را همانند شاخه تابع مقیاس‌بندی پایین‌گذر با به‌کاربردن بانک فیلتر تکرار نمود. این نوع تجزیه موجک، تجزیه بسته موجک نامیده می‌شود. به‌عنوان مثال، درخت باینری کامل برای تجزیه و بازسازی بسته موجک دو مقیاسه<sup>1</sup> - که به‌معنای به‌کاربردن دو فیلتر در هر مرحله تجزیه می‌باشد - در شکل 1 نشان داده شده است. در شکل 1،  $h_1(n)$  و  $h_0(n)$  به‌ترتیب فیلتر مقیاس‌بندی پایین‌گذر و فیلتر موجک بالاگذر آنالیز هستند، در حالی که  $g_1(n)$  و  $g_0(n)$  به‌ترتیب فیلتر مقیاس‌بندی پایین‌گذر و فیلتر موجک بالاگذر سنتز هستند. برای امکان بازسازی کامل سیگنال، این چهار فیلتر چنین روابطی با هم دارند [12]

$$h_1(n) = (-1)^n g_0(1-n) \quad (1)$$

$$g_1(n) = (-1)^n h_0(1-n) \quad (2)$$

در شکل 1 سمبل‌های  $\downarrow 2$  و  $\uparrow 2$  به‌ترتیب به عمل کاهش نرخ نمونه‌برداری با فاکتور 2 و افزایش نرخ نمونه‌برداری با فاکتور 2 اشاره دارند.

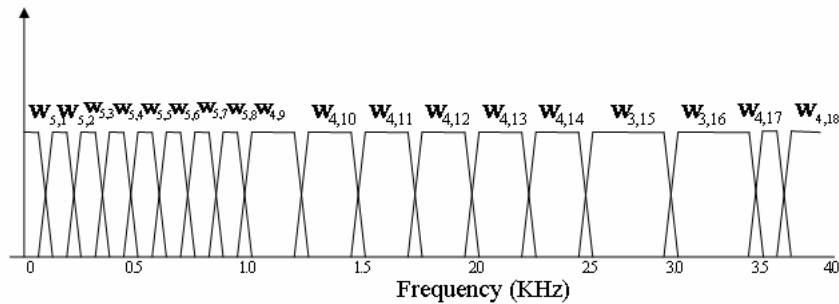
## 2-1 تجزیه بسته موجک براساس دقت شنوایی انسان در محدوده‌های فرکانسی مختلف

انسان‌ها بدون اطلاع قبلی از نویز، توانایی آشکارسازی گفتار مطلوب در

2. Bark

3. Perceptual Wavelet Packet Decomposition

1. Two-Scale



شکل 3: پهنای باندهای فرکانسی برای درخت PWPD.

استفاده شده است [14]. این الگوریتم حداقل توان نویز را از روی توان سیگنال گفتار نویزی در هر زیرباند موجک دنبال می‌کند. برای رسیدن به تخمین واقعی توان نویز، توان سیگنال نویزی با استفاده از یک سیستم مرتبه اول بازگشتی تقریب زده می‌شود [14]. ضرایب بهینه سیستم مرتبه اول بازگشتی با حداقل شدن میانگین مجذور خطای تخمین توان نویز در هر زیرباند موجک تعیین می‌شود. به همین دلیل تکنیک تخمین توان نویز عملکرد خوبی در دنبال نمودن تغییرات نویز غیرایستاد دارد.

### 5- تخمین وقتی حد آستانه و تابع آستانه گذاری

در تحقیق حاضر، از ویژگی‌های سیگنال گفتار و میزان درک صدا توسط گوش انسان، برای تخمین حد آستانه و تابع آستانه گذاری بهره گرفته شده است. گوش انسان در جاهایی که سیگنال گفتار قوی‌تر از نویز باشد، حساسیت کمتری دارد. بنابراین برای کاهش نویز گفتار پیشنهاد می‌کنیم که در زیرباندهای موجک که سیگنال گفتار دامنه بسیار زیاد دارد و نویز کمتر شنیده می‌شود (مانند مؤلفه‌های فرکانس پایین واکه‌ها)، ضرایب موجک آن زیرباند تقریباً دست‌نخورده باقی بماند. بدین منظور در این زیرباندها، از حد آستانه کوچک و تابع آستانه گذاری سخت پیوسته استفاده می‌شود. در زیرباندهای دیگر که سیگنال گفتار مؤلفه ناچیز دارد، حد آستانه بزرگ‌تر و تابع آستانه گذاری نرم استفاده می‌شود. زیرا در این زیرباندها که قسمت اعظم انرژی مربوط به نویز است، باید مؤلفه‌های بیشتری تضعیف شوند. در خیلی موارد، زیرباندهای موجک دارای مؤلفه‌های گفتاری با دامنه متوسط هستند. در این موارد، به آستانه گذاری نه نرم و نه سخت (یا به بیان دیگر، حالت بینابینی از نرم و سخت) نیاز داریم. با توجه به توضیحات داده شده معلوم است که به یک تابع آستانه گذاری احتیاج داریم که با تنظیم وقتی پارامترهای آن تمام موارد فوق برآورده شود. با جمع‌بندی موارد فوق، از تابع آستانه گذاری وقتی زیر استفاده می‌شود [15]

$$f_c(Y) = \begin{cases} Y - \text{sgn}(Y)(1-\beta)\lambda & \text{if } |Y| \geq \lambda \\ 0 & \text{if } |Y| \leq \gamma \\ \beta \left( \frac{|Y| - \gamma}{\lambda - \gamma} \right)^2 \left\{ (\beta - 3) \left( \frac{|Y| - \gamma}{\lambda - \gamma} \right) + 4 - \beta \right\} Y & \text{otherwise} \end{cases} \quad (10)$$

که در آن  $0 < \gamma < \lambda$  و  $0 \leq \beta \leq 1$  می‌باشد. شکل 4 تابع  $f_c(Y)$  را برای مقادیر مختلف  $\beta$  که تعیین‌کننده نوع تابع آستانه گذاری است - به‌ازای  $\lambda = 1$  و  $\gamma = 0.0001$  و  $\beta$  به ترتیب حد آستانه بالایی و پایینی هستند) تشریح می‌کند.

در مقاله [12]، [13] برای حذف نویز تصاویر مختلف به کار برده شد و  $\beta$  مناسب برای هر تصویر با آزمایش و خطا به دست آمد؛ ولی در تحقیق

### 3- آستانه گذاری موجک

انرژی سیگنالی مانند گفتار - که طبیعت پایین‌گذر دارد - در تعداد کمی از ابعاد موجک متمرکز شده است؛ در حالی که انرژی نویز (سفید) در همه ابعاد موجک توزیع می‌شود. بنابراین با صفر قراردادن ضرایب با دامنه کوچک‌تر از یک حد آستانه، به نحو مطلوبی می‌توان نویز را حذف نمود، در حالی که اطلاعات مهم سیگنال اصلی حفظ می‌شود. فرض کنید  $y$  دنباله مشاهده نویزی سیگنال  $x$  (با طول محدود) تخریب‌شده با نویز سفید گوسی  $n$  با متوسط صفر و واریانس  $\sigma^2$  است

$$y = x + n \quad (5)$$

در اینجا هدف بازسازی سیگنال  $x$  از مشاهدات نویزی  $y$  است. اگر  $W$  ماتریس تبدیل موجک گسسته باشد، (5) در حوزه موجک به صورت زیر نوشته می‌شود

$$Y = X + N \quad (6)$$

که در آن

$$Y = Wy, \quad X = Wx, \quad N = Wn \quad (7)$$

فرض کنید  $X_{\text{est}}$  تخمینی از سیگنال تمیز  $X$  بر اساس مشاهدات نویزی  $Y$  در حوزه موجک باشد. سیگنال تمیز  $x$  از رابطه زیر محاسبه می‌شود

$$x \cong W^{-1} X_{\text{est}} \cong W^{-1} Y_{\text{thr}} \quad (8)$$

که  $Y_{\text{thr}}$  به ضرایب موجک بعد از آستانه گذاری اشاره می‌کند. مقدار مناسب حد آستانه را به چندین روش می‌توان تعیین نمود. Donoho فرمول زیر را برای مقدار حد آستانه پیشنهاد داد [1]

$$T = \sigma \sqrt{2 \log(L)} \quad (9)$$

که در آن  $T$  مقدار حد آستانه و  $L$  طول سیگنال نویزی  $y$  است. معمول‌ترین توابع آستانه گذاری، آستانه گذاری سخت و آستانه گذاری نرم هستند. اگرچه هر دوی اینها کارایی قابل قبولی در برخی از کاربردها دارند، در مواردی کیفیت تخمین با استفاده از تابع آستانه گذاری وقتی بهبود می‌یابد [13].

### 4- تخمین توان نویز

تخمین‌گر نویز اثر مهمی روی کیفیت کلی سیستم بهسازی گفتار دارد. اگر تخمین توان نویز از توان واقعی نویز کمتر باشد، نویز باقیمانده شنیده می‌شود؛ در مقابل اگر تخمین توان نویز از توان واقعی نویز زیادتر باشد، سیگنال گفتار دچار اعوجاج شده و قابلیت درک گفتار کاهش می‌یابد. در این تحقیق، از یک الگوریتم تخمین توان نویز بدون بهره‌گیری از  $VAD^1$

1. Voice Activity Detector

که  $\beta_j^i(m)$  تعیین‌کننده تابع آستانه‌گذاری در زیرباند  $i$ -ام، سطح  $j$ -ام و فریم  $m$ -ام می‌باشد. به عبارت دیگر هنگامی که  $\alpha$  کوچک شود،  $\beta$  بزرگ می‌شود؛ در نتیجه، حد آستانه در (11) کوچک‌تر و تابع آستانه‌گذاری در (10) از نوع سخت خواهد بود. بالعکس، اگر  $\alpha$  بزرگ شود،  $\beta$  کوچک می‌شود؛ بنابراین حد آستانه بزرگ‌تر شده و تابع آستانه‌گذاری از نوع نرم انتخاب می‌شود.

## 6- ارزیابی کارایی سیستم پیشنهادی

برای انجام شبیه‌سازی‌ها، سیگنال‌های گفتار از دادگان FARSDAT [16] شامل دو گوینده زن و دو گوینده مرد انتخاب شده‌اند. از هر چهار گوینده، ده جمله در نظر گرفته شده که در مجموع 40 جمله برای ارزیابی سیستم پیشنهادی وجود دارد. سیگنال‌های گفتار نویزی از جمع سیگنال‌های گفتار تمیز با نویزهای سفید، صورتی و کارخانه حاصل شده‌اند. این نویزها از دادگان NoiseX [17] استخراج شده‌اند. برای ارزیابی جامع کارایی سیستم بهسازی، سه سطح SNR ورودی برابر با 0 dB، 5 dB و 10 dB استفاده می‌شود. فرکانس نمونه‌برداری سیگنال گفتار نویزی 8 kHz در نظر گرفته شده است. در پیاده‌سازی روش، از پنجره‌های همپینگ با همپوشانی 50% استفاده شده که طول هر پنجره 32 ms (معادل 256 نمونه) است. در ادامه پس از معرفی معیارهای اندازه‌گیری کیفیت سیگنال خروجی، نخست الگوریتم پیشنهادی تحت شرایط مختلف شامل روند رشد درخت موجک و انواع خانواده موجک در نویزهای مذکور آزمایش شده و در ادامه، با توجه به نتایج و پارامترهای به‌دست آمده از این آزمایش‌ها، روش اصلاح‌شده با دیگر روش‌های بهسازی مبتنی بر آستانه‌گذاری موجک مقایسه می‌شود.

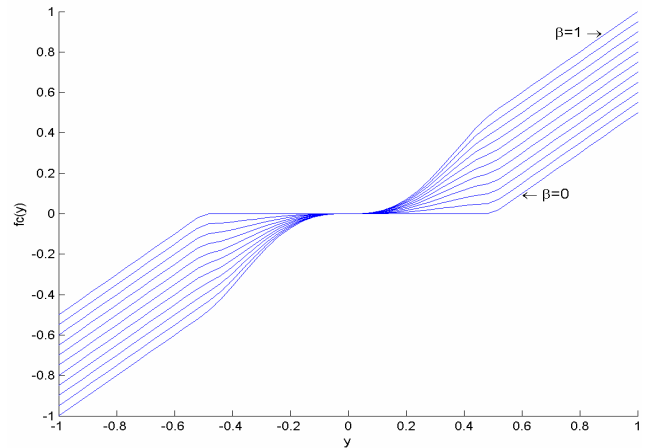
### 6-1 معیارهای اندازه‌گیری کیفیت سیگنال خروجی

اگرچه آزمایش‌های کیفی (و شخص‌گرا) معیارهای مناسب‌تری برای ارزیابی الگوریتم بهسازی می‌باشند، ولی این آزمون‌ها وقت‌گیر و هزینه‌بر هستند. از طرفی اگر تعداد پارامترهای مؤثر در آزمایش الگوریتم زیاد باشد، آزمون کیفی نمی‌تواند جوابگو باشد. به همین دلیل در عمل از اندازه‌گیری‌هایی کمی استفاده می‌شود که حتی‌الامکان همبستگی بالایی با آزمون‌های کیفی داشته باشند. در این تحقیق، از معیارهای کمی  $^1$  LAR - distance [18] و [19] و PESQ<sup>2</sup> [20] تا [22] برای اندازه‌گیری کیفیت سیگنال خروجی استفاده شده است.

LAR - distance از خانواده معیارهای LPC<sup>3</sup> بوده و از تحلیل پیشگویی خطی برای مقایسه بین دو نمونه از سیگنال گفتار استفاده می‌کند. اولین قدم در این روش سنجش گفتار، انجام یک تحلیل LPC بر روی سیگنال‌های مورد مقایسه است. با فرض در اختیار بودن  $Q$  ضریب انعکاسی مربوط به فریم  $m$ -ام (یعنی:  $K(m, q)$ ,  $q = 1, 2, \dots, Q$ ) مقادیر نسبت سطح لوله صوتی (AR)<sup>4</sup> به‌طور متناظر از رابطه

$$AR(m, q) = \frac{1 + K(m, q)}{1 - K(m, q)} \quad (14)$$

محاسبه می‌شود. در این تحقیق، با توجه به فرکانس نمونه‌برداری  $f_s = 8$  KHz، تعداد ضرایب را  $Q = 10$  فرض نموده‌ایم. با در اختیار



شکل 4: تابع آستانه‌گذاری وقتی برای مقادیر مختلف  $\beta$ .

حاضر، به‌طور وقتی و خودکار از این تابع آستانه‌گذاری استفاده می‌کنیم. در این پژوهش به‌منظور کاهش ناپوستگی زمانی فریم‌های متوالی سیگنال گفتار، حد آستانه  $\gamma$  عددی نزدیک به صفر منظور می‌شود. اما  $\lambda$  حد آستانه‌ای است که - با توجه به ابتدای بحث - باید به‌طور وقتی در هر زیرباند تعیین شود. از این رو، فرمول حد آستانه Donoho به‌فرم زیر اصلاح می‌شود

$$\lambda_j^i(m) = \alpha_j^i(m) \sigma_j^i(m) \sqrt{2 \log(L_i \log_2(L_i))} \quad (11)$$

که  $\lambda_j^i(m)$  حد آستانه در زیرباند  $i$ -ام، سطح  $j$ -ام و فریم  $m$ -ام و  $L_i$ ، تعداد ضرایب موجک در زیرباند  $i$ -ام است.  $\sigma_j^i(m)$  انحراف استاندارد نویز در زیرباند  $i$ -ام و سطح  $j$ -ام در فریم  $m$ -ام است که از الگوریتم تخمین توان نویز بخش 4 به‌دست می‌آید.  $\alpha_j^i(m)$  در واقع ضرایب بهینه سیستم مرتبه اول بازگشتی (توضیح داده‌شده در بخش 4) است که با حداقل شدن میانگین مجذور خطای تخمین توان نویز به‌فرم زیر به‌دست می‌آید [13]

$$\alpha_j^i(m) = \frac{1}{1 + ((P_j^i(m-1) / \sigma_j^{i2}(m-1)) - 1)^2} \quad (12)$$

در رابطه اخیر  $0 \leq \alpha_j^i(m) \leq 1$  است.  $P_j^i(m-1)$  توان سیگنال گفتار نویزی در زیرباند  $i$ -ام، سطح  $j$ -ام و فریم  $m-1$  و  $\sigma_j^{i2}(m-1)$  واریانس نویز در فریم  $m-1$ ، زیرباند  $i$ -ام و سطح  $j$ -ام می‌باشد.  $\alpha_j^i(m)$  با نسبت توان سیگنال گفتار نویزی به توان نویز رابطه عکس دارد. بنابراین در زیرباندی که گفتار مسلط است،  $\alpha_j^i(m)$  کوچک شده و در نتیجه حد آستانه  $\lambda_j^i(m)$  در (11) دارای مقداری کم می‌شود. در زیرباندی که نویز مسلط است،  $\alpha_j^i(m)$  زیاد شده و در نتیجه  $\lambda_j^i(m)$  بزرگ می‌شود.

پارامتر  $\beta$  در (10) تعیین‌کننده نوع تابع آستانه‌گذاری است. به‌عنوان مثال  $\beta = 1$  متناظر با آستانه‌گذاری سخت پیوسته و  $\beta = 0$  متناظر با آستانه‌گذاری نرم است. با توجه به مباحث قبلی، در روش پیشنهادی به‌دنبال آن هستیم که در نواحی (زمان - فرکانسی) با گفتار مسلط، تابع آستانه‌گذاری سخت و در نواحی (زمان - فرکانسی) با نویز مسلط، تابع آستانه‌گذاری نرم به‌کار ببریم. در نواحی با حالت بینابینی (یا  $\alpha_j^i(m)$  متوسط)، نیز باید تابع آستانه‌گذاری بین حالت نرم و سخت انتخاب شود. بنابراین می‌توان بین  $\beta$  و  $\alpha$  رابطه زیر را پیشنهاد نمود

$$\beta_j^i(m) = 1 - \alpha_j^i(m) \quad (13)$$

1. Log Area Ratio - Distance

2. Perceptual Evaluation of Speech Quality

3. Linear Prediction Coding

4. Area - Ratio

که  $M$  تعداد فریم‌های فعال گفتاری را نشان می‌دهد. به منظور حذف اثر فریم‌های با فاصله غیر معقول، مطابق با آنچه در [18] پیشنهاد شده، ابتدا فریم‌هایی که LAR - distance در آنها از 95% مقدار ماکزیمم بیشتر است، کنار گذاشته شده و متوسط‌گیری بر روی بقیه فریم‌ها انجام می‌شود. ملاک نهایی مورد ارزیابی، تفاضل Avg - segLAR گفتمان پاک‌سازی شده از Avg - segLAR گفتمان نویزی است. منفی شدن این مقدار و بزرگ شدن قدر مطلق این مقدار منفی، نشانه کم شدن فاصله میان گفتار خروجی و گفتار تمیز و به تعبیر دیگر، بهبود کیفیت سیگنال پاک‌سازی شده در مقایسه با سیگنال نویزی ورودی است. بهبود Avg - segLAR بیانگر طبیعی‌تر بودن گفتار بهسازی شده است.

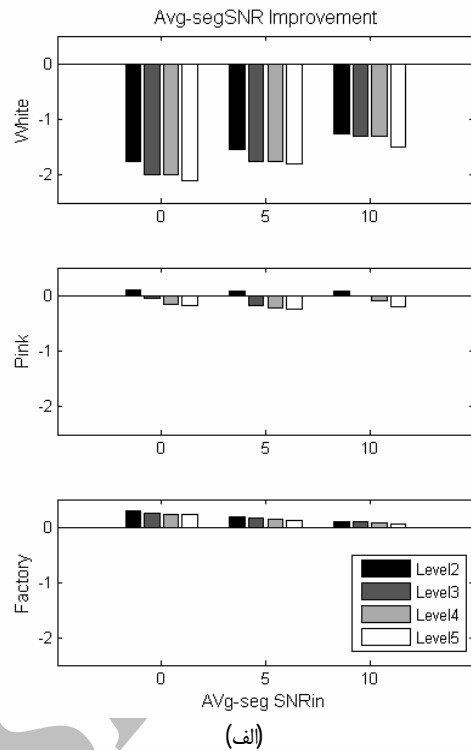
معیار کمی دیگری که برای تقریب نتایج آزمون‌های کیفی استفاده شده، PESQ است. در واقع این معیار مشابه با آزمون کیفی MOS، نمره‌ای بین صفر تا 4/5 به سیگنال می‌دهد. هرچه این نمره بزرگ‌تر باشد، بیانگر این است که میزان اعوجاج سیگنال بهسازی شده کمتر است.

اندازه‌گیری‌های کمی به‌آسانی نمی‌توانند اثر نویز باقیمانده و اعوجاج گفتار را در گفتار پاک‌سازی شده تعیین کنند. آنالیز توزیع زمان - فرکانس نویز باقیمانده و اعوجاج ایجاد شده در گفتار خروجی از اهمیت خاصی برخوردار است. از این رو، مشاهده اسپکتروگرام‌های گفتار اطلاعات مهمی در مورد نویز باقی‌مانده و اعوجاج گفتار می‌دهد.

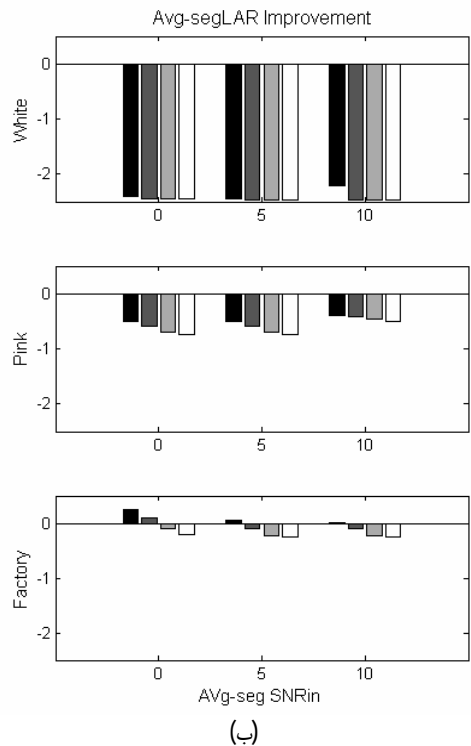
## 2-6- انتخاب سطح تجزیه مناسب در تبدیل موجک

### معمولی و تبدیل کامل بسته موجک

به‌عنوان اولین آزمایش از موجک db8 - که یک فیلتر با درجه نسبتاً بالا است - برای تعیین سطح موجک مناسب در تبدیل موجک معمولی و تبدیل کامل بسته موجک استفاده می‌شود. شکل 5- الف نتایج شبیه‌سازی مربوط به رشد درخت تبدیل موجک معمولی را نشان می‌دهد. مشاهده می‌شود در مقادیر مختلف Avg-segSNRin، در مورد همه نویزها (سفید، صورتی و کارخانه)، افزایش تعداد سطوح موجک، افزایش میزان بهبود Avg - segLAR را به دنبال دارد. از این رو افزایش سطح تجزیه، موجب بهبود خروجی سیستم می‌شود. به‌طور مشابه در شکل 5- ب اثر افزایش تعداد سطوح تجزیه در تبدیل کامل بسته موجک نشان داده شده است. در مورد نویز سفید، بهبود معیار مورد ارزیابی در اثر افزایش سطح تجزیه جزئی است، زیرا توزیع توان نویز سفید در همه باندها یکسان است. ولی در مورد نویزهای رنگی و غیرایستای میزان اثر بهبود معیار مورد ارزیابی با افزایش سطح تجزیه قابل توجه است. با توجه به این که در هر فریم 256 نمونه وجود دارد، بعد از تجزیه بسته موجک با 5 سطح در هر زیرباند موجک 8 نمونه وجود خواهد داشت. اگر سطح تجزیه موجک از 5 فراتر رود، تعداد نمونه‌ها در هر زیرباند کمتر شده و در نتیجه با این نمونه‌ها دقت تخمین توان نویز کاهش می‌یابد. با توجه به این مسأله و نیز با در نظر گرفتن شکل‌های 5- الف و 5- ب می‌توان گفت عدد 5 برای تعداد سطح تجزیه مقداری مناسب می‌باشد. به‌همین دلیل در شکل 5، نتایج آزمایش اثر رشد درخت موجک تا سطح 5 نشان داده شده است. از مقایسه قسمت‌های الف و ب شکل 5 معلوم می‌شود که تبدیل کامل بسته موجک در مقایسه با تبدیل موجک معمولی در تمام نویزها و در Avg - segSNRin های صفر تا ده dB، از بهبود بالاتری در Avg-segLAR برخوردار است. به‌طور مشابه در شکل 6 مشاهده می‌شود که تبدیل کامل بسته موجک با سطح تجزیه 5، بیشترین نمره آزمون PESQ را دارد. بنابراین با انتخاب سطح 5 برای تبدیل کامل بسته موجک و به‌کار بردن الگوریتم بهسازی مشاهده می‌شود که گفتار خروجی



(الف)

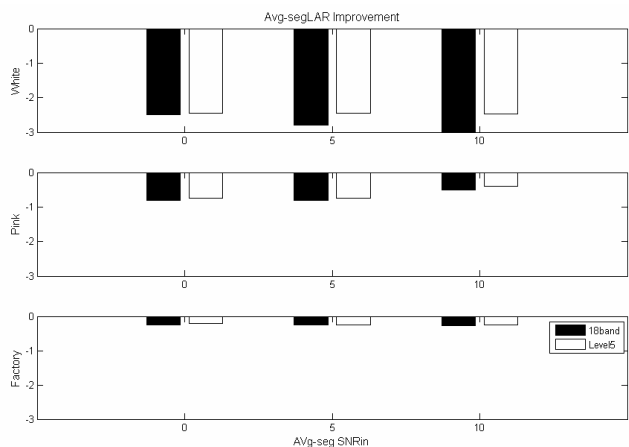


(ب)

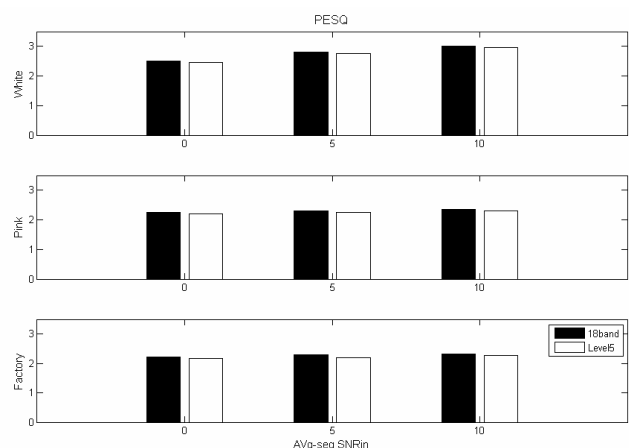
شکل 5: بررسی اثر افزایش سطوح تجزیه با در نظر گرفتن معیار بهبود Avg - segLAR در (الف) تبدیل موجک معمولی و (ب) تبدیل کامل بسته موجک با انتخاب موجک db8.

بودن پارامترهای AR مربوط به سیگنال تمیز اولیه  $(AR_s(m, q))$  و سیگنال مورد نظر (اعم از سیگنال گفتار نویزی یا سیگنال خروجی پاک‌سازی شده)  $(AR_z(m, q))$  معیار LAR - distance از متوسط‌گیری بر روی فریم‌های فعال گفتاری (به‌غیر از نواحی سکوت) به‌فرم زیر حاصل می‌شود [18] و [19]

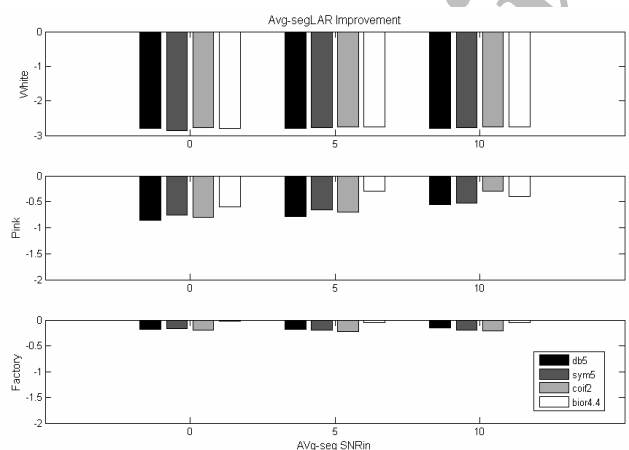
$$Avg - segLAR = \frac{1}{M} \sum_{m=1}^M \frac{1}{Q} \sum_{q=1}^Q \left| 20 \log_{10} \left[ \frac{AR_s(m, q)}{AR_z(m, q)} \right]^2 \right|^{\frac{1}{2}} \quad (15)$$



شکل 7: مقایسه تبدیل کامل بسته موجک و تبدیل PYPD با استفاده از معیار Avg - seg LAR.



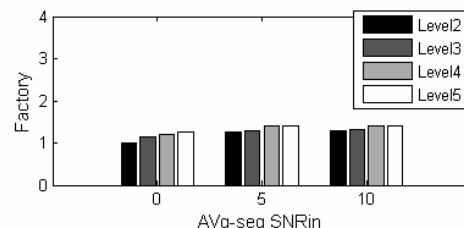
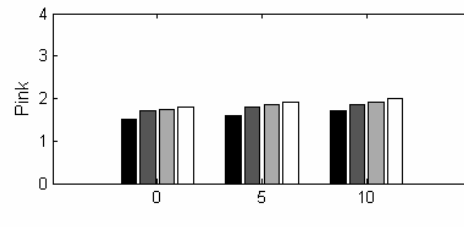
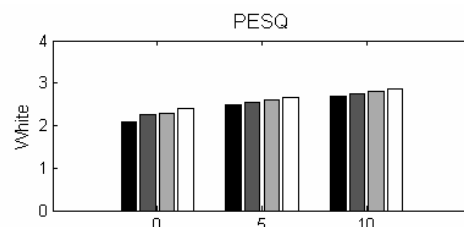
شکل 8: مقایسه تبدیل کامل بسته موجک و تبدیل PYPD با استفاده از آزمون PESQ.



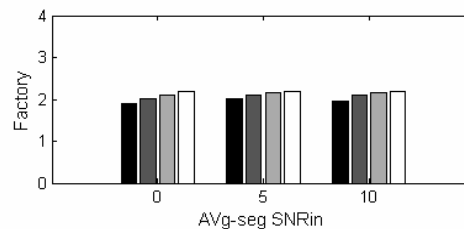
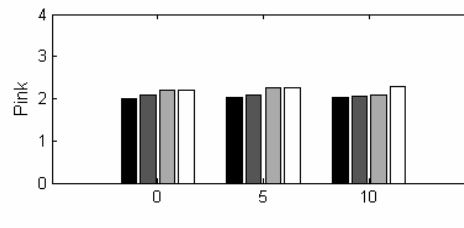
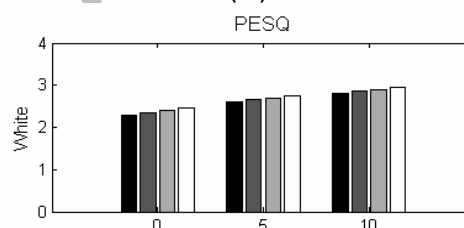
شکل 9: بررسی کارایی موجک‌ها با طول فیلتر برابر در تبدیل PYPD با استفاده از معیار Avg - seg LAR.

### 3-6 انتخاب موجک مناسب

در آزمایش قبلی مشاهده شد که تبدیل بسته موجک بر اساس باندهای بحرانی، گزینه مطلوب‌تری برای الگوریتم بهسازی گفتار می‌باشد. به منظور ارزیابی کارایی انواع مختلف موجک‌ها در بهبود خروجی سیستم بهسازی گفتار، موجک‌های db5، Bior4/4، Sym5، Coif2 که دارای طول فیلتر تقریباً مساوی هستند، در نظر گرفته می‌شود. با این انتخاب، بار محاسباتی برای موجک‌ها تقریباً یکسان می‌باشد. شکل 9 نشان می‌دهد که در نویز



(الف)

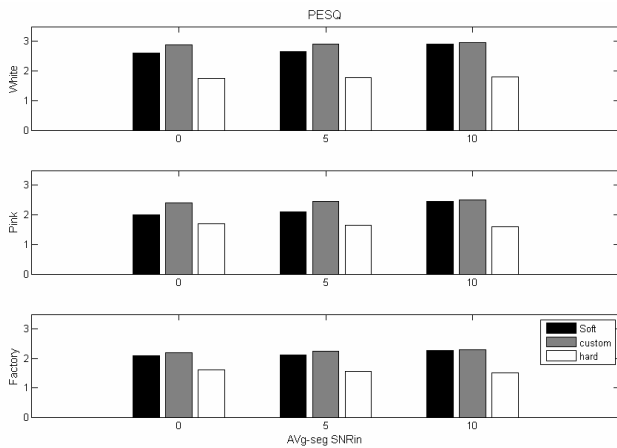


(ب)

شکل 6: بررسی اثر افزایش سطوح تجزیه با آزمون PESQ در (الف) تبدیل موجک معمولی و (ب) تبدیل کامل بسته موجک با انتخاب موجک db8.

طبیعی‌تر و درصد اعوجاج آن کمتر می‌شود.

نتایج به‌کارگیری الگوریتم پیشنهادی با تبدیل کامل بسته موجک و تبدیل بسته موجک بر پایه باندهای بحرانی در شکل 7 نشان داده شده است. مشاهده می‌شود که این دو تبدیل تقریباً نتایج یکسانی را نشان می‌دهند. بنابراین با توجه به بار محاسباتی کمتر و تطابق بیشتر مدل بارک با ساختار شنوایی انسان، استفاده از تبدیل بسته موجک بر پایه باندهای بحرانی منطقی‌تر و مطلوب‌تر است. نمرات آزمون PESQ در شکل 8 نیز نتیجه فوق را تأیید می‌کند.



شکل 12: بررسی توابع مختلف آستانه‌گذاری بر روی الگوریتم بهسازی گفتار با استفاده از آزمون PESQ.

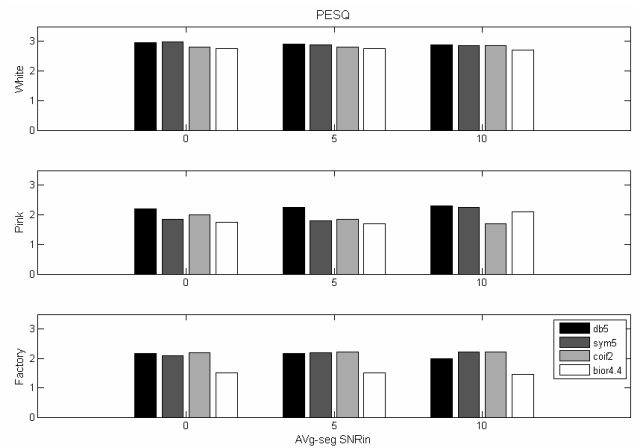
به بحث نظم موجک توجیه نمود: خانواده موجک *coiflet* تنها موجکی است که تابع موجک آن دارای ممان‌های صفر می‌باشد. اهمیت این موضوع در مورد نویز غیرایستای کارخانه - که بعد از فرآیند حذف نویز دارای نویز باقیمانده (به علت عدم ردیابی مناسب الگوریتم تخمین توان نویز) بوده و نیاز به هموارسازی دارد - بیشتر است. از مقایسه شکل 10 (حاصل آزمون PESQ) و شکل 9 به نتایجی مشابه می‌رسیم. شکل 10 نشان می‌دهد که نوع موجک بر میزان اعوجاج سیگنال گفتار بهسازی شده تأثیرگذار است. به عنوان مثال، در شکل 10 موجک *Bior4/4* در تمام نویزها کمترین نمره PESQ را به خود اختصاص داده است.

#### 4-6 بررسی کارایی تابع آستانه‌گذاری وفقی در برابر توابع آستانه‌گذاری سخت و نرم

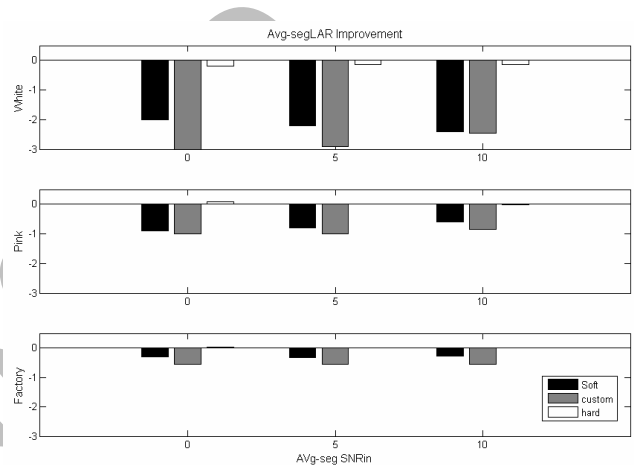
در تحقیقات قبلی [2] تا [9] تابع آستانه‌گذاری ثابت و یکسان (فقط نرم یا فقط سخت) با حد آستانه متغیر (با در نظر گرفتن ویژگی‌های گفتار و ساختار شنوایی انسان) استفاده شده است. در تحقیق حاضر، برای این که کارایی خوب تابع آستانه‌گذاری وفقی با حد آستانه متغیر نشان داده شود، سه تابع آستانه‌گذاری نرم، سخت و وفقی با حد آستانه محاسبه شده از (11) در نویزهای سفید، صورتی و کارخانه آزمایش می‌شود. تجزیه موجک بر اساس معیار بارک است. با توجه به نتایج بخش 3-6 موجک *db5* برای نویزهای سفید و صورتی و موجک *coif2* برای نویز کارخانه اعمال شده است. شکل 11 نشان می‌دهد که طبیعی شدن صدا با اعمال تابع آستانه‌گذاری وفقی حاصل می‌شود (نه با توابع آستانه‌گذاری سخت و نرم). در شکل 12، نتایج آزمون PESQ نمره بیشتری را به تابع آستانه‌گذاری وفقی اختصاص داده است. تابع آستانه‌گذاری سخت کیفیت پایین‌تری نسبت به دو تابع دیگر دارد که این ناشی از وجود نویز باقیمانده بیش از حد در خروجی سیستم می‌باشد و بعد از آن تابع آستانه‌گذاری نرم نتیجه بهتری دارد.

#### 5-6 بررسی الگوریتم پیشنهادی در مقایسه با دو روش آستانه‌گذاری موجک

برای ارزیابی مقایسه‌ای کارایی الگوریتم بهبودیافته از دو الگوریتم بهسازی گفتار در حوزه موجک که در [7] و [9] معرفی شده، استفاده می‌شود. عملکرد برتر آستانه‌گذاری نرم (به شرح بخش 4-6) مبنای انتخاب الگوریتم‌ها برای ارزیابی مقایسه‌ای بوده است. هر دو الگوریتم [7]



شکل 10: بررسی کارایی موجک‌ها با طول فیلتر برابر در تبدیل PWD با استفاده از آزمون PESQ.

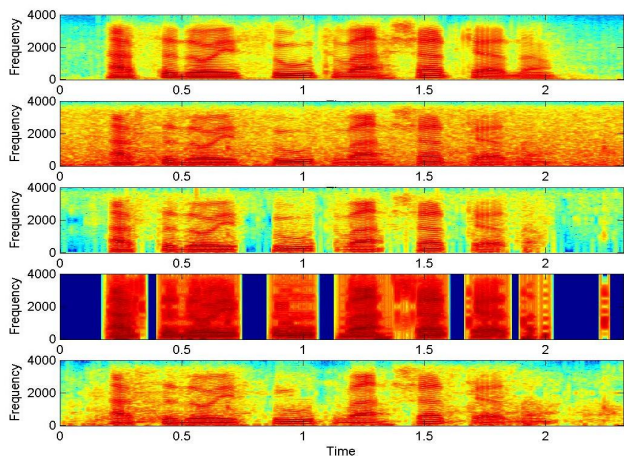


شکل 11: بررسی توابع مختلف آستانه‌گذاری بر روی الگوریتم بهسازی گفتار با استفاده از معیار Avg-segLAR.

سفید به ازای SNRهای 0 dB، 5 dB و 10 dB نتایج به دست آمده از اعمال هر یک از موجک‌ها تقریباً یکسان است (البته به نظر می‌رسد به طور جزئی موجک *db5* نتیجه بهتری دارد)، ولی در نویزهای صورتی و کارخانه تفاوت بیشتری در نتایج مشاهده می‌شود. در مورد نویزهای صورتی و کارخانه به نظر می‌رسد به ترتیب موجک‌های *db5* و *coif2* بهتر باشند. معمولاً الگوریتم‌های حذف نویز گفتار در حذف نویز سفید (در مقایسه با بحث حذف نویزهای رنگی مانند صورتی و کارخانه) توانمندتر می‌باشند. نوع موجک به کار رفته در محیط نویز سفید چندان اهمیتی ندارد. اما هدف کلی و عمده‌تر از اصلاح الگوریتم *Donoho*، حذف نویزهای رنگی و غیرایستای از گفتار است، زیرا به طور طبیعی با این گونه نویزها بیشتر سر و کار داریم. معمولاً کارایی الگوریتم در این نویزها همانند نویز سفید رضایت‌بخش نیست. در نتیجه نویز باقیمانده بعد از فرآیند حذف نویز موجب می‌شود که الگوریتم به نوع موجک اعمالی حساس باشد. مثلاً در نویزهای صورتی و کارخانه با سیستم موجک *Bior4/4* - که یک تبدیل متعامد دوتایی است - بدترین کارایی (کاهش بهبود *Avg-segLAR*) مشاهده می‌شود.

پارامتر دیگری که در تبدیل موجک اهمیت دارد، هموارشدن سیگنال بعد از فرآیند بهسازی است که اصطلاحاً به آن نظم موجک گفته می‌شود. هرچه تعداد ممان‌های صفر توابع مقیاس‌بندی و موجک بیشتر باشد، درجه نظم سیستم موجک و همواری سیگنال خروجی بالاتر می‌رود. به عنوان نمونه، در مورد نویز کارخانه به نظر می‌رسد *coif2* به طور نسبی از بهبود *Avg-segLAR* بیشتری برخوردار است؛ این نتیجه را می‌توان با توجه





شکل 15: به ترتیب از بالا به پایین: اسپکتروگرام‌های گفتار تمیز، گفتار نویزی، خروجی الگوریتم Lu، خروجی الگوریتم Chen، خروجی الگوریتم بهبودیافته.

اسپکتروگرام خروجی‌های نوعی سه الگوریتم مورد مقایسه در نویز کارخانه با  $Avg-segSNR_{in} = 0$  dB در شکل 15 نمایش داده شده است. از شکل 15 مشخص است که در خروجی دو الگوریتم Lu [7] و Chen [9]، در نواحی کم‌انرژی، نویز کاملاً حذف شده و نواحی پرانرژی هم دچار اعوجاج شده است. اثر اعوجاج در خروجی الگوریتم Chen قابل توجه‌تر است که دلیل آن نیز قبلاً بیان شد. ولی در الگوریتم بهبودیافته با وجود باقی‌ماندن کمی از نویز باقیمانده، از اعوجاج نواحی پرانرژی تا حد خوبی پرهیز شده است. آزمایش‌های شنیداری غیر رسمی نیز مؤید همین نتیجه است.

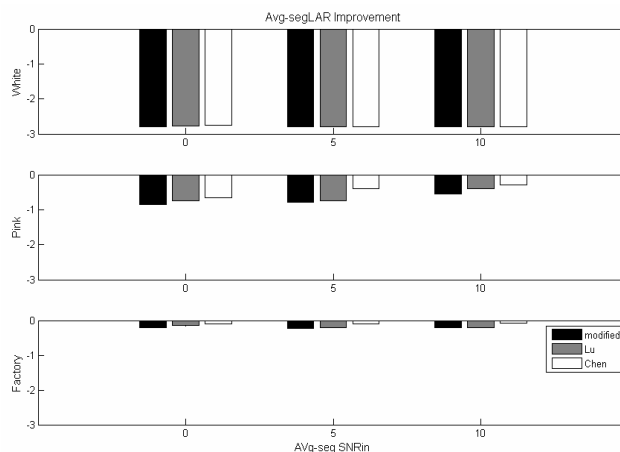
هرچند به دلیل حافظه‌دارشدن الگوریتم تخمین توان نویز [14]، بار محاسباتی روش پیشنهادی در مقایسه با الگوریتم Chen [9] کمی بیشتر است، ولی در مقابل کارایی آن در حد قابل توجهی افزایش یافته است. در مقایسه بین روش پیشنهادی و الگوریتم Lu [7] نیز باید گفت در حالی که بار محاسباتی دو روش تقریباً یکسان می‌باشد، کارایی روش پیشنهادی (بر اساس اطلاعات شکل 9) تا حدودی برتری دارد.

### 7- نتیجه‌گیری

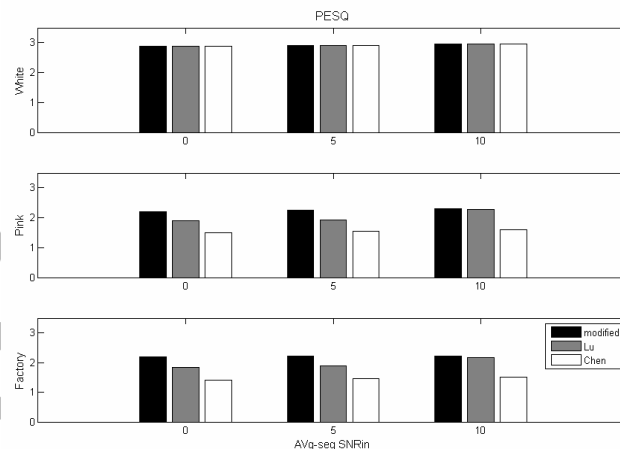
در این تحقیق عوامل مؤثر در عملکرد الگوریتم آستانه‌گذاری موجک برای گفتار نویزی مورد بررسی قرار گرفتند. این عوامل عبارتند از:

- حد آستانه و تابع آستانه‌گذاری
- نحوه رشد درخت موجک
- نوع موجک

در این پژوهش از حدود آستانه متغیر و تابع آستانه‌گذاری وقتی منطبق بر ویژگی‌های سیگنال گفتار و نویز (در هر زیرباند) استفاده شد. شبیه‌سازی‌ها نشان داد که استفاده از تابع آستانه‌گذاری وقتی از لحاظ کاهش نویز باقیمانده خروجی در مقایسه با آستانه‌گذاری سخت به مراتب بهتر بوده و از لحاظ اعوجاج خروجی نیز در برابر آستانه‌گذاری نرم عملکرد نسبی خوبی دارد. علاوه بر آن رشد درخت موجک بر مبنای معیار بارک از لحاظ بار محاسباتی و تطابق بیشتر آن با ساختار شنوایی انسان برگزیده شد. همچنین آزمایش‌ها نشان دادند که الگوریتم پیشنهادی بسته به سفید یا رنگی بودن طیف نویز به نوع موجک انتخابی حساس است. در نهایت روش بهبودیافته با دو سیستم عمده بهسازی گفتار مبتنی بر تبدیل موجک [7] و [9] مقایسه و برتری الگوریتم پیشنهادی از لحاظ بهبود کیفیت خروجی در قبال افزایش اندک بار محاسباتی تأیید شد.



شکل 13: مقایسه آستانه‌گذاری وقتی با دو روش دیگر آستانه‌گذاری ضرایب موجک با استفاده از معیار Avg - seg LAR.



شکل 14: مقایسه آستانه‌گذاری وقتی با دو روش دیگر آستانه‌گذاری ضرایب موجک با استفاده از آزمون PESQ.

و [9] از ساختارهای زیرباندی بحرانی (معیار بارک) برای تجزیه موجک و تابع آستانه‌گذاری نرم استفاده می‌کنند، با این تفاوت که در روش Lu و همکارش برای تنظیم حد آستانه از NMT (در هر زیرباند) و segSNR (در هر فریم) استفاده شده، در حالی که در روش Chen و همکارش از تکنیک عملگر انرژی تیگر برای تنظیم حد آستانه استفاده شده است.

با توجه به شبیه‌سازی‌های انجام‌گرفته در بخش‌های 2-6 و 3-6 الگوریتم بهبودیافته با تجزیه بسته موجک بر اساس باندهای بحرانی و به‌کارگیری موجک Coif2 در نویز کارخانه و موجک db5 در بقیه نویزها برای مقایسه با کارایی تکنیک‌های قبلی استفاده می‌شود. مشابه آزمایش‌های بخش‌های قبلی، هر سه الگوریتم تحت معیار بهبود Avg-segLAR و PESQ در نویزهای مختلف آزمایش شد. نتایج این بررسی در شکل‌های 13 و 14 نشان داده شده است. مشاهده می‌شود که در مقادیر مختلف نویز سفید (SNRهای 0 dB، 5 dB، 10 dB) هر سه الگوریتم نتایج مشابهی دارند، ولی در نویزهای صوتی و کارخانه الگوریتم Chen و همکارش با وجود بار محاسباتی کمتر نتیجه بدتری (به علت عدم تخمین مناسب از نویز) در مقایسه با دو روش دیگر دارد. مشاهده می‌شود که در نویزهای صوتی و کارخانه الگوریتم پیشنهادی از میزان بهبود Avg-segLAR و نمره PESQ نسبی بهتری در مقایسه با دو الگوریتم دیگر (مخصوصاً در SNRهای 0 dB و 5 dB) برخوردار است. در  $SNR = 10$  dB کارایی روش پیشنهادی با الگوریتم Lu [7] تفاوت چندانی ندارد.



## مراجع

- [15] B. J. Yoon and P. P. Vaidyanathan, "Wavelet-based denoising by customized thresholding," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'04*, vol. 2, pp. 925-928, May 2004.
- [16] <http://www.rcisp.com/FarsDat.aspx>
- [17] <http://svr-ww.eng.cam.ac.uk/comp.speech/Section1/Data/noisex.html>
- [18] S. R. Quackenbush, T. P. Barnwell, and M. A. Clements, *Objective Measures of Speech Quality*, Prentic - Hall, Englewood Cliffs, NJ, 1998.
- [19] J. H. L. Hansen and B. L. Pellom, "An effective quality evaluation protocol for speech enhancement algorithms," in *Proc. of 6th Int. Conf. on Spoken Language Processing, ICSLP'00*, vol. 7, pp. 2819-2822, Dec. 2000.
- [20] Y. Hu and P. C. Loizou, "Evaluation of objective measures for speech enhancement," in *Proc. of Interspeech*, pp. 1447-1450, Sep. 2006.
- [21] <http://www.utdallas.edu/~loizou/speech/software.htm>
- [22] A. W. Rix, M. P. Hollier, A. P. Hekstra, and J. G. Beerends, "Perceptual evaluation of speech quality (PESQ): the new ITU standard for end-to-end speech quality assessment. Part I: time-delay compensation," *J. of the Audio Engineering Society*, vol. 50, no. 10, pp. 755-764, Oct. 2002.
- فاطمه شیخ علیشاهی** به ترتیب در سالهای 1381 و 1384 مدرک کارشناسی مهندسی برق (الکترونیک) و مدرک کارشناسی ارشد مهندسی برق (مخابرات) خود را از دانشگاه یزد دریافت نمود. وی در سال 1387 به دوره دکترای مهندسی برق (مخابرات) در دانشگاه یزد وارد گردید و هم اکنون در حال تحصیل می‌باشد. زمینه‌های علمی مورد علاقه ایشان عبارتند از: پردازش سیگنال گفتار و رادار.
- حمیدرضا ابوطالبی** دوره کارشناسی و کارشناسی ارشد خود را به ترتیب در سال‌های 1374 و 1377 در رشته مهندسی برق (مخابرات) در دانشگاه صنعتی شریف گذرانده و مدرک دکترای خود را در سال 1382 در رشته مهندسی برق (مخابرات) از دانشگاه صنعتی امیرکبیر اخذ نمود. ایشان در سال 1381 در یک دوره تحقیقاتی در دانشگاه واترلو کانادا به سر برد و از سال 1382 به عنوان عضو هیأت علمی به دانشکده مهندسی برق و کامپیوتر دانشگاه یزد پیوست. زمینه‌های علمی مورد علاقه نامبرده، بهسازی سیگنال گفتار، فیلترهای وقتی، و پردازش آرایه‌ای سیگنال گفتار می‌باشد.
- محمد رضا تابان** در سال 1369 مدرک کارشناسی مهندسی برق-الکترونیک خود را از دانشگاه صنعتی اصفهان، در سال 1372 مدرک کارشناسی ارشد مهندسی برق-مخابرات خود را از دانشگاه تربیت مدرس و در سال 1377 مدرک دکترای مهندسی برق-مخابرات خود را از دانشگاه صنعتی اصفهان دریافت نمود و هم اکنون دانشیار دانشکده مهندسی برق و کامپیوتر دانشگاه یزد می‌باشد. زمینه‌های علمی مورد علاقه نامبرده متنوع بوده و شامل موضوعاتی مانند پردازش آماری سیگنال، تخمین و آشکارسازی و پردازش سیگنال‌های راداری می‌باشد.
- [1] D. L. Donoho, "Denoising by soft thresholding," *IEEE Trans. on Information Theory*, vol. 41, no. 3, pp. 613-627, May 1995.
- [2] M. Bahoura and J. Rouat, "A new approach for wavelet speech enhancement," in *Proc. of European Conf. on Speech Communication and Technology, Eurospeech'01*, vol. 3, pp. 1937-1940, Sep. 2001.
- [3] H. R. Abutalebi, F. Sheikhalishahi, and M. R. Taban, "Speech enhancement in wavelet domain by use of speech features," in *Proc. of the 14th Iranian Conf. on Electrical Engineering, ICEE'06*, Tehran, Iran, May 2006.
- [4] H. Sheikhzadeh and H. R. Abutalebi, "An improved wavelet-based speech enhancement system," in *Proc. of European Conf. on Speech Communication and Technology, Eurospeech'01*, vol. 3, pp. 1855-1858, Sep. 2001.
- [5] S. Chang, Y. Kwon, S. I. Yang, and I. J. Kim, "Speech enhancement for non-stationary noise environment by adaptive wavelet packet," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'02*, vol. 1, pp. 561-564, May 2002.
- [6] C. T. Lu and H. C. Wang, "Enhancement of single channel speech based on masking property and wavelet transform," *Speech Communication*, vol. 41, no. 2-3, pp. 409-427, Feb. 2003.
- [7] C. T. Lu and H. C. Wang, "Speech enhancement using robust weighting factors for critical-band-wavelet-packet transform," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'04*, vol. 1, pp. 721-724, May 2004.
- [8] J. W. Seok and K. S. Bae, "Speech enhancement with reduction of noise components in the wavelet domain," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'97*, vol. 2, pp. 1323-1326, Apr. 1997.
- [9] S. H. Chen and J. F. Wang, "Speech enhancement using perceptual wavelet packet decomposition and teager energy operator," *J. of VLSI Signal Processing*, vol. 36, no. 2-3, pp. 125-139, Nov. 2004.
- [10] Y. Ghanbari and M. R. Karimi-Mollaei, "A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets," *Speech Communication*, vol. 48, no. 8, pp. 927-940, Aug. 2006.
- [11] M. T. Johnson, X. Yuan, and Y. Ren, "Speech signal enhancement through adaptive wavelet thresholding," *Speech Communication*, vol. 49, no. 2, pp. 123-133, Feb. 2007.
- [12] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelet and Wavelet Transforms, A primer*, Upper Saddle River, NJ: Prentic-Hall, 1998.
- [13] X. Huang, R. A. Acero, H. Kon, and R. Reddy, *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development*, Prentic-Hall, 2001.
- [14] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 5, pp. 504-512, Jul. 2001.