

تقویت گفتار با یک فیلتر وینر تطبیقی

چکیده: این مقاله یک روش تطبیقی فیلتر وینر برای تقویت گفتار را پیشنهاد می کند. این روش بستگی به انطباق عملکرد انتقال فیلتر از نمونه به نمونه براساس آمار سیگنال گفتار؛ میانگین محلی و واریانس محلی دارد. این در حوزه زمان و نه در حوزه فرکانس اجرا میشود تا برای ماهیت متغیر زمان سیگنال های گفتار وفق داده شود. روش ارائه شده با فیلتر وینر دامنه فرکانس های سنتی، روش های تفریق و حذف نویز موجک طیفی با استفاده از معیارهای کیفیت متفاوت گفتار مقایسه میشود. نتایج شبیه سازی برتری روش فیلتر وینر پیشنهادی در مورد سر و صدا افزودنی وایت گاوسی (AWGN) و همچنین صدای رنگ شده را نشان می دهد.

1. معرفی

سیگنال های گفتار، سیگنال هایی هستند که به طور گسترده ای بین انسانها، برای انتقال پیام استفاده می شوند. از این رو، محققان توجه زیادی به پردازش گفتار داشته اند و تعداد زیادی تحقیقات در علوم گفتار و شنوایی را ارائه کرده اند. سیستم های پردازش گفتار در طیف گسترده ای از برنامه های کاربردی مانند برنامه نویسی گفتار برای ارتباطات، تشخیص گفتار برای سیستم های اطلاعات خودکار و گفتار قبل از پردازش برای کمک به افراد ناشنوا استفاده میشوند. این سیستم ها با این فرض طراحی شده اند که صداهای پس زمینه تباہ کننده وجود ندارد. در یک محیط پر سر و صدا، تقویت گفتار برای بهبود عملکرد این سیستم ها پیشنهاد میشود.

تقویت گفتار یک واژه برای توصیف الگوریتم هاست، که می تواند به منظور بهبود کیفیت، کاهش خستگی شنوایی گفتار پر سر و صدا، افزایش فهم، و بهبود عملکرد سیستم های ارتباط صوتی مورد استفاده قرار گیرد. از سوی دیگر، هیچ سیستم تقویت گفتار نمی تواند هم کیفیت گفتار وهم فهم را بهبود بخشد. در واقع، فهم گفتار را می توان به عنوان جنبه ای از کیفیت مشاهده کرد، از آنجاییکه گفتار با کیفیت بالا همیشه فهم خوب را میرساند و گفتار نامفهوم به عنوان داشتن کیفیت بالا طبقه بندی نمیشود. در بسیاری از تحقیقات قبلی، تقویت گفتار کیفیت را افزایش

میدهد اما فهم را کاهش می دهد. چندین روش برای این منظور مانند روش تفریق طیفی ، روش شبه سیگنال، روش فیلتر وینر، و روش حذف نویز موجک مطرح شده است. بهبود نسبت سخنرانی سیگنال به نویز (SNR) هدف اکثر تکنیک ها است . تفریق طیفی یکی از روش های سنتی مورد استفاده برای افزایش سخنرانی تخریب شده توسط نویز پس زمینهافزودنی ثابت است. که می تواند به عنوان یک روش غیر پارامتری طبقه بندی شود، که نیاز به یک تخمین از طیف نویز دارد. یک مشکل رایج برای روش تفریق طیفی، مشخصه سر و صدا باقی مانده به نام سر و صدا موسیقی میباشد. تفریق طیفی نیز سر و صدا را به اندازه کافی در دوره سکوت کاهش نمی دهد.

فیلتر وینر یک روش جایگزین برای تفریق طیفی برای تقویت سیگنال گفتار است. فیلتر وینر یک فیلتر خطی است که برای بازیابی سیگنال گفتار اصلی از سیگنال پر سر و صدا با به حداقل رساندن میانگین مربعات خطا (MSE) بین سیگنال برآورد و سیگنال اصلی بکار میرود. حذف صدای موجک روش دیگر بر اساس تجزیه موجک سیگنال پر سر و صدا و آستانه در حوزه موجک برای حذف سر و صداست.

تبدیل موجک برای تجزیه سیگنال پر سر و صدا به زیر گروههای موسیقی مورد استفاده قرار میگیرد ، و کاهش سر و صدا یا توسط آستانه سخت یا نرم انجام می شود. ایراد این روش این است که این امر منجر به تحریف برخی از اجزای مفید گفتار اصلی نیز میشود. همچنین برخی از روش های دیگر نیز وجود دارد که مدل آماری یک تشخیصگر را برای شناسایی ویژگی های صدا قبل از تقویت گفتار تطبیق میدهد.

4- حذف نویز موجک

از آنجاییکه آنالیز موجک بر اساس تقلید حاشیه شنوایی انتهایی-جلو است ، تلاش هایی انجام شده است تا از این ابزار پردازش سیگنال برای تقویت گفتار بهره برداری شود. بیشترین روش استفاده شده بر اساس آستانه غیر خطی از ضرایب موجک میباشد، که تجزیه و تحلیل و فیلتر غیر خطی را متصل میکند. فرآیند آستانه یک فرآیند حذف نویز است.

تبدیل موجک، سیگنال گفتار پر سر و صدا را به دو ضریب جزء ؛ تقریب یا ضرایب پایین گذر و جزئیات یا ضرایب بالا گذر تجزیه میکند. هر تقریب یا اجزای جزئیات نیمی از طول سیگنال گفتار اصلی است. بیشتر انرژی سیگنال گفتار در

جزء تقریب متمرکز شده است. بنابراین، اثر سر و صدا در جزء تقریب کوچک و در جزء جزئیات بزرگ است. اگر یک فرایند آستانه بر جزء جزئیات انجام شود، به طور قابل توجهی سر و صدا را کاهش می دهد، که انرژی سیگنال های بی پیرایه را رها میکند.

4.1 اصول آستانه

فرض کنیم که تبدیل موجک از سیگنال $X(N)$ پر سر و صدادر معادله (1) توسط X داده شده است. آستانه بر اجزای جزئیات X انجام شده است. به طور کلی دو راه آستانه شدن وجود دارد؛ آستانه سخت و آستانه نرم. آستانه سخت چنین تعریف شده است (معادله 6). که در آن T آستانه ارزش انتخابی است. و آستانه نرم به شرح زیر است

$$Thr_{Hard}(X, T) = \begin{cases} X & |X| > T \\ 0 & |X| < T \end{cases} \quad (6)$$

هر دو این روش ها تحت تاثیر تحریف گفتاراند، چرا که آنها ضرایب را تعیین میکنند که ممکن است حامل برخی اطلاعات مفید تبدیل به صفر باشند، و منجر به ناپیوستگی های فرکانس قابل مشاهده شدید بار در طیف گفتار شوند.

4.2 چگونه آستانه انتخاب کنیم

انتخاب مقدار آستانه در چندین روش انجام میشود. دونوهو فرمول زیر را بر اساس فرض سر و صدا افزودنی وایت گاوسی AWGN به دست آورد. (فرمول 9).

که در آن T مقدار آستانه است، N طول سیگنال سر و صدا و $\sigma = MAD / 0.6745$ ، با MAD دلالت کننده میانگین مطلق که در مقیاس اول ضرایب موجک برآورد شده است.

جانستون و سیلورمن یک روش آستانه سطح وابسته را برای مقابله با سر و صدا مربوطه پیشنهاد دادند، که در آن برای هر وقفه فرکانس، آستانه متناسب با انحراف استاندارد سر و صدا در آن وقفه است (فرمول 10)

با N ، $\sigma a = MADa / 0.6745$ تعداد نمونه ها در مقیاس a ، و $MADa$ میانگین مطلق برآورد شده در مقیاس a است.

اگر چه روش حذف نویز موجک نیازی به یک مدل گفتار و یا سر و صدا ندارد، و می تواند به یک

گروه گسترده تر از سیگنال ها اعمال شود، ولی صرفا یک آستانه کلی برضرایب موجک، یک عملکرد خوب را تضمین نمی کند.

5- فیلتر وینر تطبیقی

این فیلتر وینر تطبیقی پیشنهادی از آمار محلی متفاوت سیگنال گفتار سود میبرد. یک بلوک دیاگرام روش فیلتر وینر تطبیقی در شکل 1 نشان داده شده است. در فرایند فیلترینگ، MX متوسط محلی تخمینی و σ_x^2 واریانس محلی سیگنال $x(n)$ ، استفاده شده اند.

6 نتایج شبیه سازی

برای هدف ارزیابی، ما یک سیگنال گفتار برای جمله "ما سال پیش دور بودیم" را برای یک مرد و برای یک زن استفاده کردیم. ما از معیارهای کیفیت گفتار مانند سیگنال به نویز (SNR)، سگمنتال سیگنال به نویز (SNRseg)، لگاریتم درستنمایی نسبت (LLR) و انحراف طیفی (SD) استفاده کردیم. ما ابتدا با سیگنال مرد شروع کردیم و صدای افزودنی وایت گاوسی (AWGN) را به آن با سیگنال به نویزها از 5- و 5 دسی بل اضافه کردیم. نتایج تمام روش های تقویت که در بالا توضیح داده شد در سیگنال گفتار مرد برای سیگنال به نویز (SNR) از 5 دسی بل در شکل 2 تا 8 نشان داده شده است.

برای مورد سر و صدا رنگی شده، ما سر و صدا رنگی را توسط فیلتر پایین گذر از سر و صدا افزودنی وایت گاوسی قبل از اضافه کردن آن به سیگنال شبیه سازی کردیم. همچنین ما تمام روش های تقویت سخنرانی را بر روی سیگنال های مرد و زن در حضور سر و صدای رنگی تست کردیم. نتایج حاصل از این آزمایشات برای سیگنال مرد در سیگنال به نویز برابر 5 دسی بل در شکل 9 تا 14 نشان داده شده است. شکل 15 خروجی سیگنال به نویز (SNR) در مقابل ورودی سیگنال به نویز (SNR) برای تمام روش ها در سیگنال مرد را نشان می دهد. شکل 16 SNR_{seg} در مقابل ورودی (SNR) برای تمام روش ها بر روی سیگنال مرد را نشان می دهد. شکل 17 تغییر لگاریتم درستنمایی نسبت (LLR) در مقابل ورودی (SNR) برای تمام روش ها بر روی سیگنال مرد را نشان می دهد. شکل 18 تغییرات انحراف طیفی (SD) در مقابل ورودی سیگنال به نویز برای تمام روش ها بر روی سیگنال مرد را نشان می

دهد. مورد سر و صدا رنگ شده نیز درمقایسه مطالعه شده است و نتایج آن در شکل 19 به 22 داده شده است. یک مطالعه مشابه بر روی سیگنال زن تکرار شده است، و نتایج در جدول 1 تا 4 جدول بندی شده اند. نتایج همه از روش فیلتر پیشنهادی تطبیقی وینر حمایت میکنند.

7 نتیجه گیری

یک روش تطبیقی فیلتر وینر برای تقویت گفتار در این مقاله ارائه شده است. این روش بستگی به اقتباس از پاسخ ضربه فیلتر از نمونه به نمونه بر اساس آمار سیگنال گفتار دارد. نتایج نشان می دهد که روش فیلتر وینر تطبیقی پیشنهادی بهترین عملکرد را در مقایسه با تمام دیگر روش های تقویت گفتار در هر دو مقدار کم و بالایی سیگنال به نویز ذکر شده در این مقاله دارد. فیلتر پیشنهادی در هر دو مورد سر و صدا افزودنی وایت گاوسی و سر و صدا های رنگ شده نتیجه بخش است. این به ماهیت تطبیقی از پاسخ ضربه فیلتر نسبت داده شده است. این فیلتر وینر تطبیقی پیشنهاد شده دارای مزیت دیگر است. آن این است که تنها وابسته به سیگنال پر سر و صدا به عنوان یک ورودی تنهاست.