

بهسازی وفقی سیگنال گفتار در محیط‌های واقعی با استفاده از ساختار ترکیبی مبتنی بر شکل‌دهنده‌های پرتو و فیلتر پسینه

آوید آوخ^۱، استادیار؛ حمیدرضا ابوطالبی^۲، استاد

۱- باشگاه پژوهشگران جوان و نخبگان، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران، aavokh@pel.iaun.ac.ir

۲- آزمایشگاه تحقیقاتی پردازش گفتار، دانشکده مهندسی برق، دانشگاه یزد، habutalebi@yazd.ac.ir

چکیده: تنوع میدان‌های نویزی حاکم بر محیط‌های واقعی، طراحی سیستم واحدی را که قادر به حذف کامل همه اغتشاشات محیطی باشد، با مشکل مواجه می‌کند. لذا، داشتن یک شناخت اولیه از میدان‌های نویزی موجود در محیط و بررسی نوع اثر هر یک بر روی سیگنال گفتار، گام مؤثری در جهت فرآیند بهسازی محسوب می‌شود. در این مقاله، ضمن بررسی نویزهای اساسی موجود در زمینه گفتار ضبط شده در شرایط واقعی، به ارزیابی میدان‌های نویزی پایین‌گذر پخش شده و مقایسه تابع هم‌دوسی فضایی آنها در شرایط واقعی با نتایج تئوری خواهیم پرداخت. همچنین، ساختار کارآمد ترکیبی مبتنی بر آرایه‌های میکروفونی متشکل از شکل‌دهنده‌های پرتو وفقی و فیلتر پسینه پیشنهاد می‌شود. در طرح پیشنهادی، با هدف دنبال کردن شرایط پویای محیط و کنترل وفقی ضرایب وزن‌دهی شکل‌دهنده پرتو، دو الگوریتم حداقل میانگین مربع و سریع‌ترین کاهش استفاده و اثر هر یک بر بهسازی سیگنال گفتار مقایسه خواهد شد. ساختار مذکور، ضمن ایجاد صفرهای کنترل شده‌ای در الگوی تشعشعی شکل‌دهنده پرتو برای مقابله پویا با میدان‌های نویزی جهت‌دار محیط، به تعدیل نویزهای ناهمبسته و پخش شده می‌پردازد. بررسی اثر سرعت همگرایی الگوریتم‌های مورد استفاده بر میزان SNR در حضور انواع میدان‌های نویزی با طیف‌های فرکانسی مختلف از دیگر فعالیت‌های انجام شده در این تحقیق محسوب می‌شود. نتایج حاصل از شبیه‌سازی و همچنین پیاده‌سازی طرح پیشنهادی در شرایط واقعی و در سناریوهای مختلف، حاکی از نقش به‌سزای آن در بهبود عملکرد بهسازی سیگنال گفتار دارد.

واژه‌های کلیدی: آرایه میکروفونی، بهسازی گفتار، شکل‌دهی پرتو، فیلتر پسینه، نویز پخش شده.

A Hybrid Adaptive Beam-former and Post-filter for Speech Enhancement in Real Environments

A. Avokh¹, Assistant Professor; H. R. Abutalebi², Professor

1- Young Researchers and Elite Club, Najafabad Branch, Islamic Azad University, Najafabad, Iran, Email: aavokh@pel.iaun.ac.ir

2- Speech Processing Research Lab (SPRL), Electrical Eng. Dept., Yazd University, Yazd, Iran, Email: habutalebi@yazd.ac.ir

Abstract: Diversity of the noise fields in real environments makes difficult to design an efficient de-noising system. This paper addresses the problem of speech enhancement in real environments. Towards this goal, we present a hybrid scheme based on the adaptive beam-forming and post filtering. In order to capture the adaptive conditions of the environment, we adopt two algorithms: steepest descent and least mean square. We also consider the application of the post-filters on the output of the beam-former. As an advantage, the proposed structure not only produces some controlled nulls to reduce the directional noises, but also adaptively minimizes the power of diffuse and uncorrelated noises. The study of the convergence of the proposed algorithms and also the behavior of the diffuse noise in real conditions are other issues considered in this paper. Numerical results of our comprehensive implementation confirm the efficiency of the proposed scheme in suppression of the real environmental noises.

Keywords: Microphone array, Speech enhancement, Beam-forming, Post filter, Diffuse Noise.

تاریخ ارسال مقاله: ۱۳۹۶/۰۳/۰۸

تاریخ اصلاح مقاله: ۱۳۹۶/۰۸/۱۷

تاریخ پذیرش مقاله: ۱۳۹۶/۱۱/۱۷

نام نویسنده مسئول: آوید آوخ

نشانی نویسنده مسئول: اصفهان، نجف‌آباد، بلوار دانشگاه، دانشگاه آزاد اسلامی واحد نجف‌آباد، دانشکده برق، گروه مخابرات.

۱- مقدمه

به موازات بلوغ روش‌های پردازش گفتار و پیشرفت سیستم‌های مبتنی بر سیگنال‌های صوتی، مقابله با اثرات مخرب محیطی یک امر ضروری به‌شمار می‌آید. همانطور که می‌دانیم سیگنال گفتار از بدو تولید، دستخوش عوامل مختلفی از جمله نویز، تداخل و اعوجاج می‌شود. این پدیده، ضمن پایین آوردن کیفیت گفتار، عملکرد سیستم‌های پردازنده صوتی همچون سیستم‌های بازشناسی گفتار و کدکننده‌ها را نیز دچار مشکل می‌کند. از یکسو، تنوع عوامل مخرب محیطی و از سوی دیگر، محدودیت‌های عملی موجود در امر بهسازی گفتار، طراحی سیستم واحدی را که قادر به حذف کامل همه اغتشاشات محیطی باشد، با مشکل مواجه می‌نماید. از طرفی، به‌دلیل ماهیت پویای سیگنال گفتار و میدان‌های نویزی، استفاده از الگوریتم‌های بهسازی ثابت نمی‌تواند گزینه مناسبی تلقی شود. لذا، لازم است در شرایط مختلف، براساس ضوابط موجود، الگوریتم بهسازی وقتی مقتضی انتخاب شود. براساس میزان همبستگی سیگنال نویز در نقاط مختلف از فضا، نویزهای محیط را می‌توان در قالب دو دسته نویزهای همبسته و نویزهای ناهمبسته افراز نمود [۱]. در یک میدان نویزی همبسته، امواج، بدون هیچ انعکاس و یا پراکندگی، در فضای آزاد و بدون مانع عمده در پخش صدا به‌سمت آرایه‌ای از میکروفون‌ها منتشر می‌شوند؛ این درحالی است که در میدان‌های نویزی ناهمبسته، سیگنال نویز در نقاط مختلف، از همبستگی پایینی برخوردار است. در این بین، یکی از بدیهی‌ترین میدان‌های نویزی حاکم در فضاهای عمومی و وسایل نقلیه، میدان‌های نویزی پخشنده^۱ می‌باشد [۲، ۳]. میدان نویزی پخشنده، برآیند سیگنال‌های ناخواسته‌ای است که با احتمال مساوی و انرژی یکنواخت در همه جهت‌های محیط منتشر شده و همگی دارای دامنه برابر و فاز تصادفی هستند. این مدل که تلفیقی از اجزای همبسته و ناهمبسته است، به‌دلیل متداول بودن، از اهمیت بالایی برخوردار بوده و داشتن شناخت کافی از آن کمک شایانی به امر بهسازی گفتار می‌کند.

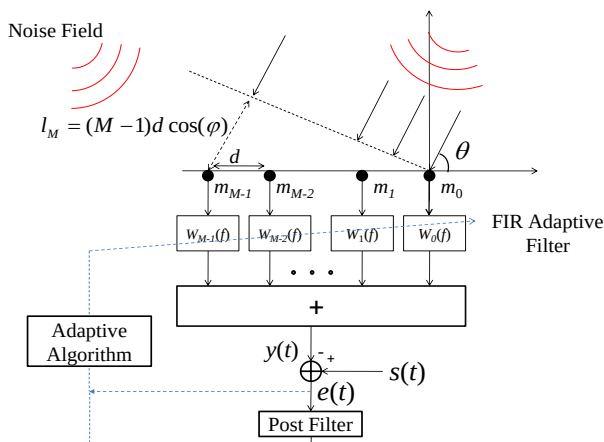
پس از آنکه پردازش آرایه‌ای، توانایی خود را در زمینه انواع سیگنال‌های مخابراتی همچون رادار، سونار و مخابرات سیار به اثبات رساند [۴]، امروزه استفاده از آرایه‌های میکروفونی در کاربردهای بهسازی گفتار [۵-۸]، بازشناسی گفتار [۹] و مکان‌یابی منابع صوتی [۱۰] رشد روزافزونی داشته است. همانطور که می‌دانیم در روش‌های چندکاناله بهسازی گفتار، با هدف ارتقای مقدار SNR^۲ و نیل به کیفیت بالاتر، با قبول افزایش پیچیدگی محاسباتی، از دو یا چند میکروفون برای ضبط صدا استفاده می‌شود؛ به‌طوری‌که می‌توان با چینش مشخصی از میکروفون‌ها و پردازش روی خروجی آنها در حوزه زمان-فضا، میزان تخریب گفتار در اثر نویزهای محیطی را کاهش داد. به‌عنوان مثال، با اعمال فیلترینگ فضایی توسط شکل‌دهنده‌های پرتو^۳، می‌توان به الگوی تشعشعی^۴ دلخواهی متناسب با طیف فرکانسی سیگنال‌ها دست یافت [۵-۸]. در این ارتباط، برخی مراجع از آرایه‌های خطی و برخی دیگر، از آرایه‌های دو یا سه‌بعدی بهره می‌گیرند.

در [۵]، مؤلفان با فرض نویز جهت‌دار و با بهره‌گیری از آرایه خطی با چیدمان غیریکنواخت، به طراحی شکل‌دهنده‌های CDB^۵ مبتنی بر الگوریتم وقتی CLMS^۶ می‌پردازند. طرح پیشنهادی آنها نه تنها توان نویزهای گذرا را در خروجی سیستم کاهش می‌دهد، بلکه با ایجاد صفرهای وقتی در الگوی تشعشعی با نویزهای جهت‌دار محیط نیز به‌شکل کارآمدی مقابله می‌کند.

مرجع [۸]، با فرض یک مدل چند-گوینده‌ای، به استفاده از شکل‌دهنده‌های پرتو LMCV^۷ در بهسازی گفتار در شبکه‌های حسگر صوتی بی‌سیم می‌پردازد. آقای تقی‌زاده و همکاران، ضمن بررسی میزان همبستگی در نویز پخشنده برای جفت میکروفون‌های یک آرایه میکروفونی، به ارزیابی مدل پیشنهادی خود هم در شرایط واقعی و هم به کمک شبیه‌سازی می‌پردازند [۱۱]. در [۱۲]، با فرض نویز پخشنده، یک فیلتر پسینه ترکیبی مبتنی بر فیلتر Zelinski اصلاح شده و فیلتر وینر پیشنهاد می‌شود. این طرح، با افراز باند فرکانسی به زیرباندهای مقتضی، همزمان نویزهای همبسته و ناهمبسته محیط را به‌شکل کارآمدی حذف می‌نماید. در [۱۳]، فیلتر پسینه روی آرایه‌ای از میکروفون‌ها پیشنهاد می‌شود که نه تنها با نویزهای پخشنده و سفید مقابله می‌کند، بلکه تداخل‌های نقطه‌ای محیط را نیز حذف می‌کند. شکل‌دهنده‌های پرتو فوق جهتی^۸ [۶، ۱۴] دسته دیگری از شکل‌دهنده‌های باندوسیع می‌باشند که با تنظیم وقتی ضرائب وزندهی در هر کانال مقدار بهره آرایه را در فرکانس‌های مختلف بیشینه می‌کنند. به‌منظور استفاده از آرایه‌های سه بعدی، مرجع [۱۵] با استفاده از آرایه کروی ۳۲ میکروفونه، شکل‌دهنده پرتویی طراحی نموده است که با تداخل‌های فضایی محیط مقابله می‌کند. طرح پیشنهادی با حذف لوب‌های جزئی و نامطلوب الگوی تشعشعی و بدون نیاز به تغییر پیکربندی فیزیکی آرایه میکروفون‌ها نقش مؤثری در امر بهسازی سیگنال گفتار ایفا می‌کند. استفاده از آرایه‌های کروی، امکان پردازش سه بعدی میدان‌های صوتی و حذف مؤثرتر نویز و تداخل را مهیا می‌کند. در [۱۶] با بهره‌گیری از یک آرایه میکروفونی به‌شکل کریستال راهکار جدیدی برای مقابله با نویزهای پخشنده محیط ارائه می‌شود. طرح پیشنهادی به کمک شکل‌دهنده پرتو MVDR^۹ و فیلتر وینر پسینه، تخمین LMMSE^{۱۰} مناسبی از سیگنال هدف می‌زند.

در راستای بهره‌گیری از فیلترهای وقتی در امر حذف نویز، مرجع [۱۷] ضمن بررسی ضعف‌های الگوریتم LMS در مقابله با نویزهای ضربه‌ای، الگوریتم وقتی جهت بهبود کارایی الگوریتم LMS ارائه می‌کند. بدین منظور، مؤلفان از معیاری مبتنی بر تئوری اطلاعات به نام کرانتروپی در تابع هزینه الگوریتم پیشنهادی خود بهره می‌گیرند. همچنین، آن‌ها برای مقابله با تنگی پاسخ ضربه کانال، از تقریب نرم صفر در تابع هزینه استفاده می‌نمایند. در مرجع [۱۸]، در راستای کاربرد حذف نویز از گفتار، یک پیش سفیدکننده جدیدی با عنوان PW-^{۱۱} TDLMS مبتنی بر نسخه متداول TDLMS ارائه می‌شود. طرح پیشنهادی مبتنی بر یک فیلتر ناهمبسته FIR ساده مرتبه اول است.

نهایت، در بخش ششم، نتیجه گیری و جمع بندی کلی از مسأله ارائه خواهد شد.



شکل (۱): شمای کلی از ساختار مورد استفاده

۲- مدل سازی و تعریف مسئله

بلوک دیاگرام نمایش داده شده در شکل (۱) را با فرض شرایط میدان دور^{۱۶} در نظر بگیرید. در این شما، سیگنال های دریافتی توسط هر یک از M میکروفون، یعنی $Y_i(f)$ ، ابتدا توسط یک فیلتر وقتی FIR^{۱۷}، وزن دهی و سپس با هم ترکیب می شوند. لذا، برای خروجی شکل دهنده پرتو $Z(f)$ در هر جزء فرکانسی^{۱۸} می توان نوشت:

$$Z(f) = \underline{W}^H(f) \underline{Y}(f), \quad (1)$$

که $\underline{W}(f)$ و $\underline{Y}(f)$ به ترتیب بردار دریافتی توسط آرایه و بردار وزن دهی شکل دهنده پرتو در هر جزء فرکانسی و $[.]^T$ عملگر ترانپوز می باشد:

$$\underline{Y}(f) = [Y_0(f) \ Y_1(f) \ \dots \ Y_{M-1}(f)]^T, \quad (2)$$

$$\underline{W}(f) = [W_0(f) \ W_1(f) \ \dots \ W_{M-1}(f)]^T. \quad (3)$$

هدف، حصول سیگنال $S(f)$ به واسطه محاسبه وقتی فیلترهای وزن دهی شکل دهنده پرتو با در نظر گرفتن یک معیار هزینه مناسب و حداقل سازی خطای حاصل است. اگر پاسخ فرکانسی کانال بین میکروفون m ام تا منبع گفتار هدف را $H_m(f)$ فرض کنیم، سیگنال دریافتی توسط این میکروفون برابر خواهد بود با:

$$Y_m(f) = H_m(f) S(f) + N_m(f), \quad (4)$$

که در آن، $N_m(f)$ برآیند سیگنال های ناخواسته موجود در محیط (شامل انواع نویزهای جهت دار، همه جهته و پخشنده) است.

با تعریف بردار پاسخ فرکانسی کانال بین میکروفون ها و منبع گفتار هدف $(\underline{h}(f))$ و با فرض ناهمبستگی مؤلفه نویز و سیگنال تمیز و همچنین، با تعریف چگالی طیف توان سیگنال تمیز $(R_{ss}(f))$ و ماتریس همبستگی بردار ورودی در حوزه فرکانس $(R_{YY}(f))$:

$$\underline{h}(f) = [H_0(f), H_1(f), \dots, H_{M-1}(f)]^T, \quad (5)$$

در مقاله [۱۹]، مؤلفان به بهبود عملکرد سیستم جداکننده گفتار دوگوشی با ارائه سیستم هایی موسوم به MESSL^{۱۲} می پردازند. در روش پیشنهادی اول، حذف وقتی نویز با استفاده از بهینه سازی ازدحام ذرات بر مبنای یادگیری LPSO^{۱۳} محقق می شود. سیستم پس پردازش پیشنهادی دوم شامل دو مرحله است. در مرحله اول آن، از روش حذف نویز تبدیل موجک به منظور حذف بخش اعظم سیگنال تداخل استفاده می شود. در مرحله دوم، روش حداقل میانگین مربع خطا (MMSE) جهت ارتقا هرچه بیش تر کیفیت سیگنال هدف جدا شده به کار می رود. اگرچه مسأله بهسازی گفتار، تحقیقات متنوعی به خود دیده است، اما محققان کمتر در محیط های واقعی به این مهم پرداخته اند. داشتن یک شناخت اولیه از میدان های نویزی حاکم بر محیط و بررسی اثر هر یک بر روی گفتار، گام مؤثری در جهت فرآیند بهسازی محسوب می شود. از مهمترین مراجع مرتبط با مقاله حاضر، می توان به [۵، ۱۱، ۱۲] اشاره کرد. توانایی مقابله همزمان و وقتی با انواع نویزهای جهت دار و پخشنده، همبسته و ناهمبسته از عمده برتری های این مقاله در مقایسه با کارهای قبلی به شمار می آید. در مقاله حاضر، ضمن بررسی انواع میدان های نویزی در شرایط واقعی، با لحاظ کردن شرایط پویای محیط و کنترل وقتی ضرایب وزن دهی شکل دهنده پرتو، ساختار مرکبی پیشنهاد می شود که نه تنها به ایجاد صف های کنترل شده در الگوی تشعشعی شکل دهنده پرتو برای مقابله با میدان های نویزی جهت دار محیط می پردازد بلکه به عنوان یک مزیت در مقایسه با [۵، ۱۱]، با اعمال فیلتر پسینه مناسب منجر به کاهش نویزهای ناهمبسته و پخشنده می شود. دیگر مزیت مقاله حاضر در مقایسه با [۵]، پیاده سازی در محیط واقعی است. برخلاف [۱۱-۱۲]، در طرح پیشنهادی، به منظور تحقق اهداف مطرح، دو الگوریتم حداقل میانگین مربع (LMS)^{۱۴} و سریع ترین کاهش (SD)^{۱۵} استفاده و اثر هر یک بر بهسازی سیگنال گفتار مقایسه خواهد شد. ملاحظات به عمل آمده باعث می شود که طرح پیشنهادی از توانایی بالایی در بهبود SNR و کیفیت سیگنال گفتار در شرایط واقعی برخوردار باشد.

ساختار مقاله بر چندین بخش استوار است: بخش دوم، به مدل سازی مسئله و معرفی توابع هزینه مورد استفاده می پردازد. در بخش سوم، ضمن بررسی میدان های نویزی پخشنده، تابع همدوسی فضایی آنها از نظر تئوری اثبات خواهد شد. طرح پیشنهادی مبتنی بر دو الگوریتم LMS و SD در بخش چهارم تشریح خواهد شد. در بخش پنجم، نتایج حاصل از پیاده سازی ساختار پیشنهادی، روی دو دسته دادگان شبیه سازی شده و دادگان واقعی به شکل مفصلی در قالب سناریوهای مختلف ارائه می گردد. در این راستا، ضمن بررسی نویزهای موجود در زمینه گفتار ضبط شده در شرایط واقعی، به ارزیابی میدان های نویزی پایین گذر پخشنده و مقایسه تابع همدوسی فضایی آنها در شرایط واقعی با نتایج تئوری خواهیم پرداخت. بررسی اثر سرعت همگرایی الگوریتم های مورد استفاده بر میزان SNR در حضور انواع میدان های نویزی با طیف های فرکانسی مختلف از دیگر سناریوهای مطرح در این بخش می باشد. در

حال، با استفاده از لم معکوس سازی [۲۲] بردار ضرایب وزن دهی بهینه به شکل رابطه (۱۲) بازنویسی می شود:

$$\underline{W}_o(f) = \frac{R_{NN}(f)^{-1} \underline{h}(f)}{\underline{h}^H(f) R_{NN}(f)^{-1} \underline{h}(f)} \times \frac{R_{ss}(f)}{R_{ss}(f) + R_{NN}(f)}, \quad (12)$$

که در آن $R_{ss}(f)$ و $R_{NN}(f)$ به ترتیب معرف چگالی طیف توان مؤلفه های تمیز و نویز در خروجی شکل دهنده پرتو هستند.

با مقایسه رابطه فوق با رابطه حاصل در بخش قبل، ملاحظه می شود که برای بهسازی مؤثرتر سیگنال گفتار، نیاز به ترکیبی از یک شکل دهنده پرتو و یک فیلتر پسینه^{۲۱} داریم. در حقیقت، شکل دهنده پرتو به حذف نویزهای جهت دار محیط پرداخته و فیلتر پسینه، ناکارآمدی شکل دهنده پرتو را در حذف نویزهای همه جهته پخشنده، نویزهای جهت دار فرکانس پایین و نویزهای جهت دار هم جهت با منبع هدف جبران می کند. بدین ترتیب، مسئله طراحی سیستم بهسازی سیگنال گفتار را می توان به دو بخش یعنی یک شکل دهنده پرتو وقتی و یک فیلتر پسینه بهینه افراز نمود. اگرچه در رابطه (۱۲) فیلتر پسینه به شکل یک فیلتر وینر وارد شده است، ولی می توان از سایر روش های تک کاناله بهسازی گفتار نیز به عنوان فیلتر پسینه بهره برد.

۳- تابع همدوسی فضایی میدان های نویزی پخشنده

در این بخش، به بررسی رفتار میدان های نویزی پخشنده می پردازیم. مطابق شکل (۱)، آرایه میکروفون خطی و یکنواختی را در حضور نویز پخشنده و در شرایط میدان دور در نظر بگیرید. در این شرایط، اختلاف دامنه سیگنال های دریافتی توسط میکروفون ها ناچیز بوده به طوری که می توان از آن چشم پوشی نمود؛ ولی به دلیل فاصله میکروفون ها، سیگنال منابع، با تأخیرهای زمانی متفاوت (τ_m) و در نتیجه، با اختلاف فاز توسط هر یک از میکروفون ها دریافت خواهند شد.

$$\tau_m = \frac{md \cos(\theta)}{c}, \quad m = 0, 1, \dots, M-1. \quad (13)$$

در این رابطه، θ زاویه امواج دریافتی با محور آرایه، d فاصله بین میکروفون ها، M تعداد میکروفون ها و c سرعت موج فیزیکی صوت است. با فرض دریافت سیگنال $x(t)$ توسط میکروفون مرجع m_0 ، برای سیگنال دریافتی توسط میکروفون m می توان نوشت:

$$y_m(t) = x(t - \frac{md}{c} \cos(\theta)), \quad m = 0, 1, \dots, M-1. \quad (14)$$

تابع همدوس فضایی، برای تعیین درجه همبستگی دو سیگنال X و Y در فرکانس های مختلف، به شکل زیر تعریف می شود [۲۳]:

$$\Gamma_{x,y}(f) = \frac{\phi_{xy}(f)}{\sqrt{\phi_{xx}(f)\phi_{yy}(f)}}, \quad (15)$$

که در آن، $\phi_{xy}(f)$ چگالی طیف توان متقابل دو سیگنال و $\phi_{xx}(f)$ و $\phi_{yy}(f)$ به ترتیب چگالی طیف توان سیگنال های x و y در فرکانس f می باشند؛ به بیان دیگر، $\Gamma_{x,y}(f)$ نرمالیزه شده دامنه چگالی طیف توان متقابل سیگنال های x و y توسط جذر حاصل ضرب چگالی طیف های این دو سیگنال است. ($|\Gamma_{x,y}(f)| \leq 1$). مسلماً، در یک فرکانس مشخص،

$$\underline{R}_{YY}(f) = E(\underline{Y}(f) \underline{Y}^H(f)) = R_{ss}(f) \underline{h}(f) \underline{h}^H(f) + \underline{R}_{NN}(f). \quad (6)$$

می توان مسئله بهینه سازی ضرایب وزن دهی را براساس یک معیار هزینه مناسب و حداقل سازی خطای مربوطه تعریف نمود. در ادامه، مسئله را با در نظر گرفتن دو تابع هزینه مختلف تحلیل می کنیم.

۲-۱- حداقل سازی توان خروجی شکل دهنده پرتو

با تعریف توان خروجی شکل دهنده پرتو به عنوان تابع هزینه $\underline{W}(f) \underline{R}_{YY}(f) \underline{W}^H(f)$ ، ضمن حفظ بهره الگوی تشعشعی در جهت دید^{۲۱}، مسئله بهینه سازی مقید، به شکل زیر تعریف می شود:

$$\begin{aligned} \underline{W}_o(f) &= \arg \min_{\underline{W}(f)} (\underline{W}^H(f) \underline{R}_{YY}(f) \underline{W}(f)) \\ \text{Subject to } \underline{W}^H(f) \underline{h}(f) &= 1, \end{aligned} \quad (7)$$

که با حل آن به روش ضرایب لاگرانژ [۲۰] می توان نوشت:

$$\underline{W}_o(f) = \frac{R_{NN}(f)^{-1} \underline{h}(f)}{\underline{h}^H(f) R_{NN}(f)^{-1} \underline{h}(f)}. \quad (8)$$

در این رابطه، $R_{NN}(f)$ ماتریس همبستگی مؤلفه نویز بردار ورودی است. مسلماً به منظور حصول فیلترهای بهینه در هر کانال، لازم است که با استفاده از یک فیلتر بانک مناسب، حل بهینه و مقید فوق در هر یک از اجزاء فرکانسی سیگنال های ورودی اعمال شود.

۲-۲- حداقل سازی خطای میانگین مربعی (MSE)

بهینه سازی بخش قبل، بر مبنای اختلاف زاویه دریافت سیگنال مطلوب و نویزهای جهت دار محیط استوار است. وجود نویز پخشنده در محیط های واقعی [۲]، نویزهای جهت دار فرکانس پایین و افزایش غیرعملی طول آرایه در حذف این دسته از نویزها، نویزهای جهت دار هم راستا با منبع هدف و نویزهای ناهمبسته، همگی بهسازی مؤثر شکل دهنده های پرتو حاصل از مسئله بهینه سازی فوق را با مشکل مواجه می کنند. برای بهسازی کارآمد در محیط های واقعی، به جای حداقل سازی توان نویز، به حداقل سازی خطای میانگین مربعی بین طیف خروجی شکل دهنده پرتو ($\underline{W}^H(f) \underline{Y}(f)$) و سیگنال مطلوب $S(f)$ می پردازیم:

$$\underline{W}_o(f) = \arg \min_{\underline{W}(f)} \left\{ \left| S(f) - \underline{W}^H(f) \underline{Y}(f) \right|^2 \right\}, \quad (9)$$

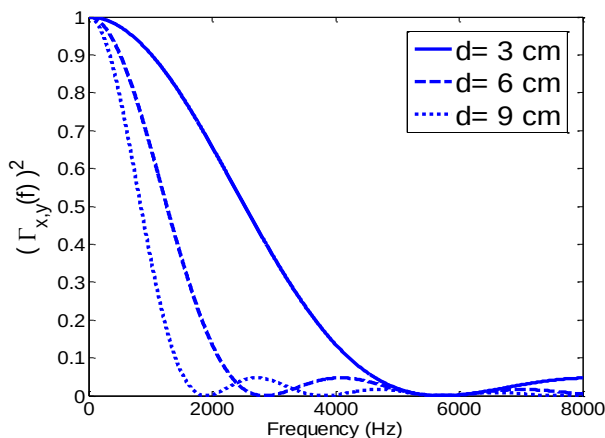
طبق معادلات وینر، با فرض ماتریس همبستگی بردار ورودی و بردار همبستگی متقابل بردار ورودی و سیگنال هدف ($r_{YS}(f)$) می توان ضرایب وزن دهی بهینه را به شکل زیر محاسبه کرد [۲۱]:

$$\underline{W}_o(f) = E\{\underline{Y}(f) \underline{Y}^H(f)\}^{-1} E\{\underline{Y}(f) S^*(f)\} = \underline{R}_{YY}^{-1}(f) r_{YS}(f) \quad (10)$$

با جایگذاری مقدار $\underline{R}_{YY}(f)$ از رابطه (۶) و با توجه به $r_{YS}(f) = \underline{R}_{SS}(f) \underline{h}(f)$ ، برای ضرایب وزن دهی بهینه می توان نوشت:

$$\underline{W}_o(f) = [R_{ss}(f) \underline{h}(f) \underline{h}^H(f) + \underline{R}_{NN}(f)]^{-1} R_{ss}(f) \underline{h}(f). \quad (11)$$

که در آن α پارامتر فراموشی می‌باشد. بدین ترتیب، با استفاده از رابطه فوق، می‌توان معکوس $\Phi_{NN}(f)$ را با محاسبات کمتری از مرتبه



شکل (۲): مربع همدوسی دو سیگنال در میدان نویزی پخشنده

$O(M^2)$ به روش لم معکوس‌سازی تخمین زد [۲۱]. این مهم، به‌خصوص در تعداد میکروفون بالا قابل تأمل است. این نکته را نیز باید در نظر گرفت که این محاسبه به‌خصوص وقتی میزان همبستگی بین مؤلفه‌های نویز بالا باشد، منجر به خطاهای گرد کردن می‌شود. از طرفی، متأسفانه ماتریس $R_{NN}(f)$ در فرکانس‌های پایین یک ماتریس بد وضع^{۲۲} است. از این رو، به‌عنوان یک روش جایگزین لازم است از الگوریتم‌های وقفی تکراری برای محاسبه پویا ضرایب استفاده کنیم. در این ارتباط، دو الگوریتم SD و LMS مطرح می‌باشد. الگوریتم SD مقایسه با الگوریتم LMS اگرچه به‌دلیل محاسبه مقدار دقیق گرادیان، بار محاسباتی بیشتری دارد، ولی، همگرایی آن سریع‌تر بوده و مهمتر آنکه به‌دلیل بهینه‌سازی ضرایب وزن دهی در هر فریم، در همه زمان‌ها الگوی تشعشعی بهتری می‌دهد. حال آنکه در الگوریتم LMS، با وجود سادگی و حجم محاسباتی کمتر، به‌دلیل اجرای تکرارهای الگوریتم به‌صورت بلوک به بلوک، در ابتدا ضرایب وزن دهی مناسب نیست و با ضرایب بهینه وینر، فاصله دارد. در نتیجه، شکل‌دهنده پرتو حاصل، الگوی تشعشعی مناسبی نداشته و طبیعتاً، بهسازی سیگنال گفتار به‌خوبی انجام نمی‌شود. ولی، به مرور بردار ضرایب وزن دهی به سمت بردار بهینه همگرا و بهسازی به‌شکل مؤثرتری اجرا خواهد شد.

در رابطه با بار محاسباتی اعمال فیلتر پسمینه نیز می‌توان گفت که با توجه فرآیند غیرتکراری این فیلتر و امکان پیاده سازی آن در حوزه فرکانس، حجم محاسبات در مقایسه با بخش شکل‌دهنده پرتو به مراتب کمتر است. لذا، در مجموع، بار محاسباتی کلی سیستم پیشنهادی در قسمت وقفی آن متمرکز شده است.

۴-۲ - شکل‌دهنده پرتو وقفی

در این بخش، به معرفی شکل‌دهنده‌های پرتو وقفی مبتنی بر دو الگوریتم سریع‌ترین کاهش و حداقل میانگین مربع می‌پردازیم.

هرچه این تابع به یک نزدیک‌تر باشد، دو سیگنال از درجه همبستگی بالاتری برخوردار خواهند بود. با استفاده از رابطه (۱۳) برای چگالی طیف توان متقابل سیگنال‌های x و y می‌توان نوشت:

$$\phi_{xy}(f) = \phi_{xx}(f) \{ \exp(-j(\frac{2\pi fmd}{c}) \cos(\theta)) \}. \quad (16)$$

از آنجایی که در میدان‌های نویز پخشنده، امواج با احتمال مساوی و انرژی یکنواخت در همه جهات منتشر می‌شوند، با در نظر گرفتن دستگاه مختصات کروی و متوسط‌گیری از رابطه (۱۶)، خواهیم داشت:

$$\phi_{xy}(f) = \frac{\phi_{xx}(f)}{4\pi} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \exp(-j(\frac{2\pi fmd}{c}) \cos(\theta)) \sin(\theta) d\theta d\phi, \quad (17)$$

$$= \frac{\phi_{xx}(f)}{2} \int_{\theta=0}^{\pi} \exp(-j(\frac{2\pi fmd}{c}) \cos(\theta)) \sin(\theta) d\theta. \quad (18)$$

حال، با تغییر متغیر $u = (\frac{2\pi fmd}{c}) \cos(\theta)$ می‌توان نوشت:

$$du = -(\frac{2\pi fmd}{c}) \sin(\theta) d\theta, \quad (19)$$

$$\phi_{xy}(f) = \frac{\phi_{xx}(f)}{2(2\pi fmd/c)} \int_{z=-2\pi fmd/c}^{2\pi fmd/c} \exp(-ju) du, \quad (20)$$

که با جایگذاری نتایج حاصل در رابطه (۱۵) می‌توان تابع همدوسی فضایی میدان‌های نویزی پخشنده را به‌شکل زیر ساده نمود:

$$\Gamma_{x,y}(f) = \text{sinc}(\frac{2fmd}{c}). \quad (21)$$

ملاحظه می‌شود از نظر تئوری، همدوسی فضایی نویز پخشنده، از تابع sinc تبعیت کرده به‌نحوی که با افزایش مقادیر f ، m و d ، همبستگی دو سیگنال کاهش می‌یابد. در شکل (۲)، مربع همدوسی سیگنال‌ها در میدان نویزی پخشنده، به‌ازای فرکانس‌های ۰ تا ۸ کیلوهرتز، در دو نقطه با فواصل ۳، ۶ و ۹ سانتی‌متر، مقایسه شده است.

۴- معرفی ساختار ترکیبی پیشنهادی

همانطور که در بخش (۲) بررسی شد، برای حصول ساختاری که به‌شکل مؤثر به امر بهسازی سیگنال گفتار بپردازد، نیاز به ترکیبی از یک شکل‌دهنده پرتو و یک فیلتر پسمینه داریم. در این بخش، ضمن ارزیابی میزان پیچیدگی محاسباتی مسأله پیش‌رو، به معرفی جامع طرح پیشنهادی مبتنی بر شکل‌دهنده پرتو وقفی و فیلتر پسمینه می‌پردازیم.

۴-۱- ارزیابی میزان پیچیدگی محاسباتی مسأله

بردار ضرایب وزن دهی بهینه در رابطه (۱۲) را در نظر بگیرید. محاسبه مستقیم معکوس ماتریس $R_{NN}(f)$ بسته به تعداد میکروفون‌های استفاده شده در آرایه، مستلزم پذیرش حجم محاسبات بسیار بالایی از مرتبه $O(M^3)$ است [۲۱]. از طرفی، در بسیاری از شرایط عملی، دسترسی مستقیم به این ماتریس امکان‌پذیر نمی‌باشد. یک راه حل ساده، استفاده از ساختار بازگشتی به فرم زیر است [۲۲]:

$$R_{NN}(f, l+1) = \alpha R_{NN}(f, l) + (1-\alpha) \underline{N}(f, l+1) \underline{N}^H(f, l+1) \quad (22)$$

۴-۲-۱- الگوریتم سریع ترین کاهش (SD)

الگوریتم SD از الگوریتم های مبتنی بر گرادینت^{۲۳} است که در هر دور تکرار، با حرکت در خلاف جهت گرادینت تابع هزینه (J) به حل بهینه وینر نزدیک می شود. در این مقاله، با تعریف تابع هزینه مقید به شکل:

$$J = \underline{W}^H(f) \underline{R}_{YY}(f) \underline{W}(f) + \lambda (\underline{W}^H(f) \underline{h} - 1), \quad (23)$$

برای ضرایب وزن دهی در هر مرحله از الگوریتم می توان نوشت:

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) + \mu \left[-\nabla_w(J(l)) \right], \quad (24)$$

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) - \mu \nabla_w \left[\underline{W}^H(f) \underline{R}_{YY}(f) \underline{W}(f) + \lambda (\underline{W}^H(f) \underline{h} - 1) \right], \quad (25)$$

که در آن μ اندازه گام و λ ضریب لاگرانژ است. حال، با استفاده از رابطه (۶) و محاسبه گرادینت تابع هزینه، رابطه (۲۵) را به شکل زیر بازنویسی می کنیم:

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) - \mu \left(\underline{\Phi}_{NN}(f) \underline{W}_l(f) + \lambda \underline{h}(f) \right), \quad (26)$$

با اعمال قید مسئله برای عبارت به دست آمده، می توان λ را محاسبه و در رابطه فوق جایگزاری نمود. در این صورت [۲۱]:

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) - \mu \left(\underline{I} - \frac{\underline{h}(f) \underline{h}^H(f)}{\|\underline{h}(f)\|^2} \right) \underline{\Phi}_{NN}(f) \underline{W}_l(f), \quad (27)$$

بدین ترتیب، با تخمین فریم به فریم $\underline{R}_{NN}(f)$ (با استفاده از رابطه (۱۳)) و جایگذاری آن در رابطه فوق، الگوریتم تکرار را برای هر فریم به طور مجزا اجرا می کنیم. از طرفی، با استفاده از بهینه سازی روی مقدار μ ، می توان سرعت همگرایی الگوریتم در هر فریم را افزایش داد:

$$\frac{\partial \left(\underline{W}_{l+1}^H \underline{\Phi}_{NN} \underline{W}_{l+1} \right)}{\partial \mu^*} = 0, \quad (28)$$

$$\mu = \frac{\underline{g}_l^H \underline{R}_{NN} \underline{W}_l(f)}{\underline{g}_l^H \underline{R}_{NN} \underline{g}_l}. \quad (29)$$

لذا، می توان الگوریتم سریع ترین کاهش را مطابق شکل (۳) خلاصه کرد. اضافه کردن پارامتر ε به عناصر قطر اصلی ماتریس همبستگی، معادل با افزودن این پارامتر به مقادیر ویژه این ماتریس بدون تغییر بردارهای ویژه آن می باشد. در واقع ما با این کار از بد وضع شدن ماتریس همبستگی در فرکانس های پایینی جلوگیری می کنیم.

۴-۲-۲ الگوریتم حداقل میانگین مربع (LMS)

الگوریتم حداقل میانگین مربع، الگوریتم تکراری است که در آن، برخلاف الگوریتم سریع ترین کاهش، به جای استفاده مستقیم از ماتریس همبستگی، از تخمین آن با استفاده از نمونه های آنی و به روز رسانی تخمین در هر فریم استفاده می شود:

$$\hat{\underline{R}}_{YY}(f) = E \left(\underline{Y}(f) \underline{Y}^H(f) \right) \cong \underline{Y}(f) \underline{Y}^H(f). \quad (30)$$

در اینجا نیز، با حرکت در خلاف جهت گرادینت تابع معیار J (تعریف شده در رابطه (۲۳))، در $l+1$ امین تکرار می توان نوشت:

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) - \mu \nabla_w \left[\underline{W}^H(f) \underline{Y}(f) \underline{Y}^H(f) \underline{W}(f) + \lambda (\underline{W}^H(f) \underline{h}(f) - 1) \right], \quad (31)$$

$$= \underline{W}_l(f) - \mu \left(\underline{Y}(f) \underline{Y}^H(f) \underline{W}(f) + \lambda \underline{h}(f) \right) \quad (32)$$

$$(1) \text{ به روز رسانی ماتریس } \underline{R}_{NN} \text{ در هر فریم مطابق رابطه (22)}$$

$$\tilde{\underline{R}}_{NN} = \underline{R}_{NN} + \varepsilon \underline{I} \quad (2)$$

$$\underline{W}_0(0) = \frac{\underline{h} \underline{h}^H}{\|\underline{h}\|^2} \quad (3)$$

$$\underline{g}_l = \left(\underline{I} - \frac{\underline{h} \underline{h}^H}{\|\underline{h}\|^2} \right) \tilde{\underline{R}}_{NN}(f) \underline{W}_l(f) \quad (4)$$

$$\mu_l = \frac{\underline{g}_l^H \tilde{\underline{R}}_{NN} \underline{W}_l(f)}{\underline{g}_l^H \tilde{\underline{R}}_{NN} \underline{g}_l} \quad (5)$$

$$\underline{W}_{l+1}(f) = \underline{W}_l(f) - \mu_l \underline{g}_l \quad (6)$$

شکل (۳): الگوریتم سریع ترین کاهش

بدین ترتیب:

$$\underline{W}_{l+1}(f) = \left(\underline{I} - \frac{\underline{h}(f) \underline{h}^H(f)}{\|\underline{h}(f)\|^2} \right) \left[\underline{W}_l(f) - \mu \underline{Z}_l \underline{X}_l \right] + \underline{W}_0(f) \quad (33)$$

توجه شود که در اینجا اندیس l شماره هر فریم است در حالی که در روش قبلی، l معرف تکرارهای الگوریتم وقتی در هر فریم بود.

۴-۳-۳ فیلتر پسینه

در تکمیل عملکرد شکل دهنده های پرتو، نیاز به سیستم مکملی با عنوان فیلتر پسینه داریم. در این مقاله، نظر به مدل پیشنهادی، از روش غیر بازگشتی Ephraim & Malah برای تخمین دامنه طیف به عنوان فیلتر پسینه استفاده می کنیم. الگوریتم مورد استفاده، با پنجره بندی سیگنال نویزی ورودی، از روش حداقل میانگین مربع خطا برای تخمین دامنه طیف سیگنال استفاده می کند. در این فرآیند تخمین، فرض می شود که ضرایب تبدیل فوریه سیگنال از نظر آماری مستقل هستند که با توجه به فرض توزیع گوسی، می توان آنها را ناهمبسته نیز قلمداد کرد. فرض کنید سیگنال خروجی شکل دهنده پرتو، تلفیقی از سیگنال تمیز $z_s(t)$ و نویز پس زمینه باقی مانده $z_n(t)$ باشد:

$$z(t) = z_s(t) + z_n(t). \quad (34)$$

با پنجره بندی خروجی شکل دهنده پرتو و گرفتن تبدیل فوریه از هر بخش، برای k امین جزء فرکانسی در l امین فریم می توان نوشت:

$$Z(k, l) = Z_s(k, l) + Z_n(k, l) \quad l = 1, 2, \dots, L, \quad (35)$$

که L تعداد کل پنجره ها است. حال، فرآیند پردازشی را به هر فریم معطوف می نماییم:

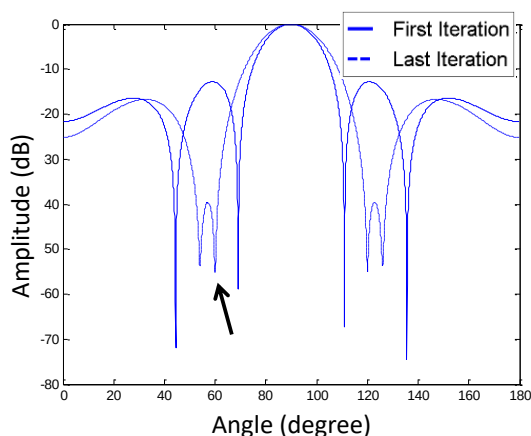
حذف نویز جهت‌دار به درستی صورت نگرفته است؛ به‌مرور با همگرایی الگوریتم و نزدیک‌تر شدن ضرایب وزن دهی به مقادیر بهینه وینر، صفری در زاویه ۶۰ درجه شکل گرفته و نویز جهت‌دار محیط به‌خوبی حذف می‌شود. این در حالی است که هر چه مقدار μ کوچکتر انتخاب شود سرعت همگرایی الگوریتم کاهش یافته و ضرایب وزن دهی دیرتر به مقدار بهینه خود می‌رسند. در واقع، رسیدن به حداقل تابع هزینه با آهنگ کندتری صورت می‌گیرد و فرآیند شکل‌گیری صفر در زاویه نویز جهت‌دار کندتر خواهد شد. از این رو، نویز جهت‌دار محیط با کیفیت پایین‌تری حذف می‌شود. این مسئله را در شکل (۶) با مقایسه نمایش زمانی و اسپکتروگرام سیگنال خروجی شکل دهنده پرتو به ترتیب برای برابر با ۰/۰۰۵، ۰/۰۰۱ و ۰/۰۱ می‌توان ارزیابی کرد.

شکل (۷) توان نویز جهت‌دار محیط را در خروجی شکل‌دهنده پرتو با فرض μ برابر با ۰/۰۰۰۱، برای یکبار اجرای الگوریتم و همچنین، میانگین‌گیری شده آن پس از ۱۰۰ بار اجرای الگوریتم نمایش می‌دهد. در LMS به دلیل تخمین ماتریس همبستگی هر فریم با استفاده از مقادیر آنی بردار ورودی آن فریم، منحنی یادگیری حاصل از یک دور تکرار دارای تغییرات شدیدی است. با میان‌گیری نتایج چندین بار اجرای الگوریتم، مثل آن است که اثر امید ریاضی را در مسئله اعمال کرده باشیم و منحنی یادگیری هموارتری به دست می‌آید.

در آزمایش دیگری به اثر اندازه گام بر منحنی یادگیری می‌پردازیم. در شکل (۸) می‌توان نتایج را یکبار با فرض نویز جهت‌دار سینوسی و بار دیگر، با فرض نویز جهت‌دار سفید مشاهده نمود. اگر چه افزایش μ منجر به افزایش سرعت همگرایی خواهد شد، ولی این مسئله از حد مشخصی به بعد، به واگرایی الگوریتم و عدم بهسازی کارآمد منجر می‌شود. همچنین، در شکل‌های (۹-الف) و (۹-ب) تغییرات الگوی تشعشی در طول زمان با فرض μ برابر با ۰/۰۰۱ و ۰/۱ مقایسه شده است.

جدول (۱): پارامترهای آزمایش روی دادگان شبیه‌سازی شده

α	نرخ نمونه برداری	مدت سیگنال (ثانیه)	d (cm)	M
۰/۹۸	۱۶۰۰۰	۶	۳ و ۵	۸



شکل (۴): مقایسه الگوی تشعشی حاصل در فرکانس ۴۰۰۰ هرتز در

تکرارهای ابتدایی و پس از همگرایی الگوریتم LMS

$$Z_k = |Z_k| \exp(j \angle Z_k), \quad (36)$$

$$Z_{S_k} = |Z_{S_k}| \exp(j \angle Z_{S_k}). \quad (37)$$

در این روابط، $|Z_k|$ و $|Z_{S_k}|$ به ترتیب STSA^{۲۴} سیگنال نویزی و مؤلفه تمیز آن می‌باشند. هدف اصلی محاسبه تخمینی از $|Z_{S_k}|$ (یعنی $|\hat{Z}_{S_k}|$) به‌نحوی است که تابع هزینه $E\{(|Z_{S_k}| - |\hat{Z}_{S_k}|)^2\}$ حداقل شود. [۲۴] با استفاده از رابطه وینر (۱۲)، نشان داده است که در SNR های بالا می‌توان از رابطه زیر برای تخمین $|\hat{Y}_{S_k}|$ بهره گرفت:

$$|\hat{Y}_{S_k}| = \frac{\xi_k}{1 + \xi_k} |Y_k|, \quad (38)$$

که در آن به ξ اصطلاحاً SNR پیشینه^{۲۵} گویند و در قالب یک ساختار بازگشتی برحسب SNR پسینه^{۲۶} $\gamma(k, l)$ تخمین زده می‌شود [۲۴]:

$$\hat{\xi}(k, l+1) = \beta G_k^2(l) \gamma(k, l) + (1 - \beta)(\gamma(k, l+1) - 1), \quad (39)$$

در این رابطه، β پارامتر فراموشی است که مقداری بین صفر و یک دارد. با عنایت به رابطه $1/(1 - \beta)$ برای حافظه سیستم، هر چه مقدار β به ۱ نزدیک‌تر باشد حافظه سیستم بیشتر شده و فریم‌های قبلی تأثیر بیشتری در تخمین فریم موردنظر خواهند داشت.

۵- شبیه‌سازی

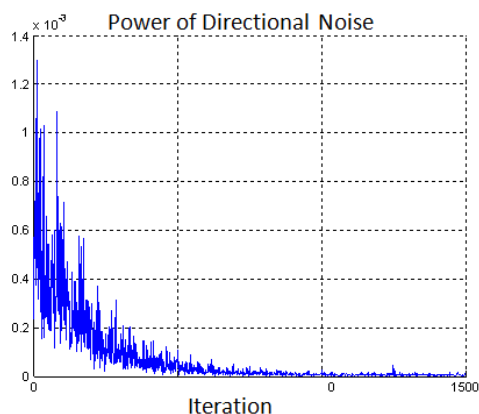
در این بخش، به‌منظور بررسی عملکرد ساختارهای پیشنهادی، به پیاده‌سازی آنها روی دو دسته دادگان شبیه‌سازی شده و دادگان واقعی ضبط شده در آزمایشگاه پردازش گفتار دانشگاه یزد [۲۵] می‌پردازیم.

۵-۱- آزمایش روی دادگان شبیه‌سازی شده

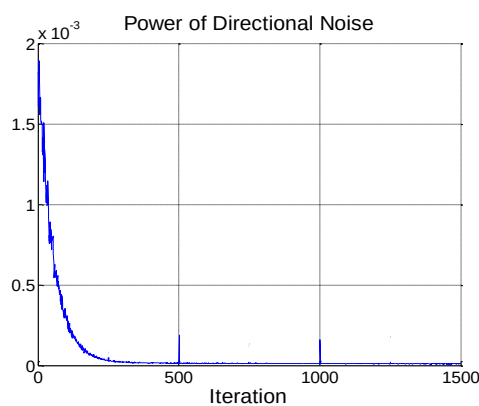
برای انجام شبیه‌سازی‌های این بخش، از سیگنال‌های استخراج شده از مجموعه دادگان فارس‌دات [۲۶] و NoiseX [۲۷]، به‌ترتیب به‌عنوان سیگنال هدف و نویز محیط استفاده می‌کنیم. در این راستا، با فرض منبع هدفی در زاویه ۹۰ درجه نسبت به محور آرایه به‌عنوان گفتار مطلوب و نویز جهت‌داری در زاویه ۶۰ درجه به اجرای ساختار ترکیبی پیشنهادی می‌پردازیم. همچنین، از نویز سفید گوسی (WGN^{۲۷}) با توان 30dB- روی همه میکروفون‌ها به‌منظور مدل‌کردن نویز ناهمبسته محیط بهره گرفته می‌شود. در جدول (۱) می‌توان مشخصات کلی پارامترهای مورد استفاده در شبیه‌سازی را مشاهده کرد.

سناریو ۱: در این سناریو، با فرض نویز جهت‌دار ثابتی با پهنای باند 1kHz حول فرکانس 4kHz، به ارزیابی عملکرد شکل‌دهنده پرتو وقتی مبتنی بر الگوریتم LMS، می‌پردازیم. در شکل (۴) می‌توان الگوی تشعشی حاصل از ضرایب وزن دهی به‌دست آمده را در فرکانس ۴۰۰۰ هرتز در تکرارهای ابتدایی و پس از همگرایی الگوریتم و با فرض μ برابر با ۰/۰۰۱ مقایسه کرد. همچنین، در شکل‌های (۵-الف) و (۵-ب)، با مقایسه نمایش زمانی و اسپکتروگرام سیگنال تمیز، سیگنال نویزی میکروفون مرجع و خروجی شکل‌دهنده پرتو، اثر الگوی تشعشی حاصل در بهسازی سیگنال گفتار نشان داده شده است. در ابتدا، به‌دلیل بهینه نبودن ضرایب وزن دهی شکل‌دهنده پرتو،

ملاحظه می‌شود که با افزایش اندازه گام، دامنه الگوی تشعشعی یک روند نوسانی پیدا می‌کند. شکل (۱۰)، بهره صفر تولیدی در زاویه نویز جهت‌دار محیط را در تکرارهای مختلف الگوریتم و به ازای مقادیر مختلف μ مقایسه می‌کند. بهره الگوی تشعشعی در این زاویه، یک روند نزولی داشته و به مرور با سرعت مشخصی کاهش می‌یابد. در حقیقت، این پدیده، نمایشی از همگرایی الگوریتم و تولید تدریجی صفر در زاویه نویز جهت‌دار محیط است. ولی در اینجا ما یک مصالحه داریم.



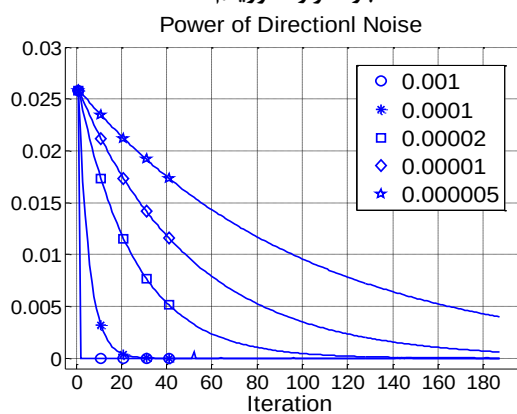
(الف)



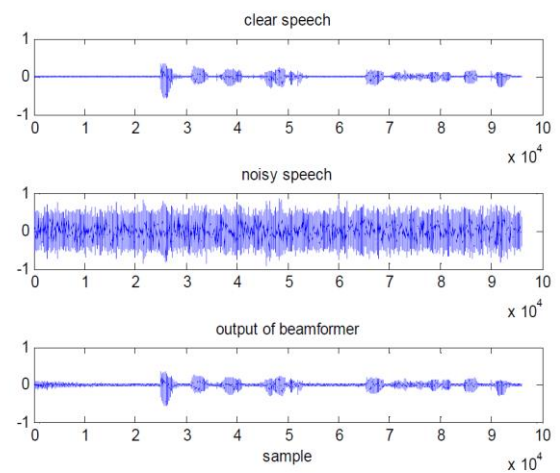
(ب)

شکل (۷): توان نویز جهت‌دار در خروجی شکل دهنده پرتو با فرض μ برابر با ۰/۰۰۰۱، (الف) یک‌بار تکرار (ب) میانگین‌گیری شده آن پس از

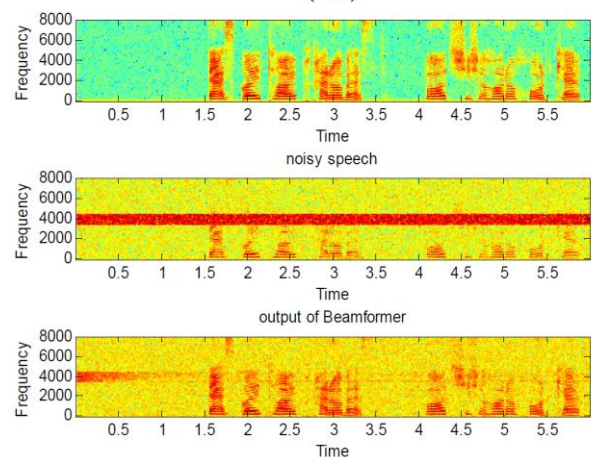
۱۰۰ بار تکرار الگوریتم LMS



(الف)

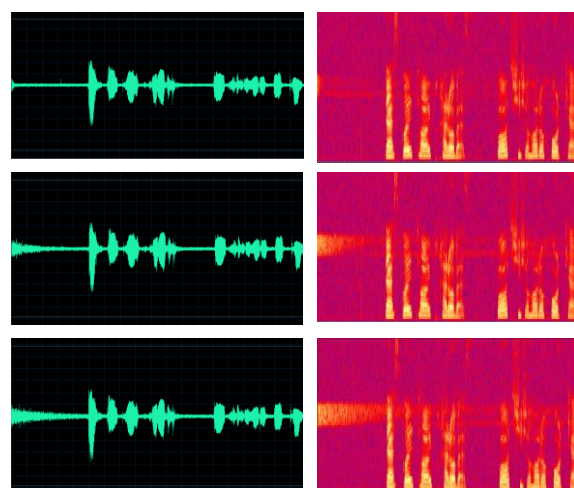


(الف)



(ب)

شکل (۵): مقایسه (الف) نمایش زمانی و (ب) اسپکتروگرام سیگنال‌های تمیز، سیگنال نویزی میکروفون مرجع و خروجی شکل دهنده پرتو LMS با فرض μ برابر با ۰/۰۰۱

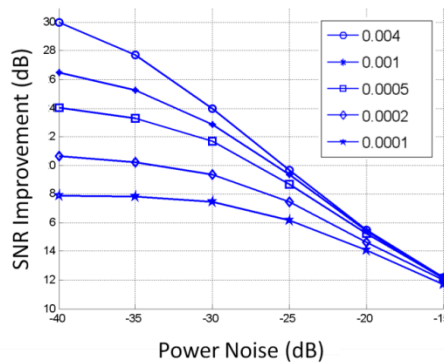


(ب)

(الف)

شکل (۶): (الف) اسپکتروگرام و (ب) نمایش زمانی سیگنال خروجی شکل دهنده پرتو LMS با فرض μ برابر با ۰/۰۰۰۵، ۰/۰۰۱ و ۰/۰۱

هر چه اندازه گام بزرگتر انتخاب شود سرعت همگرایی الگوریتم و تشکیل صفر بالا رفته ولی در مقابل میزان نوسانات دامنه الگوی تشعشعی نیز افزایش می‌یابد. در جدول (۲) می‌توان میزان بهبود SNR نویزهای جهت دار (I_{coh}) ، ناهمبسته (I_{uncoh}) و نویز کل (I) را به ازای مقادیر مختلف μ یکبار با فرض نویز جهت دار سینوسی و یکبار با فرض نویز جهت دار سفید با یکدیگر مقایسه کرد.



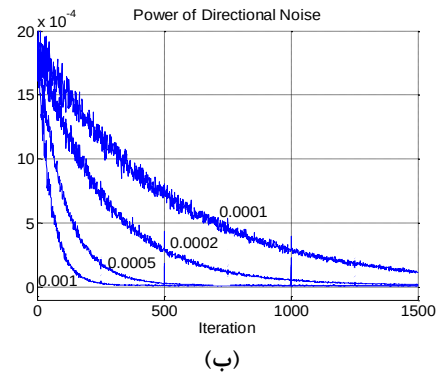
شکل (۱۱): اثر توان نویز ناهمبسته بر میزان بهبود SNR به ازای مقادیر مختلف μ

با افزایش μ ، هر چه سرعت همگرایی ضرایب به سمت مقادیر بهینه افزایش یابد، شکل دهنده پرتو از حالت بهینه برای حذف نویز ناهمبسته محیط $(\text{D}^{\text{S}}\text{B})$ با سرعت بیشتری فاصله گرفته و میزان بهبود SNR این نویز کاهش می‌یابد. در این فرآیند، هر چه توان نویز ناهمبسته میکروفون‌ها قوی‌تر باشد، این نویز، بر نویز جهت دار غالب‌تر شده و الگوریتم مفروض کارایی کمتری پیدا می‌کند. در نتیجه، میزان بهبود SNR کاهش می‌یابد. در شکل (۱۱)، اثر توان نویز ناهمبسته بر میزان بهبود SNR به ازای مقادیر مختلف μ نشان داده شده است.

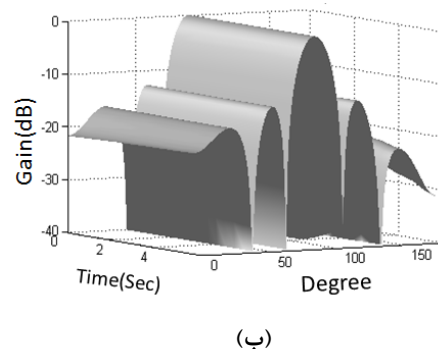
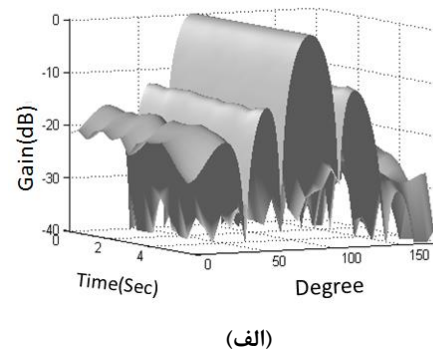
سناریو ۲- حال، به بررسی ناکارآمدی‌های شکل دهنده پرتو در حذف برخی از نویزهای محیط و جبران آن به وسیله فیلتر پسینه می‌پردازیم. شرایط توصیف شده در سناریوی قبلی را در نظر بگیرید. فرض کنید علاوه بر نویز موجود در زاویه ۶۰ درجه، نویز جهت دار دیگری نیز، در جهت دید (یا نزدیک به آن) قرار داشته باشد. به دلایل توصیف شده، شکل دهنده‌های پرتو مبتنی بر آرایه‌های یک بعدی، قادر به حذف این نویز نخواهد بود؛ از این رو، از بلوک مکمل فیلتر پسینه برای حذف آن استفاده می‌کنیم. در جدول (۳) می‌توان SNR خروجی شکل دهنده پرتو و فیلتر پسینه را به ازای نویزهای جهت دار مختلف هم جهت با سیگنال تمیز، با فرض μ برابر با ۰/۰۰۱ مشاهده کرد. همچنین، تغییرات زمانی و اسپکتروگرام سیگنال‌های تمیز، سیگنال نویزی میکروفون مرجع و خروجی شکل دهنده پرتو و خروجی فیلتر پسینه با فرض نویز هم جهت صورتی در شکل (۱۲) نمایش داده شده است.

جدول (۲): میزان بهبود SNR نویزهای جهت دار، ناهمبسته و نویز کل

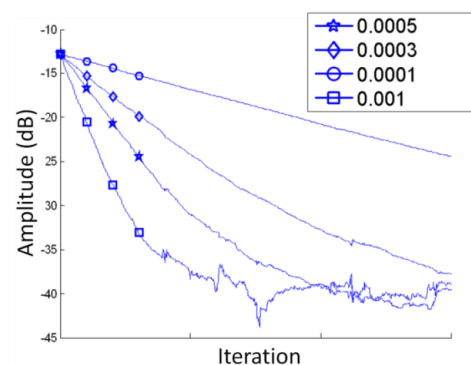
I	I_{uncoh}	I_{coh}	μ	نوع نویز جهت دار
۲۴/۰۰۵	۸/۶۰۶۲	۳۲/۰۵۲۰	۰/۰۰۴۰	۱. سیگنال
۲۲/۷۵۷۱	۸/۶۵۱۳	۲۷/۰۷۱۶	۰/۰۰۱۰	۲. نویز
۲۱/۶۹۸۴	۸/۶۵۷۸	۲۴/۴۶	۰/۰۰۰۵	۳. نویز



شکل (۸): اثر اندازه گام بر منحنی یادگیری با فرض (الف) نویز جهت دار سفید (ب) نویز جهت دار سینوسی

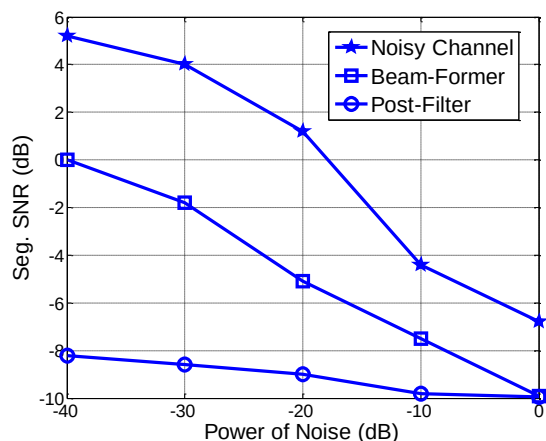


شکل (۹): تغییرات بهره الگوی تشعشعی در تکرارهای مختلف الگوریتم با فرض μ برابر با (الف) ۰/۰۰۱ (ب) ۰/۰۰۱

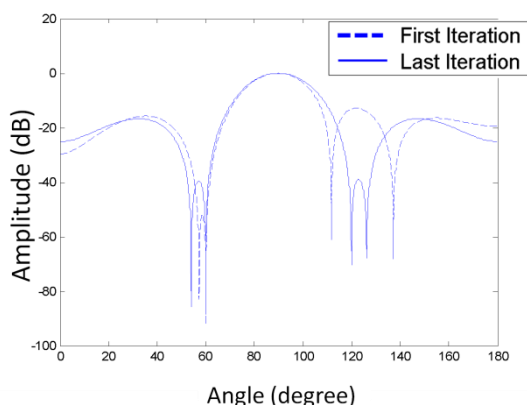


شکل (۱۰): نمایش بهره صفر تولیدی در زاویه نویز جهت دار به ازای مقادیر مختلف μ

مشاهده می‌شود که در شرایط توصیف شده، شکل دهنده پرتو نتوانسته نویز صورتی را به خوبی حذف کند ولی با اعمال فیلتر پسینه در تکمیل واحد قبلی، بهسازی به شکل مؤثرتری صورت گرفته است. همچنین، شکل (۱۳) اثر μ را در SNR خروجی شکل دهنده پرتو و فیلتر پسینه نمایش می‌دهد. دو نتیجه استخراج می‌شود: اول اینکه فیلتر پسینه، مکمل مناسبی برای شکل دهنده پرتو بوده و با حذف نویزهای پس‌مانده، منجر به ارتقاء میزان SNR خواهد شد؛ دوم آنکه



شکل (۱۴): اثر توان نویز ناهمبسته محیط بر عملکرد سیستم



شکل (۱۵): مقایسه الگوی تشعشی حاصل در فرکانس ۴۰۰۰ هرتز در تکرارهای ابتدایی و پس از همگرایی الگوریتم سریع‌ترین کاهش به دلایل بحث شده، افزایش μ منجر به افزایش SNR می‌شود. همانطور که دیده می‌شود، افزایش توان نویز ناهمبسته، اثر نامطلوبی بر عملکرد شکل دهنده پرتو می‌گذارد. در آزمایش دیگری اثر توان این نویز را بر عملکرد سیستم ارزیابی می‌کنیم. نتایج حاصل از این آزمایش در شکل (۱۴) نمایش داده شده است. واضح است که فیلتر پسینه در توان‌های مختلف نویز ناهمبسته، عملکرد تقریباً ثابتی نشان می‌دهد. شیب نزولی نمودار فیلتر پسینه به دلیل رفتار شکل دهنده پرتو است. فیلتر پسینه مورد استفاده، برخلاف شکل دهنده پرتو به نوع نویز حساس نیست بلکه با تخمین طیف نویز کل تک کانال، به بهسازی گفتار می‌پردازد.

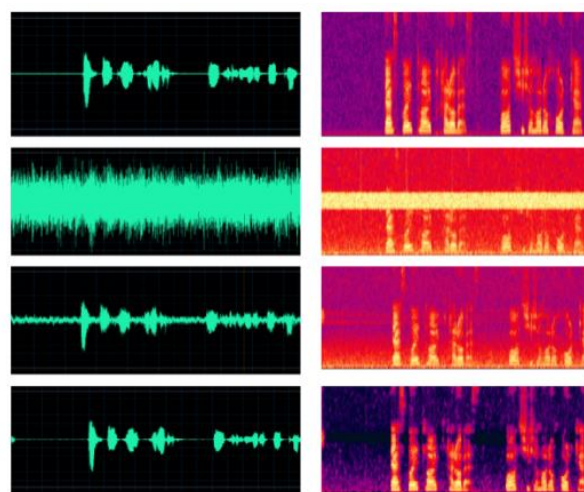
سناریو ۳- حال، پس از ارزیابی الگوریتم LMS، با فرض نویز جهت دار سفیدی با طیف 1k Hz در زاویه ۶۰ درجه، فریم‌های ۵۱۲ نمونه‌ای و $\varepsilon = 10^{-7}$ ، به ارزیابی ساختار پیشنهادی با استفاده از الگوریتم

۱۹/۴۷۳۸	۸/۸۳۳	۲۰/۸۱۰۱	۰/۰۰۰۲	مقایسه سیگنال نویز
۱۷/۳۹۲۳	۸/۹۴	۱۸/۱۴۱	۰/۰۰۰۱	
۳۲/۵۰۵۴	۸/۵۷	۳۵/۴۱۰۷	۰/۰۰۱۰	
۲۷/۶۳۸۴	۸/۶۴۵۴	۲۸/۳۷۴۷	۰/۰۰۰۱	
۲۱/۴۲۷۶	۸/۷۴۸۸	۲۱/۵۸۴۳	۰/۰۰۰۰۲	
۱۸/۶۲۰۷	۸/۹۱	۱۸/۶۹۴۵	۰/۰۰۰۰۱	
۱۶/۲۷۵۲	۸/۹۶	۱۶/۳۱۱۷	۰/۰۰۰۰۰۵	

جدول (۳): مقایسه SNR سیگنال خروجی شکل دهنده پرتو LMS و

فیلتر پسینه به ازای نویزهای جهت‌دار مختلف

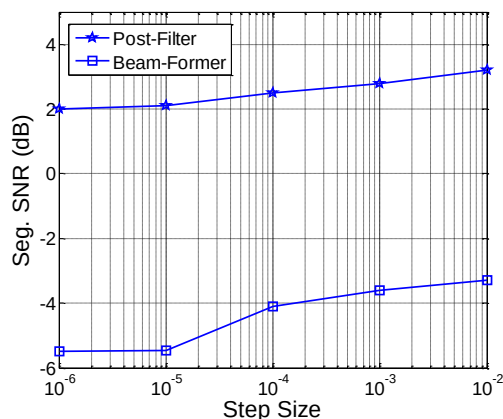
فیلتر پسینه	شکل دهنده پرتو	نوع نویز
-۲/۱۷	-۷/۷۶	قهوه‌ای
-۲/۰۵	-۶/۴	همه‌م
۳/۲۷	-۳/۵	صورتی
۴/۰۰۸	-۱/۵۷	سفید



(ب)

(الف)

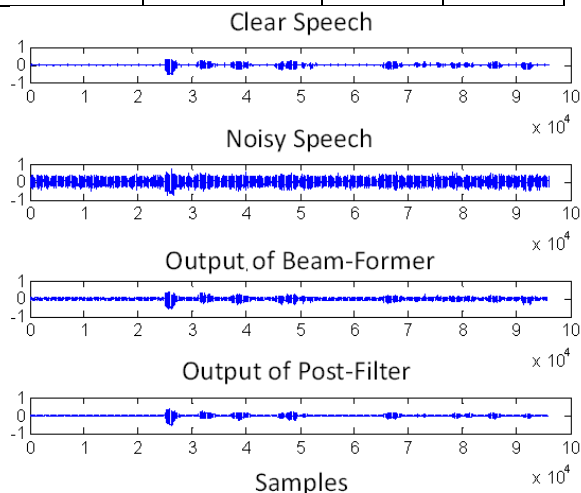
شکل (۱۶): (الف) اسپکتروگرام (ب) نمایش زمانی سیگنال تمیز، سیگنال نویزی میکروفون مرجع، خروجی شکل دهنده پرتو و خروجی فیلتر پسینه با فرض نویز هم جهت صورتی



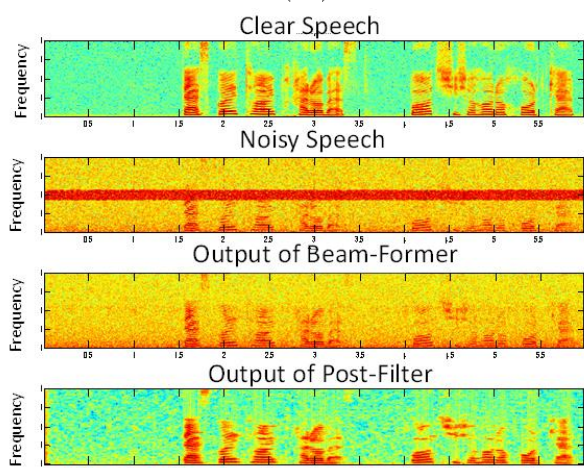
شکل (۱۷): مقایسه تغییرات Seg. SNR خروجی شکل دهنده پرتو و خروجی فیلتر پسینه با تغییر μ

جدول (۴): SNR سیگنال خروجی شکل دهنده پرتو SD و فیلتر پسینه به ازای نویزهای جهت دار مختلف

نویز	ورودی	شکل دهنده پرتو	فیلتر پسینه
نویز سفید	-۷/۳۸	-۴/۰۷	۲/۶۲
نویز صورتی	-۷/۶۸	-۳/۹۰	۲/۴۰
نویز همهمه	-۸/۳۲	-۶/۸۶	-۲/۰۱
نویز قهوه ای	-۸/۲۷	-۷/۸۰	-۲/۷۹



(الف)



(ب)

شکل (۱۷): مقایسه (الف) نمایش زمانی (ب) اسپکتروگرام سیگنال تمیز، سیگنال نویزی میکروفون مرجع، خروجی شکل دهنده پرتو SD و خروجی فیلتر پسینه با فرض نویز هم جهت صورتی

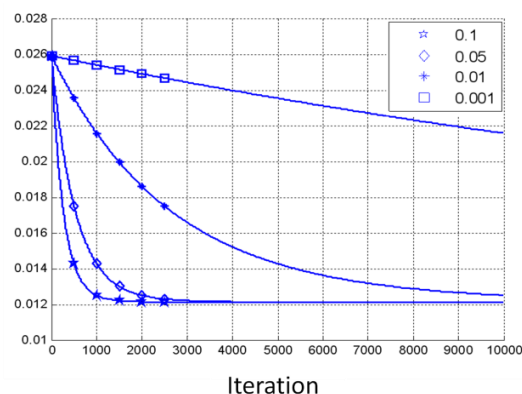
جدول (۵): SNR خروجی شکل دهنده پرتو و فیلتر پسینه

SNR ورودی	۱۵	۱۰	۵	۰	-۵	-۱۰
SNR (BF)	-۰/۴۹۹	-۱/۶۴	-۳/۴۸	-۵/۲۵	-۶/۷	-۸/۰۹
SNR (PF)	۵/۱۵۷	۴/۶	۳/۲۷	۱/۴۹	۱/۶۹	-۴/۶

۵-۲- آزمایش روی دادگان واقعی

ضبط دادگان واقعی، در آزمایشگاه پردازش گفتار دانشگاه یزد [۲۵] (اتاقی با ابعاد ۵/۷×۳/۵ متر مکعب) صورت پذیرفت. بدین منظور، ما از آرایه ای ۸ میکروفون با چیدمان خطی یکنواخت در فواصل ۵ cm و

سریع ترین کاهش می پردازیم. در شکل (۱۵) می توان الگوی تشعشی حاصل از ضرایب وزن دهی را در فرکانس ۴۰۰۰ هرتز برای تکرارهای ابتدایی و پس از همگرایی الگوریتم ملاحظه کرد. به منظور بررسی اثر اندازه گام بر سرعت همگرایی طرح پیشنهادی، در آزمایش دیگری بدون در نظر گرفتن مقدار بهینه μ در هر فریم، از مقادیر مختلفی برای آن استفاده نمودیم.



شکل (۱۶): تغییرات توان نویز جهت دار با فرض مقادیر مختلف برای μ در الگوریتم سریع ترین کاهش

در شکل (۱۶) می توان تغییرات توان نویز جهت دار را برای مقادیر مختلف μ مقایسه کرد. ملاحظه می شود که در این حالت، سرعت همگرایی پایین آمده است به طوری که برای همگرایی الگوریتم نیاز به تکرار دوره های بسیار بیشتری داریم. همچنین، برخلاف روش LMS با یک بار اجرای الگوریتم، منحنی های همواری به دست آمده است.

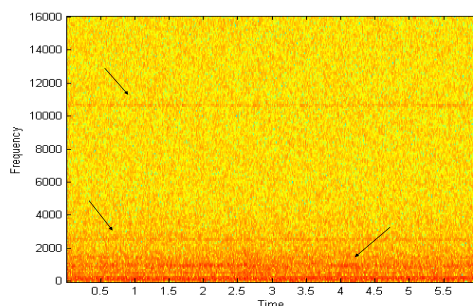
سناریو ۴- حال، به بررسی ناکارآمدی های شکل دهنده پرتو مبتنی بر الگوریتم سریع ترین کاهش و جبران آن توسط فیلتر پسینه می پردازیم. شرایط توصیف شده در سناریو قبلی را در نظر بگیرید. فرض کنید علاوه بر نویز زاویه ۶۰ درجه، نویز دیگری در جهت دید قرار داشته باشد. جدول (۴)، SNR خروجی شکل دهنده پرتو و فیلتر پسینه را به ازای نویزهای جهت دار مختلف و هم جهت با سیگنال تمیز مقایسه می کند. مقایسه نتایج حاصل با بخش قبل و همچنین، ارقام جدول اخیر با جدول (۳) مؤید آن است که SD در مقایسه با LMS اگرچه بار محاسباتی بیشتری دارد، ولی سریع تر بوده و مهمتر اینکه به دلیل بهینه سازی ضرایب وزن دهی در هر فریم، SNR بهتری می دهد. در LMS، ابتدا ضرایب وزن دهی مناسب نبوده و با ضرایب بهینه فاصله دارد. لذا، الگوی تشعشی مناسبی نداریم و بهسازی به خوبی انجام نمی شود؛ ولی، به مرور بردار ضرایب وزن دهی به سمت بردار بهینه همگرا می شود.

نمایش زمانی و اسپکتروگرام سیگنال تمیز، سیگنال نویزی میکروفون مرجع، خروجی شکل دهنده پرتو و خروجی فیلتر پسینه با فرض نویز هم جهت صورتی در شکل (۱۷) نمایش داده شده است.

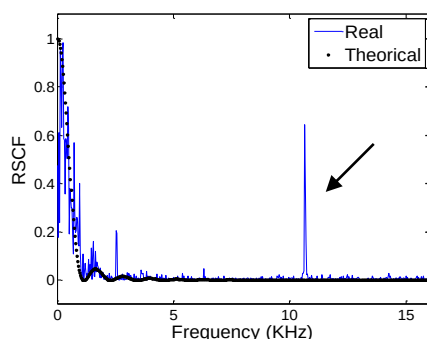
در آزمایش دیگری در این سناریو، با فرض نویز جهت داری با طیف سفید در جهت منبع هدف و با فرض SNR های مختلف به بررسی SNR سیگنال خروجی شکل دهنده پرتو و فیلتر پسینه می پردازیم. در جدول (۵) می توان مقادیر حاصل از این آزمایش را مقایسه نمود.

نویزی تولیدی به خوبی از مدل پخشنده پیروی می کند به طوری که دارای مربع همدوسی شبیه به sinc است. شکل (۲۱) تغییرات زمانی و اسپکتروگرام سیگنال نویزی کانال مرجع، خروجی شکل دهنده پرتو مبتنی بر الگوریتم LMS و خروجی فیلتر پسینه را با فرض μ برابر با ۰/۰۰۴۵ در حضور نویز جهت دار نشان می دهد.

در ابتدا، به دلیل بهینه نبودن ضرایب وزن دهی، شکل دهنده پرتو مبتنی بر LMS، حذف نویز جهت دار را به درستی انجام نداده است؛ ولی به مرور با همگرایی الگوریتم و نزدیک شدن ضرایب وزن دهی به مقادیر بهینه، نویز جهت دار محیط به خوبی حذف می شود. به دلیل ماهیت همه جهته نویز پخشنده، شکل دهنده پرتو نتوانسته آن را به خوبی حذف کند، ولی بلوک مکمل فیلتر پسینه به خوبی نویز مفروض را حذف و سیگنال تمیز را استخراج کرده است. در شکل (۲۲-الف) می توان تغییرات توان نویز جهت دار در خروجی شکل دهنده پرتو را با تغییر اندازه گام مشاهده کرد. افزایش μ منجر به افزایش سرعت همگرایی خواهد شد؛ ولی این مسئله، از حد مشخصی به بعد، به واگرایی الگوریتم منجر می شود (شکل (۲۲-ب)). این مسئله، به خوبی با مقایسه اسپکتروگرام سیگنال های نویزی مرجع، خروجی شکل دهنده پرتو LMS و خروجی فیلتر پسینه نویز در شکل (۲۳) نمایان است. واضح است عدم انتخاب صحیح اندازه گام، منجر به واگرایی الگوریتم شده به طوری که در تکرارهای مختلف توان نویز جهت دار به مرور افزایش می یابد. نتیجه تکرار آزمایش قبلی با استفاده از الگوریتم SD و مقدار بهینه μ برای هر فریم، در شکل (۲۴) نمایش داده شده است.



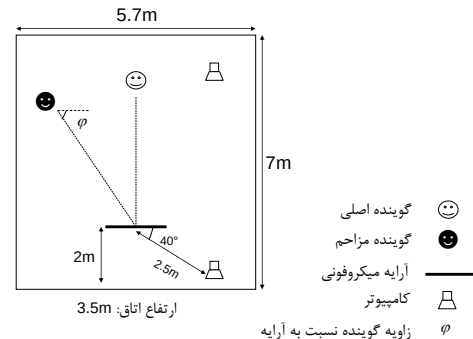
(الف)



(ب)

شکل (۱۹): (الف) اسپکتروگرام سیگنال میکروفون مرجع (ب): مقایسه مربع همدوسی سیگنال های دو میکروفون در فاصله ۱۰ سانتیمتری با مربع همدوسی یک میدان نویزی پخشنده تئوری

در ارتفاع ۱/۱ متری از سطح زمین استفاده نمودیم. سیستم ضبط صدای مورد استفاده، شامل ۸ میکروفون، ۸ تقویت کننده، منبع تغذیه، تنظیم کننده ولتاژ و مدار الکترونیکی جهت نمونه برداری از سیگنال ها است. از تمامی سیگنال ها، با نرخ ۳۲ کیلو هرتز نمونه برداری شد. شکل (۱۸) شمای کلی از پیکربندی ضبط صدا را به تصویر می کشد.



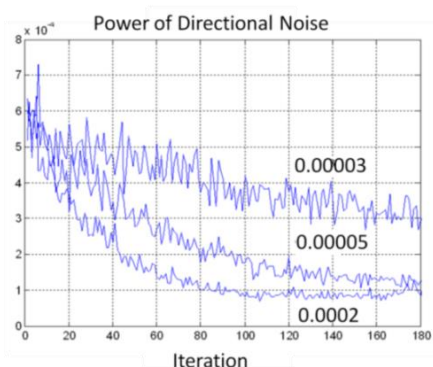
شکل (۱۸): شمای کلی از پیکربندی ضبط صدا

سناریو ۵- در گام نخست، با ضبط صدای محیط، به ارزیابی نویزهای موجود در پس زمینه محیط و مقایسه تابع همدوسی فضایی سیگنال میکروفون های مختلف پرداختیم. در این حالت، مطابق شکل (۱۹)، چهار میدان نویزی اساسی مشاهده شد: (۱) نویز باند باریکی در مجاورت فرکانس ۱۰۷۰۰ هرتز (با همبستگی بالا) که پس از بررسی های به عمل آمده، مشخص شد که ناشی از خود دستگاه ضبط صدا است. (۲) نویز باند باریکی در مجاورت فرکانس ۲۵۰۰ هرتز که به دلیل همبستگی بالای آن (شکل (۱۹-ب))، به نظر می رسد ناشی از صدای فن کامپیوتر مورد استفاده جهت ضبط صدا باشد. (۳) نویز سفید ناهمبسته (۴) نویز پخشنده با محتوای فرکانسی زیر ۲۵۰۰ هرتز.

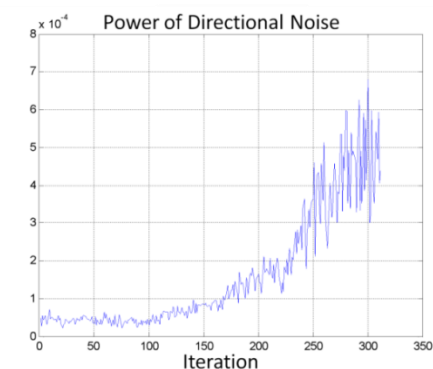
سناریو ۶- یک میدان نویزی پخشنده، برآیند سیگنال هایی است که با احتمال مساوی و انرژی یکنواخت در همه جهتهای محیط منتشر می شوند و همگی دارای دامنه برابر و فاز تصادفی هستند. در ادامه، به منظور بررسی بهتر رفتار این میدان های نویزی، به تولید نویز همهمه با رفتار پخشنده پرداختیم. هدف بررسی تابع همدوسی فضایی میدان های نویزی پخشنده است که در محیط های واقعی تا چه حد از تحلیل تئوری توصیف شده تبعیت می کند. بدین منظور، از آرایه ای در وسط اتاق و در شرایط مقارن استفاده کردیم. ضبط صدای محیط به نحوی طراحی شد که تا جای ممکن سیگنال ها با انرژی یکنواخت به آرایه میکروفونی برسند. ضمن آنکه گوینده در زاویه ۹۰ درجه نسبت به آرایه به ایراد سخن پرداخت، از ۴ گروه دو نفری در چهار گوشه اتاق خواسته شد تا به آرامی شروع به صحبت با یکدیگر کنند. نویز جهت داری نیز مطابق شکل (۱۸) در زاویه ۶۰ درجه آرایه تولید اغتشاش می کند.

در شکل (۲۰) می توان مربع همدوسی فضایی میدان نویزی ایجاد شده را با تئوری مقایسه نمود. همانطور که دیده می شود RCSF سیگنال های ضبط شده با نمودار حالت تئوری تطابق داشته و میدان

فیلتر پسینه با هدف بهسازی سیگنال گفتار در شرایط واقعی پیشنهاد شد. ساختار پیشنهادی، ضمن مقابله مؤثر با نویزهای جهت‌دار محیط، به شکل کارآمدی به تعدیل نویزهای ناهمبسته و پخشنده می‌پردازد. شکل‌دهنده پرتو استفاده شده براساس دو الگوریتم SD و LMS عمل می‌کند. استفاده از الگوریتم SD در مقایسه با LMS گرچه بار محاسباتی بیشتری دارد، ولی سریع‌تر بوده و الگوی تشعشعی بهتری نتیجه می‌دهد. در این فرآیند، با افزایش اندازه گام، سرعت هگرایی افزایش یافته و ضرایب وزن دهی زودتر به حد بهینه نزدیک می‌شوند. ولی، الگوی تشعشعی، نوسانات بیشتری پیدا می‌کند. این مسئله، منجر به افزایش خطا شده که مطلوب نیست. بررسی‌های ما مؤید آن است که شکل‌دهنده‌های پرتویی که بر مبنای اختلاف جهت دریافت سیگنال و نویز عمل می‌کنند، قادر به حذف نویزهای پخشنده محیط نیستند. همچنین، حذف نویزهای فرکانس پایین، مستلزم استفاده از آرایه‌های میکروفونی طولی بوده که خود از نظر عملی مشکل‌ساز است. توان بالای نویزهای ناهمبسته محیط دیگر عامل توجیه عدم کارایی استفاده تنها از شکل‌دهنده‌های پرتو می‌باشد. از این رو، مکمل فیلتر پسینه استفاده شد تا با تخمین مناسب نویزهای مذکور، ناکارآمدی قبلی را پوشش دهد. نتایج آزمایش‌ها روی دو دسته دادگان شبیه‌سازی شده واقعی، حاکی از بهسازی مؤثر این دو واحد در کنار همدیگر می‌باشد.

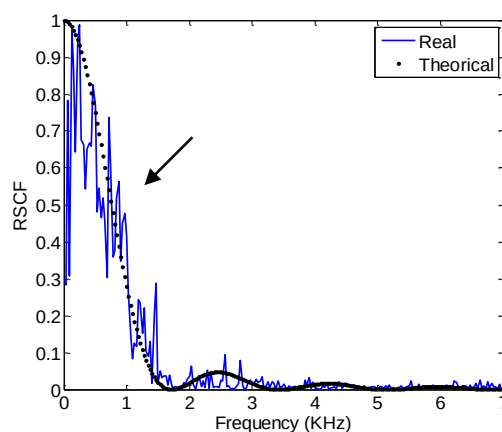
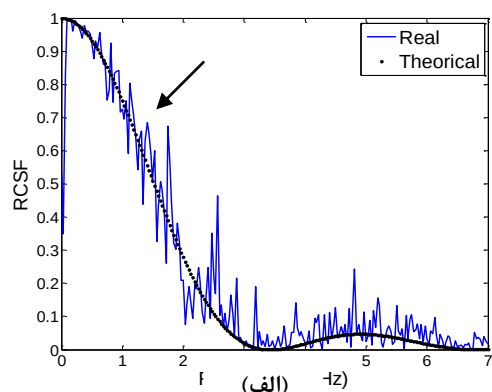


(الف)



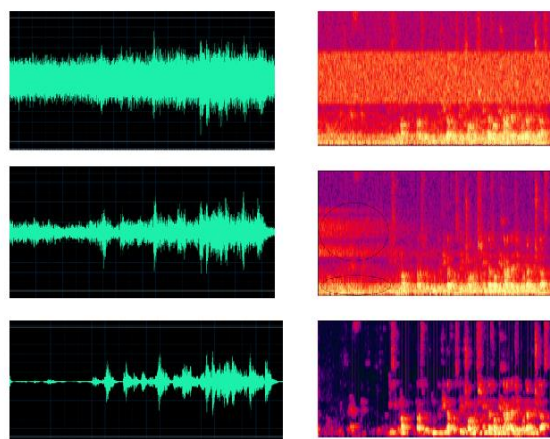
(ب)

شکل (۲۲): (الف) تغییرات توان نویز جهت‌دار در خروجی شکل‌دهنده پرتو LMS با تغییر اندازه گام در شرایط واقعی (ب) واگرایی الگوریتم LMS در اثر انتخاب ناصحیح μ



(ب)

شکل (۲۰): مقایسه مربع همدوسی سیگنال‌های دو میکروفون در فاصله (الف) ۱۵ سانتیمتری (ب) ۵ سانتیمتری با مربع همدوسی حاصل از یک میدان نویزی پخشنده تئوری



(ب)

(الف)

شکل (۲۱): (الف) نمایش زمانی (ب) اسپکتروگرام سیگنال‌های نویزی کانال مرجع، خروجی شکل‌دهنده پرتو مبتنی بر الگوریتم LMS و خروجی فیلتر پسینه در شرایط واقعی با فرض μ برابر با ۰/۰۰۴۵

۶ نتیجه گیری

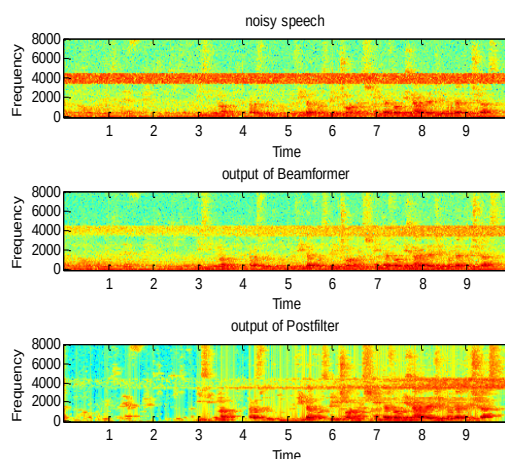
در این مقاله، ساختار مرکبی مبتنی بر شکل‌دهنده‌های پرتو وقتی و

فازی، مجله مهندسی برق دانشگاه تبریز، صفحه ۷۹-۸۴، شماره ۴، ۱۳۹۴.

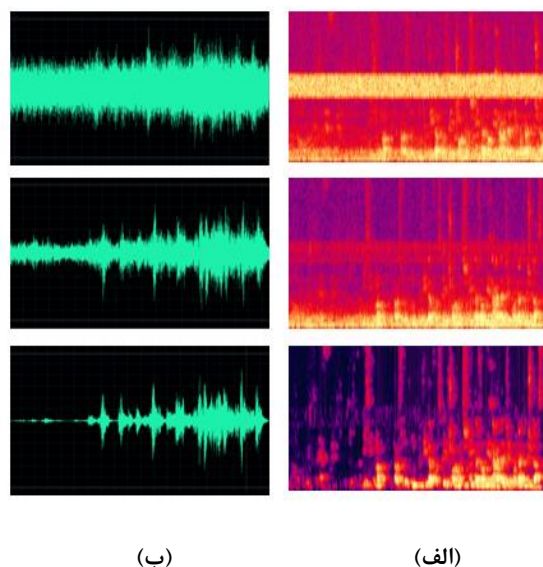
- [5] A. Avokh and H. R. Abutalebi, "Speech enhancement using linearly constrained adaptive constant directivity beamformers," *Applied Acoustics*, vol. 71, no. 3, pp. 262-268, 2010.
- [6] D. Commiello, M. Scarpiniti, R. Parisi and A. Uncini, "A novel affine projection algorithm for superdirective microphone array beamforming," in *Proc. IEEE International Symposium on Circuits and Systems*, pp. 2127-2130, Jun. 2010.
- [7] W. Dongxia, Z. Jiachao, W. Tao, "A broadband beamforming method based on microphone array for the speech enhancement," *International Conference on Signal Processing Systems*, pp. 363-366, 2010.
- [8] S. M. Golan, A. Bertrand, M. Moonen, and S. Gannot, "Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks," *Signal Processing*, vol. 107, pp. 4-20, 2015.
- [9] S. Khoubrouy and J. Hansen, "Microphone array processing strategies for distant based automatic speech recognition," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1344 - 1348, July 2016.
- [10] T. Padoisa, F. Sgard, O. Doutres, and AlainBerry, "Acoustic source localization using a polyhedral microphone array and an improved generalized cross-correlation technique," *Journal of Sound and Vibration*, vol. 386, pp. 82-99, Jan. 2017.
- [11] M. J. Taghizadeh, P. N. Garner, and H. Bourlard, "Enhanced diffuse field model for ad hoc microphone array calibration," *Signal Processing*, vol. 101, pp. 242-255, August 2014.
- [12] J. Li and M. Akagi, "A hybrid microphone array post-filter in a diffuse noise field," vol. 69, no. 6, pp. 546-557, June 2008.
- [13] Y. A. Huang, A. Luebs, J. Skoglund, and W. B. Kleijn, "Globally optimized least-squares post-filtering for microphone array speech enhancement," in *proc International Conference on Signal Processing Systems*, , pp. 380-384, 2016.
- [14] G. Huang, J. Benesty, and J. Chen, "Superdirective beamforming based on the Krylov matrix," *IEEE/ACM Trans. Audio, Speech, Language Processing*, vol. 24, pp. 2531-2543, Dec. 2016.
- [15] D. L. Alon and B. Rafaely, "Beamforming with optimal aliasing cancellation in spherical microphone arrays," *IEEE/ACM Trans. Audio, Speech & Language Processing*, vol. 24, no. 1, pp.196-210, 2016.
- [16] N. Ito, H. Shimizu, N. Ono, and S. Sagayama, "Diffuse noise suppression using crystal-shaped microphone arrays," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 19, no. 7, pp. 2101-2110, Sep. 2011.
- [17] مجتبی حاجی آبادی، عباس ابراهیمی مقدم و حسین خوش بین، "حذف نویز صوتی مبتنی بر یک الگوریتم وقتی نویز"، *مجله مهندسی برق دانشگاه تبریز*، صفحه ۱۳۹-۱۴۶، شماره ۳، ۱۳۹۵.
- [18] L. Chergui and S. Bouguezel, "A new pre-whitening transform domain LMS algorithm and its application to speech denoising," *Signal Processing*, Vol. 130, , pp. 118-128, Jan. 2017

[۱۹] مسعود گراوانچی زاده، پریا دادور و بابک بهادرینیا، "بهبود سیستم جداسازی منبع مبتنی بر آنالیز ترکیب شنیداری در زبان فارسی" شماره ۴، صفحه ۲۷۳-۲۸۳، ۱۳۹۵.

- [20] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beam-forming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313-324, Feb. 2003.
- [21] S. Haykin, *Adaptive Filter Theory*, fourth ed., Prentice Hall, 2002.
- [22] G. Dobliger "An adaptive microphone array for optimum beamforming and noise reduction," in *Proc. European Signal Processing Conference (EUSIPCO)* Florence, Italy, 2006.
- [23] P. Teal, T. D. Abhayapala, and R. A. Kennedy, "Spatial correlation in non-isotropic scattering scenarios," in *proc. International Conference on Acoustics, Speech, and Signal Processing*, pp. 2833-2836, 2002.



شکل (۲۳): اسپکتروگرام سیگنال‌های نویزی کانال مرجع، خروجی شکل‌دهنده پرتو LMS و فیلتر پسینه با انتخاب ناصحیح μ



شکل (۲۴): (الف) اسپکتروگرام (ب) نمایش زمانی کانال مرجع، خروجی شکل‌دهنده پرتو SD و خروجی فیلتر پسینه در شرایط واقعی

مراجع

- [1] Z. Yermecche, "Subband beamforming for speech enhancement in hands-free communication", *Licentiate Thesis, Blekinge Institute of Technology 372 25 Ronneby, Sweden*, Dec. 2004.
- [2] N. Ito, E. Vincent, T. Nakatani, N. Ono, S. Araki, and S. Sagayama, "Blind suppression of nonstationary diffuse acoustic noise based on spatial covariance matrix decomposition," *Signal Processing Systems*, vol. 79, no. 2, pp. 145-157, 2015.
- [3] I. McCowan and H. Bourlard, "Microphone array post-filter for diffuse noise field", In *Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. 905-908, Florida, 2002.
- [۴] زهرا حبیبی، مرتضی کازرونی، سیدحسین محسنی ارمکی و عماد حمیدی، "ارائه یک روش کاربردی جهت کالبراسیون آنتن‌های آرایه

- [26] M. Bijankhan, J. Seikhzadeghan, M. R. Roohani, Y. Samareh, K. Lucas, M. Tebyani, "FARSDAT - the speech database of farsi spoken language," in *Proc. of SST-94*, pp. 826-831, 1994.
- [27] <http://svr-ww.eng.cam.ac.uk/comp.speech/section1/data/noisex.html>
- [24] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, pp. 1109-1121, Dec. 1984.
- [25] <https://pws.yazd.ac.ir/spri>

زیر نویس ها

^{۱۵} Steepest Descent

^{۱۶} Far-field

^{۱۷} Finite Impulse Response

^{۱۸} Frequency bin

^{۱۹} Look Direction

^{۲۰} Mean Square Error

^{۲۱} Post filter

^{۲۲} Ill condition

^{۲۳} Gradient based

^{۲۴} Short Time Spectral Amplitude

^{۲۵} Prior SNR

^{۲۶} Posterior SNR

^{۲۷} White Gaussian Noise

^{۲۸} Delay and Sum Beam-former

^۱ Diffuse

^۲ Signal to Noise Ratio

^۳ Beam-former

^۴ Beam-pattern

^۵ Constant Directivity Beam-former

^۶ Constrained Least Mean Square

^۷ Linearly Constrained Minimum Variance

^۸ Superdirective Beam-formers

^۹ Minimum Variance Distortionless Response

^{۱۰} Linear Minimum Mean Square Error

^{۱۱} Pre-whitening Transform domain Least Mean Square

^{۱۲} Model-Based Expectation-Maximization Source Separation and Localization

^{۱۳} Learning-Based Particle Swarm Optimization

^{۱۴} Least Mean Square