

(پژوهشی)

جداسازی منابع سیگنال صوتی زیر آب با استفاده از الگوریتم انتقال میانگین گوسی وزن دهی شده تعمیم یافته

آذر محمودزاده^۱، حامد آگاهی^۲

mahmoodzadeh@iaushiraz.ac.ir

۱- استادیار گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران
۲- دانشیار گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

چکیده

جداسازی منابع صوتی یکی از مهم‌ترین زیر شاخه‌های بهسازی سیگنال صوتی به شمار می‌رود، که در هنگام مخلوط شدن سیگنال‌های صوتی و در حضور نویز اهمیت بیشتری پیدا می‌کند. در این مقاله، با الهام گرفتن از روش WGMS در سیستم‌های دو کاناله، این الگوریتم به سیستم‌های چند کاناله تعمیم یافته و مورد بررسی قرار می‌گیرد. جداسازی منابع صوتی به روش کور بوده و قابلیت تفکیک چندین منبع صوتی را از تعداد متفاوتی مخلوط دارد. روش پیشنهادی، با تعداد منبع صوتی و هیدروفون‌های متفاوت در شرایط نویزی، با دیگر الگوریتم‌ها قیاس شده و نتایج حاصل، نشان از کارایی موثر آن دارد.

واژگان کلیدی : سیگنال صوتی، جداسازی منابع، الگوریتم انتقال میانگین، خوشه بندی.



تاریخ دریافت مقاله: ۱۴۰۰/۱۲/۱۹

تاریخ ویرایش مقاله: ۱۴۰۱/۰۲/۱۰

تاریخ پذیرش مقاله: ۱۴۰۱/۰۲/۱۰

DOI: 10.22034/IJMST.2022.540246.1565

صص ۱۹-۱۰

مقدمه

نیز عنوان شده است که حجم محاسباتی بالایی دارد [۹]. همچنین به منظور جداسازی منابع استاتیک و دینامیک، روش فاکتورگیری نامنفی پیشنهاد می‌گردد [۱۰]. روش تجزیه‌ی تانسور^۳ نیز جهت جداسازی کور منابع نامعین مشترک، مطرح می‌باشد. این روش به برآورد ماتریس مخلوط داده‌ها پرداخته و در حالتی که تعداد منابع بیش از تعداد مشاهدات باشد، کارایی مناسبی دارد [۱۱]. از دیگر الگوریتم‌های موجود می‌توان به جداسازی کور سیگنال مخلوط صوتی با وجود نویز در حوزه‌ی زمان-فرکانس اشاره نمود [۱۲]. در این روش ابتدا ماتریس مخلوط نویزی تخمین زده شده و سپس منابع اصلی به وسیله حذف تداخل و با به کارگیری الگوریتم ژنتیک بازسازی می‌شوند. در [۱۳] روشی بر پایه‌ی دو هیدروفون مطرح شده است که از الگوریتم خوشه بندی انتقال میانگین گوسی وزن دهی شده (WGMS) جهت جداسازی منابع صوتی استفاده می‌کند. در صورت وجود بیش از دو هیدروفون، آرایه‌ای از هیدروفون-ها وجود دارد که با یک آرایش هندسی خاص در فضا قرار داده می‌شوند. افزایش تعداد هیدروفون‌ها و وجود داده‌های بیشتر، منجر به دقت بالاتر تخمین و همچنین افزایش هزینه تجهیزات، بار محاسباتی بالاتر، مصرف انرژی بیشتر و حل معادلات پیچیده‌تر خواهد شد [۱۴]. استفاده از الگوریتم EM^۵ جهت تفکیک مخلوط صوتی کانال‌ها شده از جمله روش‌های چند هیدروفون می‌باشد که نیازمند محاسبات بالایی است [۱۵].

الگوریتم ارائه شده در این مقاله، امکان جداسازی کور سیگنال‌های مخلوط شده را با استفاده از تعمیم روش WGMS برای سیستم‌های چند هیدروفون فراهم می‌کند. عمل جداسازی در حوزه زمان-فرکانس است و بدون دانش پیشین در مورد تعداد و نحوه چیدمان هندسی منابع و هیدروفون‌ها انجام می‌شود. نتایج ارزیابی، بیانگر عملکرد بهتر این روش نسبت به دیگر الگوریتم‌ها به خصوص در شرایط نویزی می‌باشد.

ساختار این مقاله به صورت زیر بیان گردیده است. در قسمت دوم به بررسی الگوریتم خوشه بندی MS و در بخش سوم به تشریح الگوریتم WGMS می‌پردازیم. در قسمت

بهبود کیفیت سیگنال صوتی و افزایش ادراک پذیری آن یکی از زمینه‌های مهم در حوزه پردازش صوت به شمار می‌رود. جداسازی منابع صوتی به ویژه در محیط‌های نویزی و یا هنگام مخلوط شدن سیگنال صوتی با یکدیگر اهمیت بیشتری پیدا می‌کند. هدف از جداسازی سیگنال منابع، تخمین صحیح سیگنال‌های ارسال شده از تعدادی منبع مختلف با استفاده از مخلوط سیگنال‌های دریافتی به وسیله-ی چندین گیرنده (هیدروفون) است. این مخلوط می‌تواند با نویز و سایر عوامل مزاحم محیطی نیز ترکیب شده باشد [۱].

مدل‌های بی‌شماری برای حل مسأله جداسازی سیگنال‌های صوتی ارائه شده است. بر اساس تعداد هیدروفون موجود، این روش‌ها به سه دسته‌ی تک کاناله، دوکاناله و آرایه‌ی هیدروفونی تقسیم می‌شوند. در حالت وجود یک هیدروفون که به روش‌های تک کاناله معروفند، تنها یک سیگنال ترکیب شده‌ی دریافتی جهت پردازش وجود دارد. اخیراً از روش‌های پردازش فازی جهت جداسازی تک کاناله استفاده شده است [۲]. در الگوریتم‌های شامل دو هیدروفون یا روش‌های دوکاناله، نیز تکنیک‌های متفاوتی بیان شده است [۳-۴].

از جمله می‌توان به الگوریتم (DUET) که یکی از روش‌های کاربردی جداسازی در یک محیط بدون پژواک می‌باشد، اشاره نمود [۵]. این الگوریتم از خاصیت تنکی سیگنال صوتی استفاده نموده و جداسازی را در حوزه زمان-فرکانس با محاسبه تأخیرها و تضعیف‌های نسبی انجام می‌دهد. با مشخص شدن پارامترهای هر منبع، نقاطی که دارای یک ویژگی هستند، بر چسب گذاری شده و خوشه بندی توسط رسم هیستوگرام دو بعدی انجام می‌شود. در نهایت با اعمال ماسک باینری و برگرداندن به حوزه زمان، عمل جداسازی کامل می‌گردد [۵-۶]. همچنین به منظور جداسازی بهتر منابع با استفاده از خاصیت تنک بودن سیگنال صوتی، روشی مبتنی بر خوشه بندی وزن دهی شده K-means مطرح شده است [۷]. استفاده از انرژی نقاط زمان-فرکانس در الگوریتم خوشه بندی انتقال میانگین (MS) از دیگر روش‌های ارائه شده در جداسازی سیگنال صوتی می‌باشد [۸]. مدل مخلوط ماتریس برای خوشه بندی به ازای هر بعد از فضای ویژگی

⁴ Weighted Gaussian Mean Shift clustering

⁵ Expectation Maximization

¹ Degenerate Unmixing Estimation Technique

² Mean Shift

³ Tensor

$$\nabla f(X) = \frac{2c}{Rh^{d+2}} \left[\sum_{i=1}^R (X_i - X) g \left(\left\| \frac{X - X_i}{h} \right\|^2 \right) \right] \quad (3)$$

$$= \left[\frac{2c}{Rh^{d+2}} \sum_{i=1}^R g \left(\left\| \frac{X - X_i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^R X_i g \left(\left\| \frac{X - X_i}{h} \right\|^2 \right)}{\sum_{i=1}^R g \left(\left\| \frac{X - X_i}{h} \right\|^2 \right)} - X \right]$$

$$\square a_h(X), m_h(X)$$

که در این رابطه $g(s) = k'(s)$ است. ترم اسکالر $a_h(X)$ وابسته به تخمین چگالی در نقطه X است که با استفاده از کرنل $G(X) = cg(\|X\|^2)$ محاسبه می‌شود و ترم $m_h(X)$ بردار انتقال میانگین است. بردار انتقال میانگین همواره به سمت مسیر رسیدن به نقطه ماکزیمم در تابع چگالی اشاره دارد. در هر تکرار الگوریتم انتقال میانگین به صورت تکراری مقدار X^t را در جهت ماکزیمم کردن گرادیان مطابق با رابطه زیر به روزآوری می‌کند:

$$X^{t+1} = X^t + m_h(X^t) \quad (4)$$

تکرارها تا زمانی ادامه می‌یابد که مقدار تغییرات کوچکتر از یک حد آستانه از پیش تعیین شده گردد. در این حالت الگوریتم به حالت پایدار رسیده و همگرا شده است. از آن جا که X^t با استفاده از قاعده گرادیان صعودی به روز می‌شود بردار همگرا شده در پایان الگوریتم، نقطه‌ای با نزدیکترین فاصله نسبت به ماکزیمم‌های محلی تخمین چگالی احتمال خواهد بود.

با توجه به کرنل گوسی بیان شده در رابطه زیر:

$$k(s) = \frac{1}{\sqrt{(2\pi)^d}} \exp\left(-\frac{s^2}{2}\right) \quad (5)$$

تابع چگالی احتمال بیان شده در رابطه (۱) با استفاده از کرنل گوسی به صورت $\hat{f}(X)$ بیان می‌گردد. برای تخمین بردار انتقال میانگین، لازم است تا گرادیان تابع چگالی احتمال محاسبه گردد. این گرادیان را می‌توان طبق رابطه زیر محاسبه نمود:

$$\nabla \hat{f}(X) = \frac{\hat{f}(X)}{h^2} m_h(X) \quad (6)$$

چهارم به شرح الگوریتم پیشنهادی برای جداسازی سیگنال صوتی منابع و چگونگی تعمیم الگوریتم WGMS پرداخته می‌شود. در قسمت پنجم نتایج حاصل از شبیه‌سازی و در نهایت نتیجه‌گیری بیان می‌گردد.

الگوریتم خوشه بندی انتقال میانگین

الگوریتم خوشه‌بندی انتقال میانگین یک تکنیک غیر پارامتریک برای پیدا کردن ماکزیمم‌های محلی و یا خوشه-بندی دادگان است که نیاز به دانش قبلی در مورد تعداد خوشه‌ها نداشته و همچنین محدودیتی در شکل توزیع داده و یا خوشه‌ها ندارد [۱۴]. این الگوریتم با یک روش تکراری ماکزیمم‌های محلی را پیدا می‌کند. به این صورت که به ازای هر نقطه داده در فضای ویژگی و قرار دادن مرکز یک پنجره جستجو در آن نقطه، بررسی می‌شود که آیا آن نقطه داده از نظر توزیع آماری در یک ماکزیمم محلی قرار گرفته است. در صورت واقع نبودن در یک ماکزیمم محلی، مرکز همسایگی جابجا شده و مجدداً بررسی انجام می‌شود. مقایسه فوق تا زمان رسیدن به یک همگرایی و پیدا نمودن نقطه داده‌ای بر روی ماکزیمم محلی ادامه می‌یابد. نقاط داده‌ای که با این ماکزیمم محلی مرتبط هستند بعنوان اعضای خوشه در نظر گرفته می‌شوند.

فرض کنید n نقطه داده $X_i, i=1, \dots, n$ ، در فضای d بعدی R^d داریم. تابع چگالی کرنل چند متغیره با استفاده از کرنل $K(X)$ و با شعاع پنجره h به صورت زیر تخمین زده می‌شود:

$$f(X) = \frac{1}{Rh^d} \sum_{i=1}^R K\left(\frac{X - X_i}{h}\right) \quad (1)$$

برای یک کرنل متقارن شعاعی کافی است که مشخصات کرنل $k(x)$ در شرط زیر برآورده شود:

$$K(X) = ck(\|X\|^2) \quad (2)$$

که در این رابطه c یک ثابت نرمالیزاسیون است که اطمینان می‌دهد انتگرال $K(X)$ برابر با یک می‌گردد و $k: \square \rightarrow \square$ یک تابع اسکالر است. اگر $\nabla f(X)$ برابر با گرادیان تابع $f(X)$ در نظر گرفته شود، ماکزیمم‌های محلی این تابع در نقاطی قرار گرفته است که $\nabla f(X) = 0$. گرادیان تابع $f(X)$ به صورت زیر محاسبه می‌شود:

رابطه (۱۰) در صورتی که پارامترهای هموارسازی h_i وابسته به کرنل فعلی نباشد، مشابه با الگوریتم انتقال میانگین معمولی خواهد بود و در نتیجه $H_i = H$ منظور می‌گردد. پس از ساده سازی رابطه (۱۰)، در نهایت رابطه اصلی نشان دهنده الگوریتم وزن‌دهی شده انتقال میانگین به فرم زیر در خواهد آمد:

$$\nabla \hat{f}(X) = \hat{f}(X) \cdot H^{-1} \cdot \left(\frac{\sum_{i=1}^n w_i k(\|X - X_i\|_{H_i^{-1}})}{\sum_{i=1}^n w_i k(\|X - X_i\|_{H_i^{-1}})} \cdot X_i - X \right) \quad (11)$$

$$= \hat{f}(X) \cdot H^{-1} \cdot \mathbf{m}(X)$$

که در این رابطه بردار انتقال میانگین وزن‌دهی شده با $\mathbf{m}(X)$ مشخص می‌شود. اختلاف اصلی الگوریتم‌های انتقال میانگین و انتقال میانگین گوسی وزن‌دهی شده در وجود ضرایب وزن‌دهی است که نشان دهنده تأثیر هر کرنل بر الگوریتم است. از آن جا که هدف از وزن‌دهی، دادن بهای بیشتر به نقاط پر انرژی‌تر در حوزه مختلط زمان-فرکانس است، بنابراین تابع وزن‌دهی مورد استفاده در این الگوریتم برابر با مربع میانگین هندسی انرژی‌های هر دو هیدروفون تعریف می‌شود. رابطه مورد استفاده برای این ضریب وزن-دهی، به فرم زیر قابل بیان است [۱۱]:

$$W(k, l) = |Y_1(k, l)| * |Y_2(k, l)| \quad (12)$$

ضرایب $W(k, l)$ یک ماتریس با سایز $K \times L$ تشکیل می‌دهد که در آن K تعداد باندهای فرکانس و L فریم‌های زمانی است. در این رابطه $Y_i(k, l)$ به ازای $i=1, 2$ تبدیل فوریه زمان کوتاه (STFT) سیگنال‌های مخلوط موجود در $\hat{\mathbf{m}}$ هیدروفون برای k امین باند فرکانسی و l امین فریم است.

برای جداسازی سیگنال صوتی با استفاده از دو هیدروفون، فضای ویژگی‌های ورودی برابر با دو ($d=2$) در نظر گرفته می‌شود. این دو ویژگی با محاسبه اختلاف سطح بین دو هیدروفون (ILD) و همچنین اختلاف زمانی بین دو هیدروفون (ITD) به صورت روابط زیر محاسبه می‌گردد:

$$\alpha(k, l) = \frac{|Y_2(k, l)|}{|Y_1(k, l)|} - \frac{|Y_1(k, l)|}{|Y_2(k, l)|} \quad (13)$$

الگوریتم خوشه بندی انتقال میانگین گوسی وزن دهی شده

ایده‌ی اصلی این مقاله در جداسازی سیگنال‌های صوتی این است که انرژی هر نقطه زمان-فرکانس سیگنال‌های مخلوط شده در هیدروفون‌ها با نقاط دیگر متفاوت است. با استفاده از وزن‌دهی نقاط زمان-فرکانس و محاسبه انرژی نقاط، این تفاوت‌ها نمایان می‌گردد. به همین منظور، در این مقاله الگوریتم خوشه بندی انتقال میانگین گوسی وزن‌دهی شده به کار گرفته شده است. در این روش یک پارامتر هموارسازی ویژه به هر ضریب هر کرنل اختصاص داده می‌شود و به جای میانگین‌گیری استاندارد از میانگین کرنل‌های وزن‌دار استفاده می‌گردد [۱۱]. برای شروع $\|X\|_{\Psi}^2 = X^T \cdot \Psi \cdot X$ برای بردار X و یک ماتریس مربعی Ψ تعریف می‌گردد. نسخه وزن‌دار رابطه (۱) را می‌توان به صورت زیر نشان داد:

$$\hat{f}(X) = \frac{\sum_{i=1}^R w_i |H_i|^{-1/2} k(\|X - X_i\|_{H_i^{-1}}^2)}{R \cdot \sum_{i=1}^R w_i} \quad (7)$$

که در این رابطه R تعداد کرنل‌هاست و برابر با نقاط داده است. برای $\hat{\mathbf{m}}$ امین نقطه داده، w_i فاکتور وزن‌دهی است و H_i ماتریس هموارسازی قطری است که به صورت زیر تعریف می‌شود:

$$H_i = \text{diag}([h_{i1}^2, \dots, h_{ii}^2]) \quad (8)$$

که در این رابطه، پارامتر هموارسازی برای i امین کرنل در $\hat{\mathbf{m}}$ امین بعد است. همچنین با توجه به $k'(s) = -s \cdot k(s)$ ، گرادیان تابع چگالی احتمال رابطه (۷) به صورت زیر محاسبه می‌شود:

$$\nabla \hat{f}(X) = \frac{\sum_{i=1}^R w_i |H_i|^{-1/2} k(\|X - X_i\|_{H_i^{-1}}^2) \cdot H_i^{-1} \cdot (X_i - X)}{R \cdot \sum_{i=1}^R w_i} \quad (9)$$

با جایگذاری رابطه (۷) در (۹)، می‌توان رابطه زیر را بازنویسی نمود:

$$\nabla \hat{f}(X) = \hat{f}(X) \frac{\sum_{i=1}^R w_i |H_i|^{-1/2} k(\|X - X_i\|_{H_i^{-1}}^2) \cdot H_i^{-1} \cdot (X_i - X)}{\sum_{i=1}^R w_i |H_i|^{-1/2} k(\|X - X_i\|_{H_i^{-1}}^2)} \quad (10)$$

⁸ Inter-sensor Time Differences

⁶ Short Time Fourier Transform

⁷ Inter-sensor Level Differences

مدل مخلوط و تعیین پارامترهای مخلوط

اگر $Y_m(k, l)$ به ازای $m = 1, \dots, M$ تبدیل فوریه کوتاه $n = 1, \dots, N$ امین سیگنال مخلوط و $S_n(k, l)$ به ازای $n = 1, \dots, N$ تبدیل فوریه کوتاه n امین سیگنال منبع باشند، می توان مدل سیگنال های مخلوط دریافتی نسبت به سیگنال های منبع را به فرم تبدیل ماتریسی زیر نمایش داد:

$$\begin{bmatrix} Y_1(k, l) \\ \vdots \\ Y_m(k, l) \\ \vdots \\ Y_M(k, l) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 & \dots & 1 \\ \vdots & & \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} & \dots & A_{mN} \\ \vdots & & \vdots & & \vdots \\ A_{M1} & \dots & A_{Mn} & \dots & A_{MN} \end{bmatrix} \begin{bmatrix} S_1(k, l) \\ \vdots \\ S_n(k, l) \\ \vdots \\ S_N(k, l) \end{bmatrix} \quad (15)$$

که در این رابطه A_{mn} اختلاف ترکیب نسبی m امین هیدروفون به ازای n امین منبع می باشد. تأخیرها و تضعیف های نسبی برای هر نقطه زمان-فرکانسی متفاوت و با استفاده از روابط زیر بیان می شوند:

$$A_{mn}(k, l) = \frac{Y_m(k, l)}{Y_{m-1}(k, l)} = \alpha_{mn} e^{-i\omega\delta_{mn}} \quad (16)$$

$$\alpha_{mn}(k, l) = \left| \frac{Y_m(k, l)}{Y_{m-1}(k, l)} \right| - \left| \frac{Y_{m-1}(k, l)}{Y_m(k, l)} \right| \quad (17)$$

$$\delta_{mn}(k, l) = -\frac{1}{\omega} \arg \left\{ \frac{Y_m(k, l)}{Y_{m-1}(k, l)} \right\} \quad (18)$$

که در این روابط دو عبارت $\alpha_{mn}(k, l)$ و $\delta_{mn}(k, l)$ به ترتیب همان ویژگی های ILD و ITD های استخراج شده از m امین هیدروفون به ازای n امین منبع برای فریم جداسازی می باشند، که برای هر جفت از هیدروفون ها محاسبه می شوند. بنابراین مدل مخلوط را می توان به صورت زیر بازنویسی نمود:

$$\delta(k, l) = -\frac{1}{\omega} \arg \left\{ \frac{Y_2(k, l)}{Y_1(k, l)} \right\} \quad (14)$$

هر کرنل در طول هر بعد شکل متفاوتی داشته و در نتیجه می توان از مقادیر مختلفی برای پارامتر هموارسازی استفاده نمود. پس از محاسبه ضرایب، ماتریس به برداری با طول $R = K.L$ تبدیل می شود. اکنون می توان وزن های محاسبه شده را در رابطه (۹) اعمال نمود. الگوریتم انتقال میانگین از یک نقطه داده که به صورت تصادفی انتخاب شده آغاز می گردد و تا به کارگیری بر روی همه بردارها در مجموعه داده ادامه می یابد. سپس نتایج خوشه بندی برای جداسازی سیگنال های صوتی مورد استفاده قرار می گیرد.

الگوریتم WGMS تعمیم یافته برای حالت بیش از دو هیدروفون

در روش WGMS، عمل تفکیک سیگنال های صوتی حاصل از دو هیدروفون با محاسبه تأخیرها و تضعیف های نسبی هیدروفون نسبت به منبع قابل انجام بوده و عملکرد مناسبی نیز در حضور نویز در مقایسه با دیگر روش ها دارد [۱۱]. با توجه به ابعاد بردار انتقال میانگین و با فرض تنک بودن سیگنال صوتی، می توان احتمال حضور سیگنال صوتی را بررسی نمود و یا همانند روش WGMS دو کاناله، افزایش تعداد هیدروفون ها را لحاظ کرد و عمل جداسازی را با وجود تعداد متفاوت منبع انجام داد.

برای وجود بیش از دو هیدروفون، ITD و ILD را باید برای هر جفت هیدروفون محاسبه نمود. این عمل باعث افزایش ابعاد فضای ویژگی و پیچیده شدن عمل خوشه بندی می شود. در این صورت نیاز به استخراج یک حالت جدید برای بردار انتقال میانگین نیست؛ زیرا بردار انتقال میانگین گوسی وزن دهی شده ذکر شده در رابطه (۱۳) برای هر تعداد از داده ها صحت دارد. در این حالت الگوریتم انتقال میانگین، تخمین های مختلف از پارامترهای مخلوط منابع را از هر جفت از هیدروفون ها به دست می آورد. در نتیجه برای تشخیص اینکه یک نقطه زمان-فرکانس متعلق به چه منبعی می باشد، در مرحله جداسازی باید این پارامترهای مخلوط، در نظر گرفته شوند.

هر مقدار از فضای ویژگی معتبر است؛ اما چون این الگوریتم $M - 1$ تخمین مختلف از پارامترهای مخلوط را فراهم می-کند، بنابراین \hat{A}_{mn} به صورت زیر تعریف می‌شود:

$$\hat{A}_{mn} = \hat{\alpha}_{mn} e^{-i\omega \hat{\delta}_{mn}^b}, \quad m = 1, \dots, M - 1 \quad (22)$$

با در نظر گرفتن $S_j(k, l)$ به عنوان منبع فعال در هر نقطه زمان - فرکانسی، می‌توان مدل مخلوط را با رابطه زیر بیان نمود:

$$Y_m(k, l) = A_{mj} S_j(k, l) + N_m(k, l), \quad m = 1, \dots, M \quad (23)$$

که در این رابطه A_{mj} اختلاف ترکیب نسبی بین منبع فعال j و هیدروفون m می‌باشد. $N_m(k, l)$ نیز نشان‌دهنده سیگنال‌های تصادفی، مستقل و با توزیع یکسان (i.i.d) بوده که آغشته به نویز سفید گوسی با میانگین صفر و واریانس σ^2 هستند.

این سیگنال‌های نویزی، در حقیقت بیانگر وجود دیگر سیگنال‌ها در نقطه زمان - فرکانسی (k, l) بوده که منبع S_j در آن نقطه فعال است. این به معنی وجود سیگنال‌های مزاحم و نامطلوب به غیر از سیگنال منبع فعال در نقطه‌ی (k, l) می‌باشد که باید حداقل شوند. بدین منظور، تابع درست‌نمایی بیشینه (ML) برای منبع $S_j(k, l)$ با توجه به مشاهدات $Y_m(k, l)$ به صورت زیر بیان می‌گردد:

$$L_j = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2} \sum_{m=1}^M |Y_m(k, l) - A_{mj} S_j(k, l)|^2} \quad (24)$$

ماکزیمم کردن رابطه‌ی (۲۴) معادل با مینیمم کردن عبارت زیر است:

$$L'_j = \sum_{m=1}^M |Y_m(k, l) - A_{mj} S_j(k, l)|^2 \quad (25)$$

برای مینیمم کردن و به حداقل رساندن رابطه‌ی فوق، باید $\frac{\partial L'_j}{\partial S_j} = 0, j = 0, \dots, N$ باشد. با حل این معادله نتیجه می‌شود:

$$\begin{bmatrix} Y_1(k, l) \\ \vdots \\ Y_m(k, l) \\ \vdots \\ Y_M(k, l) \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ A_{mj} \\ \vdots \\ A_{Mj} \end{bmatrix} S_j(k, l) \quad (19)$$

که در آن $S_j(k, l)$ منبع فعال در نقطه (k, l) است. به منظور محاسبه تأخیرها و تضعیف‌های نسبی هر جفت از هیدروفون‌های متوالی، اختلاف ترکیب نسبی مرتبط بین $m - 1$ امین هیدروفون نیز برابر رابطه زیر تعریف می‌شود:

$$B_{mn} = \alpha_{mn}^b e^{-i\omega \delta_{mn}^b} \quad (20)$$

که در این رابطه α_{mn}^b و δ_{mn}^b پارامترهای مخلوط مربوط به n امین منبع می‌باشند. با توجه به اینکه A_{mn} را می‌توان به صورت $A_{mn} = \prod_{i=2}^m B_{in}$ بیان نمود، سیگنال دریافت شده توسط m امین هیدروفون را نیز می‌توان به عنوان یک تابع از سیگنال دریافت شده توسط $(m - 1)$ امین هیدروفون مشابه فرمول زیر بیان نمود:

$$Y_m(k, l) = B_{mn} Y_{m-1}(k, l) \quad (21)$$

تخمین منابع از ضرایب مخلوط

مرحله بعد تخمین صحیح منابع صوتی از ضرایب مخلوط محاسبه شده می‌باشد. با وجود آرایه‌ای با M هیدروفون، $M - 1$ جفت هیدروفون متوالی وجود دارد و طبق رابطه (۲۱) سیگنال دریافتی هر هیدروفون را می‌توان به صورت سیگنال دریافتی هیدروفون قبلی ضرب در اختلاف ترکیب نسبی (B_{mn}) بین آن‌ها بیان نمود. در نتیجه برای هر هیدروفون $M - 1$ تخمین از پارامترهای ترکیب محلی وجود دارد. همین مسأله باعث می‌شود که فضای ویژگی ورودی برای الگوریتم انتقال میانگین دارای ابعاد $2(M - 1)$ باشد. بردار انتقال میانگین به دست آمده از رابطه (۱۱) برای

¹ Maximum Likelihood

0

⁹ Independent and Identically Distributed

جفت از هیدروفون‌ها محاسبه می‌شود. پارامترهای اختلاف ترکیب نسبی جهت تفکیک سیگنال‌ها به الگوریتم WGMS اعمال می‌گردد. الگوریتم تکراری یافتن بردارهای انتقال میانگین به صورت وزن‌دهی شده تکمیل و فرآیند به روزآوری ویژگی انجام می‌شود. سپس هر نقطه زمان - فرکانس بر اساس میزان نزدیکی به یک منبع اختصاص می‌یابد. جهت برجسب زنی ویژگی‌های به روز شده، یک ماسک باینری طراحی می‌شود تا تمامی نقاط موجود در فضای زمان- فرکانس را دسته بندی کند و منبع فعال اصلی با استفاده از رابطه (۳۰) تخمین زده می‌شود. با اعمال تبدیل فوریه معکوس زمان کوتاه، سیگنال‌ها بازسازی شده و منابع صوتی در حوزه‌ی زمان به دست می‌آیند. پس از تشریح الگوریتم پیشنهادی، در بخش بعد به توضیح نحوه شبیه سازی و نتایج آن خواهیم پرداخت.

نتایج شبیه سازی

در این مقاله جهت ارزیابی، از مجموعه دادگان صوتی [۱۵] با فرکانس نمونه‌برداری ۱۶ KHz به عنوان سیگنال‌های صوتی مرجع و نویز سفید گوسی استفاده شده است. این مجموعه شامل اصوات مختلف از جمله اصوات شناورها، زیردریایی‌ها و اژدرها می باشد. معیار ارزیابی مورد استفاده، SIR^۱ بوده که نسبت سیگنال اصلی به مقدار تداخل ناشی از سیگنال مزاحم را نشان می‌دهد. رابطه (۳۱) نحوه محاسبه معیار SIR برای جداسازی سیگنال‌های صوتی را نمایش می‌دهد [۱۶]:

$$SIR = 10 \log_{10} \frac{\|S_{target}(n)\|^2}{\|S_{int}(n)\|^2} \quad (31)$$

که در آن $S_{target}(n)$ و $S_{int}(n)$ سیگنال هدف (سیگنال منبع اصلی قبل از مخلوط شدن در هیدروفون) و سیگنال مزاحم (عامل تداخل) هستند. این معیار ارزیابی برای تک تک منابع به صورت جداگانه قابل محاسبه بوده و نشان دهنده میزان تفکیک‌پذیری آن از دیگر منابع صوتی است. به منظور ایستادن فرض کردن سیگنال، دادگان در حوزه زمان فریم بندی می‌شوند، طول هر فریم ۲۵ میلی ثانیه انتخاب شده است. پنجره‌ها از نوع همینگ بوده و دارای همپوشانی ۵۰٪ هستند. همچنین فایل‌های صوتی با توان واحد با هم مخلوط

$$\frac{\partial L'_j}{\partial S_j(k,l)} = -\sum_{m=1}^M (Y_m(k,l) - A_{mj} S_j(k,l))^* A_{mj} \quad (26)$$

در نهایت با حل معادله‌ی (۲۶)، منبع فعال در هر نقطه زمان - فرکانسی، با استفاده از تخمین گر ML به صورت زیر تخمین زده می‌شود:

$$\hat{S}_j(k,l) = \left(\frac{\sum_{m=1}^M Y_m(k,l) A_{mj}^*}{\sum_{m=1}^M |Y_{mj}|^2} \right) \quad (27)$$

با توجه به رابطه‌های (۲۵) و (۲۷)، برای تخمین منبع و جداسازی آن یک اندازه‌گیری جدید از نزدیکی هر نقطه زمان- فرکانس به منبع، تعریف می‌شود که با توجه به آن هر نقطه زمان - فرکانسی با استفاده از میزان نزدیکی، به یک منبع اختصاص می‌یابد:

$$j(k,l) = \quad (28)$$

$$\underset{n}{\operatorname{argmin}} \sum_{m=1}^M \left| Y_m(k,l) - \hat{A}_{mn} \frac{\sum_{m=1}^M Y_m(k,l) \hat{A}_{mn}^*}{\sum_{m=1}^M |\hat{A}_{mn}|^2} \right|^2$$

که \hat{A}_{mn} ، با استفاده از رابطه (۲۲) محاسبه شده و بیانگر اختلاف ترکیب نسبی برای $m-1$ امین هیدروفون به ازای منبع n می‌باشد. سپس منابع اصلی توسط ماسک باینری زیر از هم تفکیک می‌شوند:

$$M_n(k,l) = \begin{cases} 1 & j(k,l) = n \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

با ضرب ماسک باینری فوق در منبع فعال تخمین زده شده توسط رابطه (۲۷)، جداسازی کامل و منبع اصلی با فرمول زیر تخمین زده می‌شود:

$$\hat{S}_n(k,l) = M_n(k,l) \left(\frac{\sum_{m=1}^M Y_m(k,l) \hat{A}_{mn}^*}{\sum_{m=1}^M |\hat{A}_{mn}|^2} \right) \quad (30)$$

به بیان دیگر، سیگنال‌های صوتی حاصل از هیدروفون‌ها در مرحله اول فریم بندی شده و سپس از آن‌ها تبدیل فوریه زمان کوتاه گرفته می‌شود تا به حوزه مختلط زمان- فرکانس انتقال یابند. سپس به عنوان خروجی‌های STFT جهت استخراج بردارهای ویژگی مورد استفاده برای خوشه‌بندی به کار می‌روند. سپس گروه بندی هیدروفون‌ها به جفت‌های هیدروفونی انجام گرفته و اختلاف ترکیب نسبی بین هر

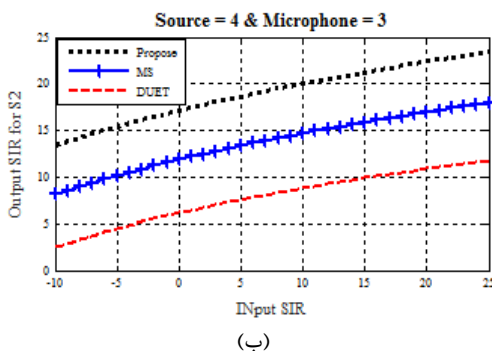
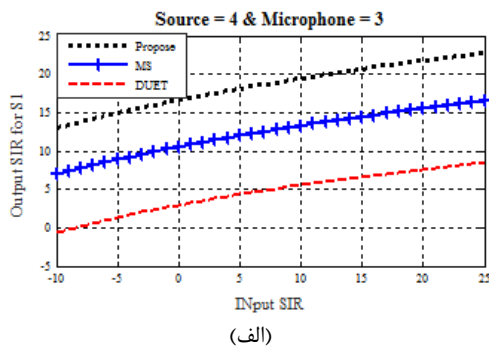
¹ Signal to Interference Ratio

جدول (۲) میزان SIR خروجی بر حسب SIR ورودی در

SNR=10dB با چهار هیدروفون و چهار منبع S_1, S_2, S_3 و S_4 .

-10dB	-5dB	0dB	5dB	10dB	15dB	20dB		
۷/۰۲	۸/۸۱	۹/۵۴	۱۰/۶۱	۱۱/۹۲	۱۲/۸۶	۱۴/۶۶	S1	روش MS
۸/۱۲	۹/۰۲	۱۰/۲۲	۱۲/۳۸	۱۳/۵۱	۱۴/۱۳	۱۵/۹۷	S2	
۷/۵۵	۹/۱۳	۱۰/۸۹	۱۱/۹۴	۱۳/۰۱	۱۴/۹۵	۱۵/۳۴	S3	
۶/۸۱	۷/۸۵	۹/۹۷	۱۰/۸۸	۱۲/۰۸	۱۳/۵۲	۱۵/۲۱	S4	
-۰/۵۸	۱/۴۶	۳/۸۱	۵/۹۸	۶/۴۲	۷/۴۷	۸/۲۱	S1	روش DUET
-۰/۸۸	۱/۹۵	۳/۵۳	۵/۳۰	۷/۰۹	۸/۱۸	۹/۱۸	S2	
۱/۲۱	۲/۵۴	۴/۹۳	۶/۵۱	۶/۸۰	۷/۱۱	۷/۹۷	S3	
۱/۰۱	۲/۲۳	۴/۷۲	۶/۸۲	۷/۱۲	۸/۷۹	۹/۰۳	S4	
۱۳/۰۱	۱۴/۹۱	۱۶/۳۵	۱۷/۹۱	۱۸/۹۱	۲۱/۷۱	۲۳/۲۸	S1	روش پیشنهادی
۱۵/۵۲	۱۵/۳۴	۱۷/۱۱	۱۸/۵۲	۱۹/۲۲	۲۱/۲۲	۲۳/۹۲	S2	
۱۳/۸۷	۱۵/۶۲	۱۷/۲۷	۱۸/۸۹	۱۹/۵۸	۲۰/۶۰	۲۲/۷۶	S3	
۱۲/۹۴	۱۴/۲۶	۱۵/۸۲	۱۷/۱۲	۱۸/۹۶	۱۹/۸۴	۲۱/۹۶	S4	

به همین ترتیب جدول های (۳) و (۴)، مقادیر SIR ناشی از تفکیک منابع صوتی در SNR های مختلف برای سه الگوریتم ذکر شده به ازای SIR ثابت 10dB به ترتیب برای سه و چهار هیدروفون و منبع را نشان می‌دهد. شکل (۲) نیز بیان کننده میزان SIR خروجی حاصل از سه الگوریتم مذکور در SNR های مختلف از -10dB تا 25dB به ازای SIR ثابت 10dB در حالت وجود سه هیدروفون و چهار منبع می‌باشد. نتایج بدست آمده در جدول های (۳) و (۴) و شکل (۲) نشان می‌دهد که درصد بهبود ناشی از اعمال الگوریتم پیشنهادی برای جداسازی اصوات، در مقایسه با تکنیک‌های دیگر قابل توجه می‌باشد. با توجه به جدول (۳) الگوریتم پیشنهادی به طور میانگین نسبت به دو الگوریتم MS و DUET در SNR=10dB به ترتیب 5.3dB و 9.83dB و نیز در SNR=5dB، 5.13 و 6.87dB بهبود داشته است.



می‌گردند. به منظور ارزیابی نتایج کمی، شبیه سازی برای تعداد مختلفی از منابع و هیدروفون‌ها به ازای SNR ها و SIR های مختلف انجام شده است؛ بدین ترتیب که تعداد هیدروفون‌ها ۳ و ۴ و همچنین تعداد منابع صوتی برابر با ۲، ۳ و ۴ انتخاب گردیدند. الگوریتم پیشنهادی، الگوریتم خوشه‌بندی انتقال میانگین و همچنین روش DUET یک بعدی [۱۷] که قابل تعمیم به چند هیدروفون می‌باشد، به تعداد دفعات انجام شده و از نتایج به دست آمده، میانگین گیری شده است. تعداد تکرارهای الگوریتم در حالتی که تعداد منابع ۲، ۳ و ۴ باشد، به ترتیب ۴۰۰، ۲۰۰ و ۱۰۰ بار انتخاب دو فایل صوتی به تصادف و سپس مخلوط کردن و جداسازی آن‌ها بوده است.

جدول های (۱) و (۲)، SIR خروجی سه الگوریتم ذکر شده، حاصل از جداسازی منابع صوتی را به ازای SIR های ورودی متفاوت از -10dB تا 20dB در SNR ثابت 10dB به ترتیب برای سه و چهار هیدروفون و منبع نشان می‌دهد. شکل (۱) نشان دهنده میزان SIR خروجی بر حسب SIR ورودی متفاوت در SNR ثابت 10dB در حالت وجود سه هیدروفون و چهار منبع است.

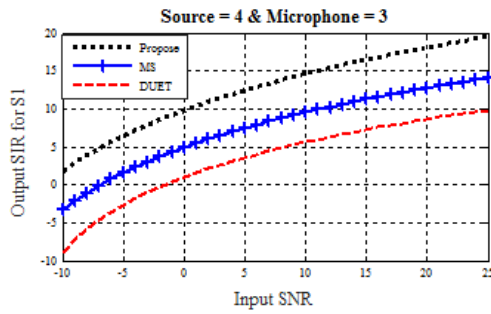
با توجه به نتایج بدست آمده در جدول های (۱) و (۲) مشاهده می‌شود که با وجود افزایش تعداد منابع و در نتیجه افزایش تعداد سیگنال های تداخلی، درصد بهبود الگوریتم پیشنهادی نسبت به دو الگوریتم MS و DUET به خصوص در SIR های پایین قابل توجه می‌باشد. با توجه به جدول (۱) الگوریتم پیشنهادی به طور میانگین نسبت به دو الگوریتم MS و DUET در SIR=-10dB به ترتیب 5.6dB و 6.7dB و نیز در SNR=5dB، 5.78 و 13dB بهبود داشته است. همچنین نتایج بدست آمده در شکل (۱) نشان می‌دهد که در حالتی که تعداد منابع از تعداد هیدروفون‌ها بیشتر باشد نیز الگوریتم پیشنهادی توانسته با درصد خوبی سیگنال منابع را از یکدیگر جدا نماید.

جدول (۱) میزان SIR خروجی بر حسب SIR ورودی در

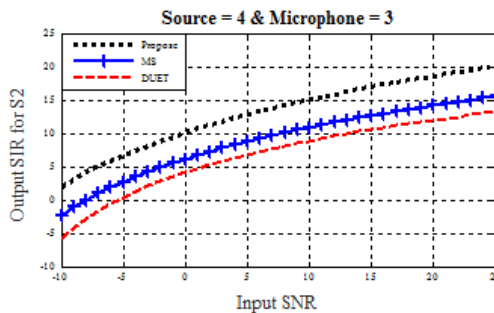
SNR=10dB با سه هیدروفون و سه منبع S_1, S_2 و S_3 .

-10dB	-5dB	0dB	5dB	10dB	15dB	20dB		
۹/۵۱	۱۱/۴۹	۱۳/۱۱	۱۴/۶۰	۱۵/۸۸	۱۷/۰۱	۱۸/۱۱	S1	روش MS
۱۰/۷۳	۱۲/۷۴	۱۴/۵۴	۱۶/۱۲	۱۷/۱۰	۱۸/۵۳	۱۹/۴۸	S2	
۱۰/۲۲	۱۲/۲۳	۱۳/۲۳	۱۵/۱۶	۱۶/۵۲	۱۷/۱۳	۱۸/۸۲	S3	
۲/۲۹	۳/۸۸	۵/۶۳	۷/۷۷	۹/۳۱	۱۰/۲۶	۱۱/۲۲	S1	روش DUET
۳/۱۳	۴/۴۸	۵/۸۸	۷/۵۶	۹/۷۴	۱۰/۲۵	۹/۵۱۰	S2	
۳/۶۶	۵/۶۱	۶/۱۳	۸/۹۰	۱۰/۰۲	۱۱/۲۴	۱۲/۴۱	S3	
۱۵/۳۱	۱۷/۵۱	۱۹/۲۵	۲۰/۷۱	۲۲/۱۳	۲۳/۲۸	۲۴/۴۸	S1	روش پیشنهادی
۱۶/۰۵	۱۸/۱۲	۱۹/۸۲	۲۱/۲۲	۲۲/۲۸	۲۳/۲۶	۲۵/۱۰	S2	
۱۶/۰۸	۱۸/۵۳	۲۰/۲۱	۲۱/۲۰	۲۳/۱۲	۲۴/۲۶	۲۵/۵۰	S3	

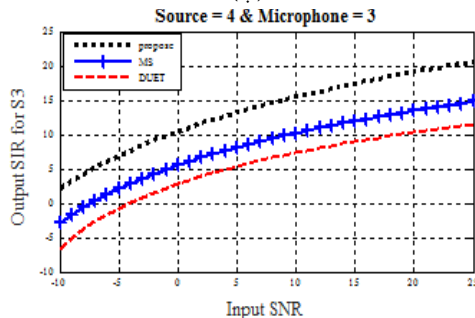
۴/۱۷	۶/۸۳	۱۰/۸۵	۱۲/۹۵	۱۴/۹۸	۱۶/۹۳	
------	------	-------	-------	-------	-------	--



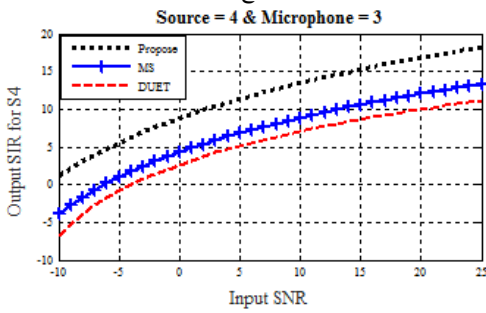
(الف)



(ب)

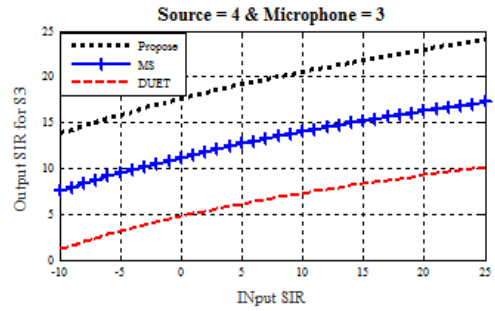


(ج)

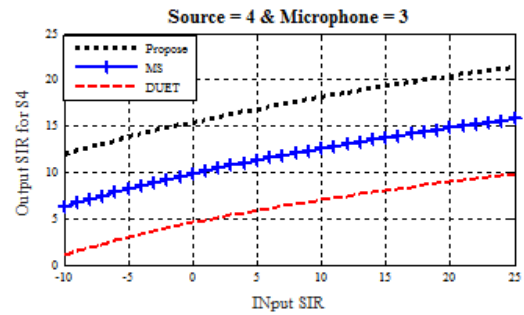


(د)

شکل (۲) میزان SIR خروجی بر حسب SNR ورودی در حالت وجود سه هیدروفون و چهار منبع. (الف) منبع S_1 . (ب) منبع S_2 . (ج) منبع S_3 . (د) منبع S_4 .



(ج)



(د)

شکل (۱) میزان SIR خروجی بر حسب SIR ورودی در حالت وجود سه هیدروفون و چهار منبع. (الف) منبع S_1 . (ب) منبع S_2 . (ج) منبع S_3 . (د) منبع S_4 .

جدول (۵) نیز مقادیر SIR ناشی از تفکیک منابع صوتی در حالت تمیز (بدون نویز) را برای تکنیک‌های MS و DUET یک بعدی و الگوریتم پیشنهادی به ازای وجود سه و چهار هیدروفون و تعداد منابع مختلف مورد مقایسه قرار داده است.

جدول (۳) میزان SIR خروجی بر حسب SNR ورودی در SIR=10dB با سه هیدروفون و سه منبع S_1, S_2, S_3 .

	-10dB	-5dB	0dB	5dB	10dB	15dB		
روش MS	-۱/۸۱	۴/۰۳	۷/۲۴	۹/۶۵	۱۱/۷۴	۱۳/۱۸	S_1	
	-۱/۵۰	۴/۷۱	۸/۹۲	۱۰/۳۱	۱۲/۳۴	۱۴/۱۷	S_2	
	-۱/۱۳	۵/۹۸	۸/۳۱	۹/۹۰	۱۱/۱۳	۱۳/۴۹	S_3	
روش DUET	-۵/۵۸	۱/۸۸	۴/۸۳	۷/۲۵	۹/۸۷	۱۰/۰۱	S_1	
	-۴/۱۴	۲/۹۵	۵/۱۷	۸/۴۱	۹/۱۸	۹/۹۲	S_2	
	-۵/۲۳	۲/۲۴	۵/۵۳	۸/۹۸	۱۰/۵۹	۱۱/۹۴	S_3	
روش پیشنهادی	۴/۶۵	۷/۴۴	۹/۸۵	۱۴/۳۸	۱۵/۷۴	۱۷/۵۸	S_1	
	۴/۸۷	۸/۵۰	۱۰/۳۱	۱۵/۶۵	۱۶/۲۳	۱۸/۹۰	S_2	
	۵/۰۴	۸/۸۳	۱۰/۷۲	۱۵/۲۳	۱۶/۴۲	۱۹/۱۳	S_3	

جدول (۴) میزان SIR خروجی بر حسب SNR ورودی در SIR=10dB با چهار هیدروفون و چهار منبع S_1, S_2, S_3, S_4 .

	-10dB	-5dB	0dB	5dB	10dB	15dB		
روش MS	-۲/۱۲	۲/۱۲	۵/۵۶	۸/۱۲	۱۰/۱۱	۱۲/۵۱	S_1	
	-۱/۸۶	۲/۸۸	۵/۶۱	۸/۸۷	۱۰/۹۳	۱۳/۰۷	S_2	
	-۲/۴۴	۳/۱۳	۶/۰۴	۷/۲۱	۹/۸۷	۱۲/۶۸	S_3	
	-۱/۵۷	۳/۰۷	۶/۶۵	۸/۵۱	۱۰/۵۰	۱۲/۹۴	S_4	
روش DUET	-۴/۵۴	-۰/۵۶	۳/۴۵	۵/۶۷	۸/۲۱	۱۰/۱۱	S_1	
	-۵/۲۲	-۰/۳۷	۴/۵۶	۶/۰۳	۹/۱۳	۱۱/۰۷	S_2	
	-۴/۸۰	۱/۲۴	۳/۸۷	۵/۱۲	۸/۰۸	۱۰/۳۹	S_3	
	-۴/۵۵	-۱/۷۸	۴/۳۰	۶/۱۰	۹/۴۵	۱۱/۴۶	S_4	
روش پیشنهادی	۳/۱۱	۷/۸۷	۱۱/۲۳	۱۳/۶۷	۱۶/۳۲	۱۷/۶۰	S_1	
	۳/۴۲	۸/۰۱	۱۱/۶۰	۱۴/۱۲	۱۵/۸۵	۱۸/۱۲	S_2	
	۳/۶۸	۸/۲۴	۱۱/۵۹	۱۴/۵۸	۱۶/۲۴	۱۸/۷۴	S_3	

[5] Y. Qingming, "Blind source separation by weighted k-means clustering," *Journal of systems engineering and electronics*, vol. 19, no. 5, pp. 882-887, 2008.

[6] D. Ayllon, G. Pita, Pj. Amores, MR. Zurera, and CL. Aguilar, "Energy-weighted mean shift algorithm for speech source separation," *IEEE Statistical Signal Processing Workshop (SSP)*, pp. 785-788, 2011.

[7] J. Jayaraman, and J. Thiagara, "Mixing matrix estimation using discriminative clustering for blind source separation," *Digital Signal Processing ELSEVIR*, vol. 23, issue. 1, pp. 9-18, 2013.

[8] P. Smaragdis, C. fevotte, G.J. Mysore, N. Mohammadiha and M. Hoffman, "Static and dynamic source separation using nonnegative factorization: Aunified view," *IEEE signal processing magazine*, vol. 31, no. 3, pp. 66-75, 2014.

[9] L. Zou, Z. J. Wang, X. Chen, and X. Ji, "Underdetermined joint blind source separation based on tensor decomposition," *Electrical and computer engineering (CCECE)*, 2016.

[10] H. Jia and X. Li, "underwater reverberation suppression based on non-negative matrix factorisation," *Journal of Sound and Vibration*, pp.116-166, 2021.

[11] D. Ayllon, G. Pita, PJ. Amores, and MR. Zurera, "Speech source separation using a generalized mean shift algorithm," *Signal Processing Elsevier*, vol. 92, pp. 2248-2252, 2012.

[12] R. P. Fernandes and J. A. Apolinário Jr, underwater target classification with optimized feature selection based on Genetic Algorithms, 2020.

[13] D. Kounades-Bastian, L. Girin, X. Alameda-pineda, S. Gannot, and R. Horaud, "An EM algorithm for joint source separation and diarisation of multichannel convolutive speech mixtures," *IEEE international conference on acoustics*, New Orleans, 2017.

[14] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, issue. 5, pp.603-619, 2002.

[15] San Francisco Maritime National Park Association, (2017), "Historic naval sound and video", Available from: <https://maritime.org/sound/index.htm>.

[16] J. Sanz-Robinson, L. Huang, and T. Moy, "Large-Area Microphone Array for Audio Source Separation Based on a Hybrid Architecture Exploiting Thin-Film Electronics and CMOS," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 4, pp. 979 – 991, 2016.

[17] S. S. Fadaei and M. H. Kahaei, "Blind separation of speech signal using one dimensional block DUET," *Int jour. Of Elect. and comp. Eng*, pp.103-110, 2008.

جدول (۵) مقایسه مقادیر SIR ناشی از تفکیک منابع صوتی در حالت تمیز در حالت وجود سه و چهار هیدروفون و تعداد منبع صوتی مختلف.

M=3			M=4			تعداد هیدروفون	
DUET	MS	Proposed method	DUET	MS	Proposed method	تعداد منابع	
۱۷/۳۹	۱۵/۸۶	۱۸/۷۳	۲۲/۰۸	۱۸/۰۳	۲۰/۳۸	S1	۲
۱۴/۷۶	۱۵/۷۴	۱۶/۴۰	۱۸/۱۳	۱۹/۵۳	۱۸/۱۶	S2	
۱۶/۰۸	۱۵/۸۰	۱۷/۵۷	۱۹/۲۲	۱۸/۷۸	۱۹/۲۲	میانگین	
۱۱/۷۹	۱۰/۸۶	۱۲/۹۹	۱۲/۰۴	۱۱/۸۹	۱۴/۶۰	S1	۳
۱۱/۲۱	۱۱/۵۸	۱۱/۶۷	۱۰/۷۲	۱۰/۷۵	۱۳/۴۶	S2	
۱۰/۵۷	۱۰/۹۲	۱۲/۱۲	۱۱/۰۵	۱۰/۸۶	۱۳/۱۴	S3	
۱۱/۲۵	۱۱/۱۲	۱۲/۲۶	۱۱/۲۷	۱۱/۱۷	۱۲/۷۴	میانگین	
۷/۰۷	۶/۹۳	۸/۵۱	۷/۴۹	۷/۶۳	۹/۸۸	S1	۴
۶/۰۱	۶/۰۴	۷/۲۵	۶/۶۳	۶/۲۱	۸/۶۰	S2	
۵/۵۸	۵/۸۱	۷/۱۷	۷/۰۵	۶/۸۸	۹/۱۲	S3	
۶/۴۸	۵/۸۲	۷/۰۱	۶/۹۱	۶/۹۲	۹/۲۸	S4	
۶/۲۶	۶/۱۵	۷/۴۹	۷/۰۲	۶/۹۱	۹/۲۲	میانگین	

نتیجه گیری

با توجه به نتایج بدست آمده مشاهده می شود که کارایی الگوریتم پیشنهادی در جداسازی منابع از یکدیگر، در حضور نویز و در جهت بهبود کیفیت شنیداری سیگنال صوتی قابل قبول می باشد. این بهبود به ویژه در شرایطی که سیگنال تحت تأثیر نویز تخریب می گردد، چشمگیرتر بوده و لذا به کارگیری آن در کاربردهای واقعی مناسبتر است. قابلیت جداسازی کاملاً کور منابع صوتی با وجود تعداد متغیر منبع و هیدروفون بدون اطلاع از نحوه چیدمان هندسی و دقت مناسب آن در قیاس با دیگر روشها نیز از مزایای این الگوریتم پیشنهادی به شمار می روند.

منابع

- [1] G. Hu, K. Wang and L. Liu, "underwater acoustic target recognition based on depthwise separable convolution neural networks," *Sensors*, vol. 21, no. 4, pp.1429, 2021.
- [2] T. Zhang, S. Liu, X. He, H. Huang and K. Hao, "underwater target tracking using forward-looking sonar for autonomous underwater vehicles," *Sensors*, vol. 20, no. 1, pp. 102, 2020.
- [۳] صیادی، حسن، سیف، محمدسعید، آب نیکی، علی اصغر، دسته بندی و شناسایی اهداف زیرآبی بر اساس اصوات منتشره، نشریه مهندسی دریا، سال سیزدهم، شماره ۲۶، پاییز و زمستان ۱۳۹۶.
- [4] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via Time-Frequency masking," *IEEE Trans on Signal Processing*, vol. 52, no. 7, pp. 1830-1847, 2004.