

# راهکار جدید استخراج ویژگی مبتنی بر

## نمونه برداری فشرده در پردازش

### سیگنال‌های صوتی

مهدی بنی‌طالبی دهکردی<sup>۱</sup>، حمیدرضا ابوطالبی<sup>۱</sup>، محمد تقی صادقی<sup>۱</sup> و جوزف کیتلر<sup>۲</sup>  
گروه پژوهشی پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، یزد، ایران  
گروه پردازش تصویر، گفتار و پردازش سیگنال، دانشکده مهندسی برق و کامپیوتر، دانشگاه ساری، انگلستان

#### چکیده

در این مقاله با استفاده از راهکار نمونه برداری فشرده، الگوریتمی برای استخراج ویژگی از سیگنال‌های صوتی معرفی می‌شود. در روش پیشنهادی، ابتدا سیگنال صوت، پنجره گذاری شده و تبدیل فوریه نمونه‌های درون هر پنجره محاسبه می‌شود. سپس مقادیر دامنه‌های هنجار شده ضرایب به دست آمده در هر پنجره با هم جمع شده و از بردار حاصله که مؤلفه‌های آن مجموع دامنه ضرایب در هر پنجره است دوباره تبدیل فوریه گرفته می‌شود؛ در نهایت با توجه به تُنکی ایجاد شده، از بردار به دست آمده به صورت تصادفی نمونه برداری می‌شود. در این پژوهش در کاربردهای مختلفی، از بردار ویژگی به دست آمده استفاده شده است. از جمله این کاربردها می‌توان به طبقه بندی اصوات و مکان یابی منابع صوت اشاره کرد. در شبیه سازی‌های صورت گرفته نشان داده می‌شود که در مقایسه با برخی طبقه بندی کننده‌های مطرح دیگر، طبقه بندی کننده مبتنی بر ویژگی ارائه شده، دقت بیش تر و بار محاسباتی کمتر دارد. همچنین در شبیه سازی‌ها نشان داده شده است که با استفاده از این الگوریتم استخراج ویژگی، موقعیت منابع را با خطای کمتر از دو درصد می‌توان، تعیین کرد.

واژگان کلیدی: نمونه برداری فشرده، استخراج ویژگی، طبقه بندی اصوات، مکان یابی منابع صوت.

#### ۱- مقدمه

در حالت کلی می‌توان ویژگی‌های آوایی را به دو دسته زمان کوتاه و زمان طولانی تقسیم کرد. ویژگی‌های زمان کوتاه به طور اساسی بر مبنای کمیت‌های مشتق شده از طیف یک سیگنال در یک بازه زمانی کوچک (به عنوان مثال در یک قاب) هستند؛ از جمله این دسته از ویژگی‌ها می‌توان به ضرایب فرکانسی کپسوترا ل مل (MFCC)<sup>۱</sup> (Tzanetakis, 2002) و اختلاف فرکانسی مبتنی بر تقابل طیفی (Zhang, 2002)، اشاره کرد. در مقابل، در ویژگی‌های زمان طولانی تغییرات شکل طیفی و یا اطلاعات مربوط به اکستریم‌های سیگنال در یک بازه زمانی بزرگ، مورد بررسی قرار می‌گیرد؛ از جمله ویژگی‌های زمان طولانی به هیستوگرام

ضرایب موجک (Daubechies Wavelet Coefficient Histograms (DWCH)<sup>۲</sup> (Ogihara, 2003) می‌توان اشاره کرد.

به طور عمومی در پردازش و طبقه بندی سیگنال‌های صوتی مانند سامانه‌های بازیابی اطلاعات موسیقی (MIR)<sup>۳</sup> از هر دو نوع ویژگی زمان کوتاه و زمان طولانی به صورت هم زمان استفاده می‌شود. در نتیجه به طور معمول در سامانه‌های طبقه بندی کننده مشابه، با تعداد زیاد ویژگی روبه رو هستیم که باعث پیچیدگی زیاد سامانه و در نتیجه تحمیل بار محاسباتی زیاد به سامانه می‌شود.

به منظور جلوگیری از پیچیدگی سامانه و کاهش بار محاسباتی، به طور معمول روش‌های کاهش تعداد ویژگی‌ها در سامانه‌های طبقه بندی کننده به کار می‌رود. از جمله

<sup>۲</sup> Daubechies Wavelet Coefficient Histograms

<sup>۳</sup> Music Information Retrieval

<sup>۱</sup> Mel Frequency Cepstral Coefficients

## ۲- نمونه برداری فشرده (CS)<sup>۳</sup>

این زمینه تحقیقاتی که از حدود چهار دهه پیش بوجود آغاز شد (Sreenivas, 2009)، امروزه نتایج کاربردی بسیار خوبی در حوزه‌های مختلف پیدا کرده است. اکثر سیگنال‌ها، به صورت ذاتی تُنک هستند؛ به این معنی که برای آنها فضای نمایشی وجود دارد که در آن فضا سیگنال حاوی تعداد زیادی نمونه صفر یا با انرژی پایین باشد (Sreenivas, 2009). در این راه کار با در نظر گرفتن بردارهای پایه مناسب سیگنال مورد مطالعه اندازه گیری می‌شود؛ سپس با ترکیب خطی وزن دار نمونه‌های اندازه گیری شده سیگنال به صورت فشرده نمونه برداری می‌شود. عمل بازسازی سیگنال از روی این نمونه‌های تصادفی منجر به حل یک معادله ماتریسی غیرخطی می‌شود. تئوری فوق را می‌توان در دسته الگوریتم‌های کاهش دهنده تعداد بُعد در نظر گرفت. هدف در نمونه برداری فشرده، به دست آوردن نمایشی با تعداد بُعد کمتر از سیگنال  $x$  به صورتی است که بتوان تخمین خوبی از سیگنال را از آن به دست آورد. این تئوری شباهت زیادی با مسئله سنتی فشرده سازی داده‌ها دارد؛ با این تفاوت که در نمونه برداری فشرده در قسمت کدکننده، نیازی به داشتن اطلاعات پیشین از چگونگی ساختار سیگنال نیست؛ فقط در قسمت کدگذاری کننده از مدل تُنک بودن برای بازسازی سیگنال استفاده می‌شود.

راه حل مرسوم برای پیاده سازی ایده فوق برای عمل نمونه برداری، مفهوم حداقل کردن نرم  $l_2$  است (که به معنای حداقل کردن انرژی در سیگنال بوده و با علامت  $\| \cdot \|_2$  نشان داده می‌شود) و عمل بازسازی، به مفهوم حداقل کردن نرم  $l_0$  است (به معنای کم کردن تعداد ضرایب صفر در بسط سیگنال که با علامت  $\| \cdot \|_0$  نشان داده می‌شود). ثابت شده که با پذیرش خطای اندک، به جای حداقل کردن نرم  $l_0$ ، استفاده از نرم  $l_1$  (جمع نمونه‌های سیگنال) منجر به مسأله ساده تری می‌شود (Wakin, 2009).

پس در یک جمع بندی می‌توان گفت تئوری نمونه برداری فشرده بیان می‌کند که اگر خطای جزئی  $\epsilon$  را بپذیریم از روی تعداد نمونه خیلی کمتر تحت شرایطی خاص

روش‌های مرسوم کاهش بُعد به PCA<sup>۱</sup> (Yap, 2010) و LDA<sup>۲</sup> (Every, 2008) می‌توان اشاره کرد. یکی دیگر از الگوریتم‌هایی که در کاهش بعد بردار ویژگی استفاده می‌شوند، الگوریتم نمونه برداری فشرده (Sreenivas, 2009) است. در این الگوریتم برخلاف سایر الگوریتم‌ها، استخراج بردار ویژگی به صورت تصادفی صورت می‌گیرد.

در این مقاله با استفاده از الگوریتم نمونه برداری فشرده، راه کار جدیدی برای استخراج بردار ویژگی با بُعد کم ارائه می‌شود. در این راه کار، با توجه به ماهیت سیگنال مورد بررسی از تبدیل‌های مختلف مانند FFT، DCT، ضرایب موجک و ...، برای تُنک سازی (کم کردن تعداد ضرایب غیر صفر بسط) سیگنال استفاده می‌شود. به عنوان نمونه برای تُنک سازی سیگنال صوت از FFT استفاده شده است.

در الگوریتم استخراج ویژگی پیشنهادی، برای تولید بردار ویژگی، سیگنال صوت، با به کارگیری پنجره همینگ پنجره گذاری شده و از نمونه‌های درون هر پنجره تبدیل فوریه (FFT) گرفته می‌شود؛ سپس دامنه‌های هنجار شده ضرایب به دست آمده در هر پنجره با هم جمع شده و از بردار حاصله که مؤلفه‌های آن مجموع دامنه ضرایب در هر پنجره است، دوباره تبدیل فوریه گرفته و در نهایت با توجه به تُنکی ایجاد شده با استفاده از قوانین حاکم بر نمونه برداری فشرده، از بردار به دست آمده به صورت تصادفی نمونه برداری می‌شود. برداری که در قسمت آخر ایجاد می‌شود حاوی اطلاعات کافی برای استفاده در طیف وسیعی از سامانه‌های طبقه بندی و پردازش سیگنال هاست. در این تحقیق برای نشان دادن کارآمدی روش پیشنهادی، کاربرد این روش در طبقه بندی اصوات و نیز در مکان یابی منابع صوت مورد بررسی قرار می‌گیرد.

در ادامه این مقاله، ابتدا در بخش دو راه کار نمونه برداری فشرده به اختصار شرح داده می‌شود. سپس در بخش سه، الگوریتم استخراج ویژگی پیشنهادی توضیح داده شده و در بخش چهار، عمل کرد ویژگی ارائه شده در دو زمینه طبقه بندی اصوات و مکان یابی منابع صوت، مورد بررسی قرار می‌گیرد. در انتها، مطالب ارائه شده جمع بندی می‌شود.

<sup>1</sup> Principal component analysis

<sup>2</sup> Linear Predictive Coding

<sup>3</sup> Compressive Sampling

شده است (Panagikas,2010). از میان این ویژگی‌ها MFCC، مرکز ثقل طیفی<sup>۳</sup> و نرخ عبور از صفر از جمله ویژگی‌های زمان کوتاه است که برای به دست آوردن آنها محاسبات در یک بازه زمانی کوتاه انجام می‌شود. از جمله ویژگی‌های زمان طولانی نیز می‌توان به ویژگی انرژی کم اشاره کرد.

(جدول ۱) - انواع ویژگی‌های به کار رفته در پردازش اصوات

ویژگی	توضیحات در مورد ویژگی مورد بررسی
MFCC	نمایش خصوصیات طیفی بر اساس ضرایب فرکانسی کپسترال میل
Spectral centroid	مرکز ثقل دامنه‌های طیفی
Spectral roll-off	فرکانس‌هایی که بیش از ۸۵ درصد انرژی در آنها ذخیره شده باشد.
Spectral flux	مجذور فاصله طیفی متوالی
Zero crossings	تعداد عبور دامنه طیفی از صفر
Low-energy	تعداد پنجره‌هایی که انرژی کمتر از میانگین انرژی کل پنجره‌ها داشته باشند.

در این تحقیق، با اعمال ویژگی‌های زمان گسترده بر روی ویژگی‌های زمان کوتاه، بردار ویژگی از سیگنال‌های صوتی استخراج می‌شود. روند نامی روش پیشنهادی به منظور استخراج بردار ویژگی که با نام اختصاری  $CS-2FFT$  نامیده شده است، در شکل (۱) آورده شده است.

همان گونه که از شکل (۱) مشخص است، در مرحله اول سیگنال مورد نظر با استفاده از پنجره همینگ با هم پوشانی پنجاه درصد، پنجره گذاری می‌شود. در مرحله بعد از نمونه‌های موجود در هر پنجره FFT گرفته می‌شود. دامنه‌های طیفی حاصله هنجار شده و با هم جمع می‌شوند. با دنبال کردن روند الگوریتم ایجاد بردار ویژگی تا این مرحله، به هر پنجره، یک عدد نظیر می‌شود که این عدد به نوعی دربرگیرنده اطلاعات مربوط به انرژی سیگنال است. از کنار هم قراردادن اعداد ایجاد شده توسط پنجره‌های شیفت داده شده بردار جدیدی ایجاد می‌شود. در مرحله بعد از بردار ایجاد شده دوباره FFT گرفته می‌شود. به منظور برجسته کردن تُنکی و صفر کردن درایه‌های با دامنه کم، خروجی مرحله قبل از یک فیلتر دامنه عبور می‌کند

داد. در  $CS$  ضرایب مورد نظر به صورت مستقیم اندازه گیری نمی‌شوند، بلکه  $M$  نمونه ( $M < N$ ) از سیگنال مورد نظر ( $x$ ) با رابطه  $y(m) = x, \phi_m^T$  (که در آن  $\phi_m$  مجموعه پایه‌هایی هستند که بر روی آنها سیگنال تُنک، و  $N$ ، طول سیگنال است)، انتخاب می‌شوند؛ فرم ماتریسی به فرم زیر است:

$$Y = \Phi x \quad (1)$$

در رابطه (۱)،  $Y$  یک بردار  $M \times 1$  ستونی و  $\Phi$  ماتریسی  $M \times N$  از پایه‌های مورد استفاده در نمونه برداری است. سطرهای ماتریس  $\Phi$  بردارهای پایه  $\phi_m$  هستند. در حالت کلی زمانی که  $M < N$  باشد بازسازی سیگنال از روی نمونه‌های برداشته شده، دچار مشکل می‌شود اما اگر فرض تُنک بودن سیگنال به فرض‌های بازسازی اضافه شود، بازسازی سیگنال  $x$  امکان پذیر خواهد بود.

فرض کنیم  $\{\psi_m\}$  مجموعه پایه‌هایی باشد که سیگنال مورد بررسی ( $x$ ) بر روی آن به صورت تُنک، افزاز شده است. تئوری نمونه برداری فشرده ( $CS$ ) بیان می‌کند زمانی که دو مجموعه پایه‌های  $\{\psi_m\}$  و  $\{\phi_m\}$  ناهمدوس<sup>۱</sup> (منظور از ناهمدوس بودن، استقلال درایه‌های مختلف از هم و عدم تأثیر آنها بر یکدیگر است) باشند و تعداد نمونه  $M$  به اندازه کافی بزرگ باشد، عملیات بازسازی ضرایب بزرگ بسط مورد استفاده در پردازش صورت گرفته بر روی سیگنال مورد نظر از تعداد نمونه خیلی کمتر با الگوریتم ( $CS$ ) امکان پذیر خواهد بود (Wakin,2009).

### ۳- روش استخراج ویژگی پیشنهادی

از جمله ویژگی‌های مورد استفاده در پردازش‌های گفتاری می‌توان به ویژگی شیوش صوتی<sup>۲</sup> (Panagikas,2010)، ویژگی‌های مبتنی بر اختلاف متقابل طیفی، ویژگی‌های مبتنی بر مدولاسیون فاصله‌های متقابل طیفی، مدولاسیون اندازه‌های طیفی مشابه و مدولاسیون اندازه‌های بزرگ طیفی اشاره کرد (Lee,2009).

برخی ویژگی‌های ساختاری شیوش صوتی که بیشتر برای طبقه بندی اصوات به کار می‌روند، در جدول (۱) آورده

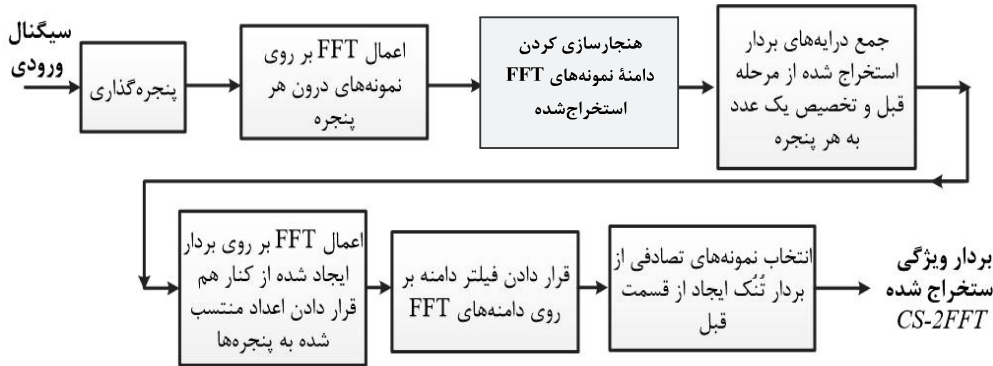
<sup>1</sup> Incoherent  
<sup>2</sup> Timbral

<sup>3</sup> Spectral Centroid

با بیان دقیق‌تر، همان‌طور که در شکل (۲) نشان داده شده است، فرض می‌شود سیگنال  $x$  بر روی ماتریس پایه‌های  $\Psi$  به صورت تُنک از درجه  $K$  باشد، یعنی اینکه  $K$  ضریب غیر صفر ( $K$  مقداری کوچک است) در بسط سیگنال روی ماتریس پایه‌های  $\Psi$  باشد، همچنین فرض می‌شود  $\Phi$  ماتریس  $M \times N$  ( $M \ll N$ )، ماتریس پایه‌های نمونه‌ها و به‌صورتی باشد که سطرهای آن با ستون‌های  $\Psi$  ناهم‌بسته شوند. (Chang,2010) و (Sainath,2010).

(آستانه‌گذاری بر روی دامنه نمونه‌ها). در مرحله آخر الگوریتم استخراج ویژگی از بردار تُنک به‌دست آمده با ضرب ماتریس نمونه بردار تصادفی گوسی عملیات نمونه‌برداری انجام می‌شود.

تعداد نمونه تصادفی برداشته شده بستگی به درجه تُنکی بردار دارد. در اینجا با توجه به ماهیت صوتی سیگنال و با آزمون و خطا این تعداد تعیین شده است. با توجه به اهمیت فرآیند نمونه‌برداری فشرده در روش پیشنهادی، در بخش آتی تئوری نمونه‌برداری فشرده، به‌اختصار بررسی خواهد شد.



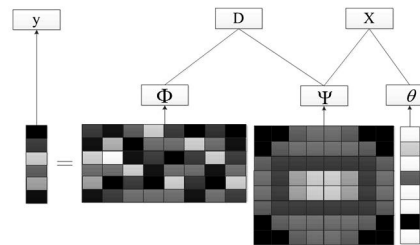
(شکل ۱) - شمای کلی چگونگی استخراج بردار ویژگی از سیگنال ورودی (در اینجا سیگنال صوت)

$$\min \|\theta\|_0 \quad S.T. \quad Y = \Phi x = D\theta \quad (3)$$

در حالت کلی حل معادله بهینه‌سازی فوق با پیچیدگی‌های زیادی همراه خواهد بود؛ در عمل برای دستیابی به یک جواب مطلوب، در معادله بالا به جای نُرم  $l_0$  ( $\|\cdot\|_0$ ) از نُرم  $l_1$  ( $\|\cdot\|_1$ ) استفاده می‌کنند [۶-۷]:

$$\min \|\theta\|_1 \quad S.T. \quad Y = \Phi x = D\theta \quad (4)$$

با توجه به مطالب ارائه شده در بالا، در پردازش سیگنال به‌جای برداشتن  $N$  نمونه از یک سیگنال از درجه تُنکی  $K$ ، تنها  $M$  نمونه متعامد از سیگنال، انتخاب می‌شود. بنابراین حس‌گرها هنگام انتقال اطلاعات تعداد بسیار کمتری نمونه را به گیرنده انتقال می‌دهند که در گیرنده می‌توان با استفاده از تعداد نمونه کمتر سیگنال را بازسازی کرد و پردازش‌های لازم را بر روی آن انجام داد. بردارهای پایه زیادی وجود دارد که می‌توان برای انتخاب  $M$  نمونه فوق از آنها استفاده کرد؛ دسته‌ای از این بردارها بردارهای تصادفی گوسی هستند. بردارهای تصادفی  $i.i.d$  گوسی یا برنولی (صفر و یک) یا برنولی ( $\pm 1$ ) پایه‌های مهم و پرکاربردی هستند که با سایر پایه‌ها متعامدند؛ بنابراین در نمونه‌برداری



(شکل ۲) - شمای کلی تئوری نمونه‌برداری فشرده

با توجه به این مطالب خواهیم داشت:  $x = \Psi\theta$  (در عبارت فوق  $\theta$  ماتریس ورودی است). تئوری نمونه‌برداری فشرده بیان می‌کند که سیگنال  $x$  با استفاده از بازایی است (Sreenivas,2009) و (Wakin,2009):

$$Y = \Phi x = \Phi \Psi \theta = D\theta \quad (2)$$

در معادله (۲)،  $Y$  برداری  $M \times 1$  و معرف بردار نمونه‌هاست،  $D = \Phi \Psi$  نیز یک ماتریس  $M \times N$  است. بازسازی سیگنال متناظر با یافتن بردار ضرایب تُنک  $\theta$  است، این مسئله باعث به‌وجود آمدن مسئله بهینه‌سازی نُرم  $l_0$  زیر می‌شود (Sreenivas,2009) و (Wakin,2009):

در این عبارت  $\theta = [0, \dots, 0, \theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,n}, 0, \dots, 0] \in R^n$  بردار ویژگی گسترده شده‌ای است که  $n = \sum_{i=1}^K n_i$  بردار ویژگی گسترده شده، بردار ضرایبی است که درایه‌های آن به جز درایه‌های مربوط به طبقه  $i$  ام، همگی صفر هستند. حال اگر  $\theta$  بردار ویژگی گسترده شده یک نمونه آزمایشی دلخواه باشد، برای به دست آوردن بردار ویژگی این نمونه باید از حل معادله (۶) استفاده کنیم (Panagikas, 2010).

در طبقه بندی کننده‌های مبتنی بر تُنکی ( $SRC^1$ )، برای یک نمونه آزمایشی  $x$  از یکی از طبقه‌ها ابتدا  $\hat{\theta}$  نمایش تُنک داده ورودی ( $x$ ) را با معادله (۳) و یا (۴) به دست می‌آوریم. در حالت ایده آل، عناصر غیر صفر  $\hat{\theta}$  همگی متناظر با ستون‌هایی از  $A$  هستند که به طبقه  $i$  مربوط می‌شوند؛ پس می‌توان به سادگی نمونه مورد نظر را مربوط به طبقه فوق دانست، اما در عمل به دلیل وجود نوفه و عدم شباهت کامل داده ورودی با داده‌های طبقه مربوطه این شرایط ایده آل برقرار نمی‌شود، بنابراین برای طبقه  $i$  ام تابع ویژگی  $\delta_i: R^n \rightarrow R^n$  را به گونه‌ای تعریف می‌کنیم که ضرایب مربوط به طبقه  $i$  ام را جدا سازد، یعنی خروجی تابع  $\delta_i(\hat{\theta}) \in R^n$ ، یک بردار جدید است که درایه‌های غیر صفر آن درایه‌هایی از  $\hat{\theta}$  هستند که به طبقه  $i$  مربوط شود. حال با فرض اینکه داده متعلق به طبقه  $i$  ام باشد، با استفاده از معادله (۶)، می‌توان  $\hat{x}_i$  را تخمین زد، یعنی:

$$\hat{x}_i = A \delta_i(\hat{\theta}) \quad (7)$$

بعد از محاسبه  $\hat{\theta}_i$  ها و  $\hat{x}_i$  ها به ازای  $i = 1, 2, \dots, K$  می‌توان گفت که  $x_i$  متعلق به طبقه‌ای است که کمترین خطای تخمین را داشته باشد، یعنی:

$$\min r_i(y) = \left\| x - \hat{x}_i \right\|_2 \quad (8)$$

در طبقه بندی کننده‌های SRC پایه، هیچ گونه کاهش بُعدی صورت نمی‌گیرد؛ ولی در الگوریتمی که در اینجا ارائه شده، با توجه به فرآیند استخراج بردار ویژگی به کار رفته، کاهش بُعد صورت می‌گیرد. بردار ویژگی پیشنهادی ضمن کاهش بُعدی که در اثر چگونگی روند استخراج بردار ویژگی

فشرده می‌توان با استفاده از یک دسته از پایه‌های تصادفی فوق، از هر سیگنال دلخواهی عمل نمونه برداری را بدون ارتباط به نوع سیگنال و پایه‌های آن انجام داد.

## ۴- بررسی کیفیت عملکرد ویژگی ارائه شده در کاربردهای مختلف

در ادامه دو مورد از کاربردهای روش استخراج ویژگی ارائه شده، در پردازش سیگنال‌های گفتاری ارائه و از طریق شبیه سازی کیفیت عملکرد سیستم‌های مربوطه بررسی می‌شود.

### ۴-۱- طبقه بندی کننده مبتنی بر نمونه برداری فشرده ( $CS-2FFT-SRC$ ):

از تئوری نمونه برداری فشرده و تئوری نمایش سیگنال‌ها به صورت تُنک می‌توان برای طراحی طبقه بندی کننده‌ها استفاده کرد.

فرض کنید  $v_{i,j} \in R^M$  بردار ویژگی استخراج شده از  $i$  امین نمونه آزمایشی سیگنال صوت در  $i$  امین طبقه باشد. علاوه بر این فرض کنید تعداد کافی نمونه آزمایشی برای  $i$  امین طبقه به صورت  $[v_{i,1}, v_{i,2}, \dots, v_{i,n_i}]$  موجود باشد.

با استفاده از فرض‌های بالا می‌توان هر نمونه جدید  $x_i \in R^M$  (بردار ویژگی استخراج شده از سیگنال آزمایش) از طبقه مشابه را به صورت تقریبی با استفاده از رابطه خطی زیر نمایش داد:

$$x_i = \sum_{j=1}^{n_i} \theta_{i,j} v_{i,j} \quad (5)$$

در عبارت بالا  $\theta_{i,j}$  ها ضرایب وزن هستند. با در نظر گرفتن بردارهای ویژگی مربوط به داده‌های آموزشی طبقه  $i$  ام ( $v_{i,j}$  ها) به عنوان ستون‌های ماتریس  $A_i$ ، ماتریس  $A$  با عمل کردی مشابه با ماتریس  $\Psi$  در رابطه (۲)، به صورت  $A = [A_1, A_2, \dots, A_K]$  تعریف می‌شود. رابطه خطی ارائه شده برای محاسبه  $x$  در معادله (۵) بر حسب همه نمونه‌های آموزشی به صورت زیر به دست می‌آید:

$$x = A\theta \in R^M \quad (6)$$

<sup>1</sup> Sparse Representation based Classifier

مبتنی بر نمونه برداری فشرده حاصل می شود، همان گونه که در شبیه سازی ها به دو مورد از کاربردهای پردازشی این بردار ویژگی اشاره شده است، اطلاعات کافی برای انجام پردازش را در خود نگه می دارد و قسمت های اضافی را حذف می کند. بنابراین در یک جمع بندی از مطالب ارائه شده می توان گفت که با توجه به خواص تئوری نمونه برداری فشرده، کاهش بُعد، ساختار ویژگی ها را تغییر می دهد و تا حد امکان قسمت های اضافی را حذف می کند.

همان گونه که گفته شد در روش های مرسوم طبقه بندی با استفاده از ویژگی های طیفی، بعد از پنجره گذاری بر روی سیگنال مورد نظر، از قسمت های به وجود آمده ویژگی های طیفی را استخراج می کنند و با استفاده از آنها طبقه بندی صورت می گیرد، در راه کار جدیدی که در اینجا ارائه داده شده، بعد از پنجره گذاری یک بردار ویژگی با طول بسیار کمتر از ویژگی های مرسوم تولید می شود. در شبیه سازی های صورت گرفته نشان داده شد که ضمن این کاهش طول، اطلاعات مربوطه در آن حفظ می شود.

خلاصه ای از الگوریتم فوق در جدول (۲) آورده شده است.

(جدول ۲) - الگوریتم ارائه شده برای طبقه بندی اصوات

گام های الگوریتم طبقه بندی کننده مبتنی بر نمونه برداری فشرده	
گام ۱	استخراج ویژگی ها برای $K$ طبقه بر روی نمونه های آموزشی
گام ۲	با استفاده از ویژگی های استخراج شده، ماتریس $A$ را ایجاد می کنیم.
گام ۳	چون ماتریس $A$ تَنُک است برای استفاده از راه کار نمونه برداری تصادفی از آن نمونه برداری می کنیم. این عمل با ضرب کردن یک ماتریس نمونه بردار ( به عنوان مثال گوسی تصادفی $(\Phi)$ در $A$ انجام می شود.
گام ۴	ماتریس تصادفی نمونه برداری شده را هنجار می کنیم تا نرم دوم آن برابر با ۱ گردد، سپس برای تمام طبقه ها بازسازی نمونه آزمایشی را انجام می دهیم.
گام ۵	خطای تخمین را برای همه طبقه ها محاسبه می کنیم
گام ۶	طبقه ای که کم ترین خطا را داشته باشد انتخاب می کنیم.

در شبیه سازی هایی که صورت گرفته است از ویژگی مطرح شده در قسمت قبل ( $CS - 2FFT$ ) در ورودی طبقه بندی کننده  $SRC$  استفاده شد. نتایج به دست آمده در ادامه ارائه شده است. لازم به ذکر است در این شبیه سازی ها

از پایگاه دادگان  $GTZAN$ <sup>۱</sup> استفاده شده است. این پایگاه داده شامل هزار قطعه سیگنال صدا با طول سی ثانیه از ده طبقه موسیقی Rock, Classic, Jazz, Pop و ... است. سیگنال های این پایگاه داده با فرکانس ۲۲۰۵۰ Hz و به صورت مونو و ۱۶ bit و در قالب فرمت Wav، ذخیره شده اند. لازم به ذکر است، این شبیه سازی ها در رایانه ای با پردازنده دو هسته ای با فرکانس کلاک ۲/۵۶ گیگاهرتز و ۲ گیگابایت رم، انجام شده است.

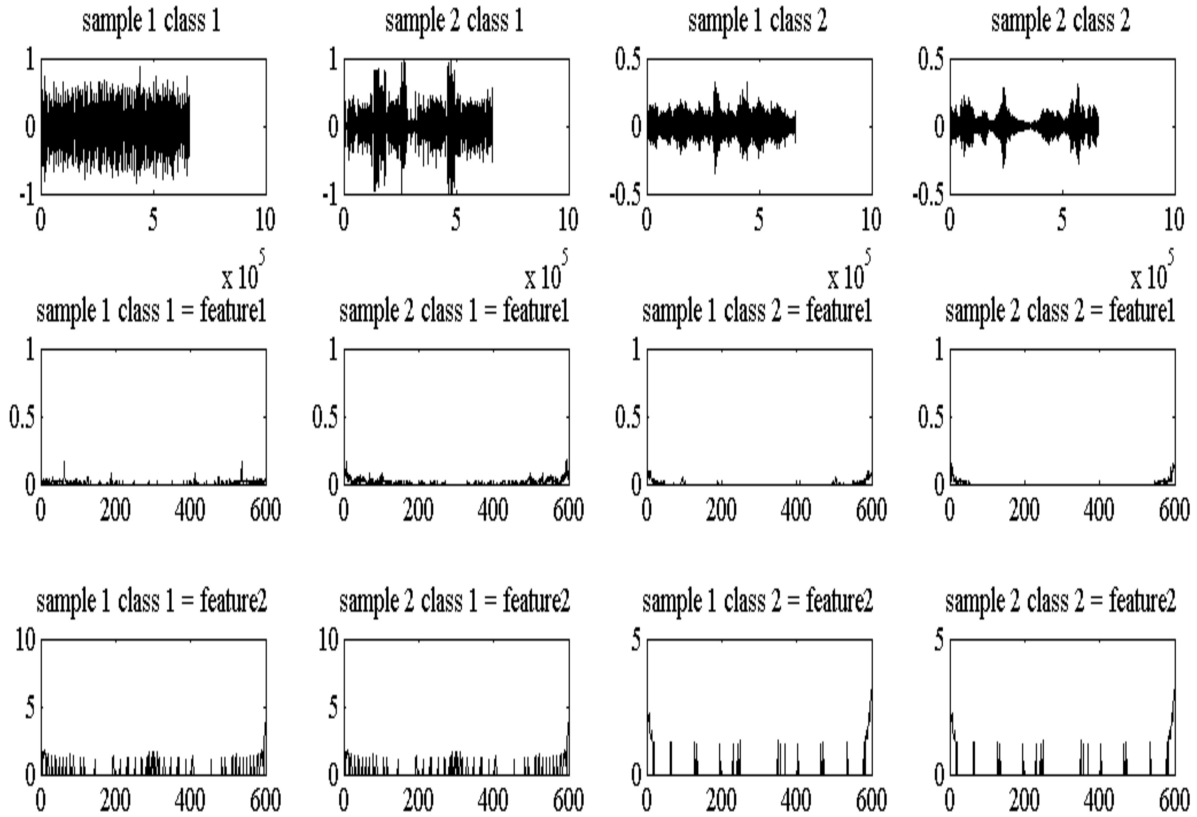
به منظور مشاهده مؤثر بودن روش استخراج ویژگی، بردار ویژگی های استخراج شده از چند سیگنال تصادفی در شکل (۳) نمایش داده شده است. در سطر اول این شکل دو نمونه از سیگنال های دو طبقه متفاوت از داده ها، در سطر دوم ضرایب FFT استخراج شده از هر یک از سیگنال ها و در سطر سوم ویژگی  $CS - 2FFT$  متناظر با آنها نمایش داده شده است، همان گونه که از سطر سوم شکل (۳) مشخص است، تعداد نمونه غیر صفر استخراج شده از سیگنال های مربوط به یک طبقه با تعداد نمونه های غیر صفر استخراج شده از سیگنال های مربوط به سایر طبقه ها تفاوت دارد، اما در سطر اول این شکل که نمونه های زمانی سیگنال آورده شده است، این تفاوت کمتر مشهود است، در نتیجه از این شکل به صورت شهودی می توان دریافت که ویژگی فوق تفاوت بین طبقه ها را به صورتی مناسب برجسته و تفاوت سیگنال های درون هر طبقه را کم می کند.

در مرحله بعدی آزمایش ها، میزان مقاومت ویژگی استخراجی در برابر نوفه مورد بررسی قرار گرفت. برای این منظور یک سیگنال ترکیبی از دو سینوسی با مرحله (فاز) متفاوت در نظر گرفته شده و ویژگی  $CS - 2FFT$  آن استخراج شد. در گام بعدی نوفه سفید گوسی به این سیگنال اضافه شده و دوباره بردار ویژگی استخراج شد. نتیجه این عملیات در شکل (۴) آورده شده است. در ستون سمت چپ این شکل، سیگنال بدون نویز (حاصل ترکیب دو سیگنال سینوسی  $(\sin(t) + \sin(2t + \pi/3))$  و در زیر آن ضرایب FFT (feature1) و در سطر سوم ویژگی  $(CS - 2FFT)$  نمایش داده شده است. در مقابل در ستون سمت راست سیگنال نوفه ای شده و متناظر با ستون سمت چپ، ضرایب FFT (feature1) و ویژگی  $(CS - 2FFT)$ ، آورده شده است. با مقایسه این دو ستون می توان دریافت که

<sup>۱</sup> [http://marsyas.info/download/data\\_sets](http://marsyas.info/download/data_sets)

طبقه‌بندی‌کننده با چند طبقه‌بندی‌کننده رایج مقایسه و نتایج در جدول (۳) ارائه شده است. در ستون‌های این جدول به ترتیب از سمت راست به چپ، عامل زمان حقیقی، تعداد بُعد بردار ویژگی به کار رفته و درصد دقت خطا نمایش داده می‌شود.

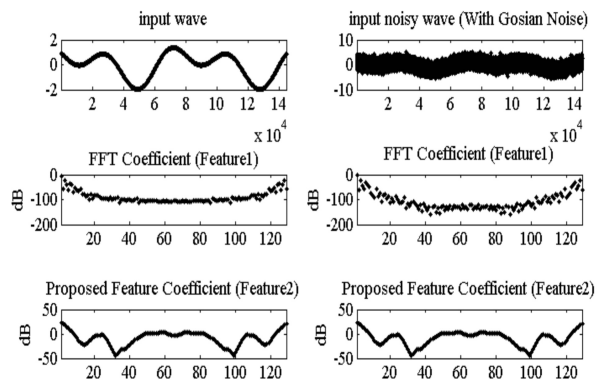
ویژگی پیشنهادی، عمل کرد خیلی مناسبی در برابر نوفه دارد. برای بررسی عمل کرد این طبقه‌بندی‌کننده از هشتاد درصد داده‌های پایگاه داده مورد استفاده به عنوان داده آموزشی و بیست درصد باقیمانده به عنوان داده آزمایشی استفاده شد؛ همچنین انتخاب داده‌های آزمایشی چندین بار به صورت تصادفی تکرار شده است. در نهایت میزان دقت



(شکل ۳) - مقایسه بردار ویژگی (CS\_2FFT) استخراج شده از ۲ طبقه مختلف.

(جدول ۳) - مقایسه نتایج طبقه‌بندی‌کننده‌های مختلف

روش به کار رفته	درصد دقت	تعداد بُعد بردار ویژگی به کار رفته	فاکتور زمان حقیقی RTF
Our Algorithm	۹۵/۷	۳۵	۳/۶
TPNMF <sup>1</sup> +SRC	۹۳/۷	۱۳۵	۷/۱۰
NTF <sup>2</sup> +SRC	۹۲	۱۳۵	۸/۹
MPCA <sup>3</sup> +SRC	۸۹/۷	۲۱۶	۱۰/۸
GTDA <sup>4</sup> +SRC	۹۲/۱	۲۱۶	۹/۶



(شکل ۴) - بررسی میزان مقاومت ویژگی استخراجی در برابر نوفه

<sup>1</sup> Topology Preserving Non-negative Matrix Factorization

<sup>2</sup> Non-negative Tensor Factorization

<sup>3</sup> Multi Linear Principal Component Analysis

<sup>4</sup> General Tensor Discriminant Analysis

عامل زمان حقیقی در واقع معیاری برای مقایسه سامانه‌های شناسایی خودکار گفتار است. این معیار به صورت نسبت  $RTF = \frac{P}{I}$  تعریف می‌شود. در این نسبت  $P$ ، مدت زمان لازم برای پردازش یک سیگنال با طول  $I$  است. با توجه به چگونگی تعریف این رابطه  $RTF$  کمیتی بدون واحد است. نتایج ارائه شده در این جدول نشان می‌دهد که الگوریتم ارائه شده با استفاده از بردار ویژگی با بُعد خیلی کمتر نسبت به سایر طبقه‌بندی‌کننده‌های مبتنی بر طبقه‌بندی‌کننده SRC متداول (ستون دوم از سمت راست)، ضمن داشتن بار محاسباتی کمتر از سایر الگوریتم‌ها، دقت طبقه‌بندی خیلی بالاتری از خود نشان می‌دهد (ستون سوم از سمت راست).

برای توجیه کمتر بودن بار محاسباتی باید بیان کرد "درست است که در روند استخراج بردار ویژگی پیشنهادی برای انجام عملیات هنجار سازی کردن و یا محاسبه تبدیل فوریه مقداری بار محاسباتی به سامانه تحمیل می‌شود، اما در مرحله آخر با انتخاب تصادفی تعداد بسیار اندک نمونه، طول بردار ویژگی مورد استفاده و در نتیجه حجم محاسبات به شدت کاهش می‌یابد."

در مورد تمامی روش‌هایی که برای مقایسه انتخاب شده‌اند در مرجع (Panagikas,2010) و (Chang,2010) اطلاعات کاملی ارائه شده است. نتایج جدول (۳) بیان می‌کند الگوریتمی که از بردار ویژگی ارائه شده به جای سیگنال اولیه در ورودی طبقه‌بندی‌کننده‌های SRC استفاده می‌کند، (در عمل دو مرتبه از تکنیک نمونه برداری فشرده استفاده شده است) سرعت و خطای کمتری در طبقه‌بندی دارد و این به دلیل آن است که بردار ویژگی علاوه بر اینکه طول بسیار کمتری دارد، اطلاعات مورد نیاز را نیز در خود دارد.

## ۲-۴- مکان‌یابی منابع گفتار با استفاده از

### ویژگی مطرح شده $CS-2FFT-L$ :

در اکثر روش‌های مختلف تخمین  $DOA^1$  (Chang,2010) و (Sainath,2010) از طریق روش‌هایی مانند محاسبه  $FFT^2$  و یا چگالی طیف و یا محاسبه ضرایب ویولت و ... از سیگنال دریافتی در میکروفون‌ها، بردار ویژگی

استخراج و بر اساس آن مقدار  $DOA$  تخمین زده می‌شود. در این روش‌ها هدف به دست آوردن دقیق‌ترین و سریع‌ترین تخمین از  $DOA$  سیگنال دریافتی در میکروفون‌هاست. برای این منظور باید ویژگی‌ای استخراج شود که علاوه بر دارا بودن تعداد بُعد کم، اطلاعات مورد نیاز را نیز در بر گیرد تا به واسطه این اطلاعات بتوان مقدار  $DOA$  را تخمین زد.

بعد از شبیه‌سازی‌های صورت گرفته بر روی ویژگی به دست آمده و نشان دادن کارایی آن در طبقه‌بندی اصوات مختلف، در مرحله بعد به دنبال استفاده از ویژگی فوق برای مکان‌یابی منابع مختلف ارسال‌کننده سیگنال هستیم. بدین منظور از ایده ارائه شده در مرجع (Cevher,2008) استفاده شد. تنها تفاوت موجود، مربوط به استفاده از ویژگی  $CS-2FFT$  به منظور تولید بردار ورودی برای الگوریتم مرجع (Cevher,2008) است. در الگوریتم جدید، ابتدا از سیگنال دریافتی در حس‌گرها بردار ویژگی مطرح شده استخراج می‌شود و در مرحله بعد با استفاده از عمل‌گر مناسب تعریف شده (در این عمل‌گر بر خلاف الگوریتم‌های قبل تنها ضریب دامنه تأثیرگذار است و به دلیل خصوصیات ویژگی مطرح شده و چگونگی استخراج آن، شیفت زمانی تأثیری در ویژگی استخراج شده ندارد)، ماتریس منبع حس‌گر<sup>۳</sup> تولید و با استفاده از آن تخمین موقعیت منابع صورت می‌گیرد.

تخمین سیگنال منابع بستگی به موقعیت منابع دارد، فرمول‌هایی که در اینجا ارائه می‌شود به صورت غیر مستقیم موقعیت اهداف را تخمین می‌زنند، بنابراین این بار محاسباتی و پهنای باند مخابراتی کاهش می‌یابد.

فرض کنیم  $K$  منبع در یک فضا به صورت یکنواخت توزیع شده باشند. همچنین فرض کنیم  $P$  حس‌گر با مکان‌های معلوم  $\xi_i = [\xi_{x_i}, \xi_{y_i}]^T$  ( $i = 1, \dots, P$ ) در یک صفحه تخت قرار داشته باشند. (در اینجا فرض بر دانستن تعداد منابع است). هدف، تعیین موقعیت منابع ارسال‌کننده سیگنال به صورت  $x_i = [x_{x_i}, x_{y_i}]^T$  با استفاده از نمونه‌های برداشته شده در حس‌گرهاست. برای این منظور فرض می‌کنیم منابع بتوانند فقط بر روی رئوس شبکه چهارخانه در نظر گرفته شده قرار داشته باشند شکل (۵)، یعنی فقط در نقاط:

<sup>3</sup> Sensor Source Matrix

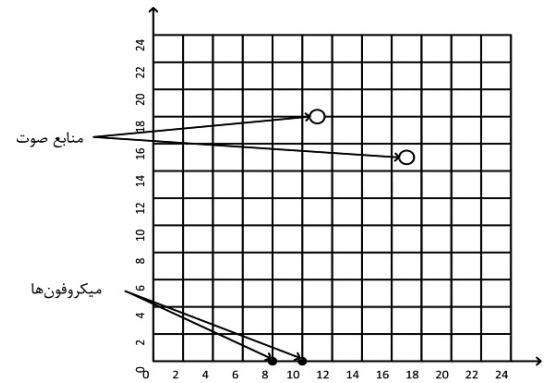
<sup>1</sup> Direction Of Arrival

<sup>2</sup> Fast Fourier Transform



$$\Phi = \{\Phi_n \mid n=1,2,\dots,N, \Phi_n = [\Phi_{x_n}, \Phi_{y_n}]^T\} \quad (9)$$

با استفاده از این گسسته‌سازی و با دانستن محدود بودن تعداد منابع، مسأله مکان‌یابی تبدیل به یک مسأله تخمین پراکندگی از سیگنال دریافتی می‌شود.



شکل ۵- شبکه چهارخانه در نظر گرفته شده

در حس گرها بردار  $\theta \in R^N$  دریافت می‌شود که در برگزیده دامنه سیگنال ارسالی از منابع موجود در  $N$  نقطه شبکه است. به دلیل این که بردار فوق تنها دارای  $K$  درایه غیر صفر است، از این مسأله به عنوان مکان‌یابی به وسیله پراکندگی فضایی یاد می‌شود.

در ادامه برای بررسی مسأله به وجود آمده و به منظور بررسی چگونگی گسترش سیگنال در فضا، عمل گر کانوولوشن خطی به صورت  $I_{x \rightarrow \xi}$  به صورتی تعریف شده که با دریافت سیگنال ارسالی از منبع موجود در مکان  $x$  و با دانستن چگونگی گسترش فضایی سیگنال در مسیرهای مختلف و نیز با در نظر گرفتن اثر مسیرهای مختلف موجود، نمونه از سیگنال دریافتی را توسط حس گری که در موقعیت  $\xi$  قرار دارد، برگرداند. به صورت مشابه عمل گر  $I_{\xi \rightarrow x}$  را به صورتی تعریف می‌کنیم که با دانستن نمونه‌های سیگنال برداشته شده توسط حس گر موجود در مکان  $\xi$ ، سیگنال ارسالی توسط منبع موجود در مکان  $x$  را باز گرداند.

به عنوان یک نمونه ساده برای عمل گر فوق می‌توان نوشت:

$$I_{x \rightarrow \xi}(S) = \frac{1}{d_{x,\xi}^\alpha} S\left(\frac{l}{F_s}\right) \quad (10)$$

در این عبارت  $d_{x,y}$  فاصله منبع  $x$  از حس گر  $\xi$ ،  $S$  سیگنال ارسالی از منبع و  $F_s$  نرخ نمونه برداری برای انتخاب  $L$  نمونه است.

فرض کنیم سیگنال ارسالی از  $k$  آمین منبع  $S_k$  باشد، بنابر این می‌توان سیگنال دریافتی در حس گر  $i$  ام را به صورت  $Z_i = X_i \theta$  نوشت که در آن  $X_i$  را ماتریس منبع حسگر  $i$  ام نامیده شده و به صورت زیر تعریف می‌شود:

$$X_i = [I_{x_1 \rightarrow \xi_i}(S_1), I_{x_2 \rightarrow \xi_i}(S_2), \dots, I_{x_N \rightarrow \xi_i}(S_N)] \quad (11)$$

به صورت مشابه می‌توان با اعمال کردن ماتریس منابع بر روی یک مجموعه پایه مشترک در حالت ماتریسی نوشت  $Z = \Psi \theta$  که در آن داریم:  $\Psi = [X_1^T X_2^T \dots X_L^T]^T$  و  $Z = [Z_1^T Z_2^T \dots Z_L^T]^T$

برای به دست آوردن تخمینی از ماتریس منبع حس گر  $i$  ام ( $X_i$ ) با استفاده از سیگنال دریافتی در حس گر  $i$  ام، با فرض دریافت سیگنال  $Z_i$  در حس گر  $i$  ام می‌توان نوشت:

$$\hat{X}_{i/i} = [I_{x_1 \rightarrow \xi_i}(\hat{l}_{\xi_i \rightarrow x_1}^i(Z_i)), I_{x_2 \rightarrow \xi_i}(\hat{l}_{\xi_i \rightarrow x_2}^i(Z_i)), \dots, I_{x_N \rightarrow \xi_i}(\hat{l}_{\xi_i \rightarrow x_N}^i(Z_i))] \quad (12)$$

در این رابطه  $\hat{X}_{i/i}$ ، تخمینی از ماتریس منبع حس گر  $i$  ام ( $X_i$ ) با استفاده از سیگنال دریافتی در حس گر  $i$  ام، سیگنال دریافتی در حس گر  $i$  ام هستند.

با قراردادن ماتریس منبع تخمین زده شده با رابطه بالا در  $\Psi = [X_1^T X_2^T \dots X_L^T]^T$  می‌توان تخمینی از  $\Psi_i$  به دست آورد. با انتقال سیگنال دریافتی در حس گر  $i$  ام به سایر حس گرها، و با بهره جویی از الگوریتم تخمین تُنکی برای به دست آوردن  $Z$  و  $\hat{\Psi}$  می‌توان تخمینی از موقعیت تُنک منابع ( $\hat{\theta}_i$ ) در حس گر  $i$  ام به دست آورد. نتایج شبیه‌سازی‌ها در ادامه آورده شده است (Cevher,2008). در شکل شماره (۶) نتیجه تخمین دو منبع مولد سیگنال گفتار توسط دو میکروفن نشان داده شده است. به صورت شهودی می‌توان عمل کرد مناسب این الگوریتم مکان‌یابی را مشاهده کرد.

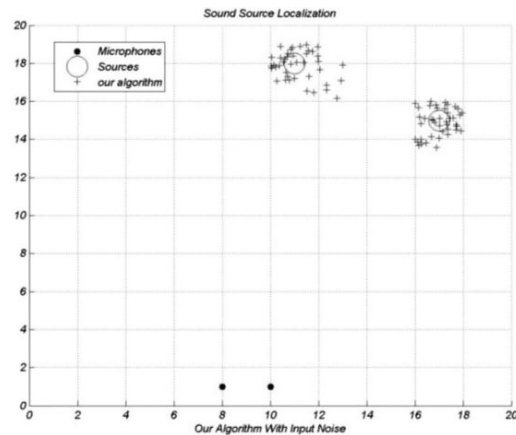
در ادامه شبیه‌سازی‌های صورت گرفته برای بررسی کیفیت عمل کرد الگوریتم مکان‌یابی مبتنی بر روش استخراج ویژگی  $CS-2FFT$  مقایسه‌ای بین نتایج تخمین موقعیت صورت گرفته توسط این الگوریتم و الگوریتم‌های ارائه شده در مراجع (Baraniuk,2008) (DTL) و (B.Dehkordi,2011) (CSNN) در جدول (۴) و در شکل (۷) آورده شده است. در الگوریتم CSNN با استفاده از نمونه برداری فشرده و گسسته‌سازی فضا برای جمع آوری داده آموزشی برای

<sup>1</sup> Sensor's Source Matrix

آموزش شبکه عصبی طراحی شده، عملیات مکان‌یابی انجام می‌شود).

همان‌گونه که از جدول (۴) مشخص است، الگوریتم  $CS - 2FFT - L$ ، عمل کرد خیلی بهتری از دو الگوریتم دیگر از خود نشان داده است. در این جدول خطای تخمین برابر با فاصله بین نقطه واقعی و نقطه تخمین زده شده توسط الگوریتم است.

در ادامه برای بررسی بیشتر عمل کرد الگوریتم، شبیه‌سازی‌هایی برای بررسی تأثیر انعکاس بر روی تخمین صورت گرفته از موقعیت منابع انجام شد. نتایج این شبیه‌سازی‌ها در جدول (۵) ارائه شده است. همان‌گونه که این جدول مشخص می‌کند، انعکاس باعث به وجود آمدن خطا در تخمین موقعیت می‌شود، اما این خطا ناچیز است و می‌توان از آن صرف نظر کرد.



(شکل ۶) - مکان‌یابی دو منبع مولد سیگنال بوسیله دو میکروفون توسط الگوریتم ارائه شده در این بخش (در حضور نوفه گوسی سفید با  $SNR=0dB$  و زمان انعکاس  $120ms$  در سیگنال ورودی به میکروفون‌ها)

(جدول ۴) - خطای تخمین مکان دو منبع مولد سیگنال به وسیله دو میکروفون با الگوریتم ارائه شده ( $CS - 2FFT - L$ ) در این بخش با الگوریتم‌های ارائه شده در مراجع [۱۳] (DTL) و [۱۴] (CSNN) (مقدار خطای مکان‌یابی بر حسب متر)

	SNR=0	SNR=5	SNR=30
$CS - 2FFT - L$	۰/۶۱	۰/۴۵	۰/۲
DTL	۰/۸	۰/۶۸	۰/۵۲
CSNN	۱/۰۳	۰/۷	۰/۶۵

در جدول (۶)، نتایج شبیه‌سازی صورت گرفته مربوط به تأثیر تغییر تعداد منابع بر روی میزان خطای تخمین موقعیت منابع ارائه شده است. همان‌گونه که این نتایج نشان می‌دهد، در مقایسه با سایر الگوریتم‌های متداول، افزایش

تعداد منابع تأثیر بسیار کمتری بر روی عمل کرد الگوریتم ارائه شده نسبت به سایر الگوریتم‌های مورد بررسی دارد.

در جدول (۷) نیز عمل کرد الگوریتم در شرایط متحرک بودن منبع، مورد بررسی قرار گرفته است. همان‌گونه که نتایج ارائه شده در این جدول نشان می‌دهند، خطای تخمین موقعیت در الگوریتم ارائه شده کمتر از سایر الگوریتم‌ها در سرعت‌های مختلف است.

در توجیه نتایج به دست آمده در جدول (۶) و (۷) باید گفت که به دلیل چگونگی الگوریتم استخراج ویژگی پیشنهادی و عدم تأثیر زمان در چگونگی استخراج این بردار ویژگی، تعداد منابع ارسال کننده سیگنال و حرکت آنها تأثیری در دقت مکان‌یابی ندارد.

(جدول ۵) - مقایسه میانگین خطای فواصل نقاط تخمین زده شده از مکان واقعی در  $SNR=5dB$  و زمان‌های انعکاس مختلف (فواصل بر حسب متر)

	T=120ms	T=350ms	T=540ms
میانگین خطا (m)	۰/۳۹	۰/۶۶	۰/۸۷

(جدول ۶) - مقایسه میانگین خطای فواصل نقاط تخمین زده شده از مکان واقعی در  $SNR=5dB$  و تعداد منابع مختلف مختلف (فواصل بر حسب متر)

		۲	۳	۵	۶
خطا (m)	$CS - 2FFT - L$	۰/۲۷	۰/۲۹	۰/۲۸	۰/۳۱
	DTL	۰/۳۹	۰/۵۱	۰/۶۸	۰/۸۰
	CSNN	۰/۴۲	۰/۶۵	۰/۸۷	۰/۹۸

(جدول ۷) - مقایسه خطای نقاط تخمین زده شده از مکان واقعی در  $SNR=5dB$  و منابع متحرک (فواصل بر حسب متر و سرعت بر حسب متر بر ثانیه)

		۰/۵	۰/۷۵	۱
خطا (m)	$CS - 2FFT - L$	۰/۱۵	۰/۲۵	۰/۲۸
	DTL	۰/۳۳	۰/۳۵	۰/۳۸
	CSNN	۰/۴۶	۰/۴۸	۰/۵۱

## ۵- جمع‌بندی و نتیجه‌گیری

در این مقاله ویژگی جدیدی ( $CS - 2FFT$ ) مبتنی بر راه کار نمونه برداری فشرده و تبدیل فوریه برای پردازش سیگنال‌های گفتار ارائه شد. در شبیه‌سازی‌های صورت گرفته نشان داده شد که این روش استخراج ویژگی به دلیل کم بودن تعداد بُعد و حفظ اطلاعات طیفی مورد نیاز در پردازش سیگنال‌ها تأثیر زیادی در سرعت و دقت الگوریتم‌های مختلف در حوزه پردازش سیگنال می‌شود.

Information Retrieval, Toronto, Canada, Jul. Aug. 2003., pp. 282-289.

[4] X. Yap, A. W.H. Khong and W-S. Gan, "Localization of acoustic source on solids: A linear predictive coding based algorithm for location template matching," in Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Dallas, Texas, USA, Mar. 2010, pp. 2490-2493.

[5] M. R. Every, "Discriminating Between Pitched Sources in Music Audio," IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 2, pp. 267-277, Feb. 2008.

[6] T.V.Sreenivas and W. B. Kleijn, "compressive sensing for sparsely excited speech signals," in Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Taipei, Taiwan, Apr. 2009, pp. 4125-4128.

[7] M. Wakin, "compressive sensing," The Connections Project and licensed under the Creative Commons Attribution License, Sep. 2009.

[8] Y. Panagakis, C. Kotropoulos, "music genre fication via topology preserving non-negative -classi tensor factorization and sparse representations," in Proceeding of Acoustics Speech and Signal Processing (ICASSP), 2010, Dallas Texas, March 2010, pp. 249-252.

[9] C-H. Lee, J-L. Shih, K-M. Yu and H-S. Lin, "Automatic Music Genre Classification Based on Modulation Spectral Analysis of Spectral and Cepstral Features," IEEE Transactions on Multimedia, vol. 11, no. 4, pp. 670-682, Jan. 2009.

[10] K. K. Chang, J. R. Jang and C. S. Iliopoulos, "music genre classification via compressive sampling," in Proceeding of International Society for Music Information Retrieval Conference (ISMIR 2010), Utrecht, Netherlands, Aug. 2010, pp. 387-392.

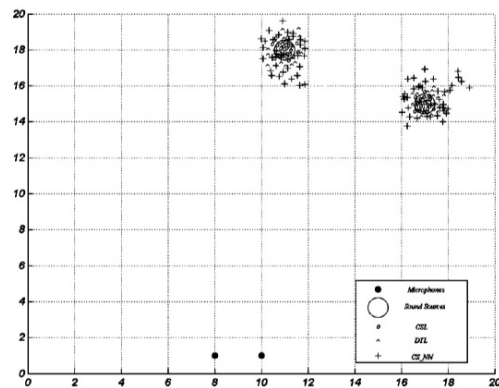
[11] T. Sainath, A. Carmi, D. Kanevsky and B. Ramabhadran, "Bayesian compressive sensing for phonetic classification," in Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Dallas, Texas, USA, Mar. 2010, pp. 4370-4373.

[12] V. Cevher, A. C. Gurbuz, J. H. McClellan and R. Chellappa, "Compressive wireless arrays for bearing estimation," in Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, Nevada, USA, Apr. 2008, pp. 2497-250.

[13] V. Cevher, M. F. Duarte and R. G. Baraniuk, "Distributed target localization via spatial sparsity," in Proceeding of European Signal Processing Conference (EUSIPCO), Lausanne, Switzerland, Aug. 2008, pp. 231-237.

در این مقاله دو کاربرد از ویژگی  $CS-2FFT$  مورد بررسی قرار گرفت. ابتدا کاربرد در طبقه‌بندی سیگنال‌ها مطرح شد و مورد آزمایش قرار گرفت. نتایج شبیه‌سازی‌ها نشان داد که با وجود مقاومت بسیار بالای الگوریتم طبقه‌بندی کننده (به دلیل نوع ویژگی استفاده شده در الگوریتم) دقت بسیار بالایی (۹۵/۷) در طبقه‌بندی طبقه‌های مختلف موسیقی دارد.

کاربرد دیگری که در این مقاله مورد توجه قرار گرفت تعیین موقعیت منابع مولد سیگنال گفتار با استفاده از ویژگی  $CS-2FFT$  بود. در این کاربرد از الگوریتم مطرح شده در مرجع (Baraniuk, 2008) و ویژگی مورد نظر به صورت هم‌زمان استفاده شد. نتایج شبیه‌سازی‌ها در این قسمت نیز قدرت بالای الگوریتم (خطای کمتر از دو درصد) در تعیین موقعیت منابع را نشان داد.



(شکل ۷) - مقایسه نتایج به دست آمده برای مکان‌یابی دو منبع مولد سیگنال بوسیله دو میکروفون توسط الگوریتم ارائه شده (CSL) در این بخش با الگوریتم‌های ارائه شده در مراجع (Baraniuk, 2008) (DTL) و (B. Dehkordi, 2011) (CSNN) (در حضور نوفه گوسی سفید با  $SNR=0dB$  و زمان انعکاس  $120ms$  در سیگنال ورودی به میکروفون‌ها)

## مراجع

[1] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," IEEE Transactions on Speech and Audio Processing, vol. 10, no. 5, pp. 293-302, Jul. 2002.

[2] L. Lu, H-J. Zhang, J-H. Tao and L-H. Cai, "Music type classification by spectral contrast feature," in Proceeding of IEEE International Conference on Multimedia & Expo (ICME), Lusanne, Switzerland, Aug. 2002, pp. 113-116.

[3] T. Li, M. Ogihara, and Q. Li, "A comparative study on content based music genre classification," in Proceeding of Annual International ACM SIGIR Conference on Research & Development on



[14] M. B. Dehkordi, H. R. Abutalebi, and H. Ghanei, "A Compressive Sensing Based Compressed Neural Network for Sound Source Localization," in Proceeding of Artificial Intelligence and Signal Processing (AISP), Tehran, Iran, Jun. 2011, pp. 6-10.



#### مهدي بنی‌طالبی دهکردی دوره

کارشناسی خود را در سال ۱۳۸۶ در رشته مهندسی برق (الکترونیک) در دانشگاه شهپرکرد گذرانده و مدرک کارشناسی ارشد خود را در سال ۱۳۹۰ در رشته مهندسی برق (مخابرات) از دانشگاه یزد اخذ نموده است. وی در سال ۱۳۹۱ به دانشگاه فنی و حرفه‌ای واحد پسران شهرکرد پیوست و هم اکنون در این آموزشکده مشغول به فعالیت است. زمینه‌های علمی مورد علاقه وی پردازش آرایه‌ای سیگنال‌های گفتار، بهسازی گفتار، مکان‌یابی منابع صوت، پردازش تصویر است. نشانی رایانامه ایشان عبارت است از:

mahdi\_banitalebi@stu.yazduni.ac.ir



#### حمیدرضا ابوطالبی دوره کارشناسی و

کارشناسی ارشد خود را به ترتیب در سال‌های ۱۳۷۴ و ۱۳۷۷ در رشته مهندسی برق (مخابرات) در دانشگاه صنعتی شریف گذرانده و مدرک دکترای خود را در سال ۱۳۸۲ در رشته مهندسی برق (مخابرات) از دانشگاه صنعتی امیرکبیر اخذ نمود. وی در جریان رساله دکترای خویش، به مدت یک سال در دوره فرصت مطالعاتی در دانشگاه واترلو کانادا به سر برد. دکتر ابوطالبی در سال ۱۳۸۲ به دانشکده مهندسی برق و کامپیوتر دانشگاه یزد پیوست و هم اکنون به عنوان دانشیار این دانشکده مشغول به فعالیت می‌باشد. وی همچنین در سال ۹۰-۱۳۸۹ یک دوره فرصت مطالعاتی را در مرکز تحقیقاتی IDIAP در سوئیس سپری نمود. زمینه‌های علمی مورد علاقه وی پردازش آرایه‌ای سیگنال گفتار، بهسازی گفتار، مکان‌یابی گوینده، فیلترهای وقتی، و آنالیز زمان-فرکانس می‌باشد.

نشانی رایانامه ایشان عبارت است از:

habutalebi@yazduni.ac.ir



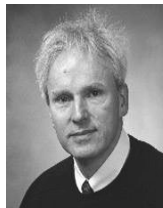
#### محمدتقی صادقی تحصیلات خود را در

مقاطع کارشناسی برق-الکترونیک و کارشناسی ارشد برق-مخابرات به ترتیب در سال‌های ۱۳۷۰ و ۱۳۷۳ در دانشگاه‌های

سال ۱۳۹۲ شماره ۱ پیاپی ۱۹

صنعتی شریف و تربیت مدرس تهران به پایان رسانید. نامبرده از سال ۱۳۷۴ الی ۱۳۷۸ در دانشکده مهندسی برق دانشگاه یزد مشغول به کار بود. پس از آن دوره دکتری مهندسی برق - مخابرات را در دانشگاه ساری انگلستان آغاز نموده و در سال ۱۳۸۱ موفق به اخذ درجه دکتری از آن دانشگاه شد. سپس به مدت دو سال در مرکز تحقیقات بینایی ماشین و پردازش سیگنال همان دانشگاه، در زمینه سیستم‌های بازشناسی چهره پژوهش نموده و از آن پس تاکنون استادیار دانشکده مهندسی برق و کامپیوتر دانشگاه یزد است. همکاری پژوهشی نامبرده با مرکز تحقیقات بینایی ماشین و پردازش سیگنال دانشگاه ساری همچنان ادامه دارد. زمینه‌های تحقیقاتی مورد علاقه نامبرده عبارتند از: بازشناسی آماری الگو، پردازش تصویر و بینایی ماشین. نشانی رایانامه ایشان عبارت است از:

m.sadeghi@yazduni.ac.ir



#### جوزف کیتلر در سال ۱۹۷۱ میلادی

تحصیلات خود را در مقاطع کارشناسی مهندسی برق در دانشگاه کمبریج به پایان رسانید. پس از آن در سال ۱۹۷۴ موفق به اخذ درجه دکتری از همان دانشگاه در زمینه بازشناسی الگو شد. نامبرده از سال ۱۹۷۳ تا ۱۹۸۶ با مراکز مختلفی همچون دانشکده مهندسی دانشگاه کمبریج، دانشگاه اکسفورد و آزمایشگاه روترفورد اپلتون IBM، دانشگاه اکسفورد و آزمایشگاه روترفورد اپلتون (RAL) همکاری پژوهشی نمود. نامبرده از سال ۱۹۸۶ به دانشگاه ساری انگلستان ملحق شده و سال‌های متمادی به‌عنوان رئیس مرکز تحقیقات بینایی ماشین و پردازش سیگنال آن دانشگاه فعالیت نموده است. همچنین در سال‌های ۱۹۹۱ و ۲۰۰۴ به ترتیب به مراتب علمی استاد تمام و استاد برگزیده در حوزه هوش ماشینی دست یافته است. زمینه‌های تحقیقاتی نامبرده جنبه‌های مختلف نظری و کاربردهای متنوع سامانه‌های تجزیه و تحلیل تصاویر و بینایی ماشین می‌باشد.

نشانی رایانامه ایشان عبارت است از:

j.kittler@surrey.ac.uk

فصلنامه  
دو ساله

شماره ۱  
پیاپی ۱۹  
پروداز و دوره

۶۸