

## استفاده از عکس STFT اصلاح شده به منظور بهبود کیفیت جداسازی منابع به روش کور

محمد حسین کهایی  
دانشیار دانشکده برق  
دانشگاه علم و صنعت ایران  
kahaei@iust.ac.ir

جعفر محمدی  
دانشجوی کارشناسی مخابرات  
دانشگاه علم و صنعت ایران  
aco.mohammadi@gmail.com

مهدی مهدی خانی  
دانشجوی کارشناسی ارشد مخابرات  
دانشگاه علم و صنعت ایران  
m\_mehdikhani2000@yahoo.com

که  $x_j(\tau, f)$  و  $s_k(\tau, f)$  به ترتیب تبدیل زمان-فرکانس سیگنال منابع به روش کور معرفی می‌شود. روش پیشنهادی مبتنی بر اصلاح بلوک عکس تبدیل فوری پنجره‌ای است.

با این روش جداسازی سیگنال‌ها با سرعت و کیفیت بهتری نسبت به روش‌های معمول انجام می‌شود. برای نمایش تاثیر روش پیشنهادی از الگوریتم DUET به عنوان الگوریتم مرجع استفاده می‌نماییم. **واژه‌های کلیدی:** جداسازی سیگنال منابع به روش کور، نمایش زمان-فرکانس، عکس تبدیل فوری پنجره‌ای.

در تمامی روشهای زمان-فرکانس مانند تجزیه و تحلیل اجزای مستقل در حوزه فرکانس<sup>۵</sup> [6] یا خوشه‌بندی<sup>۶</sup> [7] آنچه که به عنوان نقطه مشترک بلوک‌های تبدیل سیگنال از زمان به زمان-فرکانس و بالعکس از زمان-فرکانس به زمان است. تبدیل‌های مختلفی برای نمایش زمان-فرکانس سیگنال وجود دارد که به دو دسته خطی و غیر خطی تقسیم می‌شوند. بهترین تبدیل خطی‌ای که در جداسازی منابع کاربرد دارد تبدیل STFT<sup>۷</sup> است بخصوص اگر مدل مخلوط سیگنال‌ها پیچشی باشد [4].

بطور خلاصه در این الگوریتم‌ها ابتدا سیگنال سنسورها با استفاده از تبدیل خطی STFT به حوزه زمان-فرکانس برده می‌شوند. سپس سیگنال منابع از یکدیگر جدا می‌شود و در نهایت با بکاربردن عکس تبدیل STFT سیگنال‌های جدا شده به حوزه زمان باز می‌گردند. این مراحل به صورت شماتیک در شکل (۱) نمایش داده شده است.

$$x(t) \xrightarrow{\text{STFT}} X(\tau, f) \xrightarrow{\text{الگوریتم}} S(\tau, f) \xrightarrow{\text{ISTFT}} s(t)$$

شکل (۱): مراحل جداسازی سیگنال در حوزه زمان-فرکانس

در محاسبه STFT بطور معمول از پنجره‌های میرا شونده از طرفین استفاده می‌شود. وجود چنین پنجره‌هایی ضروری است زیرا سیگنال صحبت از لحاظ آماری غیرایستاد است اما در بازه‌های کوتاه می‌توان از فرض ایستاد بودن استفاده کرد. با ضرب سیگنال در پنجره‌های با طول کوتاه چنین بازه‌هایی ایجاد می‌گردد. اما از طرف دیگر برای حذف اثر پنجره در مرحله محاسبه ISTFT سیگنال بدست آمده بر مقادیر پنجره تقسیم می‌شود. در این جاست که یکی از مهمترین نقاط ضعف ISTFT خود را نشان می‌دهد. مقادیر دو انتهای پنجره نزدیک به صفر می‌باشد در نتیجه با تقسیم سیگنال بر چنین اعدادی پیک‌های ناخواسته‌ای ایجاد می‌گردد که به صورت نویز پس زمینه در سیگنال جدا شده بخوبی قابل شنیدن است. برای حل این مشکل ما از

**چکیده:** در این مقاله، روشی برای بهبود کیفیت جداسازی سیگنال منابع به روش کور معرفی می‌شود. روش پیشنهادی مبتنی بر اصلاح بلوک عکس تبدیل فوری پنجره‌ای است.

با این روش جداسازی سیگنال‌ها با سرعت و کیفیت بهتری نسبت به روش‌های معمول انجام می‌شود. برای نمایش تاثیر روش پیشنهادی از الگوریتم DUET به عنوان الگوریتم مرجع استفاده می‌نماییم.

**واژه‌های کلیدی:** جداسازی سیگنال منابع به روش کور، نمایش زمان-فرکانس، عکس تبدیل فوری پنجره‌ای.

### ۱- مقدمه

در این مقاله ما بحث جداسازی کور منابع<sup>۱</sup> را برای سیگنال‌های گفتار<sup>۲</sup> مورد بررسی قرار می‌دهیم. در جداسازی کور منابع، هدف جداکردن سیگنال منابع از مخلوط آنهاست. به بیان دیگر فرض می‌کنیم سیگنال  $N$  منبع مختلف توسط  $M$  سنسور گیرنده دریافت شده باشد. سیگنال هر سنسور مخلوطی از سیگنال منابع است که نحوه ترکیب آنها در حالت کلی به صورت پیچشی<sup>۳</sup> می‌باشد. به زبان ریاضی خواهیم داشت:

$$x_j(t) = \sum_{k=1}^N \sum_{l=1}^M h_{jk}(l) s_k(t-l) \quad j=1, \dots, M \quad (1)$$

که  $x_j(t)$  سیگنال سنسور  $j$ ام،  $s_k(t)$  سیگنال منبع  $k$ ام،  $h_{jk}(t)$  پاسخ ضربه<sup>۴</sup> از منبع  $k$ ام به سنسور  $j$ ام است.

روشهای متنوعی برای حل مساله جداسازی منابع پیشنهاد شده است. در یک دسته‌بندی عمومی این روش‌ها در حوزه‌های زمان [1]، فرکانس [2,3] یا زمان-فرکانس [4,5] تعریف می‌شوند. در حالتی که ترکیب سیگنالها به صورت پیچشی فرض می‌شود، روش‌های مبتنی بر فرکانس یا زمان-فرکانس مناسب‌تر از روش‌های زمانی هستند، زیرا در این صورت مدل ترکیب سیگنال‌ها از پیچشی به خطی تغییر می‌کند. در حقیقت با توجه به رابطه زیر در حوزه فرکانس داریم:

$$x_j(\tau, f) = \sum_{k=1}^N h_{jk}(f) s_k(\tau, f) \quad j=1, \dots, M \quad (2)$$

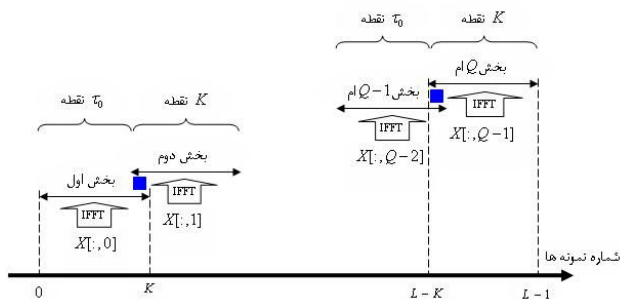
با توجه به شکل (۳) برای محاسبه تبدیل STFT، ابتدا سیگنال زمانی  $x[n]$  به  $N$  بخش  $K$  نمونه‌ای تقسیم می‌شود و سپس با استفاده از K-point DFT تبدیل فوریه هر بخش بدست می‌آید. هر بخش از ضرب سیگنال در پنجره بدست می‌آید.

چون  $KN > L$  می‌باشد لذا بخش‌های سیگنال با یکدیگر هم‌پوشانی خواهند داشت. در شکل (۳) نواحی دارای هم‌پوشانی با مربع رنگی مشخص شده‌اند. همان‌طور که با استفاده از تبدیل STFT سیگنال از حوزه زمان به حوزه زمان - فرکانس تصویر می‌شود، با اعمال تبدیل ISTFT دوباره به حوزه زمان بر می‌گردد. شکل ریاضی این تبدیل به صورت رابطه زیر است:

$$x(kT + q\tau_0) = \frac{1}{w(kT)} \sum_{m=0}^{M-1} X[m, n] e^{j2\pi(mf_0)(kT)} \quad (۴)$$

$$0 \leq k \leq K-1, \quad 0 \leq n \leq N-1$$

در عمل برای پیاده‌سازی رابطه (۴) از K-point IDFT استفاده می‌شود. نمایش تبدیل ISTFT مانند شکل (۴) است.



شکل (۴): تبدیل ISTFT

همان‌طوری که در شکل (۴) مشاهده می‌شود، با اعمال تبدیل ISTFT نمونه‌های سیگنال اصلی بدست می‌آیند اما با هم‌پوشانی زمانی و در  $N$  بخش مجزا. در نتیجه مساله اصلی چگونگی ترکیب این بخش‌ها با یکدیگر و ساختن سیگنال اولیه است. برای رسیدن به این هدف، می‌توان از میانگین‌گیری نقاط دارای هم‌پوشانی استفاده کرد. به عنوان مثال بردارهای  $X[:, 0]$  و  $X[:, 1]$  را در نظر می‌گیریم. با اعمال عکس تبدیل فوریه، نمونه‌های زیر تولید می‌شود:

$$X[:, 0] \Rightarrow x(0:K-1)w(0:K-1)$$

$$X[:, 1] \Rightarrow x(\tau_0:\tau_0+K-1)w(0:K-1)$$

که با تقسیم روابط فوق بر  $w(0:K-1)$  خواهیم داشت:

$$X[:, 0] \Rightarrow x(0:K-1)$$

$$X[:, 1] \Rightarrow x(\tau_0:\tau_0+K-1) \quad (۵)$$

با توجه به رابطه (۵) مشاهده می‌شود که نمونه‌های از شماره  $\tau_0$  تا  $K-1$  دو بار محاسبه می‌گردند. در نتیجه مقدار نهایی، مطابق رابطه (۶) میانگین این دو مقدار خواهد بود:

هم‌پوشانی  $\wedge$  پنجره‌ها به نحوی هوشمندانه استفاده کرده‌ایم که نتایج کار با الگوریتم DUET بررسی و مقایسه شده است. در بخش دوم شکل پیک‌های ناخواسته ISTFT را با بیان ریاضی بطور مشروح توضیح می‌دهیم. در بخش سوم به بیان ISTFT اصلاح شده می‌پردازیم و در نهایت در بخش آخر با ارائه نتایج شبیه‌سازی اثر مثبت روش پیشنهادی را نشان می‌دهیم.

## ۲- بررسی مشکل پیک‌های ناخواسته در ISTFT

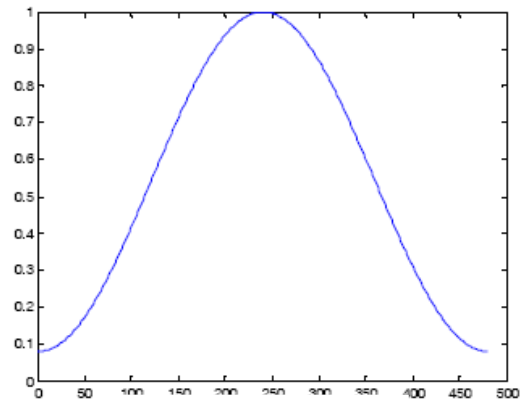
در ابتدا ضروری است مرور مختصری پیرامون تبدیل STFT و عکس آن داشته باشیم. تبدیل STFT سیگنال  $x[n]$  در حالت گسسته به صورت رابطه زیر تعریف می‌شود:

$$X(m, n) = X(mf_0, n\tau_0)$$

$$= \sum_{k=0}^{K-1} x(kT + n\tau_0)w(kT) e^{-j2\pi(mf_0)(kT)} \quad (۳)$$

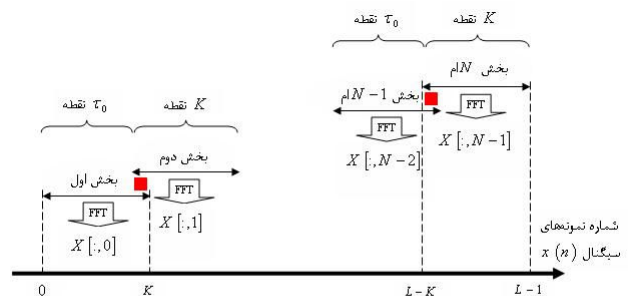
$$0 \leq m \leq M-1, \quad 0 \leq n \leq N-1$$

در رابطه (۳)  $f_0$ ،  $\tau_0$  و  $T$  به ترتیب بیانگر متغیرهای فرکانس، تاخیر و زمان در حالت گسسته است و  $m$ ،  $n$  و  $k$  شماره آنهاست.  $w(t)$  تابع پنجره است که همین‌نگ در نظر گرفته می‌شود. نمونه‌ای از این نوع پنجره با طول ۴۸۰ نمونه در شکل (۲) مشاهده می‌شود.

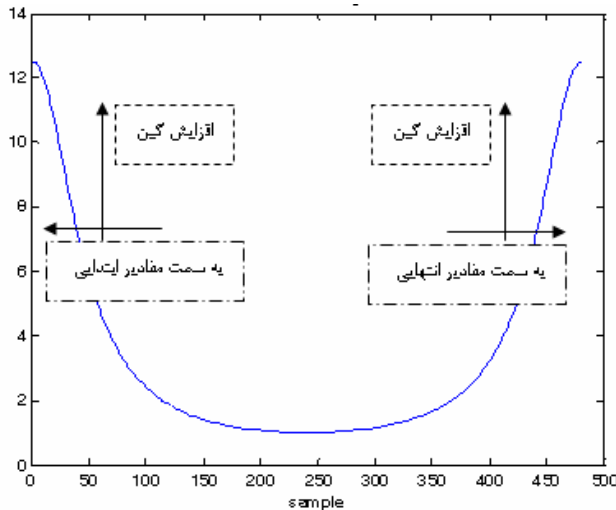


شکل (۲): پنجره همینگ

به صورت شماتیک تبدیل STFT به صورت شکل (۳) عمل می‌کند.



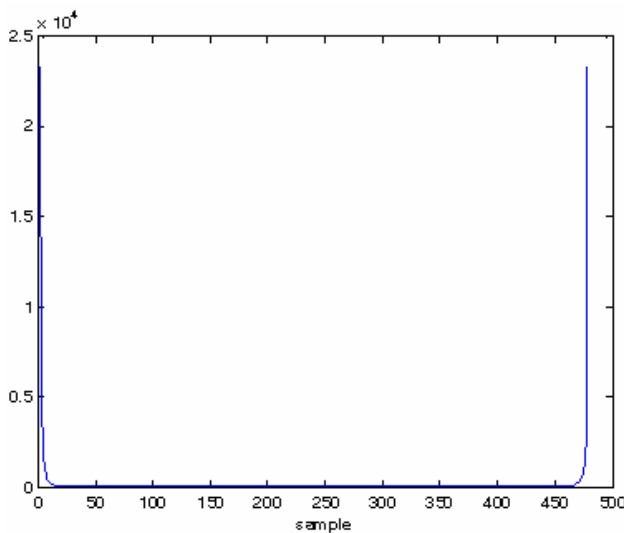
شکل (۳): تبدیل STFT



شکل (۵): عکس تابع همینگ

$$g(n) = \frac{1}{0.5 \left( 1 - \cos\left(2\pi \frac{n}{K-1}\right) \right)} \quad 0 \leq n \leq K-1$$

را نیز مطابق شکل (۶) رسم می‌کنیم.



شکل (۶): عکس تابع همینگ

مشاهده می‌گردد که با نزدیک شدن به ابتدا یا انتهای پنجره، میزان گین تقویت خطا، بسیار بزرگ می‌شود. بنابراین در این نواحی سیگنال اصلی کاملاً در نویز غرق شده و قابل شناسایی نمی‌باشد.

در عمل و در جداسازی سیگنال منابع صوتی این پدیده به صورت پیک‌های ناخواسته در ابتدا و انتهای هر بخش مشاهده می‌گردد. وجود چنین پیک‌هایی سبب افت شدید کیفیت سیگنال‌های جدا شده می‌شود و به صورت نویز شنیده می‌گردد. در بخش شبیه‌سازی این اثر را نشان خواهیم داد.

$$x(0:K-1)=A \quad \& \quad x(\tau_0:\tau_0+K-1)=B$$

$$\Rightarrow x(\tau_0:K-1) = \frac{A(\tau_0:K-1)+B(\tau_0:K-1)}{2} \quad (۶)$$

به همین ترتیب در تمام نمونه‌هایی که دو یا چند بار محاسبه می‌گردند، مقدار نهایی میانگین آنها خواهد بود. این وضعیت در ابتدا و انتهای تمام بخش‌ها وجود دارد به جز ابتدای بخش اول و انتهای بخش آخر که نمونه‌ها یک بار محاسبه می‌شوند و دارای هم‌پوشانی نیستند. این روش دارای یک اشکال عمده است که در ادامه به توضیح آن می‌پردازیم. فرض می‌کنیم در مرحله عکس تبدیل فوریه مقدار کوچکی خطا ایجاد شود. در این صورت رابطه زیر برقرار خواهد بود:

$$X[:,0] \Rightarrow x(0:K-1)w(0:K-1)+e_1(0:K-1) \quad (۷)$$

$$X[:,1] \Rightarrow x(\tau_0:\tau_0+K-1)w(0:K-1)+e_2(0:K-1)$$

$e_1$  و  $e_2$  خطای ایجاد شده است که می‌تواند ناشی از نویز محیط و یا خطای سیستم باشد. حال با اعمال مرحله تقسیم بر مقادیر پنجره، به سیگنال زیر می‌رسیم:

$$x(0:K-1)+\hat{e}_1(0:K-1) \quad (۸)$$

$$x(\tau_0:\tau_0+K-1)+\hat{e}_2(0:K-1)$$

که  $\hat{e}_1$  و  $\hat{e}_2$  عبارتند از:

$$\hat{e}_h = \frac{e_h}{w} \quad (۹)$$

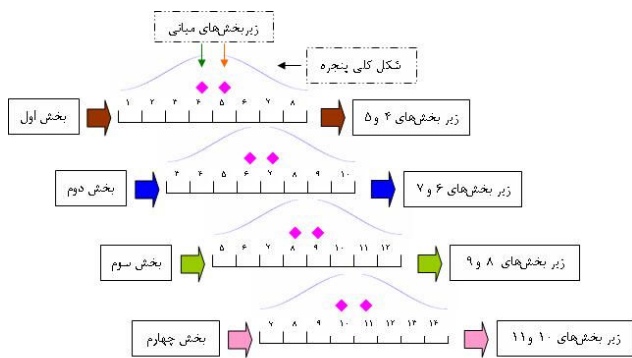
بنابراین نمونه‌های سیگنال اصلی با خطا همراه خواهند بود. از آنجایی که مقدار پنجره  $w[n]$  در ابتدا و انتهای آن کوچک است و برای برخی پنجره‌ها مانند پنجره هنینگ و بلکمن صفر می‌باشد، لذا نمونه‌های ابتدایی و انتهایی  $\hat{e}_h$  که با عکس مقادیر پنجره رابطه دارد، بزرگ خواهد شد. در شکل (۵) عکس پنجره همینگ، یعنی تابع

$$g(n) = \frac{1}{0.54 - 0.46 \cos\left(2\pi \frac{n}{K-1}\right)} \quad 0 \leq n \leq K-1$$

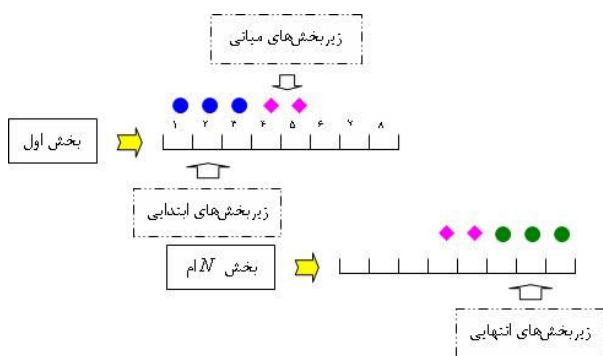
رسم شده است.

با توجه به شکل (۵) هرچه به ابتدا یا انتهای تابع  $\hat{w}$  نزدیک می‌شویم، مقدار تابع بزرگ‌تر می‌شود و در نتیجه خطای  $\hat{e}_h$  بیشتر تقویت می‌گردد. با تقویت خطا، کیفیت سیگنال بازسازی شده  $x[n]$  کاهش می‌یابد یا به عبارت دیگر نمونه‌های سیگنال اصلی بیشتر در نویز غرق (محو) می‌شوند.

پدیده محو شدن سیگنال اصلی به خصوص در نمونه‌های ابتدایی و انتهای هر بخش دیده می‌شود. برای درک بهتر این پدیده عکس پنجره هنینگ یعنی تابع



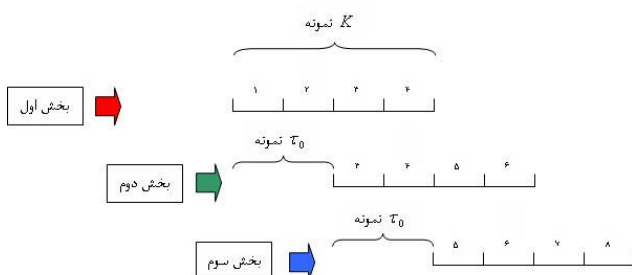
شکل (۸): انتخاب زیربخش‌های مناسب



شکل (۹): انتخاب زیربخش‌های ابتدایی و انتهایی

### ۲-۳ مثال اول

در این حالت هم‌پوشانی نمونه‌ها برابر ۵۰ درصد است. بعد از اعمال عکس تبدیل فوریه، هر بخش را به ۴ قسمت تقسیم می‌کنیم. زیربخش‌های سیگنال جدا شده به صورت شکل (۱۰) با یکدیگر هم‌پوشانی دارند. (در هر بخش ۵۰ درصد بخش قبلی تکرار می‌شود).



شکل (۱۰): ۵۰ درصد هم‌پوشانی بین بخش‌های سیگنال

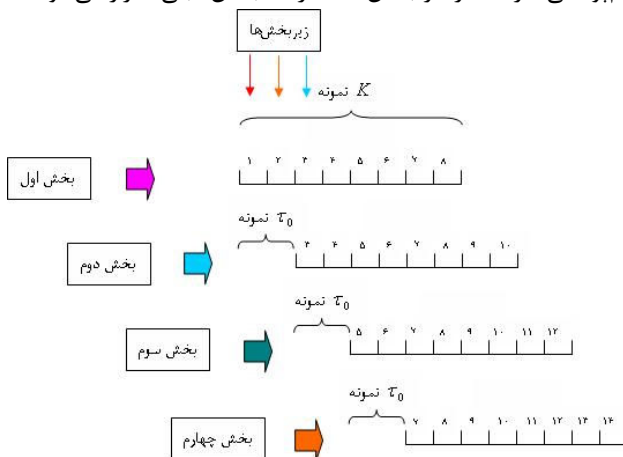
با توجه به شرایط انتخاب زیربخش‌ها، زیربخش‌های شکل (۱۱) انتخاب می‌شوند. مانند حالت قبل، در تمام بخش‌ها دو زیربخش وسطی انتخاب می‌شوند به جز اولین و آخرین بخش که مانند شکل (۱۲) زیربخش‌های ابتدایی و انتهایی هم انتخاب می‌گردند.

### ۳- پیشنهاد ISTFT اصلاح شده جهت حذف پیک‌های ناخواسته

برای حذف پیک‌های ناخواسته از قسمت‌هایی که دارای هم‌پوشانی هستند صرف‌نظر می‌گردد. در این روش هر نمونه تنها یک بار محاسبه می‌شود، بنابراین نیازی به میانگین‌گیری نیست. همچنین سعی می‌شود مقادیر نمونه‌ها از قسمت‌هایی انتخاب شود که مقدار پنجره متناظر با آنها بزرگ باشد. در نتیجه اشکال گذشته به میزان قابل توجهی کاهش می‌یابد. برای توضیح بیشتر به دو مثال توجه می‌کنیم:

#### ۱-۳ مثال اول

در این مثال فرض می‌کنیم هم‌پوشانی نمونه‌ها برابر ۷۵ درصد باشد. بعد از اعمال عکس تبدیل فوریه، هر بخش را به ۸ قسمت تقسیم می‌کنیم. زیربخش‌های سیگنال جدا شده به صورت شکل (۷) با یکدیگر هم‌پوشانی دارند. (در هر بخش، ۷۵ درصد بخش قبلی تکرار می‌شود).



شکل (۷): ۷۵ درصد هم‌پوشانی بین بخش‌های سیگنال

برای کاهش اثر منفی گین ایجاد شده در لبه‌های پنجره، زیربخش‌هایی انتخاب می‌شوند که دارای شرایط زیر هستند:

- (الف) اشتراکی نداشته باشند. (دارای هم‌پوشانی نیستند)
- (ب) اجتماع آنها تمام سیگنال اصلی را بسازد. ( $L$  نقطه)
- (ج) مقدار پنجره در آن نقاط بیشترین باشد.

با توجه به شکل (۷) زیربخش‌های وسطی بهترین گزینه هستند که در شکل (۸) نمایش داده شده‌اند.

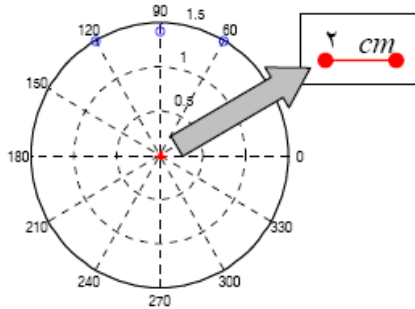
همانطوری که مشاهده می‌شود، زیربخش‌های مشخص شده اولاً با یکدیگر اشتراکی ندارند، ثانياً مجموع آنها تمام زیربخش‌های سیگنال اصلی را تشکیل می‌دهد و ثالثاً مقدار پنجره در آنها به طور نسبی از زیربخش‌های دیگر بالاتر است.

در تمام بخش‌ها، دو زیربخش وسطی انتخاب می‌شوند به جز اولین و آخرین بخش که مطابق شکل (۹) زیربخش‌های ابتدایی و انتهایی هم انتخاب می‌شوند. به این ترتیب مجموع تمام زیربخش‌ها، تمام سیگنال اصلی را می‌سازد.

زمان- فرکانس انجام می‌دهد. در این الگوریتم با محاسبه برخی ویژگی‌های منحصر بفرد هر منبع و دسته‌بندی آنها، جداسازی صورت می‌گیرد. این ویژگی‌ها شامل تاخیرها و تضعیف‌های نسبی است که در مورد سیگنال هر منبع نسبت به سنسورهای گیرنده در نظر گرفته می‌شود. بعد از مشخص شدن پارامترهای هر منبع، مرحله دسته بندی انجام می‌گیرد به عبارت دیگر در تمام نقاط زمان - فرکانس، نقاطی که دارای ویژگی‌های یکسانی هستند با یک عدد برچسب‌گذاری<sup>۱</sup> می‌شوند در نهایت با در نظر گرفتن تمام نقاطی که یک برچسب دارند، تبدیل زمان - فرکانس هر منبع مشخص می‌شود. با اعمال بلوک عکس STFT سیگنال‌های زمانی ایجاد می‌گردند [4].

#### ۵- نتایج شبیه‌سازی

در این بخش فرض می‌کنیم ۳ نفر به صورت همزمان با هم صحبت می‌کنند و سیگنال آنها توسط ۲ سنسور در یافت می‌گردد. موقعیت منابع صوتی (انسان) و سنسورها مطابق شکل (۱۴) است.

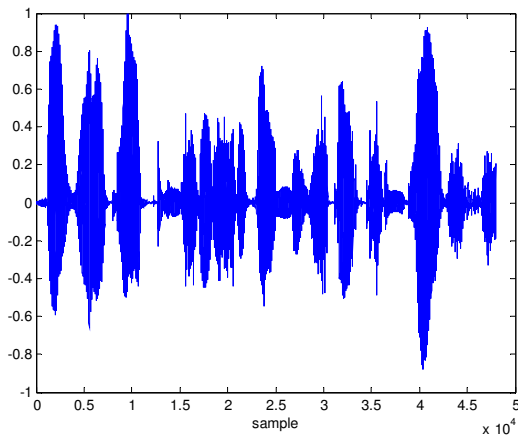


شکل (۱۴): موقعیت منابع و سنسورها

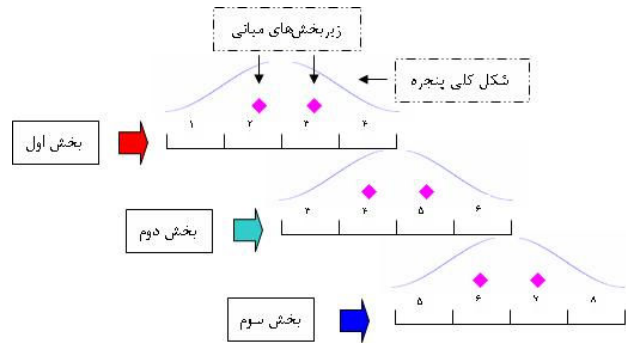
با استفاده از الگوریتم DUET سیگنال منابع را جدا می‌کنیم. در این مثال مقدار SNR را برابر 40dB فرض می‌کنیم. در اشکال زیر سیگنال اصلی هر منبع، سیگنال جدا شده بدون استفاده از روش پیشنهادی و سیگنال جدا شده با استفاده از روش پیشنهادی رسم شده‌اند.

#### ۵-۱ منبع اول

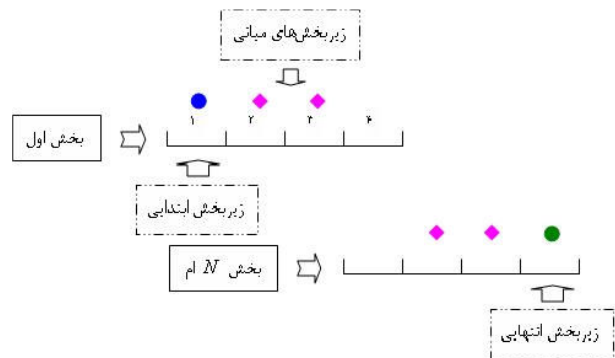
سیگنال اصلی منبع اول مطابق شکل زیر است.



شکل (۱۵): سیگنال اصلی منبع اول

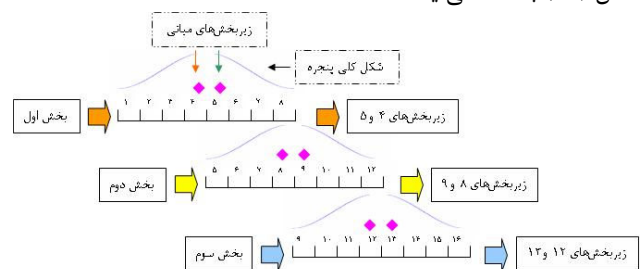


شکل (۱۱): انتخاب زیربخش‌های مناسب



شکل (۱۲): انتخاب زیربخش‌های ابتدایی و انتهایی

در این حالت نمی‌توانیم مانند حالت اول (هم پوشانی ۷۵ درصد) هر بخش را به ۸ زیربخش تقسیم کنیم زیرا در این حالت تمام زیربخش‌های سیگنال اصلی ساخته نمی‌شوند. برای توضیح بیشتر فرض می‌کنیم هر بخش به ۸ قسمت تقسیم شده باشد. در این صورت شکل (۱۳) بدست می‌آید:

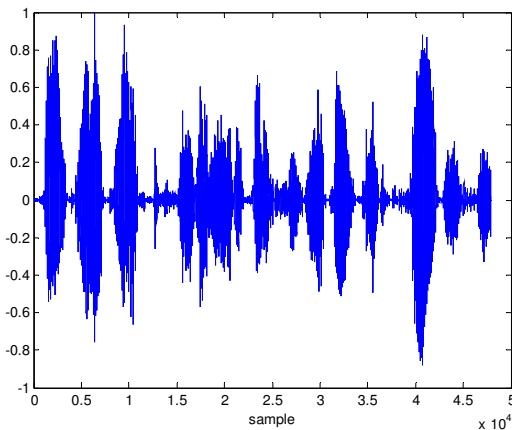


شکل (۱۳): انتخاب نادرست تعداد زیربخش‌های هر بخش

همانطوری که دیده می‌شود، با در کنار هم قرار دادن زیربخش‌های فوق تمام سیگنال اصلی ساخته نمی‌شود.

#### ۴- الگوریتم DUET

به منظور بررسی اثر روش پیشنهادی کاربرد تبدیل STFT در جداسازی منابع در نظر گرفته شده است. یکی از روش‌های مرسوم در جداسازی منابع، الگوریتم DUET است که جداسازی را در حوزه



شکل (۱۸): سیگنال جدا شده منبع اول با استفاده از روش دوم

مطلوب به توان اعوجاج، توان سیگنال‌های تداخلی، توان نویز محیطی و توان نویز حاصل از الگوریتم را محاسبه می‌کنند. نتایج مقایسه در جدول زیر آورده شده است.

جدول (۱): مقایسه پارامترهای کیفیت روش‌های عادی و پیشنهادی

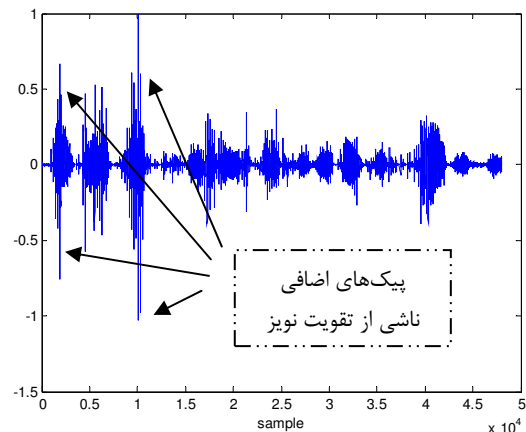
روش عادی:			
	منبع اول	منبع دوم	منبع سوم
SDR	-۱	-۵/۵	-۰/۹
SIR	۲۴	۲۵	۳۰
SNR	۴۰	۴۰	۴۵
SAR	-۱	-۵/۵	-۰/۹
روش پیشنهادی:			
	منبع اول	منبع دوم	منبع سوم
SDR	۸	۵	۷/۵
SIR	۲۶	۱۵	۲۵
SNR	۴۲	۴۳	۴۰
SAR	۸	۵	۷/۵

با مقایسه اعداد جدول (۱) بهبود کیفیت جداسازی سیگنال‌ها به خوبی مشاهده می‌شود. برای منابع دوم و سوم نیز وضعیتی مشابه وجود دارد که بطور خلاصه سیگنال‌های آنها در زیربخش‌های زیر رسم شده‌اند.

#### ۲-۵ منبع دوم

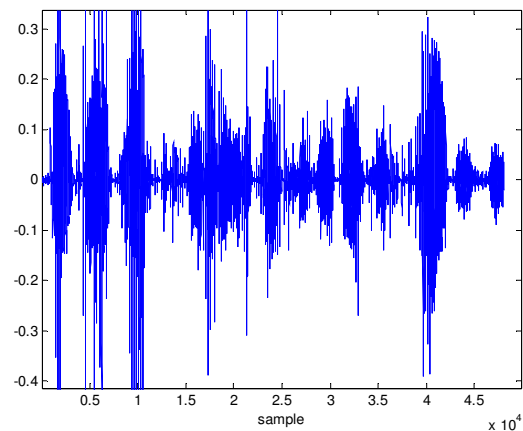
برای منبع دوم سیگنال اصلی و سیگنال جدا شده با استفاده از روش مذکور در شکل (۱۹) رسم شده‌اند.

با اعمال بلوک ISTFT به روش عادی، سیگنال جدا شده به صورت شکل (۱۶) خواهد شد. پیک‌های ناشی از تقویت خطا به قدری بزرگ هستند که سیگنال اصلی را بطور کامل تحت تاثیر قرار می‌دهند، لذا با نگاه اول در شکل (۱۶) شباهت زیادی بین آن و سیگنال اصلی مشاهده نمی‌شود.



شکل (۱۶): سیگنال جدا شده منبع اول با استفاده از ISTFT عادی

اما با بزرگ‌نمایی شکل (۱۶) به شکل (۱۷) می‌رسیم که شباهت بیشتری به سیگنال اصلی دارد.



شکل (۱۷): بزرگ‌نمایی شکل (۱۶)

وجود چنین پیک‌های بزرگی سبب ایجاد نویز زمینه‌ای می‌شود که به خوبی قابل شنیدن می‌باشد و کیفیت سیگنال جدا شده را کاملا تحت تاثیر قرار می‌دهد. اما با بکار بردن بلوک ISTFT به روش پیشنهادی، سیگنال جدا شده دارای چنین پیک‌هایی نبوده و لذا کیفیت سیگنال جدا شده بسیار بهتر از حالت قبل خواهد بود که در شکل (۱۸) نشان داده شده است.

برای مقایسه عددی، از پارامترهای SIR, SDR, SNR, SAR مطابق تعریف مرجع [8] استفاده می‌نماییم که به ترتیب میزان توان سیگنال

این روش این عیب بزرگ را دارد که در ابتدا و انتهای نقاط پنجره سبب تقویت خطای سیستم می‌شود که به صورت پیک‌های غیر ضروری در سیگنال جدا شده مشاهده می‌گردند و به صورت نویز زمینه شنیده می‌شوند. در روش پیشنهادی با حذف میانگین‌گیری بخش‌های دارای هم‌پوشانی، این مشکل تا حد قابل توجهی کاهش می‌یابد. در بخش شبیه‌سازی با ارائه پارامترهای محاسبه کیفیت جداسازی، تاثیر مثبت روش پیشنهادی نشان داده شد.

### سپاسگزاری

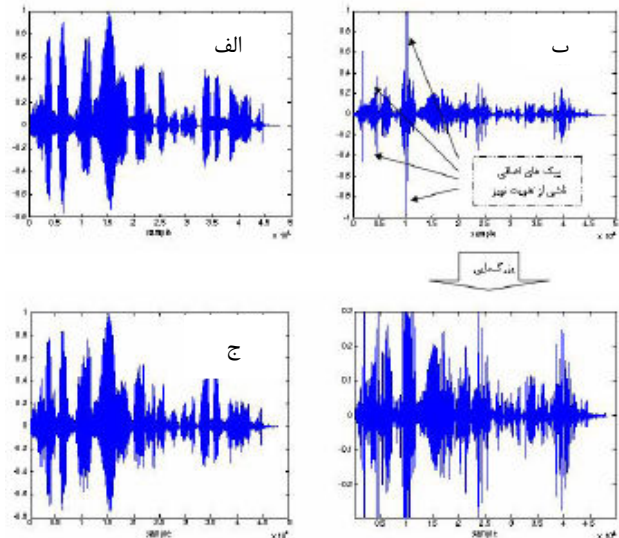
نویسندگان این مقاله از همکاری مرکز تحقیقات مخابرات ایران کمال سپاسگزاری را دارند.

### ۷- مراجع

- [1] Murata N., Ikeda S., Zeiha A., "An approach to blind source separation based on temporal structure of speech signals", *Neurocomputing*, vol. 41, pp. 1-24, Oct. 2001.
- [2] Smaragdīs P., "Blind separation of convolved mixtures in frequency domain", *Neurocomputing*, vol. 22, pp. 21-34, 1998.
- [3] Mukai R., Sawada H., Araki S., Makino S., "Frequency domain blind source separation using small and large spacing sensor pairs", in *Proc. ISCAS 2004*, vol. V, May 2004, pp. 1-4.
- [4] Yilmaz O., Rickadr S., "Blind separation of speech mixtures via time-frequency masking", *IEEE transactions on Signal Processing*, vol. 52, no. 7, pp.1830-1846, 2004.
- [5] Melia T., Rickard S., Fearon C., "Extending the DUET blind source separation technique", in *Proceedings of Signal Processing with Adaptive Sparse Structured Representation Workshop*, November 2005.
- [6] Sawada H., Araki S., Mukai R., Makino S., "Blind Extraction of Dominant Target Sources Using ICA and Time-Frequency Masking", *IEEE transactions on audio, speech and language processing*, vol. 14, no. 6, 2006.
- [7] Winter S., Kellermann W., Sawada H., Makino S., "MAP-Based Underdetermined Blind Source Separation of Convolutional Mixtures by Hierarchical Clustering and 11-Norm Minimization", *EURASIP Journal on Advances in Signal Processing*, vol. 2007, June 2006.
- [8] Fevotte C., Gribonval R., Vincet E., "BSS\_EVAL Toolbox User Guide", IRISA Tech Rep. 1706, 2005.  
[Online]. Available: [http://www.irisa.fr/metiss/bss\\_eval/](http://www.irisa.fr/metiss/bss_eval/)

### زیر نویس‌ها

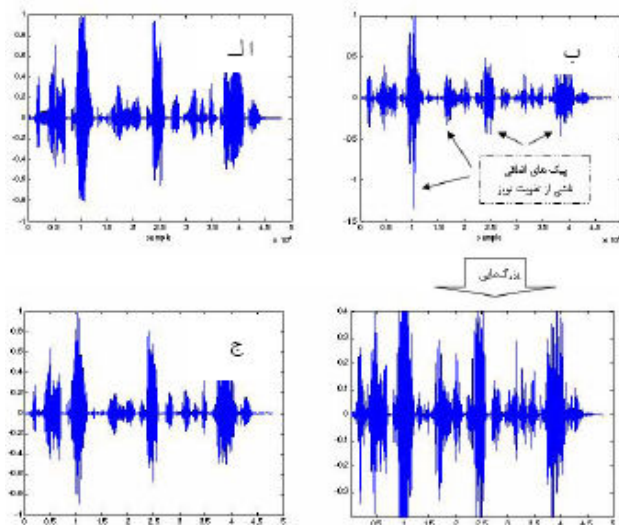
- 1 Blind Source Separation
- 2 speech
- 3 convolutive
- 4 impulse response
- 5 Frequency Domain ICA
- 6 Clustering
- 7 Short Time Fourier Transform
- 8 overlap
- 9 window
- 10 lable



شکل (۱۹): الف) سیگنال اصلی منبع دوم (ب) سیگنال جدا شده آن با روش عادی (ج) با روش پیشنهادی

### ۵-۳ منبع سوم

برای منبع سوم سیگنال اصلی و سیگنال جدا شده با استفاده از دو روش مذکور در شکل (۲۰) رسم شده‌اند.



شکل (۲۰): الف) سیگنال اصلی منبع سوم (ب) سیگنال جدا شده آن با روش عادی (ج) با روش پیشنهادی

### ۶- نتیجه‌گیری

در این مقاله، روشی جهت بهبود کیفیت جداسازی سیگنال منابع به روش کور با استفاده از بلوک عکس تبدیل فوریه اصلاح شده، مطرح شد. با توجه به مطالب بالا در روش عادی، در بدست آوردن عکس تبدیل فوریه پنجره‌ای از بخش‌های دارای هم‌پوشانی استفاده می‌شود.