

نشریه علمی پدافند غیرعامل

سال چهاردهم، شماره ۳، پیاپی ۱۴۰۲، (پیاپی ۵۵): صص ۱۳۰-۱۱۵

علمی - ترویجی

## نهان‌نگاری در صوت مبتنی بر الگوریتم تبدیل موجک درخت صفر درج شده

رضا اصفهانی<sup>۱\*</sup>، احمدرضا متین‌فر<sup>۲</sup>

تاریخ دریافت: ۱۴۰۲/۰۱/۰۸

تاریخ پذیرش: ۱۴۰۲/۰۵/۱۶

### چکیده

نهان‌نگاری در صوت همانند نهان‌نگاری در رسانه‌های دیگر (تصویر، ویدئو و ...)، از اهمیت ویژه‌ای برخوردار است. در این مقاله نهان‌نگاری در صوت مبتنی بر الگوریتم تبدیل موجک درخت صفر درج شده ارائه شده است. ارتقاء مقاومت در برابر نویز سفید و نویز افزایشی با کم‌ترین SNR یکی از موضوعات مهم در نهان‌نگاری است. الگوریتم پیشنهادی، در برابر نویز سفید نرمال بیشتر از نویز یکنواخت مقاوم بوده و در برابر SNRهای بالاتر از 10db نرخ بیت خطای کم‌تر از ۱ بیت دارد. در صورت حمله به الگوریتم پیشنهادی، با توجه به BER به دست آمده، سیگنال مخفی کاملاً تخریب می‌شود. همچنین روش پیشنهادی در برابر اضافه کردن نویز افزایشی مقاوم است. الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم و فایل‌های صوتی به فرم موزیک با تن (بلندی) ملایم، کم‌ترین تغییرات را در معیار همواری صوت دارد و افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد. الگوریتم پیشنهادی طیف فرکانسی، سیگنال صوتی را چندان تغییر نمی‌دهد و همچنین از خاصیت سطح آستانه شنوایی پیروی می‌کند و موزیک تن بالا با گفتار مرد دارای بهترین نتایج می‌باشند بنابراین از هر جهت مساعد با ساختار طیف بارک می‌باشند. همچنین الگوریتم پیشنهادی نتایج مطلوبی در حوزه زمان دارد. کم‌ترین SNR مربوط به موزیک تن بالا با گفتار زن می‌باشد که دارای SNR در حدود 13 db می‌باشد. با توجه به نتایج حاصل شده، با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت. زیرا در سیگنال گفتار زن همواری خاصی محسوس است. بنابراین با درج پیام مخفی در این نوع سیگنال صوتی، تا حدودی این فرم یکنواختی از بین خواهد رفت و با توجه به مقایسه شدن مولفه به مولفه دو سیگنال اصلی و سیگنال حاوی پیام مخفی، میزان معیار CZD افزایش می‌یابد.

**کلیدواژه‌ها:** تبدیل موجک درخت صفر درج شده (EZW)، نویز سفید، خطای افزایشی، معیار BSD و MBSD، مقیاس فاصله کپستروم CDM، فاصله ایتاکورا - سایتو ISD و فاصله COSH، تبدیل رادون زمان کوتاه STFRT، SNR و فاصله سزنا-کوشی CZD

<sup>۱</sup>استادیار گروه مخابرات دانشگاه جامع امام حسین (ع)، تهران، ایران - (resfahani@ihu.ac.ir) - نویسنده مسئول

<sup>۲</sup>استادیار دانشگاه جامع امام حسین (ع)، تهران، ایران

## ۱- مقدمه

## ۲- نهان‌نگاری در فایل صوتی

در حالت کلی برای نهان‌نگاری در فایل‌های صوتی مراحل زیر انجام می‌شود:

- نحوه عملکرد کد کننده MPEG: در کد کننده MPEG، مدل فیزیولوژیکی سیگنال‌های صوتی مورد آنالیز قرار می‌گیرد و میزان نویز ماسک‌شده موجود به عنوان تابعی از فرکانس محاسبه می‌شود.
- تطبیق (هم تراز کردن) زمانی داده‌های صوتی: در مدل فیزیولوژیکی که کم‌ترین پیچیدگی را دارد، محدوده فرکانسی به ۳۲ زیرباند فرکانسی تقسیم می‌شود.<sup>۵</sup> در فشرده‌سازی داده با استفاده از الگوریتم MPEG برای هر باند بحرانی، از تکنیک بانک فیلتر استفاده می‌شود [۶]. که به طریقه رابطه (۱) محاسبه می‌شود.

$$S_t[i] = \sum_{k=0}^{63} \sum_{j=0}^7 M[i][k] \times (C[k + 64j] \times x[k + 64j]) \quad (1)$$

$$M[i][k] = \cos \left[ \frac{(2i + 1) \times (k - 16) \times \pi}{64} \right]$$

که  $C[k]$  همان پنجره ماسک‌کننده در مدل مورد استفاده شده می‌باشد. مدل فیزیولوژیکی از یک نگاهت زمان به فرکانس مستقل و جدا به جای بانک فیلتر چند فازه استفاده می‌کند. زیرا به دقت وضوح فرکانس بیشتری در محاسبات تعیین سطح آستانه ماسک‌کننده نیاز دارد. طبق پیاده‌سازی نرم‌افزاری، یک وزن ده هانینگ<sup>۶</sup> بر روی داده‌ها قبل از تبدیل فوریه اعمال می‌شود تا اثرات لبه‌ها را در فریم صوتی کاهش دهد. تبدیل فوریه با توجه به طول پنجره آنالیز، ۵۱۲ در نظر گرفته می‌شود.

- فرایند محاسبات طیفی: در این مرحله، سطح بلندی هر طیف صوتی متناسب با زیر باند فرکانسی مورد نظر ( $Lsb^7$ )، محاسبه می‌شود.
- مجزا کردن مقادیر طیفی به مولفه‌های ضربه و غیر ضربه‌ای: هر دو مدل ضربه و مولفه‌های شبه نویزی را از سیگنال صوتی تعیین و جدا می‌کنند. این امر به این دلیل می‌باشد که قابلیت ماسک گذاری این دو نوع سیگنال متفاوت است.
- اعمال یک تابع پخش کننده: توانایی ماسک‌کنندگی یک سیگنال داده شده در محدوده ی باند بحرانی سیگنال انتشار (پخش) می‌شود.
- قراردادن باند پایین برای مقادیر سطوح آستانه: هر دو مدل شامل یک سطح آستانه ماسک‌کننده مطلق (سطح آستانه سکوت) می‌باشند که از روی آزمایشات به دست آمده است.

<sup>۵</sup> فیلتر شدن به ۳۲ زیر فضای یکسان وابسته به نرخ ناپکوئیست سیگنال

<sup>۶</sup> Hanning

<sup>۷</sup> Level subband

با توجه به وجود فرمت‌های مختلف صوتی مانند Wma, Ra, mp3, wav, Mid, Cdda, Aiff, می‌توان از این رسانه (صوت) برای هنر پنهان کردن اطلاعات در آن بهره جست. لذا بررسی ویژگی‌های صوت، فرمت‌ها و گذرهای صوتی، برای استفاده از روش‌های نهان‌نگاری در آن‌ها لازم و ضروری است [۳، ۲، ۱]. به طور کلی دو زیر فضای آنالوگ و زیر فضای دیجیتال برای محیط‌های صوت وجود دارد. برای ذخیره فایل‌های صوتی دیجیتال، مستقیماً از سیگنال صوتی آنالوگ نمونه‌برداری می‌شود. به عبارت دیگر فایل‌های صوتی دیجیتال از تبدیل صداهای آنالوگ به حوزه دیجیتال به دست آمده‌اند. فرآیند تبدیل اصوات آنالوگ به دیجیتال شامل دو مرحله چندی‌سازی کردن و نمونه‌برداری زمانی می‌باشد. الگوی چندی‌سازی<sup>۱</sup>: چندی‌سازی هر مقدار ورودی را به مقادیر گسسته تبدیل می‌کند. معمول‌ترین فرمت‌های ذخیره صدا با کیفیت بالا استفاده از فرمت‌های ۱۶ بیتی خطی چندی‌ساز شده است. همانند آن چیزی که در WAV یا AIFF<sup>۲</sup> استفاده می‌شود. مقداری از دامنه سیگنال‌های صوتی به وسیله این فرمت دچار اعوجاج می‌شود.

نرخ نمونه‌گیری زمانی: فرایندی است که مقادیر آنالوگ در تقسیمات زمانی منظم نمونه‌برداری می‌شوند. پُر استفاده‌ترین نرخ نمونه‌گیری زمانی<sup>۳</sup> برای صوت شامل 8KHz, 9.6KHz, 10KHz, 12KHz, 16KHz, 22.05KHz و 44.1KHz هستند. نرخ نمونه‌گیری تعداد نمونه‌های سیگنال صوتی است که در یک ثانیه از فریم صوت وجود دارد.

وسیله پخش یا محیط پخش در یک سیگنال صدا به محیطی گفته می‌شود که سیگنال در آن از کدکننده به کدگشا می‌رود. چهار محیط پخش شامل «محیط دیجیتالی نظیر به نظیر»، «محیطی با استفاده از کم یا زیاد کردن و نمونه‌گیری مجدد»، «پخش آنالوگ و نمونه‌گیری مجدد» و «محیط هوا» است. محیط پخش و همچنین وسیله ذخیره سیگنال، در نوع الگوریتم‌هایی که برای پنهان کردن اطلاعات به کار می‌رود، اهمیت دارند [۴، ۵].

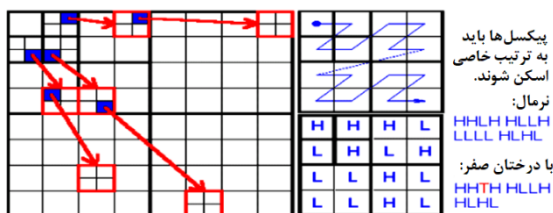
در بخش دوم موضوعات اولیه و اساسی نهان‌نگاری در فایل صوتی و تبدیل موجک درخت صفر درج شده (EZW) شرح داده شده است. ضرورت این یادآوری‌ها، به کاربرد و استفاده آن در الگوریتم پیشنهادی ارتباط دارد. در بخش سوم، روش پیشنهادی نهان‌نگاری در صوت مبتنی بر EZW ارائه شده است که شامل نحوه عملکرد سیستم، طراحی و نتایج پیاده‌سازی و ارزیابی الگوریتم پیشنهادی است.

<sup>۱</sup> Quantisation

<sup>۲</sup> Windows Audio Visual

<sup>۳</sup> Audio Interchange File Format

<sup>۴</sup> Temporal



شکل (۱): نمایش ترتیب پیمایش ضرایب تبدیل موجک [۷، ۸].

هر چند که استفاده از درخت صفر موقعیت‌های بسیاری را برای کد کردن ایجاد می‌کند، اما چندین حالت متفاوت از پیمایش به شرطی که زیرباندهای کوچک‌تر تماماً قبل از رفتن به سطح بالاتر پیمایش شوند نیز امکان‌پذیر است [۹، ۱۰].

### ۳- روش پیشنهادی

#### ۳-۱- خلاصه عملکرد سیستم در روش پیشنهادی

در این بخش هدف ارائه الگوریتم نهان نگاری جهت ذخیره داده مخفی در فایل‌های صوتی می‌باشد. اساس طراحی این الگوریتم بر پایه "الگوریتم فشرده‌سازی درخت صفر در تصاویر" می‌باشد. این الگوریتم از ویژگی طبیعی که در تصاویر وجود دارد، استفاده می‌کند. ویژگی که در تصویر طبیعی وجود دارد، تمایل تصاویر به توزیع نرمال می‌باشد. بنابراین از میان ضرایب موجود در ترکیب تصویر، ضرایبی که بزرگتر باشند مولفه‌های اصلی سازنده تصویر به شمار می‌آیند. بنابراین در بیان تصویر می‌توان از ضرایب کوچکتر صرف نظر کرد. الگوریتم درخت صفر بیانگر این موضوع است که ضرایب بزرگ در تشکیل تصویر نقش اصلی و سازنده آن را دارند و ضرایب کوچکتر تنها سازنده جزئیات تصویر می‌باشند. در الگوریتم درخت صفر ضرایب تصویر بر اساس تبدیل موجک در ۳ یا ۴ مرحله بیان می‌شوند و در ادامه ضرایبی را که در تخمین هر مرحله از مقدار قبل کوچکتر باشند را به عنوان ضرایب کم اهمیت در نظر می‌گیرد. زیرا این ضرایب در هر مرحله تخمین تبدیل موجک تنها سازنده جزئیات در بخش قبل می‌باشند و ضریب اصلی در مرحله اول تمامی این ضرایب را می‌تواند تحت پوشش قرار دهد زیرا ضریب مرحله بالاتر، بزرگتر از ضرایب مراحل پایین‌تر می‌باشد.

از این ایده که در "الگوریتم درخت صفر ضرایبی را می‌توان یافت که در تشکیل سیگنال کم‌ترین تاثیر را دارند در نهان نگاری صوتی استفاده شده است".

زیرا "همانند تصاویر، سیگنال‌های صوتی طبیعی نیز تمایل به توزیع نرمال دارند. علاوه بر آن سیگنال‌های صوتی واقعی، ترکیبی از چندین منبع صوتی مجزا می‌باشند که بر طبق اصل نرمال جمع بی‌نهایت سیگنال طبیعی با توزیع متفاوت، توزیع نرمال خواهد بود می‌توان این نتیجه را گرفت که سیگنال‌های طبیعی صوتی توزیع نرمال خواهند داشت".

- سطح آستانه سکوت، در باند پایین شنوایی صوت وجود دارد.
- یافتن سطح آستانه ماسک‌کننده برای هر زیر باند: هر دو مدل فیزیولوگوستیکی، سطح آستانه ماسک‌کننده‌ای با دقت تفکیک فرکانسی بالا برای هر زیر باند بطور مجزا محاسبه می‌کنند که توسط بانک فیلتر چند فازه به دست می‌آید.
- محاسبه نسبت سیگنال به ماسک: این نسبت به عنوان نسبتی از انرژی سیگنال در طول زیر باند به حداقل سطح آستانه ماسک‌کنندگی در آن زیر باند محاسبه می‌شود. در ادامه مولفه‌های ضربه و غیر ضربه یک سیگنال به همراه فضای ماسک‌کنندگی این مولفه‌ها در مقابل سطح آستانه مطلق ماسک‌کنندگی نمایش داده شده است.

#### ۲-۱- تبدیل موجک درخت صفر درج شده (EZW)

تبدیل موجک گسسته یک سیگنال دلخواه  $x$  به صورت رابطه (۲) بیان می‌شود:

$$DWT(\text{Scale}, \text{Position}) = \sum_{n=1}^N f(n)\Psi(\text{Scale}, \text{Position}, n) \quad (2)$$

که  $\Psi(\cdot)$  توابع متنوعی می‌باشد که می‌توان از آن‌ها بسته به کاربرد مورد نظر استفاده نمود. بهترین فیلترهای مورد استفاده در تجزیه سیگنال صوتی، استفاده از فیلترهای 'db' می‌باشد. در این مقاله از تبدیل موجک با فیلتر 'db' در چهار سطح استفاده شده است. الگوریتم تبدیل موجک درخت صفر درج شده در واقع الگوریتمی است که برای فشرده‌سازی تصویر و انتقال آن ارائه شده است. EZW کدکننده‌ای می‌باشد که براساس تبدیل موجک بنا شده است. کدکننده EZW بر پایه کد کردن پیشرو پایه‌گذاری شده است و یک تصویر را به دنباله‌ای از رشته‌ها با دقت افزایشی (پیشرو) فشرده می‌سازد. نکته مورد توجه، "استفاده از وابستگی بین ضرایب تبدیل موجک در طول کل سیگنال و بخش بزرگی از سیگنال کد شده که در زیر سطح آستانه قرار دارند" می‌باشد. در این حالت است که بحث درخت صفر مطرح می‌شود. الگوریتم EZW شرح داده شده توسط شیپپرو از روش کد کردن در سطح آستانه استفاده می‌کند [۷]. مساله مهم در این جا از دست دادن موقعیت ضرایب می‌باشد. بدون اطلاعات موقعیت، حتی اگر تمامی ضرایب هم کد شوند، ضرایب کدگشا، قادر به بازسازی سیگنال کد شده نمی‌باشد. این موضوع در نحوه انتخاب موقعیت، توسط کدکننده نهفته است. EZW از یک مرتبه پیمایش تعیین شده‌ای از موقعیت‌های ضرایب تبدیل موجک استفاده می‌کند. در شکل (۱) نمونه‌ای از ترتیب پیمایش ضرایب نمایش داده شده است.

<sup>1</sup> Embedded Zerootrees of Wavelet Transforms

درخت صفر برای یافتن مکان‌های ضرایب درخت صفر اعمال می‌شود و سپس بر طبق الگوریتم ذخیره‌سازی داده مخفی به استخراج پیام پرداخته می‌شود.

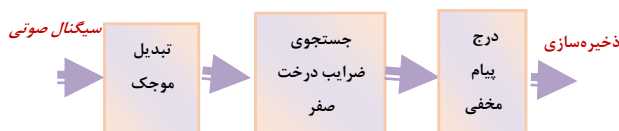
### ۲-۳- طراحی سیستم روش پیشنهادی

#### ۲-۳-۱- طراحی الگوریتم ذخیره‌سازی پیام مخفی

در این بخش به پیشنهاد الگوریتم مربوطه پرداخته می‌شود و نتایج کلی به همراه آزمون آن آورده شده است. همان‌طور که بیشتر ذکر گردید این پیشنهاد الگوریتم بر پایه تکنیک‌های تبدیل موجک و درخت صفر بنا شده است. در این قسمت روند طراحی الگوریتم بر اساس تئوری‌های موجود بیان شده، و با استفاده از الگوریتم گام به گام آزمون حدسیات ذکر می‌گردد. همان‌طور که بیان شد، الگوریتم درخت صفر بر اساس یافتن ضرایبی از تبدیل موجک می‌باشد که کم‌ترین تاثیر را در ساختار والد و اولاد تبدیل موجک دارد. بنابر تعریف درخت صفر، اگر تمامی اولاد یک والد از مقدار والد خود کم‌تر باشد آنگاه والد و اولاد آن را درخت صفر می‌نامند که با توجه به تعریف، تنها مقدار والد در ساختار سیگنال موثر بوده و می‌توان از ضرایب اولاد صرف نظر کرد. با توجه به ساختار دو بعدی الگوریتم درخت صفر، می‌بایستی این الگوریتم را برای تبدیل موجک سیگنال صوتی مناسب نمود. با توجه به الگوریتم درخت صفر که بیشترین عملکرد آن معطوف به جزئیات و مولفه‌های بالاگذر سیگنال می‌شود، بهترین مکان جهت اضافه کردن سیگنال اصلی و ضرایب تبدیل موجک مرتبه اول آن، ابتدای دنباله به دست آمده  $Xw$  می‌باشد. در ادامه نیز نحوه چیدمان ضرایب درخت صفر نمایش داده شده است. فرض شود سیگنال دلخواه یک سیگنال تصادفی  $S$  به طول ۱۶ نمونه باشد. از این سیگنال در سه سطح و با استفاده از فیلتر 'haar' تبدیل موجک گرفته می‌شود و در ادامه ضرایب درخت صفر آن محاسبه می‌شود. به کمک MATLAB موضوعات کلی زیر پیاده‌سازی شد:

- محاسبه تبدیل موجک از سیگنال تصادفی  $S$  در سه سطح با استفاده از فیلتر Haar
- چیدمان سیگنال  $X$  و بازسازی آن در دو بعد مطابق الگوریتم EZW
- محاسبه مکان ضرایب درخت صفر
- جایگذاری پیام (در این مثال آرایه ای از ۱) به جای ضرایب درخت صفر
- تبدیل معکوس گرفتن از ضرایب تبدیل موجک و بازیابی سیگنال جایگذاری شده با داده مخفی
- آزمون درستی بازیابی داده: اعمال الگوریتم درخت صفر بر روی سیگنال بازیابی شده و مقایسه مکان ضرایب درخت صفر به دست آمده با مکان ضرایب درخت صفر سیگنال اصلی

پس با پیدا کردن ضرایب کم اهمیت در سیگنال طبیعی صوتی می‌توان این ضرایب را به‌نحوی دلخواه تغییر داد که علاوه بر دارا بودن خاصیت درخت صفر حاوی اطلاعات داده مخفی نیز باشد. بلوک دیاگرام مراحل الگوریتم پیشنهادی ذخیره داده مخفی درون سیگنال صوتی در شکل (۲) آورده شده است.



شکل (۲): بلوک دیاگرام مراحل الگوریتم پیشنهادی ذخیره داده مخفی درون سیگنال صوتی

در مرحله اول اعمال تبدیل موجک و به دست آوردن ضرایب آن می‌باشد. تعداد مراتب تبدیل موجک و نوع فیلتر مورد استفاده در تبدیل موجک بسیار متنوع می‌باشد. فیلترهای تبدیل موجک برای کاربردهای متنوع و سیگنال‌های متفاوت تنوع زیادی یافته اند به طوری که با دانستن ویژگی ساختاری سیگنال می‌توان ضرایب بهتری از تبدیل موجک به دست آورد. البته ویژگی که در تمامی فیلترهای تبدیل موجک رعایت می‌شود استفاده از بردارهای پایه متعامد می‌باشد که با تکیه بر این موضوع می‌توان به این نکته دست یافت که ضرایب به دست آمده از هر مرحله برای تخمین سیگنال اصلی یکتا خواهد بود. زیرا همبستگی بردارهای متعامد صفر می‌باشد. فشردگی و کشیدگی فیلترهای مورد استفاده وابسته به کاربرد مورد نظر متفاوت می‌تواند باشد. به عنوان مثال در تخمین سیگنال‌های صوتی بیشتر از فیلترهای db استفاده می‌شود زیرا فیلترهای db دارای ویژگی اغتشاش سریع زمانی (چولگی بالا) می‌باشد. همچنین در کاربردهایی که سیگنال دارای خاصیت متقارن می‌باشد می‌توان از فیلترهای Sym استفاده نمود. مرحله دوم استخراج ضرایب درخت صفر می‌باشد. در این قسمت ضرایب سیگنال صوتی باید به نحوی تغییر یابد که متناسب با فرمت ضرایب تصویر باشد. زیرا ضرایب تصویر دو بعدی می‌باشند و ضرایب سیگنال صوتی یک بعدی هستند. پس از پیاده‌سازی سیگنال صوتی در قالب تصویر الگوریتم درخت صفر جهت یافتن ضرایب درخت صفر اعمال می‌گردد. موضوع آخر مربوط به درج پیام مخفی درون سیگنال صوتی می‌باشد. در این قسمت به تعریف روش و الگویی جهت درج پیام مخفی پرداخته می‌شود. الگوی ذخیره‌سازی نقش مهمی در بازیابی داده مخفی و همچنین پایداری الگوریتم در برابر حملات و همچنین میزان ظرفیت نهان‌نگاری دارد. و در نهایت سیگنال حامل داده‌های مخفی ذخیره می‌شود. در بازیابی داده‌های نهان‌نگاری شده از درون سیگنال حامل برعکس مراحل بالا دنبال می‌شود. در ابتدا سیگنال صوتی حامل داده مخفی به فرمت داده‌های صوتی شکل داده می‌شوند و در ادامه الگوریتم

برای به دست آوردن ضرایب مناسب می‌توان از فیلترهای دیگر تبدیل موجک استفاده نمود. اما به دلیل این که فیلترهای db سازگاری بیشتری با سیگنال‌های طبیعی دارند، این نوع فیلترها انتخاب شده‌اند. علاوه بر این، در بسیاری از مقاله‌های پژوهشی در رابطه با نهان‌نگاری و نشانه‌گذاری با استفاده از تبدیل موجک، فیلتر مورد استفاده، فیلترهای db بوده است. انتخاب درجه تبدیل موجک وابستگی زیادی به طول سیگنال در نظر گرفته شده دارد. معمولاً در پژوهش‌های انجام‌شده از مراتب ۳، ۴ و ۵ تبدیل موجک استفاده می‌شود [۱۱، ۱۲]. در این جا به دلیل این که طول سیگنال اولیه را ۱۶ در نظر گرفتیم، طول ضرایب تبدیل موجک مرتبه ۳ آن با استفاده از فیلتر Haar برابر با ۴ خواهد شد که در استفاده و تخمین سیگنال صوتی اولیه از لحاظ محاسبات و پیچیدگی بهترین انتخاب می‌باشد. زیرا انتخاب ضریبی با طول ۴ به مفهوم تخمین یک نقطه از سیگنال با همسایگی ۴ تایی می‌باشد. یعنی هر ضریب از سیگنال اصلی با پنجره‌ای به طولی ۴ مولفه از فیلتر انتخاب‌شده تخمین زده می‌شود. در ذخیره‌سازی داده باید توجه زیادی به تاثیر پیام مخفی در شکل سیگنال بازسازی شده داشت. بنابراین برای آزمون کارایی الگوریتم بدترین شرایط را برای داده مخفی در نظر می‌گیرند، حالت‌هایی که معمولاً در شرایط آزمون برای سیگنال مخفی در نظر می‌گیرند عبارتند از سیگنال مخفی با مولفه‌های تماماً یک، تماماً صفر یا یک در میان صفر و یک [۱۳، ۱۴، ۱۵، ۱۶، ۱۷]. برای درج بیت‌های پیام مخفی مکان‌هایی از درخت صفر در نظر گرفته می‌شوند که با مکان‌های قرارگیری ضرایب مولفه‌های تبدیل موجک مرتبه سوم سازگاری داشته باشند. همان‌طور که قبلاً نیز ذکر شد، این امر به دلیل این است که بازسازی سیگنال اصلی از روی مولفه‌های مرتبه سوم صورت می‌پذیرد. در این قسمت پس از تولید سیگنال تصادفی دلخواه، تبدیل موجک مرتبه سوم آن با استفاده از فیلتر db1 محاسبه می‌شود و سپس از ضرایب آن برای انتخاب درخت صفر از الگوریتم EZW استفاده می‌شود. مکان‌هایی که از الگوریتم EZW حاصل می‌شوند با مکان‌های ضرایب تبدیل موجک مرتبه سوم تطبیق داده می‌شوند و ضرایبی را که متناظر با آن‌ها در ماتریس درخت صفر، مقدار صفر دارند، به عنوان مکان‌های درج پیام مخفی انتخاب می‌شوند. در ادامه تمامی سیگنال درج‌شده را یکسان قرار داده می‌شود. در مرحله اول سیگنال درج‌شده را تماماً یک در نظر گرفته و سپس سیگنال اصلی از روی این ضرایب تغییر یافته بازسازی می‌شود. برای آزمون نتیجه عملکرد الگوریتم مجدداً از این سیگنال بازسازی‌شده تبدیل موجک در سه سطح گرفته‌شده و درخت صفر ریشه‌های آن بازبازی می‌گردد. در شکل (۳) روند شرح داده‌شده به همراه فلوجارت برنامه نمایش داده شده است.

مراحل جایگذاری داده مخفی به ۵ بخش متفاوت تقسیم می‌شود:

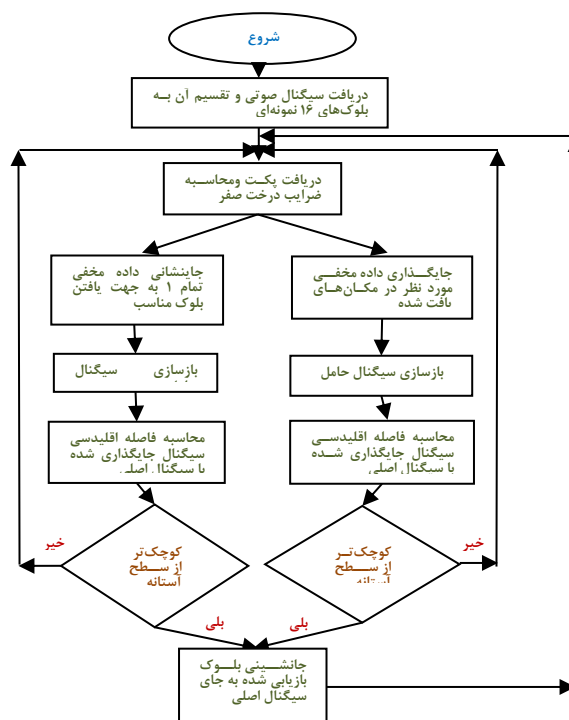
- ۱- محاسبه تبدیل موجک در سه سطح از سیگنال مورد نظر
  - ۲- چیدن ضرایب سیگنال مورد نظر به فرمت مناسب در دوبعد
  - ۳- محاسبه درخت صفر ضرایب سیگنال
  - ۴- جایگذاری داده مخفی در مکان‌های مورد نظر
  - ۵- بازبازی سیگنال با استفاده از تبدیل معکوس تبدیل موجک
- پس از جایگذاری داده مخفی آزمون درستی بازبازی داده پنهان‌شده یا همان Verification باید صورت پذیرد. یعنی آیا در بازبازی داده مخفی از ضرایب درخت صفر، مکان‌های جدید به دست آمده از اعمال تبدیل EZW بر روی سیگنال حامل اطلاعات مخفی با مکان‌های ضرایب درخت صفر سیگنال اصلی تطابق دارد یا خیر؟ برای این منظور ۳ مرحله بالا بر روی سیگنال بازبازی‌شده (سیگنال حامل اطلاعات مخفی) انجام می‌گیرد. در صورت درستی عملکرد بازبازی مکان‌های درخت صفر، میزان اغتشاش واردشده در سیگنال حامل و سیگنال اصلی با استفاده از محاسبه فاصله اقلیدسی میان این دو سیگنال محاسبه می‌شود. نتایج عددی الگوریتم سیگنال تصادفی S در فضای دو بعدی و همچنین ضرایب آن در مقیاس ۲۵۶ تا ۲۵۶ را بازسازی شد. ماتریس خروجی الگوریتم درخت صفر پس از حذف کردن ضرایب کم ارزش به دست آمد. در ادامه با استفاده از آن، و تبدیل معکوس تبدیل موجک، سیگنال اصلی بازسازی شد. در این حالت سطح آستانه را برای یافتن درخت صفر ضرایب سیگنال برابر یک شانزدهم بزرگترین مقدار سیگنال قرار دادیم. بدیهی است که برای بدست آوردن قدرت مقاومت بیشتر سطح آستانه را می‌توان بزرگتر انتخاب نمود. هرچه سطح آستانه بزرگتر باشد، مقاومت الگوریتم در برابر حملات بیشتر خواهد شد اما میزان خرابی سیگنال نسبت به قبل آن بیشتر خواهد شد.

همان‌طور که اشاره شد، پس از ترکیب سیگنال حامل با پیام مخفی باید صحت عملکرد الگوریتم در بازبازی رشته داده پیام مخفی سنجیده شود. بنابراین یک بار دیگر الگوریتم را بر روی سیگنال بازسازی‌شده اعمال شد. نتیجه حاصله از خروجی درخت صفر این فرایند باید با نتیجه درخت صفر سیگنال اصلی تطابق داشته باشد. یعنی هر دو خروجی درخت‌های صفر حاصل از الگوریتم بر روی دو سیگنال اصلی و بازسازی‌شده نتایج یکسانی را داشته است. حال که طرح اولیه طراحی الگوریتم خروجی مثبتی را در بر داشت، در ادامه، نحوه ذخیره‌سازی داده‌های پیام مخفی درون مکان‌های درخت صفر آورده شده است.

### ۳-۲-۲- پیاده‌سازی الگوریتم ذخیره‌سازی پیام مخفی

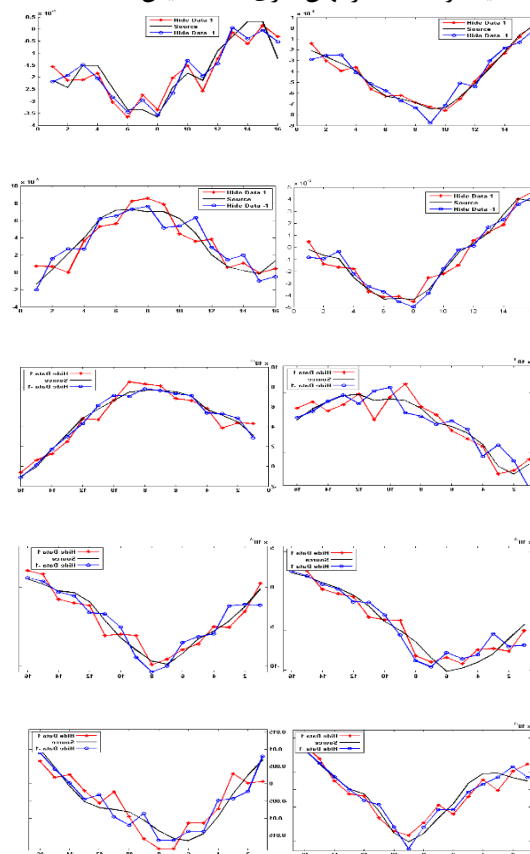
الف) الگوریتم پیشنهادی اول: در این مقاله از تبدیل موجک مرتبه سوم با فیلترهای db استفاده شده است. بدیهی است که

همان‌طور که مشاهده می‌شود، میزان تفاوت میان سیگنال‌های حامل با سیگنال اصلی ناچیز (در حدود 3-10) می‌باشد. در شکل ۴، منحنی بدون علامت و ساده مربوط به سیگنال اصلی است. در این آزمایش، سیگنال صوتی اصلی به بلوک‌های ۱۶ نمونه‌ای تقسیم‌بندی شده‌اند و در تمامی بلوک‌ها پیام مخفی به دو صورت دنباله‌ای از ۱ (منحنی با علامت \*) و ۱- (منحنی با علامت دایره‌های کوچک) نهان‌نگاری شده است. دامنه ۱ به منظور درج کد ۱ و ۱- به منظور درج کد صفر است که در هنگام نهان‌نگاری دنباله پیام مخفی دامنه‌ها در نظر گرفته شده، در حد آستانه‌ای ضرب می‌شوند. اما نکته‌ای که در این جا باید مد نظر گرفته شود این است که درج بیت پیام مخفی باعث خرابی در ساختار سیگنال صوتی نشود. زیرا خرابی و شکست در سیگنال زمانی همانند اضافه شدن تَن فرکانسی در طیف فرکانسی سیگنال و ایجاد اغتشاش و خرابی در سیگنال می‌شود. بنابراین بعد از انجام الگوریتم، فاصله اقلیدسی سیگنال بازسازی شده و سیگنال اصلی محاسبه می‌شود و در صورتی که مقدار این فاصله از یک میزان معین (حد آستانه) بیشتر نباشد، با توجه به الگوریتم درخت صفر بلوک درج‌شده را می‌توان به‌عنوان بلوک مساعد جهت نهان‌نگاری در نظر گرفت و می‌توان داده‌های بازبازی شده با استفاده از پیام مخفی را به‌جای داده‌های اصلی سیگنال قرار داد. در صورتی که میزان فاصله اقلیدسی بیشتر از حد آستانه مورد نظر باشد سیگنال اصلی بدون تغییر باقی می‌ماند. در اینجا دلیل استفاده از فاصله اقلیدسی بدین جهت می‌باشد که تغییرات صورت پذیرفته بر روی ضرایب تبدیل موجک با استفاده از فیلتر Haar صورت پذیرفته است. بنابراین بازسازی دوباره سیگنال اصلی در حوزه زمان می‌باشد. بنابراین با توجه به تغییرات زمانی صورت پذیرفته شده بر روی سیگنال ساده ترین معیار جهت اندازه‌گیری تغییرات دو سیگنال در حوزه زمان استفاده از معیار فاصله اقلیدسی می‌باشد. البته لازم به ذکر است که استفاده از چندین معیار به‌عنوان استخراج‌کننده ویژگی دارای مزیت‌های بیشتری می‌باشد اما با اضافه کردن یک معیار جدید میزان پیچیدگی محاسبات و پیاده‌سازی آن چندین برابر می‌شود. با افزایش حد آستانه حجم ذخیره داده‌های پیام مخفی افزایش می‌یابد اما در عوض سیگنال بازسازی شده بیشتر دچار خرابی و اغتشاش می‌شود. بنابراین بین میزان حد آستانه کیفیت و ظرفیت نهان‌نگاری رابطه معکوس وجود دارد یعنی با افزایش ظرفیت نهان‌نگاری، کیفیت سیگنال صوتی کاهش پیدا می‌کند. در نتایج به‌دست آمده حد آستانه  $0.1 < \text{Threshold} < 0.8$  عملکرد خوبی در نهان‌نگاری داده‌ها در هر دو وضعیت داده صفر و یک دارد. بنابراین در این مقاله بیشتر از این محدوده آستانه جهت آزمون الگوریتم استفاده شده است. بازبازی اطلاعات ذخیره‌شده عکس مراحل بالا را طی می‌کند.



شکل (۳): فلوجارت پیشنهادی اول

در شکل (۴) چند نمونه از عملکرد الگوریتم ذکر شده که با دنباله تماماً یک و تماماً صفر نهان‌نگاری شده، نمایش داده شده است.



شکل (۴): چند نمونه از عملکرد الگوریتم پیشنهادی اول، با دنباله تماماً یک و تماماً صفر نهان‌نگاری شده



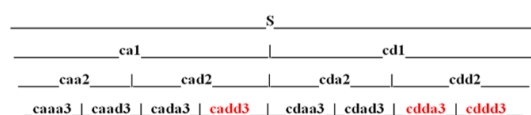
مفاهیم واقعی دارد. بنابراین با صرف نظر از این ضرایب می توان داده مخفی را در این مکان ها نهان نگاری نمود. یکی از مزیت های این کار افزایش سرعت پردازش می باشد.

د) الگوریتم سرفایل گذاری پیام مخفی: در این قسمت به شرح نحوه سرفایل گذاری بر روی داده های پرداخته می شود. سرفایل گذاری داده ها از این لحاظ دارای اهمیت می باشد که ممکن است طول پیام مخفی از ظرفیت نهان نگاری کوچک تر باشد در این صورت برای دیکد پیام درج شده با مشکل روبه رو خواهیم شد. زیرا بازبایی اطالات به صورت کاملاً کور صورت می پذیرد و در دیکد کردن داده مخفی به سیگنال اصلی اولیه که بر روی آن نهان نگاری اعمال شده است دسترسی نخواهیم داشت. بنابراین الگوریتم رمزگشایی داده مخفی تمامی طول ظرفیت نهان نگاری یک سیگنال را به عنوان داده مخفی بازبایی خواهد نمود. بنابراین باید الگوریتمی جهت مشخص نمودن ابتدا و انتهای داده مخفی مشخص نمود. همچنین از دیگر کاربردهای سرفایل گذاری بر روی پکت های داده مخفی می توان به نهان نگاری داده در چندین فایل اشاره نمود. به عبارت دیگر در صورتی که حجم یا ظرفیت نهان نگاری سیگنال صوتی کوچکتر از حجم پیام مخفی باشد، می توان پیام مخفی را به چندین پکت تقسیم نمود و هر پکت را به طور جداگانه درون فایل های متفاوت و مجزا ذخیره نمود. البته سرفایل در نظر گرفته شده باید دارای سه خصوصیت یکتایی، طول کوچک و افزونگی کم باشد. سرفایل در نظر گرفته شده نباید در طول داده های پیام مخفی وجود داشته باشد زیرا در غیر این صورت دنباله استخراجی اشتباه خواهد بود. علاوه بر این طول سر پکت نباید از یک حدی بیشتر باشد. زیرا طول زیاد سرفایل از میزان ظرفیت نهان نگاری کاسته خواهد کرد. الگوریتم استاندارد دی که برای سرفایل گذاری داده های مخفی در نظر گرفته شده است، الگوریتم HDLC می باشد که یک الگوریتم استاندارد برای ارسال داده بر روی هر نوع خط ارتباطی می باشد. در این الگوریتم دنباله "01111110" به عنوان سر فایل هر نوع پکت داده اعم از داده های مخابراتی یا غیره در نظر گرفته می شود. زیرا دو خصوصیت طول کم و افزونگی کم را دارد. به دلیل این که این دنباله ممکن است درون داده های یک پکت ظاهر شود، قبل از اضافه کردن این سرفایل به داده ها عمل Bit Stuffing بر روی داده ها صورت می پذیرد. بدین صورت که ابتدا درون دنباله رشته داده های پیام به جستجوی دنباله "1111" پرداخته می شود و صرف نظر از این که مقدار داده چه عددی باشد یک عدد "0" به انتهای این دنباله اضافه می شود. در این صورت درون دنباله داده پیام مخفی دیگر ۶ عدد ۱ پشت سر هم نخواهیم داشت و تنها با اضافه کردن سرفایل به دنباله ۶ عدد ۱ پشت سر هم خواهیم داشت که نشانگر ابتدا و انتهای فایل خواهد بود. در این مقاله پس از دریافت پکت داده مخفی و

بدین صورت که مکان های ضرایب درخت صفر سیگنال حامل، با استفاده از الگوریتم درخت صفر جستجو می شوند. در ادامه ضرایب درخت صفر به عنوان داده های نهان نگاری شده استخراج می شوند. در صورتی که داده ها بزرگتر از صفر باشند به عنوان داده نهان نگاری شده ۱ و در صورتی که داده ها کوچکتر از صفر باشند داده های به عنوان داده نهان نگاری شده صفر در نظر گرفته می شوند.

ب) الگوریتم پیشنهادی دوم: علاوه بر الگوریتم پیشنهاد شده الگوریتم دیگری نیز قابل دستیابی می باشد. همان طور که قبلاً ذکر شد داده های پیام به جای ضرایب درخت صفر در سیگنال صوتی جایگذاری می شوند. در الگوریتم پیشنهاد شده به ازای هر ضریب درخت صفر یافته شده درون هر بلوک یک عدد از دنباله پیام مخفی قرار می گیرد. بنابراین احتمال خراب شدن داده مورد نظر در اثر حملات افزایش می یابد. در الگوریتم پیشنهادی دوم، به جای تمامی ضرایب درخت صفر که از یک بلوک به دست می آید تنها یک عدد از دنباله پیام مخفی ذخیره سازی شود در این صورت با تغییر یکی از ضرایب درخت صفر پیام مخفی دچار خرابی خواهد شد.

ج) الگوریتم پیشنهادی سوم: مطابق شکل (۵) همچنین در آزمون های انجام شده مشاهده گردید که مکان هایی از درخت صفر نسبت به دیگر مکان ها بسیار تکرار می شوند. یعنی ضرایب این مکانها صرف نظر از مقادیر آنها دارای ارزش ساختاری چندانی در سیگنال صوتی نمی باشند. این مکانها عبارتند از Cadd3 و Cdda3. این مکانها در شکل زیر نمایش داده شده اند.

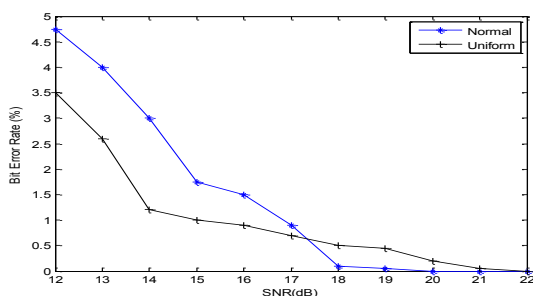


شکل (۵): مکان های درخت صفر در الگوریتم پیشنهادی سوم

همان طور که قبلاً نیز ذکر شد، مولفه های فرکانس پایین (DC) ساختار کلی سیگنال را تشکیل می دهند و مولفه های فرکانس بالا جزئیات سیگنال را تشکیل می دهد. ضرایب تبدیل موجک Cadd3 مربوط به تشکیل جزئیات فرکانس های پایین (DC) سیگنال صوتی می باشد پس بنابراین منطقی می باشد که در تشکیل ساختار اصلی سیگنال صوتی تاثیر چندانی نداشته باشد. همچنین Cdda3 و Cddd3 به ترتیب مربوط به تشکیل ساختار فرکانس پایین و فرکانس بالای جزئیات سیگنال صوتی (مولفه های فرکانس بالای سیگنال) می باشند. به دلیل این که باز جزئیات (مولفه های فرکانس بالا) تاثیر چندانی در تشکیل ساختار شکل موج صوتی ندارند بنابراین تجزیه این عوامل به زیر مجموعه فرکانس پایین و فرکانس بالا نیز اهمیت چندانی ندارند. با توجه به این مطلب در می یابیم تجزیه درخت صفر نتایج یکسانی را با

نویز کانال پرداخته شد. همان گونه که قبلا هم اشاره شد، حملات در نهان نگاری به دو مفهوم و نوع کلی تقسیم بندی می شوند. نوع اول، حملاتی می باشند که به منظور تخریب سیگنال نهان نگاری شده بر روی سیگنال حامل پیام مخفی صورت می پذیرد. هدف از این نوع حملات از بین بردن پیام مخفی درج شده درون سیگنال حامل و جلوگیری از ارسال پیام مخفی به مقصد می باشد. نوع دوم، حملاتی هستند که به منظور کشف و دیکد کردن پیام مخفی بر روی سیگنال حامل پیام مخفی صورت می پذیرد. هدف از حملات نوع دوم به دست آوردن محتوی پیام مخفی و آگاهی از آن می باشد. به دلیل این که بستر کاری در این مقاله سیگنال صوتی دیجیتال می باشد، مباحثی از قبل نویز محیط، نویز کانال و نویز گیرنده بی مفهوم خواهد بود. به عبارت دیگر، در برخورد با سیگنال های دیجیتال این فرضیات در نظر گرفته می شود که سیگنال دیجیتال از لحاظ نویز ایمن خواهد بود. زیرا در صورتی که چنین فرضیه ای در نظر گرفته نشود به هیچ یک از فایل های دیجیتال که بر روی یک بستر منطقی ذخیره می شوند نمی توان اعتماد کرد. زیرا با جابجایی و تغییر در یک بیت از فایل دیجیتال، محتوای کل فایل از بین خواهد رفت. از دیدگاهی دیگر، بررسی نویز میزان قابلیت اعتماد به الگوریتم ذخیره سازی و نهان نگاری را افزایش می دهد. بنابراین از این دیدگاه، بررسی نویز و تاثیر آن بر سیگنال نهان نگاری شده خالی از لطف نخواهد بود. از این رو در ادامه به بررسی این موضوع خواهیم پرداخت. بررسی و آزمون حملات بر روی سیگنال حامل پیام مخفی در این مقاله به دو صورت اضافه کردن نویز سفید به سیگنال و اندازه گیری میزان SNR سیگنال و همچنین حملات از نوع اضافه کردن مولفه به نمونه های سیگنال حامل پیام مخفی و اندازه گیری میزان خطای بازبازی داده مخفی صورت خواهد پذیرفت.

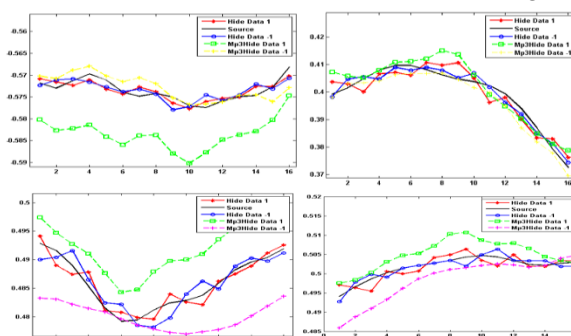
○ افزودن نویز سفید و بررسی میزان سیگنال به نویز: در این آزمون که با نرم افزار Matlab صورت انجام می شود یک سیگنال صوتی با فرکانس ۴۴۱۰۰ هرتز و کوانتیزاسیون ۱۶ بیت در نظر گرفته شده است. نویز اعمالی دو نوع نویز سفید نرمال، و نویز سفید یکنواخت انتخاب شده است. نتایج این تحقیق مطابق شکل (۷) می باشد.



شکل (۷): افزودن نویز سفید و بررسی میزان سیگنال به نویز

تبدیل آن به دنباله ای از "0" و "1" عملیات Bit Stuffing بر روی دنباله اعداد انجام می شود و در نهایت دنباله "01111110" به انتهای داده های پیام مخفی اضافه می شود.

در پیاده سازی الگوریتم پیشنهادی، در صورتی که مقدار خروجی خطای حاصل از الگوریتم پیشنهادی صفر باشد به مفهوم این است که بلوک دریافتی حاوی اطلاعات پیام مخفی می باشد. در صورتی که مقدار خطا مخالف صفر باشد بلوک مورد نظر مناسب نهان نگاری نمی باشد. در این مقاله با استفاده از این توابع فایل MP3 به فایل صوتی تبدیل می شود و سپس عملیات مورد نظر بر روی فایل صوتی انجام می پذیرد. خوبی این کار در این است که با یکبار MP3 کردن فایل صوتی مقادیر زیادی از نمونه هایی که در فایل صوتی شنیده نمی شوند و در طی پخش فایل ماسک می شوند، حذف خواهند شد. بنابراین این اطمینان را خواهیم داشت که سیگنال نهان نگاری شده در اثر MP3 شدن تغییرات چندانی نخواهد کرد. در شکل (۶) چند نمونه بلوک صوتی نهان نگاری شده که یک بار MP3 شده است و دوباره به فایل صوتی دیکد شده است را در مقایسه با سیگنال صوتی



شکل (۶): چند نمونه بلوک صوتی نهان نگاری شده یک بار MP3 شده

و دوباره به فایل صوتی دیکد شده، در مقایسه با سیگنال صوتی همان طور که در شکل (۶) مشاهده می شود، Mp3 نمودن فایل صوتی که یک بار MP3 شده است تغییرات چندانی نخواهد داشت. در این شکل، نمودارها به صورت منحنی با علامت \* برای پیام مخفی با دنباله ای از ۱، منحنی بدون علامت و ساده برای سیگنال اصلی، منحنی با علامت دایره های کوچک برای پیام مخفی با دنباله ای از ۱-، منحنی با علامت مستطیل های کوچک برای MP3 با دنباله ای از ۱ و منحنی با علامت + برای MP3 پیام مخفی با دنباله ای از ۱- مشخص شده است.

### ۳-۳-۳-۳-۳ ارزیابی الگوریتم پیشنهادی

#### ۳-۳-۳-۱-۳-۳ بررسی عملکرد الگوریتم پیشنهادی و پایداری در برابر حملات

در این بخش به کارایی الگوریتم پیشنهادی در برابر حملات و



یک نمونه داده در میان نمونه‌های سیگنال صوتی اصلی، تمامی اطلاعات مخفی شده درون سیگنال حاصل در صورتی از الگوریتم‌های تصحیح خطا استفاده نشده باشد از بین خواهد رفت. این در حالی است که رخداد خطا به صورت انتشاری در میان تمامی داده‌ها پخش خواهد شد. بنابراین با توجه به اهمیت پایداری الگوریتم در برابر حملات تخریبی، در این بخش به بررسی پایداری الگوریتم پیشنهادی در برابر نویز افزایشی خواهیم پرداخت. نحوه ایجاد نویز افزایشی به چندین صورت مختلف امکان پذیر می‌باشد. بنابراین از میان روش‌های مختلف، چند روش زیر برای آزمون این قسمت در نظر گرفته شده است:

الف) اضافه نمودن یک نمونه تصادفی در میان داده‌های سیگنال صوتی: در این قسمت بدون در نظر گرفتن محتوای سیگنال، نمونه‌هایی در میان نمونه‌های سیگنال اصلی درج می‌شود. با توجه به این که تعداد حالت‌های درج نمونه در این حالت بسیار زیاد می‌باشد، سه روش زیر برای آزمون این بخش انتخاب شد:

الف-۱) اضافه کردن یک نمونه معین به صورت یک در میان بین نمونه‌های سیگنال صوتی: با توجه به این که حملات صورت پذیرفته شده نباید کیفیت سیگنال صوتی را مخدوش کند، بنابراین در این حالت به صورت یک در میان متوسط هر دو نمونه مجاور محاسبه می‌گردد و سپس این مقدار به عنوان نمونه اضافه شده به سیگنال اصلی در نظر گرفته می‌شود. با این روش، کیفیت سیگنال صوتی تغییری نخواهد داشت و تنها حجم فایل صوتی دو برابر می‌شود. این روش بر روی ۱۵ نمونه سیگنال صوتی متفاوت به صورت تصادفی انجام شده است. رابطه (۳) نحوه به دست آوردن نمونه اضافه شده را نمایش می‌دهد.

$$YT(N) = (YS(N-1) + YS(N+1)) / 2; \quad (3)$$

در این حالت در صورتی که نرخ بیت نمونه برداری سیگنال اصلی به ۲ برابر حالت قبل تغییر کند، سیگنال صوتی از لحاظ کیفیت تغییر چندانی نخواهد کرد.

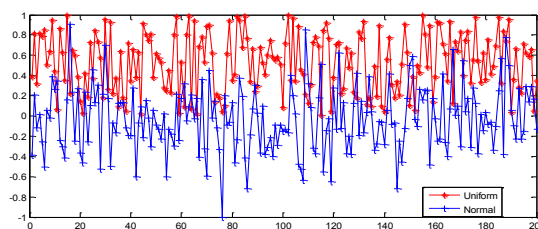
الف-۲) اضافه کردن یک نمونه معین به صورت تصادفی در بین نمونه‌های سیگنال صوتی: تفاوت این روش با روش الف در انتخاب مکان‌های جایگذاری نمونه مورد نظر می‌باشد. در این روش به تعداد ۱۰ درصد از کل نمونه‌های سیگنال صوتی، نمونه مورد نظر در بین نمونه‌های سیگنال صوتی جایگذاری می‌شود.

الف-۳) اضافه کردن یک نمونه تصادفی به صورت تصادفی در بین نمونه‌های سیگنال صوتی: در این روش، علاوه بر این که دامنه‌های نمونه درج شده به طور تصادفی انتخاب می‌گردند، مکان‌های درج نیز به طور تصادفی به تعداد ۱۰ درصد از کل نمونه‌های سیگنال صوتی انتخاب می‌شوند.

درصد نرخ خطای بیت برای سه روش اضافه نمودن یک نمونه تصادفی در میان داده‌های سیگنال صوتی به ترتیب الف-۱

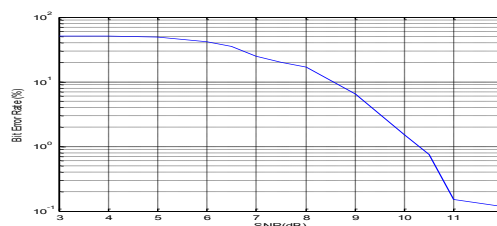
همان‌طور که از شکل (۷) مشخص است، پایداری الگوریتم پیشنهادی در برابر نویز سفید نرمال (منحنی با علامت +) بیشتر از نویز یکنواخت (منحنی با علامت \*) می‌باشد.

در شکل (۸) دو نمونه از توزیع‌های به کار برده شده نمایش داده شده است. حالت نرمال با علامت + و حالت یکنواخت با علامت \* مشخص شده است.



شکل (۸): توزیع به کار برده شده نویز سفید نرمال، و نویز سفید یکنواخت

در ادامه به دلیل کاربردی‌تر بودن روش دوم از الگوریتم پیشنهادی، به بررسی بیشتر میزان تاثیر نویز بر روی این الگوریتم و آزمون‌های آن پرداخته شده است. نمونه‌های مورد استفاده برای آزمون آماری ۱۰ نمونه صوتی متفاوت می‌باشد. نمونه‌های صوتی شامل ۲ نمونه‌ی صوتی با گفتار مرد، ۲ نمونه‌ی صوتی با گفتار زن، ۲ نمونه‌ی صوتی با گفتار کودک، ۲ نمونه‌ی صوتی موسیقی و ۲ نمونه‌ی صوتی مختلط می‌باشد. نرخ خطای بیت (BER) برای ۱۰ بار تکرار به صورت متوسط برای هر کدام از نمونه‌ها بر حسب لگاریتم محاسبه شده است. نتایج محاسبات خطا در شکل (۹) آورده شده است.



شکل (۹): نتایج محاسبات خطا

همان‌طور که مشاهده می‌شود، الگوریتم پیشنهادی در برابر SNR های بالاتر از 10db نرخ بیت خطای کم‌تر از ۱ بیت دارد.

○ **بررسی عملکرد الگوریتم پیشنهادی در برابر خطای افزایشی:** در این قسمت مقاومت الگوریتم پیشنهادی را در برابر حملات افزایشی بررسی می‌کنیم. نحوه عملکرد حملات افزایشی بدین صورت می‌باشد که در بین داده‌های سیگنال صوتی بلوک‌هایی به صورت تصادفی اضافه می‌گردد. بدین شکل، در صورتی که الگوریتم به کار رفته نسبت به مکان چه در حوزه زمان و چه در حوزه فرکانس ایمن نباشد، بازیابی اطلاعات مخفی به طور صحیح صورت نخواهد پذیرفت. به عنوان مثال در صورتی که از روش LSB برای نهان نگاری استفاده شود، با اضافه کردن

سیگنال صوتی: در این روش بلوک‌هایی با توزیع یکنواخت در میان بلوک‌های سیگنال صوتی حامل پیام مخفی اضافه می‌شود. به دلیل این که بلوک‌های سکوت تقریباً دارای شکل یکنواخت و مقدار دامنه کوچک هستند، در اینجا برای بررسی بیشتر الگوریتم پیشنهادی، از توزیع یکنواخت در انتخاب نمونه‌های بلوک درج‌شده استفاده شده است. با توجه به این که معمولاً بلوک‌های سکوت پیش و یا پس از یک گام صوتی در سیگنال صوتی ظاهر می‌شوند، بنابراین بهترین مکان جهت اضافه کردن بلوک درج‌شده پیش و یا پس از بلوک‌های سکوت به نظر می‌رسد. در ادامه برای یافتن بلوک‌های سکوت از دو تکنیک ساده آماری میانگین و واریانس نمونه‌ها استفاده شده است. در اینجا، بلوک‌هایی که میانگین آماری آن‌ها از یک سطح آستانه کوچکتر باشند و همچنین واریانس نمونه‌های آن بزرگتر از  $0.73$  باشد به عنوان بلوک سکوت در نظر گرفته می‌شود. درصد نرخ خطای بیت برای حالت اضافه نمودن بلوک‌های سکوت در میان پیچ‌های صوتی سیگنال صوتی در روش بلوک سکوت نوین یکنواخت با طول ۱۶ نمونه برابر با  $1.25$  به دست آمد.

نتیجه این بخش نیز مانند بخش قبلی مشابه می‌باشد. زیرا بلوک‌هایی که دارای توزیع یکنواخت هستند، مستعد درج پیام مخفی با استفاده از روش الگوریتم پیشنهادی نخواهند بود.

### ۳-۲-۳- بررسی عملکرد الگوریتم پیشنهادی در برابر فشرده‌سازی و MP3

در این قسمت به پایداری الگوریتم پیشنهادی در برابر فشرده‌سازی و MP3 پرداخته شده است. همان‌طور که در قبل شرح داده شد، پیام مخفی بر روی نمونه‌های PCM صوت درج می‌شوند. نحوه ذخیره‌سازی اطلاعات بدین صورت می‌باشد که در ابتدا فایل MP3 انتخاب‌شده را به فرمت WAV تبدیل کرده و سپس الگوریتم پیشنهادی بر روی آن اعمال می‌شود.

در ادامه خروجی الگوریتم که به فرمت WAV می‌باشد به MP3 تبدیل می‌شود. در انتها برای اندازه‌گیری و محاسبه میزان خطای حاصله، فایل MP3 به دست آمده، دوباره به WAV تبدیل می‌شود و با استفاده از این فایل الگوریتم دیکدکننده پیشنهادی بر روی آن اعمال می‌شود. در نهایت میزان خطای حاصله بر حسب بیت محاسبه می‌شود. برای انجام این آزمون از فایل‌های صوتی با نرخ بیت ۱۲۸ کیلو بیت بر ثانیه استفاده شده است.

نمونه‌های صوتی شامل ۲ نمونه ی صوتی با گفتار مرد، ۲ نمونه ی صوتی با گفتار زن، ۲ نمونه ی صوتی با گفتار بچه، ۲ نمونه ی صوتی بی کلام و ۲ نمونه ی صوتی مختلط می‌باشد. در جدول (۱) نتایج آزمون فشرده‌سازی MP3 برای مقادیر متفاوت سطح آستانه آورده شده است.

برابر با ۵۷، الف-۲ برابر با ۲،۵ و الف-۳ برابر با ۱۱ حاصل شد. همانطور که مشاهده می‌شود، در صورتی که به طور یک در میان به نمونه‌های سیگنال اصلی متوسط دو نمونه مجاور اضافه گردد، درصد خطا به نحو چشمگیری افزایش می‌یابد. این روش، حمله موفقی خواهد بود در صورتی که نرخ بیت نمونه برداری سیگنال صوتی به دو برابر افزایش یابد که با این وجود کیفیت سیگنال اصلی تغییر چندانی نخواهد کرد. اما به دلیل این که دیکد داده‌های پیام مخفی با نیمی از داده‌های اصلی درج‌شده صورت می‌پذیرد، درصد خطا به طور قابل ملاحظه ای افزایش می‌یابد. در صورتی که نرخ بیت سیگنال صوتی بدون تغییر در نظر گرفته شود با توجه به یک درمیان اضافه کردن نمونه‌ها می‌توان این نتیجه را گرفت که با این کار سیگنال صوتی را به طور کل تخریب نموده ایم. در این دو حالت به میزان صد درصد به سیگنال، داده افزوده می‌شود. و همانطور که BER نشان می‌دهد، آشکارسازی پیام مخفی کاملاً تصادفی می‌شود. به عبارت دیگر سیگنال مخفی کاملاً تخریب می‌شود.

ب) اضافه نمودن یک بلوک داده تصادفی در میان بلوک‌های سیگنال صوتی: در این روش یک بلوک با داده‌های تصادفی در میان داده‌های سیگنال صوتی اضافه می‌گردد. با توجه به این نکته که الگوریتم پیشنهادی بر پایه بلوک‌های ۱۶ نمونه ای از سیگنال صوتی استوار است، بنابراین بلوک‌های نمونه درج‌شده را نیز با همین طول در نظر می‌گیریم. تعداد بلوک‌های در نظر گرفته‌شده در این قسمت برابر ۱۰ درصد تعداد کل بلوک‌های ۱۶ تایی موجو در سیگنال می‌باشد. همچنین دامنه نمونه‌های موجود درون بلوک‌های تصادفی به صورت نویز سفید گوسی انتخاب شده است. درصد نرخ خطای بیت اضافه نمودن یک بلوک داده تصادفی در میان بلوک‌های سیگنال صوتی در روش بلوک تصادفی نویز گوسی با طول ۱۶ نمونه، برابر با ۲ به دست آمد. همان‌طور که از نتایج مشهود است، روش پیشنهادی در برابر اضافه‌کردن نویز افزایشی به صورت بلوک مقاوم به نظر می‌رسد. علت این امر بدین خاطر می‌باشد که بلوک‌های اضافه‌شده به صورت نویز گوسی فاقد ویژگی بلوک‌هایی است که الگوریتم پیشنهادی بر روی این بلوک‌ها عملیات نهان‌نگاری را پیاده‌سازی می‌کند. به عبارت دیگر نویزهای گوسی شکل قله مانند خود را تاحدی حفظ می‌کنند.

با توجه به الگوریتم درخت صفر و نتایج آزمایشات، چنین به نظر می‌رسد که بلوک‌هایی مساعد نهان‌نگاری می‌باشند که دارای شیب تقریباً ثابتی باشند و در ابتدا یا انتهای بلوک دارای پیک کوچکی باشند. از اینرو هرچند بلوک نویز گوسی که سیگنال صوتی اضافه شود، تاثیر چندانی بر الگوریتم درج پیام مخفی پیشنهادی نخواهد داشت.

ج) اضافه نمودن بلوک‌های سکوت در میان پیچ‌های صوتی

جدول (۱): نتایج آزمون فشرده‌سازی MP3 برای مقادیر متفاوت سطح آستانه

نرخ خطای بیت (%)			سیگنال
۶۴kbps	۹۶kbps	۱۲۸kbps	نرخ نمونه برداری سطح آستانه
۲۴,۱	۲۲,۴	۱۵,۳	۰,۵
۲۰,۳	۲۱,۸	۹,۸	۰,۷
۲۶,۶	۱۸,۹	۶,۶	۰,۹
۱۹,۷	۱۷,۱	۷,۲	۱,۱

$$CD = \frac{\sum_{m=1}^M w(m)d(c_x, c_y, m)}{\sum_{m=1}^M w(m)} \quad (۴)$$

$$d(c_x, c_y, m) = \left[ [c_x(0) - c_y(0)]^2 + 2 \sum_{k=1}^L [c_x(k) - c_y(k)]^2 \right]^{\frac{1}{2}}$$

همان‌طور که در رابطه (۴) مشاهده می‌شود، آنالیز کیفیت صوت به روش CDM، برای هر قطعه از صوت (فریم) به‌طور جداگانه در نظر گرفته می‌شود. یعنی صوت را به فریم‌های با طول دلخواه تقسیم می‌کنند و فاصله اقلیدسی بین فریم نهان‌نگاری شده و اصلی را محاسبه می‌کنند همچنین به دلیل این که انرژی هر فریم (بلندی) متفاوت می‌باشد، هر فریم را در متوسط انرژی آن که به‌صورت  $w$  در فرمول فوق نمایش داده شده است، ضرب می‌کنند. در اینجا به دلیل آن که فایل‌های صوتی از قبل دسته‌بندی شده‌اند و طول هر فایل صوتی کوتاه در نظر گرفته شده است، نیازی به فریم‌بندی فایل صوتی نداریم. علاوه بر این برای درک محسوس تر و همچنین مقایسه بین فایل‌های صوتی متفاوت باید معیارهای به دست آمده را نرمالیزه کرد. به‌منظور این هدف معیار CDM به دست آمده را بر طول هر فایل صوتی تقسیم کرده و میانگین تعداد فایل‌های مشابه از یک نوع را محاسبه می‌کنیم تا در مقایسه بین فایل‌های صوتی مختلف، دید بهتری به ما دهد. نتایج بررسی این روش در جدول (۲) آورده شده است.

### ۳-۳-۳- بررسی الگوریتم پیشنهادی از لحاظ آماری

از آنجایی که صوت به دسته‌های گوناگونی طبقه‌بندی می‌شود، بنابراین برای بررسی جامع تر الگوریتم پیشنهادی ۲۰ قطعه صوتی در نظر گرفته شد. این مجموعه شامل ۵ قطعه صوتی آهنگ (بی کلام)، ۵ قطعه صوتی گفتار زن، ۵ قطعه صوتی گفتار مرد و ۵ قطعه موسیقی آمیخته گفتار و آهنگ می‌باشد. در ادامه تمامی روش‌های آزمون بر روی این مجموعه اعمال می‌گردد.

### ○ بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس

مقیاس فاصله کپستروم CDM: همانطور که می‌دانیم این معیار بیانگر میزان همواری طیف سیگنال می‌باشد. در رابطه (۴) ضرایب کپستروم سیگنال اصلی را  $C_x(k)$  و ضرایب کپستروم سیگنال نهان‌نگاری شده را  $C_y(k)$  مینامیم.  $d(c_x, c_y, m)$  فاصله بین ضرایب کپستروم فریم  $m$  نامیده می‌شود.

جدول (۲): نتایج بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم

معیار CDM				شکل موج	دنباله پیام مخفی نوع صوت
یکی در میان یک و صفر	تماماً صفر	تماماً یک			
$7.321 \times 10^{-4}$	$6.431 \times 10^{-4}$	$4.542 \times 10^{-4}$		موزیک با تن بالا	
$2.456 \times 10^{-5}$	$7.338 \times 10^{-5}$	$4.554 \times 10^{-5}$		موزیک با تن ملایم	
$7.22 \times 10^{-7}$	$5.331 \times 10^{-7}$	$7.342 \times 10^{-7}$		گفتار زن	
$7.342 \times 10^{-7}$	$6.324 \times 10^{-7}$	$5.331 \times 10^{-7}$		گفتار مرد	
$4.342 \times 10^{-4}$	$1.324 \times 10^{-3}$	$1.442 \times 10^{-4}$		موزیک تن بالا با گفتار مرد	
$4.231 \times 10^{-4}$	$1.008 \times 10^{-2}$	$1.349 \times 10^{-4}$		موزیک تن پایین با گفتار مرد	
$3.837 \times 10^{-7}$	$1.541 \times 10^{-4}$	$5.221 \times 10^{-7}$		موزیک تن بالا با گفتار زن	
$8.874 \times 10^{-7}$	$9.534 \times 10^{-7}$	$8.324 \times 10^{-7}$		موزیک تن پایین با گفتار زن	
$7.612 \times 10^{-4}$	$8.333 \times 10^{-7}$	$2.635 \times 10^{-7}$		مختلط	

همان طور که از متوسط معیار محاسبه شده در جدول (۲) به نظر می‌رسد، فایل‌های صوتی به فرم موزیک با تن (بلندی) ملایم کم‌ترین تغییرات را در معیار همواری صوت دارد. این امر بدین خاطر است که موزیک با تن ملایم همواری ملایمی دارد. فاصله ایتاکورا- سایتو ISD و فاصله COSH: معیار ISD فاصله بین طیف توان سیگنال نهان‌نگاری شده و سیگنال اصلی را محاسبه می‌کند و COSH نسخه‌ی مقترنی از فاصله ISD می‌باشد. فرمول آن‌ها

$$\text{COSH} = \int_{-\pi}^{\pi} \left[ \frac{1}{2} \left( \frac{Y(w)}{X(w)} + \frac{X(w)}{Y(w)} \right) - 1 \right] \frac{dw}{2\pi}$$

$$\text{IS} = \int_{-\pi}^{\pi} \left( \log \frac{Y(w)}{X(w)} + \frac{X(w)}{Y(w)} - 1 \right) \frac{dw}{2\pi} \quad (5)$$

جدول (۳): نتایج بررسی فاصله ایتاکورا - سایتو

ISD		COSH		معیار
یکی در میان یک و صفر		تماماً صفر	تماماً یک	دنباله پیام مخفی نوع صوت
-۳,۰۱۴	-۲,۵۳۱	-۲,۳۶۱	-۲,۳۶۱	موزیک با تن بالا
-۱,۲۳۳	-۲,۹۸۸	-۳,۵۶۸	-۳,۵۶۸	
-۱۰,۳۴۴	۶,۲۳۱	۶,۲۴۷	۶,۲۴۷	موزیک با تن ملایم
-۸,۳۴۲	-۳,۲۲۵	۷,۴۳۲	۷,۴۳۲	
۱۵,۵۶۷	-۱,۲۲۵	-۱,۲۴۵	-۱,۲۴۵	گفتار زن
۸,۳۳۲	-۱,۸۸۶	۰,۱۱۱	۰,۱۱۱	
۲۰,۱۳۴	-۲۱,۷۸۸	-۲۳,۲۲۷	-۲۳,۲۲۷	گفتار مرد
۷,۰۰۱	-۲۳,۸۴۴	۱۲,۶۳۵	۱۲,۶۳۵	
-۴,۵۲۱	-۱,۳۴۹	-۱,۴۶۳	-۱,۴۶۳	موزیک تن بالا با گفتار مرد
-۱,۳۲۴	-۱,۲۳۱	-۱,۱۶۵	-۱,۱۶۵	
-۱۷,۰۲۳	۱۲,۱۴۵	۱۳,۲۷۴	۱۳,۲۷۴	موزیک تن پایین با گفتار مرد
-۳,۲۲۳	۷,۳۲۱	۷,۲۴۱	۷,۲۴۱	
-۱۵,۳۳۶	-۳,۱۹۵	-۴,۵۴۹	-۴,۵۴۹	موزیک تن بالا با گفتار زن
-۶,۲۴۲	-۱,۴۵۹	-۵,۳۲۸	-۵,۳۲۸	
-۱,۴۳۸	-۱,۴۳۴	-۱,۵۱۹	-۱,۵۱۹	موزیک تن پایین با گفتار زن
-۴,۵۴۸	-۱,۶۵۶	-۱,۵۵۸	-۱,۵۵۸	
۷,۷۵۲	-۷,۲۲	-۵,۳۲۶	-۵,۳۲۶	مختلط
۵,۹۵۳	-۲,۹۸۲	-۳,۶۵۲	-۳,۶۵۲	

تبدیل رادون یک سیگنال در واقع تصویر کردن سیگنال مورد نظر در زاویه مورد نظر می باشد. بنابراین این معیار فاصله میزان تشابه تبدیل فوریه سیگنال را با چرخش های متفاوت آن در زوایای

جدول (۴): نتایج بررسی معیار تبدیل فوریه- رادون زمان کوتاه

۹۰ درجه		۴۵ درجه		معیار
یکی در میان یک و صفر	تماما صفر	تماما یک	دنباله پیام مخفی نوع صوت	
۰,۰۱۱۳	۰,۰۱۰۸	۰,۰۱۰۱	موزیک با تن بالا	
۲,۹۷۳*۱۰ <sup>-۳</sup>	۱,۹۷۱*۱۰ <sup>-۳</sup>	۳,۹۴۰*۱۰ <sup>-۳</sup>		
۹,۹۲۸۴*۱۰ <sup>-۳</sup>	۰,۰۱۲۲	۰,۰۱۵۴	موزیک با تن ملایم	
۰,۰۰۲۹	۰,۰۱۱۷	۰,۰۱۱۳		
۸,۰۰۷۵*۱۰ <sup>-۴</sup>	۹,۵۹۹۹*۱۰ <sup>-۴</sup>	۰,۰۱۸	گفتار زن	
۷,۸۲۲۷*۱۰ <sup>-۴</sup>	۲,۰۹۷۳*۱۰ <sup>-۴</sup>	۷,۴۸۲۲*۱۰ <sup>-۳</sup>		
۱,۷۷۰۲*۱۰ <sup>-۱</sup>	۸,۱۷۸۸*۱۰ <sup>-۱</sup>	۰,۰۰۳۲	گفتار مرد	
۶,۹۹۲۱*۱۰ <sup>-۳</sup>	۵,۰۰۴۸*۱۰ <sup>-۳</sup>	۰,۰۲۳۴		
۴,۹۹۲*۱۰ <sup>-۳</sup>	۸,۲۲۳۸*۱۰ <sup>-۳</sup>	۳,۸۳۸۶*۱۰ <sup>-۳</sup>	موزیک تن بالا با گفتار مرد	
۷,۳۴۹*۱۰ <sup>-۳</sup>	۳,۳۲۲۶*۱۰ <sup>-۳</sup>	۷,۰۷*۱۰ <sup>-۳</sup>		
۱,۴۳۵۰*۱۰ <sup>-۳</sup>	۹,۲۳۱۴*۱۰ <sup>-۳</sup>	۰,۰۱۲۳	موزیک تن پایین با گفتار مرد	
۳,۱۹۹۷*۱۰ <sup>-۳</sup>	۱,۶۰۱۶*۱۰ <sup>-۳</sup>	۰,۱۰۰۲		
۰,۰۰۱۲	۶,۱۹۲*۱۰ <sup>-۳</sup>	۰,۸۶۴*۱۰ <sup>-۳</sup>	موزیک تن بالا با گفتار زن	
۹,۰۳۴*۱۰ <sup>-۳</sup>	۳,۲۹۱*۱۰ <sup>-۳</sup>	۳,۲۱۹*۱۰ <sup>-۳</sup>		
۲,۳۴۱*۱۰ <sup>-۳</sup>	۶,۳۲۹*۱۰ <sup>-۳</sup>	۰,۰۰۲۴	موزیک تن پایین با گفتار زن	
۴,۴۶۹*۱۰ <sup>-۳</sup>	۵,۸۶*۱۰ <sup>-۳</sup>	۰,۰۱۰۳		
۶,۲۵۵۳*۱۰ <sup>-۴</sup>	۷,۵۴۲۶*۱۰ <sup>-۴</sup>	۹,۵۸۳۰*۱۰ <sup>-۴</sup>	مختلط	
۹,۳۴۲*۱۰ <sup>-۳</sup>	۳,۳۳۱*۱۰ <sup>-۳</sup>	۹,۳۳۲*۱۰ <sup>-۳</sup>		

نمودار آستانه شنوایی سطح آستانه به ازای فرکانس های مختلف میزانی متفاوت دارد. بنابراین در صورتی که میانگین تبدیل رادون یک سیگنال به صفر نزدیک باشد به مفهوم این است که سیگنال مورد آزمون از همواری نسبتا یکنواختی برخوردار است. در مثال جدول (۷) با توجه با این که تمامی ضرایب در یک رنج و نزدیک به مقدار صفر می باشند، می توان این نتیجه را گرفت که افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد.

○ بررسی آماری الگوریتم پیشنهادی در حوزه ادراک

معیار BSD و MBSD: طبق بحث های قبل، معیار BSD بر پایه محاسبه اختلاف دامنه بین دو سیگنال در حوزه طیف بارک<sup>۱</sup> می باشد. بر اساس مدل فیزیوآکوستیک<sup>۲</sup>، طیف فرکانسی صوت به

با توجه به خصوصیت الگوریتم تبدیل رادون و این که تبدیل رادون سیگنال داده شده را به فضایی با زاویه داده شده تبدیل می کند، نمونه های جدید به فضایی با بردار واحد وابسته نگاشته می شوند. به عبارت دیگر در این فضای جدید می توان میزان پراکندگی نمونه ها را صرف نظر از وابستگی بردارهای پایه نسبت به یکدیگر به دست آورد. با توجه به در نظر گرفتن دو زاویه ۴۵ درجه و ۹۰ درجه برای تبدیل رادون، بدین صورت می توان در مورد سیگنال های مورد آزمایش بحث نمود که با این آزمون می توان میزان پراکندگی سیگنال را حول دامنه صفر به دست آورد. در ادامه با توجه به این که سیگنال استفاده شده در تبدیل رادون، ضرایب تبدیل فوریه سیگنال صوتی می باشد، مفهوم این آزمون به معنی اندازه گیری میزان فرکانس های سیگنال با دامنه کوچک می باشد. این معیار تا حدی مشابه معیار اندازه گیری تن های زیر سطح آستانه شنوایی می باشد با این تفاوت که در

<sup>۱</sup> طیف بارک یک تقسیم بندی فرکانسی است که با توجه به مدل مطلق شنوایی محدوده فرکانسی صوت را به نواحی بحرانی مختلفی تقسیم می نماید.

<sup>۲</sup> Physioacoustic



بلندی) با توجه به باند فرکانسی متناظر با آن دارای سطح تاثیر شنوایی متفاوت می‌باشند. بنابراین برای هر زیر باند وزن خاصی با توجه به طیف فرکانسی متناظر با آن نسبت می‌دهند. وزن هر زیر باند با توجه به سطح آستانه بارک و سطح دامنه سیگنال تعیین می‌گردد. رابطه محاسبه معیار MBSD طبق رابطه (۷) است:

$$MBSD = \sum_{k=1}^C M(i)D_{xy}(i) \quad (7)$$

که  $M(i)$  وزن‌های اختصاص داده شده به هر زیر باند است. نتایج این دو روش در جدول (۵) آورده شده است.

نواحی یا زیر باندهایی تقسیم می‌شود که سطح بلندی شنوایی در آن نواحی متفاوت می‌باشد. برای محاسبه معیار BSD، طیف فرکانسی صوت مورد نظر را متناظر با طیف بارک، ناحیه‌بندی کرده و سپس فاصله اقلیدسی میان مجموعه مولفه‌های فرکانسی سیگنال اصلی با سیگنال نهان‌نگاری شده در آن نواحی محاسبه می‌شود. فرمول محاسبه آن طبق رابطه (۶) است:

$$D_{xy} = |S_x(i) - S_y(i)| \quad (6)$$

که  $S$  طیف توان زیر باند مورد نظر می‌باشد.

تنها تفاوت معیار MBSD با BSD در نسبت دادن وزن‌های متفاوت به هر زیر باند می‌باشد. با توجه به این موضوع که هر تن

جدول (۵): نتایج بررسی معیار MBSD و BSD

MBSD		BSD		معیار
یکی در میان یک و صفر	تماماً صفر	تماماً یک	دنباله پیام مخفی نوع صوت	
۰,۰۹۵۲	۰,۱۰۳۵	۰,۲۵۲۱	موزیک با تن بالا	
۰,۰۲۷۱	۰,۰۵۰۶	۰,۲۰۵۷	موزیک با تن ملایم	
۰,۱۰۲۱	۰,۲۹۰۱	۰,۱۲۲۹	گفتار زن	
۰,۰۲۴۳	۰,۰۱۷۷	۰,۱۰۹۰	گفتار مرد	
۰,۱۱۹۸	۰,۱۱۱۳	۰,۱۳۵۹	موزیک تن بالا با گفتار مرد	
۰,۰۰۳۲	۰,۰۹۵۰	۰,۰۹۳۲	موزیک تن پایین با گفتار مرد	
۰,۳۷۲۱	۰,۴۳۲۲	۰,۴۶۱۸	موزیک تن بالا با گفتار زن	
۰,۰۹۹	۰,۰۲۳۹	۰,۰۲۷	موزیک تن پایین با گفتار زن	
۰,۰۲۴	۰,۰۲۸	۰,۰۲۴	مختلط	
۰,۰۰۲	۰,۰۸۷	۰,۰۱۵		
۰,۰۱۳	۰,۰۶۱	۰,۰۵۱		
۰,۰۱۳	۰,۰۱۴	۰,۰۱۱		
۰,۰۹۸	۰,۱۷۶	۰,۳۴۶		
۰,۰۱۰	۰,۰۱۱	۰,۰۱۶		
۰,۱۱۹	۰,۲۲۸	۰,۱۶۳		
۰,۰۶۴	۰,۰۱۳	۰,۰۸۲		
۰,۰۲۲	۰,۱۱۴۳	۰,۰۶۷		
۰,۰۲۹	۰,۱۹۲	۰,۲۹۰		

گرفت که درج پیام مخفی درون سیگنال حامل علاوه بر این که در حوزه فرکانس اغتشاش وارد نمی‌کند، بلکه خاصیت سطح آستانه شنوایی صوت را نیز بهم نمی‌زند. همانطور که از جدول بالا پیداست الگوریتم پیشنهاد شده طیف فرکانسی سیگنال صوتی را چندان تغییر نمی‌دهد و همچنین از خاصیت سطح آستانه

آزمون BSD و MBSD مشابه آزمون اندازه‌گیری STFRT، اندازه‌گیری میزان اغتشاش سیگنال در حوزه فرکانس می‌باشد. با این تفاوت که در معیار BSD یا MBSD طیف فرکانسی از فیلتر سطح آستانه شنوایی عبور داده می‌شود. در صورتی که معیار BSD یا MBSD به مقدار صفر نزدیک باشد می‌توان این نتیجه را

حوزه زمان نیز دارد. کمترین SNR مربوط به موزیک تن بالا با گفتار زن می باشد که دارای SNR ی در حدود 13 db می باشد. دلیل این نتیجه می تواند به علت اغتشاش شدید زمانی سیگنال مربوطه باشد.

فاصله سزنا-کوشی CZD: این معیار که فرمول آن در رابطه (۹) آورده شده است، یک همبستگی بر پایه استاندارد متریک است که به طور مستقیم بردار نمونه ها را در حوزه ی زمان مقایسه می کند. این معیار، میزان شباهت بین دو سیگنال را اندازه گیری می کند.

$$C = \frac{1}{K} \sum_{i=0}^{K-1} \left( 1 - \frac{2\min(x(i), y(i))}{x(i) + y(i)} \right) \quad (9)$$

نتایج این آزمون در جدول (۷) آورده شده است.

جدول (۷): نتایج بررسی فاصله سزنا-کوشی

معیار CZD			
یکی در میان یک و صفر	تماما صفر	تماما یک	دنباله پیام مخفی نوع صوت
-۰,۰۰۲۴	-۱,۱۴*۱۰ <sup>-۳</sup>	-۲,۲۲*۱۰ <sup>-۳</sup>	موزیک با تن بالا
۰,۱۶۳۴	-۰,۰۵۲۸	۰,۶۱۹۲	موزیک با تن ملایم
-۰,۴۶۲۱	۰,۰۳۴۳	۰,۷۵۴۲	گفتار زن
۰,۱۰۳۲	۰,۰۱۵۴	-۰,۰۰۹۵	گفتار مرد
-۰,۰۱۸۳۰	۰,۰۹۷۳	-۸,۳۴*۱۰ <sup>-۳</sup>	موزیک تن بالا با گفتار مرد
۰,۱۹۳۹	-۰,۱۲۴۸	۰,۱۰۲۲	موزیک تن پایین با گفتار مرد
۰,۰۲۷۷	۰,۰۲۱۴	۹,۲۷۳۱*۱۰ <sup>-۴</sup>	موزیک تن بالا با گفتار زن
-۰,۰۹۸۵	-۰,۰۱۳۶	-۰,۰۶۵۷	موزیک تن پایین با گفتار زن
۱,۴۳*۱۰ <sup>-۴</sup>	-۸,۳۷۷*۱۰ <sup>-۴</sup>	۰,۰۲۲۳	مختلط

در این معیار در صورتی که دو سیگنال مورد آزمون یکسان باشند، مقدار این معیار برابر با صفر خواهد شد. این معیار به بررسی تک تک عناصر دو سیگنال مورد آزمون نسبت به یکدیگر می پردازد. همانطور که از رابطه بالا پیداست در صورتی که مولفه های سیگنال نهان نگاری شده اختلاف کمی با مولفه های سیگنال اصلی داشته باشند، میزان این معیار به سمت صفر میل می کند. با توجه به جدول بالا می توان دریافت که الگوریتم پیشنهادی در حوزه زمان اغتشاش چندانی را نیز بر سیگنال صوتی متحمل نمی کند. از نتایج حاصله در جدول بالا می توان دریافت که با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت. زیرا در سیگنال گفتار زن همواری خاصی محسوس است. بنابراین با درج پیام مخفی در این نوع سیگنال صوتی، تا حدودی این فرم یکنواختی از بین خواهد رفت و با توجه به مقایسه شدن مولفه به مولفه دو سیگنال اصلی و سیگنال حاوی پیام مخفی، میزان معیار CZD افزایش می یابد.

شنوایی پیروی می کند. علت این امر نیز واضح می باشد زیرا الگوریتم پیشنهادی بر روی سیگنال صوتی اعمال شده است که حاصل خروجی تبدیل MP3 می باشد. بنابراین سیگنال صوتی مبرا از داده هایی است که درون سطح آستانه شنوایی قرار می گیرند. همچنین با توجه به جدول بالا موزیک تن بالا با گفتار مرد دارای بهترین نتایج می باشند. زیرا با توجه به ساختار فرکانسی این نوع سیگنال ها، تن هایی با دامنه کوچک و یا غیر تن کم تر در طیف سیگنالی آن ها موجود می باشد. بنابراین از هر جهت مساعد با ساختار طیف بارک می باشند.

○ معیارهای حوزه زمان

نسبت سیگنال به نویز SNR: معیار SNR معروفترین مقیاس اندازه گیری اغتشاش در حوزه ی زمان می باشد که سیگنال اصلی و سیگنال نهان نگاری شده را نمونه به نمونه با یکدیگر مقایسه می کند. این معیار حاوی اطلاعاتی در مورد اغتشاشات جمع شونده بر روی سیگنال ایستادن می باشد. رابطه محاسبه SNR طبق رابطه (۸) است.

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2} \quad (8)$$

نتایج این روش در جدول (۶) آورده شده است.

جدول (۶): نتایج بررسی نسبت سیگنال به نویز

معیار SNR			
یکی در میان یک و صفر	تماما صفر	تماما یک	دنباله پیام مخفی نوع صوت
۳۰,۰۰۱۰	۲۰,۱۹۰۰	۳۲,۲۰۱۰	موزیک با تن بالا
۱۹,۹۶۴۲	۱۹,۳۱۴۳	۱۸,۷۲۰۵	موزیک با تن ملایم
۳۰,۱۷۸۲	۳۰,۶۵۰۱	۳۰,۳۲۹۱	گفتار زن
۱۳,۰۲۲۲	۱۱,۹۳۸۸	۱۵,۸۱۲۰	گفتار مرد
۱۷,۳۴۲۱	۱۴,۹۵۲۶	۱۸,۱۲۱۱	موزیک تن بالا با گفتار مرد
۱۵,۳۰۷۵	۱۶,۹۲۳۱	۱۷,۲۳۲۰	موزیک تن پایین با گفتار مرد
۱۴,۹۲۷۳	۱۴,۳۴۷۲	۱۳,۳۸۳۹	موزیک تن بالا با گفتار زن
۲۱,۳۲۶۵	۲۲,۸۸۲۴	۱۹,۶۲۰۵	موزیک تن پایین با گفتار زن
۱۴,۹۹۳۸	۱۶,۸۸۷۲	۱۵,۲۳۷۷	مختلط

با توجه به رابطه SNR که در بالا ذکر شده است، مفهوم SNR بدین معنی است در صورتی که سیگنال نهان نگاری شده شدیداً در حوزه زمان مورد اغتشاش واقع شود فاصله آن تا سیگنال اصلی بسیار زیاد خواهد شد. بنابراین مخرج کسر رابطه SNR دچار رشد فزاینده ای می شود. در این صورت مقدار SNR به سمت صفر و حتی منفی بی نهایت میل می کند. بنابراین هرچه در این معیار میزان SNR بیشتر باشد، دلیل بر تطابق بیشتر سیگنال اصلی با سیگنال نهان نگاری شده دارد. با توجه به نتایج جدول بالا این نتیجه را می توان گرفت که الگوریتم پیشنهادی نتایج مطلوبی در

and Systems, Hiroshima, Japan, pp. 473-480, 1996. doi: 10.1109/MMCS.1996.535015.

[3] Cvejic, Nedeljko, "Algorithms For Audio Watermarking and Steganography, 2004.

[4] S. S. Agaian, D. Akopian, O. Caglayan, and S. A. D'Souza, "Lossless Adaptive Digital Audio Steganography," Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, 2005., Pacific Grove, CA, USA, pp. 903-906, 2005. doi: 10.1109/ACSSC.2005.1599886.

[5] W. C. CHU, "Speech Coding Algorithms-Foundation and Evaluation of Standardized Coder, 2003.

[6] J. Herre and S. Dick, "Psychoacoustic Models for Perceptual Audio Coding—A Tutorial Review," Applied Sciences, vol. 9, no. 14, p. 2854, Jul. 2019, doi: 10.3390/app9142854.

[7] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," in IEEE Transactions on Signal Processing, vol. 41, no. 12, pp. 3445-3462, Dec. 1993, doi: 10.1109/78.258085.

[8] V. Ralph Algazi and Robert R. Estes Jr. "Analysis-based coding of image transform and subband coefficients", Proc. SPIE 2564, Applications of Digital Image Processing XVIII, 22 August 1995.

[9] K. L. Narasihimhasprasad, M. V. Nagabhushanam, V. V. Satyanarayana Tallapragada, and J. Krishna Sunkara, "Embedded Zero-Tree Wavelet Coding with Selective Decomposition Bands," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, pp. 0445-0449, 2019. doi: 10.1109/ICCSP.2019.8698062.

[10] J. Antonio Alvarez-Cedillo, T. Alvarez-Sanchez, M. Aguilar-Fernandez, and J. Sandoval-Gutierrez, 'Many-Core Algorithm of the Embedded Zerotree Wavelet Encoder', Coding Theory. IntechOpen, Mar. 11, 2020. doi: 10.5772/intechopen.89300.

[11] Rezik, Siwar et al. "Speech steganography using wavelet and Fourier transforms." EURASIP Journal on Audio, Speech, and Music Processing, pp. 1-14, 2012.

[12] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura, "A digital watermark based on the wavelet transform and its robustness on image compression," Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269), Chicago, IL, USA, vol. 2, pp. 391-395, 1998. doi: 10.1109/ICIP.1998.723388.

[13] Z. Xu, K. Wang, and X. -h. Qiao, "Digital Audio Watermarking Algorithm Based on Quantizing Coefficients," 2006 International Conference on Intelligent Information Hiding and Multimedia, Pasadena, CA, USA, 2006, pp. 41-46, doi: 10.1109/IIH-MSP.2006.265115.

[14] Xiang, Shijun, Hyoung Joong Kim, and Jiwu Huang. "Audio watermarking robust against time-scale modification and MP3 compression." Signal Processing 88.10, pp. 2372-2387, 2008.

[15] M. Baziyad, et al., "Maximizing embedding capacity for speech steganography: a segment-growing approach," Multimedia Tools and Applications, vol. 80, pp. 24469-24490, 2021.

[16] F. J. Farsana and K. Gopakumar, "Speech encryption algorithm based on nonorthogonal quantum state with hyperchaotic keystreams." Advances in Mathematical Physics 2020, pp. 1-12, 2020.

[17] A. Kuznetsov, et al., "Direct Spread Spectrum Technology for Data Hiding in Audio,." Sensors 22.9, 3115, 2022.

#### ۴- نتیجه گیری

الگوریتم نهان نگاری بایستی مناسب طراحی شود و ضمن هنر حفظ سیگنال پوشانه، در صورت حمله، سیگنال مخفی آن نیز آشکار نشود. در الگوریتم پیشنهادی این موضوعات وجود دارد و به طور کلی نتایج و ویژگی های زیر در آن به دست آمده است:

- پایداری الگوریتم پیشنهادی، در برابر نویز سفید نرمال بیشتر از نویز یکنواخت است و در برابر SNR های بالاتر از 10db نرخ بیت خطای کم تر از ۱ بیت دارد.
- در صورت حمله به الگوریتم پیشنهادی، با توجه به BER به دست آمده، آشکارسازی پیام مخفی کاملاً تصادفی شده و سیگنال مخفی کاملاً تخریب می شود.
- همچنین روش پیشنهادی در برابر اضافه کردن نویز افزایشی به صورت بلوک مقاوم به نظر می رسد و مستعد درج پیام مخفی با استفاده از روش الگوریتم پیشنهادی نخواهند بود.
- نتایج بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم همان طور که از متوسط معیار محاسبه شده در نتایج به دست آمده، فایل های صوتی به فرم موزیک با تن (بلندی) ملایم کم ترین تغییرات را در معیار همواری صوت دارد.
- در الگوریتم پیشنهادی، افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد.
- الگوریتم پیشنهاد شده طیف فرکانسی سیگنال صوتی را چندان تغییر نمی دهد و همچنین از خاصیت سطح آستانه شنوایی پیروی می کند.
- همچنین با توجه به نتایج حاصل شده، موزیک تن بالا با گفتار مرد دارای بهترین نتایج می باشند بنابراین از هر جهت مساعد با ساختار طیف بارک می باشند.
- الگوریتم پیشنهادی با توجه به جداول (۶) و (۷)، نتایج مطلوبی در حوزه زمان دارد.
- کم ترین SNR مربوط به موزیک تن بالا با گفتار زن می باشد که دارای SNRی در حدود 13db می باشد. دلیل این نتیجه می تواند به علت اغتشاش شدید زمانی سیگنال مربوطه باشد. با توجه به نتایج حاصل شده، با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت.

#### ۵- مراجع

- [1] C. Kratzer, J. Dittmann, T. Vogel and R. Hillert, "Design and evaluation of steganography for voice-over-IP," 2006 IEEE International Symposium on Circuits and Systems, Kos, Greece, p. 4, 2006. doi: 10.1109/ISCAS.2006.1693105.
- [2] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," Proceedings of the Third IEEE International Conference on Multimedia Computing

## Steganography Audio Based on Zero-Tree Wavelet Transform Algorithm

R. Esfahani\*, Ahmadreza Matinfar

### Abstract

Audio steganography is very important, equally steganography in other media (image, video, etc.). In this paper is presented steganography of audio based on embedded zero-tree wavelet transform algorithm. Improving the resistance against white noise and additive noise with the lowest SNR is one of the important topics in steganography. The proposed algorithm is more robust against normal white noise than uniform noise and has a bit error rate of less than 1 bit against SNRs higher than 10db. According to the obtained BER, if the proposed algorithm is attacked, the hidden signal is lost completely. Also, the proposed method is resistant against additive noise. The proposed algorithm has the least changes in the sound smoothness criterion in frequency domain with Capstrom distance scale and audio files in the form of music with soft tone (loudness), and the increase of secret message does not have much effect on creating disturbances in the frequency domain. The proposed algorithm of the frequency spectrum does not change the audio signal much, and it also follows the property of the hearing threshold level, and high-pitched music with male speech has the best results, so it is favorable to the spectrum structure of Bark. Also, the proposed algorithm has favorable results in the time domain. The lowest SNR is related to high-pitched music with female speech, which has an SNR of about 13db. According to the obtained results, we will have the worst case of embedding a secret message by choosing the audio signal with female speech. Because there is a certain smoothness in the female speech signal. Therefore, this uniform state will be lost to some extent by embedding a secret message in this type of audio signal, and the CZD criterion will increase according to the component-by-component comparison of the two main signals and the signal containing the secret message.

**Key Words:** *Embedded Zero Tree Wavelet Transform (EZW), White Noise, Additive Error, BSD and MBSD Criteria, CDM Kapstrom Distance Scale, ISD and COSH Distance, Short Time Radon Transform STFRT, SNR and CZD*

---

\*Assistant Professor- Imam Hossein Comprehensive University - (resfahani@ihu.ac.ir) - writer - in - change