

Transparent and Robust Audio Watermarking Using Synergy LU Decomposition and the Fibonacci Sequence in GBT-DCT-DWT Transforms

M. Mosleh , S. Setayeshi* , B. Barekatin, M. Mosleh

*Amirkabir University of Technology

(Received: 15/04/2020, Accepted: 05/08/2020)

ABSTRACT

Digital watermarking technology is presented as one of the best solutions for solving unauthorized copying, content identification and authentication of digital media. Digital watermarking can be applied in the fields of image, text, audio, and video contents. Audio watermarking has recently attracted the attention of researchers because the human auditory system is much more sensitive than his vision system. Therefore, insertion of the data into audio signals in a transparent way is much more difficult than other watermarking species. An efficient audio watermarking system should be able to improve reconciliation of the three measures of transparency, strength and capacitance. Improving the compromise between these three measures is a challenging problem, due to the fact that increasing the capacity of the input signal, causes distortions in the signal which in turn causes a reduction in transparency and robustness. In this paper, a new method is proposed for watermarking of audio signals that is able to improve the three criteria of transparency, robustness and capacity in an appropriate manner. In order to improve the resilience of the hidden signal against signal processing attacks, the synergy of three efficient transforms in the field of signal processing namely, the graph-centric conversion (GBT), the discrete cosine transform (DCT) and the discrete wavelet transform (DWT), has been used. In addition, in order to maintain the transparency of the watermarked signal, the watermarked data in the high triangle matrix resulting from the LU decomposition of the approximation coefficients of the discrete wavelet are entered in an order which is based on the Fibonacci sequence. The procedure is based on the value of the watermarked bit, the non-zero values of the upper triangular matrix are replaced by the nearest even or odd numbers in the Fibonacci sequence. The latent extraction operation is completely blind. The results of the evaluation of the proposed method on audio files with Blue, Electronic, Classic and Jazz styles show that the proposed method, despite good resistance to various signal processing attacks, has an average signal rate of 45.13 dB and a placement rate of 625.75 bits per second.

Keywords: Copyright protection, Audio watermarking, Graph-based Transform(GBT), Discrete Cosine Transform (DCT), Discrete wavelet transform (DWT) LU decomposition.

* Corresponding Author Email: setayesh@aut.ac.ir

علمی - پژوهشی

نهان نگاری صوتی شفاف و مقاوم با استفاده از هم‌افزایی تجزیه LU و دنباله فیبوناچی در بستر تبدیلات GBT-DCT-DWT

مهدی مصلح^۱، سعید ستایشی^{۲*}، بهرنگ برکتین^۳، محمد مصلح^۴

۱- دانشجوی دکتری، دانشکده مهندسی کامپیوتر، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران، ۲- دانشیار، دانشکده مهندسی انرژی و

فیزیک، دانشگاه صنعتی امیرکبیر، تهران، ایران، ۳- استادیار، مرکز تحقیقات مه داده، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران،

۴- دانشیار، گروه مهندسی کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران

(دریافت: ۱۳۹۹/۰۱/۲۷، پذیرش: ۱۳۹۹/۰۵/۱۵)

چکیده

فناوری نهان نگاری دیجیتال به عنوان یکی از بهترین راه‌حل‌ها جهت حل مسئله کپی غیرمجاز، شناسایی محتوا و احراز هویت رسانه‌های دیجیتال ارائه شده است. نهان نگاری دیجیتال می‌تواند در حوزه تصویر، متن، صوت ویدئو به کار گرفته شود. نهان نگاری صوت، توجه بسیاری از محققین را به خود جلب کرده است چرا که سامانه شنوایی انسان بسیار حساس‌تر از سامانه بینایی می‌باشد. بنابراین، درج داده‌ها در سیگنال‌های صوتی به شیوه‌ای شفاف، بسیار سخت‌تر از سایر گونه‌های نهان نگاری است. سامانه نهان نگاری صوت کارآمد می‌بایست قادر باشد سه معیار شفافیت، مقاوم‌پذیری و ظرفیت جاده‌ی را بهبود بخشد. بهبود هم‌زمان این سه معیار به عنوان مساله‌ای چالش برانگیز در حوزه نهان نگاری است چرا که افزایش ظرفیت جاده‌ی باعث افزایش اعوجاج سیگنال شده و این امر موجب کاهش شفافیت و مقاوم‌پذیری می‌گردد. در این مقاله رویکردی جدید به منظور نهان نگاری صوت پیشنهاد می‌گردد که قادر است سه معیار مقاوم‌پذیری، شفافیت و ظرفیت را به نحو مناسبی بهبود بخشد. به منظور بهبود مقاوم‌پذیری سیگنال نهان نگاره شده در برابر حملات پردازش سیگنال از هم‌افزایی سه تبدیل کارآمد در حوزه پردازش سیگنال به نام‌های تبدیل گراف محور (GBT)، تبدیل کسینوسی گسسته و تبدیل موجک گسسته (DWT) استفاده شده است. علاوه بر این، به منظور حفظ شفافیت سیگنال نهان نگاری شده، داده‌های نهان نگاره در ماتریس بالا مثلثی حاصل از تجزیه LU ضرایب تقریب تبدیل موجک گسسته، بر اساس دنباله فیبوناچی درج می‌شوند. رویه کار بدین صورت است که بر اساس مقدار بیت نهان نگاره، مقادیر غیرصفر ماتریس بالا مثلثی با نزدیک‌ترین اعداد زوج و یا فرد در دنباله فیبوناچی جایگزین می‌شوند. عملیات استخراج نهان نگاره به صورت کاملاً کور انجام می‌شود. نتایج حاصل از ارزیابی روش پیشنهادی بر روی فایل‌های صوتی با سبک‌های Electronic، Blue، Classic و Jazz نشان می‌دهند که روش ارائه شده علیرغم مقاوم‌پذیری مناسب در برابر حملات مختلف پردازش سیگنال، به طور متوسط دارای نرخ سیگنال به نویز ۴۵/۱۳ دسی‌بل و نرخ جاده‌ی ۶۲۵/۷۵ بیت بر ثانیه است.

کلیدواژه‌ها: حق کپی‌رایت، نهان نگاری صوت، تبدیل گراف محور (GBT)، تبدیل کسینوسی گسسته (DCT)، تبدیل موجک گسسته (DWT)، تجزیه LU

۱- مقدمه

(به عنوان مثال دوربین‌های دیجیتال^۱، سی‌دی‌های قابل حمل، دی‌وی‌دی^۲، رایانه‌های قابل حمل^۳ و رایانه شخصی^۴) این امکان را برای کاربران در سراسر جهان ایجاد کرده است که بتوانند داده‌های چندرسانه‌ای خود را به صورت ساده و ارزان برای دیگران به اشتراک بگذارند [۱-۳]. در چنین شرایطی به کارگیری روش‌های جهت تضمین امنیت داده‌ها و نیز حفاظت از حق نشر

رشد سریع اینترنت و استفاده از اطلاعات به صورت دیجیتال باعث تغییرات مهمی در جامعه، اقتصاد جهان و همچنین راه‌های ارتباطی مردم شده است. شبکه‌های ارتباطی و اطلاعات چندرسانه‌ای در قالب‌های دیجیتال (عکس، صوت، تصویر)، ضمن ایجاد فرصت‌هایی مناسب برای پیشرفت و نوآوری، چالش‌های متعددی را نیز پیش روی محققین قرار داده است. استفاده از نرم‌افزارهای مختلف و کاهش قیمت دستگاه‌های دیجیتال

¹ Digital camera

² Digital video disc

³ Laptap

⁴ Personal computer

* رایانامه نویسنده پاسخگو: Setayesh@aut.ac.ir

در این مقاله، روشی جدید جهت نهان‌نگاری داده‌های صوتی ارائه می‌شود که قادر است با هم‌افزایی سه تبدیل قدرتمند GBT، DWT و DCT معیار مقاوم‌پذیری در برابر حملات پردازش سیگنال را به نحو قابل توجهی بهبود بخشد. علاوه بر این، به‌منظور بهبود معیار شفافیت، داده‌های نهان‌نگاره داخل ماتریس بالا مثلثی حاصل از تجزیه LU ضرایب DWT بر اساس دنباله فیبوناچی درج می‌شوند. این عمل موجب می‌گردد کمترین تغییرات در سیگنال صوتی رخ دهد و در نتیجه شفافیت سیگنال نهان‌نگاره افزایش یابد. نتایج حاصله نشان می‌دهند که به‌کارگیری هم‌زمان روش‌های اشاره‌شده در روش نهان‌نگاری پیشنهادشده موجب بهبود مصالحه بین سه پارامتر شفافیت، مقاوم‌پذیری و ظرفیت جاده می‌شود.

ادامه این مقاله به‌صورت زیر سازمان‌دهی شده است. در بخش دوم به مرور کارهای پیشین در زمینه نهان‌نگاری صوت پرداخته می‌شود. در بخش سوم تبدیلات ریاضی GBT، DCT و DWT معرفی می‌شوند. در بخش چهارم، روش نهان‌نگاری صوت پیشنهادی تشریح می‌گردد. در بخش پنجم به ارائه نتایج حاصل از ارزیابی کارایی روش پیشنهادی پرداخته می‌شود. در نهایت مقاله به بخش نتیجه‌گیری به پایان می‌رسد.

۲- مروری بر روش‌های نهان‌نگاری مقاوم پیشین

تاکنون روش‌های متعددی جهت نهان‌نگاری مقاوم صوت ارائه‌شده است که در ادامه به برخی از مهم‌ترین آن‌ها اشاره خواهیم کرد.

Hsu و Hu به معرفی طرحی نهان‌نگاری صوتی شفاف، مقاوم و با ظرفیت بالا در حوزه DCT پرداختند [۱۰]. داده‌های نهان‌نگاره با استفاده از دست‌کاری ضرایب DCT انجام شده است. آن‌ها با حفظ توازن در باندهای فرکانسی توانستند داده‌های نهان‌نگاره را به‌صورت کور بازیابی نمایند.

Pourhashemi و همکاران روشی نهان‌نگاری صوتی مقاوم را با استفاده از تبدیل لوکاس و تبدیل FFT ارائه نمودند [۱۱]. ایده اصلی در این روش جادهی داده‌های نهان‌نگاره در انرژی زیر فریم‌ها با استفاده از توالی ریاضی لوکاس می‌باشد که قابلیت ذخیره کردن دو بیت در هر فریم را داراست. استخراج دنباله نهان‌نگار به‌صورت کور انجام می‌گیرد. نتایج شبیه‌سازی نشان می‌دهد که ظرفیت بالا و شفافیت و مقاوم‌پذیری نسبتاً قابل قبولی دارد.

Lei و همکاران به ارائه طرحی نهان‌نگاری صوتی مقاوم بر مبنای تبدیل SVD و الگوریتم تکامل تفاضلی (DE)^۴ با استفاده از الگوریتم کوانتیزاسیون DM^۵ پرداختند [۱۲]. در این روش برای

آن‌ها ضروری است. برای این منظور روش‌های رمزنگاری^۱ و نهان‌نگاری^۲ به‌صورت مکمل یکدیگر توسعه یافتند. روش‌های رمزنگاری قادرند با استفاده از رویکردهای خاصی داده‌ها را به‌گونه‌ای کدگذاری نمایند که طرف‌گیرنده تنها با داشتن کلید کدگشایی قادر است اطلاعات را رمزگشایی نماید. در این میان اگرچه به‌کارگیری روش‌های رمزنگاری جهت انتقال امن اطلاعات به کار می‌روند ولیکن هیچ تضمینی برای تأمین امنیت اطلاعات پس از رمزگشایی و در زمان استفاده ارائه نمی‌دهند [۴]. برای رفع این چالش رویکردهای نهان‌نگاری پیشنهاد شدند که قادرند با درج اطلاعات نامحسوس در رسانه دیجیتال حفاظت از آن را در برابر کپی غیرمجاز، ادعای مالکیت نادرست و پخش غیرمجاز تأمین نمایند. نهان‌نگاری به شرکت‌ها قدرت به اجرا در آوردن حق انحصار را می‌دهد و می‌تواند در حوزه تجارت به شدت مورد استفاده قرار بگیرد. وارد کردن نهان‌نگاره نباید در شفافیت سیگنال تأثیر مخرب داشته و همچنین نهان‌نگاره نباید در اثر عملیات مختلف از بین برود. تقریباً تمامی مشکل طراحی نهان‌نگاری به خاطر دو مورد بالاست. بدیهی است که در بسیاری از موارد نیازهای خاص، کاربردهای خاصی را برای نهان‌نگاری به وجود می‌آورد و طراحی خاص می‌طلبد [۶-۵].

روش‌های نهان‌نگاری را می‌توان از نظر ماهیت و کاربرد به دو دسته کلی مقاوم و شکننده تقسیم‌بندی نمود [۸-۷]. در روش‌های نهان‌نگاری مقاوم، الگوریتم استخراج می‌بایست قادر باشد با وجود حملات خرابکارانه، داده نهان‌نگاره را به‌درستی بازیابی نماید. در نهان‌نگاری شکننده، الگوریتم استخراج قادر است هرگونه تغییری را در سیگنال نهان‌نگاری شده تشخیص دهد. از شکننده بودن بیشتر برای مقاصد تعیین اعتبار^۳ استفاده می‌شود، به این صورت که اگر الگوی شکننده به‌درستی از سیگنال الگوگذاری شده استخراج شد، می‌توان نتیجه گرفت که اطلاعات دست‌کاری نشده‌اند [۹].

در یک سامانه نهان‌نگاری سه معیار شفافیت، مقاوم‌پذیری و ظرفیت جادهی با یکدیگر در تقابل هستند که رسیدن به یک مصالحه قابل قبول بین آن‌ها همواره به‌عنوان مسئله‌ای اساسی مطرح می‌باشد.

نهان‌نگاری صوتی یکی از پرچالش‌ترین گونه‌های نهان‌نگاری است چرا که سامانه شنوایی انسان بسیار حساس‌تر از سامانه بینایی‌اش می‌باشد. تاکنون رویکردهای مختلفی جهت نهان‌نگاری صوت ارائه‌شده است که سعی کرده‌اند مصالحه بین سه پارامتر شفافیت، مقاوم‌پذیری و ظرفیت را بهبود بخشند. در بخش دوم به برخی از مهم‌ترین روش‌های نهان‌نگاری صوتی پرداخته خواهد شد.

¹ Cryptography

² Steganography

³ Authentication

⁴ Differential evolution (DE)

⁵ dither modulation (DM)

چندین بیت نهان‌نگاره درون بخشی از سیگنال ظرفیت جاده‌ی را افزایش دهند.

Hsu و Hu به معرفی دو طرح نهان‌نگاری صوتی کور بر اساس مدلاسیون^۶ RDM پرداختند [۱۸]. روش ارائه‌شده در زیر باند تقریب سطح چهارم تبدیل موجک گسسته عمل می‌نماید. طرح معرفی‌شده اول از فیلتری پایین‌گذر FIR برای پوشش دادن نویز استفاده می‌کند در حالی که طرح ارائه‌شده دوم به‌منظور جبران اعوجاج ناشی از مدلاسیون بردار استفاده می‌شود.

Jeyhoon و همکاران الگوریتمی نهان‌نگاری صوت کور پیشنهاد دادند که داده‌های نهان‌نگاره را بر اساس دست‌کاری ضرایب DCT جاده‌ی می‌کنند [۱۹]. برای این منظور باند فرکانسی DCT انتخاب‌شده به مجموعه‌ای از فریم‌های کوچک تقسیم شده و داده‌های نهان‌نگاره در هر یک از فریم‌ها درج می‌شوند. به‌منظور استخراج بیت‌های نهان‌نگاره از رگرسیون خطی و انحراف معیار استفاده شده است.

Mosleh و همکاران به معرفی طرحی نهان‌نگاری صوتی هوشمند با استفاده از هم‌افزایی تبدیل SVD و ماشین بردار پشتیبان^۷ (SVM) پرداختند [۲۰]. عملیات درج بر اساس دست‌کاری مقادیر در تبدیل SVD انجام می‌شود. به‌منظور استخراج داده‌های نهان‌نگاره از تشخیص‌گری یادگیر بر اساس ماشین بردار پشتیبان استفاده شده است.

Li و همکاران به معرفی طرحی نهان‌نگاری صوتی بر اساس مدلاسیون طیف گسترده پرداختند [۲۱]. بر خلاف طرح‌های نهان‌نگاری طیف گسترده مرسوم که از همبستگی ساده جهت تشخیص نهان‌نگاره استفاده می‌کنند، در این طرح، پیش از محاسبه‌ی پارامتر همبستگی، مشخصه‌های ادراکی سیگنال صوتی نهان‌نگاری شده استخراج می‌شود. استخراج داده نهان‌نگاره به‌صورت کاملاً کور انجام شده است و نیازی به سیگنال اصلی نمی‌باشد.

Bartic و همکاران به معرفی طرحی نهان‌نگاری صوتی دیجیتال بر مبنای آنالیز زمانی-فرکانسی پرداختند [۲۲]. در این طرح، دامنه زمانی-فرکانسی مناسب برای درج نهان‌نگاره انتخاب می‌شود. فرآیند استخراج دنباله نهان‌نگاره بر مبنای گرادیان است.

Mosenfar و همکاران به معرفی طرحی نهان‌نگاری صوتی هوشمند بر مبنای تجزیه QR و الگوریتم ژنتیک^۸ پرداختند [۲۳].

تجزیه سیگنال میزبان و یافتن ضرایب تخمینی علاوه بر تبدیل DCT، جهت کاهش انرژی از تبدیلات^۱ LWT و DWT استفاده شده است. برای به‌دست آوردن مقادیر منفرد و افزایش مقاومت طرح از SVD استفاده شده است. برای کوانتیزه کردن مقادیر منفرد و درج در نهان‌نگاره از کوانتیزاسیون DM استفاده شده است. همچنین از الگوریتم بهینه‌سازی DE برای حل تصادم بین پارامترهای مقاومت و ادراک‌ناپذیری استفاده شده است.

Yuan و همکاران به معرفی طرحی نهان‌نگاری صوتی بر مبنای تشخیص ویژگی ضرایب کپسترال فرکانس مل (MFCC)^۲ و تبدیل موجک پیچیده دو درختی (DTCWT)^۳ پرداختند [۱۳]. برای استخراج سگمنت‌های ویژگی که باید پس از حملات مخرب به سیگنال میزبان، جابجا شوند از بخش تشخیص ویژگی ضرایب کپسترال فرکانس مل و برای درج نهان‌نگاره در ضرایب پایین‌گذر از بخش تبدیل موجک پیچیده‌ی دو درختی با استفاده از روش‌های طیف گسترده استفاده شده است.

Hu و همکاران به معرفی طرحی نهان‌نگاری صوتی کور بدون هیچ گونه اعوجاج قابل درک پرداختند [۱۴]. در این روش از تبدیل بسته موجک گسسته^۴ (DWPT) و تبدیل DCT استفاده شده است. از روش مادولاسیون شاخص کوانتیزاسیون^۵ (QIM) بر مبنای ادراک برای توازن بین پارامترهای مقاومت و غیر قابل ادراک بودن و برای درج بیت‌های نهان‌نگاره استفاده شده است. از الگوریتم شبکه عصبی پس انتشار نیز برای یافتن بهترین مکان برای درج نهان‌نگاره، استفاده شده است.

Al-Hajj به معرفی الگوریتمی نهان‌نگاری صوتی دیجیتال غیر کور پرداخت [۱۵]. در این طرح از تبدیل DWT و روش SVD استفاده شده

است. برای دستیابی به مقاومت بالا از طریق تجزیه و پخش بیت‌های نهان‌نگاره از DWT و برای دستیابی به ادراک‌ناپذیری قابل قبول از طریق درج بیت‌های نهان‌نگاره در مقادیر منفرد قطری ماتریس S از SVD استفاده شده است.

Huang و همکاران به ارائه یک طرح نهان‌نگاری صوت کور با استفاده از دست‌کاری دامنه فرکانس‌های پایین تبدیل موجک گسسته پرداختند [۱۶]. برای تغییر دامنه فرکانس‌های پایین با در نظر گرفتن بهترین میزان شفافیت، سامانه ارائه‌شده سعی می‌کند اختلاف بین ضرایب اصلی و درج شده را حداقل نماید.

Xiang و همکاران به ارائه روشی نهان‌نگاری صوت پر ظرفیت بر اساس طیف گسترده پرداختند [۱۷]. آن‌ها سعی کردند با درج

^۶ Rational dither modulation (RDM)

^۷ Support vector machine (SVM)

^۸ Genetic Algorithm (GA)

^۱ Lifting wavelet transform (LWT)

^۲ Mel-Frequency Cepstral coefficients

^۳ Dual-tree complex wavelet transform

^۴ Discrete wavelet packet transform (DWPT)

^۵ Quantization Index Modulation (QIM)

سپس می‌توان از A و D ماتریس گراف-لاپلاسیان را به صورت رابطه (۳) محاسبه کرد [۲۸]:

$$L = D - A \quad (3)$$

ماتریس L ماتریسی متقارن از مقادیر حقیقی خواهد بود و بر اساس تئوری طیفی، تجزیه مقادیر ویژه (EVD) این ماتریس دارای مجموعه‌ای از مقادیر ویژه حقیقی غیرمنفی خواهد بود که با $\Lambda = \{\lambda_1, \dots, \lambda_N\}$ و مجموعه‌ای از بردارهای ویژه مستقل از هم و متناظر با مقادیر ویژه که با $V = \{v_1, \dots, v_N\}$ مشخص می‌شود، خواهد بود [۲۸].

$$LL = V\Lambda V^T \quad (4)$$

سپس می‌توانیم از این بردارهای ویژه متعامد استفاده کنیم تا همبستگی فریم صوتی ورودی را کاهش دهیم [۲۸].

$$c = V^T s \quad (5)$$

به طوری که c ماتریس ضرایب تبدیل گراف محور است و s فریم صوت ورودی است.

۳-۲- تبدیل کسینوسی گسسته

تبدیل کسینوسی گسسته روشی برای برگرداندن سیگنال درون مؤلفه‌های فرکانسی مقدماتی و ابتدایی است. DCT می‌تواند به صورت یک بعدی برای سیگنال‌هایی مانند صوت و به صورت دوبعدی برای سیگنال‌هایی چون تصویر به کار رود. تبدیل کسینوسی گسسته، تبدیل طیفی‌ای است که خواص تبدیل فوریه گسسته را دارا می‌باشد. با استفاده از این تبدیل تنها توابع کسینوسی مقدار موج گوناگون از توابع پایه است و روی سیگنال‌های مقداردهی شده حقیقی و ضرایب طیفی اعمال می‌گردد. از سیگنالی یک بعدی می‌توان تبدیل کسینوسی گسسته گرفت و سپس با گرفتن معکوس این تبدیل ضرایب اولیه را بازیابی کرد. تبدیل کسینوسی گسسته سیگنالی را درون مؤلفه‌های فرکانس پایین و بالا تقسیم می‌کند. بیشتر سیگنال ورودی در مؤلفه‌های فرکانس پایین است که DC نامیده می‌شود، تجمع می‌یابد. اما توان کمی در مؤلفه‌های فرکانس بالا یا مؤلفه‌های AC باقی می‌ماند. بازیابی سیگنال توسط معکوس تبدیل کسینوسی گسسته صورت می‌گیرد. تغییرات در مؤلفه‌های AC تأثیر کمی بر روی سیگنال بازیابی شده دارد اما اصلاح و تغییر در مؤلفه DC یا مؤلفه‌های فرکانس پایین می‌تواند نقش مهمی در بازیابی سیگنال اصلی داشته باشد. بنابراین، اطلاعات ذخیره‌سازی شده در مؤلفه AC می‌تواند بر روی مقاومت سیگنال تأثیرگذار باشد به نحوی که ممکن است یک پیام رمز را حذف کند [۲۹].

در این روش عملیات درج بر اساس دست‌کاری مقادیر تجزیه QR فریم‌ها و یافتن بهترین مکان جهت درج داده‌های نهان‌نگاره با استفاده از الگوریتم ژنتیک انجام می‌شود.

Mosleh و همکاران روشی نهان‌نگاری صوت مقاوم، شفاف و پرمزفیت در بستر تبدیل DCT بر اساس الگوریتم ژنتیک پرداختند [۲۴]. در این طرح الگوریتم ژنتیک قادر است با یافتن طول فریم مناسب و بهترین محدوده فرکانسی جهت درج داده‌ها، مصالحه معیارهای شفافیت، مقاوم‌پذیری و ظرفیت را بهبود بخشد.

Attari و همکاران به ارائه روشی نهان‌نگاری صوتی کور در حوزه تبدیل موجک پرداختند [۲۵]. این روش داده‌های نهان‌نگاره را در ضرایب فرکانس پایین تبدیل موجک که سامانه شنوایی انسان حساسیت کمتری دارد، بر اساس دنباله فیبوناچی درج می‌نماید.

Karajeh و Magableh به ارائه روشی نهان‌نگاری صوتی بر اساس تبدیل DCT و تجزیه Schur پرداختند [۲۶]. آنها از تبدیل DCT به منظور افزایش مقاوم‌پذیری و از تبدیل Schur برای بهبود شفافیت استفاده کردند.

Dutta و Kaur به ارائه روشی نهان‌نگاری صوت پرمزفیت مبتنی بر هم‌افزایی الگوریتم ژنتیک و تجزیه LU پرداختند [۲۷]. الگوریتم ژنتیک جهت یافتن تعداد بهینه نمونه‌های صوتی که برای درج هر بیت نهان‌نگاره مورد نیاز است، استفاده شده است. به منظور افزایش شفافیت، درج داده‌ها در تجزیه LU ضرایب تبدیل موجک انجام شده است.

۳- مبانی تحقیق

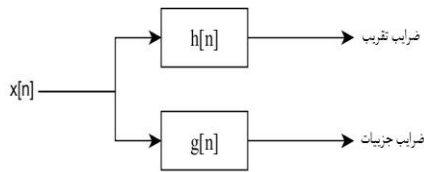
۳-۱- تبدیل گراف محور

با داشتن فریمی با N نمونه، می‌توان گرافی $G = \{V, E, s\}$ تعریف کرد به طوری که V و E گره‌ها و یال‌های گراف هستند و s سیگنال صوتی است. برای این گراف، عنصرهای ماتریس مجاورت به صورت رابطه (۱) تعیین می‌گردد [۲۸]:

$$A_{ij} = \begin{cases} a_{ij} & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

که a_{ij} وزن یال‌های بین گره‌ها i و j در گراف است. ماتریس درجه D نیز ماتریسی قطری است که به صورت رابطه (۲) تعریف می‌شود [۲۸]:

$$D_{ij} = \begin{cases} \sum a_{ij} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$



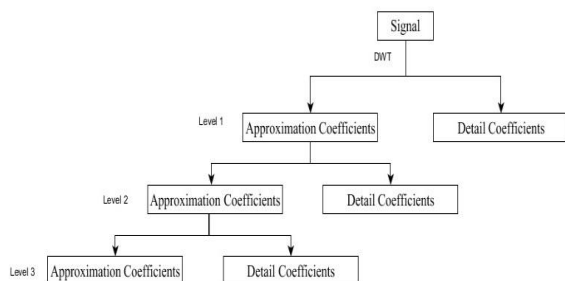
شکل (۱): فیلتر بالاگذر و پایین‌گذر در تبدیل موجک گسسته

در نتیجه خروجی اولین مرحله از اعمال تبدیل موجک، دو نسخه، یکی بالاگذر و دیگری پایین‌گذر، با طول کاهش‌یافته (نصف شده) از سیگنال اولیه به‌صورت روابط (۸ و ۹) به‌دست می‌آیند [۳۰]:

$$y_{high}[k] = \sum_n x[n] \cdot g[2k - n] \quad (۸)$$

$$y_{low}[k] = \sum_n x[n] \cdot h[2k - n] \quad (۹)$$

در روابط (۸ و ۹) y_{low} و y_{high} به ترتیب خروجی فیلتر بالاگذر g و فیلتر پایین‌گذر h می‌باشند. تعداد ضرایب به‌دست‌آمده با تعداد نقاط ورودی برابر است. انواع مختلف توابع موجک مطرح شده است. این یک سطح از تبدیل موجک گسسته بر روی سیگنال ورودی است. می‌توان موجک گسسته انجام داد و این عمل را می‌توان تا رسیدن به سطح مورد نظر ادامه داد. این روند در شکل (۲) نشان داده شده است. در این شکل تجزیه سه سطحی یک سیگنال را می‌بینید.



شکل (۲): تجزیه سه سطحی سیگنالی با تبدیل موجک گسسته

سیگنال اصلی با تبدیل معکوس تبدیل موجک گسسته به‌دست می‌آید. در رابطه (۱۰) این تبدیل معکوس آورده شده است [۳۰]:

$$\begin{aligned} S &= A1 + D1 \\ &= A2 + D2 + D1 \\ &= A3 + D3 + D2 + D1 \end{aligned} \quad (۱۰)$$

با افزایش تعداد مراحل تبدیل، میزان جزئیات نیز کاهش می‌یابد.

۴-۳- تجزیه LU

می‌توان نشان داد که هر ماتریس مربعی A می‌تواند به‌صورت ضرب ماتریسی پایین مثلثی L و ماتریسی بالا مثلثی U بیان شود،

تبدیل کسینوسی گسسته می‌تواند سیگنالی صوتی را به‌صورت مجموعه‌ای از توابع کسینوسی در فرکانس‌های مختلف نمایش دهد. فرض کنید $x[n] = x(t) | t = nT$ سیگنالی نمونه‌گیری شده با دوره نمونه‌برداری T و طول n را نشان می‌دهد.

DCT این سیگنال $x[n]$ به‌صورت رابطه (۶) نمایش داده می‌شود [۲۹]:

$$y[n] = w[k] \sum_{n=0}^{N-1} x[n] \cos\left(\frac{k}{2N}(2n+1)\right) \quad (۶)$$

$$k = 0, 1, \dots, N-1$$

$$w[k] = \begin{cases} 1/\sqrt{N} & k = 0 \\ \sqrt{2/N} & 1 \leq k \leq N-1 \end{cases} \quad (۷)$$

یکی از ویژگی‌های تبدیل DCT این است که تراکم انرژی در چند نمونه می‌باشد. این ویژگی باعث کاهش اعوجاج در سیگنال صوتی می‌گردد که در نهان‌نگاری‌های صوتی بسیار مفید می‌باشد.

۳-۳- تبدیل موجک گسسته

تبدیل موجک پیوسته ابزار مناسبی برای تجزیه و تحلیل زمان-فرکانس است اما بار محاسباتی زیادی دارد به همین دلیل از تبدیل موجک گسسته استفاده می‌شود. این تبدیل سیگنال را از فیلتری پایین‌گذر عبور می‌دهد، و سیگنال تقریب را تشکیل می‌دهد. سپس سیگنال را از فیلتر بالاگذر عبور می‌دهد و سیگنال جزئیات به‌دست می‌آید. سیگنال تقریب شبیه سیگنال اصلی است با این تفاوت که دارای تغییرات کمتری است چون مقداری از این تغییرات به علت عبور سیگنال از فیلتر پایین‌گذر از بین رفته است. طول سیگنال‌های تقریب و جزئیات نصف طول سیگنال اصلی است.

در تبدیل موجک گسسته همان‌طور که در شکل (۱) نشان داده شده است ابتدا سیگنال از فیلتری دیجیتال پایین‌گذر نیم باند با پاسخ ضربه $h[n]$ عبور می‌کند. در نتیجه این عمل فیلترینگ، تمام مؤلفه‌های فرکانسی که بیشتر از نصف بزرگ‌ترین فرکانس موجود در سیگنال باشند حذف می‌شوند. از آنجا که بیشترین فرکانس موجود در سیگنال خروجی فیلتر برابر است با $\pi/2$ رادیان، نیمی از نمونه‌ها قابل حذف‌اند. لذا با حذف یکی در میان نمونه‌ها، طول سیگنال نصف خواهد شد بدون اینکه اطلاعاتی را از دست داده باشیم. روند مشابهی نیز با استفاده از فیلتری دیجیتال بالاگذر نیم باند با پاسخ ضربه $g[n]$ انجام می‌پذیرد.

مرحله اول: خواندن سیگنال صوت ورودی S و تصویر نهان نگاره W و تبدیل تصویر نهان نگاره به دنباله‌ای دودویی

مرحله دوم: فریم‌بندی سیگنال ورودی به مجموعه‌ای از فریم‌های غیر همپوشان (Frames)

مرحله سوم: به‌کارگیری تبدیل GBT بر روی هر فریم با استفاده از ساختاری مناسب گرافیکی.

مرحله چهارم: اعمال تبدیل DCT بر روی k ضرایب اولیه GBT که دارای بیشترین مقادیر می‌باشند.

مرحله پنجم: به‌کارگیری تبدیل موجک دو سطحی هار بر روی ضرایب فرکانس پایین حاصل از تبدیل DCT

مرحله ششم: تبدیل بردار ضرایب تقریب سطح دوم به فرم ماتریس دوبعدی

مرحله هفتم: محاسبه تجزیه LU بر روی ماتریس دوبعدی مرحله قبل به دو ماتریس پایین مثلثی (L) و بالا مثلثی (U)

مرحله هشتم: درج بیت‌های دنباله نهان نگاره در نمونه‌های غیرصفر

ماتریس بالا مثلثی U.

اگر بیت نهان نگاره دارای مقدار صفر باشد، مقادیر غیرصفر ماتریس U به نزدیک‌ترین اعداد زوج و در غیر این صورت به نزدیک‌ترین اعداد فرد در دنباله فیبوناچی تبدیل می‌شوند.

برای مثال اگر ماتریس دوبعدی X به صورت زیر باشد:

$$X = \begin{bmatrix} 79 & -40 \\ -154 & 113 \end{bmatrix}$$

آن‌گاه با محاسبه تجزیه LU ماتریس دوبعدی X خواهیم داشت:

$$LU(X) = \begin{bmatrix} 0.5130 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -154.73 & 113 \\ 0 & 17.96 \end{bmatrix}$$

$$L = \begin{bmatrix} 0.5130 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } U = \begin{bmatrix} -154.73 & 113 \\ 0 & 17.96 \end{bmatrix}$$

اگر فرض شود که مقدار بیت نهان نگاره برابر مقدار یک است، بنابراین، ماتریس بالا مثلثی U به صورت زیر تغییر می‌کند:

$$U' = \begin{bmatrix} -89 & 89 \\ 0 & 21 \end{bmatrix}$$

به طوری که دنباله فیبوناچی به صورت زیر می‌باشد:

Fibo sequence

$$= \{1. 1. 2. 3. 5. 8. 13. 21. 34. 55. 89. 144. 233. \dots\}$$

مرحله نهم: بازیابی ماتریس دوبعدی اصلاح شده X' از طریق حاصل ضرب ماتریس پایین مثلثی I و ماتریس بالامثلثی اصلاح شده U'

مرحله دهم: تبدیل ماتریس دوبعدی X' به فرم برداری

به عبارتی: $A=LU$ فرآیند محاسبه L و U برای ماتریسی معلوم A به عنوان تجزیه LU یا فاکتورگیری LU شناخته می‌شود. تجزیه LU عملیاتی با مقدار مشخص نیست، به عبارت دیگر بدون اعمال قیود خاص بر روی L و U، بی‌نهایت حالت برای ماتریس‌های بالا مثلثی و پایین مثلثی ایجاد شده وجود دارد. این قیود عملی نوعی تجزیه را از نوع دیگر متمایز می‌کند. [۳۱].

$$a_{ij} = \sum_{k=0}^{N-1} l_{ik} u_{kj}, \quad i, j = 0, 1, \dots, N-1 \quad (11)$$

ماتریس L و ماتریس U به صورت روابط (۱۲) و (۱۳) محاسبه می‌گردد [۳۱]:

$$u_{ij} = a_{ij} - \sum_{k=0}^{i-1} l_{ik} u_{kj}, \quad \forall i \leq j \quad (12)$$

$$l_{ij} = \frac{1}{u_{jj}} \left(m_{ij} - \sum_{k=0}^{j-1} l_{ik} u_{kj} \right), \quad \forall i > j \quad (13)$$

۴- روش پیشنهادی نهان نگاری صوتی

به دلیل حساسیت بالای سامانه شنوایی انسان، نهان نگاری اطلاعات در سیگنال‌های صوتی نسبت به سایر گونه‌های نهان نگاری از پیچیدگی بالاتری برخوردار است. همان طوری که اشاره شد، تاکنون روش‌های متعددی به منظور نهان نگاری صوت ارائه شده است که در بسیاری از آن‌ها همواره سعی شده است با توجه به کاربرد مورد نظر، یکی از معیارهای شفافیت، مقاوم‌پذیری و یا ظرفیت بهبود یافته و سایر معیارها حداقل مقادیر ممکن را داشته باشند. در این مقاله سعی شده است با الهام از مزایای روش‌های پیشین و اجتناب از معایب آن‌ها روشی نهان نگاری صوت جدید ارائه گردد. مقاوم‌پذیری رویکرد پیشنهادی در برابر حملات پردازش سیگنال از طریق هم‌افزایی تبدیلات DCT، GBT و DWT فراهم می‌شود. علاوه بر این به منظور تضمین شفافیت سیگنال نهان نگاری شده از تجزیه LU و دنباله فیبوناچی بهره گرفته می‌شود. در ادامه جزئیات روش پیشنهادی ارائه می‌گردد.

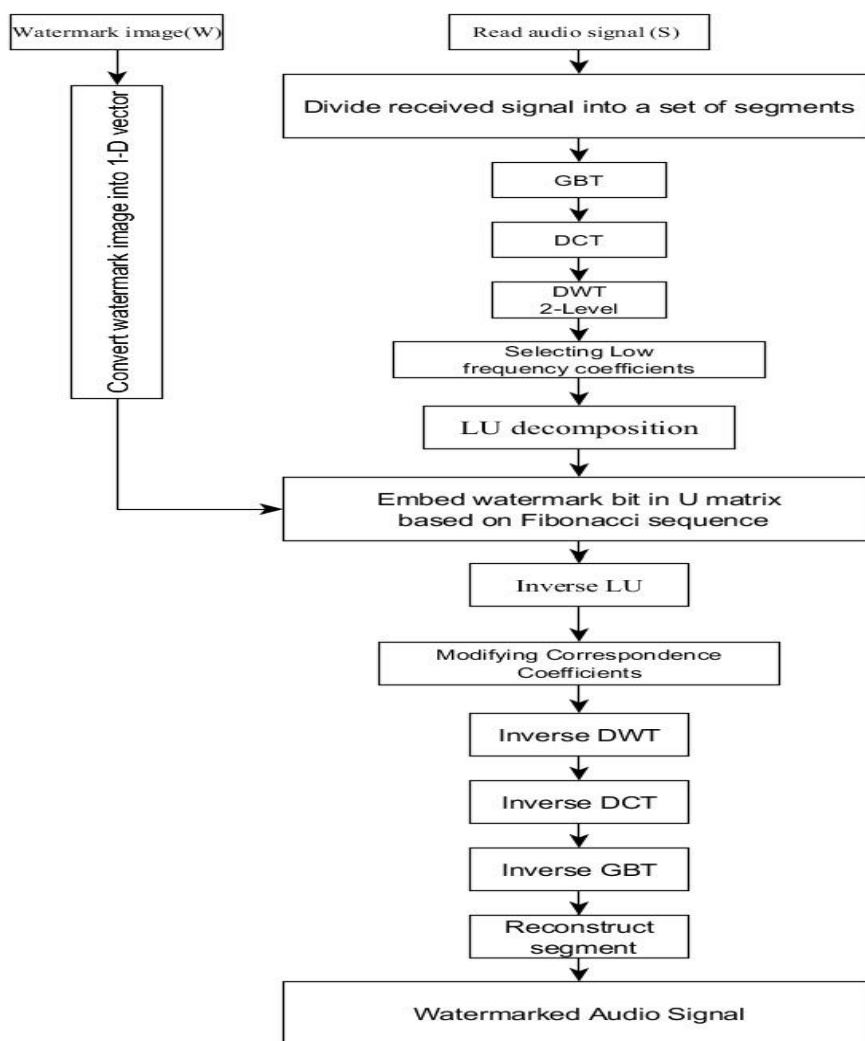
۴-۱- روش جاسازی پیشنهادی

به منظور جاسازی دنباله نهان نگاری، در ابتدا سیگنال میزبان به مجموعه‌ای از فریم‌های غیرهمپوشان تقسیم می‌شود. در ادامه به منظور افزایش مقاوم‌پذیری طرح پیشنهادی سه تبدیل GBT، DCT و DWT به صورت متوالی بر روی هر بلوک اعمال می‌گردد. در نهایت، بیت‌های نهان نگاره در ماتریس بالامثلثی حاصل از تجزیه LU ضرایب موجک، بر اساس دنباله فیبوناچی، درج می‌شوند. شکل (۳) دیاگرام بلوکی روش درج پیشنهادی را نشان می‌دهد. علاوه بر این، جزئیات روش درج پیشنهادی در ادامه ارائه شده است.

مرحله سیزدهم: بازیابی سیگنال نهان‌نگاری شده از طریق فریم‌های اصلاح‌شده.

مرحله یازدهم: محاسبه وارون تبدیل موجک دو سطحی بر روی بردار اصلاح‌شده ضرایب تقریب

مرحله دوازدهم: محاسبه وارون تبدیلات DCT و GBT و به‌دست آوردن فریم نهان‌نگاری شده



شکل (۳): فرآیند جاسازی نهان‌نگاره در روش پیشنهادی

۲-۴- روش استخراج پیشنهادی

در شکل (۴) دیاگرام بلوکی فرآیند استخراج پیشنهادی جهت بازیابی دنباله نهان‌نگاره نشان داده شده است. علاوه بر این جزئیات فرآیند استخراج نهان‌نگاره در روش پیشنهادی به شرح زیر می‌باشد:

مرحله سوم: به‌کارگیری تبدیل GBT بر روی هر فریم با استفاده از ساختاری مناسب گرافی.

مرحله چهارم: اعمال تبدیل DCT بر روی k ضرایب اولیه GBT که دارای بیشترین مقادیر می‌باشند.

مرحله پنجم: به‌کارگیری تبدیل موجک دو سطحی هار بر روی ضرایب فرکانس پایین حاصل از تبدیل DCT

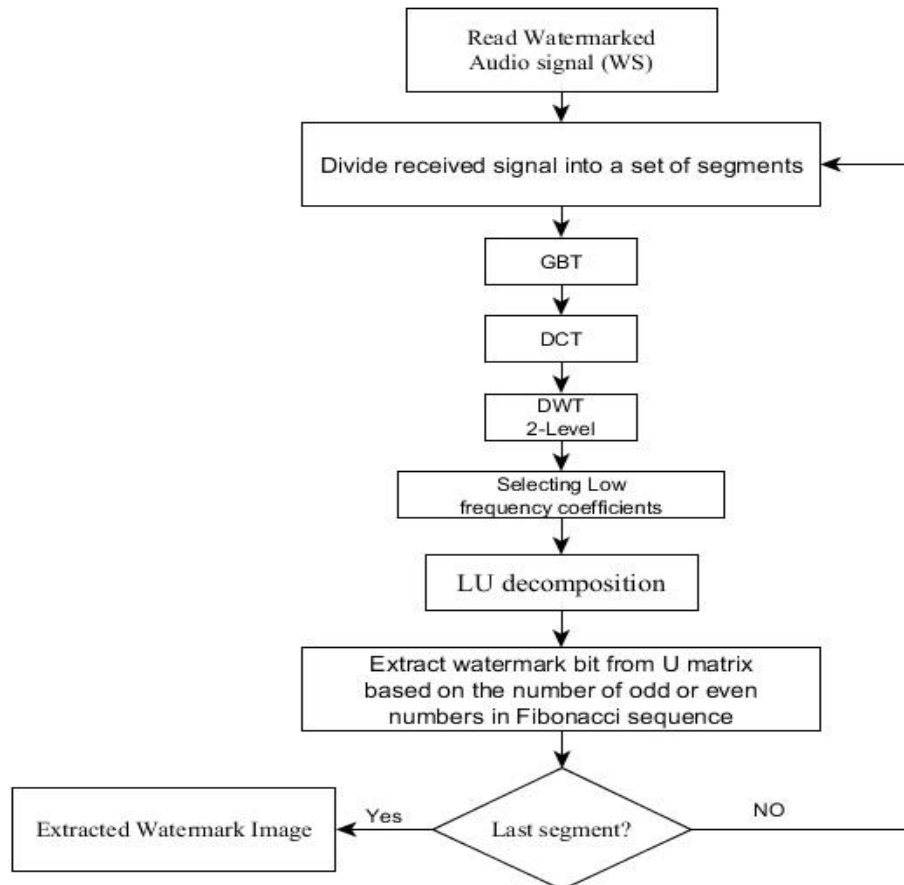
مرحله ششم: تبدیل بردار ضرایب تقریب سطح دوم به فرم ماتریس دوبعدی

مرحله اول: خواندن سیگنال صوت نهان‌نگاری شده S'

مرحله دوم: فریم‌بندی سیگنال ورودی S' به مجموعه‌ای از فریم‌های غیر همپوشان

مرحله هشتم: در صورتی که در ماتریس بالا مثلثی U تعداد اعداد زوج از تعداد اعداد فرد بیشتر باشند، بیت نهان نگاره صفر و در غیر این صورت یک تشخیص داده می‌شود.

مرحله هفتم: محاسبه تجزیه LU بر روی ماتریس دوبعدی مرحله قبل به دو ماتریس پایین مثلثی (L) و بالا مثلثی (U)

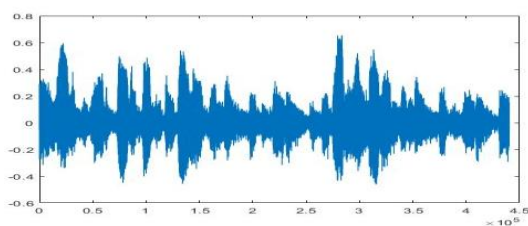


شکل (۴): فرآیند استخراج نهان نگاره روش پیشنهادی

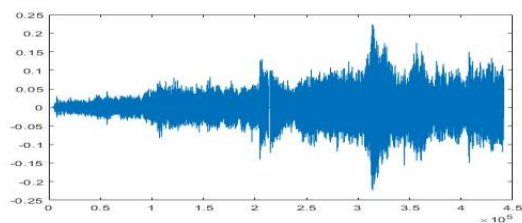
۵- نتایج شبیه‌سازی‌ها

به منظور پیاده‌سازی رویکرد نهان‌نگاری صوتی پیشنهادی از نرم‌افزار MATLAB ۲۰۱۷ استفاده شده است. تمامی شبیه‌سازی‌ها بر روی یک سامانه کامپیوتری با پردازنده Core i7 با میزان حافظه RAM برابر ۶ GB انجام شده‌اند.

به منظور ارزیابی طرح پیشنهادی از پنج نوع فایل صوتی با سبک‌های متفاوت (Jazz و Classic ، Electronic ، Blues) با مشخصات Mono 16-bit و فرکانس نمونه‌برداری 44100 هرتز استفاده شده است. شکل (۵) نمایش چهار فایل صوتی را نشان می‌دهد.



(الف)



(ب)

که $S[n]$ و $S'[n]$ به ترتیب مبین سیگنال اصلی و سیگنال نهان‌نگاره شده می‌باشند. لازم به ذکر است که بر اساس استاندارد IFPI حداقل مقدار این پارامتر ۲۰ دسی‌بل می‌باشد.

پارامتر مهم دیگر برای نشان دادن کیفیت شنیداری به نام ODG وجود دارد، که با نرم‌افزار EAQUAL اندازه‌گیری می‌شود و می‌تواند عددی در بازه بین (۰ تا -۴) به خود اختصاص دهد به طوری که مقدار ۰ مبین بهترین شفافیت از لحاظ شنیداری و مقدار -۴ به معنای غیرقابل تحمل از لحاظ شنیداری می‌باشد [۳۲].

ب- معیار میزان ظرفیت:

ظرفیت داده مبین میزان بیت‌هایی که در واحد زمان در سیگنال صوتی جاده می‌شوند. بر اساس استاندارد IFPI حداقل ظرفیت روش نهان‌نگاری 20 bps می‌باشد.

ظرفیت جاده می‌باشد بر اساس رابطه ۱۵ محاسبه می‌گردد [۳۴]:

$$C = \frac{F_s}{N} \text{bps} \quad (15)$$

به طوری که F_s فرکانس نمونه‌برداری و N تعداد نمونه‌ها در هر فریم می‌باشد.

ج- معیارهای مقاوم‌پذیری

مقاومت روش نهان‌نگاری صوت می‌تواند به‌عنوان توانایی آن در استخراج دنباله نهان‌نگاره از سیگنال دریافت شده تصور گردد که بر روی آن حملات پردازش سیگنال همچون نویز اضافه، فشرده‌سازی فیلترینگ و غیره صورت گرفته است. معیارهای همبستگی نرمال شده (NC) و نرخ خطای بی‌تی (BER) به‌عنوان دو معیار شناخته‌شده جهت محاسبه میزان مقاوم‌پذیری طرح نهان‌نگاری مورد استفاده قرار می‌گیرند که به‌صورت روابط (۱۶) و (۱۷) معرفی می‌شوند [۳۳]:

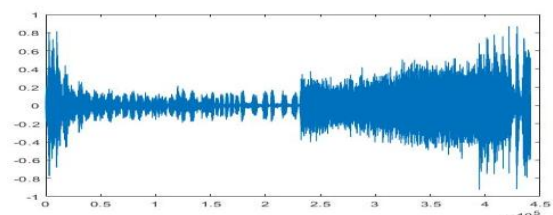
$$BER(W, \bar{W}) = \frac{\sum_{i=1}^M \sum_{j=1}^M W(i, j) \oplus \bar{W}(i, j)}{M \times M} \quad (16)$$

$$NC = \frac{\sum_{i=1}^M \sum_{j=1}^M W(i, j) \cdot \bar{W}(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^M W(i, j)} \cdot \sqrt{\sum_{i=1}^M \sum_{j=1}^M \bar{W}(i, j)}} \quad (17)$$

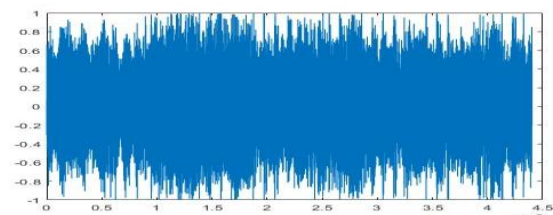
که W و \bar{W} به ترتیب دنباله نهان‌نگاری اصلی و دنباله نهان‌نگاره استخراج‌شده می‌باشند. همچنین نماد \oplus مبین یا انحصاری می‌باشد.

۵-۲- نتایج آزمایش‌ها

نتایج حاصل از ارزیابی میزان شفافیت (بر اساس معیارهای SNR و ODG)، مقاوم‌پذیری (بر اساس معیارهای BER و NC) و ظرفیت



(ج)



(د)

شکل (۵): نمایش چهار فایل صوتی مختلف با سبک‌های (الف) Blues، (ب) Classic، (ج) Electronic و (د) Jazz

اطلاعات نهان‌نگاره تصویر Lena می‌باشد که در شکل (۶) نشان داده شده است.



شکل (۶): تصویر نهان‌نگاره

۵-۱- معرفی معیارهای ارزیابی کارایی

الف- معیار شفافیت ادراکی

روشی نهان‌نگاری صوت دارای شفافیت ادراکی است اگر قادر به جاسازی دنباله‌ای نهان‌نگاره در سیگنال اصلی به‌گونه‌ای باشد که تغییرات بین سیگنال اصلی و نهان‌نگاره شده قابل نشود. از جمله مهم‌ترین معیارهای ارزیابی عینی شفافیت ادراکی می‌توان به معیار نسبت سیگنال به نویز (SNR) اشاره نمود. معیار SNR، یک معیار آماری است که توسط استاندارد IFPI به‌منظور محاسبه میزان شباهت بین سیگنال اصلی و نهان‌نگاره شده معرفی شده است. معیار SNR توسط رابطه (۱۴) محاسبه می‌گردد [۳۳]:

$$SNR = 10 \log_{10} \frac{\sum S^2[n]}{\sum_n [S[n] - S'[n]]^2} \quad (14)$$

جاده‌ی، با طول فریم‌های صوتی مختلف ۴، ۸، ۱۶ و ۳۲ نمونه بر روی فایل‌های صوتی با سبک‌های (Classic, Electronic, Blues, Jazz) در جدول (۱-۴) نشان داده شده است. بنابراین، همان‌طوری که از نتایج به‌دست‌آمده مشاهده می‌شود با افزایش طول فریم صوتی، اگر چه ظرفیت جاده‌ی کاهش می‌یابد ولیکن میزان مقاوم‌پذیری و شفافیت افزایش می‌یابند. از طرف دیگر تداوم افزایش طول فریم‌های صوتی موجب کاهش چشمگیر ظرفیت

جاده‌ی شده در حالی که میزان مقاوم‌پذیری و شفافیت تغییر چندانی نخواهند کرد. بنابراین، می‌توان نتیجه گرفت به‌منظور رسیدن به مصالحه‌ای مناسب بین سه معیار ظرفیت، شفافیت و مقاوم‌پذیری بهترین طول فریم ۱۶ می‌باشد. بنابراین، در جدول (۵) نشان داده شده است. همان‌طوری که مشاهده می‌شود روش نهان‌نگاری پیشنهادی به‌طور متوسط دارای SNR=45/13 dB و ODG=-0/39 در متوسط نرخ جاده‌ی ۶۲۵/۷۵ bps می‌باشد.

جدول (۱): نتایج حاصل از ظرفیت، مقاوم‌پذیری و شفافیت با تغییر اندازه فریم بر روی سیگنال صوتی Blue

طول فریم	شفافیت		مقاوم‌پذیری		ظرفیت جاده‌ی (برحسب میزان بیت در هر ثانیه)
	SNR(dB)	ODG	NC	BER	
۴	۱۹/۴۱	-۱/۳۱	۰/۹۸۰۱	۰/۲۷۵۳	۲۳۷۲
۸	۳۲/۱۴	-۰/۵۶	۰/۹۸۴۵	۰/۲۱۹۳	۱۲۱۹
۱۶	۴۲/۲۱	-۰/۴۱	۰/۹۹۳۴	۰/۱۴۷۲	۶۲۹
۳۲	۴۶/۱۹	-۰/۳۷	۰/۹۹۴۱	۰/۱۳۷۹	۳۰۸

جدول (۲): نتایج حاصل از ظرفیت، مقاوم‌پذیری و شفافیت با تغییر اندازه فریم بر روی سیگنال صوتی Classic

طول فریم	شفافیت		مقاوم‌پذیری		ظرفیت جاده‌ی (برحسب میزان بیت در هر ثانیه)
	SNR(dB)	ODG	NC	BER	
۴	۱۷/۹۲	-۱/۴۴	۰/۹۷۹۰	۰/۳۱۵۲	۲۳۶۱
۸	۳۰/۳۵	-۰/۶۹	۰/۹۸۱۱	۰/۲۵۹۳	۱۲۰۴
۱۶	۴۱/۲۴	-۰/۴۸	۰/۹۹۱۱	۰/۲۳۴۲	۶۱۸
۳۲	۴۵/۱۱	-۰/۳۷	۰/۹۹۳۱	۰/۱۷۹۹	۳۰۲

جدول (۳): نتایج حاصل از ظرفیت، مقاوم‌پذیری و شفافیت با تغییر اندازه فریم بر روی سیگنال صوتی Electronic

طول فریم	شفافیت		مقاوم‌پذیری		ظرفیت جاده‌ی (برحسب میزان بیت در هر ثانیه)
	SNR(dB)	ODG	NC	BER	
۴	۲۳/۳۷	-۰/۸۶	۰/۹۷۸۲	۰/۳۱۲۳	۲۳۸۳
۸	۳۶/۱۴	-۰/۵۷	۰/۹۸۱۵	۰/۲۹۹۳	۱۲۲۳
۱۶	۵۲/۸۳	-۰/۲۴	۰/۹۹۰۴	۰/۲۶۹۷	۶۳۴
۳۲	۵۵/۳۲	-۰/۲۱	۰/۹۹۲۱	۰/۱۷۸۹	۳۱۲

جدول (۴): نتایج حاصل از ظرفیت، مقاوم‌پذیری و شفافیت با تغییر اندازه فریم بر روی سیگنال صوتی Jazz

طول فریم	شفافیت		مقاوم‌پذیری		ظرفیت جاده‌ی (برحسب میزان بیت در هر ثانیه)
	SNR(dB)	ODG	NC	BER	
۴	۲۰/۳۱	-۰/۹۷	۰/۹۸۲۱	۰/۲۷۵۳	۲۳۱۱
۸	۳۵/۹۱	-۰/۷۱	۰/۹۸۴۵	۰/۲۱۹۳	۱۲۱۴
۱۶	۴۲/۲۵	-۰/۴۳	۰/۹۹۲۹	۰/۱۷۰۳	۶۲۲
۳۲	۴۴/۱۹	-۰/۳۲	۰/۹۹۴۷	۰/۱۴۷۹	۳۰۲

جدول (۵): نتایج حاصل از ارزیابی روش پیشنهادی از نظر شفافیت و ظرفیت جادهی

ODG	SNR(dB)	ظرفیت جادهی (برحسب میزان بیت در هر ثانیه)	فایل صوتی
-۰/۴۱	۴۳/۲۱	۶۲۹	Blues
-۰/۴۸	۴۱/۲۴	۶۱۸	Classic
-۰/۲۴	۵۳/۸۳	۶۳۴	Electronic
-۰/۴۳	۴۲/۲۵	۶۲۲	Jazz
-۰/۳۹	۴۵/۱۳	۶۲۵,۷۵	میانگین

علاوه بر این، مقاوم‌پذیری روش پیشنهادی در برابر حملات Stirmark بر اساس معیارهای نرخ خطای بی‌تی (BER) و همبستگی نرمال در جداول (۶ و ۷) نشان داده شده است.

جدول (۶): نتایج حاصل از نرخ خطای بی‌تی روش پیشنهادی در برابر حملات Stirmark

حملات	فایل صوتی با سبک			
	Blues	Electronic	Classic	Jazz
Nothing	۰/۱۴۷۲	۰/۲۶۹۷	۰/۲۳۴۲	۰/۱۷۰۳
Addbrumm_1100	۰/۷۵۸۱	۰/۷۶۶۹	۰/۹۱۱۸	۰/۴۵۶۳
Addbrumm_2100	۰/۶۸۷۸	۰/۶۶۵۱	۰/۶۱۳۵	۰/۷۱۲۴
Addsinus	۰/۸۴۵۳	۱/۱۱۴۲	۰/۹۶۳۵	۰/۹۸۹۹
Compressor	۱/۱۵۶۲	۱/۵۸۹۹	۱/۲۳۱۹	۱/۱۹۰۳
Extrastereo_50	۰/۴۵۷۱	۰/۴۲۵۴	۰/۴۴۴۲	۰/۴۵۰۳
FFT_invert	۰/۶۹۶۲	۰/۶۵۴۹	۰/۴۹۹۶	۰/۷۴۲۱
FFT_real_reverse	۱/۳۶۵۲	۱/۷۴۵۷	۱/۸۵۸۹	۱/۶۳۶۲
Invert	۰/۱۷۴۵	۰/۲۳۵۶	۰/۲۱۷۲	۰/۳۶۴۴
Lsbzero	۰/۷۸۹۱	۰/۹۶۵۸	۰/۸۶۶۶	۰/۹۹۶۵
Rc_highpass	۰/۸۵۸۷	۰/۷۴۸۹	۰/۹۸۷۲	۰/۸۹۹۴

جدول (۷): نتایج حاصل از همبستگی نرمال شده روش پیشنهادی در برابر حملات Stirmark

حملات	فایل صوتی با سبک			
	Blues	Electronic	Classic	Jazz
Nothing	۰/۹۹۳۴	۰/۹۹۰۴	۰/۹۹۱۱	۰/۹۹۲۹
Addbrumm_1100	۰/۹۸۹۲	۰/۹۸۹۵	۰/۹۸۸۶	۰/۹۹۱۰
Addbrumm_2100	۰/۹۹۱۵	۰/۹۹۰۸	۰/۹۹۱۴	۰/۹۹۰۳
Addsinus	۰/۹۸۵۵	۰/۹۸۱۴	۰/۹۸۶۲	۰/۹۸۹۱
Compressor	۰/۹۸۱۶	۰/۹۸۰۲	۰/۹۸۳۳	۰/۹۸۴۰
Extrastereo_50	۰/۹۹۵۶	۰/۹۹۶۹	۰/۹۹۸۳	۰/۹۹۶۱
FFT_invert	۰/۹۹۰۱	۰/۹۹۶۵	۰/۹۹۱۴	۰/۹۹۳۴
FFT_real_reverse	۰/۹۸۰۹	۰/۹۸۷۴	۰/۹۸۴۰	۰/۹۸۳۹
Invert	۰/۹۹۴۸	۰/۹۹۲۷	۰/۹۹۸۳	۰/۹۹۱۴
Lsbzero	۰/۹۸۸۹	۰/۹۸۷۳	۰/۹۸۶۵	۰/۹۸۲۴
Rc_highpass	۰/۹۸۹۱	۰/۹۸۴۶	۰/۹۸۴۰	۰/۹۶۴۸

در جدول (۸) نتایج حاصل از ارزیابی جامع روش پیشنهادی با سایر روش‌های پیشین مرتبط از نظر میزان شفافیت (بر اساس ثابته) نشان داده شده است.

جدول (۸): مقایسه روش پیشنهادی با کارهای پیشین از نظر معیارهای شفافیت و میزان درج داده نهان‌نگاره

روش - مرجع	ظرفیت (بیت بر ثانیه)	نرخ سیگنال به نویز (دسی بل)	ODG
DCT [۱۰]	۶۰۲/۹۳	۲۰/۸۹	-۰/۰۴۰
DE + DM [۱۲]	۱۳۶	۳۵ to ۳۹	Not reported
MFFCFD+ DTCWT [۱۳]	۶۴	۲۹,۷	Not reported
DWPT+ DCT+ QIM [۱۴]	۴۵ to ۴۷۳	۱۹ to ۲۵	-۰/۰۵ to -۱/۵
DWT + SVD [۱۵]	۲۵۸	۳۸/۱۷	-۰/۷۶
DWT + KKT [۱۶]	۶۶۶ to ۱۳۳۳	۱۴/۱ to ۱۹/۱	Not reported
SS + DCT [۱۷]	۸۴	Not reported	-۰/۷
DWT + RDM [۱۸]	۱۰۲ to ۶۸۹	۱۹/۸ to ۳۰/۳	-۰/۰۱۵ to -۱/۶
DCT + LR & SD [۱۹]	۱۶۴ to ۳۰۰۰	۱۴/۵ to ۳۵	-۰/۳۶ to -۱/۶۴
SVD + SVM [۲۰]	۱۷۲	۴۲/۰۷	-۰/۳۳
SS [۲۱]	۴۳	Not reported	-۰/۶۲ to -۳/۶۴
MDWT + DCT [۲۲]	۸۶/۱۳	۱۹/۹۲	-۰/۳۳۴
Blind-DWT[۲۵]	۶۸۶	۱۴/۹۳۸	-۰/۶۲۳
DCT-Schur[۲۶]	۵۱۶/۲۶	۷۷/۹۵	-۰/۱۷۹۳
DWT-Schur[۳۵]	۳۱۹/۲۹	۸۱/۴۳	-۰/۱۸۴
میانگین روش پیشنهادی	۶۲۵/۷۵	۴۵/۱۳	-۰/۳۹

حملات مختلف پردازش سیگنال Stirmark زیر ۱/۵ درصد است که نشان می‌دهد طرح پیشنهادی از مقاوم‌پذیری بسیار خوبی برخوردار است. بنابراین، می‌توان نتیجه گرفت که طرح پیشنهادی توانسته است به نحو مناسبی سه معیار شفافیت، مقاوم‌پذیری و ظرفیت را بهبود بخشد.

به‌عنوان کارهای آتی می‌توان با انتخاب مناسب محل‌های درج توسط قوانین فازی و نیز به‌کارگیری یک استخراجگر هوشمند مبتنی بر شبکه‌های عصبی کارایی سامانه‌های نهان‌نگاری صوت را بهبود بخشید.

۷- مراجع

- [1] Petitcolas, A. P. Fabien, R. J. Anderson, and M. G. Kuhn, "Information hiding-a survey," Proceedings of the IEEE 87, no. 7, pp. 1062-1078, 1999.
- [2] J. C. Ingemar, Matthew L. Miller, J. A. Bloom, and C. Honsinger, "Digital watermarking," vol. 1558607145, Springer, 2002.
- [3] J. C. Ingemar, Matt L. Miller, and A. L. McKellips, "Watermarking as communications with side information," published in the Proceedings of the IEEE, vol. 87, no. 7, pp. 1127-1141, 1999.
- [4] A. Brickman, "Literature survey on audio watermarking," EE381K-Multidimensional Signal Processing, March 2003.

۶- نتیجه‌گیری

در این مقاله، روشی نهان‌نگاری صوتی جدید بر مبنای هم‌افزایی سه تبدیل کآمد به نام‌های تبدیل گراف محور (GBT)، تبدیل کسینوسی گسسته (DCT) و تبدیل موجک گسسته (DWT) پیشنهاد شد. به‌کارگیری قابلیت‌های موجود در هر یک از این سه تبدیل موجب شد تا رویکرد پیشنهادی از مقاوم‌پذیری بسیار خوبی در برابر حملات پردازش سیگنال برخوردار باشد. علاوه بر این، به‌منظور دستیابی به شفافیت بالا و نیز استخراج کور دنباله نهان‌نگاره، درج بیت‌های نهان‌نگاره در ماتریس بالامثلثی حاصل از تجزیه LU ضرایب تقریب تبدیل موجک گسسته، بر اساس دنباله فیبوناچی انجام شد بدین صورت که اگر بیت نهان‌نگاره دارای مقدار صفر باشد، مقادیر غیر صفر ماتریس بالا مثلثی U با نزدیک‌ترین اعداد زوج و در صورتی که یک باشد، با نزدیک‌ترین اعداد فرد در دنباله فیبوناچی جایگزین می‌شوند. نتایج حاصل از ارزیابی‌ها بر روی چهار فایل صوتی با سبک‌های مختلف Blues, Electronic, Jazz و Classic نشان داد که روش پیشنهادی به‌طور متوسط دارای نرخ سیگنال به نویز ۴۵/۱۳ دسی‌بل و ODG برابر ۰/۳۹- در متوسط نرخ جادهی ۶۲۵/۷۵ بیت بر ثانیه می‌باشد. علاوه بر این متوسط میزان خطای طرح پیشنهادی در برابر

- [20] M. Mosleh, H. Latifpour, M. Kheyrandish, M. Mosleh, and N. A. Hosseinpour, "robust intelligent audio watermarking scheme using support vector machine," *Journal of Zhejiang University SCIENCE*, pp. 11-08, 2016.
- [21] R. Li, S. Xu, and H. Yang, "Spread spectrum audio watermarking based on perceptual characteristic aware extraction," *IET Signal Processing*, vol. 10(3), pp. 266-273, 2016.
- [22] D. Bratić, F. Vešović, and V. Mijanović, "Audio watermarking under gradient-based 37 reconstruction attack," *In Embedded Computing (MECO), 2016 5th Mediterranean Conference on IEEE*, pp. 148-152, 2016.
- [23] S. M. Mohsenfar, M. Mosleh, and A. Barati, "Audio watermarking method using QR decomposition and genetic algorithm," *Multimedia Tools and Applications*, vol. 74(3), pp. 759-779, 2015.
- [24] M. Mosleh, et al., "High-capacity, transparent and robust audio watermarking based on synergy between DCT transform and LU decomposition using genetic algorithm," *Analog Integrated Circuits and Signal Processing*, vol. 100(3), pp. 513-525, 2019.
- [25] A. A. Attari and A. Beheshti Shirazi, "A. Robust and Blind Audio Watermarking in Wavelet Domain," *In: Proceedings of the International Conference on Graphics and Signal Processing - ICGSP '17. Singapore, Singapore*, pp. 69-73, 2019.
- [26] H. Karajeh and M. Maqableh, "An imperceptible, robust, and high payload capacity audio watermarking scheme based on the DCT transformation and Schur decomposition," *Analog Integr Circ Sig Process*, 2018.
- [27] A. Kaur and M. K. Dutta, "An optimized high payload audio watermarking algorithm based on LU-factorization," *Multimedia Systems*, vol. 24(3), pp. 341-353, 2018.
- [28] M. Farzaneh, R. M. Toroghi, and M. Asgari, "Audio Compression Using Graph- based Transform," *In 2018 9th International Symposium on Telecommunications (IST)*, pp. 410-415, 2018.
- [29] G. Strng, "The discrete cosine transform," *SIAM Review*, vol. 41(1), pp. 135-147, 1999.
- [30] A. Al-Haj, A. A. Mohammad, and L. Bata, "DWT based audio watermarking," *Int. Arab J. Inf. Technol.*, vol. 8(3), pp. 326-333, 2011.
- [31] M. Mosleh, et al., "Presenting a systematic method for LU decomposition of a matrix with linear systolic arrays," *In book presenting a systematic method for LU decomposition of a matrix with linear systolic arrays, Series Presenting a systematic method for LU decomposition of a matrix with linear systolic arrays, IEEE*, pp. 123-127, 2008.
- [32] A. Zplane Lerch, "Development, Eequal Evaluate Audio Quality," version: 0.1. 3alpha, 2002.
- [33] A. Kaur and M. K. Dutta, "An optimized high payload audio watermarking algorithm based on LU-factorization," *Multimedia Systems*, vol. 24(3), pp. 341-353, 2018.
- [34] H. Latifpour, M. Mosleh, and M. Kheyrandish, "An intelligent audio watermarking based on KNN learning algorithm," *International Journal of Speech Technology*, vol. 18(4), pp. 697-706, 2015.
- [35] H. Karajeh, et al., "A robust digital audio watermarking scheme based on DWT and Schur decomposition," *Multimedia Tools and Applications*, vol. 78(13), pp. 18395-18418, 2019.
- [5] Arnold M, "Audio Watermarking: features, applications and algorithms," *IEEE International conference Multimedia and Expo*, vol. 2, pp. 1033-1066, 2000.
- [6] L. Miller Matt, J. C. Ingemar, Jean-Paul, M. G. Linnartz, and Ton Kalker, "A review of watermarking principles and practices," published in *Digital Signal Processing in Multimedia Systems*, Ed. K. K. Parhi and T. Nishitani, Marcell Dekker Inc., pp. 461-485, 1999.
- [7] S. Prabhishak and R. S. Chadha, "A Survey of Digital Watermarking Techniques, Applications and Attacks," *International Journal of Engineering and Innovative Technology (IJET)*, vol. 2, Issue 9, pp. 165-175, 2013.
- [8] S. Amit, J. Susheel, and J. Anurag, "A Survey: Digital Video Watermarking," *International Journal of Scientific & Engineering Research*, vol. 4, Issue 7, July- ISSN 2229-5518 pp. 1261-1265, 2013.
- [9] H. Farooq, "A Survey of Digital Watermarking Techniques for multimedia data," *MIT International Journal of Electronics and communication Engineering*, vol. 2, no.1, ISSN 2230-7672, pp. 37-43, Jan. 2012.
- [10] H. T. Hu and L. Y. Hsu, Robust, "Transparent and high-capacity audio watermarking in DCT domain," *Signal Processing*, vol. 109, pp. 226-235, 2015.
- [11] S. M. Pourhashemi, M. Mosleh, and Y. Erfani, "Audio watermarking based on synergy between Lucas regular sequence and Fast Fourier Transform," *Multimedia Tools and Applications*, vol. 78(16), pp. 22883-22908, 2019.
- [12] B. Lei, Y. Soon, and E. L. Tan, "Robust SVD-based audio watermarking scheme with differential evolution optimization," *IEEE transactions on audio, speech, and language processing*, vol. 21(11), pp. 2368-2378, 2013.
- [13] X. C. Yuan, C. M. Pun, and C. P. Chen, "Robust Mel-Frequency Cepstral coefficients feature detection and dual-tree complex wavelet transform for digital audio watermarking," *Information Sciences*, vol. 298, pp. 159-179, 2015.
- [14] H. T. Hu, L. Y. Hsu, and H. H. Chou, "Perceptual-based DWPT-DCT framework for selective blind audio watermarking," *Signal Processing*, vol. 105, pp. 316-327, 2014.
- [15] A. Al-Haj, "A dual transform audio watermarking algorithm," *Multimedia tools and applications*, vol. 73(3), pp. 1897-1912, 2014.
- [16] H. N. Huang, S. T. Chen, M. S. Lin, W. M. Kung, and C. Y. Hsu, "Optimization-based embedding for wavelet-domain audio watermarking," *Journal of Signal Processing Systems*, vol. 80(2), pp. 197-208, 2015.
- [17] Y. Xiang, I. Natgunanathan, Y. Rong, and S. Guo, "Spread spectrum-based high embedding capacity watermarking method for audio signals," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 23(12), pp. 2228-2237, 2015.
- [18] H. T. Hu and L. Y. Hsu, "Supplementary Schemes to Enhance the Performance of DWT-RDM Based Blind Audio Watermarking," *Circuits, Systems, and Signal Processing*, pp. 1-22, 2016.
- [19] M. Jeyhoon, M. Asgari, L. Ehsan, and S. Z. Jalilzadeh, "Blind audio watermarking algorithm based on DCT, linear regression and standard deviation," *Multimedia Tools and Applications*, pp. 1-17, 2016.

Transparent and Robust Audio Watermarking Using Synergy LU Decomposition and the Fibonacci Sequence in GBT-DCT-DWT Transforms

M. Mosleh , S. Setayeshi* , B. Barekatin, M. Mosleh

*Amirkabir University of Technology

(Received: 15/04/2020, Accepted: 05/08/2020)

ABSTRACT

Digital watermarking technology is presented as one of the best solutions for solving unauthorized copying, content identification and authentication of digital media. Digital watermarking can be applied in the fields of image, text, audio, and video contents. Audio watermarking has recently attracted the attention of researchers because the human auditory system is much more sensitive than his vision system. Therefore, insertion of the data into audio signals in a transparent way is much more difficult than other watermarking species. An efficient audio watermarking system should be able to improve reconciliation of the three measures of transparency, strength and capacitance. Improving the compromise between these three measures is a challenging problem, due to the fact that increasing the capacity of the input signal, causes distortions in the signal which in turn causes a reduction in transparency and robustness. In this paper, a new method is proposed for watermarking of audio signals that is able to improve the three criteria of transparency, robustness and capacity in an appropriate manner. In order to improve the resilience of the hidden signal against signal processing attacks, the synergy of three efficient transforms in the field of signal processing namely, the graph-centric conversion (GBT), the discrete cosine transform (DCT) and the discrete wavelet transform (DWT), has been used. In addition, in order to maintain the transparency of the watermarked signal, the watermarked data in the high triangle matrix resulting from the LU decomposition of the approximation coefficients of the discrete wavelet are entered in an order which is based on the Fibonacci sequence. The procedure is based on the value of the watermarked bit, the non-zero values of the upper triangular matrix are replaced by the nearest even or odd numbers in the Fibonacci sequence. The latent extraction operation is completely blind. The results of the evaluation of the proposed method on audio files with Blue, Electronic, Classic and Jazz styles show that the proposed method, despite good resistance to various signal processing attacks, has an average signal rate of 45.13 dB and a placement rate of 625.75 bits per second.

Keywords: Copyright protection, Audio watermarking, Graph-based Transform(GBT), Discrete Cosine Transform (DCT), Discrete wavelet Transform (DWT) LU Decomposition.

* Corresponding Author Email: setayesh@aut.ac.ir