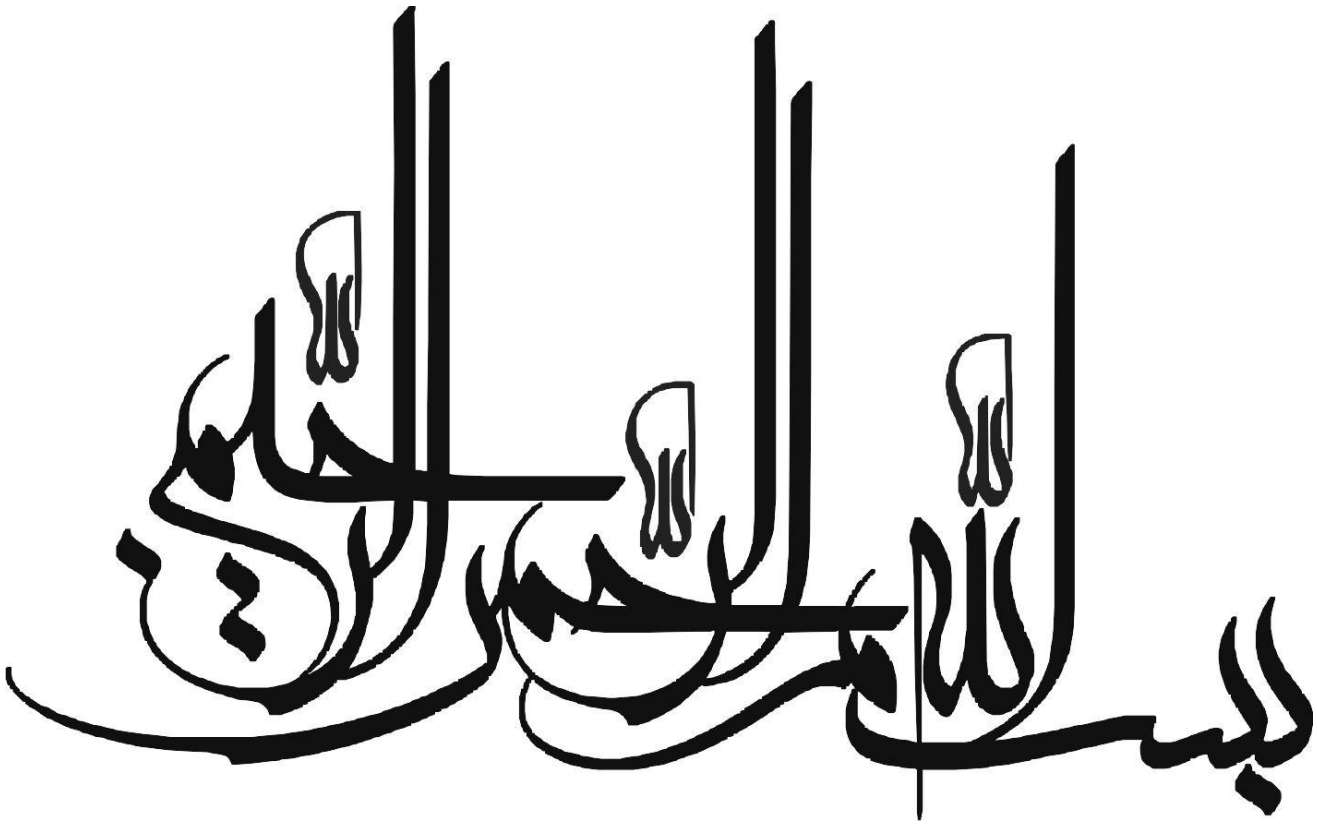

بسم الله الرحمن الرحيم

دانشکده برق و رباتیک
گروه الکترونیک

استفاده از توابع بسل جهت پنهان سازی اطلاعات درون گفتار

دانشجو : ایمان خسرویان چم‌پیری

استاد راهنما :
دکتر حسین مروی

استاد مشاور:
دکتر هادی گرایلو

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

شهریور ۱۳۹۱

دانشگاه صنعتی شاهرود

دانشکده :

گروه :

پایان نامه کارشناسی ارشد آقای ایمان خسرویان چم پیری

تحت عنوان: استفاده از توابع بسل جهت پنهان سازی اطلاعات درون گفتار

در تاریخ ۹۱/۶/۲۷ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجه **عـالی** مـورد پذیرش قرار گرفت.

اساتید راهنما	امضاء	اساتید مشاور	امضاء
حسین مروی		هادی گرایلو	

اساتید داور	امضاء	نماینده تحصیلات تکمیلی	امضاء
مرتضی زاهدی		مرتضی رحیمیان	
سید علی سلیمانی			

تقدیم به

پدرم

مادرم

برادرم

به نام آنکه به قلم سوگند خورد و به انسان آموخت.

حال که این پایان نامه به پایان رسیده و به تمام مراحل آن از ابتدا تا کنون می‌اندیشم بیشتر به معنای این سخن پی می‌برم که گفت: هلاک شد آنکه استادی نداشت تا راهنماییش کند.

و به تحقیق اگر راهنمایی استاد بزرگوار دکتر حسین مروی نبود این پایان نامه راه به جایی نداشت.

همچنین از مهندس محسن زارعیان و نیز همه دوستانم در دانشگاه صنعتی شاهرود و تمامی اساتیدی که برای تعلیم دانشجویان زحمات زیادی متقبل شدند تشکر می‌نمایم.

تعهد نامه

اینجانب **ایمان خسرویان چم‌پیری** دانشجوی دوره کارشناسی ارشد رشته **مهندسی برق الکترونیک** گرایش **سیستم** دانشکده برق و رباتیک **دانشگاه صنعتی شاهرود** نویسنده پایان‌نامه **استفاده از توابع بسل جهت پنهان سازی اطلاعات درون گفتار** تحت راهنمائی **دکتر حسین مروی** متعهد می شوم .

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

پنهان سازی به معنای مخفی کردن نامحسوس اطلاعات سیگنال پیام (یا سیگنال الگو) در داخل سیگنال میزبان است به نحوی که سیگنال الگوگذاری شده‌ی حاصل از سیگنال اصلی یا میزبان قابل تمیز نباشد. این کار می‌تواند به روش‌های مختلفی انجام شود. سیگنال میزبان و سیگنال پیام هر کدام می توانند یکی از انواع محصولات چند رسانه‌ای از قبیل تصویر، صوت و متن باشند.

تبدیل فوریه – بسل تبدیلی است که سیگنال را به صورت مجموعی از توابع بسل نمایش می‌دهد. از آنجا که توابع بسل با سیگنال گفتار از نظر ساختاری تشابه دارند، این تبدیل می‌تواند در زمینه پردازش سیگنال‌های گفتاری مفید واقع شود.

روش‌های پیشنهاد شده در این پایان نامه، استفاده از تبدیل فوریه – بسل، جهت پنهان‌سازی اطلاعات در سه روش جایگزینی بیت کم ارزش، طیف گسترده و مدولاسیون اندیس کوانتیزاسیون می‌باشد. همچنین جهت مقایسه‌ی کارآیی این روش‌ها، در کنار آن از سه تبدیل فوریه، کسینوسی و موجک نیز استفاده می‌گردد. برای مقایسه بین این چهار تبدیل از معیارهای نرخ خطای بیت، هم بستگی بین نهان نگاره اصلی و استخراج شده و شفافیت استفاده می‌شود. برای دو معیار اول روابطی معرفی شده و برای معیار شفافیت از روش PESQ کمک گرفته خواهد شد. جهت تست پایداری در برابر حملات، سیگنال در برابر دو حمله فشرده سازی و افزودن نویز قرار داده می‌شود. در فشرده سازی، سیگنال به دو روش MP3 و GSM 6.10 فشرده و سپس به فرمت اصلی خود یعنی WAV باز گردانده شده و دو پارامتر نرخ خطای بیت و هم بستگی بین نهان نگاره اصلی و استخراج شده محاسبه می‌گردد. در افزودن نویز، به سیگنال نویز سفید افزوده شده تا میزان نسبت سیگنال به نویز در مقادیر ۱۰ dB و ۱۵ dB حاصل شود و سپس دو پارامتر نرخ خطای بیت و هم بستگی بین نهان نگاره اصلی و استخراج شده محاسبه می‌شود.

نتایج انجام شده بر روی پایگاه داده TIMIT نشان دهنده‌ی عملکرد خوب این تبدیل در مقایسه با سایر تبدیل‌های متداول کسینوسی، فوریه و موجک می‌باشد.

لیست مقالات استخراج شده از پایان نامه

1. Speech Watermarking using Fourier-Bessel Transform, accepted in the 2012 5th international conference on computer and electrical engineering (ICCEE 2012)

2. Increasing Accuracy of Automatic Language Identification by Using Spectral Features and WRBF Neural Network, accepted in the 2012 5th international conference on computer and electrical engineering (ICCEE 2012)

3. Speech watermarking using Fourier-Bessel transform and quantization index modulation, accepted in the 2012 4th International Conference on Signal Processing Systems (ICSPS 2012)

۴. معرفی تبدیل بسل به عنوان حوزه‌ی جدید جهت پنهان‌سازی اطلاعات درون سیگنال گفتار، ارائه شده به مجله پردازش علائم و داده‌ها

فهرست مطالب

عنوان **صفحه**

فهرست شکل‌ها

فهرست جدول‌ها

فصل اول:

مقدمه‌ای بر پنهان سازی اطلاعات

۱-۱ مقدمه

نمایش دادن اطلاعات به صورت دیجیتال مزایای بسیاری نسبت به سایر نمایش‌های معمول دارد که یکی از آنها قابلیت تهیه‌ی تعداد نامحدودی کپی بی‌عیب اصلی از یک محتوای دیجیتال می‌باشد. همچنین استفاده وسیع از اینترنت، به عنوان محیطی برای به اشتراک گذاشتن و تبادل ساده و سریع انواع محصولات چند رسانه‌ای، و پیشرفت‌های نو در فشرده‌سازی اطلاعات، در سال‌های پایانی قرن گذشته‌ی میلادی، توزیع موثر محتوای دیجیتال را آسان کرده و سبب تحول چشم‌گیری در عرصه‌ی ارتباطات دیجیتال شده است. تمامی این تحولات موجب رشد کمی و کیفی تولید انواع محصولات دیجیتال متنی، صوتی، تصویری و فیلم، و فراگیر شدن استفاده از این محصولات شد. با وجود این، این مزایا عامل اصلی نگرانی مالکان فکری و تولیدکنندگان محتوا شده و صاحبان این آثار به تدریج نگران حفاظت از حقوق خود شدند. به این ترتیب بحث حفاظت از اطلاعات دارای حق کپی[۱] برای جلوگیری از پایمال شدن حقوق این موئلفان ضروری به نظر می‌رسید. از طرف دیگر و در همین سال‌ها، محدودیت‌های اعمالی از سوی دولت‌ها برای استفاده‌ی عمومی از سیستم‌های رمزنگاری[۲] بیشتر می-شد و نیاز به روش جایگزینی برای مخابرات مخفی[۳] محسوس بود.

به این ترتیب ایده پنهان سازی اطلاعات[۴] راه حلی برای حل این مشکلات به وجود آورد. پنهان سازی به معنای مخفی کردن نامحسوس اطلاعات سیگنال پیام (یا سیگنال الگو[۵] یا نهان نگاره) در داخل سیگنال میزبان[۶] است به نحوی که سیگنال الگوگذاری[۷] شده‌ی حاصل از سیگنال اصلی یا میزبان قابل تمیز نباشد. سیگنال میزبان می تواند یکی از انواع محصولات چند رسانه‌ای از قبیل تصویر، صوت، گفتار و متن باشد.

---

۱ Copyright
۲ Cryptography
۳ Covert Communication
۴ Data Hiding
۵ Watermark
۶ Host Signal
۷ Watermarked Signal

روش‌های بسیاری برای پنهان‌سازی تصویر و فیلم پیشنهاد شده است،که بیشتر از نقص سیستم بینایی انسان برای پنهان‌سازی اطلاعات استفاده کرده‌اند. در مقایسه با پنهان‌سازی تصویر و فیلم، پنهان‌سازی گفتار از چالش‌های ویژه‌ای برخوردار است زیرا سیستم شنوایی انسان بسیار حساس‌تر از سیستم بینایی انسان است. برای فائق آمدن بر این چالش‌ها می‌بایست نهان‌نگاره را به گونه‌ای درون گفتار درج کرد که برای شنونده نامحسوس بوده و همچنین در مقابل پردازش‌ها مقاوم باشد.

یکی از مباحث بسیار نزدیک به پنهان‌سازی اطلاعات، نهان‌نگاری می‌باشد که این دو دارای تفاوت‌هایی در کاربرد می‌باشند. در نهان‌نگاری، نهان‌نگاره حاوی اطلاعات مهم است و مورد اهمیت می‌باشد، حال آنکه در پنهان سازی سیگنال اصلی مورد اهمیت است و نهان‌نگاره برای حفاظت از آن مورد استفاده قرار می‌گیرد. البته لازم به ذکر است که دو مفهوم نهان‌نگاری و پنهان سازی اطلاعات را نمی‌توان به طور کامل از هم جدا نمود و در بسیاری از اوقات به دلیل نزدیکی به جای هم استفاده می‌شوند.

۱-۲ کاربردهای پنهان‌سازی اطلاعات

در این قسمت بعضی از کاربردهای رایج برای سیستم‌های پنهان‌سازی اطلاعات ارائه می شود[۱]. هر کدام از این کاربردها به یک سیستم پنهان‌سازی اطلاعات با ویژگی‌های خاصی احتیاج دارند که در ادامه بحث می شود.

۱-۲-۱ حفاظت از حق کپی[۱]

یکی از مهم‌ترین کاربردهای پنهان سازی اطلاعات حفاظت از حق کپی می باشد. در این کاربرد به منظور نشان دادن مالک سیگنال اصلی و جلوگیری از ادعای دیگران برای مالکیت همان محصول، نهان‌نگاری انجام می‌شود. روش‌های مقاوم برای استفاده در این کاربرد مناسب میباشند. زیرا افراد غیرمجاز با هدف تخریب سیگنال نهان‌نگاره، محصول را مخدوش می‌کنند. بنابراین این روش‌ها می-بایست نسبت به انواع حملات مقاوم باشند.

[۱] Copyright Protection

۱-۲-۲ تعیین اعتبار

در کاربردهای تعیین اعتبار، هدف، تشخیص دست‌کاری نشدن اطلاعات است. این کار از طریق روش-های شکننده، که مقاومت کمی در برابر حملات دارند انجام می‌شود. برای مثال نمونه‌ای از کاربردهای تعیین اعتبار، پنهان کردن شماره‌ی شناسایی افراد در عکس آن‌ها می‌باشد. اگر شماره‌ی موجود در عکس با شماره‌ی شناسایی تطبیق نداشته باشد کارت شناسایی دست‌کاری شده و فاقد اعتبار می-باشد.

۱-۲-۳ کنترل کپی

با استفاده از پنهان سازی اطلاعات می‌توان وضعیت کپی را در داخل اطلاعات پنهان کرد. برای مثال محصولی که در آن الگوی غیرقابل کپی پنهان شده، توسط هیچ دستگاه رایت سی‌دی و دی‌وی‌دی قابل کپی شدن نیست. البته این کار مستلزم هماهنگی شرکت‌های سخت افزاری و نرم افزاری می-باشد.

۱-۲-۴ مخابرات مخفی

پنهان‌سازی اطلاعات درون یک محصول چند رسانه‌ای و ارسال آنها به صورت محرمانه می‌باشد. در این روش‌ها ظرفیت بالایی مورد نیاز است. همچنین مقاومت این گونه روش‌ها در مقابل حملات محدود است. برای این کاربرد تنها از روش‌هایی با آشکارسازی کور می توان استفاده کرد.

۱-۲-۵ نظارت بر بخش برنامه‌ها[۱]

شرکت‌ها می‌بایست بر اجرای قراردادی که با کانال‌های تلویزیونی یا رادیویی برای پخش آگهی بسته-اند نظارت کنند. یک راه برای این کار گذاشتن فردی برای بررسی کانال‌های تلویزیونی و نظارت بر مدت زمان و دفعات پخش آگهی می‌باشد که البته این کار پرهزینه و با خطای زیاد همراه است. راه دیگر استفاده از پنهان‌سازی اطلاعات می باشد. به این صورت که شرکت‌ها می توانند از یک سیگنال

[۱] Broadcast Monitoring

پنهان شده در آگهی و زمان سنجی وجود آن (به صورت اتوماتیک) روی امواج دریافتی، استفاده کنند. در این کاربرد از سیستم‌های پنهان سازی اطلاعات مقاوم استفاده می شود و پایداری از اهمیت بالایی برخوردار است. به همین دلیل می‌توان از روش‌های نیمه‌کور که پایداری بالاتری نسبت به روش‌های کور دارند استفاده کرد.

۱-۲-۶ پنهان نگاری اختصاصی

در این حالت به هر فرد دریافت کننده اطلاعات، سیگنال الگوی مخصوص به خود فرد داده می شود. از این طریق می توان به منبع نشت احتمالی اطلاعات پی برد.

۱-۳ انواع مختلف پنهان سازی اطلاعات بر اساس میزان پایداری نهان‌نگاره

روش‌های ارائه شده برای مخفی سازی را می‌توان از جنبه‌های مختلف طبقه بندی کرد. یک سیگنال همواره مورد حملات مختلف قرار می‌گیرد و این حملات بر نهان‌نگاره تاثیر می‌گذارد. حملات زیادی در سیستم‌های پنهان‌سازی وجود دارند که تمام سیستم‌های پنهان سازی به پایداری در مقابل تمامی حملات نیاز ندارند. برخی حملات رایج در سیستم‌های نهان نگاری صوتی عبارتند از:

- اضافه کردن نویز جمعی و ضربی
- فیلتر کردن شامل انواع فیلتر پایین گذر، بالا گذر و میان گذر
- فشرده سازی توسط روش‌های مختلف مانند MP3 و GSM 6.10
- Dynamics: که در آن دامنه سیگنال صوتی در طول زمان افزایش و کاهش می یابد
- اضافه کردن اکو
- حذف نمونه‌های سیگنال صوتی
- تغییر گام: در این حالت گام سیگنال صوتی تغییر می کند، در حالی که طول سیگنال تغییر نمی-کند.

- تغییر نرخ نمونه برداری: برای مثال کاهش آن از 8KHZ به 6KHZ

بر اساس میزان پایداری نهان‌نگاره، طبقه بندی سیستم‌های پنهان سازی اطلاعات را می‌توان به صورت زیر بیان کرد.

۱-۳-۱ سیستم‌های مقاوم

روش‌هایی هستند که نسبت به انواع حملات مقاوم می‌باشند. در این روش‌ها حتی اگر حمله کننده از وجود پیام در سیگنال میزبان مطلع باشد، از بین بردن پیام با انجام حملات متداول نباید امکان پذیر باشد.

۱-۳-۲ سیستم‌های شکننده

این روش‌ها دارای مقاومت محدودی در برابر حملات هستند. از این روش‌ها برای پی بردن به وجود دست‌کاری در اطلاعات استفاده می شود و سیگنال پیام در این روش‌ها با کوچکترین دست کاری مخدوش می شود. از شکننده بودن بیشتر برای مقاصد تعیین اعتبار استفاده می شود. به این صورت که اگر سیگنال پیام به درستی از سیگنال استخراج شد، می توان نتیجه گرفت که اطلاعات دست کاری نشده اند.

۱-۳-۳ سیستم‌های نیمه شکننده

این روش‌ها در برابر بعضی از حملات مانند فشرده سازی با نرخ کم، پایدار می باشند. این گونه روش‌ها می توانند برای تشخیص محل دست کاری در اطلاعات به کار روند.

۱-۴ انواع مختلف پنهان سازی اطلاعات از نقطه نظر آشکارسازی نهان‌نگاره

همچنین جهت تقسیم بندی از نقطه نظر نوع آشکارسازی سیگنال نهان‌نگاره، روش‌های موجود را می توان به سه دسته زیر تقسیم کرد.

۱-۴-۱ آشکارسازی کور

در این روش‌ها نیازی به سیگنال میزبان اصلی در سمت گیرنده نیست و پیام مستقیما از سیگنال حاوی پیام استخراج می شود.

۱-۴-۲ آشکارسازی نیمه کور

در این روش‌ها برای آشکارسازی تنها داشتن بعضی ویژگی‌های سیگنال میزبان کافی می باشد.

۱-۴-۳ آشکارسازی بینا

در این روش‌ها برای آشکارسازی به سیگنال اصلی نیاز است و بدون داشتن سیگنال میزبان اصلی استخراج پیام در سمت گیرنده عملی نیست.

روش‌هایی که از آشکارسازی کور استفاده می کنند کاربردهای بیشتری نسبت به روش‌های بینا و نیمه کور دارند. اما با توجه به اینکه در آشکارسازی بینا و نیمه کور از اطلاعات سیگنال اصلی برای استخراج پیام استفاده می شود این گونه روش‌ها در برابر حملات مقاوم ترند.

۱-۵ ویژگی‌های مورد نیاز در سیستم‌های پنهان سازی ونحوه‌ی اندازه‌گیری آنها

در طراحی سیستم‌های پنهان سازی اطلاعات چندین ویژگی مختلف باید در نظر گرفته شود. اهمیت هر یک از این ویژگی‌ها در کاربردهای مختلف متفاوت است. چند مورد از این ویژگی‌ها در ادامه بیان شده است.

۱-۵-۱ شفافیت[۱]

اگر در یک سیستم مخفی سازی اطلاعات، مخفی کردن پیام تاثیری بر کیفیت سیگنال اصلی نداشته باشد یعنی تفاوت محسوسی قبل و بعد از درج پیام در سیگنال اصلی ایجاد نشود، آن سیستم دارای شفافیت است. به عنوان مثال اگر سیگنال میزبان سیگنال گفتار باشد کیفیت شنیداری سیگنال گفتار

[۱] Transparency

اصلی و سیگنال گفتار حاوی پیام نباید برای شنونده معمولی تفاوت محسوسی داشته باشد. این خصوصیت در تمام کاربردها از اهمیت بالایی برخوردار است.

اندازه گیری شفافیت یکی از مشکل‌ترین مسائل می باشد. دلیل این امر این است که تشخیص شفافیت به عهده‌ی انسان است و مدل کردن سیستم پیچیده‌ی بینایی و شنوایی انسان بسیار دشوار می‌باشد. روش‌های ارزیابی را به دو دسته‌ی کلی کیفی و کمی می‌توان تقسیم نمود.

۱-۵-۱-۱ روش‌های کیفی

یکی از این گونه روش‌ها میانگین امتیاز نظرسنجی[۱] می‌باشد که بیشتر برای سیگنال‌های گفتار به کار می‌رود[۲]. در این روش از تعدادی افراد خواسته می‌شود که به سیگنال گفتار پنهان سازی شده و اصلی گوش دهند و به سیگنال پنهان سازی شده از ۱ تا ۵ امتیازی را نسبت دهند. سپس مقادیر به دست آمده را میانگین گیری می‌کنند و نتیجه‌ی نهایی را به عنوان کیفیت سیگنال پنهان سازی شده گزارش می‌کنند. ویژگی‌های هر امتیاز در جدول ۱-۱ آمده است.

جدول ۱-۱ نحوه امتیاز دهی در روش MOS

امتیاز	اختلال	کیفیت
۵	غیر قابل درک	عالی
۴	قابل درک، آزار دهنده	خوب
۳	کمی آزار دهنده	متوسط
۲	آزار دهنده	ضعیف
۱	خیلی آزار دهنده	بد

همچنین آزمون دیگری تحت عنوان تست تمیز دهنده در هر دو سیستم پنهان سازی صوت و تصویر کاربرد دارد. در این آزمون از تعدادی افراد خواسته می‌شود سیگنال پنهان سازی شده و اصلی را دیده یا شنیده و تشخیص دهند کدام پنهان سازی شده و کدام اصلی است. معمولا نرخ ۵۰٪ در این مورد نشان دهنده کیفیت بالای سیگنال پنهان سازی شده است.

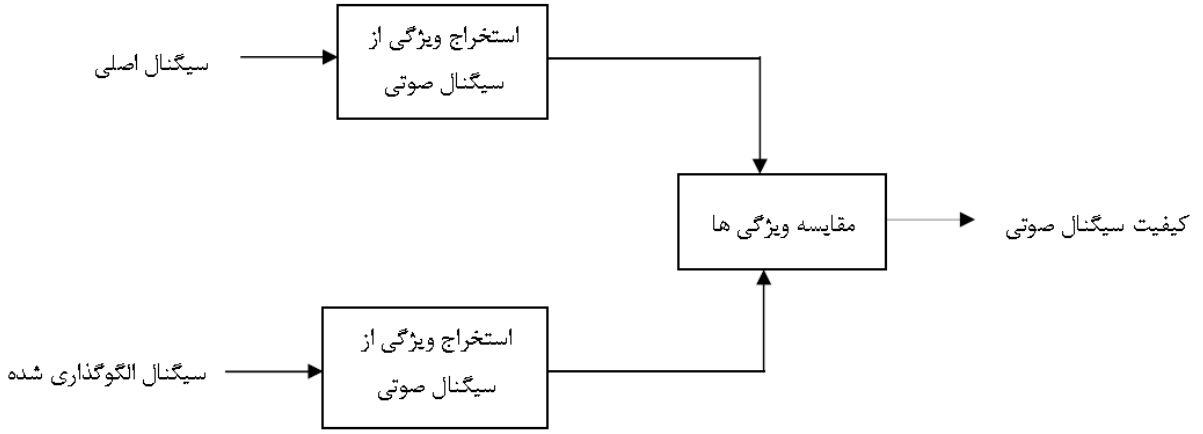
[۱] Mean Opinion Score (MOS)

۱-۵-۱-۲ روش‌های کمی

از بین معیارهای سنجش کمی، SNR و PSNR در سیستم‌های پنهان سازی رایج‌تر می‌باشند. در سیستم‌های صوتی بیشترSNR و در سیستم‌های تصویری PSNR بیشتر به کار می روند. ولی با این وجود هیچ کدام از این معیارها مقدار دقیقی برای ارزیابی شفافیت نمی‌دهند. دلیل این امر این است که در هیچ کدام خصوصیات بینایی و شنوایی انسان در نظر گرفته نشده است.

برای مثال در نظر بگیرید که نمونه‌های یک سیگنال صوتی را قرینه کنید. این کار هیچ تأثیری در نحوه‌ی به گوش رسیدن و کیفیت صوت نخواهد داشت. در صورتی که اگر از معیار SNR استفاده شود اغتشاش بزرگی نشان داده خواهد شد. در صورتی که از نظر سیستم شنوایی انسان اغتشاشی وجود نداشته است.

به همین دلیل روش‌هایی پیشنهاد شده‌اند که مقادیر نهایی آنها به نتایج ارزیابی کیفی نزدیک‌تر باشد. بلوک دیاگرام این روش‌ها در شکل ۱-۲ مشاهده می‌شود. در این گونه روش‌ها تعدادی ویژگی (بسته به روش) از سیگنال‌های اصلی و تخریب شده استخراج می شود. سپس این ویژگی‌ها با هم مقایسه شده و نهایتاً خروجی اعلام می‌گردد. ویژگی‌های شنیداری و دیداری انسان در این گونه روش‌ها در نظر گرفته می شود و به این دلیل نتایج آنها به نتایج روش‌های کیفی نزدیک‌تر است.



شکل ۱-۲ بلوک دیاگرام ارزیابی کیفیت

نمونه‌ای از این روش‌ها در صوت و تصویر Perceptual Evaluation of audio Quality(PEAQ) [۳]و Perceptual Evaluation of Speech Quality(PESQ) [۴] می‌باشد. اساس عملکرد این دو روش در بلوک دیاگرام نشان داده شده در شکل ۱-۲ می‌باشد. هر دو روش خروجی را به صورت امتیاز کیفی تفاضل می‌دهند. روش PEAQ مقادیری بین ۰ تا ۴- می‌دهد که ۰ نشان دهنده ی یکسان بودن کیفیت دو سیگنال و ۴- نشان دهنده ی متمایز بودن آنهاست. همچنین PESQ مقادیری بین۰/۵ تا ۴/۵ می دهد که ۴/۵ نشان دهنده ی یکسان بودن، و اعداد کوچکتر نشان دهنده‌ی کیفیت های متمایز است[۵].

۱-۵-۲ پایداری[۱] (مقاومت)

پایداری یک سیستم پنهان سازی بدین معنی است که پیام پنهان شده، در مقابل اعمال تغییرات ناخواسته و غیرعمدی مانند نویز در طول مسیر انتقال و یا اعمال تغییرات عمدی که به منظور تغییرات پیام یا از بین بردن آن توسط حمله کننده فعال انجام می گیرد مانند فشرده سازی مقاومت لازم را داشته باشد.

در بعضی از کاربردهای پنهان سازی اطلاعات، وظیفه واحد تشخیص پیام، اعلام وجود یا عدم وجود پیام معرفی شده در درون یک داده است. حال زمانی که پیام معرفی شده در داده مورد نظر، وجود نداشته باشد و واحد تشخیص پیام، به غلط وجود پیام در داده را اعلام نماید اعلام مثبت غلط روی داده است. هم چنین زمانی که پیام معرفی شده در داده موجود باشد و واحد تشخیص به وجود آن پی نبرد و به اشتباه عدم وجود پیام در داده را اعلام نماید اعلام منفی غلط، رخ داده است. محاسبه احتمال بروز این دو اشتباه برای روش‌های مختلف به صورت تحلیلی انجام می شود و روش‌هایی که این دو احتمال برای آنها کمتر محاسبه شود، روش‌های دقیق تری خواهند بود.

[۱] Robustness

در بعضی از کاربردهای دیگر هدف پنهان سازی بیت در داده می باشد. در این سیستم‌ها برای تست پایداری پس از اعمال حملات پیام را استخراج می کنند و با پیام اصلی مقایسه می کنند. سپس خطای نرخ بیت را به صورت زیر محاسبه می کنند.

$$BER = \frac{\text{تعداد بیت های تشخیص داده شده به طور نادرست}}{\text{تعدادکل بیت ها}} \quad (۱-۱)$$

همچنین گاهی اوقات به جای اعلام BER میزان همبستگی پیام استخراج شده و اصلی را به دست می آورند. که به صورت زیر محاسبه می شوند:

$$NC = \frac{w.\hat{w}^T}{w.w^T} \quad (۱-۲)$$

که در رابطه‌ی بالا w پیام اصلی و $\hat{w}$ پیام استخراج شده می باشد.

۱-۵-۳ ظرفیت[۱]

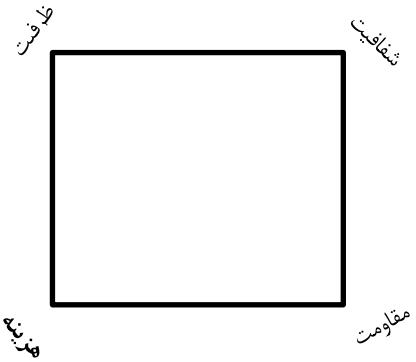
در یک سیستم مخفی سازی هر چقدر بتوان پیام بیشتری را در میزبان مخفی نمود این سیستم مناسب‌تر خواهد بود. در سیستم‌های صوتی به تعداد بیت‌های نهان شده در یک ثانیه از سیگنال صوتی گفته می شود. برای مثال می‌توان گفت، نرخ پنهان سازی داده در این سیستم ۱۲۸ بیت بر ثانیه می‌باشد.

۱-۵-۴ هزینه‌ی پردازش

هزینه‌ی پردازش به معنای مدت زمان لازم برای نهان سازی و استخراج داده می باشد. این ویژگی یکی از عامل‌های مهم در طراحی یک سیستم پنهان سازی می باشد.

همواره مصالحه‌ای بین ویژگی‌های ذکر شده در بالا وجود دارد و افزایش یکی موجب کاهش بقیه می-شود به طوری که می توان آنها را در چهار راس یک مربع قرار داد. هر بار بسته به کاربرد به یکی از رئوس این مربع نزدیک می شویم. البته لازم به ذکر است که هدف سیستم‌های پنهان سازی بهبود تمام ویژگی‌های فوق است.

[۱] Bit rate or Capacity



شکل ۱-۱ مصالحه بین شفافیت، ظرفیت، مقاومت و هزینه در یک سیستم پنهان سازی

۱-۶ سازماندهی پایان نامه

در این پایان نامه در فصل‌های آینده به صورت زیر عمل خواهد شد.

در فصل دوم، چند سیستم مختلف پنهان سازی اطلاعات درون گفتار، روش‌های مورد استفاده در این سیستم‌ها و روابط موجود در آنها مورد بررسی قرار گرفته خواهد شد. این سیستم‌ها عبارتند از جایگزینی بیت کم ارزش، طیف گسترده، پنهان سازی در پژواک، مدولاسیون اندیس کوانتیزاسیون، کدینگ فاز، تغییر گام سیگنال گفتار و مدل سینوسی.

در فصل سوم تبدیلی به نام تبدیل فوریه - بسل معرفی می‌شود. همچنین روابط مربوط به آن و چگونگی محاسبه آن به کمک تبدیل فوریه آورده شده است. سپس کاربردهایی که تا کنون از تبدیل در مقالات سال های اخیر وجود داشته نیز آورده شده است.

در فصل چهارم به طرح سه روش پیشنهادی پرداخته می‌شود. در این سه روش از هر چهار تبدیل فوریه – بسل، کسینوسی، فوریه و موجک به جهت مقایسه استفاده شده است. برای مقایسه این چهار تبدیل از میزان نرخ خطای بیت، همبستگی بین نهان نگاره اصلی و استخراج شده، شفافیت و پایداری در برابر حملات (فشرده سازی و افزودن نویز) کمک گرفته می‌شود.

در نهایت در فصل آخر به بیان نتیجه گیری کلی و طرح پیشنهاداتی برای ادامه کار در آینده پرداخته خواهد شد.

فصل دوم:

معرفی روش‌های پنهان سازی اطلاعات

۲-۱ مقدمه

در این بخش بعضی از روش‌های پایه‌ای پنهان سازی اطلاعات را شرح داده می‌شود. البته کارهای انجام شده در این زمینه زیاد است، که به ذکر تعدادی از آنها را بسنده می‌گردد.

۲-۲ پنهان‌سازی به روش جایگزینی بیت کم ارزش[۱]

این روش برای پنهان کردن هر نوع از اطلاعات چندرسانه‌ای در نوع دیگر قابل استفاده است. در این روش کم ارزش‌ترین بیت یا بیت‌های هر نمونه از سیگنال اصلی با یک یا چند بیت از اطلاعات نهان-نگاره جایگزین می‌شود. واضح است که در این روش و روش‌هایی که در ادامه گفته می‌شود اطلاعات از نوع دیجیتال است. آشکارسازی به صورت کور و با خواندن بیت یا بیت‌های کم ارزش سیگنال پنهان-سازی شده‌ی دریافتی انجام می‌شود.

از مزیت‌های این روش توانایی پنهان‌سازی حجم اطلاعات زیاد می‌باشد. به طور ایده‌آل در هر یک کیلوهرتز از نمونه‌های سیگنال Kbps ۱ اطلاعات پنهان می‌شود. با این حال این روش نسبت به انواع حملات غیر مقاوم است که این می‌تواند با توجه به کاربرد هم عیب هم حسن محسوب شود. در کاربرد اعتبارسنجی می‌توان کوچکترین دستکاری در اسناد و محل دستکاری را تشخیص داد. در سایر کاربردها که پایداری در برابر حملات مهم است، با تکرار نهان‌نگاره و در نتیجه کاهش نرخ پنهان‌سازی می‌توان مقاومت را تا اندازه‌ای افزایش داد.

اخیراً الگوریتمی[۶] نیز برای افزایش پایداری این روش برای نرخ بیت بالا پیشنهاد شده است.

۲-۳ پنهان سازی به روش طیف گسترده[۲]

این روش [۷] نسبت به روش قبل از پایداری بیشتری برخوردار است. دراین روش یک دنباله‌ی شبه نویز مدوله شده با سیگنال پیام، در طول سیگنال اصلی گسترده می‌شود. سیگنال اصلی و سیگنال

---

[۱] Least Significant Bit

[۲] Spread Spectrum

پیام می‌توانند حوزه‌ی زمان و یا هر حوزه‌ی تبدیلی باشد که مرسوم‌ترین تبدیل‌ها که تا کنون استفاده شده، تبدیل کسینوسی گسسته، تبدیل فوریه‌ی گسسته و تبدیل موجک گسسته می‌باشد.

لازم به ذکر است دنباله شبه نویز دنباله‌ای است که از نظر ظاهر شبیه به نویز بوده ولی وابسته به پارامترهایی می‌باشد که با دانستن آنها می‌توان دوباره آن را تولید کرد. دنباله شبه نویز می‌تواند یک سیگنال آشوب باشد و یا توسط یک کلید مخفی تولید شود.

۲-۳-۱ پنهان‌سازی نهان‌نگاره

اطلاعات بیتی نهان‌نگاره یا معادل دو قطبی آن $b = \{-1,1\}$ ابتدا توسط یک دنباله‌ی شبه نویز r(n) طبق رابطه‌ی (۲-۱) مدوله می‌شود. سپس توسط یک شاخصه‌ی کنترل α به سیگنال اصلی اضافه می‌شود.

مقدار α بر طبق یک رابطه‌ی دوطرفه بین دو معیار مقاومت و کیفیت سیگنال پنهان‌سازی شده تنظیم می‌شود. هرچه مقدار α بیشتر باشد، مقاومت سیستم بالاتر و شفافیت پایین‌تر بوده و هرچه مقدار α کمتر باشد، مقاومت سیستم پایین‌تر و شفافیت بالاتر می‌رود. رابطه‌ی سیگنال پنهان‌سازی شده به صورت رابطه‌ی (۲-۲) می باشد[۸].

$$w(n) = br(n) \tag{۱-۲}$$

$$x(n) = s(n) + \alpha w(n) \tag{۲-۲}$$

۲-۳-۲ آشکارسازی نهان‌نگاره

برای آشکارسازی نهان‌نگاره از رابطه‌ی همبستگی خطی استفاده می‌شود. در آشکارساز دنباله‌ی شبه نویز توسط کلید مخفی که در اختیار داریم دوباره تولید می‌شود. استخراج نهان‌نگاره توسط رابطه‌ی همبستگی بین سیگنال پنهان‌سازی شده و دنباله‌ی شبه نویز طبق رابطه‌ی (۲-۳) بدست می‌آید.

$$c = \frac{1}{N}\sum_{i=1}^{N} x(i)r(i) \tag{۳-۲}$$

که N نشان دهنده‌ی طول سیگنال می‌باشد. رابطه‌ی (۲-۳) را با جایگذاری رابطه‌ی (۲-۲) در آن، می‌توان به صورت رابطه‌ی (۲-۴) نوشت.

$$c = \frac{1}{N}\sum_{i=1}^{N} s(i)r(i) + \frac{1}{N}\sum_{i=1}^{N} \alpha b r^2(i) \qquad (۴-۲)$$

اگر بین سیگنال میزبان و دنباله‌ی شبه نویز وابستگی وجود نداشته باشد بخش اول رابطه‌ی (۲-۴) حذف می‌شود. اگر چه در عمل این بخش کاملاً حذف نمی‌شود. یک راه حل ممکن برای حذف بخش این است که با یک پیش پردازش، سیگنال میزبان را از دنباله‌ی شبه نویز جدا می‌کنیم. روش‌های پیش پردازش شامل فیلتر بالا گذر، کدینگ پیش بینی خطی و استفاده از فیلترهای whitening در این زمینه ارائه شده است[۹].

چنین پیش پردازش‌هایی باعث می‌شود که بخش اول رابطه‌ی (۲-۴) تقریباً از بین برود. استخراج اطلاعات نهان‌نگاره توسط یک آزمون فرض با استفاده از پارامترهای اندازه‌ی همبستگی C و مقدار آستانه یτ ، طبق رابطه‌ی (۲-۵) صورت می‌گیرد.

$$m = \begin{cases} 1 & if\ c > \tau \\ 0 & if\ c \leq \tau \end{cases} \qquad (۵-۲)$$

مقدار آستانه تأثیر مستقیم بر روی احتمال مثبت غلط و منفی غلط دارد. مثبت غلط، خطایی است که سیگنال حامل اطلاعات نهان‌نگاره نباشد ولی ما آن را یک سیگنال پنهان‌سازی شده تشخیص دهیم از طرف دیگر منفی غلط، خطایی است که آشکارساز وجود نهان‌نگاره در یک سیگنال پنهان‌سازی شده را تشخیص ندهد.

۲-۴پنهان‌سازی به روش مدولاسیون اندیس کوانتیزاسیون

به طور کلی این روش [۱۰] برای سیگنال میزبان کوانتیزاسیون مجدد انجام می‌دهد. گام این کوانتیزاسیون مجدد با توجه به نهان نگاره تعیین می‌شود. یعنی اگر بیت نهان‌نگاره صفر بود، از یک گام کوانتیزاسیون استفاده می‌شود و اگر یک بود از گام دیگری استفاده می‌شود. در مجموع دو دسته-

ی گسسته‌ساز[۱] انتخاب می‌شود که هر دسته به یکی از بیت‌های صفر یا یک از نهان‌نگاره اختصاص داده می‌شود.

برای آشکارسازی نیز با توجه به کوتاهترین فاصله‌ی سیگنال پنهان‌سازی شده از سطوح گسسته‌سازها، تصمیم گیری برای استخراج بیت صفر یا یک انجام می‌شود. این روش در تصویر نیز کاربرد دارد.[۱۱]

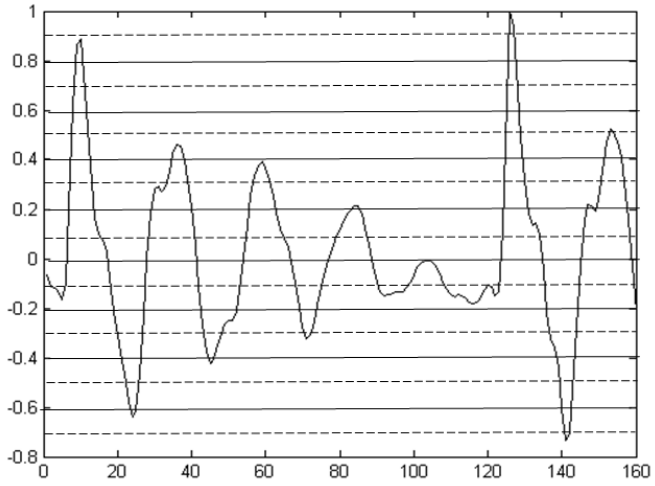
۲-۴-۱ پنهان‌سازی نهان‌نگاره

به طور کلی پنهان‌سازی در این روش را می‌توان با رابطه‌ی زیر نشان داد.

$$y_i = Q_{b_i}(x_i) = \Delta round\left(\frac{x_i+\frac{b_i\Delta}{2}}{\Delta}\right) - \frac{b_i\Delta}{2} \qquad (۱۶-۲)$$

که در رابطه‌ی فوق $b_i \in \{0,1\}$، هر یک از بیت‌های نهان‌نگاره می‌باشد. x_i یک نمونه از سیگنال میزبان است که با توجه به پیام کوانتیزه می‌شود و Δ گام کوانتیزاسیون نامیده می‌شود.

شکل ۲-۱ نمایشی برای توضیح فرمول بالاست. این شکل یک فریم از سیکنال گفتار می‌باشد. در این شکل دو گام کوانتیزاسیون به صورت خط و نقطه‌چین وجود دارد. برای مثال در هر نمونه جهت پنهان‌سازی بیت یک، از خطوط، و برای پنهان‌سازی بیت صفر، از نقطه‌چین‌ها جهت کوانتیزاسیون مجدد استفاده می‌گردد.



شکل ۲-۱ درج پیام به روش مدولاسیون اندیس کوانتیزاسیون

[۱] Quantizer

۲-۴-۲ آشکارسازی نهان‌نگاره

جهت آشکارسازی نهان‌نگاره، به ازای بیت صفر و یک، دو سیگنال $r_{i,0}$ و $r_{i,1}$ از سیگنال پنهان‌سازی شده r_i محاسبه می‌شوند.

$$r_{i,0} = Q_0(r_i) = \Delta ruond\left(\frac{r_i}{\Delta}\right) \quad (۱۷-۲)$$

$$r_{i,1} = Q_1(r_i) = \Delta ruond\left(\frac{r_i+\frac{\Delta}{2}}{\Delta}\right) - \frac{\Delta}{2} \quad (۱۸-۲)$$

پس از محاسبه‌ی دو سیگنال فاصله‌ی اقلیدسی هرکدام از آنها با سیگنال دریافتی محاسبه می‌شود. آشکارسازی با مقایسه‌ی فاصله‌های به دست آمده انجام می‌شود. اندیس بردار با فاصله‌ی کمتر نشان دهنده‌ی بیت پنهان شده می‌باشد. در حالتی که پنهان ساز در هر نمونه از سیگنال اصلی یک بیت پنهان کرده باشد، آشکارسازی را می‌توان با رابطه‌ی زیر نمایش داد.

$$\hat{b}_i = \arg min_{b_i}\left\|r_i - r_{i,b_i}\right\|^2 \quad (۱۹-۲)$$

باید توجه داشت که این کوانتیزاسیون خطی بوده و فاصله سطوح برای همه نمونه‌ها یکسان است. در [۱۲] با استفاده از کوانتیزاسیون غیر خطی پارامترهای نرخ خطای بیت و شفافیت بهبود داده شده است.

۲-۵ پنهان سازی در پژواک[۱]

در این روش [۱۳] با استفاده از اضافه کردن پژواک با تاخیرهای متفاوت، پنهان‌سازی نهان‌نگاره در سیگنال گفتار انجام می‌شود. به این صورت که برای پنهان کردن بیت‌های صفر و یک نهان‌نگاره از تاخیرهای متفاوتی استفاده می‌شود.

۲-۵-۱ پنهان‌سازی نهان‌نگاره

به طور کلی اضافه کردن پژواک به صورت رابطه ۲-۶ می‌باشد.

[۱] Echo hiding

$$X_w(t) = X_0(t) + \alpha . X_0(t - \Delta t) \quad (۶-۲)$$

در رابطه‌ی (۲-۶)، Δt برای پنهان کردن اطلاعات به کار می‌رود. همان گونه که می‌بینیم Δt و α پارامترهایی هستند که با تنظیم صحیح آنها می توان از شنیده نشدن سیگنال پیام مطمئن شد. به طور کلی رابطه‌ی (۲-۶) می تواند به شکل زیر نوشته شود.

$$X_w(t) = \sum_{k=0}^{N} \alpha_k . X_0(t - \Delta t_k) \quad (۷-۲)$$

که در رابطه‌ی (۲-۷)، Δt_0 برابر صفر و α_0 برابر یک می باشد و N برابر با تعداد پژواک‌های اضافه شده می‌باشد. رابطه‌ی(۲-۷) را می‌توان به صورت رابطه‌ی (۲-۸) در زیر نوشت.

$$X_w(t) = X_0(t) * h(t) \quad (۸-۲)$$

که در آن تابع $h(t)$ به صورت رابطه‌ی (۲-۹) تعریف شده است.

$$\mathrm{h}(t) = \sum_{k=0}^{N} \alpha_k . \delta(t - \Delta t_k) \quad (۹-۲)$$

اگر رابطه‌ی (۲-۸) را به حوزه‌ی فرکانس ببریم، کانولوشن به ضرب تبدیل می شود.

$$X_w(\omega) = X_0(\omega) H(\omega) \quad (۱۰-۲)$$

برای پنهان کردن در این روش ابتدا سیگنال گفتار به M بلوک به نام X_j تقسیم می‌شود. هر بیت از سیگنال پیام در هر بلوک پنهان می شود. برای هر بلوک، پژواک برای پنهان کردن صفر و یک، با تاخیر متناظر و ضریب تضعیف متناظر به صورت رابطه‌ی (۲-۱۱) ساخته می شود.

$$w_k(t) = \alpha_k . X_0(t - \Delta t_k) \qquad k = 0,1 \quad (۱۱-۲)$$

سپس دو سیگنال مدوله کننده برای بیت‌های صفر و یک ساخته می شود.

$$m_0(t) = \sum_{j=0}^{M-1} (1 - b_j) rect_j(t) \quad (۱۲-۲)$$

$$m_0(t) + m_1(t) = 1 \quad \forall t \quad rect_j(t) = \begin{cases} 1 & t_j < t < t_{j+1} \\ 0 & otherwise \end{cases} \quad b_j = I(j) \quad (۱۳-۲)$$

که در رابطه‌ی (۲-۱۳)، $I(j)$ نمونه ی jام پیام است. با استفاده از سیگنال‌های به دست آمده، سیگنال پنهان سازی شده به صورت زیر به دست می آید.

$$X_w(t) = X_0(t) + m_0(t) w_0(t) + m_1(t) w_1(t) \quad (۱۴-۲)$$

۲-۵-۲ آشکارسازی نهان‌نگاره

برای آشکارسازی تابع h(t) محاسبه می‌شود تا با استفاده از آن Δt_k و در نتیجه‌ی آن بیت‌های پنهان شده به دست آید. با توجه به رابطه‌ی (۲-۱۰) با تقسیم $X_w(\omega)$ بر $X_0(\omega)$، $H(\omega)$ محاسبه شده و سپس با گرفتن معکوس تبدیل فوریه، h(t) محاسبه می‌شود. آشکارسازی به این صورت به سیگنال اصلی نیاز دارد. آشکارسازی در این روش می تواند به شیوه‌ی deconvolution homomorphic نیز انجام شود که در این صورت برای جداسازی سیگنال اصلی و پژواک نیازی به سیگنال اصلی نمی‌باشد. در این روش ابتدا هر بلوک به حوزه‌ی کپستروم[۱] برده می شود یعنی با به کار بردن تابع لگاریتم، ضرب در رابطه‌ی (۲-۱۰) به جمع تبدیل می‌شود.

$$C_w(q) = IFFT(\log|X_0(\omega)H(\omega)|) = X_0(q) + H(q) \qquad (۱۵-۲)$$

که عبارت بالا تابعی از کیوفرنسی[۲] می باشد. با توجه به رابطه‌ی (۲-۱۵) سیگنال اصلی و پژواک نهان شده روی محور کیوفرنسی جدا شده‌اند.

سپس خود همبستگی C_w در حوزه کپستروم محاسبه می شود. با استفاده از پیک خود همبستگی اندازه Δt تعیین شده و به این صورت تصمیم گیری روی بیت پنهان شده انجام می گیرد.

۲-۶ پنهان سازی به کمک تغییر گام سیگنال گفتار

این روش [۱۴] بر مبنای تغییر گام قسمت‌های صدادار از سیگنال گفتار است که دارای خاصیت شبه پریودیک است. اطلاعات گام معمولا توسط کدک‌های نظیر AMR ، GSM-06.10 و QCELP محافظت می شوند. این الگوریتم با استفاده از شکل ۲-۲ که شامل قسمتی از سیگنال گفتار است توضیح داده خواهد شد. این سیگنال از سه بخش سکوت[۳] بی صدا[۴] و صدادار[۵] تشکیل شده است. بخش صدادار را می توان به عنوان بخشی که خاصیت پریودیک داشته و انرژی آن از یک حد آستانه‌ی

---

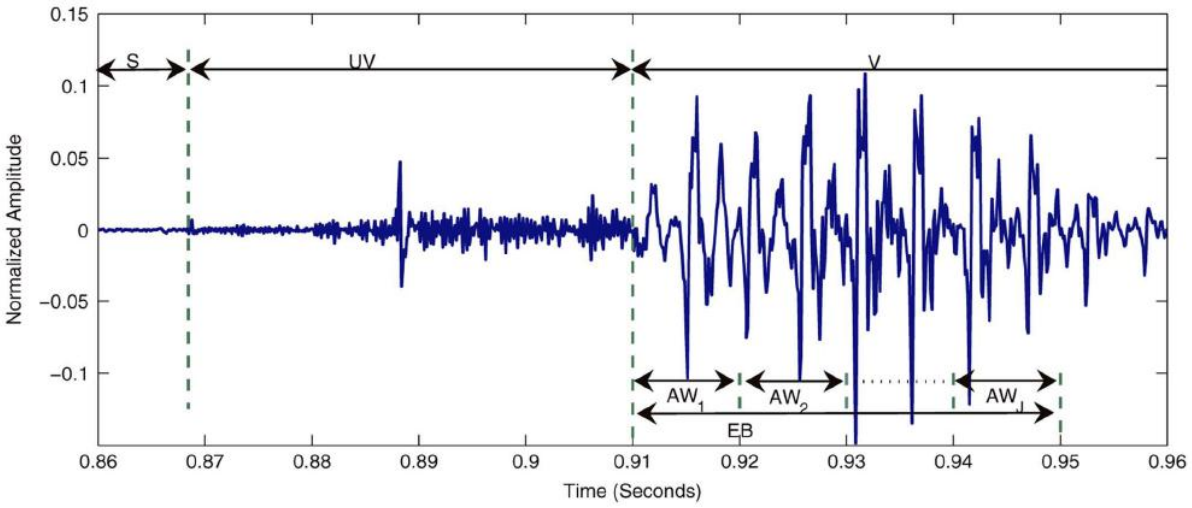
۱ Cepstrum
۲ Quefrency
۳ Silence (S)
۴ Unvoiced (UV)
۵ Voiced (V)

مشخص بالاتر است، تشخیص داد. در سگمنت صدادار، با پردازش پنجره‌های آنالیز[1] گام پریودها تخمین زده می شود. این پنجره‌ها بدون همپوشانی و دارای طول L نمونه هستند. برای پنهان کردن یک بیت در هر سگمنت صدادار، ابتدا این سگمنت به J پنجره ی آنالیز تقسیم بندی می شود. گام پریودها $\{p_j\}_{i=1}^{j}$ برای هر پنجره‌ی آنالیز تعیین می‌شود. سپس میانگین گام پریودها برای هر سگمنت محاسبه می شود.

$$p_{avg} = \sum_{i=1}^{J} \frac{p_i}{J} \qquad (۲۰-۲)$$

سپس میانگین گام پریودها با استفاده از روش کوانتیزاسیون اندیس مدولاسیون برای پنهان کردن بیت iام کوانتیزه می شود.

$$p'_{avg} = Q_{b_i}(p_{avg}) \qquad (۲۱-۲)$$



شکل ۲-۲ قسمتی از یک سیگنال گفتار که به بخش‌های سکوت، صدادار و بی‌صدا تقسیم شده است.[۱۲]

در رابطه‌ی ۲-۲۱ ، $b_i \in \{0,1\}$ و $Q_{b_i}(x)$ نیز رابطه کوانتیزه کننده است که سیگنال اصلی را با توجه به اطلاعات پیام b_i ، کوانتیزه می‌کند. سپس گام پریودها با استفاده از رابطه زیر تغییر داده می شود.

$$p'_m = p_m + (p'_{avg} - p_{avg}) \qquad (۲۲-۲)$$

[1] Analysis Windows (AW)

در نهایت با استفاده الگوریتمی به نام PSOLA سگمنت‌ها کنار هم قرار گرفته و سیگنال پردازش شده بازسازی می‌شود. در گیرنده بدون استفاده از سیگنال اصلی گفتار، مراحل بالا تکرار می‌شود و برای هر سگمنت مقادیر میانگین گام پریود تغییر داده شده مشخص می‌شود. سپس با استفاده از مقادیر میانگین گام پریود تغییر داده شده بیت‌های اطلاعات استخراج می شود. همانطور که گفته شد چون اطلاعات گام معمولا توسط کدک‌هایی نظیر AMR و GSM-06.10 و QCELP محافظت می شوند این روش به کدک‌های گفتار با نرخ بیت پایین مقاوم است اما ظرفیت پنهان‌سازی در این روش حدود هشت بیت بر ثانیه می‌باشد که بسیار پایین است.

۲-۷ پنهان‌سازی به روش کدینگ فاز

از آنجا که شنوایی انسان به دامنه سیگنال گفتار حساس بوده نسبت به اندازه‌ی دقیق فاز حساس نمی‌باشد، می‌توان از فاز سیگنال گفتار به عنوان یک میزبان مناسب برای پنهان کردن اطلاعات نهان-نگاره استفاده کرد. در این روش ابتدا سیگنال گفتار فریم بندی شده سپس با یک تبدیل مناسب اطلاعات نهان‌نگاره در بخش فاز ذخیره می‌شود. دو روش برای این کار وجود دارد که خلاصه آن [۱۲] در ادامه توضیح داده می‌شود.

۲-۷-۱ روش بِندِر[۱]

این روش[۱۵]، اولین روشی است که از کدینگ فاز برای وارد کردن اطلاعات نهان‌نگاره استفاده کرده است که خلاصه آن همان گونه که در آورده شده است بیان می‌شود.

۱- ابتدا سیگنال گفتار میزبان $S[i]\,,(0 \le i \le I-1)$ به فریم‌های $S_n[i]\,,(0 \le i \le N-1)$ تقسیم می شود.

۲- برای هر فریم تبدیل فوریه K نقطه ای انجام می شود که $K = I/N$ می باشد و ماتریسی از فاز و اندازه ی دامنه برحسب فرکانس بدست می آید.$(0 \le k < K-1)$

۳- اختلاف فاز برای هر دو فریم متوالی طبق رابطه ی زیر بدست می آید.

[۱] Bender

$$\Delta\varphi_{n+1}[\omega_k] = \varphi_{n+1}[\omega_k] - \varphi_n[\omega_k] \quad (۲۳-۲)$$

۴- اطلاعات بیتی نهان‌نگاره به صورت فاز ثابت $\frac{\pi}{2}$ یا $-\frac{\pi}{2}$ در فریم اول ذخیره می شود.

$$\varphi_0 = \varphi'_{data} \quad (۲۴-۲)$$

۵- برای تولید دوباره فاز در فریم‌هایی بعدی روابط زیر را انجام می دهیم.

$$\begin{aligned}
&\varphi'_1[\omega_k] = \Delta\varphi_1[\omega_k] + \varphi'_0[\omega_k] \\
&\varphi'_n[\omega_k] = \Delta\varphi_n[\omega_k] + \varphi'_{n-1}[\omega_k] \quad (۲۵-۲) \\
&\varphi'_N[\omega_k] = \Delta\varphi_N[\omega_k] + \varphi'_{N-1}[\omega_k]
\end{aligned}$$

۶- دامنه‌های اصلی $A_n(w_k)$ و فاز اصلاح شده $\varphi'_n(w_k)$ با استفاده از تبدیل معکوس به حوزه‌ی زمان بر می‌گردند و سیگنال پنهان‌سازی شده بدست می‌آید.

برای آشکارسازی، سیگنال دریافتی را به بلوک‌هایی با طول N تقسیم می‌کنیم و فریم اول را به حوزه‌ی فرکانس می‌بریم سپس با خواندن فاز فریم اول اطلاعات نهان‌نگاره استخراج می‌شود.

۲-۷-۲ روش گرول[۱]

در این روش [۱۵] نشان داده شده اگر تغییر فاز باعث تغییر زیاد در پوش سیگنال گفتار نشود این تغییرات توسط گوش انسان قابل درک است. در روش مدولاسیون فاز، ابتدا سیگنال وارد حوزه فرکانس می‌شود. سپس فاز نمونه‌هایی که می‌خواهیم اطلاعات نهان‌نگاره را در آنها درج کنیم، تغییر می‌دهیم. این عمل نباید باعث تغییرات زیاد در پوش سیگنال شود. در روشی که ارائه شده تعدادی نمونه‌ی فرکانسی به صورت تصادفی انتخاب می شوند سپس از بین آنها تعدادی نمونه‌ی فرکانسی با مقدار دامنه‌ی کمتر انتخاب می‌شوند و با استفاده از روش دو جزئی، اطلاعات در بخش فاز آنها وارد می‌شود.

[۱] Gruhl

برای درج نهان‌نگاره ابتدا سیگنال گفتار را فریم بندی می کنیم و از آن تبدیل فوریه می گیریم، تا اطلاعات فاز سیگنال را بدست آوریم. الگوریتم درج نهان‌نگاره در هر فریم را می توانیم به صورت زیر خلاصه کنیم:

۱- توسط یک تولید کننده الگو شبه تصادفی، N نمونه به صورت تصادفی بین صفر تا ۰/۵ انتخاب می کنیم.

۲- اعداد تولید شده را در فرکانس نمونه برداری ضرب می کنیم تا فرکانس‌های تصادفی بین صفر تا حداکثر فرکانس سیگنال بدست آوریم.

۳- فرکانس‌ها را به ترتیب اندازه دامنه از کوچک به بزرگ مرتب می کنیم.

۴- تعداد M نمونه اول را انتخاب می کنیم. نیمه‌ی اول نمونه‌ها را در یک گروه و نیمه‌ی دیگر را در گروه دیگر قرار می دهیم.

۵- برای اینکه مقدار بیت صفر را وارد کنیم فاز نمونه‌های گروه اول را α و فاز نمونه‌های گروه دوم را $-\alpha$ قرار می دهیم. برای وارد کردن بیت یک برعکس این کار را انجام می دهیم.

در مرحله استخراج نهان‌نگاره سه مرحله اول مشابه مراحل درج نهان‌نگاره می باشد. مراحل بعد به صورت زیر خلاصه می شود:

۴- تعداد M نمونه‌ی اول را انتخاب می کنیم. میانگین فاز را برای نمونه‌های گروه اول و نمونه‌های گروه دوم بدست می آوریم.

۵- میانگین فاز گروه اول را از میانگین فاز گروه دوم کم می کنیم.

۶- این مقدار را با یک آستانه مقایسه می کنیم. و طبق رابطه‌ی زیر نتایج بدست می آید.

در روش‌های معمول برای فریم بندی سیگنال گفتار، از پنجره‌ها با طول یکسان استفاده می کنند. اما در این روش از پنجره‌ها با طول متغیر استفاده می‌شود. ابتدا و انتهای پنجره‌ها بگونه‌ای طراحی می-شوند که شامل محتویات فرکانسی با دامنه‌ی بالا باشند. این کار می تواند با تبدیل فوریه‌ی زمان کوتاه

انجام گیرد. زیرا این تبدیل تغییر محتویات فرکانسی را در حوزه زمان نشان می‌دهد. استفاده از این پنجره‌ها باعث مقاومت این روش در برابر حملات غیر همزمانی می شود.

۲-۸ پنهان‌سازی به روش مدل سینوسی

در این روش [۱۶] از مدولاسیون مسیرهای فرکانسی مدل سینوسی سیگنال گفتار، برای وارد کردن اطلاعات نهان‌نگاره استفاده می‌شود. اما برخلاف مدل سینوسی که در بخش قبل توضیح دادیم، در این روش از یک مدل شبه هارمونیکی استفاده می‌شود که تنها در بخش‌های صدادار گفتار می‌توان استفاده کرد. بدین صورت که ابتدا فرکانس پیچ سیگنال گفتار برحسب زمان بدست می‌آید و مسیرهای فرکانسی دیگر، به صورت مضربی از فرکانس پیچ سیگنال گفتار در نظر گرفته می‌شود.

فرض کنید یک بخش صدادار از سیگنال گفتار را در اختیار داریم. ابتدا گفتار فریم بندی می‌شود و فرکانس پیچ برای هر فریم بدست می‌آید. تطبیق فرکانسی بین دو فریم برای مسیر فرکانسی pام به صورت رابطه‌ی خطی زیر می‌باشد.

$$\omega_p(n) = p\left(\omega_0^k + \frac{\omega_0^{k+1} - \omega_0^k}{N} n\right) \qquad (۲۶-۲)$$

برای سنتز سیگنال از اطلاعات فاز در مرحله‌ی آنالیز استفاده نمی‌شود سیگنال طبق رابطه‌ی زیر سنتز می‌شود.

به خاطر استفاده نکردن از اطلاعات فاز، شکل سیگنال سنتز شده نسبت به سیگنال اصلی تفاوت دارد. در مورد گفتار صدادار سیگنال سنتز شده، از لحاظ شنیداری نزدیک به سیگنال اصلی بوده اما کاملاً شبیه به سیگنال اصلی نمی‌باشد. لذا برای معیار شفافیت، سیگنال سنتز شده‌ی پنهان‌سازی شده با سیگنال سنتز شده‌ی بدون اطلاعات نهان‌نگاره مقایسه می‌شود و با سیگنال اصلی مقایسه نمی‌شود که یک ضعف این روش است. نحوه پنهان‌سازی و آشکارسازی نهان‌نگاره با توجه به [۱۲] آورده شده است.

۲-۸-۱ پنهان‌سازی نهان‌نگاره

برای وارد کردن اطلاعات از یکی از مسیرهای فرکانسی $p > 1$ استفاده شده و اطلاعات با یک مدولاسیون مناسب که قابل شنیده شدن نباشد در مسیر فرکانسی وارد می شود. از بردار سینوسی زیر برای وارد کردن یک بیت اطلاعات نهان‌نگاره استفاده می شود.

$$w_0 = \frac{1}{2} - \frac{1}{2} cos\left(2\pi \frac{n}{M-1}\right) \quad for\ 0 \le n \le M-1 \qquad (۲۷-۲)$$

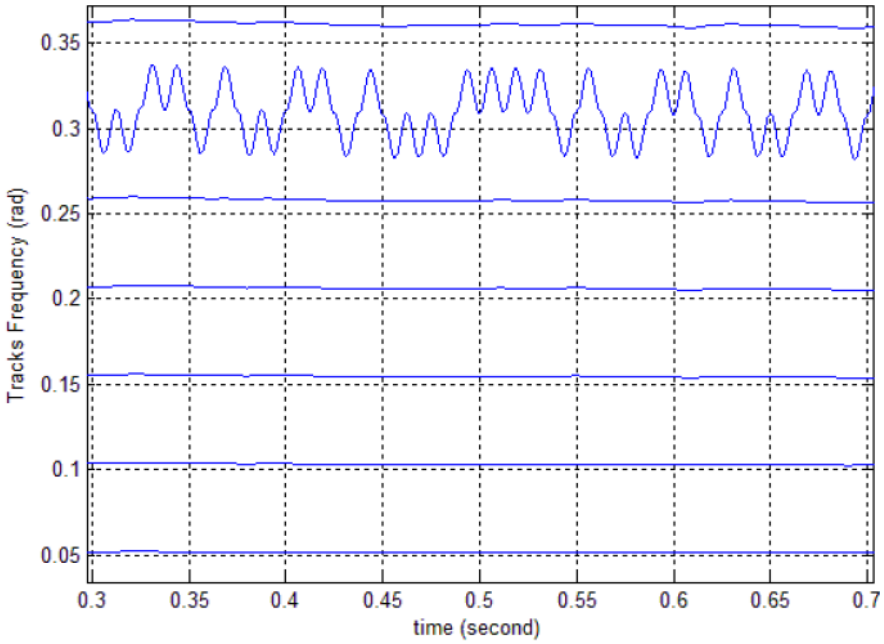
که M طول بردار سینوسی است. ابتدا اطلاعات نهان‌نگاره از حالت (۰ و ۱) به یک رشته ی قطبی b_k با مقادیر (۱- و ۱) تبدیل می شود. سپس توسط رابطه زیر بردار نهایی نهان‌نگاره ساخته می شود.

$$w_n = \sum_{k=0}^{L-1} b_k w_0(n - kM) \qquad (۲۸-۲)$$

که L تعداد نمونه‌های نهان‌نگاره b_k می‌باشد. برابر با N_s/M می‌باشد که طول سیگنال میزبان است. در نهایت اطلاعات نهان‌نگاره طبق رابطه‌ی زیر به مسیر فرکانس اضافه می شود.

$$\omega_p^w(n) = \omega_p + \lambda \omega_0^{min} \quad with \quad \omega_0^{min} = \min_{0 \le k \le N_s - 1} \omega_0(k) \qquad (۲۹-۲)$$

که λ یک شاخصه کنترل می باشد. $\omega_0^{\min}$ برابر با مینیمم فرکانس‌های پیچ آنالیز شده در فریم‌ها می باشد. در نهایت سیگنال طبق رابطه ی۱-۲۷ سنتز می شود و سیگنال پنهان‌سازی شده$S^w(n)$ به دست می آید. λ به دست آمده بگونه‌ای تنظیم می‌شود که تغییرات ایجاد شده با مسیرهای فرکانسی دیگر تداخل نداشته باشد. در مراحل شبیه سازی مقدار λ بین ۰/۱ تا ۰/۵ انتخاب شده است. یک گفتار صدادار با فرکانس نمونه برداری ۱۶ KHz را با این روش پنهان‌سازی شده است. مقدار λ برابر ۰/۵ می‌باشد. شکل ۲-۳ تغییرات حاصل از وارد شدن اطلاعات نهان‌نگاره در مسیر فرکانسی ششم از این گفتار را نشان می دهد.



شکل۲-۳ تغییر مسیر فرکانسی ششم در اثر وارد کردن اطلاعات نهان‌نگاره

۲-۸-۲ آشکارسازی نهان‌نگاره

در مرحله‌ی آشکارسازی باید دانست که اطلاعات نهان‌نگاره در کدام مسیر فرکانسی (p) پنهان شده است. برای آشکارسازی ابتدا مسیر فرکانسی پیچ $\widehat{\omega}_0^w(n)$ ، مشابه روشی که در بخش قبل توضیح دادیم برای سیگنال پنهان‌سازی شده بدست می‌آید. سپس با دانستن p، مسیر فرکانسی بدون مدولاسیون pام را طبق رابطه‌ی زیر تخمین می زنیم.

$$\widehat{\omega}_p(n) = p\widehat{\omega}_0^w(n) \qquad (۳۰-۲)$$

در مرحله‌ی بعد برای استخراج مسیر فرکانسی pام به همراه اطلاعات نهان‌نگاره طبق رابطه‌ی زیر عمل می‌کنیم.

$$S_p^w(n) = \left(h(n) * s^w e^{-j\sum_{k=0}^{n}\widehat{\omega}_p(k)}\right)e^{j\sum_{k=0}^{n}\widehat{\omega}_p(k)} \qquad (۳۱-۲)$$

h(n) یک فیلتر پایین گذر با فرکانس قطع $\omega_0^{min}/2$ می باشد. در رابطه‌ی ۱-۲۷ ابتدا توسط $e^{-j\sum_{k=0}^{n}\widehat{\omega}_p(k)}$ مسیر فرکانسی pام به مبدأ فرکانسی منتقل می‌شود. سپس با فیلتر پایین گذر h(n)

مسیرهای فرکانسی دیگر حذف می‌شوند در انتها نیز توسط $e^{j\sum_{k=0}^{n}\widehat{\omega}_p(k)}$مسیر فرکانسی pام، به محل اصلی خود باز می‌گردد.

$S_p^w(n)$ یک سیگنال مختلط مطابق رابطه‌ی (۲-۳۲) حاوی اطلاعات مسیر فرکانسی pام، می‌باشد. این سیگنال یک سیگنال تحلیلی بوده و قسمت موهومی آن از روی قسمت حقیقی آن با تبدیل هیلبرت[۱] بدست می آید. فرکانس لحظه‌ای[۲] $\widehat{\omega}_p^w(n)$ ، از طریق رابطه‌ی (۲-۳۳) بدست می آید که حالتی از مشتق فاز می باشد.

$$S_p^w(n) = S_p^r(n) + jS_p^j(n) \qquad (۳۲-۲)$$

$$\widehat{\omega}_p^w(t) = \frac{S_p^r(n)S_p^i(n+1) - S_p^r(n+1)S_p^i(n)}{S_p^r(n)^2 + S_p^i(n)^2} \qquad (۳۳-۲)$$

$\widehat{w}(n)$ با استفاده از اختلاف $\widehat{\omega}_p^w(n)$ از $\widehat{\omega}_p(n)$ بدست می‌آید در نهایت تصمیم گیری برای بیت استخراج شده توسط رابطه‌ی (۲-۳۴) انجام می‌گیرد.

$$b_k = \begin{cases} 1 & if \ \ \sum_{m=0}^{M-1}\widehat{w}(kM+m) \geq 0 \\ -1 & otherwise \end{cases} \quad , \quad for \ \ 0 \leq k \leq L-1 \qquad (۳۴-۲)$$

در این روش قرار دادن سیگنال در مسیرهای فرکانسی دوم تا چهارم باعث می شود کیفیت سیگنال پنهان‌سازی شده کاهش یابد. در نتیجه برای پنهان‌سازی اطلاعات باید از مسیرهای فرکانسی پنجم به بالا استفاده کنیم.

بهترین نتایج برای نرخ بیت خطا برای مسیر فرکانسی پنجم و ششم و مقدار $\lambda = 0.5$ بدست آمده است. از آنجایی که اطلاعات مهم سیگنال گفتار در مسیرهای فرکانسی پایین‌تر قرار دارد لذا این روش در برابر حملات نویز جمع شونده مقاومت کمتری خواهد داشت.

---

[۱] Hilbert
[۲] Instantaneous Frequency

۲-۹ پنهان‌سازی تهی[۱] یا بدون اتلاف

در یک روش[۱۷] که به پنهان‌سازی تهی مشهور است، نهان‌نگاره در خود سیگنال پنهان نمی‌شود بلکه در مجموعه‌ای از ویژگی‌های سیگنال پنهان می‌گردد. برای آشکارسازی نهان‌نگاره نیز با استفاده از همین ویژگی‌ها نهان‌نگاره بازیابی می‌شود.

۲-۹-۱ پنهان‌سازی نهان‌نگاره

برای مثال فرض کنید در هر فریم از گفتار یک بیت از نهان‌نگاره پنهان می‌شود. برای این کار داده‌های یک فریم به عنوان ورودی، و بیت نهان‌نگاره به عنوان خروجی یک سیستم طبقه‌بند در نظر گرفته می‌شود. سیستم طبقه‌بند می‌تواند یک شبکه عصبی یا یک ماشین بردار پشتیبان باشد. به این ترتیب، طبقه‌بند با استفاده از داده‌های ورودی و خروجی آموزش داده شده و پارامترهای آن محاسبه می‌گردد.

۲-۹-۲ آشکارسازی نهان‌نگاره

برای آشکارسازی نهان‌نگاره، هر فریم از گفتار به طبقه‌بند داده شده تا خروجی آن که همان بیت نهان‌نگاره است، به دست آید. بدیهی است که باید پارامترهای طبقه‌بند به همراه گفتار ارسال گردد. به همین جهت این آشکارسازی به صورت نیمه‌کور می‌باشد.

با توجه به اینکه اطلاعات نهان‌نگاره در خود سیگنال پنهان نمی‌شود، در سیگنال هیچ تغییری حاصل نمی‌گردد و شفافیت این روش از همه روش‌های دیگر بالاتر است. عیب این روش در این است که عمل پنهان‌سازی به صورت برخط[۲] انجام نمی‌شود. یعنی همزمان با تولید گفتار عمل پنهان‌سازی صورت نمی‌گیرد. بلکه برای پنهان سازی به کل داده‌های گفتار جهت آموزش طبقه‌بند نیاز می‌باشد.

---

۱ Zero Watermarking
۲ Online

فصل سوم:

معرفی تبدیل فوریه - بسل و کاربردهای آن در پردازش گفتار

۳-۱ مقدمه

می‌دانیم هر سیگنال برای ذخیره سازی، انتقال و یا پردازش نیازمند آن است که به طور مطلوبی نمایش داده شود و یا به اصطلاح کد شود. در بسیاری از موارد نمایش یک سیگنال مستقیماً به وسیله‌ی نمونه‌های آن و یا یک تابع تحلیلی (اگر تابع مناسب وجود داشته باشد) مطلوب نیست. برای مثال سیگنالی که در حوزه‌ی زمان تعریف شده است ممکن است در حوزه‌ی فرکانس فشردگی بیشتری داشته باشد. همچنین بسیاری از سیگنال‌های عملی دارای اطلاعات اضافه می‌باشند. سیگنال-های تصویر و گفتار نیز در این گروه قرار دارند و بهتر است و تا حدی ضرورت دارد که با تعداد نمونه‌ی کمتری نمایش داده شوند تا برای ذخیره سازی و یا پهنای باند انتقال، صرفه‌ی اقتصادی وجود داشته باشد. همچنین پردازش سیگنال می‌تواند در یک حوزه بهتر از حوزه‌ی دیگر باشد. مثال واضح آن پردازش سیگنال در حوزه‌ی فرکانس است که به دلیل پیدایش قطعات سخت‌افزار تبدیل فویه گسسته (FFT) پردازش سیگنال در حوزه‌ی فرکانس راحت‌تر و سریع‌تر انجام می‌گیرد. در صورتی که طیف وسیعی از این سیگنال‌ها در حوزه‌ی زمان اتفاق می‌افتند.

علاوه بر این، تکیه‌ی بیشتر تکنیک‌های شناسایی الگو بر توانایی تولید مجموعه ضرایبی از یک ردیف داده (نمونه‌های حوزه‌ی زمان) است که فشرده‌تر باشند تا بتوان از آنها به عنوان ویژگی برای آموزش سیستم‌های شناسایی استفاده نمود.

در این فصل بسط یک سیگنال گفتار به سری‌های فوریه - بسل بررسی خواهد شد. ضرایب به دست آمده می‌توانند به عنوان بردار ویژگی برای شناسایی الگو استفاده شوند. در ضمن کاربردهایی از بسط سیگنال گفتار به سری فوریه - بسل در زمینه‌هایی چون شناسایی گفتار، شناسایی گوینده، جداسازی گوینده، جداسازی گفتار و شناسایی زبان گوینده و موارد دیگر وجود دارد که در این فصل به ذکر بعضی از آنها پرداخته شده است.

۳-۲ نمایش سری‌ها

احتمالاً اولین نمایش داده‌ها به صورت بسط مجموعی از توابع به سال ۱۷۵۳ باز می‌گردد[۱۸]. در آن زمان برنولی[۱] برای توصیف مکانی و زمانی تارهای مرتعش موفق به کشف رابطه‌ی (۳-۱) شد.

$$f(x,t) = A_1 \sin x \cos(at) + A_2 \sin(2x) \cos(2at) + \cdots \qquad (۱-۳)$$

که در رابطه‌ی (۳-۱) ، a طول تار مرتعش می‌باشد.

ایده‌ی نمایش هر تابع به عنوان مجموعی از سینوسی‌ها تقریباً جدید و بحث برانگیز بود تا اینکه تئوری جامعی برای نمایش سری هر تابع دلخواه به وجود آمد.

تئوری نمایش سری هر تابع دلخواه جامع‌تر از نوشتن سیگنال به صورت مجموعی از سینوسی‌ها بود. در حقیقت هر مجموعه‌ی متعامد[۲] از توابع پایه می‌تواند برای نمایش هر تابع دلخواه به کار رود. اگر مجموعه‌ی توابع متعامد را به صورت زیر تعریف کنیم:

$$\sum_{-\infty}^{+\infty} \varphi_m(t)\varphi_n(t) = \begin{cases} 1 & , m = n \\ 0 & , m \neq n \end{cases} \qquad (۲-۳)$$

تابع $f(t)$ را می‌توان به صورت زیر بسط داد:

$$f(t) = \sum_{n=-\infty}^{+\infty} C_n \varphi_n(t) \qquad (۳-۳)$$

که در رابطه‌ی (۳-۳)، C_n به صورت رابطه‌ی (۳-۴) تعریف می‌شود.

$$C_n = \int_{-\infty}^{+\infty} f(t) \varphi_n(t) dt \qquad (۴-۳)$$

این نتیجه را می‌توان با ضرب کردن طرفین رابطه‌ی (۳-۳) در $\varphi_n(t)$ و انتگرال گرفتن از طرفین در تمام بازه‌ی زمانی t نشان داد. اگر $f(t)$ را به سیگنال‌های انرژی محدود و باند محدود منحصر کنیم خاصیت جالبی را می توان بیان کرد. انرژی تابع $f(t)$ به صورت زیر است:

$$E = \int_{-\infty}^{+\infty} f^2(t) dt = \sum_{n=-\infty}^{+\infty} C_n^2 < \infty \qquad (۵-۳)$$

---

[۱] D.Bernoulli
[۲] orthogonal

رابطه‌ی (۳-۵) همان فرمول پارسوال در ضرایب فوریه می‌باشد. همچنین فرم عمومی نمایش سری‌ها برای راه حل‌های ریاضی مفید است. به طور واضح انتخاب توابع پایه به صورت رابطه‌ی (۳-۶) سری-های فوریه را نتیجه می‌دهد.

$$\varphi_n(t) = e^{jn\omega t} \qquad (۳-۶)$$

اما بسیاری توابع دیگر یافت شده است که مفید می‌باشند. اگر $f(t)$ تنها در بازه‌ی محدودی از زمان $(-\tau, \tau)$ مقدار داشته باشد (یک فرض واقع گرایانه) مطلوب می باشد که انرژی آن در همین بازه متمرکز شده باشد. انرژی متمرکز شده با نسبت کسری انرژی طبق رابطه‌ی (۳-۷) نشان داده می‌شود.

$$E = \frac{\int_{-\tau}^{\tau} |\varphi_n(t)|^2 dt}{\int_{-\infty}^{+\infty} |\varphi_n(t)|^2 dt} \qquad (۳-۷)$$

می‌توان نشان داد که E برای $\varphi_n(t)$ مرتبط با توابع کروی[۱] کشیده ماکزیمم می شود.

گزینه‌ی دیگر برای $\varphi_n(t)$ خانواده‌ی توابع بسل می‌باشد که بسط فوریه - بسل را نتیجه می‌دهد. سری‌های فوریه - بسل در اپتیک و آکوستیک کاربردهایی یافته است. به علاوه خاصیت بی‌نظیری از این سری‌ها برای پردازش گفتار نیز مفید است. خواص سری‌های فوریه - بسل در ادامه‌ی این فصل معرفی خواهند شد.

۳-۳ سری فوریه - بسل

توابع بسل با حل معادله‌ی دیفرانسیل زیر به دست می آیند.

$$t^2 y'' + ty' + (t^2 - n^2)y = 0 \quad n > 0 \qquad (۳-۸)$$

که معادله‌ی دیفرانسیل بسل نامیده می شود. حل کلی این معادله جواب زیر را می‌دهد.

$$y = C_1 J_n(t) + C_2 Y_n(t) \qquad (۳-۹)$$

که $J_n(x)$ را تابع بسل نوع اول مرتبه n و $Y_n(x)$ را تابع بسل نوع دوم مرتبه n می نامند. توابع بسل قابل ارائه به فرم سری هستند. برای مثال $J_n(x)$ را می توان این گونه نوشت:

[۱] Prolated Spheroidal Functions

$$J_n(t) = \sum_{r=0}^{\infty} \frac{(-1)^r (t/2)^{n+2r}}{r!\,\Gamma(n+r+1)} \qquad (۱۰-۳)$$

و در حالت خاص:

$$J_0(t) = 1 - \frac{t^2}{2^2} + \frac{t^4}{2^2 4^2} - \frac{t^6}{2^2 4^2 6^2} + \cdots \qquad (۱۱-۳)$$

می‌توان نشان داد توابع بسل با در نظر گرفتن تابع وزن‌دهی[۱] t متعامد هستند. به این ترتیب می توان دید:

$$\int_0^a tJ_n(\alpha t)J_n(\beta t)dt = \frac{a[\beta J_n(a\alpha)J'_n(a\beta) - \alpha J_n(a\beta)J'_n(a\alpha)]}{\alpha^2 - \beta^2} \qquad \alpha \neq \beta \qquad (۱۲-۳)$$

و اگر $\beta \to \alpha$ به وسیله قاعده هوپیتال خواهیم داشت:

$$\int_0^a tJ_n^2(\alpha t)dt = a^2/2 \left[J'^2_n(a\alpha) + (1 - \frac{n^2}{a^2\alpha^2})J_n^2(a\alpha)\right] \qquad (۱۳-۳)$$

حال اگر α و β ریشه‌های مختلف معادله‌ی $J_n(at) = 0$ باشند می‌توان نوشت:

$$\int_0^a tJ_n(\alpha t)J_n(\beta t)dt = 0 \qquad \alpha \neq \beta \qquad (۱۴-۳)$$

و در نتیجه $J_n(\alpha t)$ و $J_n(\beta t)$ با در نظر گرفتن تابع وزن دهی t متعامد هستند. همچنین با توجه به اینکه $J_n(a\alpha) = 0$ می‌توان رابطه‌ی (۳-۱۳) را به صورت زیر نوشت:

$$\int_0^a tJ_n^2(\alpha t)dt = a^2/2 \left[J'^2_n(a\alpha)\right] \qquad (۱۵-۳)$$

با در نظر گرفتن فرمول زیر:

$$J'_n(t) = \tfrac{1}{2}(J_{n-1}(t) - J_{n+1}(t)) \qquad (۱۶-۳)$$

در نهایت می‌توان رابطه‌ی (۳-۱۳) را به صورت زیر نوشت:

$$\int_0^a tJ_n^2(\alpha t)dt = \frac{a^2}{8}(J_{n-1}(a\alpha) - J_{n+1}(a\alpha))^2 \qquad (۱۷-۳)$$

حال با دانستن این مطلب می توان هر تابع دلخواه روی بازه‌ی $(0, a)$ را به صورت جملاتی از توابع بسل نوشت.

$$f(t) = \sum_{i=1}^{\infty} C_i J_n(\lambda_i t) \qquad (۱۸-۳)$$

[۱] Weighting Function

که در رابطه‌ی (۳-۱۸)، λ_1 ، λ_2 و... ریشه‌های مثبت معادله‌ی $J_n(at)=0$ هستند. برای به دست آوردن ضرایب C_m طرفین رابطه‌ی (۳-۱۸) را در $tJ_n(\lambda_m t)$ ضرب می‌کنیم.

$$tf(t)J_n(\lambda_m t)=\sum_{i=1}^{\infty} C_i J_n(\lambda_i t)J_n(\lambda_m t) \qquad (۳-۱۹)$$

با فرض اینکه $f(t)$ روی بازه‌ی $(0,a)$ تعریف شده است، از رابطه‌ی (۳-۱۹) در این بازه انتگرال می‌گیریم. با توجه به خاصیت تعامد، همه‌ی جملات سمت راست رابطه، به جز $i=m$ می‌شوند. در نتیجه داریم:

$$\int_0^a tf(t)J_n(\lambda_m t)=C_m\int_0^a tJ_n^2(\lambda_m t) \qquad (۳-۲۰)$$

با استفاده از رابطه‌ی (۳-۱۷) داریم:

$$\int_0^a tf(t)J_n(\lambda_m t)=\frac{a^2}{8}C_m[J_{n-1}(a\lambda)-J_{n+1}(a\lambda)]^2 \qquad (۳-۲۱)$$

با ساده کردن این رابطه، ضرایب C_m به صورت زیر به دست می‌آیند.

$$C_m=\frac{8}{a^2C_m[J_{n-1}(a\lambda_m)-J_{n+1}(a\lambda_m)]^2}\int_0^a tJ_n(\lambda_m t)f(t)dt \qquad (۳-۲۲)$$

در حالت خاص اگر بخواهیم $f(t)$ را روی هر بازه‌ی دلخواه $(0,a)$ با استفاده از تابع بسل مرتبه‌ی صفر بسط دهیم:

$$f(t)=\sum_{m=1}^{\infty} C_m J_0(\lambda_m t) \quad , \quad 0<t<a \qquad (۳-۲۳)$$

با در نظر گرفتن فرمول زیر:

$$J_{-n}(\mathrm{t})=(-1)^n J_n(t) \qquad (۳-۲۴)$$

ضرایب C_m به صورت زیر محاسبه می شوند:

$$C_m=\frac{2\int_0^a tf(t)J_0(\lambda_m t)dt}{a^2[J_1(\lambda_m a)]^2} \qquad (۳-۲۵)$$

و λ_m به ازای $m=1,2,...$ ریشه‌های مثبت معادله‌ی $J_0(a\lambda_m)=0$ می‌باشد که به ترتیب صعودی مرتب شده‌اند. انتگرالی که در رابطه‌ی (۳-۲۵) یا (۳-۲۲) وجود دارد، تبدیل هنکل[۱] محدود نامیده

[۱] Hankel Transform

می‌شود. الگوریتم‌های محاسبه‌ی تبدیل هنکل سریع[۱] که بر پایه‌ی تبدیل فوریه‌ی سریع[۲] می‌باشد محاسبه‌ی رابطه‌ی (۳-۲۲) را با سرعتی اندکی کمتر از سرعت محاسبه‌ی ضرایب فوریه ممکن کرده است. با بسط سیگنال‌های گفتار با توابع بسل مجموعه ویژگی‌هایی با خصوصیت‌های مختلف تولید می‌شود. از آنجا که محاسبه‌ی ضرایب سری فوریه - بسل نیازمند محاسبه‌ی تبدیل هنکل می‌باشد، بعضی از خصوصیت‌های تبدیل هنکل و نحوه‌ی محاسبه‌ی آن در ادامه ارائه خواهد شد.

۳-۳- ۱ خصوصیت‌های تبدیل هنکل

تبدیل هنکل به صورت رابطه‌ی (۳-۲۶) تعریف می شود.

$$F(\lambda) = \int_0^\infty t f(t) J_0(\lambda t) dt \qquad (۲۶-۳)$$

و فرمول تبدیل معکوس آن به صورت رابطه‌ی (۳-۲۷) است:

$$f(t) = \int_0^\infty \lambda F(\lambda) J_0(\lambda t) d\lambda \qquad (۲۷-۳)$$

اگر در نظر بگیریم :

$$f_1(t) \overset{h}{\leftrightarrow} F_1(\lambda) \quad \text{و} \quad f_2(t) \overset{h}{\leftrightarrow} F_2(\lambda)$$

سپس رابطه‌ی (۳-۲۸) را خواهیم داشت که فرمول پارسوال می‌باشد.

$$\int_0^\infty t f_1(t) f_2^*(t) dt = \int_0^\infty \lambda F_1(\lambda) F_2^*(\lambda) d\lambda \qquad (۲۸-۳)$$

سایر فرمول‌ها نیز به این صورت است:

$$f(at) \overset{h}{\leftrightarrow} \frac{1}{a^2} F\left(\frac{\lambda}{a}\right) \qquad (۲۹-۳)$$

$$f(t) + g(t) \overset{h}{\leftrightarrow} F(\lambda) + G(\lambda) \qquad (۳۰-۳)$$

در تبدیل هنکل خاصیت شیفت وجود ندارد زیرا توابع بسل نسبت به شیفت تغییر پذیر هستند.

---

[۱] Fast Hankel Transform
[۲] Fast Fourrier Transform

۳-۳-۲ الگوریتم‌های محاسبه‌ی تبدیل هنکل

الگوریتم‌هایی برای محاسبه‌ی تبدیل هنکل ارائه شده‌اند. این الگوریتم‌ها باعث استفاده از سری‌های فوریه - بسل بدون هزینه‌ی زیاد محاسباتی شده است. همچنین الگوریتم‌های تبدیل هنکل سریع بر اساس الگوریتم‌های سریع و قابل دسترس تبدیل فوریه سریع می‌باشد.

اولین الگوریتم محاسبه‌ی تبدیل هنکل توسط سیگمن[۱] ارائه شد. این الگوریتم بر اساس تبدیلی به نام گاردنر[۲] بود که توابع بسل را به توابعی تبدیل می‌کرد که نسبت به شیفت تغییر ناپذیر شوند. عیب این روش در این بود که طی این تبدیل، زمان نمونه‌برداری به صورت نمایی تغییر می‌کرد که این نمونه-برداری غیر خطی با واقعیت فیزیکی داده‌ها مغایر بود.

الگوریتم مشابهی توسط یوهانسن و سورنسن[۳] پیشنهاد شد که انتگرال را به انتگرال پیچش[۴] نگاشت می‌کرد که این روش هم همان مشکل روش سیگمن را داشت.

یک روش سریع توسط کاونگ و کوک[۵] پیشنهاد شد که عبارت بود از بسط تابع تحت تبدیل مجموعه-ای از چند جمله‌ای‌های گوسی- لگر[۶] ، که به عنوان تبدیل‌های تحلیلی مشهورند. الگوریتم آنها مشکل همگرایی داشت و به این خاطر به این روش توجهی نشد.

الگوریتمی توسط کندل[۷] با نمونه برداری یکنواخت پیشنهاد شد. با توجه به اینکه بیشتر سری‌های زمانی به طور یکنواخت در زمان نمونه برداری می‌شوند این روش بسیار مورد توجه قرار گرفت. این الگوریتم برای پیاده سازی ساده بوده و بر اساس محاسبه‌ی FFT می باشد.

گوپالان[۸] نیز مشابه این روش را آورده است که به توضیح آن می‌پردازیم.

با توجه به رابطه‌ی (۳-۲۶) داریم:

---

۱ Seigman
۲ Gardner Transform
۳ Johansen & Sorensen
۴ Convolution Integral
۵ Cook & Cavangh
۶ Gaussian-Laguerre Polynomials
۷ Candel
۸ Gopalan

$$F(\lambda) = \int_0^{\infty} tf(t) J_0(\lambda t) dt \qquad (۳۱-۳)$$

فرم انتگرالی بسل مرتبه‌ی صفر به صورت رابطه‌ی (۳-۳۲) می‌باشد[۱۹].

$$J_0(\lambda t) = \frac{1}{2\pi} \int_0^{2\pi} e^{-j\lambda t.cos\theta} d\theta \qquad (۳۲-۳)$$

با جایگزینی رابطه‌ی (۳-۳۲) در رابطه‌ی (۳-۳۱)، رابطه‌ی زیر را نتیجه می‌گیریم.

$$F(\lambda) = \frac{1}{2\pi} \int_0^{a} tf(t) \int_0^{2\pi} e^{-j\lambda t.cos\theta}\, d\theta dt \qquad (۳۳-۳)$$

حال انتگرال‌ها را با مجموع[۱] تخمین می‌زنیم. برای این منظور بازه‌ی $(0,2\pi)$ را به K قسمت مساوی به طول $\Delta\theta = \frac{2\pi}{K}$ ، و بازه‌ی صفر تا a را به N قسمت مساوی به طول $\Delta T = \frac{a}{N}$ تقسیم می‌کنیم. سپس رابطه‌ی (۳-۳۱) را بازنویسی می‌کنیم.

$$F(\lambda) \simeq \frac{(\Delta T)^2}{K} \sum_{n=0}^{N-1} nf(n\Delta T) \sum_{k=0}^{K-1} e^{-jn\Delta T\lambda \cos(k\Delta\theta)} \qquad (۳۴-۳)$$

$$F(\lambda) \simeq \frac{(\Delta T)^2}{K} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} nf(n\Delta T)\, e^{-jn\Delta T\lambda \cos(k\Delta\theta)} \qquad (۳۵-۳)$$

که رابطه‌ی بین λ با نرخ نمونه برداری (ΔT) و تعداد نمونه‌ها (N) به صورت زیر می‌باشد.

$$\lambda = \frac{\ell.2\pi}{N.\Delta T} \qquad (۳۶-۳)$$

با استفاده از رابطه‌ی (۳-۳۶) در رابطه‌ی (۳-۳۵) خواهیم داشت:

$$F(\lambda) \simeq \frac{(\Delta T)^2}{K} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} nf(n\Delta T)\, e^{-j\frac{2\pi}{N} n\ell \cos(k\Delta\theta)} \qquad (۳۷-۳)$$

در رابطه‌ی (۳-۳۷) مجموع داخلی به عنوان تبدیل فوریه‌ی گسسته‌ی دنباله‌ی $nf(n)$ با وضوح فرکانسی $\omega\Delta = \frac{2\pi}{N.\Delta T}$ شناخته می‌شود که در مقادیر $\ell\cos(k\Delta\theta)$ محاسبه شده است و به سادگی با تکنیک‌های FFT به دست می آید. پس تقریب تبدیل هنکل محدود به صورت زیر است.

$$F(\lambda) \simeq \frac{(\Delta T)^2}{K} \sum_{k=0}^{K-1} P[\ell\cos(k\Delta\theta)] \qquad (۳۸-۳)$$

که تابع $P(x)$ به صورت زیر تعریف می‌شود.

$$P(x) = \sum_{n=0}^{N-1} nf(n) e^{-j\frac{2\pi}{N} nx} \qquad (۳۹-۳)$$

[۱] Summation

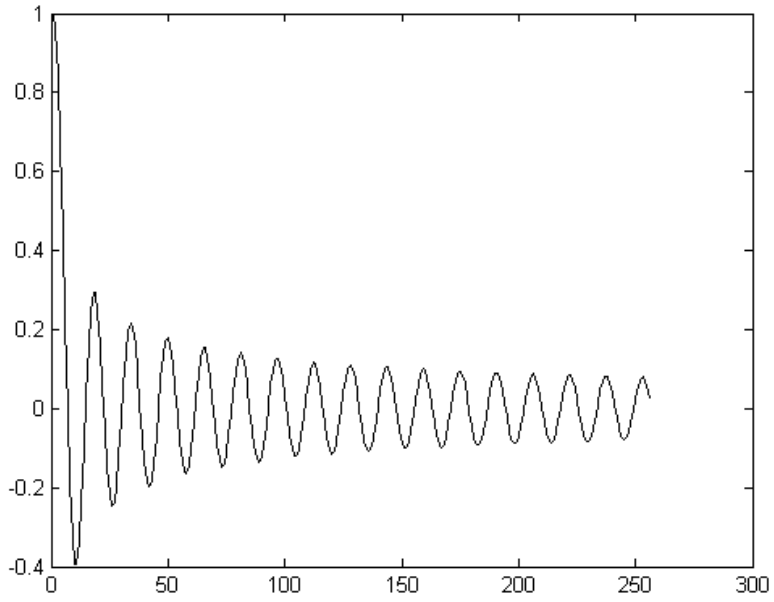
البته باید توجه داشت، هر چند که تابع $P(x)$ مختلط می‌باشد اما می‌توان نشان داد که برای سیگنال-های حقیقی مجموع قسمت‌های موهومی برای $k=0,1,\ldots,K-1$ برابر صفر است. بنابراین برای سیگنال-های حقیقی فقط محاسبه‌ی قسمت حقیقی تبدیل فوریه‌ی گسسته مورد توجه است.

برای محاسبه‌ی تبدیل فوریه در مقادیر $\ell\cos(k\Delta\theta)$ می توان نزدیکترین مقدار صحیح به آن را در نظر گرفت. با این حال برای بالاتر رفتن دقت در محاسبه می‌توان به تعداد کافی صفر به تابع $nf(n)$ اضافه نمود تا وضوح تبدیل بالاتر رود.

۳-۴ نتایج عددی نمایش یک سیگنال با سری فوریه - بسل

در این قسمت برای محاسبه‌ی ضرایب سری فوریه – بسل مثال‌های عددی آورده شده است. برای این کار دو تابع را در فاصله‌ی محدود در نظر گرفته و طبق رابطه‌ی (۳-۲۶) و الگوریتم گفته شده در بخش ۳-۳-۲ ضرایب سری فوریه – بسل محاسبه شده است.

به عنوان اولین مثال تابع $f(t)=J_0(t)$ در بازه‌ی صفر تا ۱۰۲/۸۸ (ریشه سی و سوم تابع بسل مرتبه صفر) انتخاب شده که با فاصله زمانی ۰/۴۰۲ نمونه‌برداری شده و ۲۵۶ نمونه از آن گرفته شده است. نمودار تابع در شکل ۳-۱ قابل مشاهده می‌باشد.

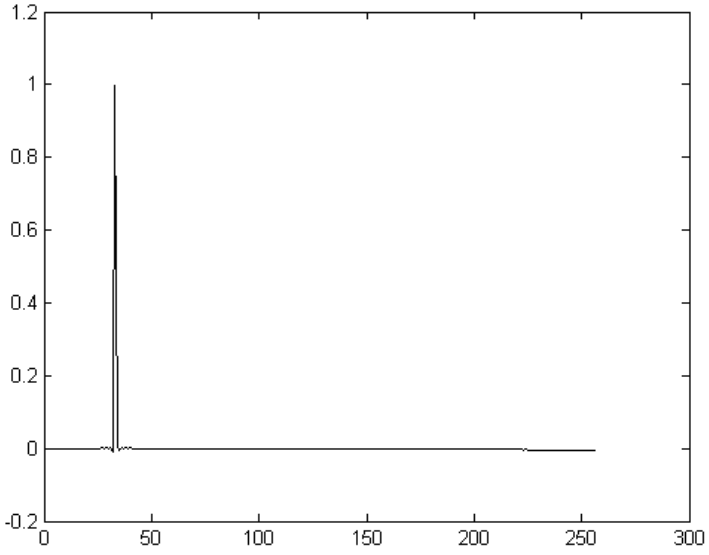


شکل ۳-۱ نمودار تابع $\boldsymbol{f(t)=J_0(t)}$

با محاسبه از راه فرمول می‌توان ضرایب را به صورت رابطه‌ی (۳-۴۰) به دست آورد.

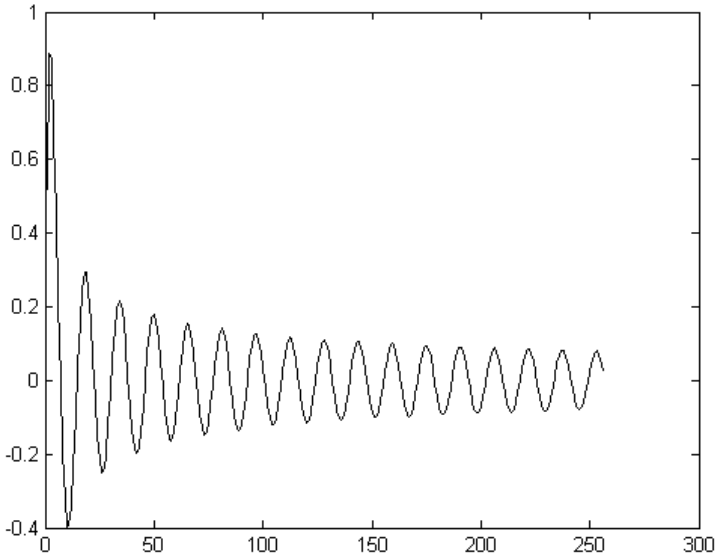
$$C_m = \begin{cases} 1 & , \quad m = 33 \\ 0 & , \quad m \neq 33 \end{cases} \qquad (۴۰-۳)$$

سپس طبق الگوریتم ضرایب محاسبه شده است. نمودار اندازه‌ی ضرایب در شکل ۳-۲ قابل مشاهده می‌باشد.



شکل ۳-۲ نمودار اندازه‌ی ضرایب فوریه - بسل تابع $\boldsymbol{f(t) = J_0(t)}$

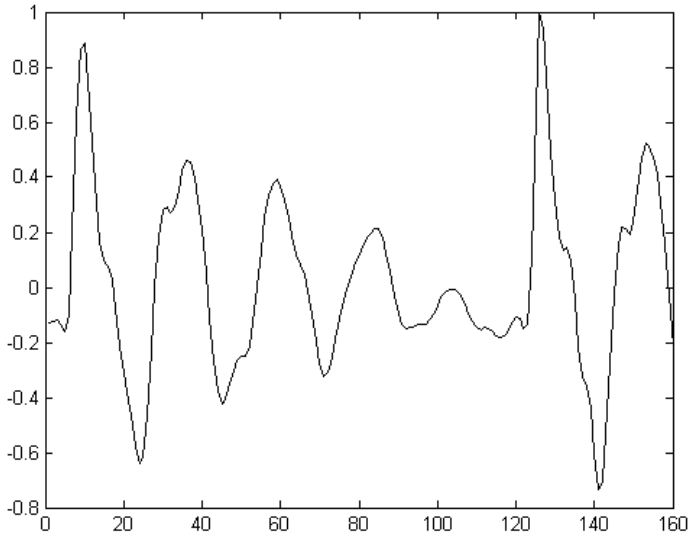
همان گونه که انتظار می‌رفت، الگوریتم همگرایی خوبی دارد. حال با استفاده از رابطه‌ی (۳-۲۳) تابع را بازسازی می‌کنیم. نمودار تابع بازسازی شده در شکل ۳-۳ قابل مشاهده است.



شکل ۳-۳ نمودار تابع بازسازی شده

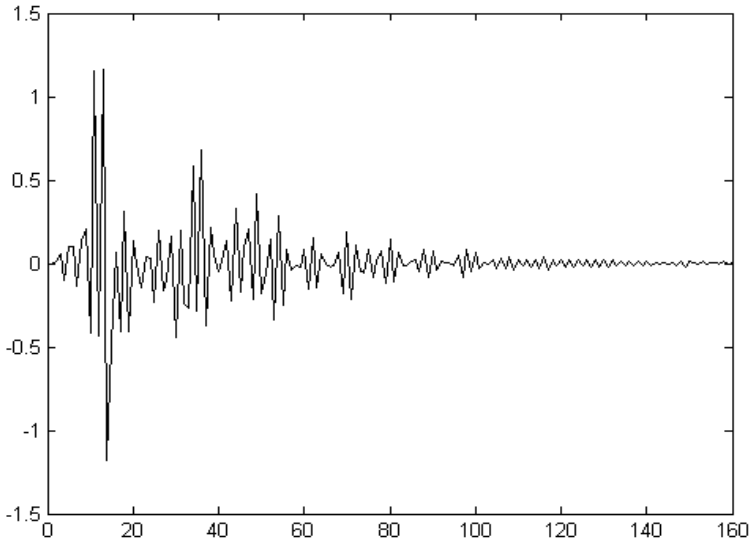
میزان خطای محاسبه شده با استفاده از معیار میانگین مربعات خطا[1] برابر ۰/۰۰۲۷ می‌باشد. ناچیز بودن این مقدار نشان از کارایی بالای الگوریتم دارد.

به عنوان مثال دوم یک فریم گفتار را مورد بررسی قرار می‌دهیم. این فریم از یکی از فایل‌های پایگاه داده تیمیت[2] انتخاب شده است. نمودار این فریم را در شکل ۳-۴ مشاهده می کنید.



شکل ۳-۴ نمودار یک فریم گفتار

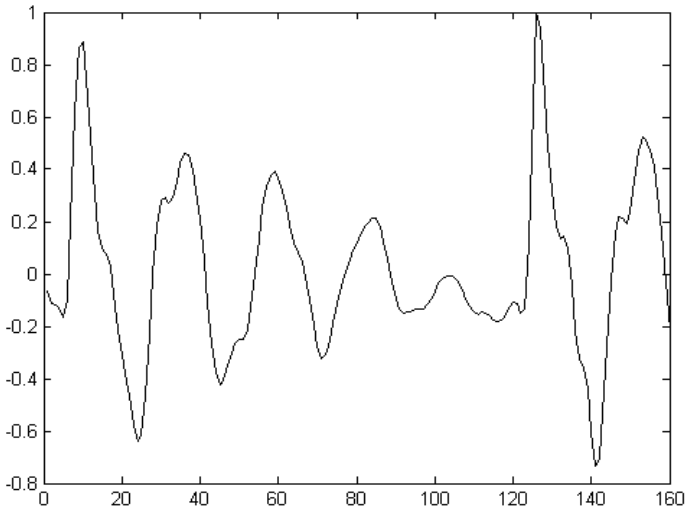
حال ضرایب سری فوریه – بسل به وسیله الگوریتم محاسبه می‌شوند. نمودار اندازه‌ی ضرایب در شکل ۳-۵ قابل مشاهده می‌باشد.



شکل ۳-۵ نمودار اندازه‌ی ضرایب فوریه - بسل یک فریم گفتار

[1] Mean Squared Error (MSE)
[2] Timit

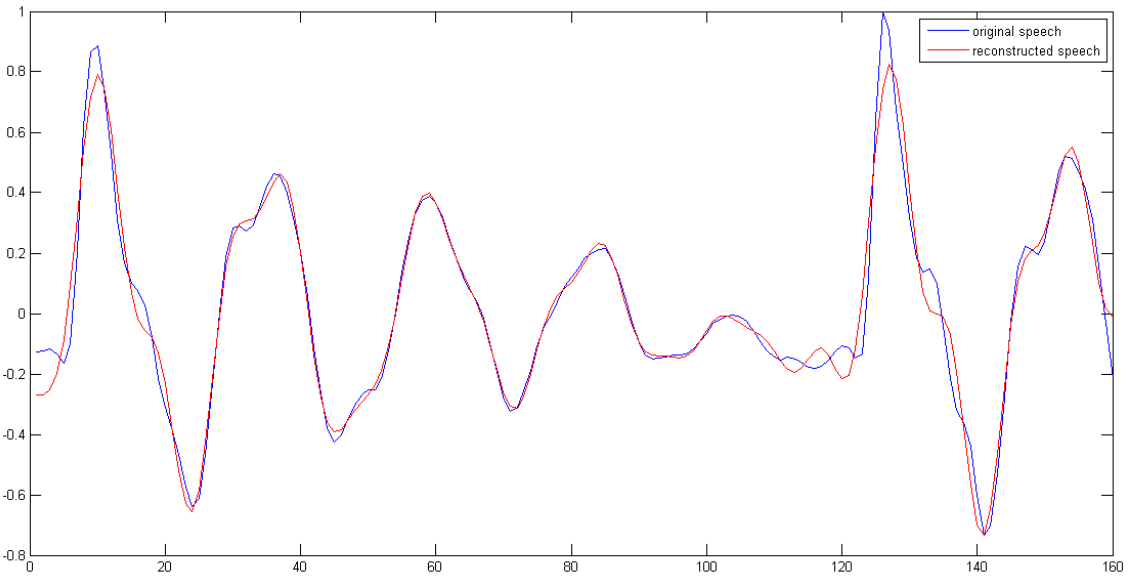
حال فریم گفتار با استفاده از رابطه‌ی (۳-۲۳) بازسازی می‌شود. نمودار فریم گفتار بازسازی شده در شکل ۳-۶ قابل مشاهده است.



شکل ۳-۶ نمودار فریم گفتار بازسازی شده

میزان خطای محاسبه شده با استفاده از معیار میانگین مربعات خطا برابر ۰/۰۰۰۵۳۸۹۶ می‌باشد. این مقدار حتی از مقدار محاسبه شده در مثال قبل هم بسیار کمتر است. این حاکی از توانایی بالای توابع بسل برای بسط سیگنال‌های گفتار می‌باشد.

حال سیگنال را فقط با ۴۰ ضریب اول بازسازی می‌کنیم. نمودار فریم گفتار بازسازی شده در شکل ۳-۷ قابل مشاهده است.



شکل ۳-۷ نمودار فریم گفتار بازسازی شده با ۴۰ ضریب در کنار فریم گفتار اصلی

همان گونه که در شکل ۳-۷ پیداست حتی با یک چهارم ضرایب هم می‌توان سیگنال گفتار را به خوبی بازسازی نمود. میزان خطای محاسبه شده با استفاده از معیار میانگین مربعات خطا برابر ۰/۰۴۱۲ می‌باشد.

۳-۵ کاربردهای تبدیل فوریه - بسل در پردازش گفتار

در این بخش به برخی از کاربردهای تبدیل فوریه - بسل در پردازش گفتار اشاره می نماییم. البته همان گونه که قبلا نیز گفته شد توابع بسل در زمینه‌های دیگری همچون تصویر و سیگنال‌های قلب نیز کاربرد دارد.

۳-۵-۱ تشخیص جنسیت گوینده

در مرجع [۲۰] برای تشخیص جنسیت گوینده، روشی با استفاده از توابع مد ذاتی[۱] و بسط سری فوریه - بسل ارائه شده است. در این روش سیگنال گفتار با استفاده از یک تکنیک تجزیه‌ی غیرخطی به نام تجزیه‌ی مد تجربی[۲] به دو مجموعه از سیگنال‌های با مدولاسیون دامنه و مدولاسیون فرکانس تجزیه می‌شوند که توابع مد ذاتی نامیده می شوند.

با محاسبه‌ی انرژی توابع، توابع مد ذاتی که دارای بیشترین اطلاعات از سیگنال گفتار هستند شناسایی شده و از آنها برای بازسازی سیگنال گفتار استفاده می شود. سپس این سیگنال بازسازی شده فریم بندی می گردد و ضرایب فوریه - بسل آنها محاسبه می شود. این ضرایب به عنوان ویژگی برای کلاسه‌بند ماشین بردار پشتیبان[۳] برای دو کلاس مذکر و مونث می‌باشند. کارآیی این کلاسه بند ۷۲/۹۲٪ محاسبه شده است.

توابع بسل به شکل توابع نمایی میرا شونده در زمان می‌باشند. که معادل توابع با باند باریک در حوزه‌ی فرکانس است. از آنجا که توابع مد ذاتی هم باند باریک می‌باشند، به همین دلیل از ضرایب فوریه – بسل برای استخراج ویژگی آنها استفاده شده است تا توصیف بهتری از توابع مد ذاتی داشته باشند.

---

[۱] Intrinsic Mode Functions
[۲] Empirical Mode Decomposition
[۳] Supported Vector Machine (SVM)

۳-۵-۲ شناسایی گوینده

در زمینه‌ی شناسایی گوینده کارهای مختلفی صورت گرفته است که چند مورد از آنها به شرح زیر می‌باشد.

الف: در مرجع [۲۱]، از ضرایب فوریه - بسل و کلاسه‌بند شبکه‌ی عصبی پس انتشار برای شناسایی گوینده استفاده شده است. در این روش، ضرایب فوریه - بسل با استفاده از تابع بسل مرتبه اول محاسبه شده و به عنوان ورودی در شبکه عصبی استفاده شده است. مشاهده می‌شود که هر چه تعداد ضرایب بالاتر رود کارآیی روش نیز بالاتر می رود. البته به دلیل اینکه از شبکه‌ی عصبی آماده‌ی موجود در نرم‌افزار متلب استفاده شده و این نرم‌افزار برای ضرایب پیش‌بینی خطی[۱] نوشته شده بود، ضرایب پیش‌بینی خطی بهتر از ضرایب فوریه – بسل نتیجه می‌دادند.

ب: در مرجع [۲۲]، از ضرایب فوریه – بسل با استفاده تابع بسل مرتبه اول و نوعی کلاسه‌بند تجاری گسسته‌ساز بردار استفاده شده و بر روی پایگاه داده‌ی گرین فلگ[۲] و ناتو[۳] کار صورت گرفته است. در این روش برای هر فریم گفتار ضرایب فوریه – بسل محاسبه شده است. سپس ضرایب به صورت مجموعه‌های پنج‌تایی در آمده و حاصل جمع این مجموعه‌ها به عنوان ورودی کلاسه‌بند استفاده شده است. هدف از این کار نوعی توصیف انرژی سیگنال گفتار در بازه‌های فرکانسی مختلف می‌باشد. کارآیی این روش با توجه به تعداد ضرایب بین ۵۰ تا ۷۵ درصد محاسبه شده است. استفاده از ضرایب فوریه – بسل در این روش به دلیل شباهت ساختار سیگنال گفتار و توابع بسل می‌باشد. زیرا هر دو غیر سینوسی، شبه متناوب و میرا شونده هستند.

ج: در مرجع [۲۳] نیز مقایسه‌ای بین ویژگی‌های بسل و کپستروم[۴] انجام گرفته است. در این روش، استفاده از ضرایب فوریه – بسل به عنوان ویژگی، پایگاه داده و کلاسه‌بند، همانند روش قسمت ب می‌باشد. مقادیر ۶۵٪ و ۷۶٪ برای ویژگی‌های بسل و مقادیر ۸۰٪ و ۸۶٪ برای ویژگی‌های کپستروم به

---

[۱] Linear Predictive Coefficients
[۲] Green Flag
[۳] NATO
[۴] Cepstrum

دست آمده است. با توجه به نتایج درج شده در این مقالات می‌توان نتیجه گرفت که توابع بسل توابع مناسبی برای استفاده در تحلیل و سنتز سیگنال‌های گفتار و شناسایی گوینده می باشند.

۳-۵-۳ تخمین تعداد گویندگان

در مرجع [۲۴] تحقیقی صورت گرفته است که در آن از سیگنال‌های ضبط شده با دو میکروفن استفاده شده است. که با استفاده از تاخیری که بین سیگنال‌های گویندگان وجود دارد و همبستگی متقابل آنها، تعداد گویندگان تخمین زده می شود. در این روش ابتدا ضرایب فوریه - بسل سیگنال هر میکروفن محاسبه شده و با یک چهارم ضرایب بازسازی می شوند تا به صورت باند محدود در آیند. سپس همبستگی متقابل بین آنها محاسبه می‌شود. به واسطه‌ی تاخیری که بین سیگنال‌ها وجود دارد در محل تاخیر قله‌هایی مشاهده می‌شود که تعداد آنها همان تعداد گویندگان می‌باشد. با مقایسه‌ی نتایج گرفته شده با نتایج تجربی نشان داده شده است که این روش نسبت به روش‌های قبلی کارآیی خوبی دارد.

۳-۵-۴ بهسازی گفتار

همان طور که می‌دانیم یکی از روش‌های مهم بهسازی گفتار روش تفریق طیفی می‌باشد. در این روش برای تخمین سیگنال بدون نویز، ضرایب فوریه نویز تخمین زده شده از ضرایب فوریه سیگنال نویزی شده تفریق می گردد. در مرجع [۲۵] به جای ضرایب فوریه از ضرایب فوریه - بسل استفاده شده است. سیگنال تخمین زده شده در نهایت بازسازی می شود.

در این روش از توابع بسل نوع یک مرتبه اول استفاده شده است. همچنین از خاصیت ذاتی توابع بسل که همچون یک فیلتر پایین گذر می باشد نیز کمک گرفته شده است تا نویز فرکانس بالا را حذف نماید.

۳-۵-۵ اصلاح و بازسازی گفتار

در مرجع[۲۶] به اصلاح سیگنال گفتار تغییر شکل یافته با استفاده از ضرایب فوریه - بسل پرداخته شده است. در این روش نیز از تابع بسل نوع یک مرتبه اول استفاده شده است. با توجه به اینکه طیف سیگنال بسل میان‌گذر می‌باشد از این خاصیت ضرایب فوریه - بسل برای به دست آوردن کیفیت سیگنال مورد نظر استفاده شده است. در ابتدا با تعدادی از ضرایب که تقریباً در میانه‌ی همه‌ی ضرایب بودند سیگنال گفتار بازسازی شد. سپس فقط ضرایب خاصی انتخاب و با آنها سیگنال گفتار بازسازی شد. در یک آزمایش دیگر چند ضریب از ابتدا حذف، و چند ضریب نزدیک به آخر تقویت شده و طیف فرکانسی آنها باهم مقایسه شد.

در این تحقیق نشان داده شده است که با استفاده از ضرایب انتخاب شده می‌توان هویت گوینده را پنهان کرد و در عین حال جمله‌ی گفته شده را ادا نمود.

۳-۵-۶ کدینگ گفتار

در مرجع [۲۷] آورده شده است که به صورت تجربی تعدادی از ضرایب انتخاب شده است. با محاسبه-ی میزان شباهت سیگنال اصلی و بازسازی شده می توان فهمید با داشتن ۱۵ تا ۳۰ ضریب از ضرایب فوریه بسل می توان با کیفیت خوب سیگنال گفتار را بازسازی نمود و اگر ۳۰ ضریبی که در قله قرار می‌گیرند انتخاب نمود کیفیت بسیار خوبی می‌توان به دست آورد. این روش تعداد بیت های کد را به شدت پایین می آورد.

از آنجا که توابع سینوسی توابع ویژه‌ای برای تارهای مرتعش بوده و توابع بسل توابع ویژه‌ای برای لوله های مرتعش می‌باشند و گفتار از درون یک لوله‌ی صوتی تولید می‌شود، می‌توان نتیجه گرفت با تعداد کمی از ضرایب فوریه – بسل، می‌توان گفتار را با کیفیت خوبی بازسازی نمود. این امر باعث می‌شود که بیشترین کیفیت در کمترین نرخ بیت ذخیره شود.

۳-۵-۷ تشخیص لحظه شروع گفتار[۱]

در کاربردهای عمومی پردازش گفتار یکی از مسائل مهم لحظه‌ی شروع ابتدای گفتار از غیر گفتار می-باشد. لحظه‌ی شروع گفتار متشکل از ساختار زمانی و فرکانسی، در بازه‌ی خیلی کوتاه است. که باعث می‌شود تشخیص آن سخت باشد. با این حال این ویژگی کاربرد زیادی در حفاظت گفتار، شناسایی گفتار و شناسایی لهجه دارد.

در تحقیق انجام شده در مراجع [۲۸و۲۹] جهت تشخیص لحظه شروع گفتار از ضرایب فوریه - بسل استفاده شده است. در این تحقیق با انتخاب تعداد خاصی از ضرایب سیگنال گفتار بازسازی می‌شود تا فقط نواحی صدادار و بی‌صدا باقی بمانند. سپس با الگوریتمی به نام جداسازی انرژی گسسته[۲] تابع پوش دامنه‌ی سیگنال گفتار محاسبه می‌شود. که با استفاده از این تابع می‌توان لحظه شروع گفتار را محاسبه نمود.

---

[۱] Voice Onset Time (VOT)
[۲] Discrete Energy Separation Algorithm (DESA)

فصل چهارم:

استفاده از تبدیل فوریه – بسل در پنهان سازی اطلاعات درون سیگنال گفتار

۴-۱ مقدمه

پنهان سازی اطلاعات می‌تواند هم در حوزه‌ی زمان و هم در حوزه‌ی تبدیل سیگنال صورت گیرد. حوزه‌ی تبدیلی که برای پنهان سازی اطلاعات مورد استفاده قرار می‌گیرد بسیار مهم است. اگر اطلاعات در نمونه‌های زمانی سیگنال پنهان شود، با هر گونه تغییری در سیگنال، اطلاعات پنهان شده به طور مستقیم دستخوش تغییر قرار می‌گیرند. اما اگر اطلاعات در حوزه‌ی تبدیل سیگنال پنهان شود تغییرات به صورت غیر مستقیم بر نهان نگاره اثر می‌گذارد. در نتیجه باید فضایی انتخاب شود که در برابر حملات مقاومت لازم را داشته باشد. از جمله تبدیل‌های معروف و متداول برای سیگنال می‌توان به تبدیل فوریه، تبدیل کسینوسی و تبدیل موجک نام برد.

در فصل سوم تبدیل دیگری نیز به نام تبدیل فوریه – بسل معرفی شد که این تبدیل، سیگنال را براساس مجموعی از توابع بسل نمایش می‌دهد. همچنین کاربردهایی از این تبدیل در پردازش سیگنال گفتار ذکر گردید.

در این فصل ابتدا نحوه‌ی پیاده سازی تبدیل فوریه – بسل توضیح داده شده است. سپس نحوه‌ی پنهان سازی اطلاعات در فضای تبدیل فوریه – بسل مورد بررسی قرار گرفته می‌شود. همچنین جهت بررسی کارایی این تبدیل، مقایسه‌ای نیز بین این تبدیل و تبدیل‌های معروف دیگری همچون تبدیل فوریه، تبدیل کسینوسی و تبدیل موجک، صورت گرفته است.

۴-۲ پایگاه داده

پایگاه داده‌ای که در این فصل استفاده می‌شود TIMIT نام دارد. این پایگاه داده شامل ۶۳۰۰ داده‌ی آموزش و آزمایش می‌باشد. این داده‌ها توسط ۶۳۰ گوینده و هر کدام ۱۰ جمله بیان شده است. این افراد با ۸ لهجه مختلف انتخاب شده‌اند. همه این داده‌ها با فرکانس ۱۶ کیلوهرتز نمونه برداری، و با ۱۶ بیت کوانتیزه شده‌اند. در کنار فایل‌های گفتار فایل‌های دیگری برای کاربردهای تشخیص گفتار وجود دارد که برای کاربرد پنهان سازی استفاده نمی‌شود.

۳-۴ پیاده سازی تبدیل فوریه - بسل

همان گونه که در فصل سوم گفته شد هر تابعی مانند $f(t)$ را که روی بازه‌ی $(0,a)$ تعریف شده است، تحت شرایطی می توان به صورت زیر نوشت:

$$f(t)=\sum_{m=1}^{\infty} C_m J_0(\lambda_m t) \qquad (۱-۴)$$

که در آن λ_m به ازای $m=1,2,...$ ریشه‌های مثبت معادله‌ی $J_0(a\lambda_m)=0$ می‌باشد که به ترتیب صعودی مرتب شده‌اند. ضرایب C_m نیز که ضرایب فوریه – بسل نامیده می‌شوند، با استفاده از رابطه‌ی زیر به دست می‌آیند.

$$C_m=\frac{2\int_0^a tf(t)J_0(\lambda_m t)dt}{a^2[J_1(\lambda_m a)]^2} \qquad (۲-۴)$$

انتگرالی که در رابطه‌ی (۴-۲) وجود دارد، تبدیل هنکل محدود نامیده می‌شود. هدف در این بخش به دست آوردن ضرایب C_m می‌باشد. در رابطه‌ی (۴-۲) محاسبه‌ی مخرج کسر نیاز به الگوریتم خاصی ندارد و تابع بسل مرتبه‌ی یک به راحتی قابل محاسبه است. پس مهم محاسبه‌ی صورت کسر است و در نتیجه برای محاسبه‌ی ضرایب C_m کافی است فقط تبدیل هنکل محاسبه شود. الگوریتم محاسبه‌ی تبدیل هنکل که بر پایه‌ی تبدیل فوریه سریع می باشد، محاسبه ضرایب فوریه – بسل را ممکن کرده است.

همان گونه که در فصل سوم گفته شد تقریب تبدیل هنکل محدود به صورت زیر است.

$$F(\lambda)\simeq\frac{(\Delta T)^2}{K}\sum_{k=0}^{K-1}P[\ell\cos(k\Delta\theta)] \qquad (۳-۴)$$

که تابع $P(x)$ به صورت زیر تعریف می‌شود.

$$P(x)=\sum_{n=0}^{N-1} nf(n)e^{-j\frac{2\pi}{N}nx} \qquad (۴-۴)$$

برای این منظور بازه‌ی $(0,2\pi)$ را به K قسمت مساوی به طول $\Delta\theta=\frac{2\pi}{K}$ ، و بازه‌ی $(0,a)$ را به N قسمت مساوی به طول $\Delta T=\frac{a}{N}$ تقسیم می‌کنیم. که رابطه‌ی بین λ با نرخ نمونه برداری (ΔT) و تعداد نمونه‌ها (N) به صورت زیر می‌باشد.

$$\lambda = \frac{\ell.2\pi}{N.\Delta T} \qquad (۵-۴)$$

حال با استفاده از این مطالب می‌توان مراحل الگوریتم محاسبه‌ی تبدیل هنکل را به شرح زیر بیان کرد.

۱. ابتدا سیگنال گفتار ورودی به طول N و هم‌پوشانی دلخواه پنجره بندی می‌شود. هر پنجره را $f(n)$ می‌نامیم.

۲. از روی سیگنال $f(n)$، سیگنال $nf(n)$ به ازای $n = 0\,,1\,,\ldots,N-1$ محاسبه می‌گردد.

۳. تبدیل فوریه‌ی گسسته‌ی سیگنال $nf(n)$ محاسبه می‌شود. سیگنال به دست آمده را $P(n)$ می‌نامیم.

۴. به ازای K دلخواه، $\Delta\theta = \frac{2\pi}{K}$ قرار داده می‌شود. (هر چه K بزرگتر باشد دقت الگوریتم بالاتر است)

۵. دنباله‌ی x با تعریف $x = \ell\cos(k\Delta\theta)$ به ازای $k = 0,1,2,\ldots,K-1$ محاسبه می‌گردد.

۶. سیگنال به دست آمده $P(n)$ در نقاط x درون‌یابی می‌گردد و تمام مقادیر محاسبه شده با هم جمع می‌شود.

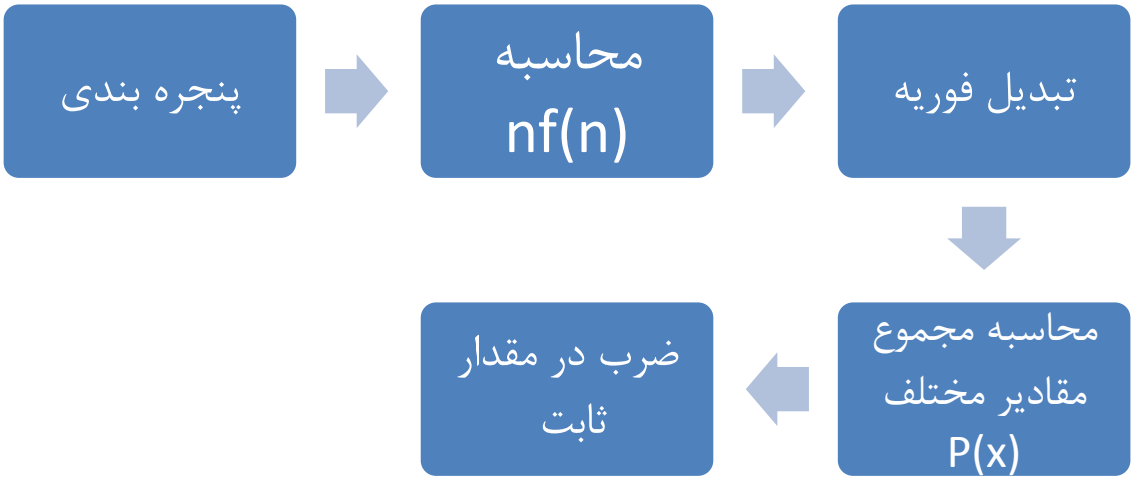
۷. مقدار نهایی در مرحله ششم در مقدار ثابت $\frac{(\Delta T)^2}{K}$ ضرب می‌شود که همان مقدار تبدیل هنکل می‌باشد.

۸. مراحل پنجم تا هفتم برای λ های مختلف تکرار می‌شود تا تعداد دلخواه از ضرایب فوریه – بسل به دست آید. به ازای هر λ یک ضریب فوریه – بسل به دست می‌آید.

در مرحله‌ی ششم باید در نظر داشت که دو روش برای درون‌یابی وجود دارد. روش اول اینکه نزدیکترین عدد صحیح به مقدار x انتخاب شود. روش دوم آن است که هنگام محاسبه‌ی تبدیل فوریه، جهت بالاتر بردن وضوح تبدیل، سیگنال $P(n)$ با تعداد نقاط بیشتری محاسبه شود. سپس نزدیکترین مقدار به x انتخاب می‌شود. با استفاده از روش دوم مقدار دقیق‌تری به دست می‌آید.

همچنین اگر مقدار x از تعداد نقاط $P(n)$ بیشتر بود با استفاده از خاصیت تناوبی بودن تبدیل فوریه گسسته، می‌توان مقدار $P(n)$ را در مقادیر بزرگتر محاسبه نمود.

شکل ۴-۱ بلوک دیاگرام مراحل الگوریتم محاسبه‌ی تبدیل هنکل را به اختصار نشان می‌دهد.



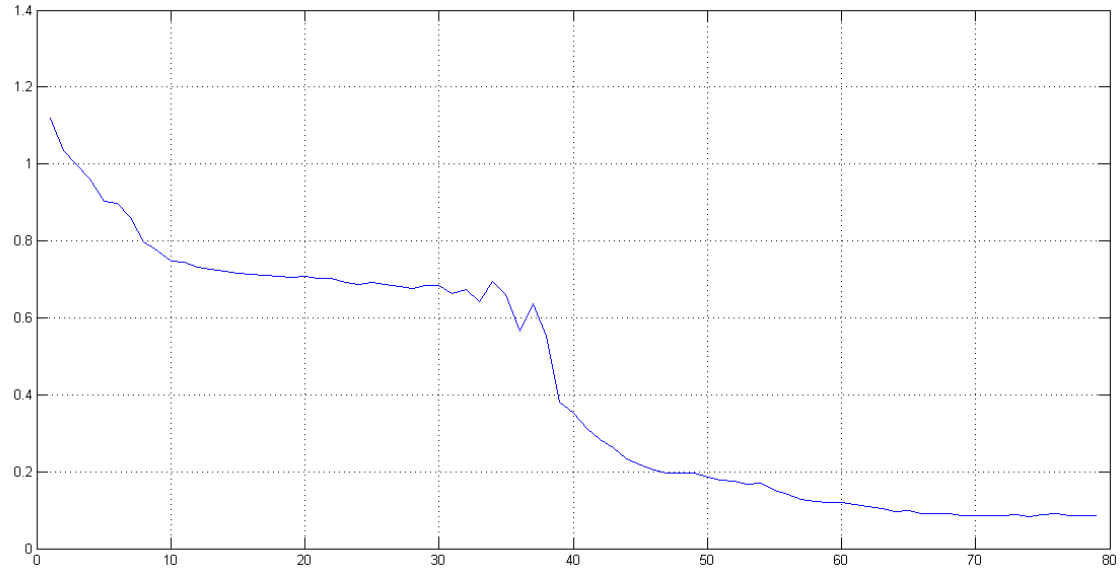
شکل ۴-۱ بلوک دیاگرام مراحل الگوریتم محاسبه‌ی تبدیل هنکل

در بخش ۳-۴ مثالی از تبدیل فوریه بسل یک فریم از سیگنال گفتار ارائه شد که در آن از همین الگوریتم استفاده شده است.

۴-۴ بحث در تعداد ضرایب سری فوریه - بسل

همان گونه که در رابطه‌ی ۴-۱ آمده است می‌توان هر تابعی را به صورت مجموع نامحدودی از توابع بسل نوشت. اما در عمل به دست آوردن تعداد نامحدودی از ضرایب امکان پذیر نیست. معقول‌ترین مقدار برای تعداد ضرایب همان تعداد نمونه‌های خود سیگنال می‌باشد. اما در کاربردهایی همچون فشرده سازی، تعداد ضرایب کمتر از تعداد نمونه‌های سیگنال انتخاب شده است. لذا در این بخش میزان خطایی که با تعداد مختلف از ضرایب به وجود می‌آید مورد بررسی قرار می‌گیرد.

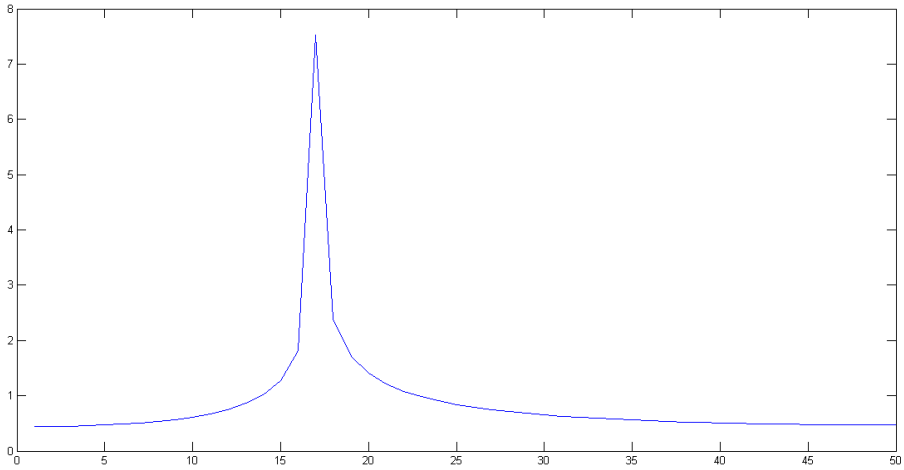
آزمایشی برای این کار در نظر گرفته می‌شود. در این آزمایش یکی از فایل‌های پایگاه داده تیمیت به تصادف انتخاب شده است. این سیگنال گفتار به فریم‌های با طول ۸۰ نمونه تقسیم می‌شود. هر فریم گفتار با تعداد ضرایب مختلف فوریه - بسل، از یک تا ۸۰ ضریب بازسازی می گردد. سپس طبق معیار MSE خطای بازسازی برای هر تعداد از ضرایب محاسبه می‌شود که در نتیجه یک دنباله ۸۰ تایی از خطا به دست می‌آید. این کار برای همه فریم‌ها انجام می‌شود. در نهایت میانگین همه دنباله‌های به دست آمده محاسبه می‌شود. نتایج را در نمودار شکل ۴-۲ مشاهده می کنید.



شکل ۴-۲ نمودار میانگین خطا به ازای تعداد مختلف ضرایب

همان گونه که در نمودار شکل ۴-۳ مشاهده می‌شود، میزان خطای بازسازی با زیاد شدن تعداد ضرایب کاهش می یابد. این کاهش وقتی تعداد ضرایب از نصف تعداد نمونه‌ها می‌گذرد بسیار شدیدتر است. برای توجیه این مطلب باید گفت که هر ضریب نماینده‌ی یک تابع بسل است که قسمتی از اطلاعات سیگنال گفتار را در خود جای داده است. هر چه ضرایب بیشتری به آن اضافه می‌شود سیگنال گفتار کامل‌تر شده و میزان خطای آن کمتر می‌گردد. ضرایبی که در نیمه‌ی اول قرار گرفته-اند حاوی اطلاعات اصلی، و ضرابیب که در نیمه‌ی دوم قرار دارند شامل جزئیات سیگنال می‌شوند. به همین دلیل است که از نیمه به بعد در میزان خطا نزول چندانی مشاهده نمی‌شود. از دید فرکانسی

نیز می‌توان به این موضوع نگاه کرد. در شکل ۴-۳ اندازه‌ی تبدیل فوریه‌ی تابع بسل مرتبه‌ی صفر مشاهده می‌شود. همان گونه که در شکل دیده می‌شود این تابع یک تابع باند باریک می‌باشد. ضرایب مختلف شیفت یافته‌های این تابع را به آن اضافه می‌کنند. چون در تبدیل فوریه‌ی سیگنال گفتار فرکانس‌های بالا وجود ندارد ضرایب اولیه بزرگتر و ضرایب انتهایی کوچک‌تر می‌باشند. به همین دلیل حذف کردن آنها خطای کمی به سیگنال وارد می‌کند.



شکل ۴-۳ اندازه تبدیل فوریه تابع بسل

۴-۵ روش پیشنهادی اول: جایگزینی بیت کم ارزش در ضرایب بسل

همان طور که در بخش توضیح داده شد در روش جایگزینی بیت کم ارزش در هر نمونه از خود سیگنال یا تبدیل یافته سیگنال یک بیت پنهان می‌شود. در این بخش با استفاده از روش جایگزینی بیت کم ارزش به پنهان سازی اطلاعات در فضای تبدیل فوریه – بسل خواهیم پرداخت.

۴-۵-۱ الگوریتم روش پیشنهادی اول

مراحل پیشنهادی جهت پنهان سازی نهان نگاره در قسمت فرستنده به صورت زیر است.

۱- ابتدا سیگنال گفتار مربوطه به فریم‌های با طول زمانی مشخص تقسیم می‌شود. که در اینجا طول هر فریم ۱۰میلی ثانیه در نظر گرفته می‌شود.

۲- از فریم‌های ۱۰میلی ثانیه‌ای سیگنال گفتار، تبدیل فوریه – بسل گرفته می‌شود.

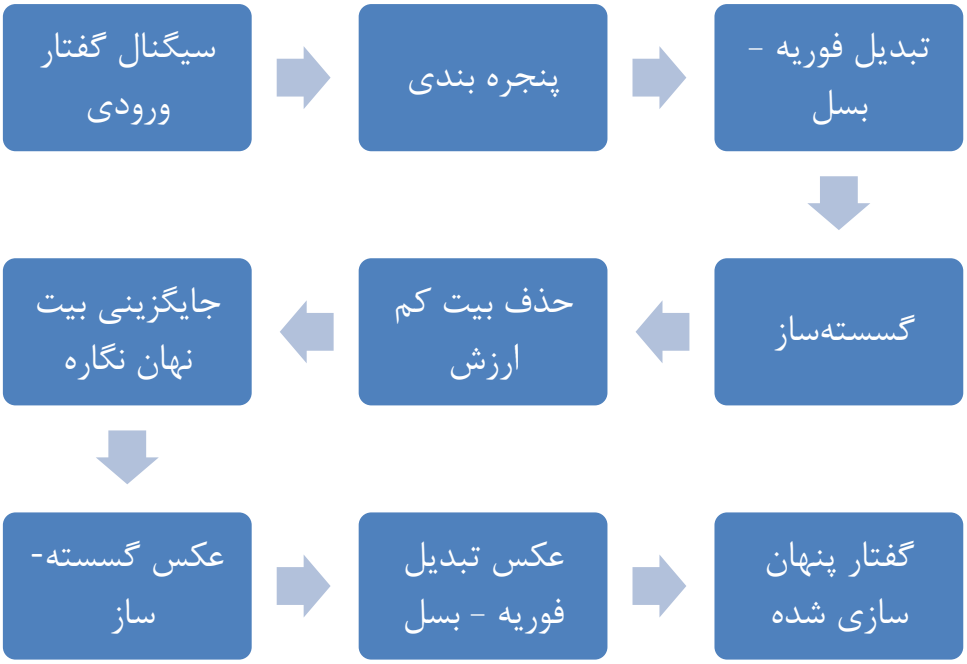
۳- ضرایب فوریه – بسل با تعداد بیت دلخواه کوانتیزه می‌شود که در این پایان نامه ۱۶بیت بـرای هرنمونه در نظر گرفته شده است.

۴- کم ارزش‌ترین بیت هر نمونه از سیگنال اصلی شناسایی شده و صفر می‌گردد.

۵- اطلاعات دودویی نهان نگاره به ضرایب فوریه – بسل اضافه می‌گردد.

۶- فریم‌های سیگنال گفتار با استفاده از تبدیل معکوس فوریه - بسل بازسازی می‌شوند.

شکل ۴-۴ بلوک دیاگرام الگوریتم پنهان سازی گفته شده را نمایش می دهد.



شکل ۴-۴ بلوک دیاگرام الگوریتم پنهان سازی نهان نگاره با استفاده از روش LSB در فضای فوریه - بسل

مراحل تشخیص و بازیابی نهان نگاره نیز در زیر آورده شده است:

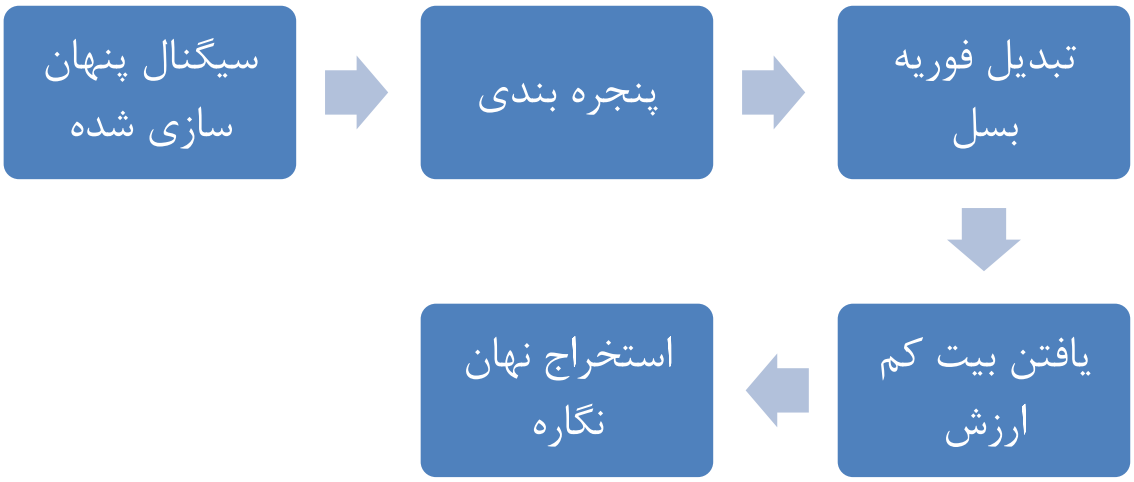
۱. ابتدا سیگنال گفتار پنهان سازی شده به فریم‌های با همان طول زمانی تقسیم می‌شود.

۲. از فریم‌های سیگنال گفتار، تبدیل فوریه – بسل گرفته می‌شود.

۳. ضرایب فوریه – بسل با همان تعداد بیت کوانتیزه می‌شود.

۴. کم ارزش‌ترین بیت هر نمونه از سیگنال گفتار پنهان سازی شناسایی شده که این بیت‌ها همان بیت‌های نهان نگاره می‌باشند.

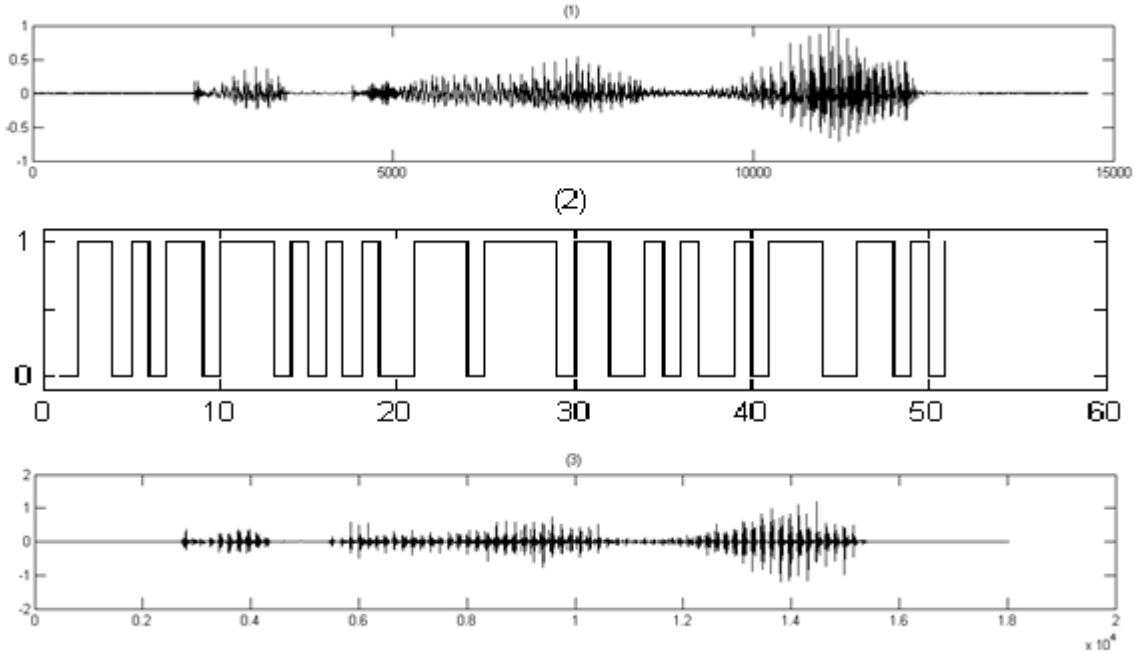
شکل ۴-۵ بلوک دیاگرام الگوریتم تشخیص و بازیابی نهان نگاره را نمایش می دهد.



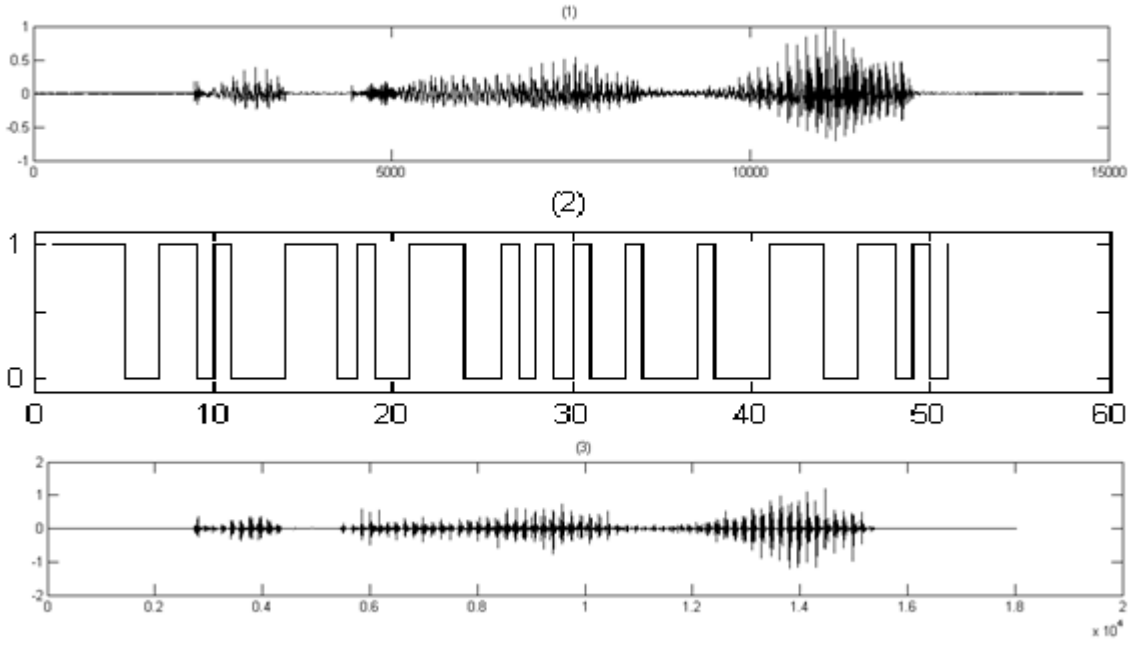
شکل ۴-۵ بلوک دیاگرام الگوریتم تشخیص و بازیابی نهان نگاره با استفاده از روش LSB در فضای فوریه - بسل

۴-۵-۲ نتایج شبیه سازی

الگوریتم گفته شده در بخش ۴-۴-۱ در نرم افزار متلب پیاده سازی شده و بر روی فایل‌های پایگاه داده‌ی تیمیت انجام شد. در شکل ۴-۶ سیگنال گفتار اصلی، قسمتی از نهان نگاره و ضرایب فوریه – بسل سیگنال گفتار اصلی نمایش داده شده است. همچنین در شکل ۴-۷ سیگنال گفتار پنهان سازی شده، قسمتی از نهان نگاره استخراج شده و ضرایب فوریه – بسل سیگنال گفتار اصلی، قابل مشاهده می‌باشد. این کار برای هم پوشانی‌های مختلف از صفر درصد تا ۵۰ درصد انجام شد. برای هر هم پوشانی نرخ خطای بیت و همبستگی بین نهان نگاره‌ی اصلی و استخراج شده با توجه به روابط (۱-۱) و (۱-۲) محاسبه شد. جهت مقایسه بین تبدیل‌های مختلف همین کار برای تبدیل‌های کسینوسی، فوریه و موجک نیز انجام شد. نتایج آن در شکل‌های ۴-۸ و ۴-۹ مشاهده می‌شود.



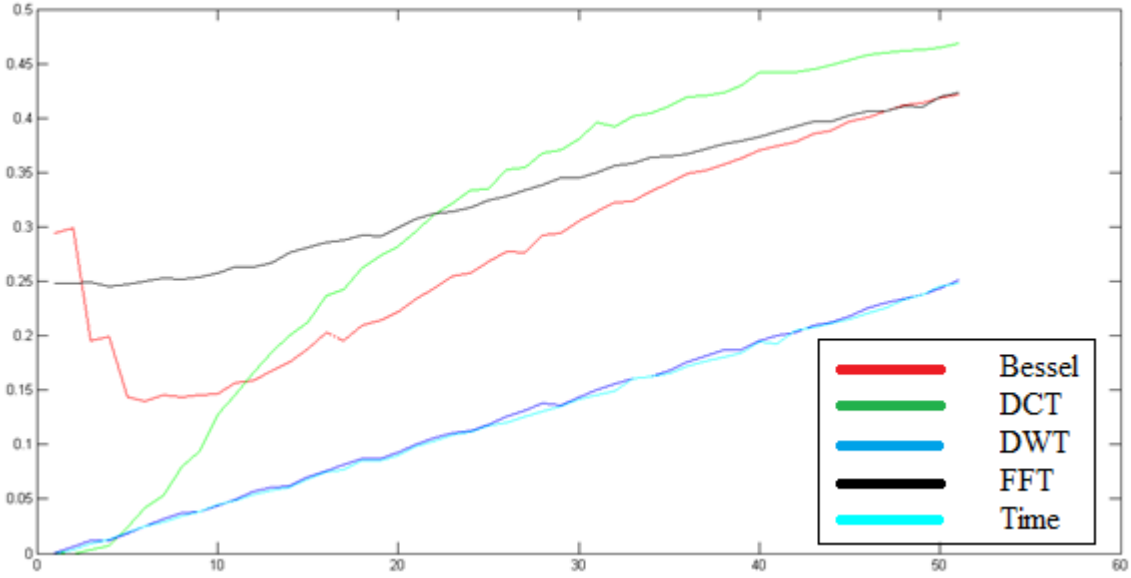
شکل ۴-۶ (۱) سیگنال گفتار اصلی (۲) قسمتی از نهان نگاره (۳) ضرایب فوریه – بسل سیگنال گفتار اصلی



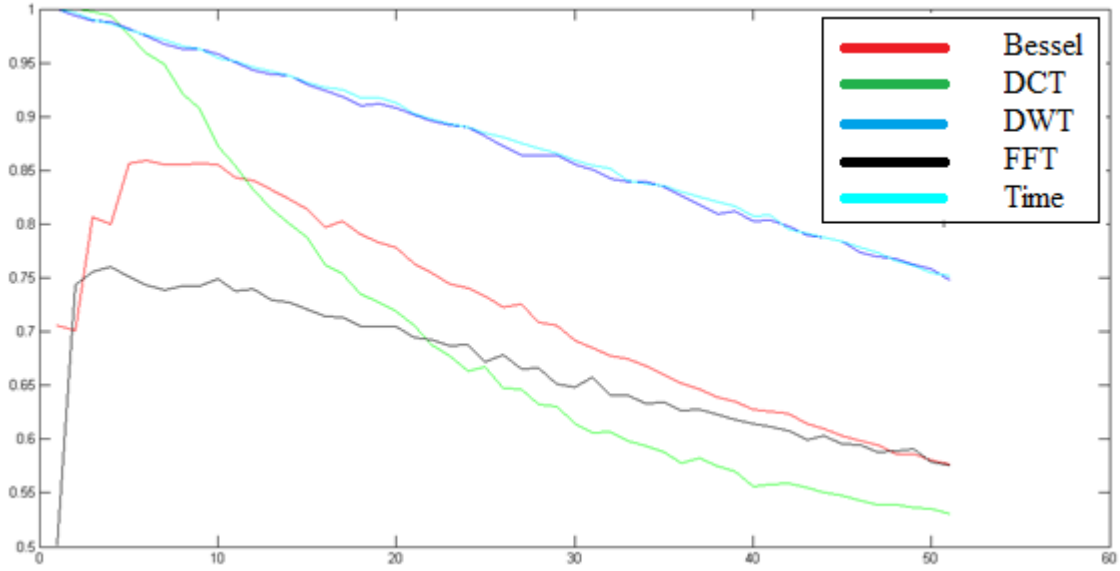
شکل ۴-۷ (۱) سیگنال گفتار پنهان سازی شده (۲) قسمتی از نهان نگاره استخراج شده (۳) ضرایب فوریه – بسل سیگنال گفتار پنهان سازی شده

همان گونه که در شکل ۴-۸ مشاهده می‌شود نرخ خطای بیت برای تبدیل موجک از همه پایین تر است. پس از آن تبدیل فوریه بسل قرار دارد. بین تبدیل کسینوسی و فوریه نیز در ابتدا کسینوسی و سپس فوریه در رتبه سوم قرار دارد. علت اینکه در نمودار تبدیل فوریه – بسل ابتدا نزول و سپس

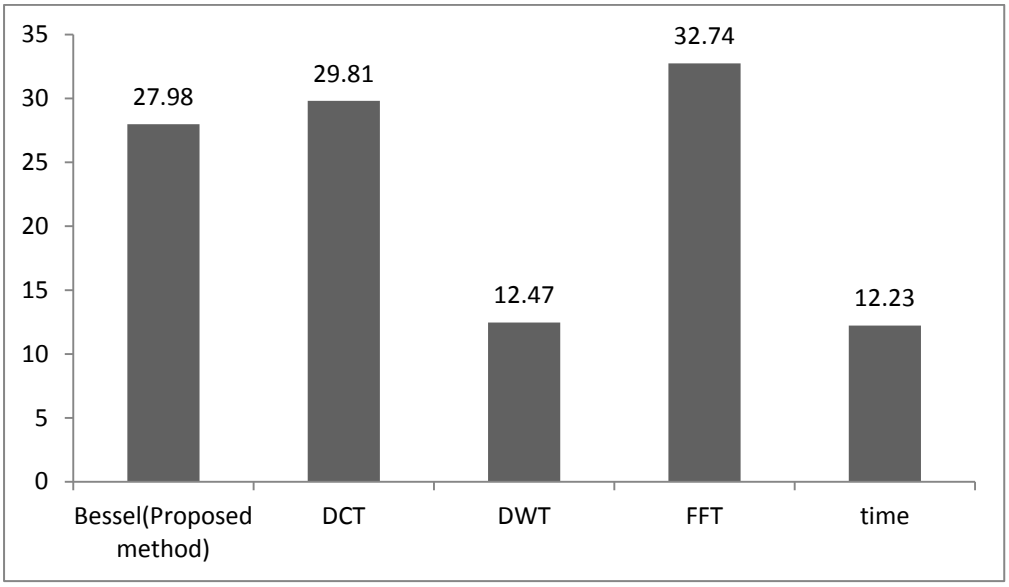
صعود مشاهده می‌شود آن است که هنگام تبدیل فوریه – بسل و بازسازی مجدد سیگنال در چند نمونه اول تا حدی خطا وجود دارد که با در نظر گرفتن کمی هم پوشانی این خطا از بین می رود. در حقیقت خطای اولیه نمودار مربوط به تبدیل می‌باشد نه به نوع سیستم پنهان سازی. برای اینکه به طور کلی مقایسه‌ای بین حوزه‌های مختلف صورت گیرد میانگین خطا برای هر تبدیل محاسبه شده و در نمودار شکل ۴-۱۰ قابل مشاهده می‌باشد. با توجه به این نمودار تبدیل فوریه – بسل در رتبه دوم قرار دارد.



شکل ۴-۸ نمودار نرخ خطای بیت در روش LSB

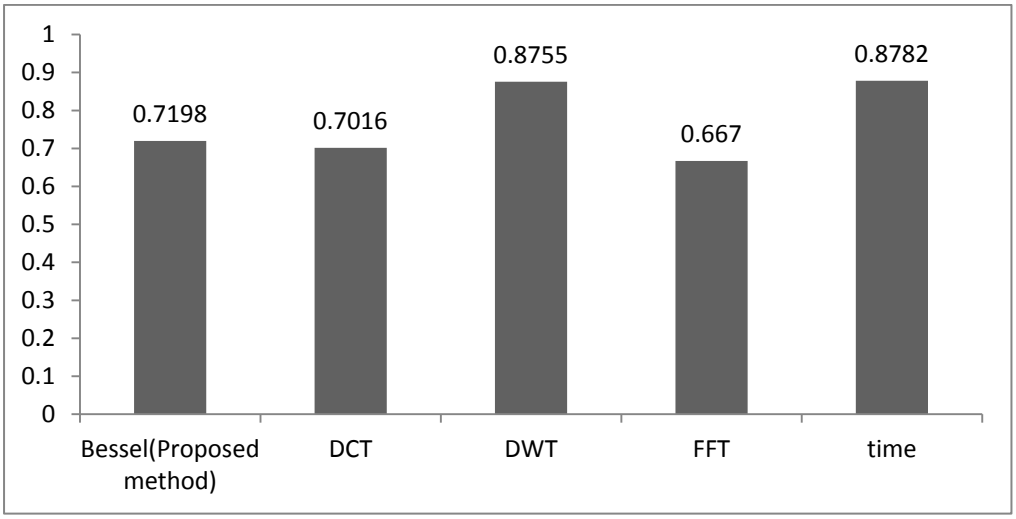


شکل ۴-۹ نمودار همبستگی بین نهان نگاره‌ی اصلی و استخراج شده در روش LSB



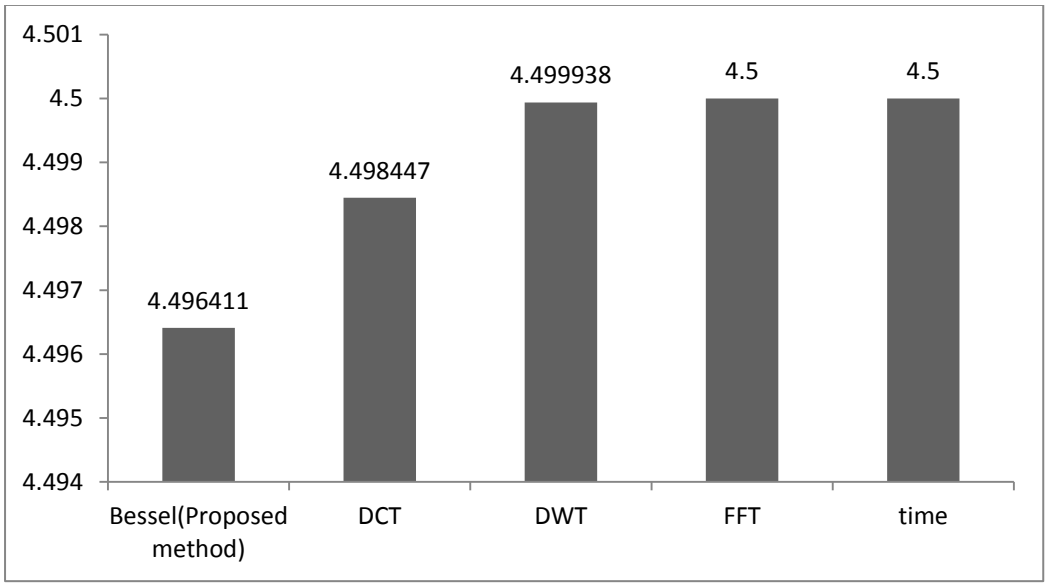
شکل ۴-۱۰ نمودار میانگین خطای نرخ بیت برای تبدیل‌های مختلف در روش LSB

همان گونه که در شکل ۴-۹ مشاهده می‌شود بیشترین میزان همبستگی بین نهان نگاره‌ی اصلی و استخراج شده برای تبدیل موجک می‌باشد. هر چه این میزان به یک نزدیکتر باشد به منزله پایدارتر بودن تبدیل می‌باشد. پس از آن تبدیل فوریه - بسل قرار دارد. علت صعود و نزول نمودار تبدیل فوریه - بسل همان خطای بازسازی در چند نمونه اول می‌باشد که قبلاً توضیح داده شده است. مجدداً برای اینکه مقایسه‌ی کلی بین حوزه‌های مختلف صورت گیرد میانگین همبستگی برای هر تبدیل محاسبه شده و در نمودار شکل ۴-۱۱ قابل مشاهده می‌باشد. با توجه به این نمودار نیز تبدیل فوریه – بسل در رتبه دوم قرار دارد.



شکل ۴-۱۱ نمودار میانگین همبستگی نهان‌نگاره‌ی استخراج شده و اصلی در روش LSB

برای مقایسه‌ی میزان شباهت بین سیگنال گفتار اصلی و پنهان سازی شده از معیار PESQ استفاده شده است. هر چه این میزان به ۴/۵ نزدیکتر باشد نشان دهنده شفافیت بالاتر سیستم است. نتایج این مقایسه در نمودار شکل ۴-۱۲ مشاهده می‌گردد.



شکل ۴-۱۲ مقایسه میزان شباهت سیگنال اصلی و نهان‌نگاری شده با معیار PESQ در روش LSB

با توجه به شکل میزان شباهت در تبدیل فوریه – بسل از سایر تبدیل‌ها پایین‌تر است. البته باید توجه داشت که مقادیر بسیار به یکدیگر نزدیک می‌باشند.

۴-۵-۳ پایداری در برابر فشرده سازی

یکی از حملات مرسومی که به سیگنال وارد می‌شود و می‌تواند موجب تخریب نهان نگاره شود، فشرده‌سازی است. از جمله روش‌های متداول فشرده‌سازی ساختارهای MP3 و GSM 6.10 می-باشد. برای شبیه سازی این حمله، فایل گفتار پنهان سازی شده به دو صورت MP3 و GSM 6.10 در نرم افزار Audacity فشرده سازی شده و سپس دوباره به فرمت اصلی خود یعنی WAV بازگردانده می‌شود. سپس الگوریتم تشخیص نهان نگاره روی آن انجام شده و دو پارامتر نرخ خطای بیت و همبستگی بین نهان نگاره‌ی اصلی و استخراج شده محاسبه می گردد. نتایج این کار در دو جدول ۴-۱ و ۴-۲ مشاهده می‌شود. شایان ذکر است که در فرمت MP3 میزان کیفیت فشرده سازی 128 kbps می‌باشد.

جدول ۴-۱ مقایسه میزان پارامتر BER پس از فشرده سازی در روش LSB

BER	اولیه	ثانویه (MP3)	ثانویه(GSM 6.10)
Bessel	27.1	57.48	50.07
DCT	23.7	76.49	50.24
DWT	7.8	29.50	50.52
FFT	28.2	60.49	49.58
Time[6]	7.46	50.06	50.19

جدول ۴-۲ مقایسه میزان پارامتر NC پس از فشرده سازی در روش LSB

NC	اولیه	ثانویه (MP3)	ثانویه(GSM 6.10)
Bessel	0.72	0.5134	0.4992
DCT	0.77	0.4989	0.4946
DWT	0.92	0.4855	0.4881
FFT	0.72	0.5039	0.5006
Time[6]	0.93	0.4950	0.4993

همان طور که در جداول ۴-۱ و ۴-۲ پیداست تمامی مقادیرثانویه نزدیک به ۵۰ درصد بوده و این به آن معناست که نهان نگاره تقریباً از بین رفته است. زیرا با انتخاب تصادفی بین صفر و یک نیز می‌توان به همین درصدها رسید. قبلاً در بخش ۲-۲ ذکر شده بود که این روش نسبت به انواع حملات غیر مقاوم است. نتایج ارائه شده در جداول ۴-۱ و ۴-۲ نیز تا حدی این مطلب را تأیید می نماید.

۴-۵-۴ پایداری در برابر افزودن نویز

یکی دیگر از حملات رایج که می‌تواند باعث از بین رفتن نهان نگاره شود، اضافه شدن نویز است. برای شبیه سازی این حمله مقداری نویز سفید به سیگنال اضافه کرده تا نسبت سیگنال به نویز به میزان دلخواه برسد. پس از آن با الگوریتمی که در بخش ۴-۴-۱ گفته شد نهان نگاره بازیابی می‌شود. سپس دو پارامتر نرخ خطای بیت و همبستگی بین نهان نگاره‌ی اصلی و استخراج شده محاسبه می‌گردد. این کار برای دو مقدار سیگنال به نویز ۱۰ dB و ۱۵ dB انجام شد. مقادیر به دست آمده در دو جدول ۴-۳ و ۴-۴ ارائه شده است.

با توجه به جداول ۴-۳ و ۴-۴ ، تمامی مقادیر ثانویه نزدیک به ۵۰ درصد بوده و این به آن معناست که نهان نگاره تقریباً از بین رفته است و روش جایگزینی بیت کم ارزش همانند فشرده سازی نسبت به افزودن نویز نیز ناپایدار است.

جدول ۴-۳ مقایسه میزان پارامتر BER پس از افزودن نویز در روش LSB

BER	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	27.1	28.50	16.50
DCT	23.7	30.49	00.50
DWT	7.8	12.50	89.49
FFT	28.2	76.49	75.50
Time[6]	7.46	51.11	50.67

جدول ۴-۴ مقایسه میزان پارامتر NC پس از افزودن نویز در روش LSB

NC	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	0.72	0.5052	0.4950
DCT	0.77	0.5057	0.5016
DWT	0.92	0.4934	0.5013
FFT	0.72	0.5108	0.4967
Time[6]	0.93	0.4926	0.4898

۴-۶ روش پیشنهادی دوم: پنهان سازی طیف گسترده در ضرایب بسل

همان گونه که در بخش ۲-۳-۱ آمده است، این روش یک دنباله‌ی شبه نویز را در طول سیگنال گسترده می‌کند. اگر اطلاعات بیتی الگو را با معادل دو قطبی آن $b=\{-1,1\}$ در نظر گرفته و $r(n)$ دنباله‌ی شبه نویز باشد، رابطه‌ی سیگنال الگوگذاری شده به صورت رابطه‌ی (۴-۶) می باشد.

$$x(n) = s(n) + \alpha br(n) \quad (۴-۶)$$

که $x(n)$ سیگنال الگو گذاری شده و $s(n)$ سیگنال اصلی می‌باشد. مقدار α بر طبق یک رابطه‌ی دوطرفه بین دو معیار پایداری و شفافیت تنظیم می‌شود که با آزمون و خطا به دست می‌آید. در واقع در این روش یک بیت در هر فریم پنهان می‌شود.

استخراج الگو با محاسبه‌ی همبستگی بین سیگنال الگوگذاری شده و دنباله‌ی شبه نویز و سپس آستانه گذاری بدست می‌آید.

۴-۶-۱ الگوریتم روش پیشنهادی دوم

مراحل پیاده سازی این روش در قسمت پنهان سازی به شرح زیر می‌باشد.

۱- ابتدا سیگنال گفتار مربوطه به فریم‌های با طول زمانی مشخص تقسیم می‌شود. که در اینجا طول هر فریم ۱۰میلی ثانیه در نظر گرفته می‌شود.

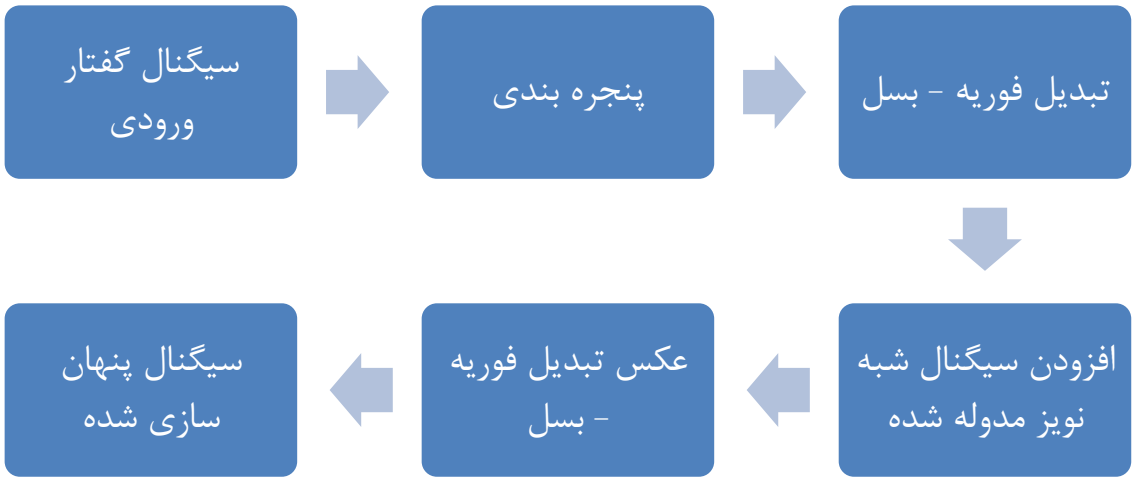
۲- از فریم‌های ۱۰میلی ثانیه‌ای سیگنال گفتار، تبدیل فوریه – بسل گرفته می‌شود.

۳- نهان نگاره از حالت دودیی به حالت باینری تبدیل می‌شود.

۴- سیگنال شبه نویز در هر بیت از نهان نگاره ضرب شده (مدوله شده) و به هرکدام از فریم‌ها افزوده می‌گردد. در این الگوریتم از یک سیگنال آشوب به عنوان سیگنال شبه نویز استفاده شده است. پارامتر α در رابطه‌ی (۴-۶) نیز باید طوری تعیین شود که هم کیفیت سیگنال پس از پنهان سازی تغییر زیادی نکند و هم خطای بازیابی نهان نگاره پایین باشد.

۵- فریم‌های سیگنال گفتار با استفاده از تبدیل معکوس فوریه - بسل بازسازی می‌شوند.

مراحل الگوریتم پنهان سازی به اختصار در بلوک دیاگرام شکل ۴-۱۳ نشان داده شده است.



شکل ۴-۱۳ بلوک دیاگرام الگوریتم پنهان سازی به روش طیف گسترده

مراحل لازم جهت بازیابی و تشخیص نهان نگاره به صورت زیر می‌باشد.

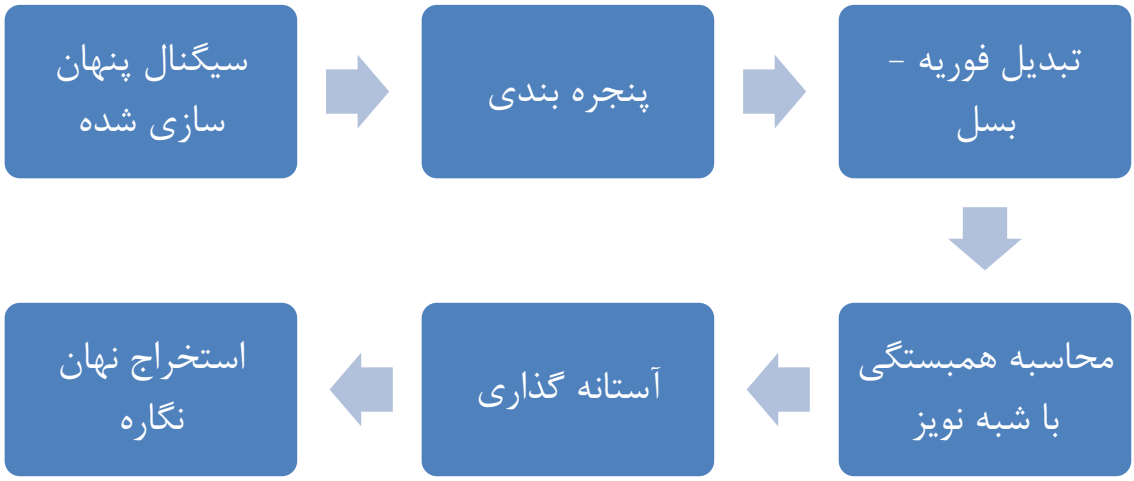
۱- ابتدا سیگنال گفتار مربوطه به فریم‌های با همان طول زمانی تقسیم می‌شود.

۲- از فریم‌های سیگنال گفتار، تبدیل فوریه – بسل گرفته می‌شود.

۳- همبستگی بین سیگنال الگوگذاری شده و دنباله‌ی شبه نویز طبق رابطه‌ی (۲-۳) محاسبه می-گردد.

۴- با استفاده از آستانه گذاری بیت پنهان سازی شده در هر فریم استخراج می‌گردد. میزان آستانه با استفاده از ازمون و خطا تعیین می‌گردد.

در بلوک دیاگرام شکل ۴-۱۴ مراحل الگوریتم بازیابی و تشخیص نهان نگاره به اختصار نشان داده شده است.

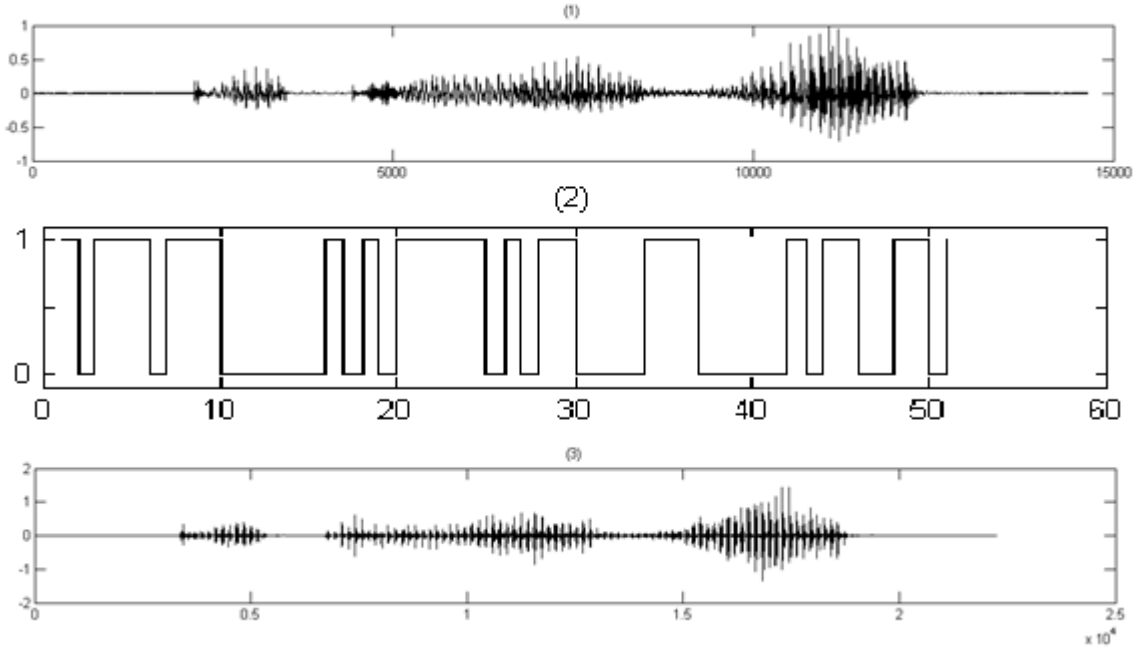


شکل ۴-۱۴ بلوک دیاگرام مراحل الگوریتم بازیابی و تشخیص نهان نگاره در روش طیف گسترده

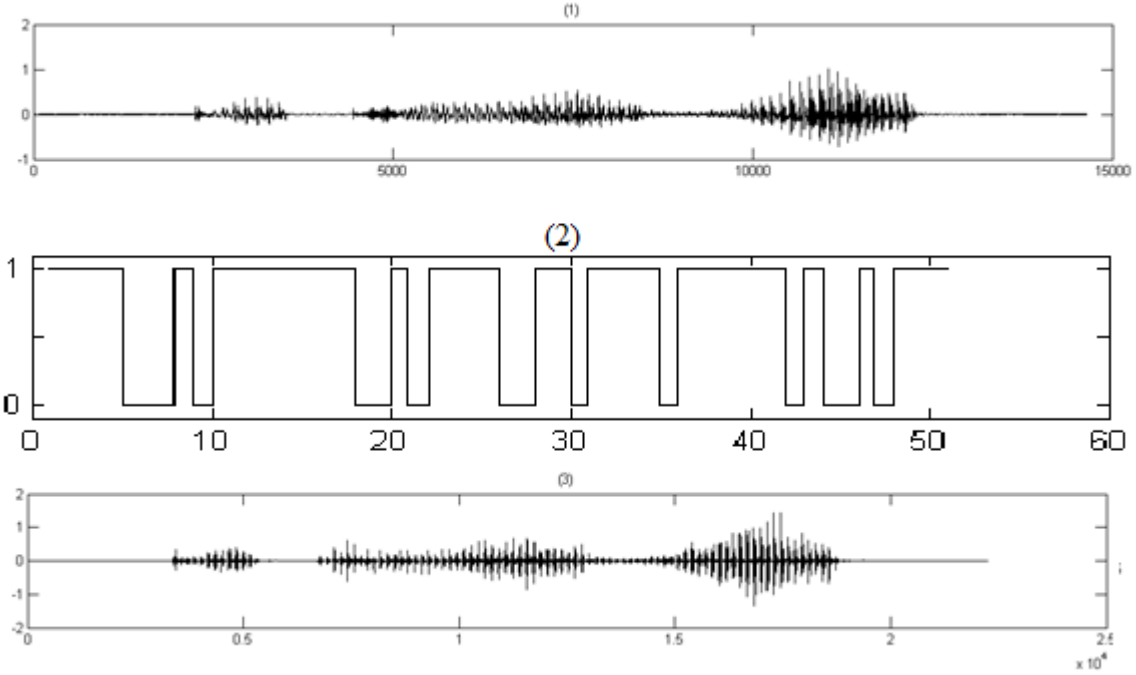
۴-۶-۲ نتایج شبیه سازی

الگوریتم گفته شده در بخش ۴-۵-۱ در نرم افزار متلب پیاده سازی شده و بر روی فایل‌های پایگاه داده‌ی تیمیت انجام شد. در شکل ۴-۱۵ سیگنال گفتار اصلی، قسمتی از نهان نگاره و ضرایب فوریه –

بسل سیگنال گفتار اصلی نمایش داده شده است. همچنین در شکل ۴-۱۶ سیگنال گفتار پنهان سازی شده، قسمتی از نهان نگاره استخراج شده و ضرایب فوریه – بسل سیگنال گفتار اصلی، قابل مشاهده می‌باشد.

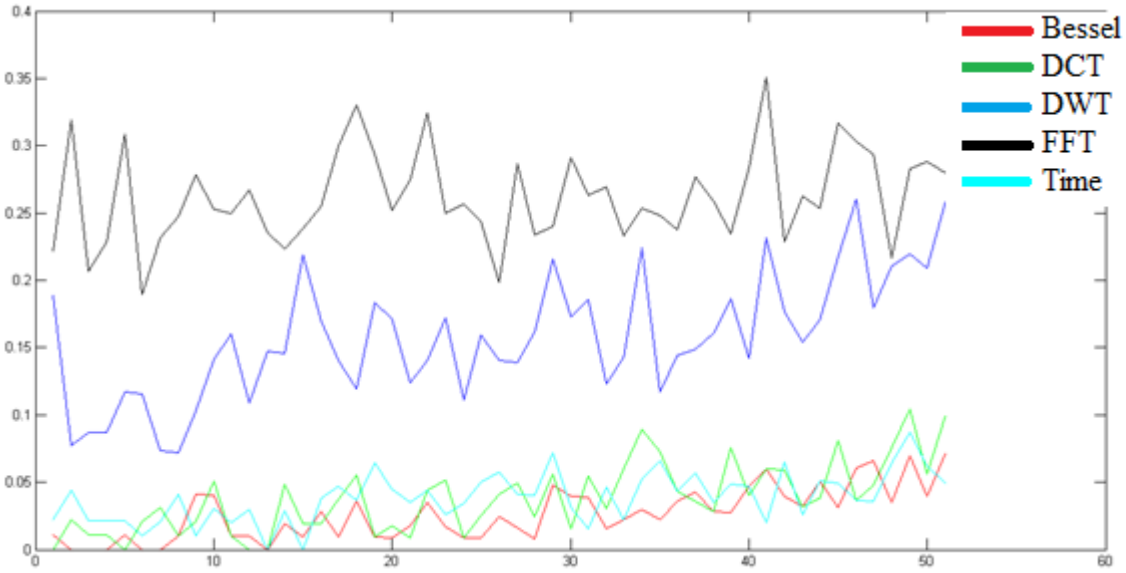


شکل ۴-۱۵ (۱) سیگنال گفتار اصلی (۲) قسمتی از نهان نگاره (۳) ضرایب فوریه – بسل سیگنال گفتار اصلی

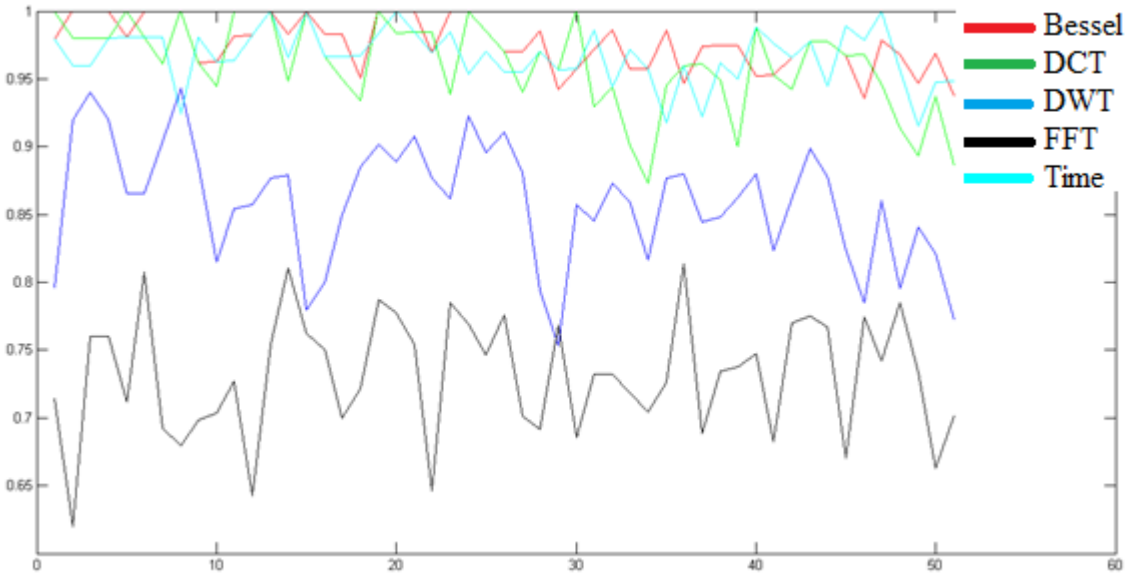


شکل ۴-۱۶ (۱) سیگنال گفتار پنهان سازی شده (۲) قسمتی از نهان نگاره استخراج شده (۳) ضرایب فوریه – بسل سیگنال گفتار پنهان سازی شده

همچون قبل، این کار برای هم پوشانی‌های مختلف از صفر درصد تا ۵۰ درصد انجام شد. برای هر هم پوشانی نرخ خطای بیت و همبستگی بین نهان نگاره‌ی اصلی و استخراج شده با توجه به روابط (۱-۱) و (۱-۲) محاسبه شد. نتایج آن در شکل ۴-۱۷ و ۴-۱۸ مشاهده می‌شود.



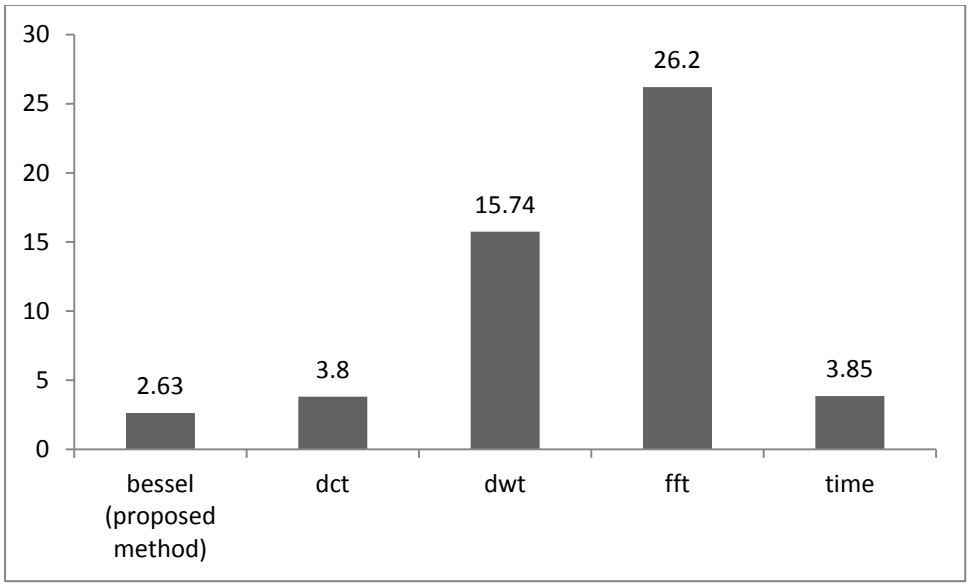
شکل ۴-۱۷ میزان نرخ خطای بیت در روش طیف گسترده



شکل ۴-۱۸ همبستگی بین نهان نگاره‌ی اصلی و استخراج شده در روش طیف گسترده

همان طور که در شکل ۴-۱۷ مشاهده می‌شود کمترین میزان نرخ خطای بیت در اکثر هم پوشانی‌ها متعلق به تبدیل فوریه - بسل می‌باشد. تبدیل‌های کسینوسی، موجک و فوریه به ترتیب در رتبه‌های

بعدی قرار گرفته‌اند. برای مقایسه کلی بین حوزه‌های مختلف میانگین مقادیر مختلف نرخ خطای بیت محاسبه شده و در شکل ۴-۱۹ قابل مشاهده می‌باشد.



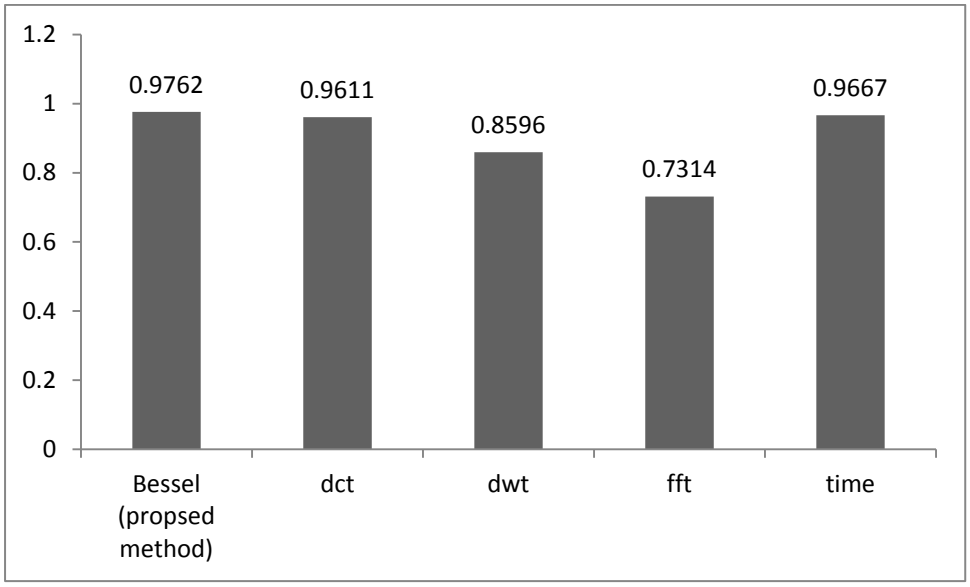
شکل۴-۱۹میانگین نرخ خطای بیت برای روش طیف گسترده

با توجه به شکل ۴-۱۹ تبدیل فوریه – بسل از نظر نرخ خطای بیت از سایر تبدیل‌ها بهتر می‌باشد.

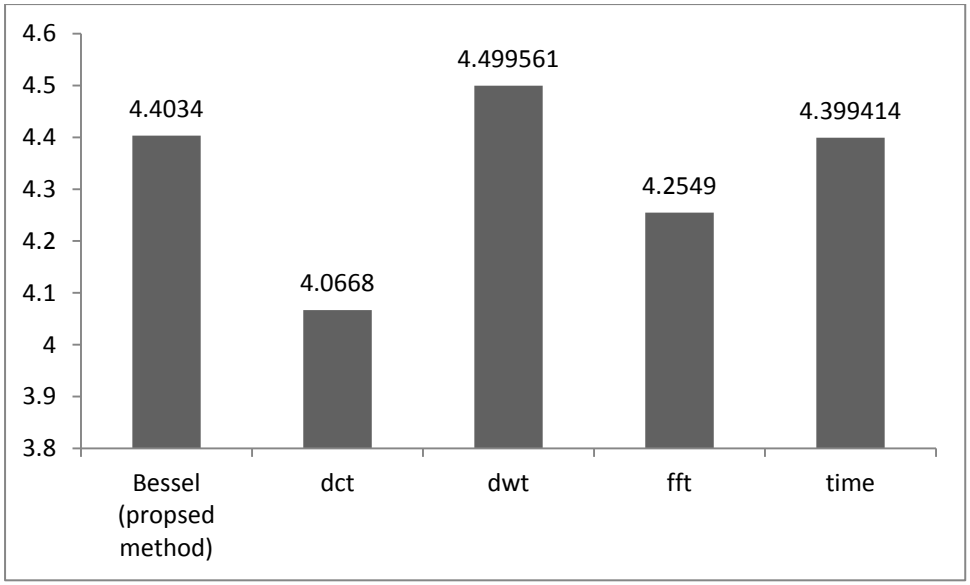
همچنین با توجه به شکل ۴-۱۸ بیشترین میزان هم بستگی بین نهان نگاره‌ی اصلی و استخراج شده، در اکثر هم پوشانی‌ها متعلق به تبدیل فوریه - بسل می‌باشد. میزان هم بستگی در سایر تبدیل‌ها کمتر می‌باشد. میانگین مقادیر مختلف همبستگی برای مقایسه کلی بین حوزه‌های مختلف محاسبه شده و در شکل ۴-۲۰ قابل مشاهده می‌باشد.

با توجه به شکل ۴-۲۰ به طور میانگین تبدیل فوریه – بسل بیشترین میزان همبستگی را در بین سایر تبدیل‌ها دارد.

همچنین برای مقایسه‌ی میزان شباهت بین سیگنال گفتار اصلی و پنهان سازی شده از معیار PESQ استفاده شده است. نتایج این مقایسه در نمودار شکل ۴-۲۱ مشاهده می‌گردد.



شکل ۴-۲۰ نمودار میانگین همبستگی نهان نگاره‌ی استخراج شده و اصلی در روش طیف گسترده



شکل ۴-۲۱ مقایسه میزان شباهت بین سیگنال اصلی و پنهان سازی شده در روش طیف گسترده

۴-۶-۳ پایداری در برابر فشرده سازی

در این بخش نیز همچون بخش ۴-۴-۳ سیگنال گفتاری که با روش طیف گسترده پنهان سازی شده است، به فرمت‌های MP3 و GSM 6.10 تبدیل شده و سپس به فرمت WAV بازگردانده می‌شود.

سپس با استفاده از الگوریتم بازیابی نهان نگاره گفته شده در بخش ۴-۵-۱ نهان نگاره شناسایی می-شود و در نهایت پارامترهای نرخ خطای بیت و همبستگی بین نهان نگاره اصلی و بازیابی شده محاسبه می‌گردد. نتایج حاصله را در جداول ۴-۵ و ۴-۶ مشاهده می‌نمایید.

جدول ۴-۵ مقایسه میزان پارامتر BER پس از فشرده سازی در روش طیف گسترده

BER	اولیه	ثانویه (MP3)	ثانویه (GSM 6.10)
Bessel	3.6	33.81	43.88
DCT	4.3	23.02	29.50
DWT	14.4	48.92	55.40
FFT	23.7	41.01	41.73
Time[7]	41.01	51.80	50.36

جدول ۴-۶ مقایسه میزان پارامتر NC پس از فشرده سازی در روش طیف گسترده

NC	اولیه	ثانویه (MP3)	ثانویه (GSM 6.10)
Bessel	0.95	0.72	0.5200
DCT	0.96	0.80	0.7467
DWT	0.88	0.56	0.5200
FFT	0.81	0.80	0.7600
Time[7]	0.64	0.51	0.4667

با توجه به جدول ۴-۵ در فشرده سازی MP3 تبدیل کسینوسی با کمترین خطا بیشترین پایداری را داشته و پس از آن تبدیل فوریه - بسل و تبدیل فوریه قرار می‌گیرند. در تبدیل موجک نیز نهان نگاره تقریباً از بین رفته و پایداری آن بسیار پایین است.

در فشرده سازی GSM 6.10 نیز تبدیل کسینوسی با کمترین میزان همبستگی بیشترین پایداری را داشته و پس از آن تبدیل فوریه و تبدیل فوریه - بسل قرار می‌گیرند. همانند قبل در تبدیل موجک نهان نگاره از بین رفته و پایداری آن بسیار پایین است.

مشابه همین وضعیت از نظر میزان همبستگی در جدول ۴-۶ مشاهده می‌گردد. از بین رفتن نهان نگاره در تبدیل موجک در این جدول نیز قابل رؤیت است.

۴-۶-۴ پایداری در برابر افزودن نویز

در این روش نیز جهت بررسی پایداری در برابر نویز به سیگنال پنهان سازی شده نویز سفید اظافه شده است تا سیگنال به نویزهای dB ۱۰ و dB ۱۵ به دست آید. سپس با الگوریتم گفته شده در بخش ۴-۵-۱ نهان نگاره از سیگنال نویزی شده بازیابی می‌گردد. در نهایت نرخ خطای بیت و هم بستگی بین نهان نگاره اصلی و استخراج شده محاسبه می‌شود. نتایج به دست آمده در جداول ۴-۷ و ۴-۸ قابل رؤیت است.

جدول ۴-۷ مقایسه میزان پارامتر BER پس از افزودن نویز در روش طیف گسترده

BER	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	3.6	48.92	45.32
DCT	4.3	27.34	23.74
DWT	14.4	52.52	49.64
FFT	23.7	32.37	35.25
Time[7]	41.01	53.24	51.08

جدول ۴-۸ مقایسه میزان پارامتر NC پس از افزودن نویز در روش طیف گسترده

NC	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	0.95	0.6133	0.6400
DCT	0.96	0.7733	0.7867
DWT	0.88	0.5733	0.5733
FFT	0.81	0.8267	0.8133
Time[7]	0.64	0.4933	0.5067

با توجه به جدول ۴-۷ فقط تبدیل کسینوسی و فوریه در مقابل افزودن نویز مقاومت داشته و در تبدیل موجک و فوریه – بسل نهان نگاره تقریباً از بین رفته است. مشابه این مطلب در جدول ۴-۸ نیز مشاهده می‌شود.

۴-۷ روش پیشنهادی سوم: پنهان سازی به روش مدولاسیون اندیس کوانتیزاسیون در ضرایب بسل

با توجه به مطالب گفته شده در بخش ۲-۵ این روش به این صورت است که دو دسته‌ی گسسته‌ساز انتخاب می‌شود. با توجه به سیگنال الگو، سیگنال میزبان بر یکی از سطوح گسسته‌ساز قرار می‌گیرد. در آشکارساز نیز با توجه به فاصله‌ی سیگنال میزبان به دو دسته‌ی سطوح گسسته‌ساز ، تصمیم گیری برای استخراج بیت صفر یا یک انجام می‌شود.

در این روش نیز همچون روش طیف گسترده یک بیت در هر فریم پنهان می‌شود.

۴-۷-۱ الگوریتم روش پیشنهادی سوم

مراحل پیشنهادی جهت پنهان سازی نهان نگاره در قسمت فرستنده به صورت زیر است.

۱- ابتدا سیگنال گفتار مربوطه به فریم‌های با طول زمانی مشخص تقسیم می‌شود. که در اینجا طول هر فریم ۱۰میلی ثانیه در نظر گرفته می‌شود.

۲- از فریم‌های ۱۰میلی ثانیه‌ای سیگنال گفتار، تبدیل فوریه – بسل گرفته می‌شود.

۳- اولین داده از هر فریم با استفاده از رابطه ۲-۱۶ الگوگذاری می‌شود.

۴- از هر فریم عکس تبدیل فوریه – بسل گرفته می‌شود.

مراحل الگوریتم پنهان سازی به اختصار در بلوک دیاگرام شکل ۴-۲۲ نشان داده شده است.

مراحل بازیابی و تشخیص نهان نگاره نیز به صورت زیر می‌باشد.

۱- ابتدا سیگنال گفتار پنهان سازی شده مربوطه به فریم‌های با همان طول زمانی تقسیم می‌شود.

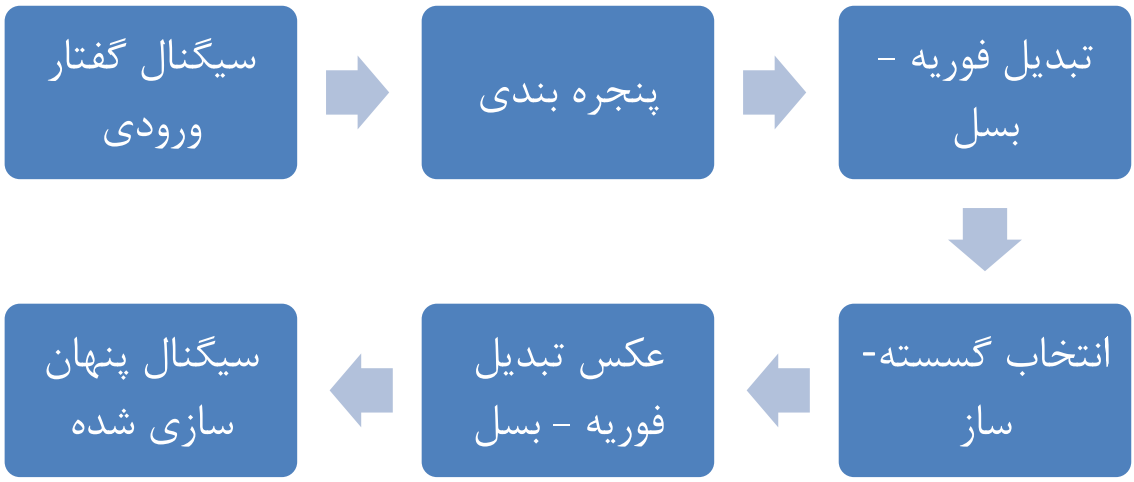
۲- از فریم‌های سیگنال گفتار پنهان سازی شده، تبدیل فوریه – بسل گرفته می‌شود.

۳- اولین داده از هر فریم، در دو رابطه‌ی ۲-۱۷ و ۲-۱۸ قرار داده می‌شود. یکی متعلق به بیت صفر و دیگری متعلق به بیت یک می‌باشد.

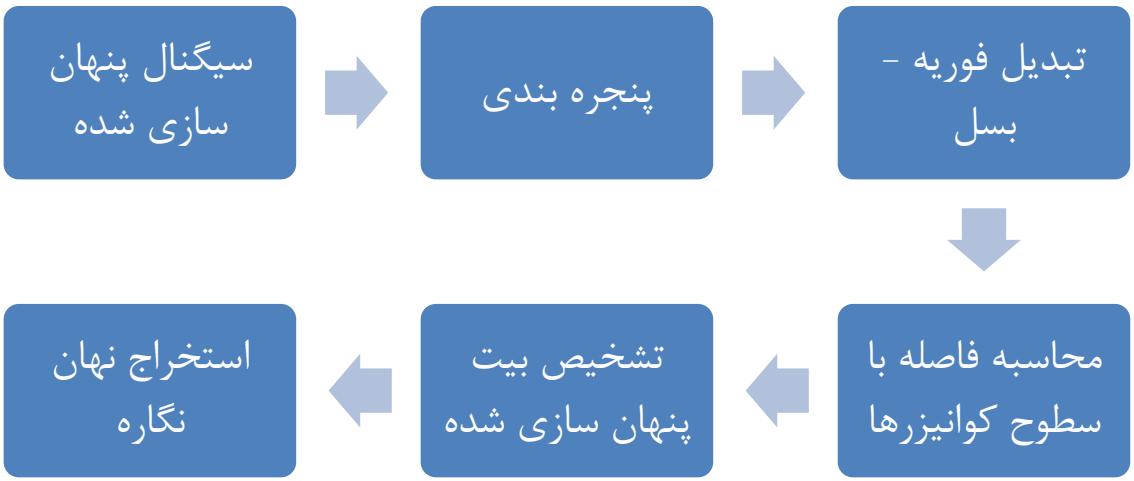
۴- فاصله اقلیدسی هر دو مقدار جدید به دست آمده، با داده اصلی حساب می‌شود.

۵- کوتاه‌ترین فاصله نشانگر بیت پنهان شده می‌باشد.

بلوک دیاگرام شکل ۴-۲۳ مراحل الگوریتم بازیابی و تشخیص نهان نگاره را به اختصار نشان می‌دهد.



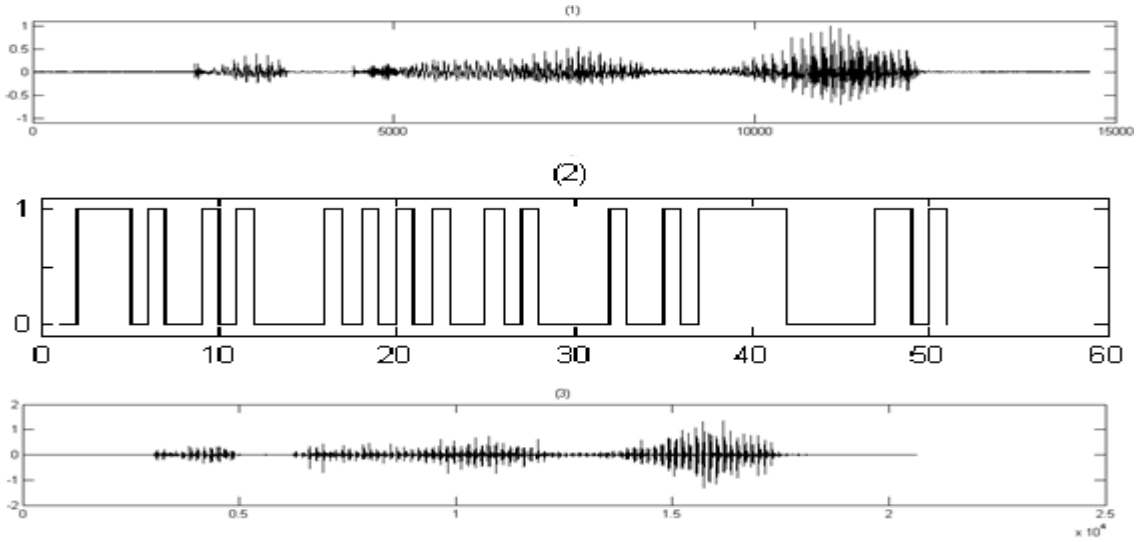
شکل ۴-۲۲ بلوک دیاگرام الگوریتم پنهان سازی در روش QIM



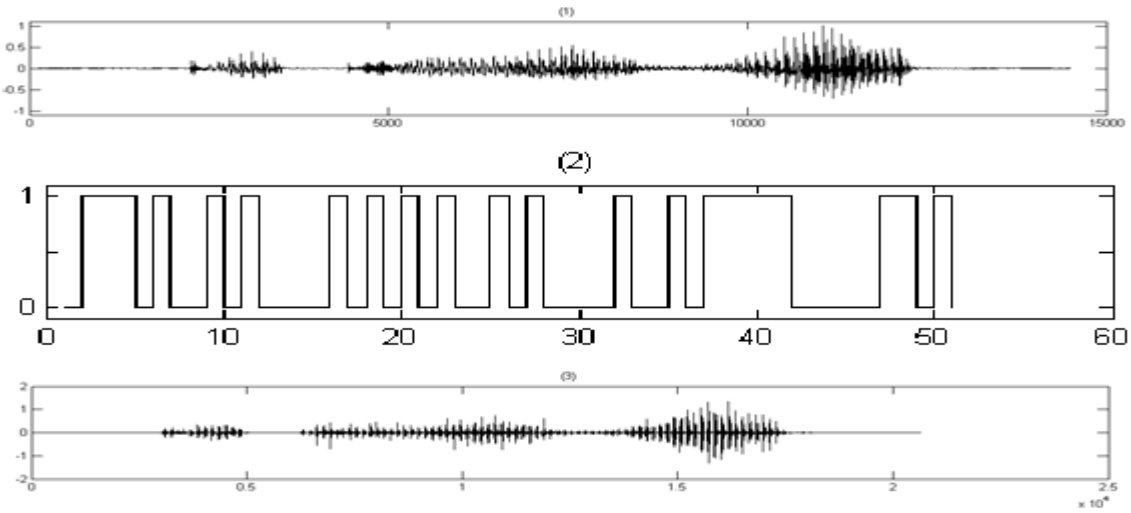
شکل ۴-۲۳ بلوک دیاگرام مراحل الگوریتم بازیابی و تشخیص نهان نگاره در روش QIM

۴-۷-۲ نتایج شبیه سازی

الگوریتم پنهان سازی و بازیابی گفته شده در بخش ۴-۶-۱ در نرم افزار متلب پیاده سازی شده و بر روی فایل‌های پایگاه داده‌ی تیمیت انجام شد. در شکل ۴-۱۵ سیگنال گفتار اصلی، قسمتی از نهان نگاره و ضرایب فوریه – بسل سیگنال گفتار اصلی نمایش داده شده است. همچنین در شکل ۴-۱۶ سیگنال گفتار پنهان سازی شده، قسمتی از نهان نگاره استخراج شده و ضرایب فوریه – بسل سیگنال گفتار اصلی، قابل مشاهده می‌باشد.



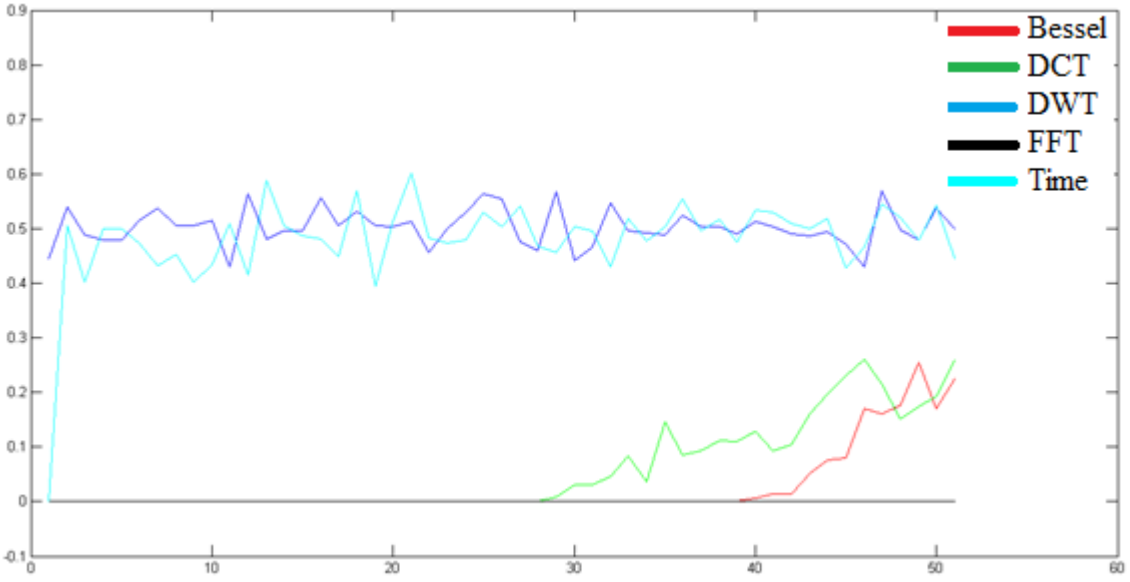
شکل ۴-۲۴ (۱) سیگنال گفتار اصلی (۲) قسمتی از نهان نگاره (۳) ضرایب فوریه – بسل سیگنال گفتار اصلی



شکل ۴-۲۵ (۱) سیگنال گفتار پنهان سازی شده (۲) قسمتی از نهان نگاره استخراج شده (۳) ضرایب فوریه – بسل سیگنال گفتار پنهان سازی شده

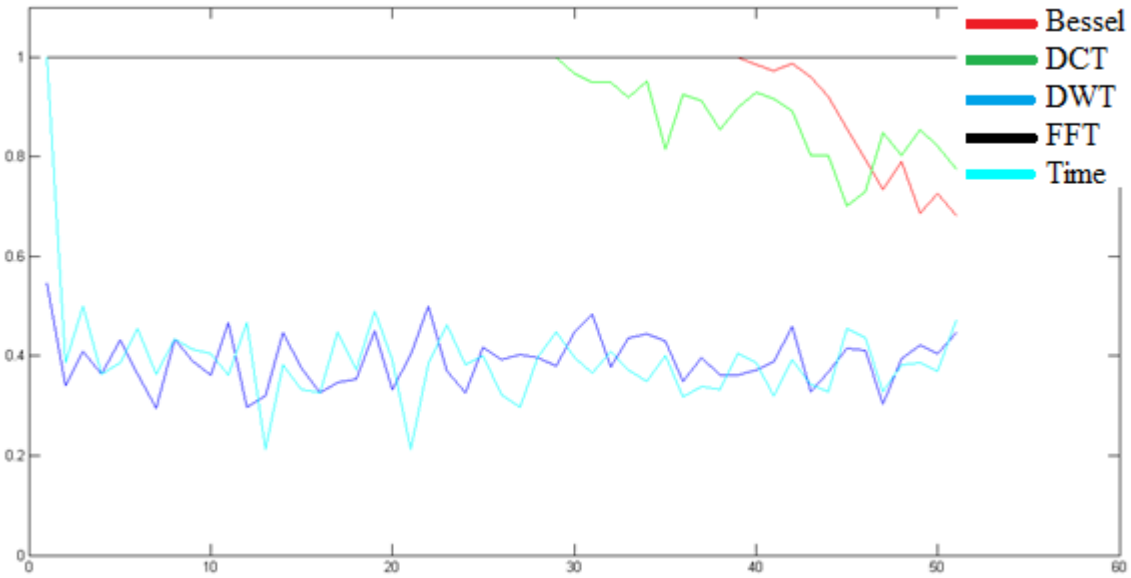
این الگوریتم برای هم پوشانی‌های مختلف از صفر درصد تا ۵۰ درصد اجرا شد. برای هر هم پوشانی نرخ خطای بیت و همبستگی بین نهان نگاره‌ی اصلی و استخراج شده با توجه به روابط (۱-۱) و (۱-۲) محاسبه شد. نتایج آن در شکل ۴-۲۶ و ۴-۲۷ مشاهده می‌شود.

باتوجه به شکل ۴-۲۶ میزان نرخ خطای بیت برای تبدیل‌های فوریه همواره صفر می‌باشد. در تبدیل فوریه – بسل این مقدار در ابتدا صفر بوده و از هم‌پوشانی حدود ۴۰ درصد به بعد به میزان خطا افزوده می‌گردد. در تبدیل کسینوسی در ابتدا میزان خطا صفر بوده و از هم‌پوشانی حدود ۲۵ درصد به بعد به این مقدار بالا می‌رود. تبدیل موجک رفتار خوبی از خود نشان نمی‌دهد و نمودار آن حاکی از نابودی نهان نگاره است. شاید به این دلیل است که سایرتبدیل‌ها حوزه‌ی سیگنال را تغییر می‌دهند ولی تبدیل موجک تقریبی از خود سیگنال ارائه می‌دهد. برای مقایسه‌ی کلی بین حوزه‌های مختلف

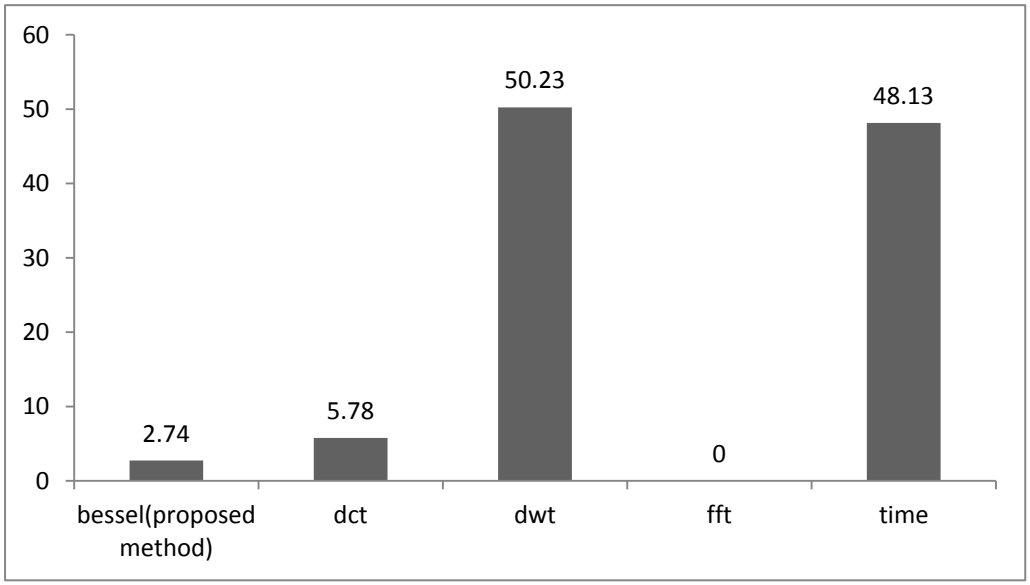


شکل ۴-۲۶ میزان نرخ خطای بیت در روش QIM

میانگین مقادیر مختلف نرخ خطای بیت محاسبه شده و در شکل ۴-۲۸ قابل مشاهده می‌باشد. با توجه به این شکل تبدیل فوریه – بسل از نظر نرخ خطای بیت در رتبه دوم قرار دارد.



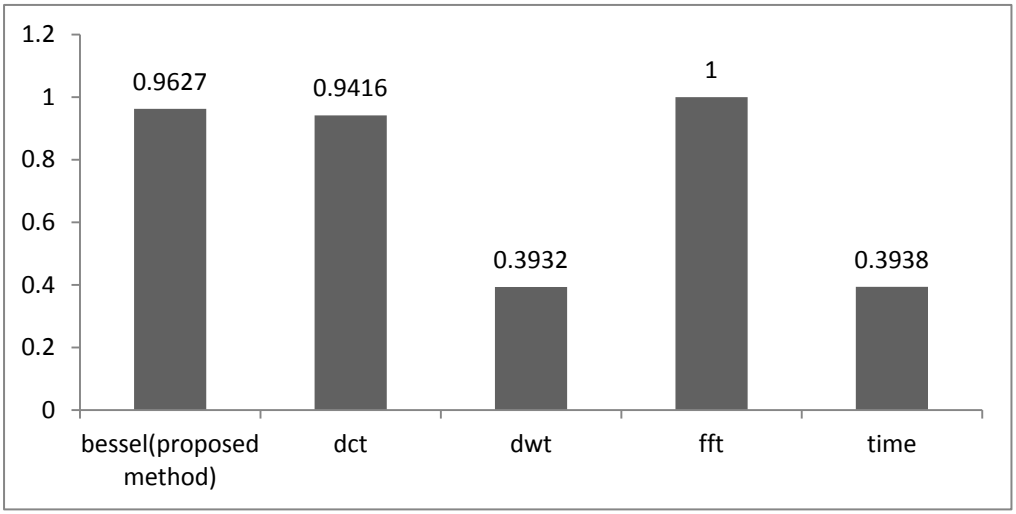
شکل ۴-۲۷ همبستگی بین نهان نگاره‌ی اصلی و استخراج شده در روش QIM



شکل ۴-۲۸ میانگین نرخ خطای بیت در روش QIM

همان گونه که در شکل ۴-۲۷ پیداست میزان همبستگی بین نهان نگاره‌ی اصلی و استخراج شده، برای تبدیل فوریه همواره برابر یک می‌باشد. در تبدیل فوریه – بسل این مقدار در ابتدا یک بوده و از هم‌پوشانی حدود ۴۰ درصد به بعد از میزان همبستگی کاسته می‌گردد. در تبدیل کسینوسی در ابتدا میزان همبستگی یک بوده و از هم‌پوشانی حدود ۲۵ درصد به بعد به این مقدار پایین می‌آید. تبدیل

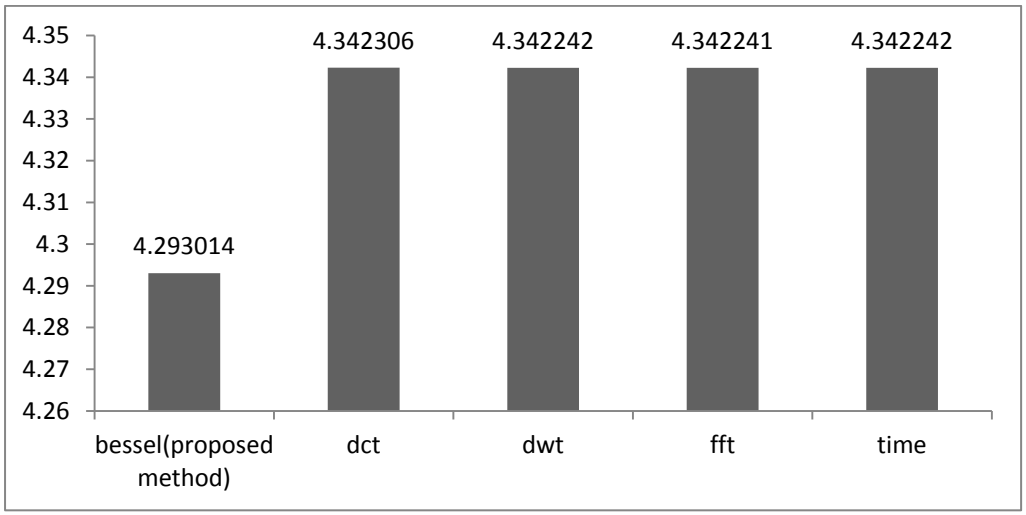
موجک در این نمودار نیز نتایج ضعیفی از خود نشان می‌دهد. میانگین مقادیر مختلف همبستگی برای مقایسه کلی بین حوزه‌های مختلف محاسبه شده و در شکل ۴-۲۹ نمایش داده شده است.



شکل ۴-۲۹ میانگین مقادیر هم بستگی در روش QIM

با دقت در نمودار شکل ۴-۲۹ ، تبدیل فوریه – بسل از نظر میزان هم بستگی نیز پس از تبدیل فوریه قرار گرفته است.

همچنین مقایسه‌ای بین تبدیل‌های مختلف از نظر میزان شباهت بین سیگنال گفتار اصلی و پنهان سازی شده با استفاده از معیار PESQ انجام شده است. نتایج این مقایسه در نمودار شکل ۴-۳۰ مشاهده می‌گردد.



شکل ۴-۳۰ مقایسه میزان شفافیت در روش QIM

همان طور که در نمودار شکل ۴-۳۰ پیداست تبدیل فوریه – بسل از پایین‌ترین میزان شفافیت برخوردار است. این نتیجه حاکی از این است که باید در جهت الگوریتم محاسبه تبدیل فوریه – بسل کارهای بیشتری صورت گیرد. در این نمودار بیشترین میزان شفافیت متعلق به تبدیل کسینوسی بوده و تبدیل‌های فوریه و موجک در ردیف‌های بعدی قرار دارند.

۴-۷-۳ پایداری در برابر فشرده سازی

جهت محاسبه میزان پایداری در برابر فشرده سازی، سیگنال گفتاری که با روش مدولاسیون اندیس کوانتیزاسیون پنهان سازی شده است، به فرمت‌های MP3 و GSM 6.10 تبدیل شده و سپس به فرمت WAV بازگردانده می‌شود. سپس با استفاده از الگوریتم بازیابی نهان نگاره گفته شده در بخش ۴-۶-۱ بازیابی نهان نگاره انجام می‌شود و در نهایت پارامترهای نرخ خطای بیت و همبستگی بین نهان نگاره اصلی و بازیابی شده محاسبه می‌گردد. نتایج حاصله را در جداول ۴-۹ و ۴-۱۰ مشاهده می‌نمایید.

جدول ۴-۹ مقایسه میزان پارامتر BER پس از فشرده سازی در روش QIM

BER	اولیه	ثانویه(MP3)	ثانویه(GSM 6.10)
Bessel	0	0	27.91
DCT	8.6	28.68	55.81
DWT	52.5	48.06	52.71
FFT[30]	0	44.19	51.94
Time	49.61	51.94	48.06

جدول ۴-۱۰ مقایسه میزان پارامتر NC پس از فشرده سازی در روش QIM

NC	اولیه	ثانویه(MP3)	ثانویه(GSM 6.10)
Bessel	1	1.0000	0.6833
DCT	0.92	0.7000	0.4333
DWT	0.35	0.3833	0.4000
FFT[30]	1	0.5167	0.6000
Time	0.3667	0.3000	0.4000

با توجه به نتایج جدول ۴-۹ مشاهده می‌شود، در فشرده سازی به صورت MP3، تبدیل فوریه - بسل همچنان نرخ خطای بیت خود را حفظ کرده است. که این، پایداری بالای تبدیل فوریه – بسل را نشان می‌دهد. پس از آن تبدیل کسینوسی قرار دارد. در تبدیل فوریه تقریباً نهان نگاره از بین رفته است. در تبدیل موجک با توجه با اینکه نهان نگاره قبلاً از بین رفته بود میزان آن به طور تصادفی پایین آمده است. در فشرده سازی به صورت GSM 6.10 نیز تنها تبدیلی که نهان نگاره را حفظ کرده است تبدیل فوریه - بسل می‌باشد و در سایر تبدیل‌ها نهان نگاره از بین رفته است.

با دقت در جدول ۴-۱۰ نیز مطلب فوق تأیید می‌شود. یعنی در فشرده سازی MP3 ، تبدیل فوریه – بسل همچنان مقدار همبستگی خود را حفظ کرده است. در فشرده سازی GSM 6.10 نیز مقدار همبستگی برای تبدیل فوریه – بسل بیشترین مقدار است.

۴-۷-۴ پایداری در برابر افزودن نویز

برای تست پایداری در برابر افزودن نویز، به سیگنال گفتار پنهان سازی شده نویز افزوده تا مقادیر سیگنال به نویزهای dB ۱۰ و dB ۱۵ به دست آید. سپس با استفاده از الگوریتم بازیابی نهان نگاره گفته شده در ۴-۶-۱ نهان نگاره شناسایی شده و پارامترهای نرخ خطای بیت و همبستگی بین نهان نگاره اصلی و استخراج شده محاسبه می‌گردد. نتایج به دست آمده در جداول ۴-۱۱ و ۴-۱۲ آمده است.

جدول ۴-۱۱ مقایسه میزان پارامتر BER پس از افزودن نویز در روش QIM

BER	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	0	97.17	0
DCT	8.6	56.51	44.48
DWT	52.5	88.46	03.57
FFT[30]	0	56.51	88.46
Time	49.61	61.72	50.78

جدول ۴-۱۲ مقایسه میزان پارامتر NC پس از افزودن نویز در روش QIM

NC	اولیه	ثانویه (SNR=10 dB)	ثانویه(SNR=15 dB)
Bessel	1	0.8475	1.0000
DCT	0.92	0.4407	0.4915
DWT	0.35	0.4746	0.3220
FFT[30]	1	0.4915	0.4915
Time	0.3667	0.3729	0.4068

با توجه به نتایج نشان داده شده در جدول ۴-۱۱ ، تبدیل فوریه – بسل تنها تبدیلی است که نهان نگاره در آن پایدار مانده است. در سیگنال به نویز dB ۱۵ خطا صفر مانده است و در سیگنال به نویز dB ۱۰ خطا حدود ۱۸ درصد شده است. این به معنای پایداری بالای این تبدیل در برابر افزودن نویز می‌باشد. در سایر تبدیل‌ها نهان نگاره از بین رفته است.

جدول ۴-۱۲ نیز همین مطالب را تأیید می‌نماید که تبدیل فوریه – بسل نسبت به افزودن نویز به میزان خوبی پایداری نشان می‌دهد.

فصل پنجم:

نتیجه گیری کلی و پیشنهادات آینده

با توجه به عنوان پایان نامه تمرکز اصلی تحقیقات بر روی استفاده از تبدیل فوریه – بسل در سیستم-های پنهان سازی درون سیگنال گفتار قرار داده شد. هدف اصلی این پایان نامه بررسی کارآیی این تبدیل و مقایسه آن با سایر تبدیل‌های متداول همچون تبدیل کسینوسی، فوریه و موجک بوده است.

۵-۱ نتیجه گیری کلی

در این پایان نامه از تبدیل فوریه - بسل در سه روش جایگزینی بیت کم ارزش، طیف گسترده و مدولاسیون اندیس کوانتیزاسیون استفاده شد. همچنین جهت مقایسه، در کنار آن از سه تبدیل فوریه، کسینوسی و موجک نیز استفاده گردید. برای مقایسه بین این چهار تبدیل از معیارهای نرخ خطای بیت ، هم بستگی بین نهان نگاره اصلی و استخراج شده و شفافیت استفاده شد. برای دو معیار اول روابطی معرفی شده و برای معیار شفافیت از روش PESQ کمک گرفته شد. جهت تست پایداری در برابر حملات، سیگنال در برابر دو حمله فشرده سازی و افزودن نویز قرار گرفت.

در فشرده سازی، سیگنال به دو روش MP3 و GSM 6.10 فشرده و سپس به فرمت اصلی خود یعنی WAV باز گردانده شده و دو پارامتر نرخ خطای بیت و هم بستگی بین نهان نگاره اصلی و استخراج شده محاسبه گردید.

در افزودن نویز، به سیگنال نویز سفید افزوده شد تا میزان نسبت سیگنال به نویز در مقادیر ۱۰ dB و ۱۵ dB حاصل شود. و دو پارامتر نرخ خطای بیت و هم بستگی بین نهان نگاره اصلی و استخراج شده محاسبه گردید.

با توجه به نتایج به دست آمده، در روش جایگزینی بیت کم ارزش، از نظر نرخ خطای بیت و هم بستگی، تبدیل فوریه – بسل در رتبه دوم و پس از تبدیل موجک قرار گرفت. از نظر میزان شفافیت نیز این تبدیل پایین ترین مقدار را داشت. در تست پایداری در برابر حملات هیچ تبدیلی نهان نگاره را حفظ نکرد که این خود گواهی بر شکننده بودن روش بود.

در روش طیف گسترده، از نظر نرخ خطای بیت و هم بستگی، تبدیل فوریه – بسل بهترین نتایج را کسب کرده و کمترین میانگین خطا و بهترین میانگین هم بستگی را داشت. از نظر میزان شفافیت نیز

بهترین مقدار در روش PESQ را کسب کرد. اما در تست پایداری در برابر فشرده سازی و افزودن نویز تبدیل کسینوسی نتایج بهتری کسب کرد.

در روش مدولاسیون اندیس کوانتیزاسیون، از نظر نرخ خطای بیت و هم بستگی، تبدیل فوریه – بسل پس از تبدیل فوریه در جایگاه دوم قرار گرفت. البته لازم به ذکر است که مقادیر بسیار نزدیک به هم بود. در مقایسه میزان شفافیت تبدیل فوریه – بسل مقداری کمتر از سایر تبدیل‌ها داشت. در تست پایداری در برابر فشرده سازی و افزودن نویز، تبدیل فوریه – بسل بهترین نتیجه را داشت. به طوری که در برابر فشرده سازی MP3 و افزودن نویز با نسبت سیگنال به نویز dB ۱۵ کاملاً مقاوم بوده و نرخ خطای بیت و هم بستگی، بر خلاف سایر تبدیل‌ها هیچ تغییری نکرد. در برابر فشرده سازی GSM 6.10 و افزودن نویز با نسبت سیگنال به نویز dB ۱۰ این تبدیل بسیار بهتر از بقیه توانست نهان نگاره را حفظ کند. با توجه به این نتایج می‌توان گفت بهترین روش پنهان سازی از نظر پایداری، روش مدولاسیون اندیس کوانتیزاسیون در حوزه فوریه – بسل می‌باشد.

با توجه نتایج نشان داده شده به صورت کلی می‌توان گفت تبدیل فوریه بسل عملکرد خوبی در مقایسه با سایر تبدیل‌های متداول دارد.

۵-۲ پیشنهادات آینده

برای کارهای آینده پیشنهاداتی وجود دارد که به شرح زیر می‌باشد.

۱. می‌توان با تمرکز بر سیستم‌های پنهان سازی و تغییراتی بر روی آنها میزان پایداری و شفافیت سیستم‌ها را بالا برد. سپس حوزه های مختلف را در این سیستم‌ها بررسی کرد.

۲. در بررسی تحلیلی پارامترهایی از سیستم پنهان سازی چون نرخ خطای بیت، باید مدلی از توزیع آماری ضرایب فوریه – بسل سیگنال گفتار در دست داشت. در صورت یافتن چنین مدلی کارهای ارزشمندی می‌توان در این زمینه انجام داد.

۳. حمله‌های دیگری از قبیل فیلترینگ، افزودن پژواک، کوانیزاسیون مجدد، نمونه برداری مجدد، تغییر گام سیگنال، افزودن انواع نویزهای غیر سفید و فشرده سازی با نرخ بیت‌های مختلف وجود دارد. شبیه سازی و بررسی پایداری در برابر این حملات نیز مفید است.

۴. از تبدیل فوریه – بسل می‌توان در زمینه‌های دیگر غیر از گفتار نیز بهره برد. این ضرایب به عنوان ویژگی‌هایی برای سیستم‌های شناسایی الگو می‌توانند مفید واقع شوند.

۵. در الگوریتم پیاده سازی تبدیل فوریه – بسل دو ایراد وجود دارد. اول اینکه این الگوریتم فقط برای توابع بسل مرتبه صفر تعریف شده و برای مراتب بالاتر قابل تعمیم نیست. دوم اینکه در چند نمونه ابتدایی سیگنال هنگام بازسازی سیگنال مقداری خطا به چشم می‌خورد. در صورت رفع این دو مشکل بررسی کلی‌تری در تبدیل فوریه – بسل می‌توان انجام داد.

مراجع

[1] Cox J., Miller L., Bloom A., 1999, *Watermarking application and their properties*, NEC research institute.

[2] Chun Shein Lu, 2005, *Multimedia security: steganography and digital watermarking techniques for protection of intellectual property*, Idea Group Publishing.

[3] ITU-T Recommendation, Feb. 2001, *Perceptual Evaluation of Speech Quality (PESQ, An objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs,)* P.862.

[4] *Perceptual Evaluation of Speech Quality (PESQ, An objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs,)* ITU-T Recommendation P.862, Feb. 2001.

[5] Y. Lin W. Abdulla, 2011, " Objective quality measures for perceptual evaluation in digital audio watermarking", *IET Signal Processing*, ISSN 1751-9675

[6] Nedeljko Cvejic, Tapio Seppänen, 2004, " Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method", Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04)

[7] I. J. Cox, J. Killian, F. T. Leighton, and T. Shamoon, Dec. 1997, "Secure spread spectrum watermarking" IEEE Trans. Image Process. Vol. 6, no. 12, pp. 1673- 1678.

[۸] فتاح پور ف.، ۱۳۸۹، روشی جدید در واترمارکینگ صوت با استفاده از طیف گسترده تعمیم یافته و بالاترین ماسک آستانه شنیداری، دانشکده فنی و مهندسی، دانشگاه سمنان

[۹] دوست ر.، ۱۳۸۳، بهبود نهان نگاری صدا در مخابرات، دانشکده برق و الکترونیک، دانشگاه امیرکبیر

[10] B. Chen and G. W. Wornell, 2001, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1423-1443.

[۱۱] خادمی کلانتری ن.، ۱۳۸۸، طراحی و بررسی سیستم‌های نقش آب زنی مقاوم، دانشکده برق و الکترونیک، دانشگاه امیرکبیر

[۱۲] زارعیان م، ۱۳۹۰، بهبود روش‌های مخفی سازی اطلاعات در سیگنال میزبان گفتار باند باریک، دانشکده برق و الکترونیک، دانشگاه امیرکبیر

[10] Gruhl D. A. L, and W. Bender, May 1996, "Echo Hiding", pp. 295- 315, *Information Hiding: First International Workshop,* Vol. 1174, R. J. Anderson, Springer-Verlag, Lecture Notes in Computer Science, Cambridge, U.K.

[14] David J. Coumou, Gaurav Sharma, 2008 " Insertion, Deletion Codes with Feature-Based Embedding: A New Paradigm for Watermark Synchronization, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 3, NO. 2.

[15] Bender W., Gruhl D., Morimoto N., and A. Lu, 1996, "Techniques for data hiding", *IBM Syst*. J, vol. 35, no. 3/4, pp. 313-336.

[16] L. Girin & S. Marchand, Watermarking of speech signals using the sinusoidal model and frequency modulation of the partials, *Proc. Int. Conf. on Acoustics, Speech & Signal Proc*., Montréal, Canada, 2004.

[17] N. Chen, J. Zhu, 2008, "A robust zero-watermarking algorithm for audio". EURASIP Journal on Advances in Signal Processing (103).

[18] Schroeder J, 1993, Signal processing via fourier-Bessle series expansion, digital signal processing 3, pp. 112-124.

[19] Gopalan K., Chan C. S., 1983, Numerical Evaluation of Fourier-Bessel Series Expansion. ICASSP83 IEEE.

[20] Spoorthy S., Ramamurthy G., 2011, Gender Identification using Significant Intrinsic Mode Functions and Fourier-Bessel Expansion, Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN 2011).

[21] Gopalan K., Anderson T.R., 1999, Speaker Identification using Bessel function representation and a back-propagation neural network. IEEE Catalog Number: 95TH8081.

[22] Gopalan K., 1997, Speaker Identification using Features based on First Order Bessel function Expansion of Speech, 0-7803-3905-3/97/$10.00@1997 IEEE.

[23] Gopalan K., Timothy R. Anderson, MAY 1999, A Comparison of Speaker Identification Results Using Features Based on Cepstrum and Fourier–Bessel Expansion, IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 7, NO. 3.

[24] Kumar A., Prakash CH., 2011, Bessel Features for Estimating Number of speakers from Multispeaker Speech Signals, Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference IEEE**.**

[25] Gurgen F.S., Chen C.S., 1990, Speech enhancement by Fourier-Bessel coefficients of speech and noise, Communications, Speech and Vision, IEE Proceedings I, vol. 137, pp. 290-294.

[26] Gopalan K., 1999, Speech Modification by Selective Fourier-Bessel Series Expansion of Speech Signals, 0-7803-5582-2/99/$10.000 1999 IEEE.

[27] Gopalan K., 2001, Speech Coding using Fourier-Bessel Expansion of Speech Signals, IECON'O1: The 27th Annual Conference of the IEEE Industrial Electronics Society.

[28] Pachori R. B., Suryakanth, 2010, Detection Of Voice Onset Time Using FB Expansion and AM-FM Mode, 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010)

[29] Prakash Ch., Dhananjaya N., Suryakanth, 2011, Bessel Features for Detection of Voice Onset Time using AM-FM Signal, International Conference on Systems, Signals and Image Processing (IWSSIP), 2011 18th IEEE.

[30] Zareian M, Sayadiyan A, Sheikhzadeh H, 2011, " A novel quantization-based data hiding approach"2011 3rd IEEE International Conference on Signal Processing Systems (ICSPS 2011)