

به نام خدا

فشرده سازی صوت

1)

تبدیل های DCT و MDCT هر دو تبدیل های مرتبط با تبدیل فوریه هستند، اما به گونه ای طراحی شده اند که در زمینه های خاصی کاربرد بیشتری دارند. در اینجا به توضیح این تبدیل ها و تفاوت هایشان با تبدیل فوریه خواهیم پرداخت.

۱. تبدیل گسسته کسینوسی (DCT) یک تبدیل ریاضی است که به طور عمده برای تحلیل سیگنال های پیوسته و گسسته استفاده میشود. این تبدیل، مانند تبدیل فوریه، سیگنال را به مجموعه ای از ضرایب فرکانسی تبدیل میکند، اما تنها از تابع کسینوس استفاده میکند، نه ترکیب های سینوس و کسینوس.

انواع: DCT

• DCT نوع ۱: معمولاً در برخی از پردازش های سیگنال های دیجیتال استفاده میشود.

• DCT نوع ۲: این نوع در پردازش های تصویر و فشرده سازی تصویر، مانند JPEG، و نیز در فشرده سازی و پردازش صدا به کار میرود.

ویژگیهای: DCT

• DCT توانایی تبدیل سیگنال های غیر ایستا به فرکانس های مختلف را دارد.

• تبدیل DCT بیشتر در مواردی که سیگنال ها به طور طبیعی در حوزه فرکانسی کم فرکانس غالب هستند، مانند تصاویر و سیگنال های صوتی، به کار میرود.

۲. تبدیل گسسته کسینوسی اصلاح شده (MDCT) نوعی خاص از DCT است که در بسیاری از فرمت های فشرده سازی صدا مثل MP3 و AAC استفاده میشود . MDCT در واقع نسخه ای از DCT است که برای کاهش آرتیفکت های ناپایدار در هنگام فشرده سازی داده ها طراحی شده است.

ویژگیهای MDCT

- MDCT تغییراتی در نحوه تقسیم بندی سیگنال به بخش های کوچکتر دارد که به کاهش افت کیفیت کمک میکند.
- این تبدیل به طور خاص در پردازش سیگنال های صوتی کاربرد دارد، جایی که به سیگنال ها در طول زمان توجه میشود.
- MDCT معمولاً در قال بی که سیگنال ها به "بخش های همپوشانی" تقسیم میشوند، عمل میکند. این ویژگی باعث میشود تا در فرآیند فشرده سازی صدا از آرتیفکت ها یی مانند "پچکاری" جلوگیری شود.

۳. تفاوتها و شباهتها با تبدیل فوریه

شباهتها:

- تحلیل فرکانسی: هم DCT و هم MDCT و هم تبدیل فوریه سیگنال ها را به حوزه فرکانس تبدیل میکنند و امکان تجزیه و تحلیل محتویات فرکانسی یک سیگنال را فراهم می آورند.
- الگوریتم های تبدیل: هر سه تبدیل در حالت کلی برای تجزیه سیگنال های پیچیده به ترکیب های مختلف فرکانسی استفاده میشوند.

تفاوتها:

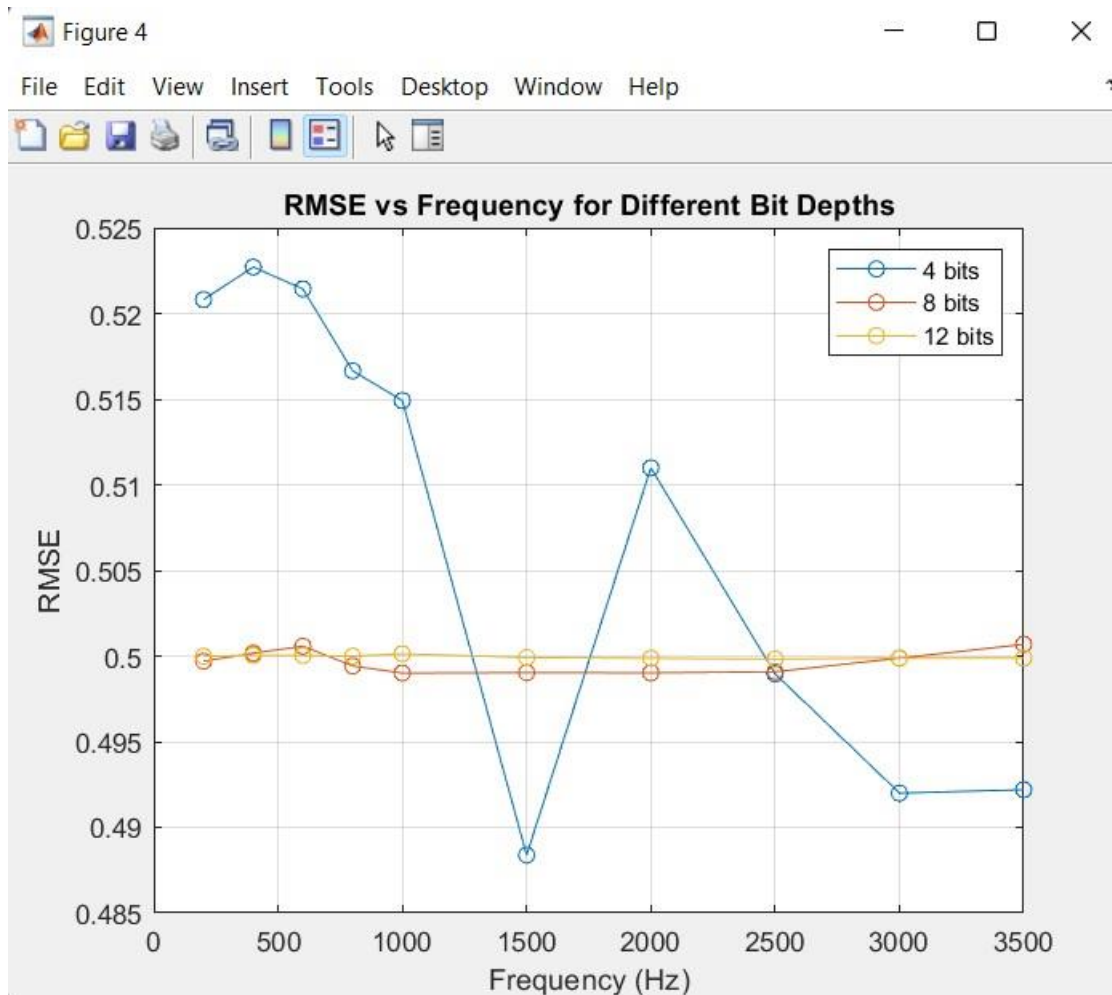
- استفاده از سینوس و کسینوس: در حالی که تبدیل فوریه از ترکیب سینوس و کسینوس استفاده میکند، DCT تنها از تابع کسینوس بهره میبرد. در MDCT نیز از تغییرات خاصی در بخش های کسینوسی استفاده میشود که به نوعی به کاهش آرتیفکتهای فشرده سازی کمک میکند.
- دوره ای بودن: تبدیل فوریه به طور طبیعی به سیگنال های دوره ای نگاه میکند، اما DCT برای سیگنال های غیردوره ای بهتر عمل میکند.
- محدوده کاربرد DCT: بیشتر در پردازش تصویر و صدا استفاده میشود، در حالی که MDCT عمدتاً در فشرده سازی صدا کاربرد دارد. تبدیل فوریه نیز به طور عمومی در تحلیل فرکانسی و پردازش سیگنال ها استفاده میشود.

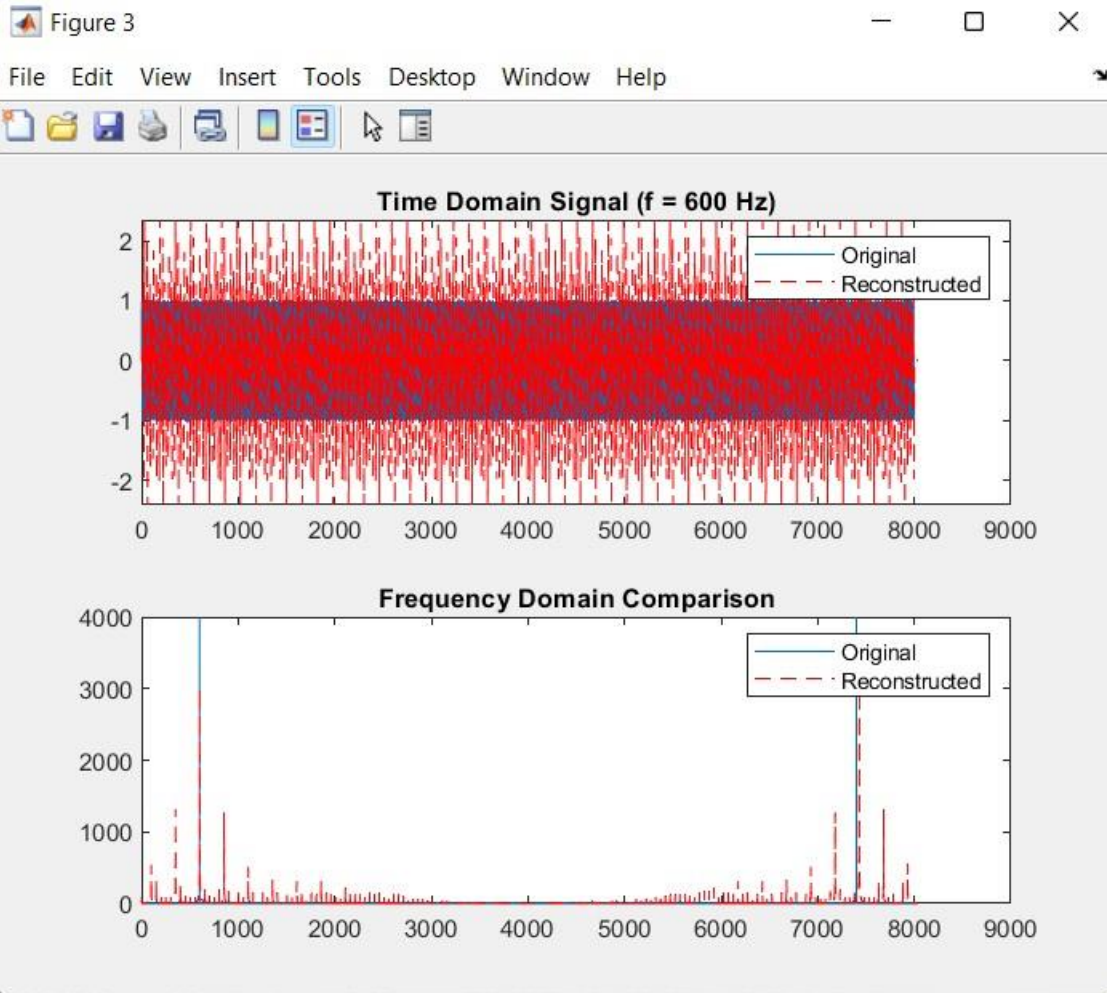
خلاصه تفاوتها:

1. DCT برای تبدیل سیگنال های گسسته به حوزه فرکانسی طراحی شده و فقط از کسینوس استفاده میکند.
 2. MDCT نسخهای اصلاح شده از DCT است که در زمینه فشرده سازی صدا کاربرد دارد و باعث کاهش آرتیفکت های فشرده سازی میشود.
 3. تبدیل فوریه شامل ترکیب سینوس و کسینوس است و برای تحلیل سیگنال های دوره ای و گسسته به کار میرود.
- این تفاوت ها باعث میشود که DCT و MDCT در زمینه های خاصی مثل پردازش تصویر و فشرده سازی صدا بیشتر کاربرد داشته باشند، در حالی که تبدیل فوریه به طور عمومی تر در انواع مختلف پردازش سیگنال ها و تحلیل ها استفاده میشود.

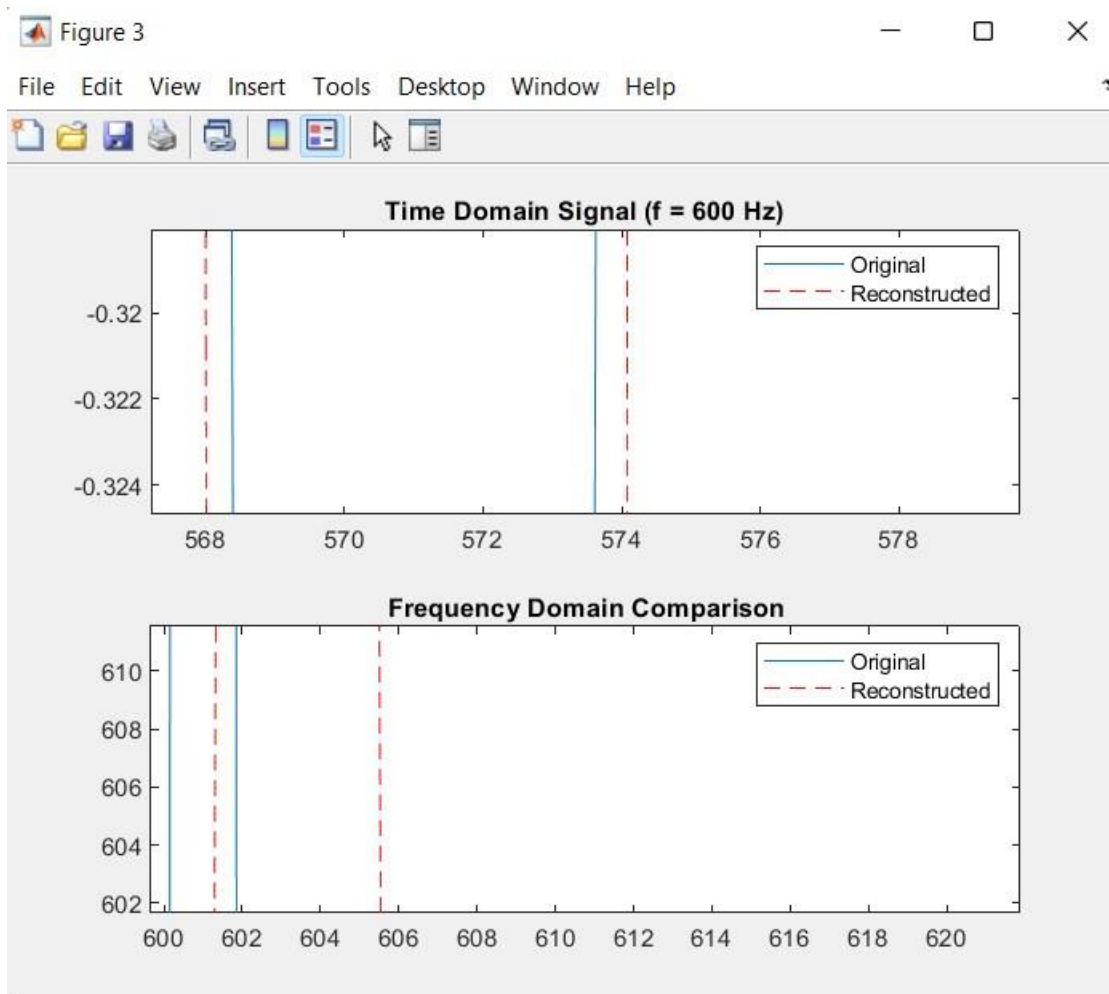
)2

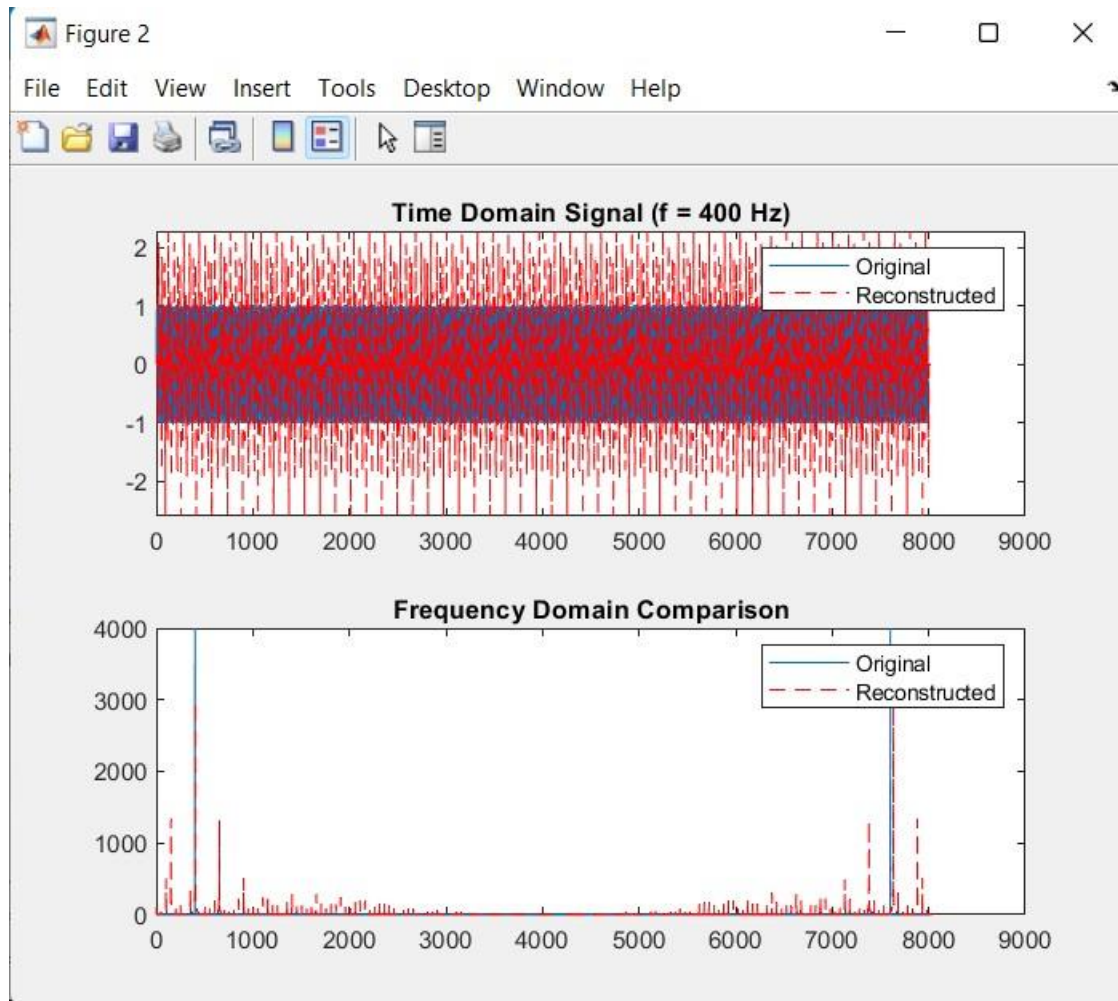
خروجی:

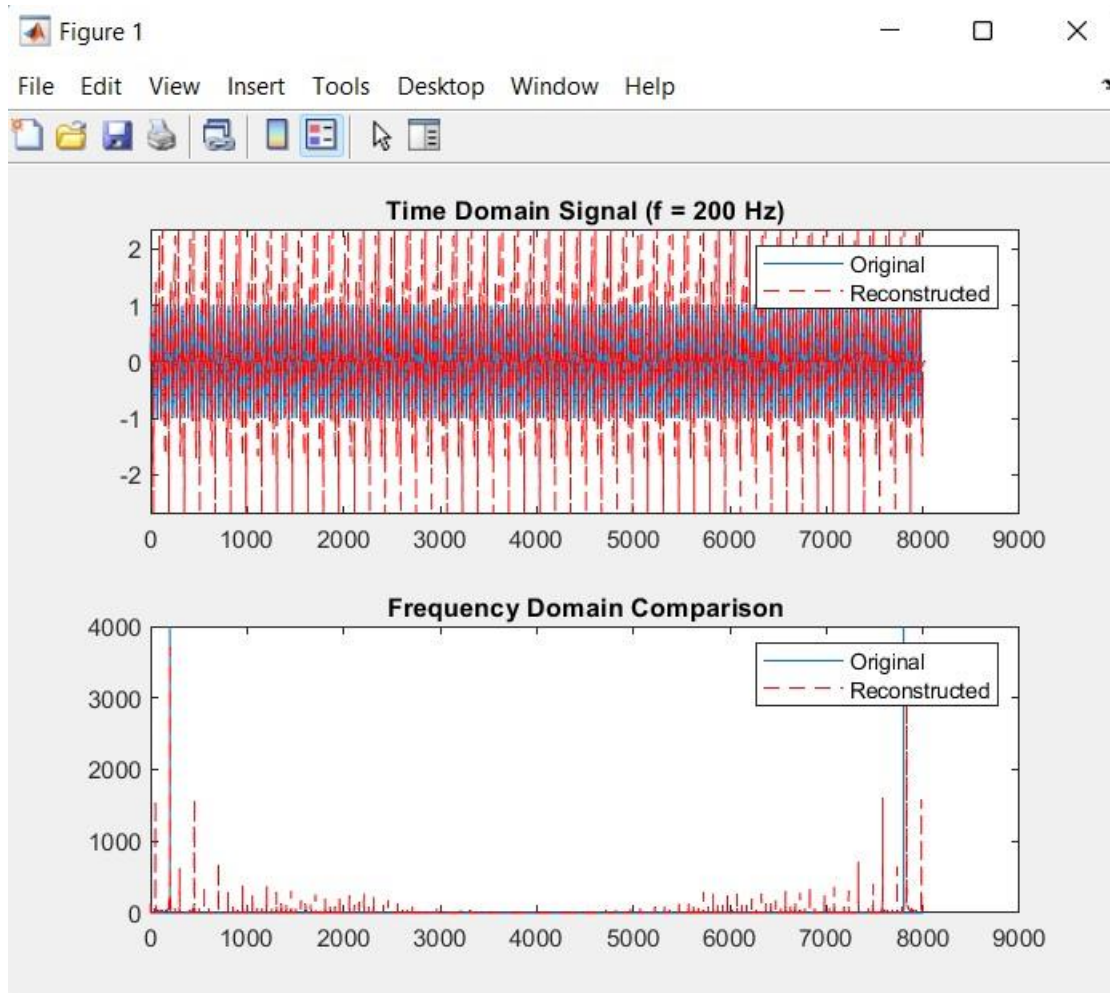




نمونه زوم شده:







در قسمت اول فرکانس نمونهبرداری ۸ کیلوهرتز تنظیم شده است.

اندازه بلوکها برای تبدیل MDCT مشخص شده است.

3 مقدار مختلف بیت کوانتیزاسیون برای تست تأثیر فشردهسازی در نظر گرفته شده است.

لیستی از فرکانسهای مختلف برای آزمایش تعریف شده است. یک ماتریس برای ذخیره مقادیر RMSE در طول آزمایش مقاردهی اولیه شده است.

در قسمت بعدی یک تابع تعریف شده است. این تابع پنجره، یک پنجره سینوسی است که به منظور کاهش مصنوعات طیفی و بهبود دقت بازسازی استفاده میشود. این پنجره باعث میشود که مقدار سیگنال در ابتدا و انتها کاهش یابد، که منجر به کاهش spectral leakage میشود.

سپس حلقه فرکانس های مختلف را بررسی میکند و یک سیگنال سینوسی خالص (pure tone) با فرکانس f تولید میشود.

سپس MDCT برای بلوکهای همپوشان اعمال میشود. مقدار MDCT کوانتیزه میشود تا فشرده سازی با مقدار بیت مشخص (۴، ۸ یا ۱۲ بیت) انجام شود.

سپس، سیگنال از ضرایب کوانتیزه شده بازسازی میشود.

سپس خطای RMSE بین سیگنال اصلی و بازسازی شده محاسبه میشود تا میزان دقت بازسازی مشخص شود.

در مرحله بعد سیگنال اصلی و بازسازی شده در حوزه زمان نمایش داده میشود. دامنه فوریه (FFT) برای هر دو سیگنال رسم میشود تا بررسی شود که چقدر سیگنال بازسازی شده به سیگنال اصلی نزدیک است.

در قسمت بعد نمودار RMSE بر حسب فرکانس tone برای بیت های مختلف رسم میشود.

این نمودار نشان میدهد که تعداد بیت‌های کوانتیزاسیون چگونه روی خطای بازسازی تأثیر می‌گذارد.

`function X = mdct_transform1(x, N, window)`

در این تابع سیگنال ورودی پدینگ میشود تا از خطای مرزی جلوگیری شود. سپس، سیگنال به بلوکهای همپوشان تقسیم میشود. MDCT نوع ۴ (DCT-IV) برای تبدیل فرکانسی اعمال میشود.

`function x_reconstructed = imdct_transform1(X, N, window)`

در این تابع ضرایب MDCT با استفاده از IDCT-IV به حوزه زمان بازمیگردند. از اصل همپوشانی و جمع‌آوری (OLA) برای ترکیب بلوکها و بازسازی سیگنال استفاده میشود.

`function X_quant2 = quantize1(X, b)`

در این تابع مقدار بیشینه MDCT نرمالسازی میشود. ضرایب در سطوح b^2 تقسیم‌بندی شده و گرد میشوند. که باعث کاهش دقت داده‌ها و در نتیجه فشردگی میشود.

تغییر تعداد بیت‌های کوانتیزاسیون تأثیرات مختلفی مانند تأثیر بر RMSE، تأثیر بر کیفیت شنیداری و بر اندازه داده و فشردگی دارد. بیت‌های کمتر باعث افزایش خطای کوانتیزاسیون و نویز در سیگنال بازسازیشده میشوند.

بیتهای بیشتر کیفیت صوتی بهتری فراهم میکنند، اما داده بیشتری مصرف میشود.
برای کاربردهای عملی، باید تعادلی بین فشردهسازی و کیفیت حفظ شود.

3)

در پردازش سیگنال، تابع پنجره برای کاهش اثرات نامطلوب تبدیل فوریه و کاهش اعوجاج در سیگنالهای زمان-محدود استفاده میشود. سیگنالهایی که در حوزه زمان محدود هستند، در حوزه فرکانس انتشار وسیعی دارند و میتوانند باعث ایجاد لکهای فرکانسی (spectral leakage) شوند.

هنگامی که یک سیگنال را به بلوکهای کوچک تقسیم میکنیم مثل کاری که در MDCT انجام میدهیم، مرزهای این بلوکها میتوانند باعث ناپیوستگی شوند. اعمال تابع پنجره باعث میشود که تغییرات در مرز بلوکها نرمتر شود، و به این ترتیب اعوجاج و خطا کاهش یابد. بدون اعمال پنجره، سیگنالهای بریده شده در مرزهای بلوکها ناپیوستگی دارند که باعث ایجاد مؤلفههای فرکانسی ناخواسته میشود. تابع پنجره مقدار سیگنال را در مرزهای بلوک به صفر نزدیک میکند و این ناپیوستگی را کاهش میدهد. اگر یک سیگنال کاملاً متناوب نباشد، تبدیل فوریه آن ممکن است شامل مؤلفههای فرکانسی ناخواسته باشد. پنجرهها، با اعمال یک ضریب ملایم، از بروز این نشت طیفی جلوگیری میکنند.

در سیستمهای کدگذاری مانند MDCT، اعمال پنجرههای هموار باعث بهبود کیفیت بازسازی و کاهش خطای کوانتیزاسیون میشود. با کاهش ناهمواری در حوزه فرکانس، پنجرهها باعث کاهش تداخل بین بلوکها و افزایش دقت بازسازی سیگنال میشوند.

پنجره‌ی سینوسی دارای خاصیت **Perfect Reconstruction** است، به این معنی که وقتی از MDCT و IMDCT برای فشردسازی و بازسازی سیگنال استفاده کنیم، سیگنال اصلی بدون خطا (در شرایط ایده‌آل و بدون کوانتیزاسیون) بازسازی میشود. این پنجره، نسبت به پنجرههای مستطیلی، نشت طیفی کمتری دارد و در نتیجه دقت بیشتری در حوزه‌ی فرکانس ارائه میدهد.

این پنجره در کدکهای مبتنی بر MDCT به طور خاص برای همپوشانی ۵۰٪ طراحی شده است. این یعنی وقتی بلوکهای متوالی در پردازش MDCT روی هم قرار میگیرند، خطای ناشی از بریدگی سیگنال حذف میشود.

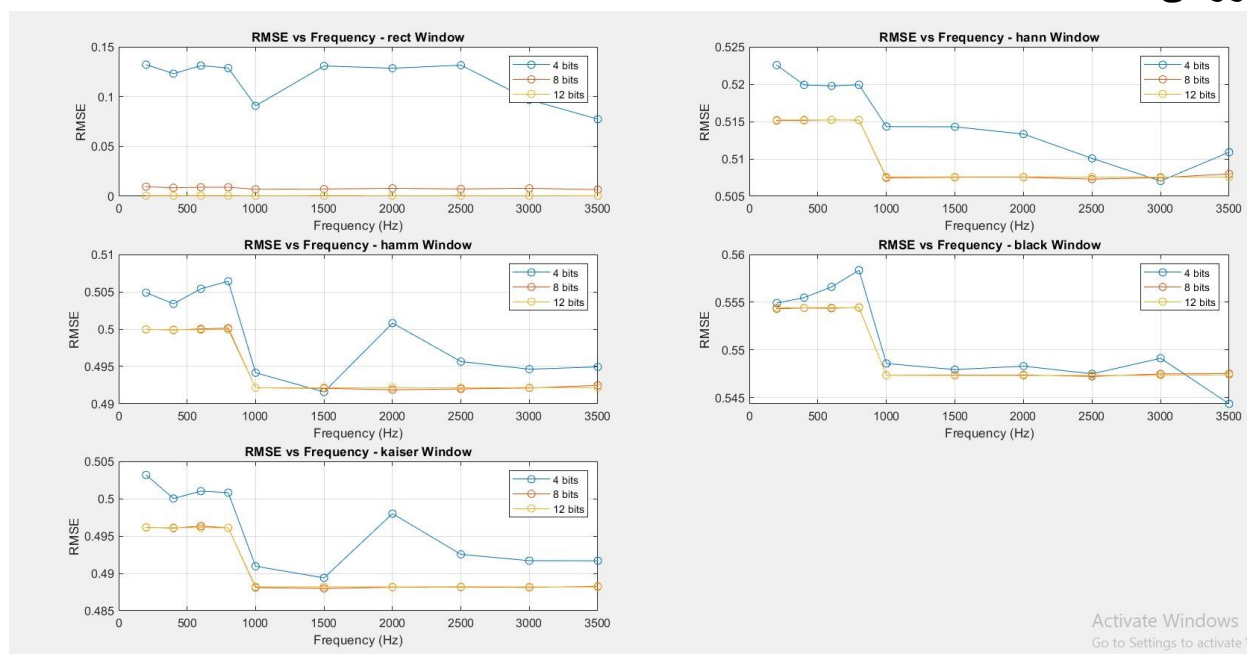
در کدکهای صوتی، استفاده از این پنجره باعث میشود که خطای اعوجاج و تداخل بین بلوکها حداقل شود، که برای کیفیت صدای بهتر و نرخ فشردسازی بالاتر بسیار مهم است. این پنجره محاسبات پیچیده‌ای نیاز ندارد و تنها از یک تابع سینوسی ساده تشکیل شده که در مقایسه با پنجرههای پیچیده‌تر مثل Blackman یا Kaiser سریعتر اجرا میشود.

کد این قسمت ۵ پنجره مخ تلف را بررسی میکند و تفاوت هایی با کد سوال قبل دارد. مثلاً کد سوال اول برای سه فرکانس اول، سیگنال اصلی و بازسازیشده را نمایش میدهد اما کد سوال دوم این بخش را حذف کرده است و فقط RMSE را محاسبه میکند.

کد سوال اول یک نمودار RMSE کلی برحسب فرکانس رسم میکند کد سوال دوم چندین نمودار RMSE جداگانه برای هر نوع پنجره رسم میکند.

در کد دوم ، هنگام اعمال IMDCT مقدار همپوشانی بهطور صریحتر مشخص شده است در کد اول این بخش کمتر بهینه بود و ممکن است برخی اعوجاجها در بازسازی ایجاد شوند.

خروجی:



4)

فرمت WAV (Waveform Audio File Format) یک فرمت صوتی بدون فشرد هسازي است که توسط مایکروسافت و IBM توسعه داده شده است. این فرمت معمولاً داد ههای صوتی را بهصورت PCM (Pulse Code Modulation) ذخیره میکند، که کیفیتی بدون افت دارد.

ویژگیهای اصلی WAV:

- کیفیت بالا، اما حجم فایل زیاد
- پشتیبانی از نرخ نمونهبرداری (Sample Rate) و عمق بیت (Bit Depth) متنوع
- سازگار با ویندوز و بسیاری از نرمافزارهای صوتی
- مناسب برای ویرایش و پردازش صوتی قبل از فشرد هسازي

توضیح کد:

در قسمت اول ، نرخ نمونه برداری، اندازه بلوک MDCT ، تعداد بیتهای کوانتیزاسیون و لیست فرکانسهای مورد آزمایش مشخص میشود. همچنین، چندین نوع تابع پنجره تعریف میشوند.

در قسمت بعد برای هر نوع تابع پنجره، مقدار مناسب آن انتخاب میشود. تابع پنجره روی بلوکهای $2 \times \text{block_size}$ اعمال میشود.

س پس سیگنال سینوسی تولید میشود.

MDCT روی سیگنال اعمال میشود.

ضرایب کوانتیزه شده و سپس بازسازی (IMDCT) میشوند.

خطای RMSE بین سیگنال اصلی و بازسازی شده محاسبه و ذخیره می شود

سپس نتایج RMSE برای هر نوع تابع پنجره و تعداد بیت‌های مختلف رسم میشود .

`function X = mdct_transform3(x, N, window)`

MDCT روی سیگنال با استفاده از پنجره انتخاب شده اعمال میشود.

`function x_reconstructed = imdct_transform3(X, N, window)`

تبدیل معکوس MDCT روی ضرایب کوانتیزه شده اعمال شده و سیگنال بازسازی میشود.

`function X_quant3 = quantize(X, b)`

کوانتیزاسیون ضرایب MDCT برای کاهش حجم داده انجام میشود.

مقایسه دو نسخه از کد:

نسخه دوم نسخه اول ویژگی

از چندین نوع تابع پنجره (مستطیلی، هانینگ، فقط از تابع سینوسی استفاده تابع پنجره
هامینگ، بلکمن، کایزر) استفاده میکند میکند RMSE برای هر تابع پنجره در
نمودارهای فقط یک نمودار RMSE تحلیل

جداگانه بررسی میشود برحسب فرکانس رسم میشود RMSE برای انواع مختلف
توابع پنجره آزمایش شده فقط روی یک تابع پنجره و عمومیت
و مقایسه میشود یک سیگنال اجرا میشود الگوریتم

خروجی:

Channel 1:

	none	hann
4	0.091215	0.25333
8	0.010841	0.24754
12	0.0042607	0.24751