

پروژه

۲ فشرده‌سازی

صوت

در این پروژه فشرده‌سازی صوت را به عنوان کاربرد مهم از تبدیل کسینوس گسسته اصلاح شده^۱ (MDCT) بررسی می‌کنیم.

۱-۲ مقدمه

همانطور که می‌دانید فایل‌های صوت را می‌توان با فرمت‌های متنوع همچون^۲ MP3،^۳ AAC و^۴ WMA فشرده ساخت. وجه اشتراک این فرمت‌ها در آن است که بخش اصل فشرده‌سازی بر مبنای گونه‌ایاز تبدیل^۵ DCT انجام می‌شود که MDCT نام دارد. Audio codec سیستم است که از آن برای کدگذاری/کدبرداری فایل‌های صوت استفاده می‌شود. در این پروژه می‌خواهیم ی طرح پایه‌ای برای Audio codec را بر مبنای MDCT پیاده‌سازی کنیم و به بررسی کیفیت صوت گذشته بپردازیم.

شما در این درس با تحلیل حوزه فرکانس سی نال‌ها آشنا شدید و تبدیل‌های مختلف را فراگرفتید. یاز تبدیل‌های حوزه زمان به حوزه فرکانس، تبدیل کسینوس گسسته (DCT) است که پایه اصل فشرده‌سازی سیاری از دادگان محسوب می‌شود. MDCT با تغییری در DCT حاصل شده است تا برای فشرده‌سازی صوت مناسب باشد.

^۱Modified Discrete Cosine Transform (MDCT)

^۲MPEG-1 Audio Layer 3 (MP3)

^۳Advanced Audio Coding (AAC)

^۴Windows Media Audio (WMA)

^۵Discrete Cosine Transform (DCT)

پروژه

۲-۲ توضیحات

سی نال صوت به بلوک هایی به طول n تقسیم می شود و در ادامه تبدیل MDCT اعمال می گردد. درال وریتم های فشرده سازی صوت، MDCT را بر روی بردارهایی از داده اعمال می کنند که همپوشان دارند. دلیل این کار پرهیز از پدیدار آمدن مصنوعات است که منجر به خطاهایی متناوب می شوند. گوش انسان بیشتر از چشم او به خطاهایی متناوب حساسیت دارد و به شنیدن آن تمرکز خواهد کرد. برداری به طول $2n$ که در بردارنده دو بلوک پیایی است در ورودی تبدیل MDCT در نظر گرفته می شود. حاصل تبدیل برداری به طول n خواهد بود و از خاصیت هم پوشان می توان برای بازسازی بخش مشترک دو بلوک پیایی استفاده کرد. اولین و آخرین بلوک قابل بازیابی نخواهند بود، بنابراین لازم است که به شکل مناسب از Padding استفاده شود.

پس از اجرای تبدیل MDCT به فشرده سازی سی نال می پردازیم. ایده اصل آن است که می توانیم برخاز مؤلفه های فرکانس را با دقت بیشتری نگه داریم و تأکید کمتری برای برخی دی داشته باشیم. برای فشرده سازی سی نال صوت می توان حاصل تبدیل MDCT را به b بیت کوانتیزه کرد که منجر به فشرده سازی با اتلاف و کاهش فضای لازم برای ذخیره سازی می شود. برای بازسازی سی نال صوت کارهایی در جهت معکوس انجام می شوند.

برای بهبود کیفیت سی نال بازسازی شده می توان از window function استفاده کرد که سی نال ورودی x را تغییر مقیاس می دهد به نحوی که در ابتدا و انتها پنجره مقادیری نزدی به صفر داشته باشد. سی نال x دقیقاً متناوب نیست و این روی رد تا حدی اثر آن را جبران می سازد. می توان از تابع h برای پنجره ای به طول $2n$ باشد که مؤلفه i آن در رابطه ۲-۱ تعریف شده است. از آن قبل از اعمال تبدیل MDCT و بعد از تبدیل وارون استفاده می شود.

$$h_i = \sqrt{2} \sin \left(\frac{\left(i - \frac{1}{2}\right) \pi}{2n} \right) \quad (2-1)$$

۲-۳ خواسته های پروژه

۱. در مورد تبدیل های DCT و MDCT تحقیق کنید و تفاوت ها و شباهت های آنها را با تبدیل فوریه بررس کنید.

۲. بر اساس توضیحات گفته شده ی Audio codec ساده را پیاده سازی کنید. می توانید از بلوک های پی به طول 32 یا 64 استفاده کنید. از ضرایب کوانتیزه شده سی نال صوت را بازسازی کنید. این

ال وریتم را بر روی چند pure tone اعمال کنید و کیفیت صوت بازسازی شده را بصورت کیف و کم بررسی کنید. معیار RMSE را در نظر داشته باشید. در مثالهایی سی نال های اصل و بازسازیشده را در حوزه زمان و فرکانس مقایسه کنید. نمودار RMSE بر حسب فرکانس tone را رسم کنید. تأثیر تغییر تعداد بیت های کوانتیزاسیون را بررسی کنید.

۳. در مورد window function ها مطالعه کنید و بررسی کنید که چرا اعمال این توابع م تواند باعث کاهش خطا شود. تأثیر استفاده از تابع معرف شده در رابطه ۱-۲ را بررسی کنید. در این حالت هم نمودار RMSE بر حسب فرکانس tone را رسم کنید. چه تغییری کرده است؟

۴. ی فایل با فرمت WAV را در نظر ب پرید. (به اختصار در مورد این فرمت تحقیق کنید) اگر صوت انتخاب شده بصورت stereo است، هر ی از کانالها را جداگانه پردازش کنید. ال وریتم فشرده سازی را اجرا سازید و کیفیت صوت بازسازی شده را به صورت کیف و کم بررسی کنید. در جدول معیار RMSE را برای مقادیر متفاوت از تعداد بیت های کوانتیزاسیون و با فرض حضور window function مقایسه کنید. چه نتیجه ای م گیرید؟

پروژه