**تمرین 3 (**Mel-Frequency Cepstral Coefficients (MFCC)**)**

Mel-Frequency Cepstral Coefficients (MFCC) یک تکنیک استخراج ویژگی است که به طور گسترده در پردازش سیگنال گفتار و صوتی استفاده می‌شود. هدف MFCC‌ها نمایش ویژگی‌های طیفی سیگنال‌های گفتاری و صوتی به روشی فشرده و از نظر ادراکی معنادار است. در اینجا یک تفکیک از MFCC‌ها آمده است:

**1. مقیاس مل:**

مقیاس مل یک مقیاس ادراکی است که پاسخ دستگاه شنوایی انسان به فرکانس‌های مختلف را تقریب می‌زند. این بر اساس مشاهده است که انسان‌ها تفاوت بین فرکانس‌های پایین تر را با دقت بیشتری نسبت به فرکانس‌های بالاتر درک می‌کنند. مقیاس Mel یک نگاشت غیر خطی فرکانس را به مقیاسی ارائه می‌دهد که با ادراک انسان همسوتر باشد.

**2. محاسبه MFCC:**

فرآیند محاسبه MFCC شامل چندین مرحله است:

**آ. پیش پردازش:** سیگنال صوتی به فریم‌های کوتاه، معمولاً 20 تا 40 میلی ثانیه با یک همپوشانی کوچک بین فریم‌ها تقسیم می‌شود.

**ب. Windowing:** هر فریم در یک تابع پنجره مانند پنجره Hamming ضرب می‌شود تا نشت طیفی کاهش یابد.

**ج. تبدیل فوریه:** تبدیل فوریه، معمولاً تبدیل فوریه گسسته (DFT)، برای به دست آوردن طیف قدر بر روی هر قاب پنجره ای اعمال می‌شود.

**د. فیلتر مِل:** طیف قدر از میان مجموعه ای از فیلترهای مثلثی با فاصله یکنواخت در مقیاس Mel عبور می‌کند. هر فیلتر نشان دهنده یک باند فرکانسی خاص است.

**ه. لگاریتم:** لگاریتم خروجی فیلتر برای تبدیل مقادیر بزرگی به مقیاس لگاریتمی‌گرفته می‌شود که با درک انسان از بلندی صدا همسو می‌شود.

**و. تبدیل کسینوس گسسته (DCT):** انرژی‌های log-filter bank با استفاده از DCT تبدیل می‌شوند. ضرایب به دست آمده به عنوان ضرایب فرکانس Mel-Cepstral (MFCCs) شناخته می‌شوند.

**3. ویژگی‌های MFCC:**

MFCCها پوشش طیفی سیگنال صوتی را می‌گیرند و نمایشی فشرده از ویژگی‌های طیفی ارائه می‌دهند. ضرایب مرتبه پایین نشان دهنده شکل کلی طیف هستند، در حالی که ضرایب مرتبه بالاتر تغییرات طیفی با جزئیات بیشتری را ثبت می‌کنند.

**4. کاربردها:**

MFCCها ثابت کرده‌اند که در برنامه‌های مختلف پردازش گفتار و صدا، از جمله تشخیص گفتار، شناسایی گوینده، طبقه بندی ژانر موسیقی، تشخیص احساسات، سنتز گفتار و غیره موثر هستند. آنها به عنوان ویژگی‌های آموزنده ای عمل می‌کنند که ویژگی‌های مهم سیگنال صوتی را ثبت می‌کنند.

MFCCها به دلیل توانایی آن‌ها در گرفتن اطلاعات طیفی مرتبط به روش ادراکی معنی دار، به یک نمایش ویژگی استاندارد در بسیاری از سیستم‌های پردازش گفتار و صدا تبدیل شده‌اند.