



هدف تمرین:

- آشنایی با استخراج ویژگی گفتار
- آشنایی با تحلیل LPC
- آشنایی با استخراج ضرایب MFCC و طراحی فیلتربانک
- ویژگی‌های جدید مبتنی بر MFCC

مراحل:

۱. ابتدا عبارت “ظاهرش کُنه ضمیرش را لو داد” را با استفاده از wavesurfer با فرکانس نمونه‌برداری ۱۶ کیلوهرتز، بصورت ۱۶ بیتی و مونو در فرمت wav، با صدای خودتان ضبط کنید.
۲. برای دو فرم واکدار با دو واژه مختلف در فایل ضبط شده، مراحل زیر را انجام دهید. می‌خواهیم با استفاده از سه الگوریتم مختلف، ضرایب LPC را محاسبه کنیم.
 - a. ابتدا ضرایب LPC را با الگوریتم معکوس ماتریس اتو کوریلیشن در روش اتو کوریلیشن (ضرایب یول-واکر) بدست آورید. (راهنمایی: می‌توانید از دستور toeplitz کتابخانه scipy در پایتون، یا دستورات معادل در Matlab استفاده کنید).
 - b. از الگوریتم لوینسون برای محاسبه ضرایب LPC استفاده کنید. (می‌توانید از دستور levinson در کتابخانه spectrum در پایتون، یا دستورات معادل در Matlab کمک بگیرید).
 - c. از دستورهای built-in برای بدست آورد ضرایب LPC استفاده کنید (می‌توانید از دستور lpc در کتابخانه spectrum در پایتون، یا دستور معادل در Matlab استفاده کنید).
۳. مقدار G (گین) در رابطه $H(z) = \frac{G}{A(z)}$ را به کمک lpc بدست آورید.
۴. سیگنال باقی مانده یا تحریک را با استفاده از ضرایب LPC بدست آورید.
۵. پاسخ فرکانسی فیلتر مجرای گفتار را ترسیم کنید. (راهنمایی: می‌توانید از دستور freqz از کتابخانه scipy در پایتون، یا دستورات معادل در Matlab استفاده کنید).
۶. پوش طیف را با استفاده از ضرایب LPC ترسیم کنید و با پاسخ فرکانسی قسمت قبل مقایسه کنید. همچنین بر روی پوش طیف، محل رخداد فرم‌نت‌ها را نیز مشخص کنید.
۷. همراه این تمرین، یک فایل گفتاری تلفنی برای شما ارسال شده است. آنرا باز کرده و مراحل پیش پردازش را با همان مقادیر گفته شده در تمرین دوم بر روی این فایل انجام دهید.

۸. برای یک فریم واگذار، طیف فرکانسی آنرا ترسیم کرده و تفاوت مهم آن با طیف فرکانسی فریم‌های سیگنال گفتار قبلی بیان کنید.

۹. می‌خواهیم یک الگوریتم VAD ساده پیاده‌سازی کنیم. مراحل زیر را انجام دهید.

- a. لگاریتم انرژی کوتاه مدت هر فریم را محاسبه کنید و آنرا ترسیم کنید.
- b. یک آستانه ثابت به گونه‌ای انتخاب کنید که با اعمال آن بر منحنی لگاریتم انرژی، بیشتر فریم‌های گفتاری از غیرگفتاری متمایز شود.
- c. از یک فیلتر میانه با طول ۵ فریم برای هموارسازی فریم‌های گفتاری آشکار شده از مرحله قبل استفاده کنید. (خروجی این مرحله باید شماره فریم‌های گفتاری باشد)
- d. نتیجه الگوریتم خود را با الگوریتم VAD کتابخانه librosa مقایسه کنید.
- e. طیف فرکانسی هر فریم را محاسبه کنید.
- f. ۲۰ فیلتر مثلی با دامنه یکسان، توزیع شده به صورت یکتواخت با همپوشانی ۵۰ درصد از هر طرف بر روی محور mel طراحی کنید. فرکانس شروع فیلتر مثلی اول ۱۵۰ هرتز و فرکانس پایان فیلتر مثلی آخر ۳۸۰۰ هرتز است. مراحل را به صورت کامل توضیح دهید. (دقت شود که طراحی بانک فیلترها به صورت کامل توسط خودتان انجام شود و از کتابخانه‌های آماده استفاده نکنید.) (راهنمایی: از فرمول‌های زیر کمک بگیرید)

$$f_{mel}(f_{hz}) = 2595 \times \log_{10} \left(1 + \frac{f_{hz}}{700} \right)$$

$$f_{hz}(f_{mel}) = 700 \times \left(10^{\frac{f_{mel}}{2595}} - 1 \right)$$

- g. لگاریتم خروجی بانک فیلتر را محاسبه کنید (حاصل باید یک بردار ۲۰ بعدی باشد).
- h. از نتیجه حاصل، DCT بگیرید و تنها ۱۳ ضریب اول را برای هر فریم استخراج کنید.
- i. به جای ضریب صفرم کپسترا، لگاریتم انرژی کوتاه مدت فریم را جایگزین کنید.
- j. بر روی فریم‌های گفتاری نتیجه شده از VAD، نرمالیزه‌سازی CMVN را انجام داده و بردارهای میانگین، انحراف استاندارد و MFCC نرمالیزه شده را بدست آورید.
- k. ضرایب مشتق مرتبه اول و دوم و ضرایب MFCC نرمالیزه شده مرحله قبل را با استفاده از رابطه بیان شده در اسلایدهای درس با مقدار $\tau = 2$ محاسبه کنید. (از کد آماده استفاده نکنید).

۱۰. مقالات مختلف را بررسی کنید و حداقل ۳ مقاله (پس از سال ۲۰۱۵) بیابید که نوآوری‌های اصلی آن بر روی بهبود ضرایب MFCC بوده است.

- a. مشخص کنید که این مقاله بر کدام حوزه پردازش گفتار متمرکز است.

b. در هر مقاله چه تغییری در محاسبه ضرایب MFCC اعمال شده است؟

- نتیجه نهایی را در قالب یک گزارش PDF و کدهای مربوط به تمرین، به ایمیل ui.speech.course.2021@gmail.com ارسال کنید. توجه کنید که در عنوان ایمیل حتما نام، نام خانوادگی و شماره تمرین (تمرین 3) را قرار دهید و در فایل گزارش نیز نام و نام خانوادگی و شماره دانشجوییتان را ذکر کنید. در غیر اینصورت تکلیف شما دریافت نخواهد شد. همچنین در نظر داشته باشید که کامل بودن توضیحات، مرتب بودن گزارش ارسالی و درستی توضیحات همه در امتیاز نهایی تاثیرگذار هستند.

موفق باشید