

به نام خدا

علی انصاری ۹۸۳۱۱۳۶

بخش اول: تشخیص صدا

۱. طیف توان به عنوان مجذور قدر مطلق تبدیل فوریه گسسته محاسبه می شود. تبدیل فوریه سریع یک روش ریاضی برای تبدیل تابع زمان به تابعی از فرکانس است. این تابع توسعه‌ای از تبدیل فوریه گسسته است که عبارات تکراری را در الگوریتم ریاضی حذف میکند که با این عملیات می توان از تعداد زیادی نمونه بدون کاهش سرعت تبدیل استفاده کرد زیرا تبدیل فوریه سریع محاسبات را با ضرب $n/\log_2(n)$ کاهش میدهد.

۲. ابتدا فایل صوتی را میخوانیم و سپس فرکانس نایکوئیست را محاسبه میکنیم. سپس طیف توان را که با فرمول توان دو قدر مطلق تبدیل فوریه گسسته است به دست می‌آوریم و آنرا با تقسیم کردن بر طول داده نرمال‌سازی میکنیم. سپس با استفاده از `fft_freq` فرکانس `power` را با فرمول $1/n * fs$ محاسبه میکنیم.

۳. با استفاده از تابع زیر بیشترین میزان فرکانس را محاسبه میکنیم.

```
def freq_for_maximum_power(filename):  
    frequency, power_spectrum = create_power_spectrum(filename)  
    max = frequency[np.argmax(power_spectrum)]  
    print(filename[7:-4] + ' ' + str(max))  
    return(max)
```

۴. طبق گفته دستورکار اگر ماکسیمم فرکانس صدا از ۱۸۰ کمتر بود جنسیت صدا مرد است و در غیر اینصورت زن است. نمودار صداها به همراه جنسیت آنها (عنوان نمودار) در پوشه `output` بخش اول قرار داده شده است.

همچنین با کوچک‌تر کردن بازه جستجو یا چک کردن همزمان ماکسیمم در هنگام ساختن فرکانس‌ها کد را بهینه‌تر کرد

```
C:\Users\MY_PC\Desktop\SS_final\PART1>python part1.py  
v0 211.0202202322809  
v1 110.43235723920863  
v10 218.0506359746872  
v11 128.9516661222912  
v2 50.047482262404  
v3 131.32155104194143  
v4 283.72423274871784  
v5 205.48636114308582  
v6 86.5617607694223  
v7 181.34491467004406  
v8 149.89958292986054  
v9 180.00235293348715  
v0: woman  
v1: man  
v10: woman  
v11: man  
v2: man  
v3: man  
v4: woman  
v5: woman  
v6: man  
v7: woman  
v8: man  
v9: woman
```

۱. الگوریتم های subtraction spectral : تفریق طیفی یکی از اولین الگوریتم‌هایی است که برای تقویت گفتار تک کانالی پیشنهاد شده است. در این روش، طیف نویز در حین مکث گفتار تخمین زده می شود و برای تخمین گفتار تمیز، از طیف گفتار پر سر و صدا کم می شود. این نیز با ضرب طیف گفتار پر سر و صدا با تابع افزایش و ترکیب آن با فاز گفتار پر سر و صدا به دست می آید. اشکال این روش وجود اعوجاج های پردازشی است که به آن نویز باقیمانده می گویند. تعدادی از تغییرات این روش در طول سال های گذشته برای رفع اشکال توسعه یافته است. این گونه ها خانواده های از الگوریتم های طیفی تفریق را تشکیل می دهند. الگوریتم های زیادی همچون تفریق طیفی پایه، تفریق بیش از حد طیفی، تفریق طیفی چند بانندی، فیلتر وینر، تفریق طیفی تکراری و تفریق طیفی بر اساس خواص ادراکی ساخته شده اند. برای آزمایش عملکرد الگوریتم های نوع کسر، معیارهای هدف SNR و PESQ، طیف نگارها و تست های شنیداری غیررسمی برای انواع نویزهای ثابت و غیر ثابت در سطوح مختلف SNR انجام می شود. از نتایج مشهود است که مشکل روش تفریق طیفی نویز باقیمانده را به میزان قابل توجهی کاهش می دهد و گفتار تقویت شده دارای حداقل تحریف گفتار است.

۲. ورودی های این تابع فرکانس سیگنال اصلی و ورودی دوم سیگنال نویز دار و ورودی سوم خود نویز است. ابتدا تبدیل فوریه سیگنال نویز دار را حساب می کنیم سپس قدر مطلق از تبدیل فوریه سیگنال نویز دار قدر مطلق می گیریم و همین کار را برای نویز انجام می دهیم.

هر کدام از اعضای `signal_noisy_fft` یک عدد مختلط به فرم $iy + x$ است که میتوان به فرم قطبی آنرا نوشت. یک متغیر b را به صورت $e^{i(\theta)}$ تعریف می کنیم که تتای ما همان زاویه به دست آمده است. سپس میانگین قدر مطلق تبدیل فوریه نویز را بدست می آوریم. در نهایت سیگنال نویزگیری شده را محاسبه می کنیم که برابر است با عکس تبدیل فوریه که میشود تفاضل میانگین اندازه تبدیل فوریه سیگنال نویز و اندازه تبدیل فوریه سیگنال نویز دار ضربدر b و سیگنال نویزگیری شده ما آماده میشود.

۳. ورودی ها فرکانس و فایل تست است. SNR نسبت نرخ سیگنال به نویز است. هر چه مقدار SNR کمتر باشد نویز بیشتر است.

ابتدا `signal square mean root` را برای سیگنال اصلی بدست می آوریم. `Square mean Root` جذر میانگین اعضای لیست به توان دو است. `Mean` یک لیست به عنوان آرگومان میگیرد و میانگین آنرا برمیگرداند. یعنی ابتدا اعضا به توان دو میرسن و سپس از آنها میانگین میگیرد و در نهایت از آن جذر میگیرد. سپس `RMS` را برای نویز بدست می آوریم. جذر `RMS` سیگنال اصلی به توان دو تقسیم بر ۱۰ به توان `SNR/10` سپس یک لیست تصادفی به طول صدای ورودی از توزیع نرمال که بین ۰ تا `noise square mean root` تولید میکند. در نهایت نیز سیگنال اصلی را به علاوه این نقاط تصادفی ایجاد شده میکند و برمیگرداند و نویزها ایجاد میشوند.

نمودار فایل اصلی، فایل نویزی و فایل نویزگیری شده و همچنین فایل صدای آنها ضمیمه شده اند.