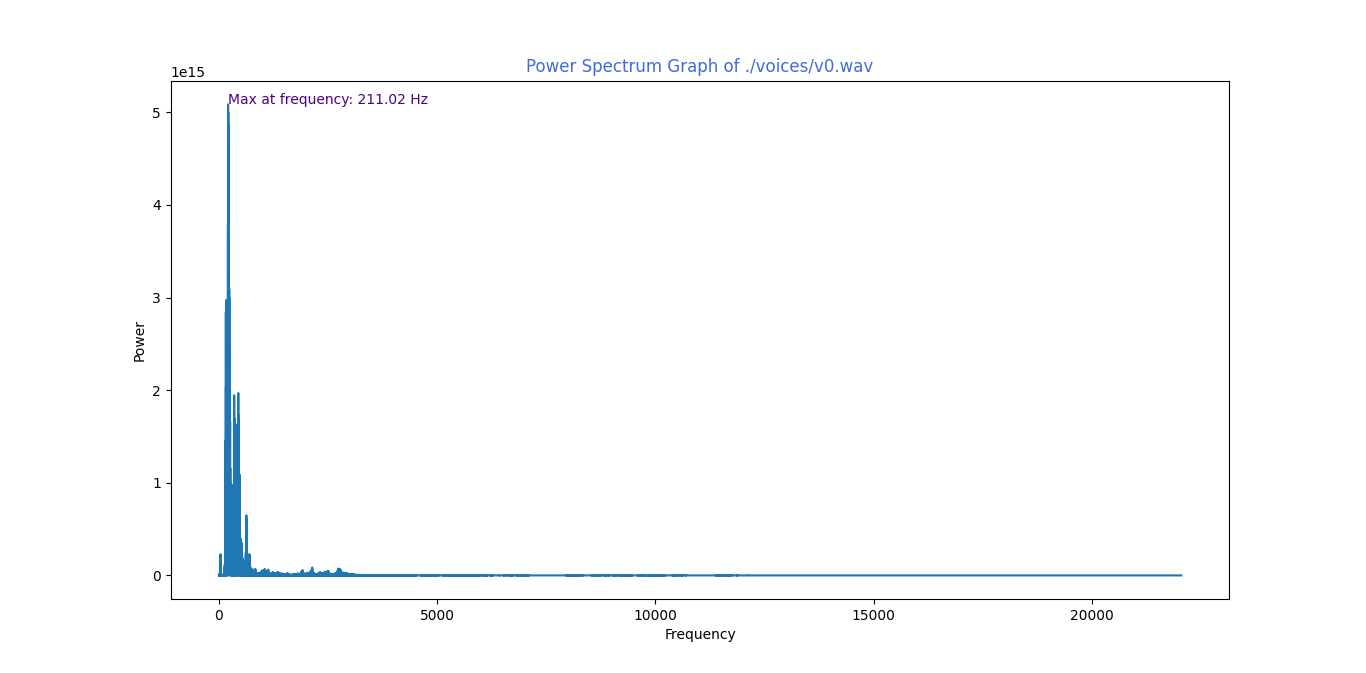
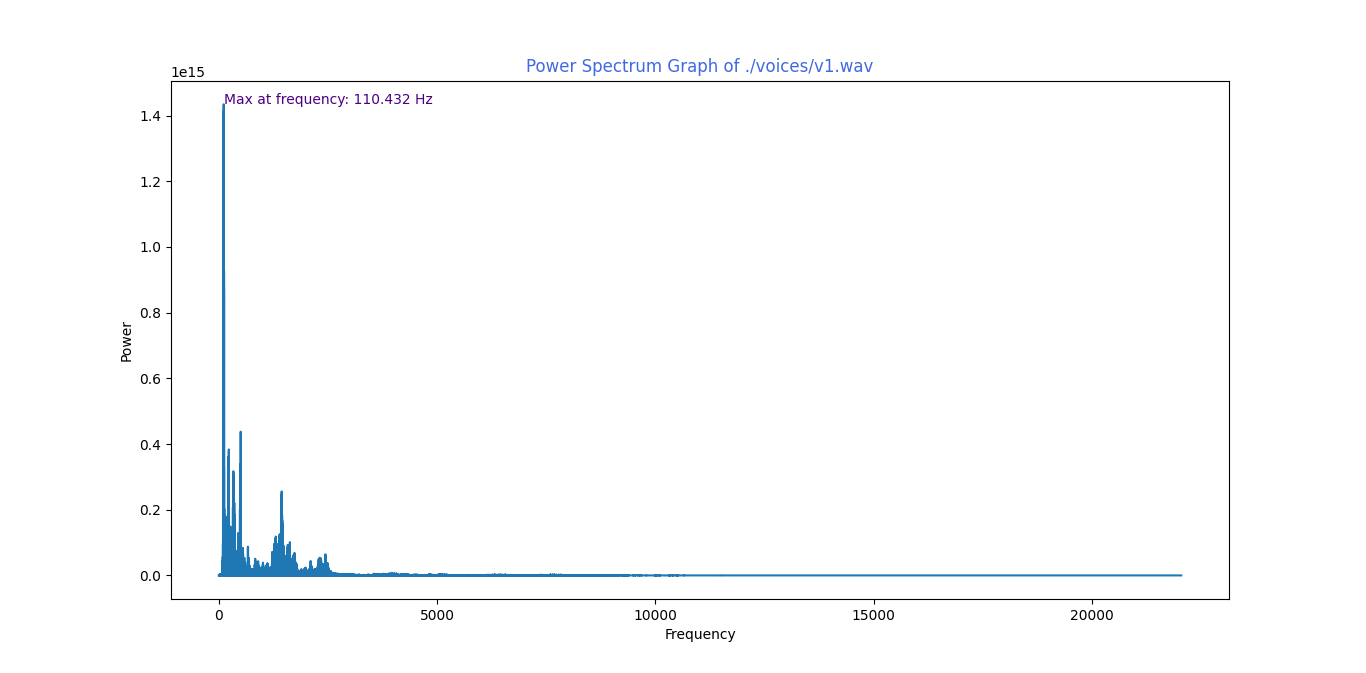
**پروژۀ امتیازی درس سیگنال­ها و سیستم­ها**

نگار موقتیان، 9831062

**بخش اول – تشخیص صدا**

می­دانیم اکثر سیگنال­ها را با استفاده از تبدیل فوریه می­توان به صورت جمعی از توابع سینوسی نوشت و می­توان این توابع سینوسی را اجزاء تشکیل دهندۀ سیگنال دانست. همچنین می­دانیم تبدیل فوریۀ هر سیگنال یک تابع یکتاست، بنابراین با بردن سیگنال به حوزۀ زمان در حقیقت دادۀ یکسانی را از دو دید متفاوت بررسی می­کنیم. بنابراین در تبدیل فوریه برای تجزیه و تحلیل طیفی، به جای تحلیل مستقیم سیگنال این اجزاء سازنده و نسبت آن­ها با استفاده از تبدیل فوریه و در حوزۀ فرکانس بررسی می­شوند. از این طریق می­توان به اطلاعاتی دست یافت که استخراج آن­ها در حوزۀ زمان دشوار است.

شکل­های زیر نمودار power spectrum مربوط به یکی از صداهای زن و یکی از صداهای مرد را نشان می­دهند.



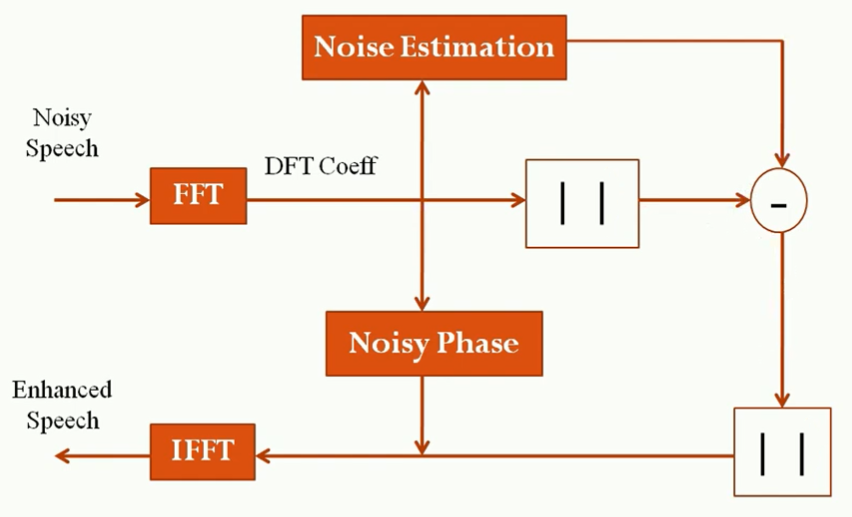
همانطور که در این دو شکل مشاهده می­شود، فرکانسی که به ازای آن توان مقدار ماکسیمم را دارد در صدای زنان (شکل اول – 211 هرتز) از صدای مردان (شکل دوم – 110 هرتز) به طور معمول بیش­تر است. بنابراین از این طریق و با استفاده از تحلیل طیفی می­توان میان صدای مردان و زنان تفکیک قائل شد.

برای بهینه­سازی زمان اجرا و کاهش پیچیدگی­های این تابع می­توان بازۀ فرکانسی مورد بررسی را محدود کرد. همانطور که در دستور کار نیز اشاره شده صدای معمول انسان میان 50 تا 300 هرتز متغیر است، بنابراین برای تشخیص جنسیت صدا نیازی نیست که طیف فرکانسی­ای به گستردگی شکل بالا را بررسی کرده و تبدیل فوریۀ تابع را به ازای تمام این بازه محاسبه کنیم. بلکه کافیست برای مثال فرکانس­هایی میان 0 تا 400 را مورد بررسی قرار دهیم.

همچنین برای افزایش دقت این برنامه می­توان از یکسری دادۀ training استفاده کرد که از پیش برچسب­گذاری شده­اند و از این طریق می­توان بازه­ای که برای صدای مردان و زنان مشخص شده را دقیق­تر کرد.

**بخش دوم - نویزگیری**

Spectral subtraction روشی برای بازیابی توان و یا اندازۀ یک سیگنال نویزدار در حوزۀ فرکانس است. برای این کار ابتدا تخمینی از میانگین نویز زده شده و سپس این مقدار از سیگنال نویزدار کم می­شود. روند کلی این روش در شکل زیر آمده­است.



*در ابتدا از سیگنال نویزدار تبدیل فوریه گرفته شده و اندازه و فاز آن جدا می­شود. از طرف دیگر تخمینی از میانگین تبدیل فوریۀ نویز نیز زده می­شود (در این روش فرض می­شود نویز یکنواخت بوده و تبدیل فوریۀ آن در بازه­های مورد بررسی تغییر زیادی ندارد). حال در صورتی که هدف ما بازیابی توان سیگنال باشد مربع میانگین نویز از مربع اندازۀ تبدیل فوریۀ سیگنال کم شده و از آن جذر گرفته می­شود و اگر هدف بازیابی اندازۀ سیگنال باشد میانگین نویز از اندازۀ تبدیل فوریۀ سیگنال کم می­شود (در برنامۀ نوشته شده از روش دوم استفاده شده­است). در نهایت نیز فازی که از سیگنال نویزدار بدست آورده بودیم در این اندازه ضرب می­شود و با گرفتن تبدیل فوریه معکوس از سیگنال حاصل، سیگنال بدون نویز تولید می­شود.*

الگوریتم پیاده­سازی شده به طور کلی مانند توضیحات قسمت قبل است. در پیاده­سازی این الگوریتم به جای FFT از STFT استفاده شده­است. این تبدیل نوعی از تبدیل فوریه است که در آن به جای گرفتن تبدیل فوریه از کل سیگنال، سیگنال به بازه­ها یا پنجره­هایی (که با یکدیگر همپوشانی دارند) تقسیم شده و تبدیل فوریه در هر یک از این قسمت­ها محاسبه و بررسی می­شود. STFT به طور معمول در تحلیل سیگنال­های صوتی استفاده می­شود، زیرا در این سیگنال­ها تغییرات زیادی در کل طول سیگنال داریم و لذا تبدیل فوریه را در بازه­های کوچک­تری بررسی می­کنیم.

در برنامۀ نوشته شده این تبدیل با اندازۀ پنجرۀ برابر با 30 × 10-3 و همپوشانی برابر با 50% از طریق تابع audio\_to\_frames اعمال شده­است.

پس از آن با استفاده از تابع estimate noise میانگین نویز سیگنال در هر پنجره تخمین زده می­شود. در این تابع بر روی سیگنال پیمایش شده و نویز هر پنجره طبق رابطۀ زیر از روی میزان نویز پنجره­های قبلی محاسبه می­شود.

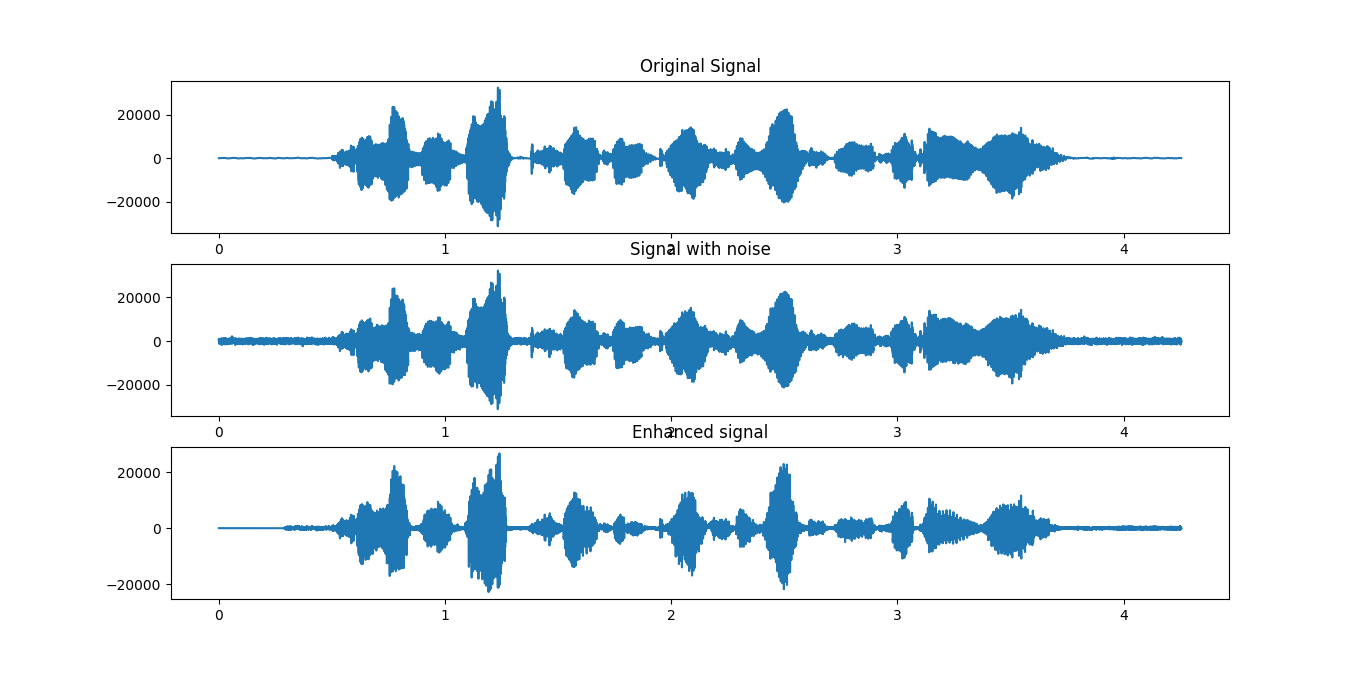
که در آن y[n] مقدار تخمین زده شده و x[n] خروجی بدست آمده از قسمت (با اعمال STFT) قبل است.

همچنین برای بدست آوردن مقدار آلفا در هر مرحله از رابطۀ زیر استفاده می­شود.

*که در آن* a *برابر با 25 و* T *برابر با 1.5 در نظر گرفته شده و گاما حاصل تقسیم مربع اندازۀ سیگنال نویزدار در هر مرحله بر روی مربع میانگین اندازۀ نویز تخمین زده شده در چند (در اینجا 20) فریم قبلی است. همچنین مقدار نویز تخمین زده شده در 20 فریم اول برابر با اندازۀ خود سیگنال قرار داده می­شوند (فرض می­شود که در ابتدای صوت تنها نویز داریم و شخص سکوت کرده است).*

*پس از تخمین نویز در تابع* spectral­\_subtraction *مراحلی که در سوال 1 توضیح داده شده بود طی شده و در نهایت با استفاده از تابع* frames\_to\_audio *توسط* STFT *معکوس به حوزۀ زمان برگردانده می­شود.*

*شکل زیر مثالی از اجرا روند گفته شده می­باشد.*



*همانطور که مشاهده می­شود نویزها در شکل سوم کاهش داده شده­اند. در چند فریم اول مقدار نویز برابر با مقدار خود سیگنال نویزدار قرار داده شده بود، لذا در ابتدای سیگنال خروجی هیچ نویزی نداریم. اما پس از آن مقدار نویز تخمین زده می­شود و هنوز مقداری نویز داریم.*

*این روش برای نویزهایی با* SNR *ها مختلف کار می­کند اما در صورتی که مقدار نویز بالا باشد به دلیل اینکه آن را تخمین می­زنیم ممکن است شکل سیگنال اصلی پس از حذف نویز تغییر زیادی بکند و صدای فرد کیفیت اولیه را نداشته باشد.*