

实验二：语音信号的频分复用

陈炫中 2019011236

1. 实验目的

- 将多段采样率 8000Hz 的信号在频域上合并为一段采样率 48000Hz 的信号验证“频分复用原理”。

2. 实验过程

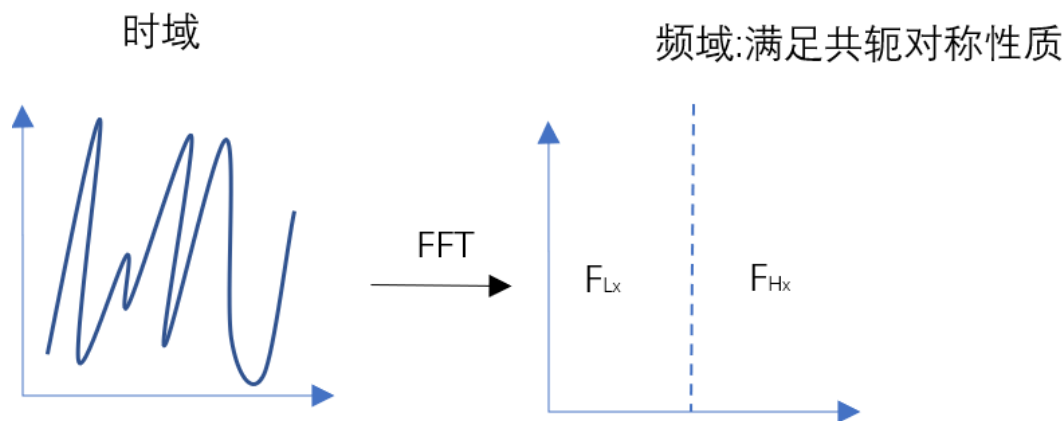
收集的 4 段不同的 30s 原始音频的内容来自普通话测试朗读示范语音，见 ./data 文件夹。

对这四段语音信号进行频分复用，将四段 30 s 音频频谱拼接到一起，并保证拼接后频谱的共轭对称性，这就将四段音频调制到了一段音频中；将调制后的信号分解到单独的频段，进行 IFFT，就解调得到原来的四段音频，具体过程如下：

首先对读入的音频进行预处理：重采样至 8000 Hz, 滤去 3400 Hz 以上的音频信号

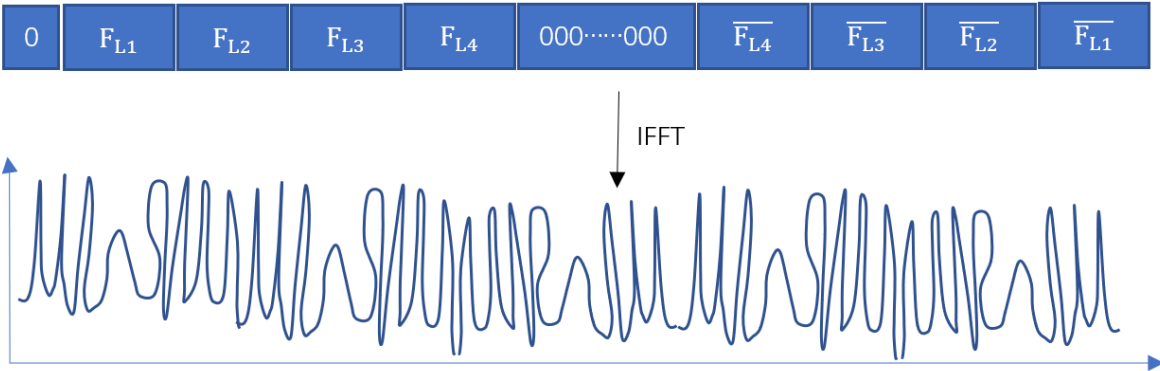
```
# 预处理
raw_data, ori_sample_rate = librosa.load(f'./data/{i + 1}.wav',
duration=DURATION)
b, a = signal.butter(10, FILTER_FREQ, fs=ori_sample_rate)
filtered_data = signal.filtfilt(b, a, raw_data)
resample_data = librosa.resample(filtered_data, ori_sample_rate, SAMPLE_RATE)
```

每一段音频经过傅里叶变换后为频谱，每一段频谱在频率范围上共轭对称，将 4 段频谱按滤波后拼接到一起，并使其满足频谱共轭对称的性质，完成音频的调制，如下示意图所示：



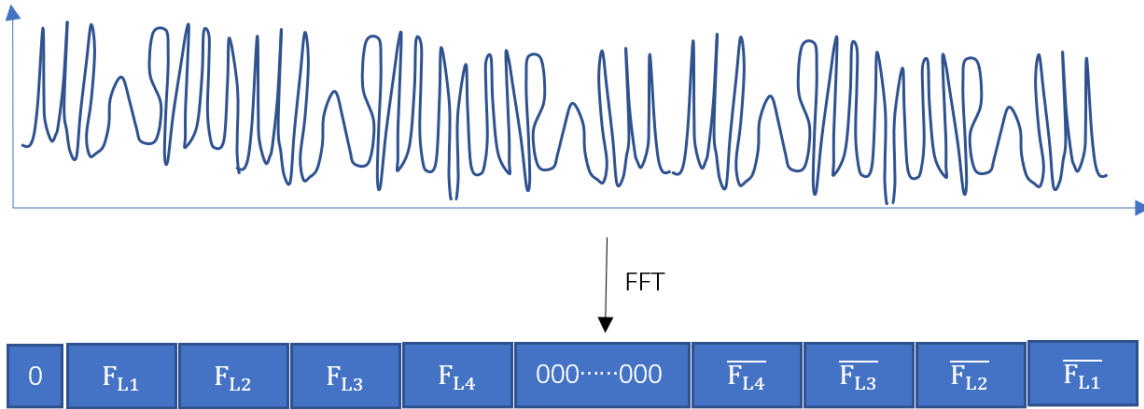
注意到：对于一段长度为 N 的信号 (N 为偶数)， $X(\frac{N}{2} + K) = X^*(\frac{N}{2} - K)$ ，即 $x(1)$ 与 $x(N-1)$ 共轭对称，并非 $x(0)$ 与 $x(N-1)$ 共轭对称。因此我将 $x(0)$ 置为 0，频谱信息从 $x(1)$ 开始拼接，中间部分全部置为 0，使其满足 48000Hz 的采样率，得到调制后的频谱，如下所示：

合并后的频谱使其满足共轭对称性质： $\overline{F_{Lx}}$ 是 F_{Lx} 共轭并反转



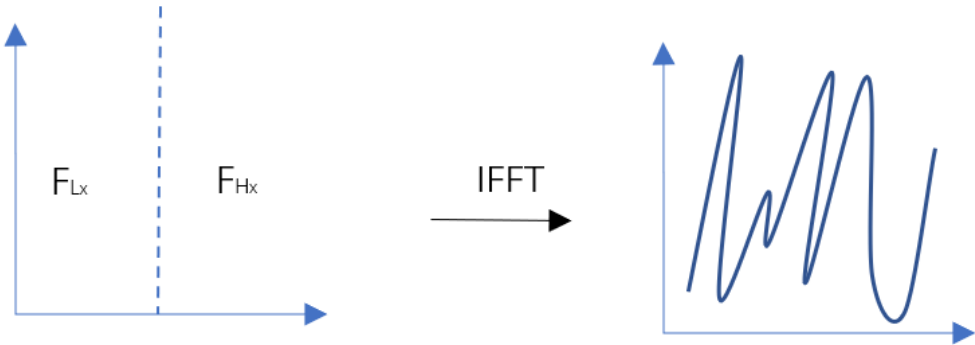
调制后的频谱进行 ifft 变换，得到合并后的音频时域信号，丢弃虚部，通过 python 的 `soundfile` 库输出到 `encode.wav`。

之后将调制后的信号解调到各段音频，保存为 `decode_{i}.wav` ($i = 1, 2, 3, 4$)，如下示意图所示：



频域

时域



而对于 $N = 1s, 2s, 5s, 10s$ 的分帧频分复用来说，在这个任务中，上述对整段音频进行频分复用即相当于此处 $N = 30s$ 的结果，所以对 $N = 1s, 2s, 5s, 10s$ 的分帧频分复用同理，将音频切为多个等长的片段，每一个片段的长度为 N 秒，分别对每一片段进行频分复用，最终再解码。解码后将每一帧进行拼接，组合为完整的音频即可，符合实际的通信场景。

$N = 1s, 2s, 5s, 10s, 30s$ 频分复用所得到的一段编码后的音频以及解码后的四段音频详见 `./output` 文件夹，解码后的各段音频几乎与原始音频完全一致，声音效果也很好。

实验结果的复现通过如下命令即可：

```
python fdm.py
```