

# 信号处理原理 语音信号的频分复用

刘泓尊 2018011446 计 84

Department of Computer Science, Tsinghua University

2020 年 11 月 17 日

## 目录

<b>1 实验原理</b>	<b>1</b>
<b>2 实验过程</b>	<b>1</b>
2.1 原始音频信号与预处理	1
2.2 调制与频分复用	2
2.3 解调与 MSE 误差	3
2.4 分帧频分复用与 MSE 误差	5
<b>3 实验小结</b>	<b>5</b>

**文件说明** 项目文件夹根目录下 main.py 是程序脚本，默认 Frame Duration 为 1s. /audio 文件夹下存放了原始音频和复原后的音频。/figure 文件夹下存放了不同 Frame Duration 下的频谱图、时域图以及 FDM 之后的频谱、时域图，还保留了 MSE 计算结果。

## 1 实验原理

频分复用：语音频带通常在 300Hz 到 3400Hz, 8KHz 采样之后的数字音频信号可以看做是频带受限信号。将多路音频的频带互不重叠地放在频域，可以实现多路语音同时传输的目的。

## 2 实验过程

我选择了周杰伦的 5 首歌曲（稻香、夜曲、青花瓷、七里香、告白气球）作为原始音频。

### 2.1 原始音频信号与预处理

下面是 5 段音频的时域波形，每个时长约 30s. 所有信号的时域截取到 30s.

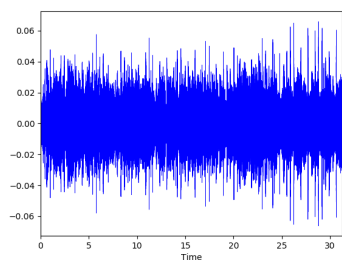


图 1: 稻香

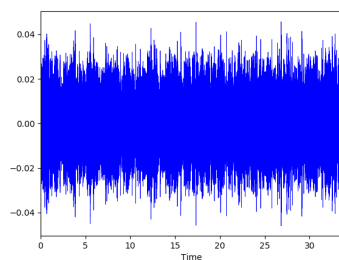


图 2: 七里香

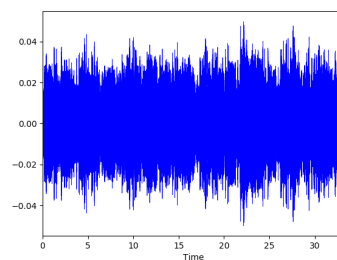


图 3: 青花瓷

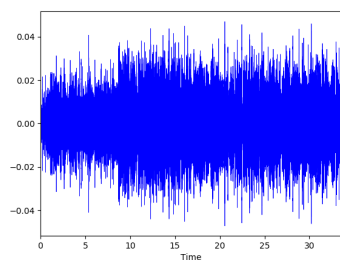


图 4: 夜曲

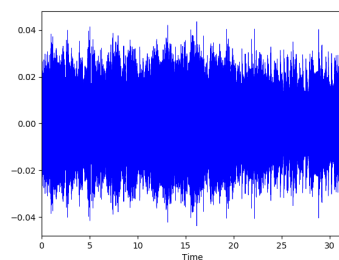


图 5: 告白气球

下面是每段音频的频域信号，8K 采样。

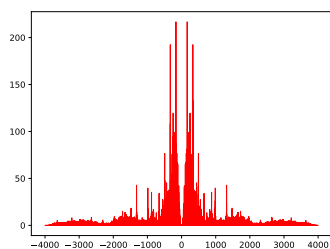


图 6: 稻香

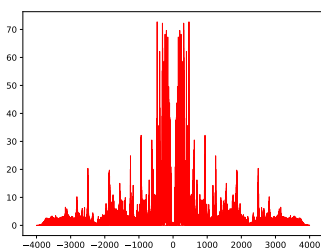


图 7: 七里香

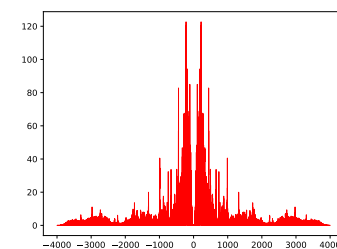


图 8: 青花瓷

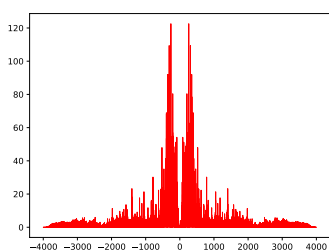


图 9: 夜曲

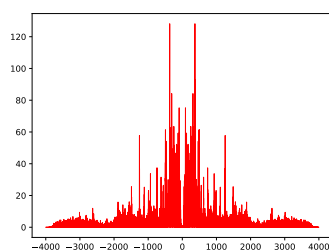


图 10: 告白气球

之后使用低通滤波将 3400Hz 以上的频率滤掉，作为频分复用模块的输入。这里就不再展示滤波后的信号频谱了。

## 2.2 调制与频分复用

之后将 5 路信号放在同一个频谱上，以 4000Hz 为间隔排布，合成之后的频域信号如下。注意保持频域共轭对称的性质。

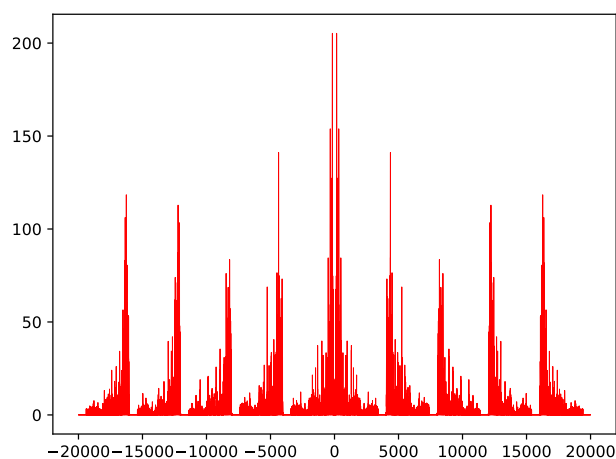


图 11: FDM 之后的频域信号

将上述频域信号 IDFT 之后得到时域信号，存储到磁盘，之后再从磁盘中读出来（模拟了信道传输过程）。时域信号如下。

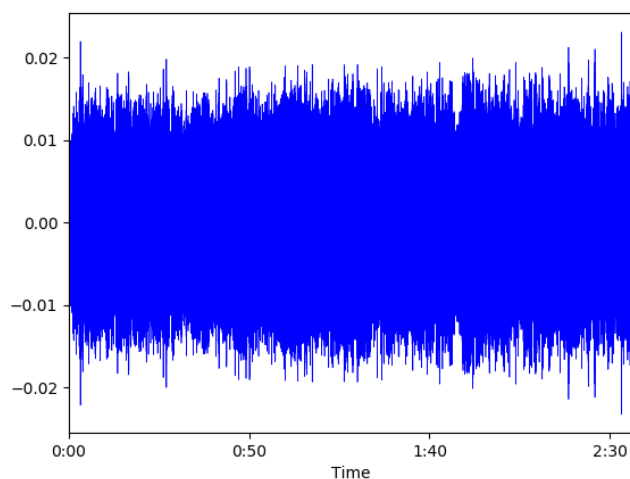


图 12: FDM 之后的时域信号

## 2.3 解调与 MSE 误差

读取 FDM 之后的音频，进行 DFT 获得频谱，然后按照调制阶段的规则将每段的信号拆分出来，恢复到原始的  $[-4000, 4000]$ Hz 的频带，下图展示了恢复之后的频带信号。可以看到恢复出来的信号频域和原始频域几乎完全一致。（注意在上一步中 3400Hz 以上的频率已经滤掉了）

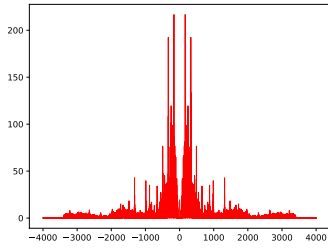


图 13: 稻香

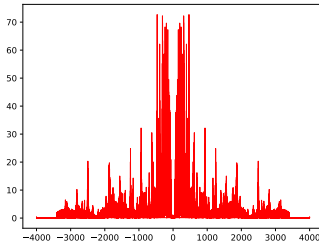


图 14: 七里香

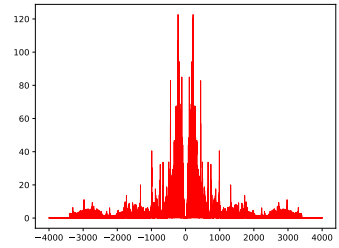


图 15: 青花瓷

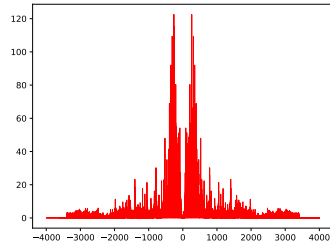


图 16: 夜曲

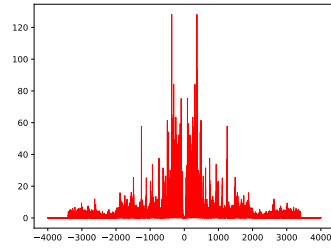


图 17: 告白气球

将上述频谱进行 IDFT 变到时域即可复原出原始信号，完成解调的全部过程。下面是恢复出来的时域信号。音频几乎没有损失。

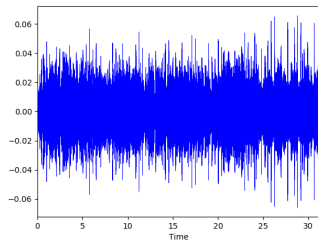


图 18: 稻香

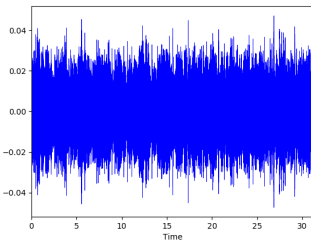


图 19: 七里香

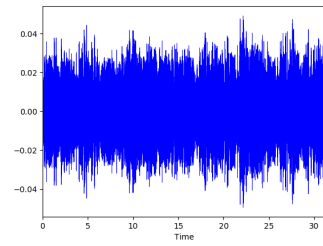


图 20: 青花瓷

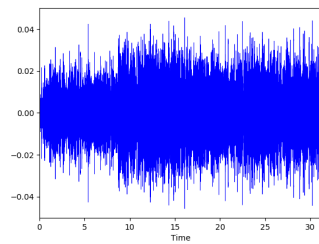


图 21: 夜曲

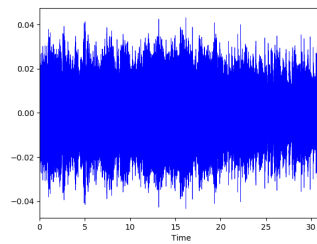


图 22: 告白气球

下面是每路信号的 MSE 误差。值得注意的是，因为我使用了歌曲作为音频，在采样和滤波过程中滤掉了乐器的声音，所以 MSE 较大。经过亲自实验，如果采用纯人声信号做实验，MSE 会在 0.01 量级左右。

音频	稻香	七里香	青花瓷	夜曲	告白气球
MSE	0.3136	0.3130	0.3178	0.3201	0.3155

## 2.4 分帧频分复用与 MSE 误差

接下来我进行了分帧的频分复用，每帧长度分别为  $N = 1s, 2s, 5s, 10s, 20s$ 。

下面以  $N = 10s$  为例。将原始信号每  $10s$  分为一帧，时长不足则补 0。下图展示了  $N = 10s$  时，分成 3 帧分别进行频分复用之后得到的频带信号。可以看到频带随着时间变化而不同。

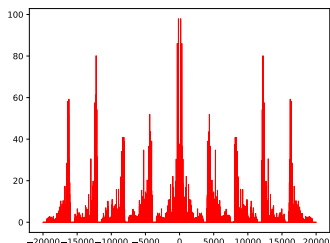


图 23: 0s-10s

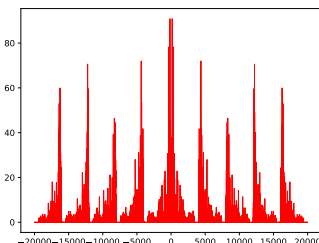


图 24: 10s-20s

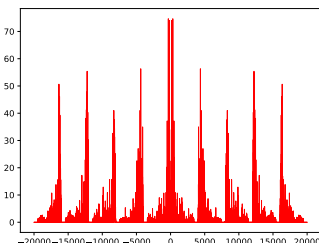


图 25: 20s-30s

然后将每帧频带信号分别 IDFT 之后拼接成混合音频，存入磁盘。之后再从磁盘中读取处理后的音频，按相同的切分规则取出每段音频的频带，IDFT 之后按帧拼接即可得到原始信号。在这里不再展示复原后的音频，效果与上一部分是一样的，几乎没有失真。

我对不同帧时长的 MSE 误差进行了统计，结果如下：

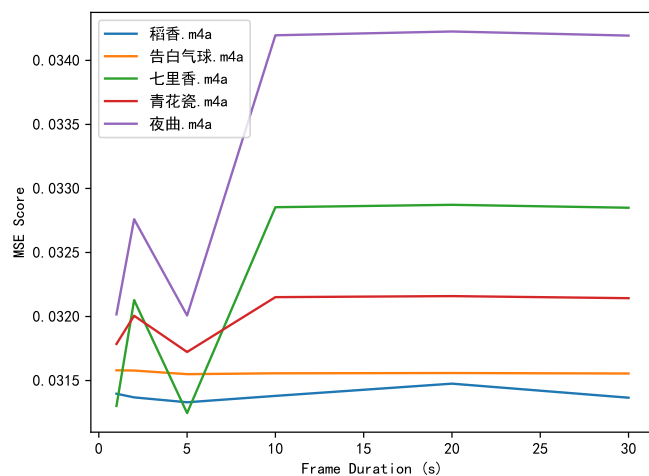


图 26: 采用不同帧时长的分帧频分复用的 MSE

可以看到，当采用较小的 Frame Duration 时 (如  $1s, 2s$ )，复原信号的 MSE 较低，有更好的复原效果。较长的 Frame Duration 的误差相差不大。而且高音乐器较多的歌曲误差会更大。

## 3 实验小结

本次实验我深入理解了频分复用方法。在合成多路的频率信号时，我一开始没有考虑到要保持信号共轭对称的性质，只是将信号的幅值划分到不同频带，导致复原出的信号失真严重，经过认真排查终于解决。可见信号频谱的幅值和相位都携带了很重要的信息，不能忽略任何一个。

感谢老师和助教的悉心指导！