



中国科学技术大学
University of Science and Technology of China

语音信号处理基础

课程实验报告

学生姓名：姚苏航

学号：PB22061220

提交时间：

报告评分：

说明：

1. 每一道实验题目都需要按照第 2-3 页的模板填写实验报告；例如，第一次编程实验有 3 道题目，需要填写 3 次；
2. 中英文皆可；
3. 不是每一项都需要填写，根据实验难度选择。但实验分析，实验步骤，实验结果不可少。

Task One

Analysis of the Experimental Task

1. What does the question mean? What's the input and expected output of the experiment?

这个问题要求编写一个 MATLAB 程序，能够将语音文件加速 2 倍。具体的处理方法是通过重叠加法（overlap-add）分析信号的短时傅里叶变换（STFT），在每个分析帧中使用矩形窗口（512 个样本），帧之间有 256 个样本的重叠。然后，抛弃每隔一个帧，最后重新合成加速后的语音信号。

输入：一个语音文件（例如 test_16k.wav），采样频率为 16kHz。

输出：加速后的语音文件及其对应的窄带谱图。

2. Can you give theoretical analysis on the question?

STFT：短时傅里叶变换将信号从时域转换到频域，通过窗口函数对信号进行分帧，得到频域信息。

overlap-add：是合成处理后信号的方法，通过重叠并合并各个处理后的帧，以减少帧之间的边界效应

3. What's the expected result for the experiment?

原始语音信号的时域波形图。

加速后的语音信号的时域波形图。

原始语音的窄带谱图（STFT 谱图）。

加速后的语音的窄带谱图（STFT 谱图）

Details of Experiment

1. Configuration of the experiment, e.g. Length of FFT, sampling frequency.

窗口长度：512 样本（每帧）。

重叠长度：256 样本（50%重叠）。

采样频率：16 kHz（文件 test_16k.wav 的采样频率）。

FFT 点数：1024（用于频域分析）

2. Procedure for experiment

a) Step1 : 使用 audioread 函数读取输入的语音文件（例如 test_16k.wav）

b) Step2 : 将信号分帧，每帧使用矩形窗口（rectwin）

c) Step3 : 使用 fft 或 spectrogram 计算信号的短时傅里叶变换，得到频域表示。

d) Step4 : 选择每两个分析帧中保留一个帧，减少帧数，达到加速效果。

e) Step5 : 绘制结果

Results of Experiment

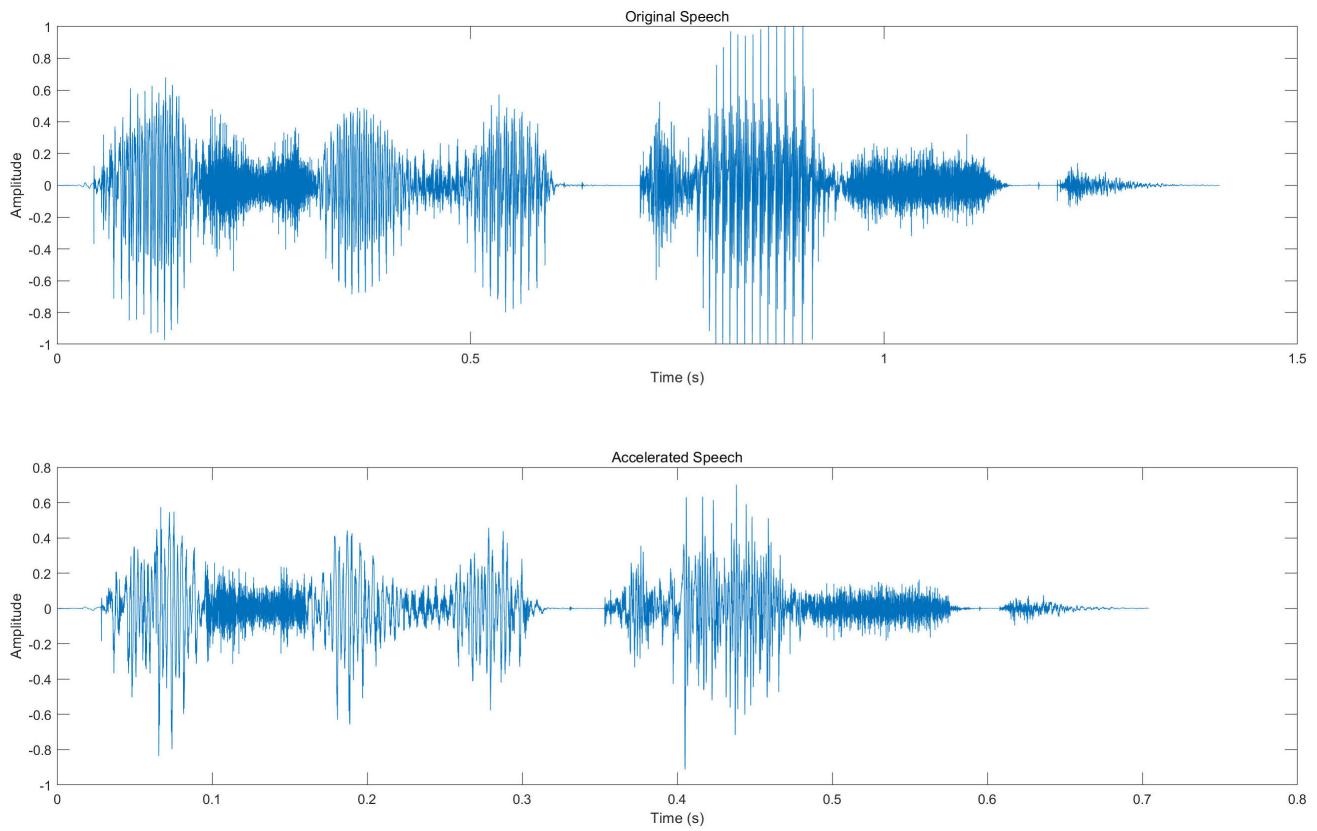


Fig1: 信号时域波形图

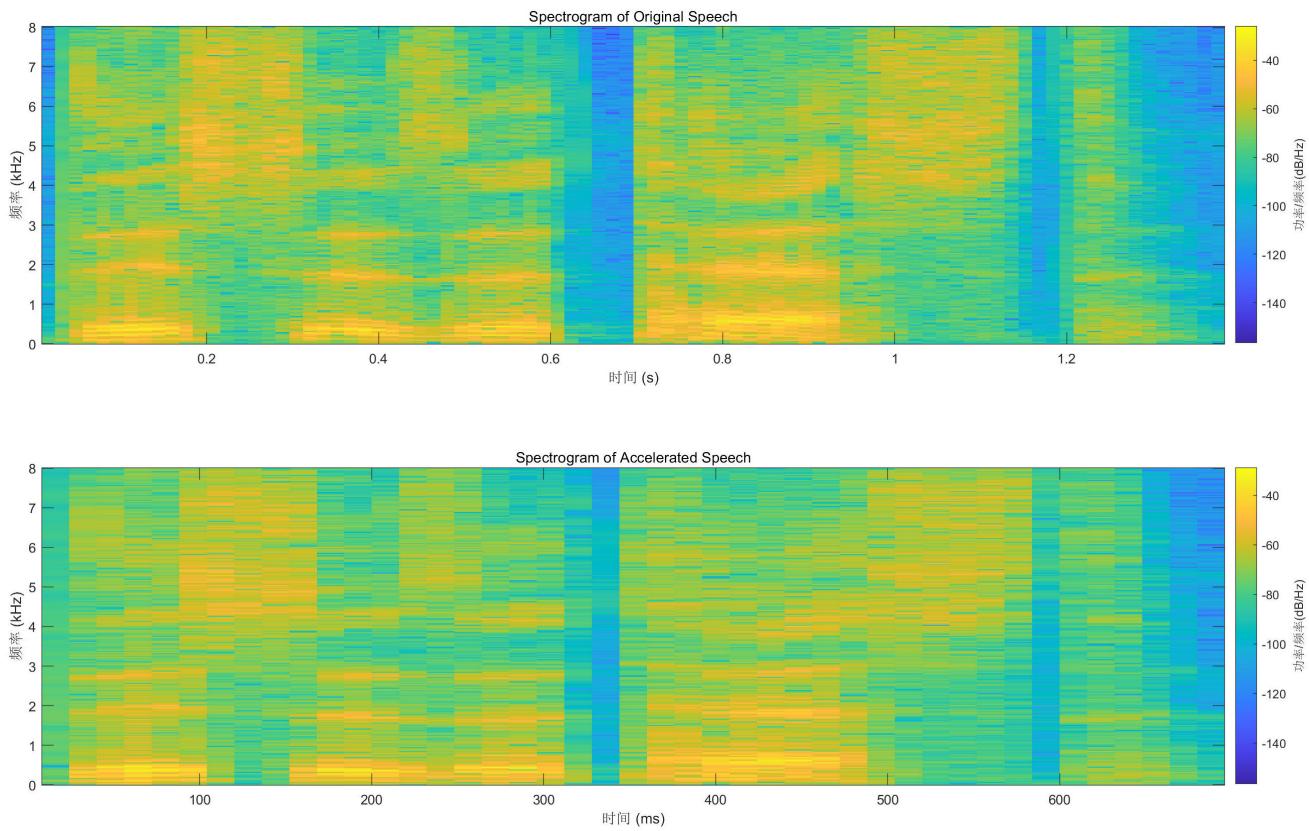


Fig2: STFT 谱图

1. Analysis of the results:
 - a) Is the result in accord with the expected result?
 - b) If so, provide further analysis on the results. For example, provide some statistics, data, and other evidence to testify the effectiveness of your experiment.
 - c) If not, explain the unexpected the result and give possible reason for it.

结果和预期保持一致，加速后的语音波形会被压缩，整体时长会变短。由于去除了每隔一帧的频域信息，语音的节奏变快，周期较短，波形的间隔会显得更为密集。虽然语音的频率变高，但波形的形态保持类似，主要的区别是时间轴上的压缩，并且加速后的语音在听觉上也更加短促。

Task Two

Analysis of the Experimental Task

1. What does the question mean? What's the input and expected output of the experiment?
这个问题要求计算语音信号的实数倒谱（real cepstrum），并应用两种类型的倒谱加窗（低频倒谱加窗和高频倒谱加窗），以此来展示表达发音道响应（低频倒谱加窗）和源激励（高频倒谱加窗）之间的差异。

输入：

- 语音文件（如 test_16k.wav）
- 起始样本（例如，开始样本为 13,000 对于有声帧，1,000 对于无声帧）
- 语音帧的持续时间（例如 40 毫秒，需转化为样本数）
- 用于计算倒谱的 FFT 大小
- 倒谱滤波的截止点（低频滤波或高频滤波的截止点）
- 滤波类型（低频滤波或高频滤波）

输出：

- 有声和无声语音帧的处理结果，主要包括以下图像：
- 加权汉明窗口后的语音部分
- 语音信号的对数幅度谱和经过倒谱平滑的对数幅度谱（低频滤波）
- 语音信号的真实倒谱
- 经过倒谱滤波后的对数幅度谱（低频滤波与高频滤波）

2. Can you give theoretical analysis on the question?

倒谱是将语音信号进行傅里叶变换后，取其对数幅度谱并反变换得到的结果。倒谱的一个重要应用是分离源和传递响应，它通过区分低频和高频来提取语音的发声源和传递路径（声道特征）。

3. What's the expected result for the experiment?

加权窗后的时域信号：显示窗函数（汉明窗）加权后的语音信号，它应该是原始语音帧的一个局部窗口，窗口两端的振幅会逐渐减小。

对数幅度谱和倒谱平滑后的对数幅度谱：对数幅度谱反映语音信号在频域的特征，平滑后的对数幅度谱可以减少不必要的高频噪声。低频提升会使平滑后的谱更加突出低频部分。

实倒谱：实倒谱可以看作是源和传递响应的分离，通常会看到低频部分与高频部分的分开。

倒谱提升后的对数幅度谱：低频提升后的谱图将会更加突出语音信号中的低频成分，反映声道特性；而高频提升后的谱图则更强调信号的激励部分。

Details of Experiment

Configuration of the experiment, e.g. Length of FFT, sampling frequency.

采样频率：16 kHz

FFT 长度：512（可根据需要调整）

语音帧的长度：40 毫秒，转换为样本数：frame_length = 40e-3 * fs。

使用汉明窗对每帧进行加窗。

Step1 :读取语音文件并根据给定的起始样本和帧长度提取语音帧。

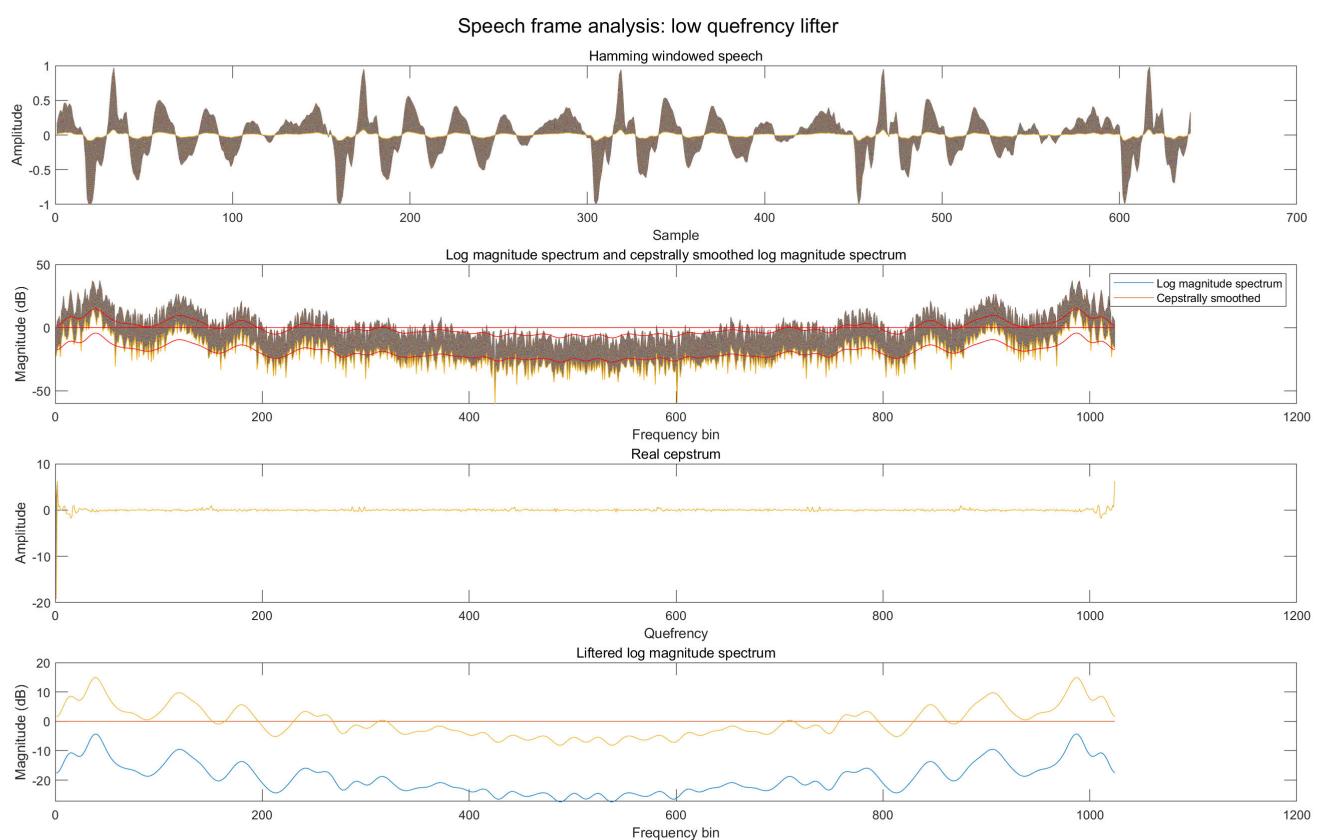
Step2 :对提取的帧进行 STFT 和对数幅度谱计算。

Step3 :对对数幅度谱进行反傅里叶变换，得到实倒谱

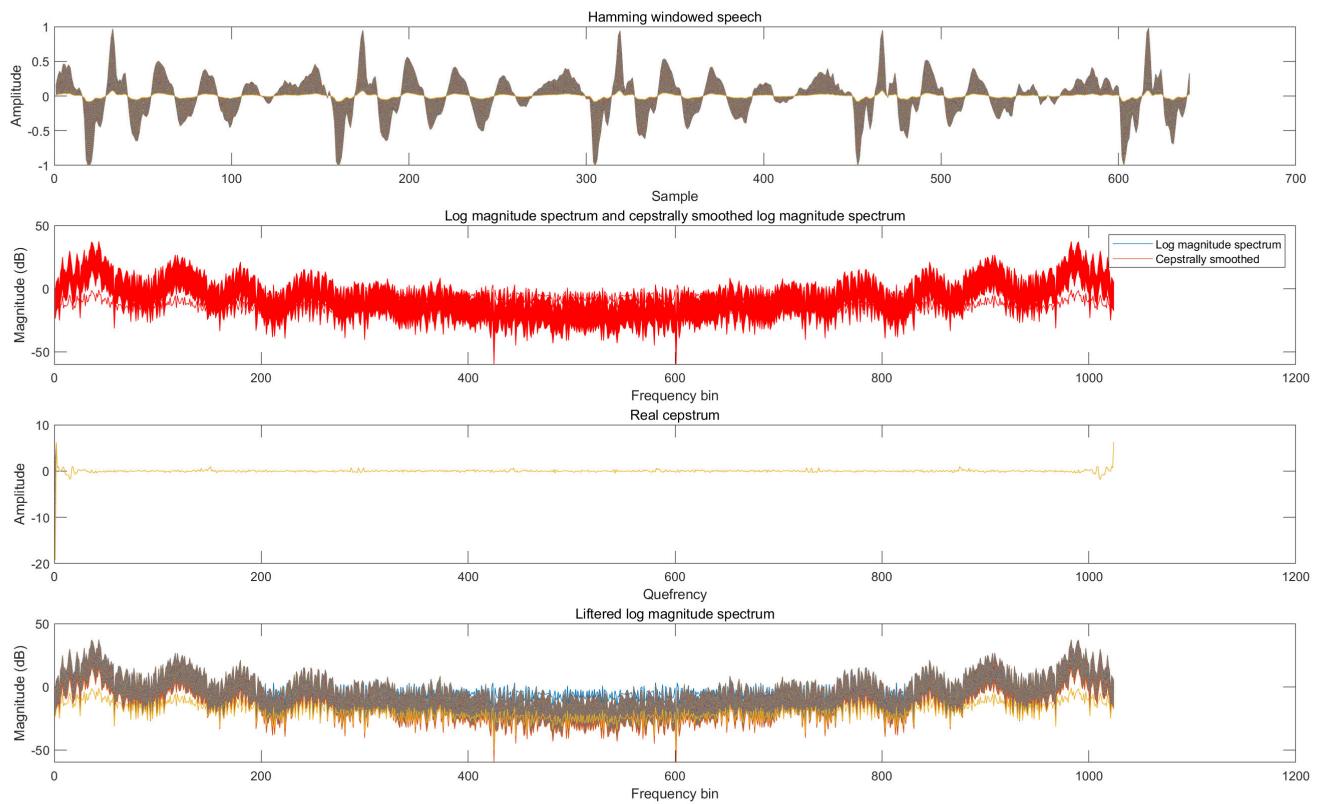
Step4 :根据低频或高频的要求，进行倒谱滤波。

Step5 :绘制加窗后的语音信号、对数幅度谱、倒谱及其提升后的频谱图。

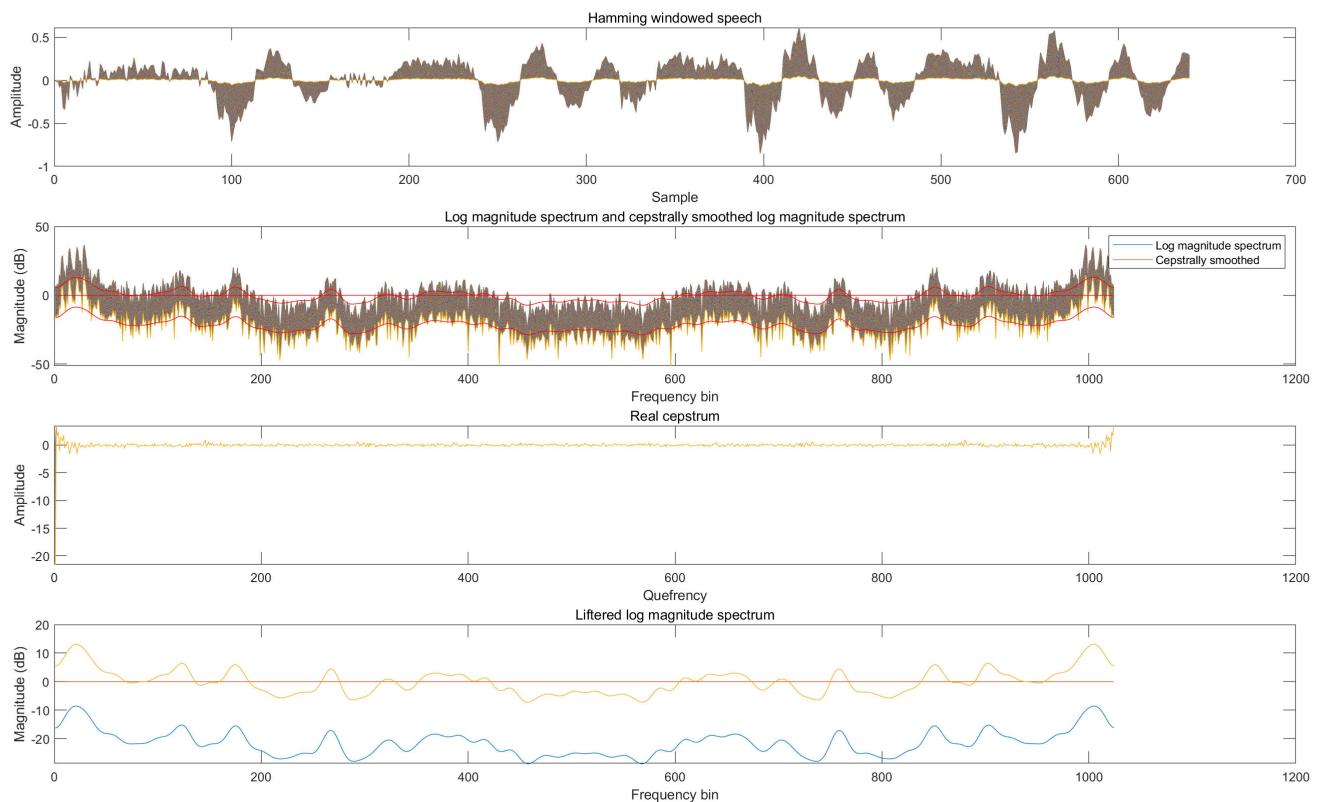
Results of Experiment

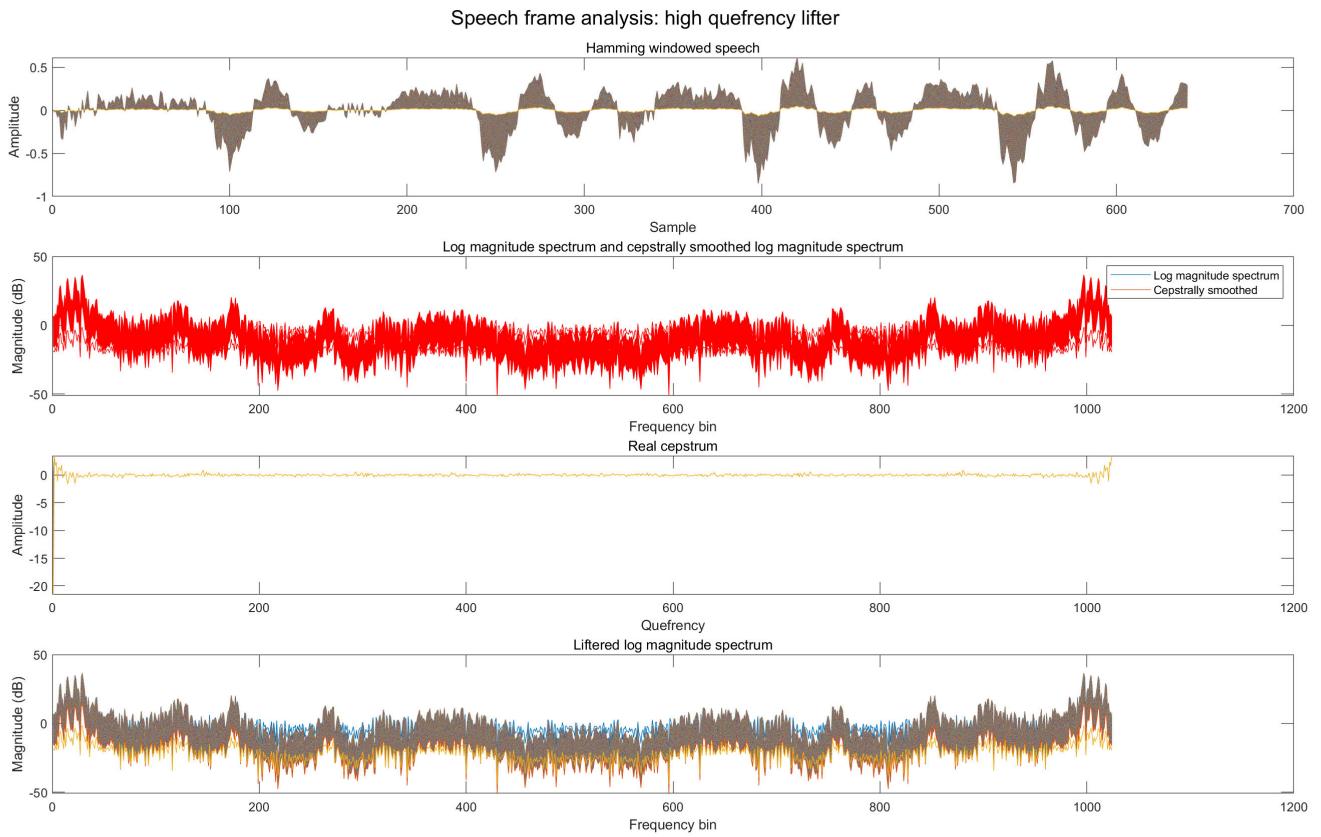


Speech frame analysis: high quefrency lifter



Speech frame analysis: low quefrency lifter





Analysis of the results:

Is the result in accord with the expected result?

If so, provide further analysis on the results. For example, provide some statistics, data, and other evidence to testify the effectiveness of your experiment.

If not, explain the unexpected the result and give possible reason for it.

结果和预期保持一致，显示窗函数（汉明窗）加权后的语音信号，它应该是原始语音帧的一个局部窗口，窗口两端的振幅会逐渐减小，低频滤波会使平滑后的谱更加突出低频部分。低频滤波后的谱图将会更加突出语音信号中的低频成分，反映声道特性；而高频滤波后的谱图则更强调信号的激励部分。

Task Three

Analysis of the Experimental Task

1.What does the question mean? What's the input and expected output of the experiment?

这个问题要求我们通过自相关法（Autocorrelation Method）计算语音信号的一帧的线性预测系数（LPC）。对于有声的稳态元音（如“ah.wav”）和无声音（如“test_16k.wav”），从中提取一个语音帧。提取该帧并使用汉明窗进行加窗，计算LPC系数。使用LPC分析法和

短时傅里叶变换（STFT）分析法分别绘制对数幅度谱，并将两者进行对比。

输入:

语音文件名：有声帧使用 ah.wav，无声帧使用 test_16k.wav

起始样本：对于两个帧，起始样本都为 3000

帧长：有声帧为 300 个样本，无声帧为 480 个样本

窗口类型：均为汉明窗

LPC 阶数：均为 12

输出:

LPC 对数幅度谱 STFT 对数幅度谱

2. Can you give theoretical analysis on the question?

倒谱是将语音信号进行傅里叶变换后，取其对数幅度谱并反变换得到的结果。倒谱的一个重要应用是分离源和传递响应，它通过区分低频和高频来提取语音的发声源和传递路径（声道特征）。

3. What's the expected result for the experiment?

LPC 对数幅度谱：通过 LPC 分析得到的对数幅度谱，能够展示语音信号的共振峰（代表声道的形状），其特点是相对平稳，适合描述发音道响应。

STFT 对数幅度谱：STFT 分析得到的对数幅度谱能够揭示语音信号的时频特性，包括突出的语音成分和不规则的噪声部分。

Details of Experiment

Configuration of the experiment, e.g. Length of FFT, sampling frequency.

帧长：有声帧 300 个样本（10kHz 采样率），无声帧 480 个样本（16kHz 采样率）

LPC 阶数：12

采样频率：10kHz（有声），16kHz（无声）

FFT 大小：用于 STFT 的 FFT 长度一般选择 1024（足够分辨频率成分）

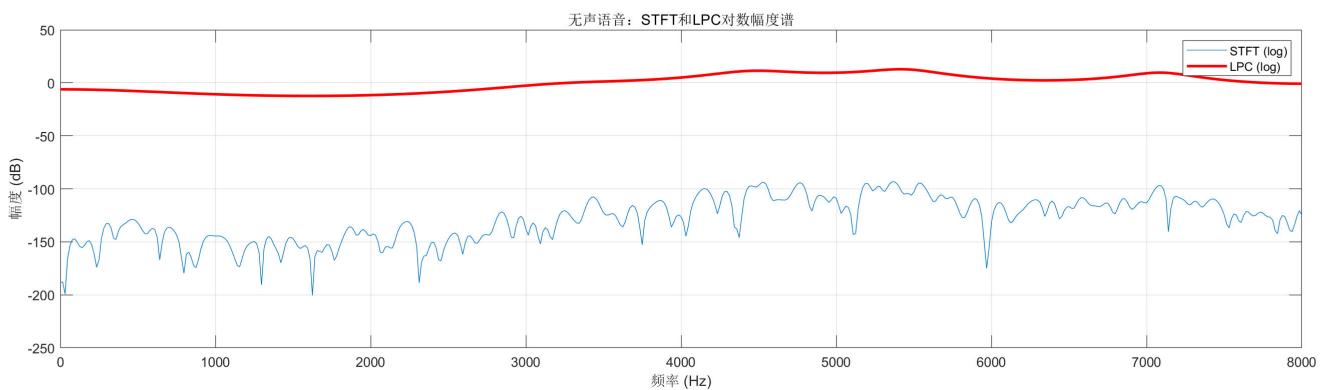
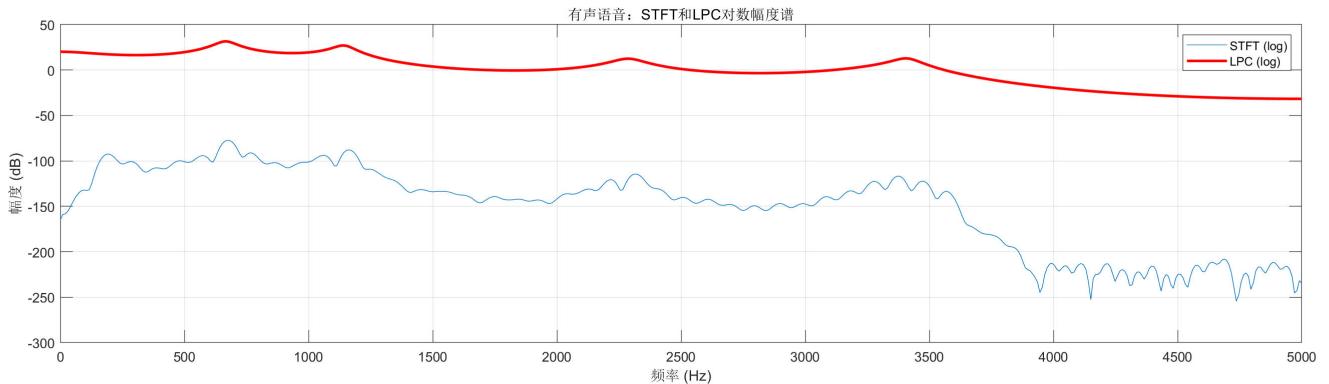
Step1 : 读取语音文件并根据给定的起始样本和帧长度提取语音帧。

Step2 : 计算 LPC 系数

Step3 : 对每帧应用汉明窗，计算 STFT

Step4 : 绘制 STFT 对数幅度谱和 LPC 对数幅度谱

Results of Experiment



Analysis of the results:

Is the result in accord with the expected result?

If so, provide further analysis on the results. For example, provide some statistics, data, and other evidence to testify the effectiveness of your experiment.

If not, explain the unexpected the result and give possible reason for it.

结果和预期保持一致，通过LPC分析得到的对数幅度谱，能够展示语音信号的共振峰（代表声道的形状），其特点是相对平稳，适合描述发音道响应。STFT对数幅度谱包括突出的语音成分和不规则的噪声部分。