

信息隐藏技术实验报告

Lab2 语音信号的处理

网络空间安全学院 信息安全专业

2112492 刘修铭 1028

一、题目

1. 学习慕课：2.2 语音信号处理基础
2. FFT
3. DWT
4. DCT

在 matlab 中调试完成课堂上的例题，练习使用常用的语音信号处理方法。

二、实验要求

编程实现，提交实验报告。

三、实验原理

(一) FFT

- 离散傅里叶变换

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi k \frac{n}{N}}$$
$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi k \frac{n}{N}}$$

- 把时域波形分解成许多不同频率的正弦波的叠加和
- 傅里叶变换反映的是整个时间段中信号的频谱特性，信号任何时刻的微小变化都会影响到整个频谱
- 任何有限频段上的信息都无法反映时域的某一个时间段的信号
- 不合适：时变的、非平稳信号
- 短时傅里叶变换：对于时变的、非平稳信号，用窗口取得某一时间段的信号，再分析其频谱。

$$STFT_x(t, f) = \int [x(t') g^*(t' - t)] e^{-j2\pi f t'} dt'$$

- 取较窄的时间窗，可以得到较好的时间分辨率，但是其频率分辨率降低

- 取较宽的时间窗，可以得到较高的频率分辨率，但是时域信号的细节特性就无法分辨
- 短时傅里叶变换无法同时满足时间分辨率和频率分辨率的要求

(二) DWT

小波变换（数学显微镜）：窗口大小（即窗口面积）固定但形状可变的时频局部化分析方法，引入了两个可变因子，平移与伸缩

$$CWT_x(\tau, a) = \frac{1}{\sqrt{|a|}} \int x(t) h^*\left(\frac{t - \tau}{a}\right) dt$$

$$x(t) = \frac{1}{C_H} \iint \frac{1}{a^2} CWT_x(\tau, a) \frac{1}{\sqrt{|a|}} h\left(\frac{t - \tau}{a}\right) da d\tau$$

$\sqrt{|a|}h\left(\frac{t - \tau}{a}\right)$ 为小波的平移与伸缩

- 当尺度 α 增加时，表示以伸展了的 $h(t)$ 的波形取观察整个信号 $x(t)$
- 当尺度 α 减小时，表示以缩窄了的 $h(t)$ 的波形取观察信号 $x(t)$ 的局部
- 平移因子 τ 则是将扩展和缩窄了的小波沿着时间轴平移，以达到对整个时间段的覆盖和分析
- 在低频部分具有较高的频率分辨率，较低的时间分辨率
- 在高频部分具有较高的时间分辨率，较低的频率分辨率
- 多分辨率分析的直观含义：人观察目标，设他所处的尺度为 j ，观察目标所获得的信息为 V_j
 - 当走近目标，即尺度增加到 $j + 1$ ，观察目标所获得的信息应该比尺度 j 下获得的信息更为丰富，即 $V_j \subset V_{j+1}$ 。尺度越大，距离越近，信息越丰富；

(三) DCT

离散余弦变换 (DCT)

- 正变换: $T(u) = \sum_{x=0}^{N-1} f(x)g(x, u) \quad u = 0, 1, 2, \dots, N - 1$
 - 正变换核: $g(x, 0) = \frac{1}{\sqrt{N}} \quad g(x, u) = \sqrt{\frac{2}{N}} \cos \frac{(2x+1)u\pi}{2N}$
- 反变换: $f(x) = \frac{1}{\sqrt{N}} C(0) + \sqrt{\frac{2}{N}} \sum_{u=1}^{N-1} C(u) \cos \frac{(2x+1)u\pi}{2N} \quad x = 0, 1, \dots, N - 1$

四、实验过程（含主要源代码）

开始实验前，本人录制了一段音频作为本次实验的语料数据。

(一) FFT

FFT 是快速傅里叶变换（Fast Fourier Transform）的缩写，是一种将信号从时域转换到频域的数学工具，可以将信号表示为一系列频率成分的组合。

按照慕课说明，代码实现如下：

wavread 函数已无法使用，转用 audioread 函数替代

```
1  clc;
2  clear all;
3  close all;
4
5  [x, fs] = audioread('./media/lxmliu2002.wav');
6  fx = fft(x); % fft 函数——快速傅里叶变换
7
8  subplot(2, 1, 1);
9  plot(x);
10 title('Original Signal');
11 subplot(2, 1, 2);
12 plot(abs(fftshift(fx))); % fftshift 函数——将频谱移到中心
13 title('Frequency Spectrum');
14 saveas(gcf, './pic/FFT.png');
```

(二) DWT

DWT 是离散小波变换（Discrete Wavelet Transform）的缩写，通过将信号分解成不同频率的小波组分，使得可以在不同分辨率下分析信号。

1. dwt

使用 'db4'（Daubechies 4）小波基对信号进行一级小波变换。其中，cd1 是细节分量，ca1 是近似分量，a0 是重构后的结果。

```
1  clc;
2  clear all;
3  close all;
4
5  [a, fs] = audioread('./media/lxmliu2002.wav');
6  plot(a)
7
8  [ca1, cd1] = dwt(a(:, 1), 'db4');
9  a0 = idwt(ca1, cd1, 'db4', length(a(:, 1)));
10 subplot(4, 1, 1);
11 plot(a(:, 1));
12 title('Original Signal');
13 subplot(4, 1, 2);
14 plot(cd1);
15 title('Detail Coefficients');
16 subplot(4, 1, 3);
17 plot(ca1);
18 title('Approximation Coefficients');
19 subplot(4, 1, 4);
20 plot(a0);
```

```

21 title('Reconstructed Signal');
22
23 saveas(gcf, './pic/DWT.png');

```

2. wavedec

使用'db4' (Daubechies 4) 小波基对信号进行三级小波变换。其中, cd1 是细节分量, ca1 是近似分量, a0 是重构后的结果。

```

1  clc;
2  clear all;
3  close all;
4
5  [a, fs] = audioread('./media/lxmliu2002.wav');
6  plot(a)
7  title('Original Signal')
8
9  [ca1, cd1] = wavedec(a(:, 1), 1, 'db4');
10 a0 = waverec(ca1, cd1, 'db4');
11 subplot(4, 1, 1);
12 plot(a(:, 1));
13 title('Original Signal')
14 subplot(4, 1, 2);
15 plot(cd1);
16 title('Detail Coefficients')
17 subplot(4, 1, 3);
18 plot(ca1);
19 title('Approximation Coefficients')
20 subplot(4, 1, 4);
21 plot(a0);
22 title('Reconstructed Signal')
23
24 saveas(gcf, './pic/wavedec.png');

```

3. wavedec_3

使用'db4' (Daubechies 4) 小波基对信号进行一级小波变换。其中, cd1 是一级分解细节分量, cd2 是二级分解细节分量, cd3 是三级分解细节分量, ca3 是三级分解近似分量, a0 是重构后的结果。

```

1  clc;
2  clear all;
3  close all;
4
5  [a, fs] = audioread('./media/lxmliu2002.wav');
6  plot(a)
7
8  [c, l]=wavedec(a(:,2),3 , 'db4') ;
9  ca3=appcoef(c,l,'db4',3) ;
10 cd3=detcoef(c,l,3) ;
11 cd2=detcoef(c,l,2) ;

```

```

12 cd1=detcoef(c,1,1) ;
13 a0=waverec(c,1,'db4') ;
14
15 subplot(6, 1, 1);
16 plot(a(:, 2));
17 title('Original signal');
18 subplot(6, 1, 2);
19 plot(cd1);
20 title('Detail Coefficient 1');
21 subplot(6, 1, 3);
22 plot(cd2);
23 title('Detail Coefficient 2');
24 subplot(6, 1, 4);
25 plot(cd3);
26 title('Detail Coefficient 3');
27 subplot(6, 1, 5);
28 plot(ca3);
29 title('Approximation Coefficient 3');
30 subplot(6, 1, 6);
31 plot(a0);
32 title('Reconstructed signal');
33
34 saveas(gcf, './pic/wavedec_3.png');

```

(三) DCT

DCT 是离散余弦变换 (Discrete Cosine Transform) 的缩写，将时域信号转换为频域信号。

```

1 clc;
2 clear all;
3 close all;
4
5 [a, fs] = audioread('./media/lxmliu2002.wav');
6 plot(a)
7 title('Original signal');
8
9 a1 = dct(a(:, 1));
10 a0 = idct(a1);
11 subplot(3, 1, 1);
12 plot(a(:, 1));
13 title('Original signal');
14 subplot(3, 1, 2);
15 plot(a1);
16 title('DCT Coefficients');
17 subplot(3, 1, 3);
18 plot(a0);
19 title('Reconstructed signal');
20
21 saveas(gcf, './pic/DCT.png');

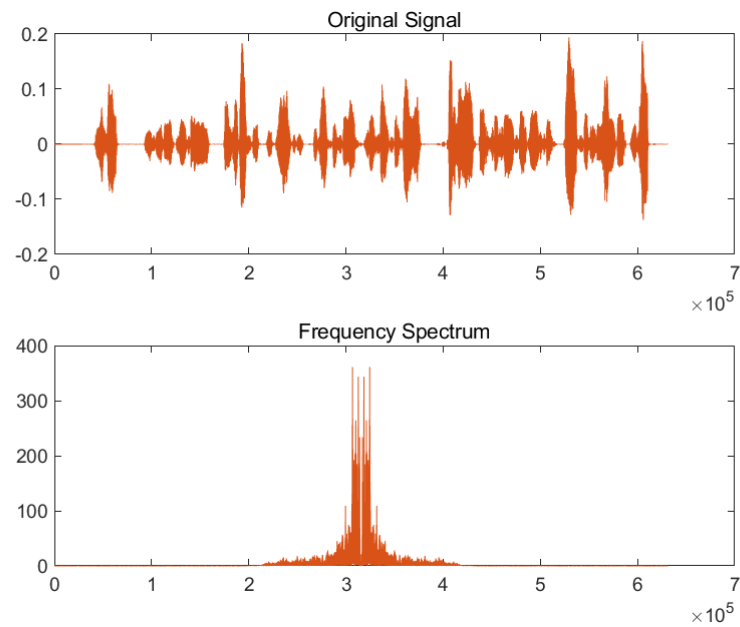
```

五、实验结果及分析

(一) FFT

如图，可以看到，经 FFT 处理后，语音信号的能量主要集中在 30kHz ~ 33kHz 附近。

1. 原始语音信号
2. FFT 变换信号

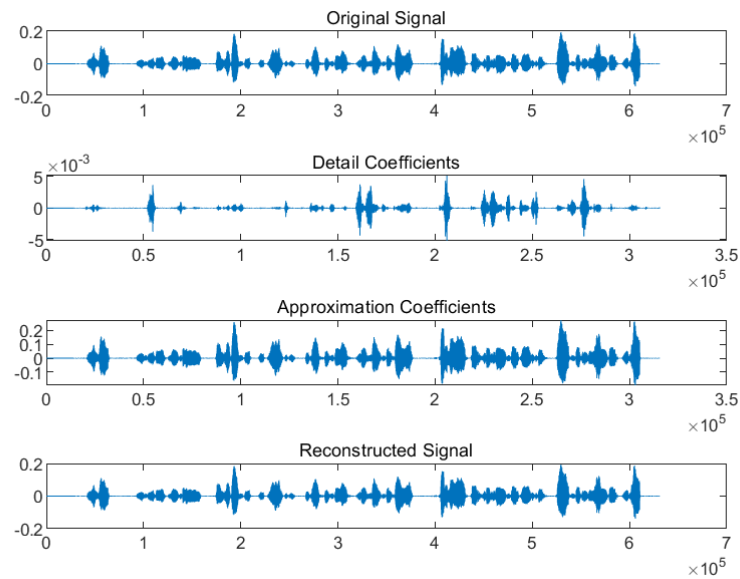


(二) DWT

1. DWT

运行程序，得到如下的输出。

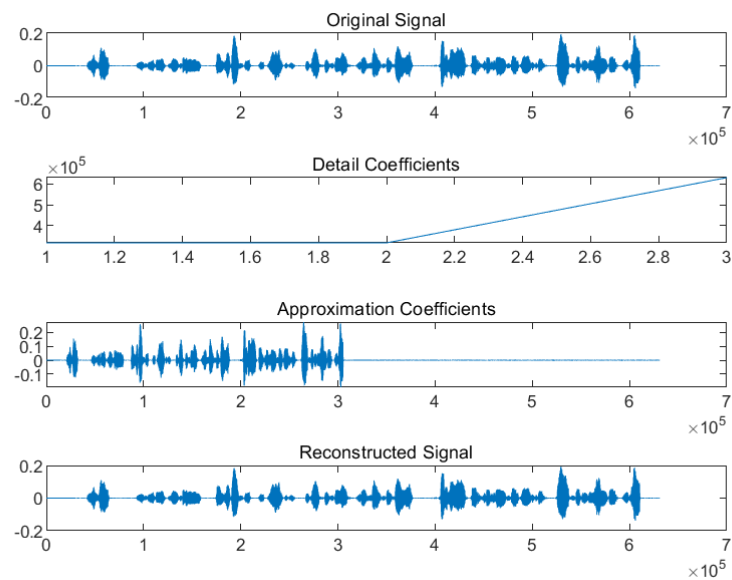
1. 原始语音信号
2. 一级分解的细节分量
3. 一级分解的近似分量
4. 重构信号



2. Wavedec

运行程序，得到如下的输出。

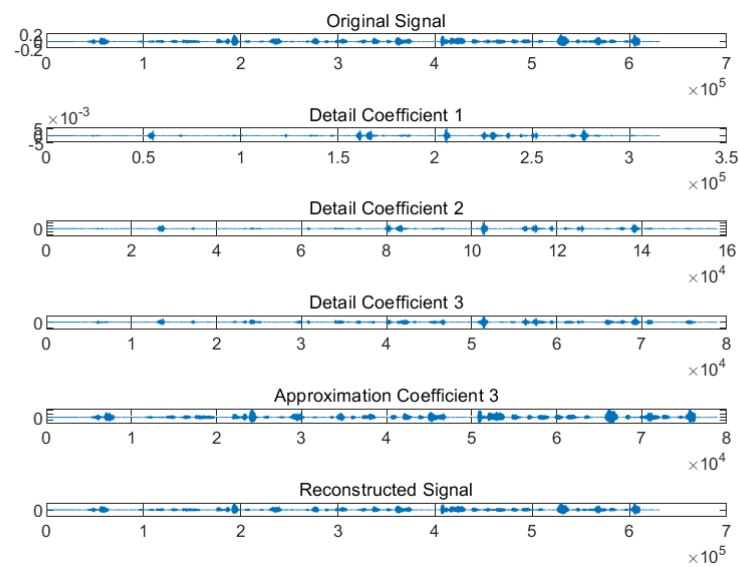
1. 原始语音信号
2. 一级分解的细节分量
3. 一级分解的近似分量，分解后数据的长度减半
4. 重构信号



3. Wavedec_3

运行程序，得到如下的输出。

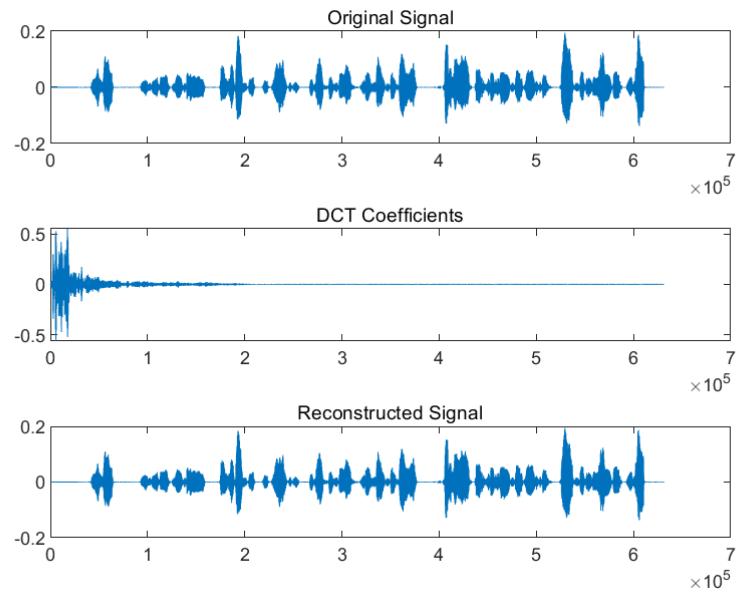
1. 原始语音信号
2. 一级分解的细节分量
3. 二级分解的细节分量
4. 三级分解的细节分量
5. 按级分解的近似分量
6. 重构信号



(三) DCT

运行程序，得到如下的输出。

1. 原始语音信号
2. DCT 变换信号
3. 重构信号



六、参考

本次实验主要参考慕课完成。

七、说明

本次实验所有代码均已放在 `codes` 文件夹下。

```
1 FFT.m
2 DWT.m
3 wavedec.m
4 wavedec_3.m
5 DCT.m
```

本人录制的音频为 `codes/media/lxmliu2002.wav`，本次实验所有图片均位于 `codes/pic` 文件夹