



语音信号处理基础



语音信号处理基础

语音编码的分类

波形编码

参数编码

常用处理算法

傅立叶变换与短时傅立叶变换

小波变换

离散余弦变换 (DCT)

文A

1. 语音编码分为两大类(1)

波形
编码

力图使重建的语音波形保持原语音信号的波形形状。

特点

话音质量好，但编码速率比较高 (64—16kb/s)。

文A

1. 语音编码分为两大类(2)

参数 编码

通过对语音信号特征参数的提取并编码，力图使重建的语音信号具有较高的可懂度，而重建的语音信号波形与原始语音波形可以有很大的差别。

特点

编码速率低，它可以达到2.4kb/s甚至更低，能够达到听懂话音，但是其主要问题是语音的自然度较低。

1. 常见的波形编码

波形
编码

脉冲编码调制 (PCM)

自适应增量调制 (ADM)

自适应差分编码 (ADPCM)

自适应预测编码 (APC)

自适应子带编码 (ASBC)

自适应变换编码 (ATC)

1. 参数编码

参数
编码

线性预测编码 (LPC) 等

文A

填空题 2分

语音信号的编码方式可以分为两大类：[填空1]和[填空2]，前者力图使重建的语音波形保持原语音信号的波形形状。

文A

2. 常用语音处理算法



傅立叶变换与短时傅立叶变换



小波变换



离散余弦变换



离散傅立叶变换

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j 2 \pi k n / N}$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j 2 \pi k n / N}$$



离散傅立叶变换的特点

- 01 把时域波形分解成许多不同频率的正弦波的叠加和
- 02 傅立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。
- 03 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
- 04 不适合：时变的、非平稳信号



短时傅立叶变换 (STFT)

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

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

$g^*(t)$ 为窗函数。

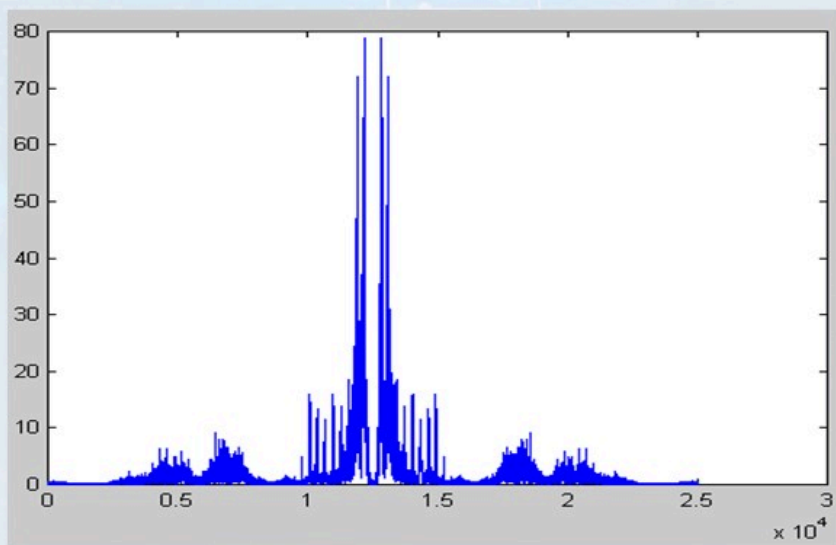


短时傅立叶变换的特点

- 01 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。
- 02 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。
- 03 短时傅立叶变换无法同时满足时间分辨率和频率分辨率的要求。



演示：快速傅里叶变换



声音信号的傅氏变换

```
[x,fs,nbits]=wavread('xxxx.wav');  
fx=fft(x);  
%fft函数--快速离散傅里叶变换  
plot(abs(fftshift(fx)));  
%fftshift函数——调整fft函数的输出顺序，  
将零频位置移到频谱的中心
```

文A



小波变换(Wavelet Transform)

小波分析法

是一种窗口大小（即窗口面积）固定但其形状可变的时—频局部化分析方法。

01

在低频部分具有较高的频率分辨率，较低的时间分辨率

02

在高频部分具有较高的时间分辨率，较低的频率分辨率

这正符合低频信号变化缓慢而高频信号变化迅速的特点，因此，被称为数学显微镜。

文A



小波变换(Wavelet Transform)

引入了两个可变因子：平移，伸缩

连续小波变换

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

$h(t)$ 为小波母函数

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



小波变换(Wavelet Transform)

$$\frac{1}{\sqrt{|a|}} h\left(\frac{t - \tau}{a}\right)$$

为小波的平移与伸缩

- 01 当尺度 a 增加时，表示以伸展了的 $h(t)$ 波形去观察整个信号 $x(t)$
- 02 当尺度 a 减小时，表示以缩窄了的 $h(t)$ 波形去观察信号 $x(t)$ 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。

3. 多分辨分析的直观含义



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

当他走近目标，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；

反之，如果远离目标，即尺度减小到 $j-1$ ，则观察目标所获得的信息比尺度 j 下获得的信息更为减少。尺度越小，距离越远，信息越稀少。



几种常用的小波函数

Haar小波

$$\psi(x) = \begin{cases} 1 & 0 \leq x < 1/2 \\ -1 & 1/2 \leq x < 1 \\ 0 & \text{其它} \end{cases}$$

Daubechies小波系

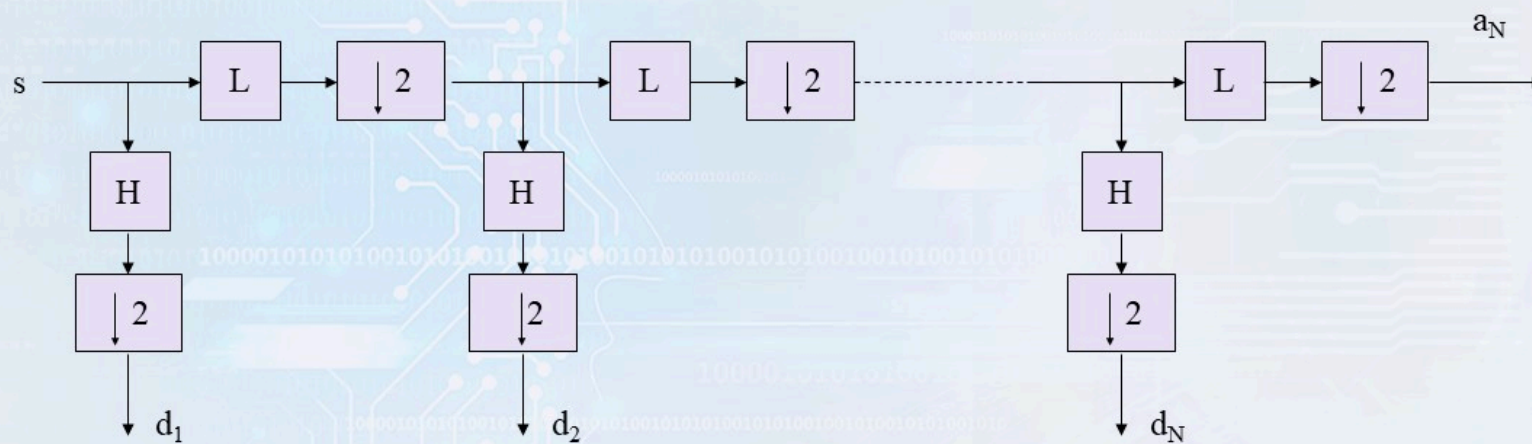
db1-db10

db1为Haar小波

其他，没有解析式



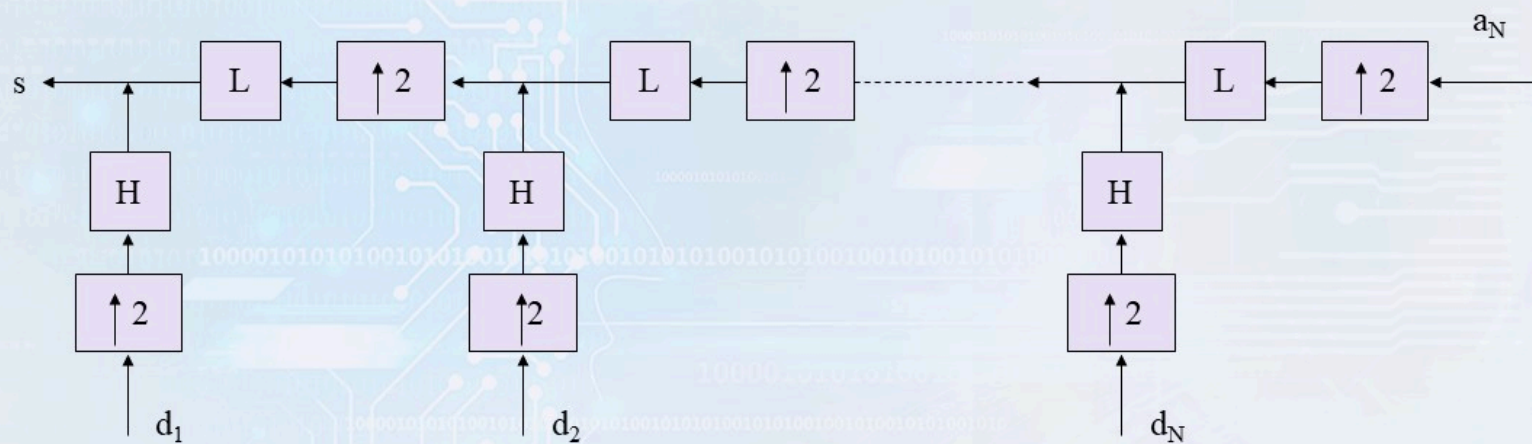
小波分解的实现



文A



小波重构的实现

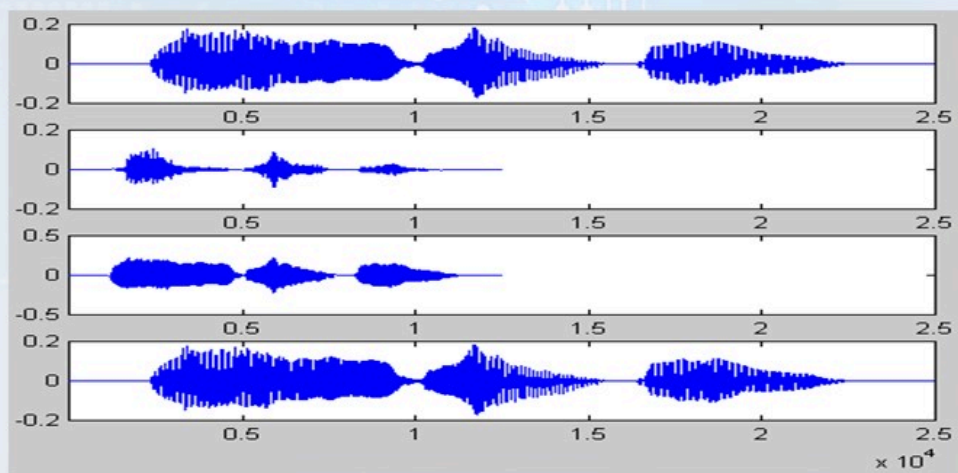


文A



语音信号的小波变换

(一级分解)

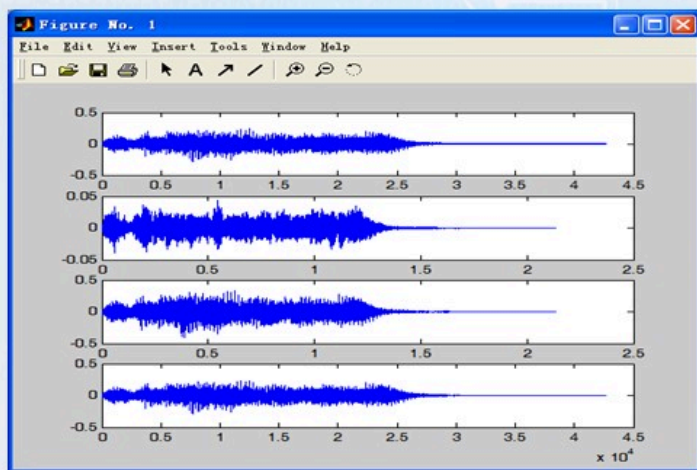


```
[a,fs,nbits]=wavread('myg.wav');  
[ca1,cd1]=dwt(a(:,1),'db4');  
a0=idwt(ca1,cd1,'db4',length(a(:,1)));  
subplot(4,1,1);plot(a(:,1));  
subplot(4,1,2);plot(cd1);  
subplot(4,1,3);plot(ca1);  
subplot(4,1,4);plot(a0);
```

文A



语音信号的小波变换



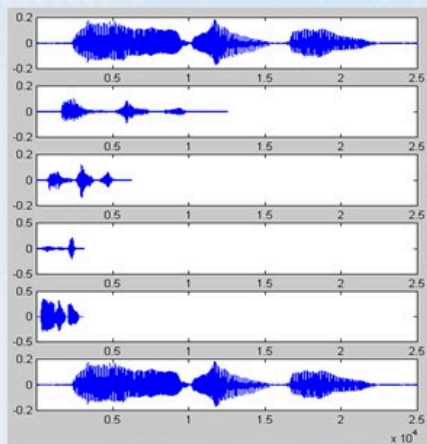
```
[a,fs,nbits]=wavread('tada.wav');  
[ca1,cd1]=dwt(a(:,1),'db4');%ca1近似, cd1细节  
a0=idwt(ca1,cd1,'db4',length(a(:,1)));  
subplot(4,1,1);plot(a(:,1));  
subplot(4,1,2);plot(cd1);  
subplot(4,1,3);plot(ca1);  
subplot(4,1,4);plot(a0);
```

文A



语音信号的小波变换

(三级分解)



```
[a,fs,nbits]=wavread('myg.wav');  
[c,l]=wavedec(a(:,2),3,'db4');  
ca3=appcoef(c,l,'db4',3);%提取第3层的近似系数  
cd3=detcoef(c,l,3);%提取第3层的细节系数  
cd2=detcoef(c,l,2);  
cd1=detcoef(c,l,1);  
a0=waverec(c,l,'db4');
```

```
subplot(6,1,1);plot(a(:,2));  
subplot(6,1,2);plot(cd1);  
subplot(6,1,3);plot(cd2);  
subplot(6,1,4);plot(cd3);  
subplot(6,1,5);plot(ca3);  
subplot(6,1,6);plot(a0);
```




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

$$g(x,0) = \frac{1}{\sqrt{N}} \quad g(x,u) = \sqrt{\frac{2}{N}} \cos \frac{(2x+1)u\pi}{2N}$$

文A



离散余弦变换 (DCT)

DCT变换的系数

$$C(0) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)$$

$$C(u) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x) \cos \frac{(2x+1)u\pi}{2N}$$

$$u = 1, 2, \dots, N-1$$



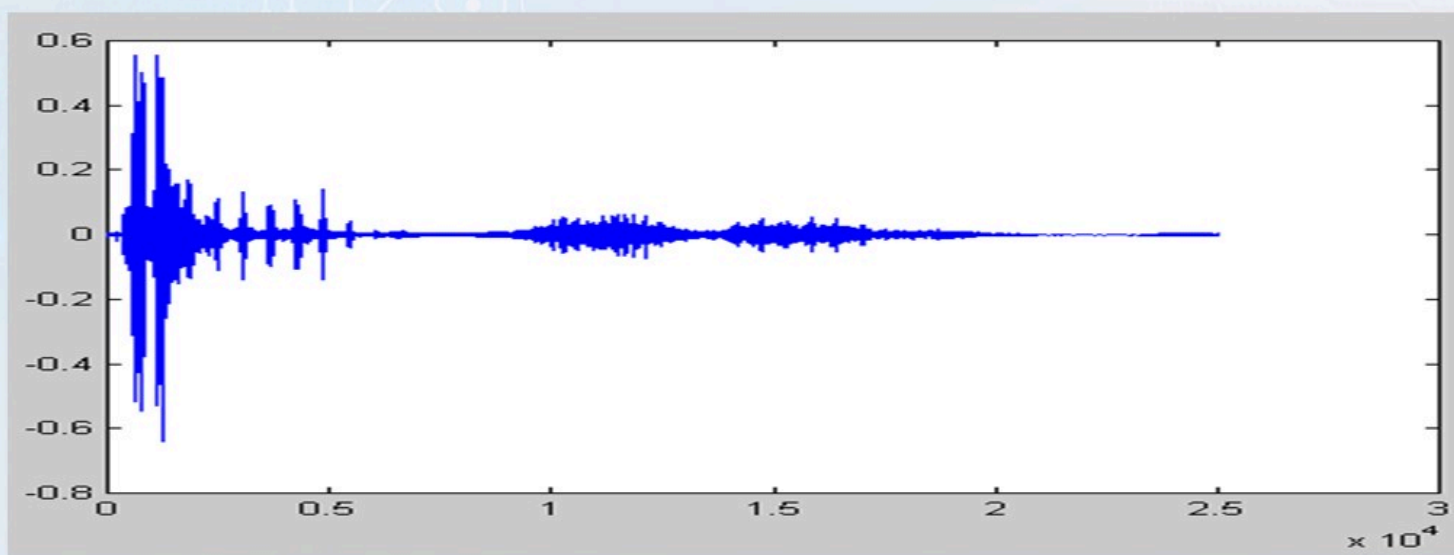
离散余弦变换 (DCT)

反变换

$$f(x) = \frac{1}{\sqrt{N}} C(0) + \sqrt{\frac{2}{N}} \sum_{u=1}^{N-1} C(u) \cos \frac{(2x+1)u\pi}{2N}$$

$$x = 0, 1, \dots, N-1$$

语音信号的DCT变换





小结

几种变换

从时域信号
转换到变换域（频域、
时-频域、小波域等）

从压缩角度

能量集中，
有利于压缩编码

从信息隐藏角度

考虑隐藏的位置

文A

填空题 1分

[填空1] 分析法是一种窗口大小（即窗口面积）固定但其形状可变的时——频局部化分析方法。

文A

《信息隐藏技术》

实验2：语音信号的常用处理方法上机实验

内容：

- 1、FFT
- 2、DWT
- 3、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。

要求：

编程实现，提交实验报告。

提交方法：qq群作业。

时间：2024-3-19