



数字信号处理基础知识

雷敏

北京邮电大学 网络空间安全学院

leimin@bupt.edu.cn

考核

成绩评分说明

第一、平时成绩。课程一共分为四次实验，每次实验报告满分为**25分**，总分为**100分**。每次课程课堂互动会随机抽点一部分学生，如果课堂互动抽点问题未到课，每次扣除平时成绩**5分**。平时成绩占总成绩的**60%**。

第二、期末会安排一个期末综合大作业，大作业满分**100分**，综合大作业占总成绩的**40%**。综合大作业大作业包含填空题，判断题，名词解释题和综合设计题，有详细的评分标准。其中动手实践的操作题目占**60%**，基本知识题目占**40%**，要求同学们在随堂**90分钟**内完成。

第三、最终本课程的总成绩为平时成绩***0.6**+期末大作业成绩***0.4**。

实验报告成绩

- 1、是否按照要求完成实验任务，并按时提交实验报告（总分**15**分）。每次实验报告要求是下周课程上课前一天**24**点之前在教学云平台上提交。凡逾期提交实验报告扣**3**分。
- 2、实验报告文档的规范性，包括格式和内容的规范性（总分**4**分）。如果实验报告的名称命名不规范，扣**1**分，实验报告命名规范要求是：学号+姓名+第一次实验报告；学号+姓名+第二次实验报告；学号+姓名+第三次实验报告；学号+姓名+第四次实验报告；如果实验报告的封面不规范扣**1**分，实验报告的封面必须规范，要求是学号+姓名，不是姓名+学号；实验报告封面上实验的名称必须正确，实验的名称严格要求实验指导书中的名称；如果实验内容的格式不规范，扣**1**分；

实验报告成绩

3、实验遇到的问题及解决方法：注重记录自己本人完成实验过程中遇到的问题及解决方法，每个同学在实验过程中遇到的问题不一定完全相同，请记录本人在实验过程中遇到的问题及解决这些问题的方法。此部分总分为3分。

4、实验心得体会和思考：本人通过完成本次实验有哪些心得体会和思考，通过本次实验得到的收获，此部分总分为3分。

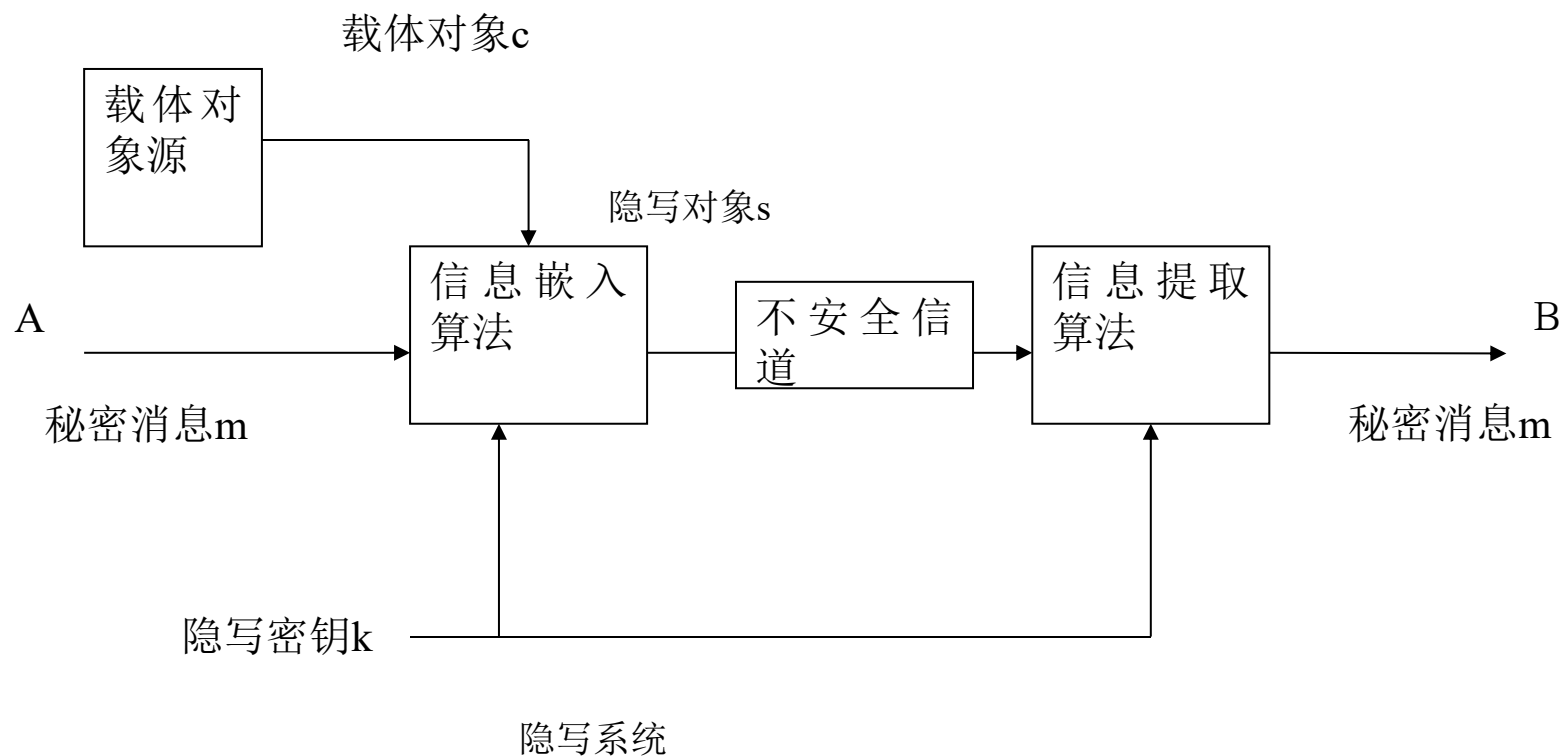
5、实验报告中的实验遇到的问题及解决方法，实验心得体会和思考必须是本人的，如果不是，需要扣分；或者实验心得体会和思考，实验遇到的问题及解决方法写的过于简单也会扣分。

每次实验报告的成绩会在云平台上给出来，每次实验报告在报告结束后一周内会给出，也就是在提交报告后一周内给出成绩，同时还会给出扣分原因，便于后续改进。如果满分，无扣分原因。

注意事项

- 1、请在QQ群中按照要求标注自己的身份，便于助教核对是否所有选课的同学都在QQ群中的信息。
- 2、有两个助教，王壮博士负责信息安全专业（含双培生），张力元硕士负责网络空间安全专业（含实验班）
- 3、期中会有一个调查问卷，麻烦各位填写一些。
- 4、去年本课程的期末学生评教为全院第五名（49门课程），我们今年也共同努力，希望各位能有所收获。
- 5、期末大作业计划在云平台上完成，到时候可能会提前有一个测试，请各位到时候配合一下。预计会在课堂完成。

隐写系统



隐写算法性能指标

- 容量（Capacity）
 - 负载（Payload），载体数据利用率，嵌入效率（Embedding Efficiency）
- 不可感知性（Imperceptibility）
 - 透明性（Transparency）保真性（Fidelity）
- 稳健性（Robustness）
 - 鲁棒性、健壮性
- 安全性（Security）
 - 统计不可检测性（Statistical Undetectability）

隐写算法性能指标

○ 容量

- 隐写算法容量指，算法在载体对象中能嵌入的消息总数。
- 容量也常用平均每样点能嵌入的信息比特来衡量算法容量，即，载体数据利用率=嵌入消息总比特数/样点总数，单位为比特每样点(bit per sample, bps)。对于图像，样点即像素，量纲可写为bpp(bit per pixel)。

○ Capacity

- This refers **to the number of bits** of information that are embedded in the host signal.
- The payload is often **normalized by the number of samples** of the host signal, resulting in a **bit rate R per sample** of the host.

隐写算法性能指标

○ Transparency (Fidelity, Imperceptibility)

- In most applications, embedding of information **should not cause perceptual degradation** of the host signal.
- 透明性（保真性，不可感知性）
 - 透明性指算法对载体对象感官质量的影响程度。
 - 通常，算法引入的失真应该是不可感知的。

隐写算法性能指标

○ 透明性（保真性，不可感知性）

- 实例，以峰值信噪比衡量透明性，

$$PSNR = XY \max_{x,y} \frac{p_{x,y}^2}{\sum_{x,y} (p_{x,y} - \tilde{p}_{x,y})^2}$$



（替换最低比特平面）PSNR=51dB



（替换第7比特平面）PSNR=14dB

隐写算法性能指标

○ Robustness

- This refers to the ability of the embedding algorithm to **survive common signal processing** operations.
- 稳健性（鲁棒性，健壮性）
 - 稳健性指算法抵抗常规信号处理操作的能力。
 - 隐写对象经滤波等操作处理后，会产生失真。如果信息提取算法仍能从这样的载体中提取消息，那么就称算法对滤波等操作稳健。

隐写算法性能指标

○ 稳健性（鲁棒性、健壮性）

略过深入讨论

- 案例：**LSB**算法对噪声添加处理不稳健。



LSB
psnr=50



误码率
为0

高斯
白噪
psnr=30

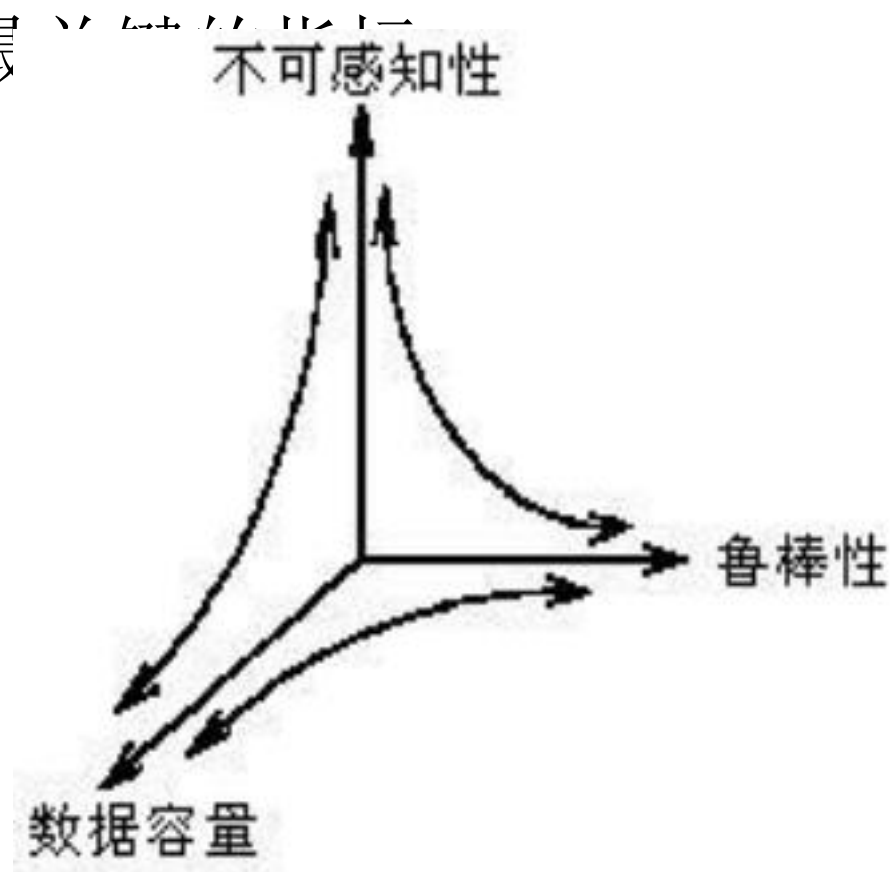


误码率
为0.5

隐写算法性能指标

○ 隐写算法三个最

- 不可感知性
- 鲁棒性
- 容量



音频信号处理基础

- 数字音频的概念
- 人类听觉系统(HAS)特点
- 语音质量评价
- 语音信号处理基础

音频基础——数字音频概念

○ 采样频率

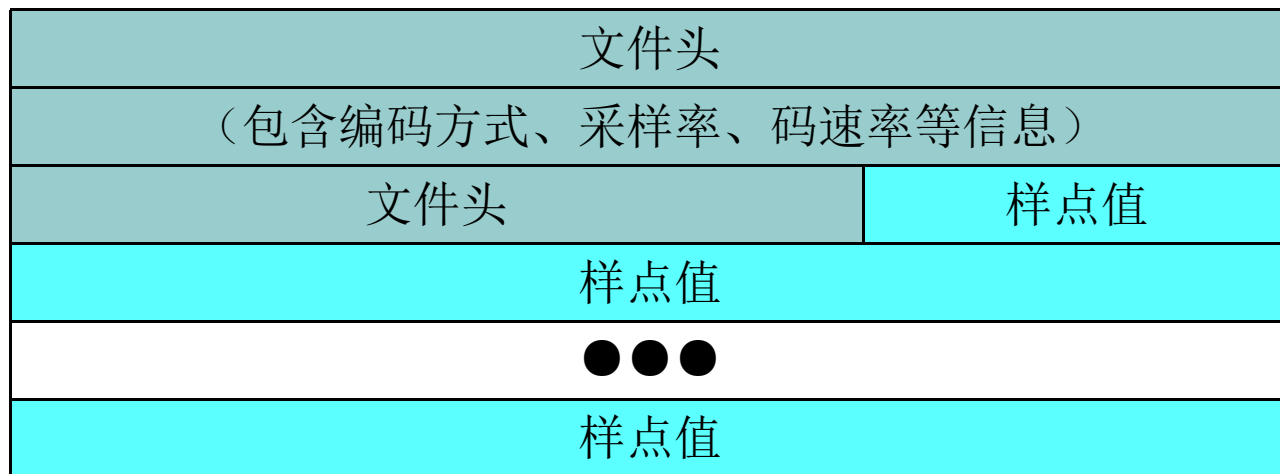
- “观察”信号的频度，单位为Hz。

○ 码速率

- 我们需要对采样所得信息进行编码，产生的码流单位为b/s(bit per second)。
- PCM码速率：64K、128K、256K b/s

音频基础——数字音频文件

- 音频文件格式
 - WAV文件格式



音频基础——数字音频文件

○ 音频文件实例

- WAV文件: s1.wav (Fs:8KHz, Nbits=16)



音频基础——数字音频文件

○ 音频文件实例

- WAV文件: s1.wav (Fs:8KHz, Nbits=16)

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f	
00000000h:	52	49	46	46	24	E8	07	00	57	41	56	45	66	6D	74	20	; RIFF\$?.WAVEfmt
00000010h:	10	00	00	00	01	00	01	00	40	1F	00	00	80	3E	00	00	;@...<.>.
00000020h:	02	00	10	00	64	61	74	61	00	E8	07	00	F4	FE	B2	FE	;data.?.酷猫
00000030h:	1B	FE	94	FD	CA	FD	2B	FE	BC	FE	82	FF	50	FF	66	FF	; .鳍 ? 鳍 P

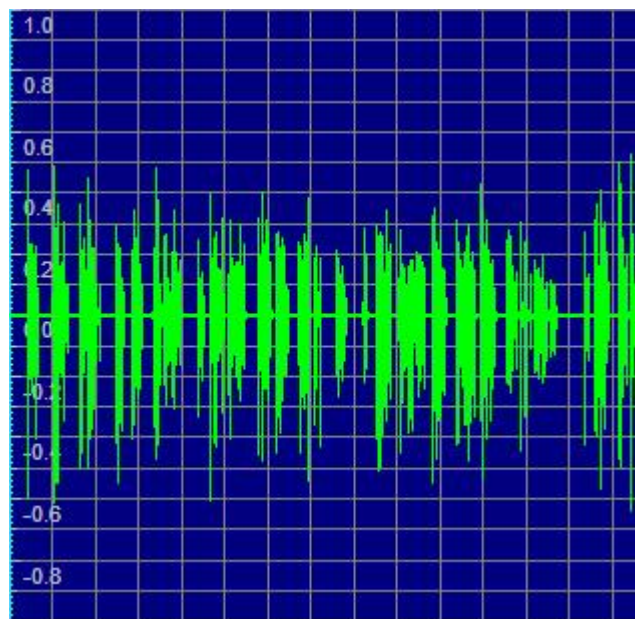
前44个字节为音频文件头

第45、46个字节为第一个
样点量化幅值

音频基础——单双声道

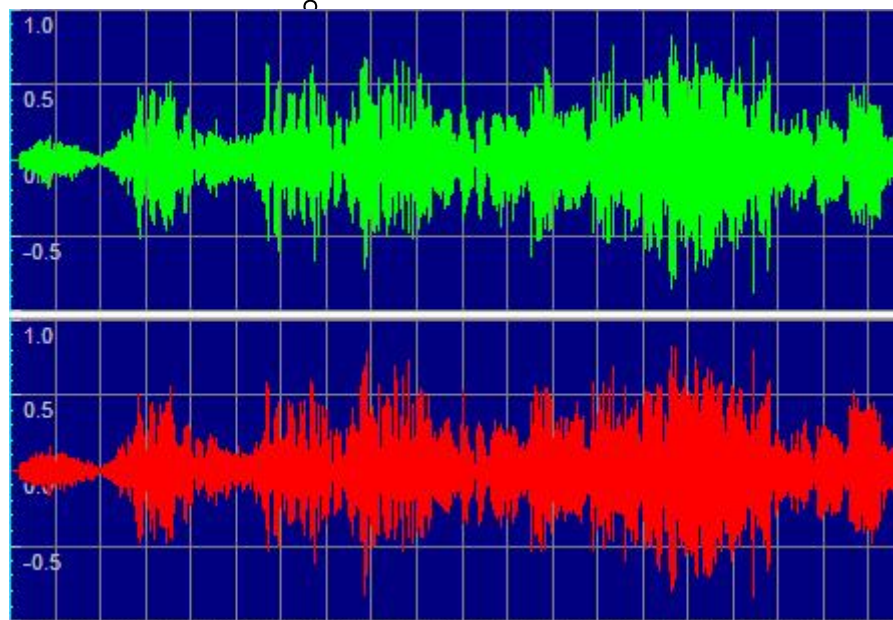
○ 单声道

- 一个声音通道



○ 双声道

- 二个声音通道，能够形成立体声效果



人耳听觉感知特性

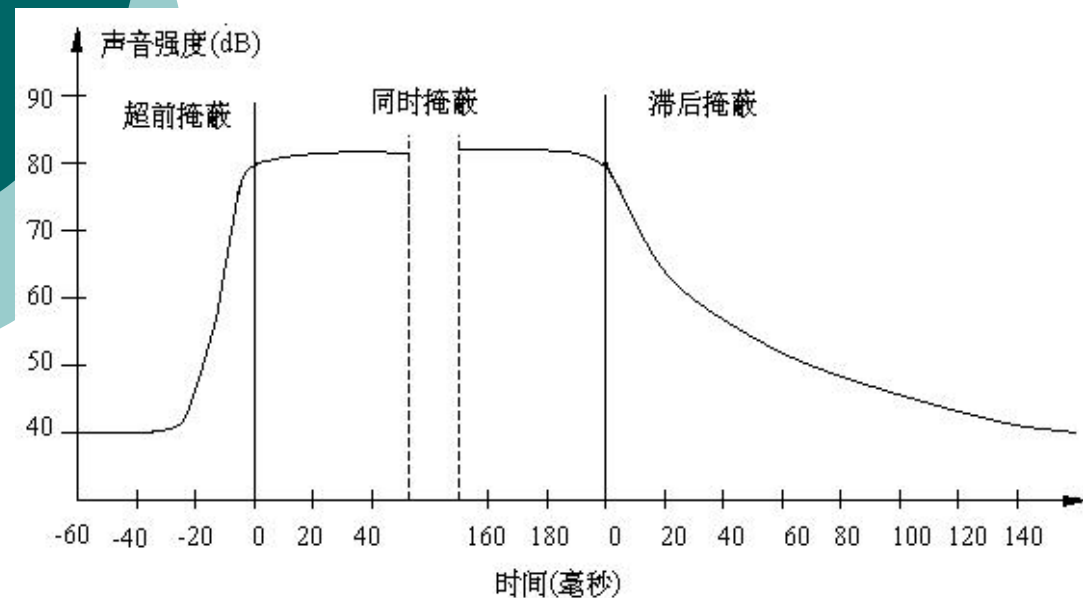
- 人类能感知声音的频率范围 **16Hz-16kHz**
- 年轻人, **20kHz**
- 老年人, **10kHz**
- 案例: 公园“驱逐器”

人耳听觉感知特性——掩蔽效应

○ 掩蔽效应

- 心理声学实验表明：人耳难以感知位于强信号附近的弱信号。
- 这种声音心理学现象称为掩蔽。
- 强信号称之为掩蔽音，弱信号称之为被掩蔽音
- 掩蔽音和被掩蔽音同时出现所产生的掩蔽效应称为同时掩蔽或频域掩蔽，
- 否则称之为异时掩蔽或时域掩蔽。

人耳听觉感知特性——掩蔽效应



- 同时（频域）掩蔽
 - 低音容易掩蔽高音，而高音掩蔽低音较难。
- 异时（时域）掩蔽
 - 超前掩蔽(pre-masking)
 - 滞后掩蔽(post-masking)

语音的质量评价

○ 主观评价

- 主观评价由人对语音的质量进行评价。

○ 客观评价

- 客观评价不以人为主体的。
- 通常提取处理前后的语音的某类特征参量，根据某种规则计算这两组参量之间的差别，是一种“误差”度量。

主观评价方法

- 平均意见分（MOS: Mean Opinion Score）
 - 用5级评分标准来评价语音的质量。

MOS	质量标准	受损程度
5	极好	不可察觉
4	较好	可察觉，但不影响听觉效果
3	一般	轻微影响听觉效果
2	较差	影响听觉效果
1	极差	严重影响听觉效果

- 实验者的分别打分评价语音质量。
- 全体实验者的平均分就是所测语音质量的MOS分。

MOS分评价

○ MOS分

- 4.0-4.5（网络质量）：高质量数字化语音。
- 3.5分左右（通信质量）：能感觉到语音质量有所下降，但不妨碍正常通话。
- 3.0分以下（合成语音质量）：具有足够高的可懂度，但自然度不够好，并且不易进行讲话人识别。

主观评价的优缺点

○ 优点：

- 真实：反映了人对语音质量的真实感觉。

○ 缺点：

- 需要组织大量的试听人员，欠缺灵活性。
- 评分结果重复性和稳定性较低。
- 评分结果受试听者的主观影响较大。

客观评价方法

○ 比较波形的差别：

- 基于**SNR**的评价方法：信噪比、分段信噪比、变频分段信噪比。

$$SNR = 10 \times \lg_{10} \left(\frac{\sum_{i=0}^{N-1} x^2(i)}{\sum_{i=0}^{N-1} (y(i) - x(i))^2} \right)$$

客观评价与主观评价一致性的问题

- 研究方便、快捷、准确的，与主观评价相一致的客观评价标准，是该领域研究的一个方向。
- 如何衡量一个客观评价与主观评价的一致性？

客观评价与主观评价一致性的问题

- 通过在客观评价和主观评价之间建立的函数关系，可以用客观评价值求出对主观评价值的预测值。
- 该预测值和实测的主观评价值之间的相关度就作为该客观评价方法与主观评价方法之间的相关度。

客观评价的优缺点

○ 优点:

- 简单，可重复性强，与试听者主观感觉无关。

○ 缺点:

- 无法达到与主观评价完全一致的效果。

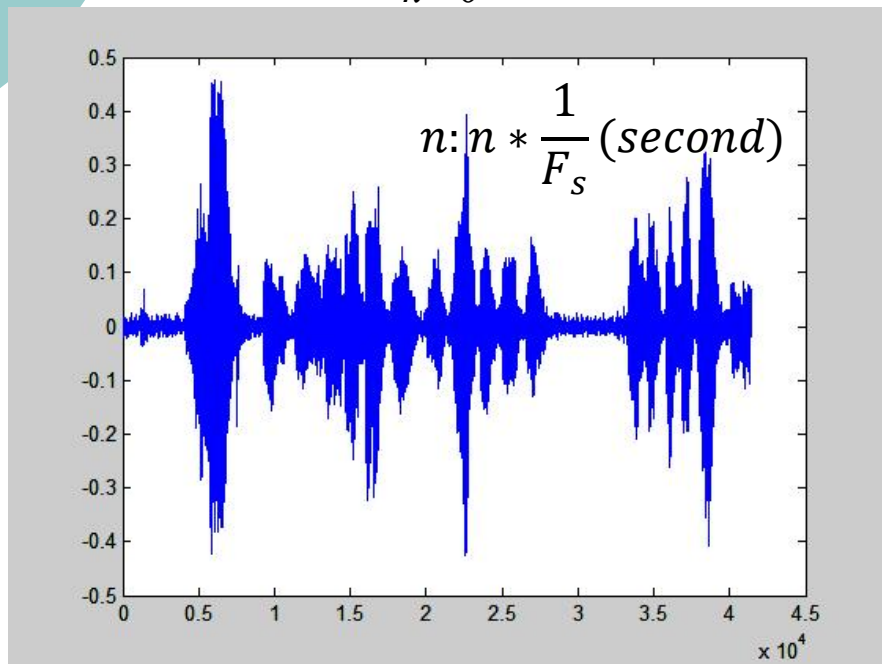
常用语音信号变换

- 常用变换域操作
 - DFT: 离散傅里叶变换
 - DCT: 离散余弦变换
 - DWT: 离散小波变换

常用语音信号变换

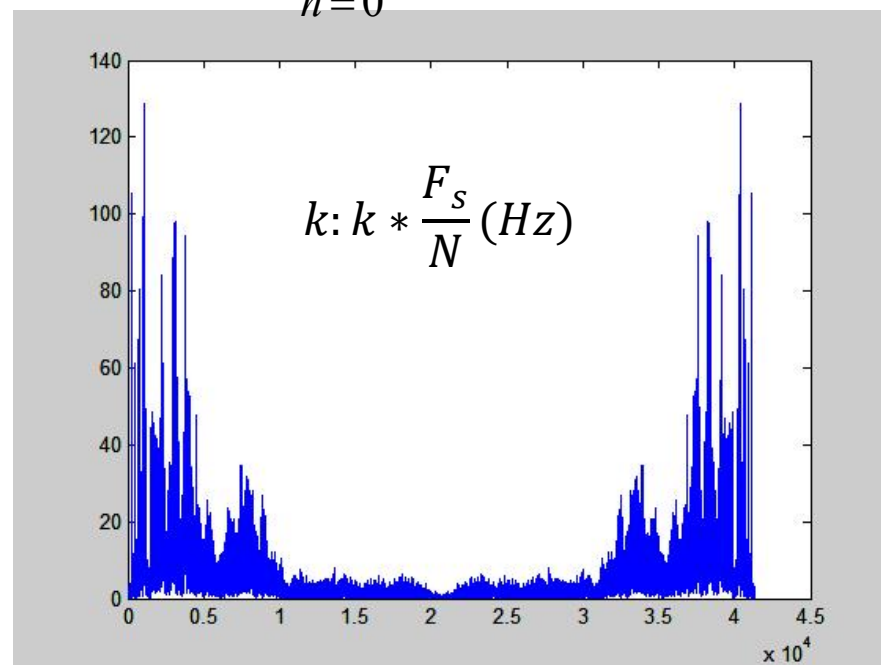
音频（时域）

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



音频（**DFT**）

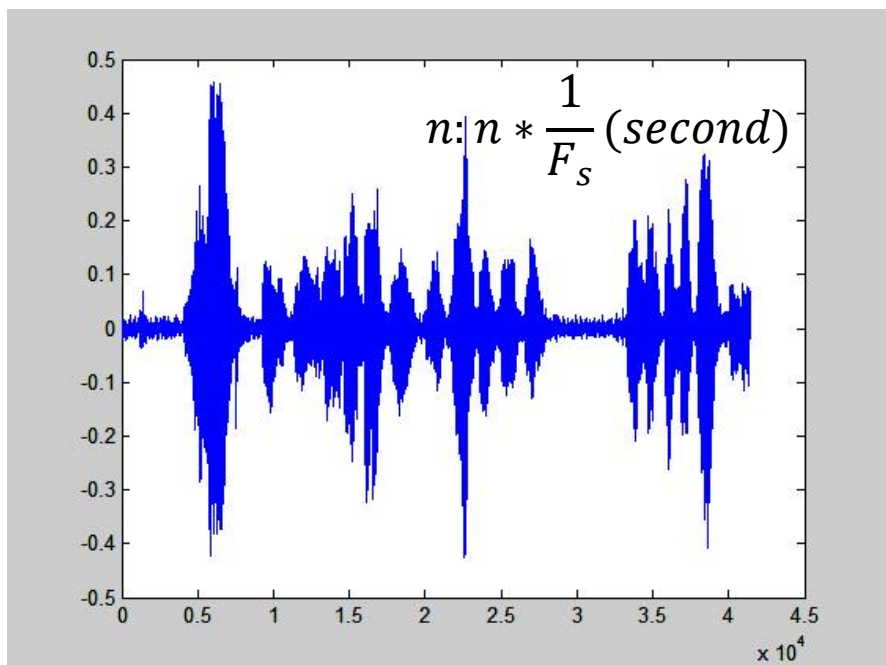
$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}$$



常用语音信号变换

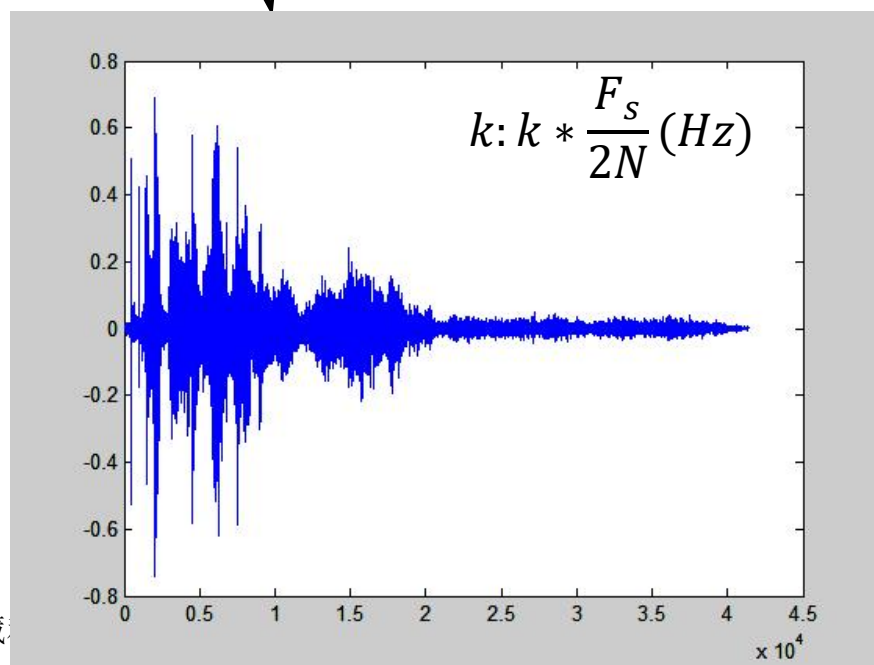
音频（时域）

- $x[n] = \frac{1}{\sqrt{N}} X[0] + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} X[k] \cos \frac{(2n+1)k\pi}{2N}$



音频（**DCT**）

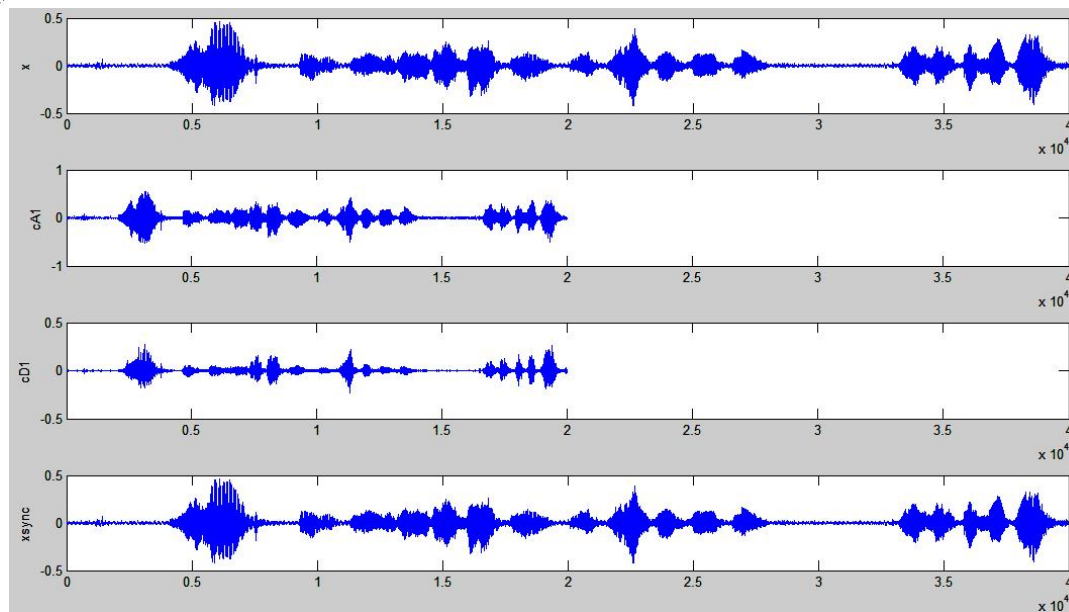
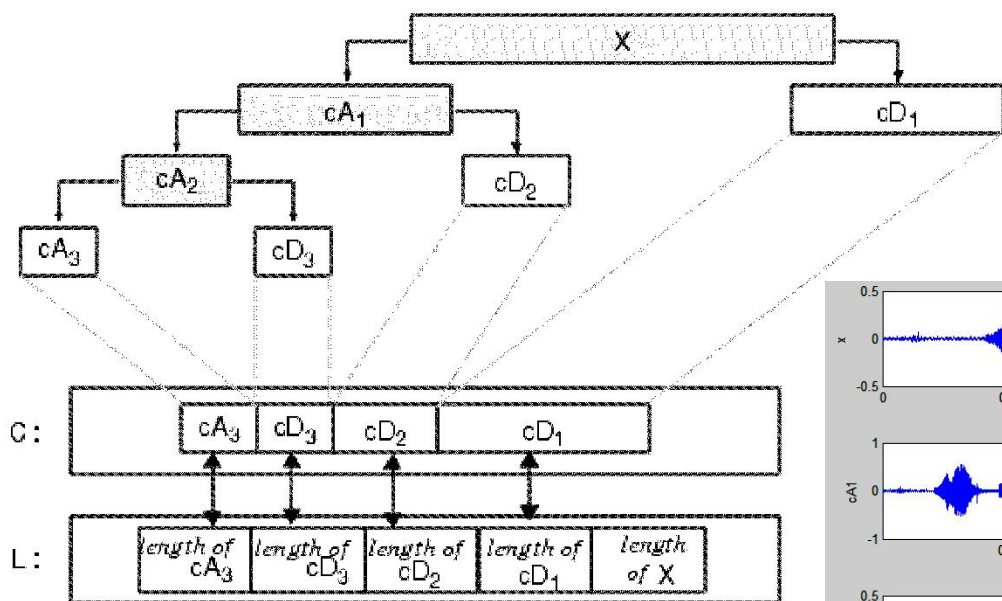
- $X[0] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n]$
- $X[k] = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x[n] \cos \frac{(2n+1)k\pi}{2N}$



常用语音信号变换

DWT原理和系数

音频 (DWT)



常用语音信号变换

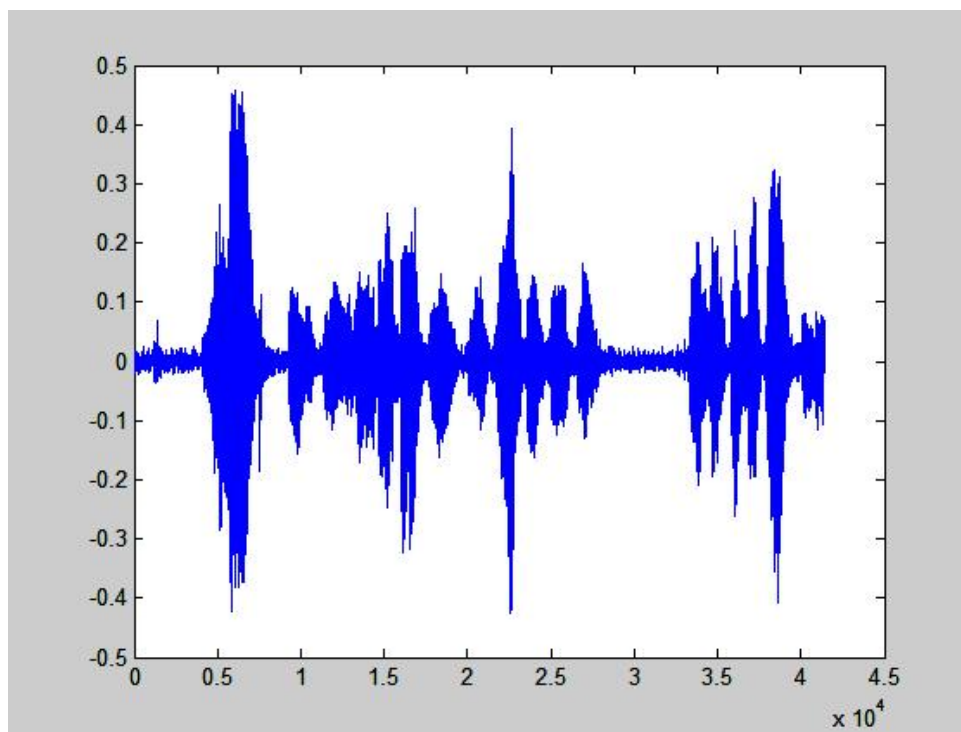
○ 案例二：DWT系数特性研究

- DWT可迭代地对信号进行多层次分解，其中近似系数反映信号的平缓变换情况，可视为信号低频成份，细节系数反映信号剧烈变换情况，可视为信号高频成份。
- 案例尝试修改语音信号DWT细节系数，观察修改前后语音质量变换情况。

常用语音信号变换

○ 案例二：DWT系数特性研究

● 1. 准备语音信号

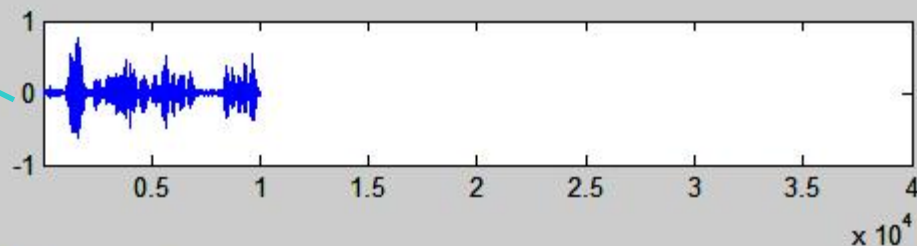


常用语音信号变换

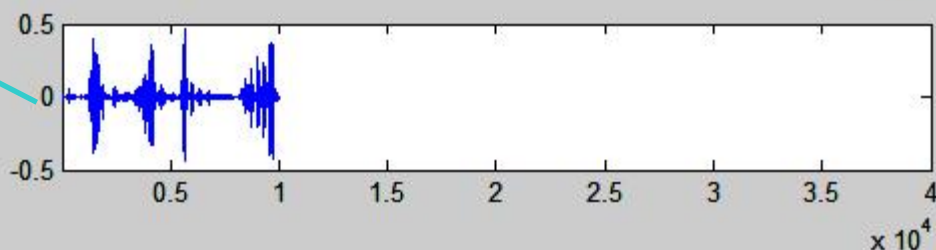
○ 案例二：DWT系数特性研究

- 2. 对语音信号进行2级小波分解，得到各级小波系数。

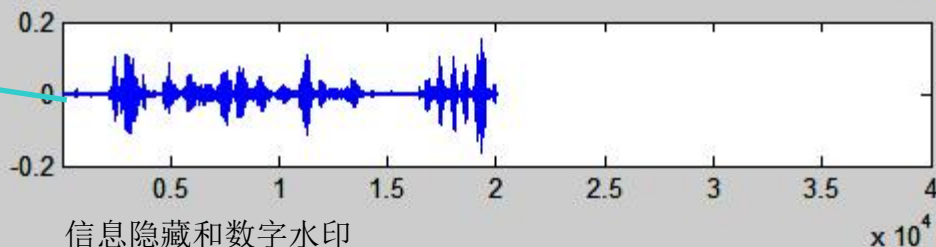
2级近似系数



2级细节系数



1级细节系数

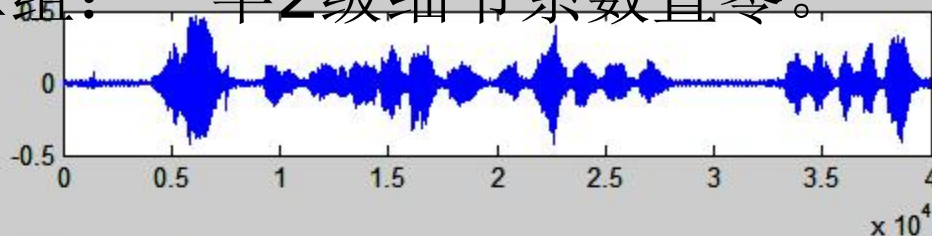


常用语音信号变换

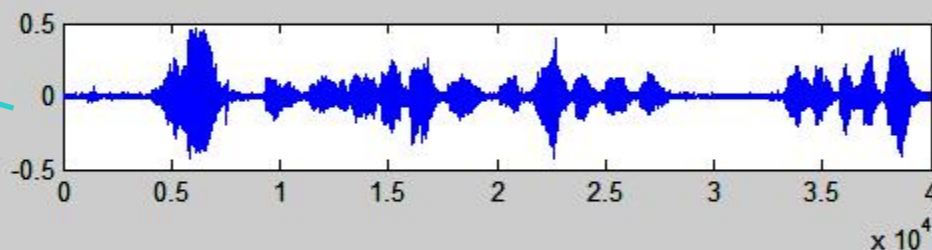
○ 案例二：DWT系数特性研究

- 3. 分别用两组系数重构语音。第一组：原始系数，第二组：一半2级细节系数置零。

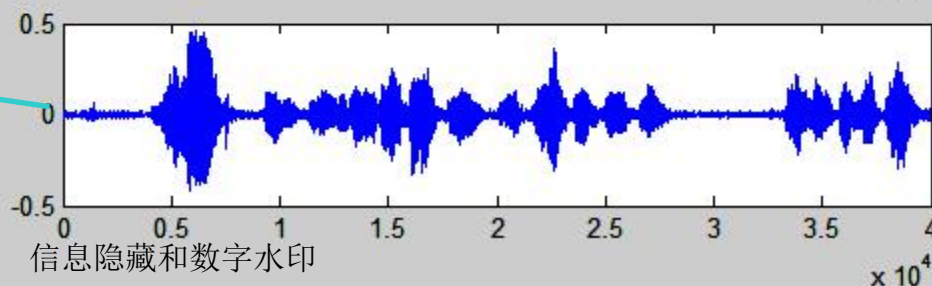
原始音频



原始系数重构音频



修改系数重构音频



图像信号处理基础

- 数字图像的概念
- 人类视觉系统(HVS)特点
- 图像质量评价
- 图像信号处理基础

图像信号处理基础

○ 图像的基本表示

- 图像由像素（**pixel**）组成，像素存储为矩阵。



图像信号处理基础



- 图像分辨率（image resolution）
 - 即图像中像素的数目，常记为：水平*垂直，分别表示水平和垂直方向的像素个数。
 - 例如：显示屏分辨率通常为 $1024*768$ ，表示图像每行有 1024 个像素，每列有 768 个像素。
 - 示例图像为 $350*258$

图像信号处理基础

- 常见图像类型包括

- 灰度图像

- 图像每个像素为大小在0~255间灰度值（gray value）。

- 每个像素用1个字节表示。

- 0:黑色, 255:白色



170	170	169	170	169	166	166	166
170	170	169	170	169	166	166	166
170	170	169	170	169	166	166	166
170	170	169	170	169	166	166	166
170	170	169	170	169	166	166	166
170	170	168	166	168	166	165	166
169	169	168	165	166	168	167	165
166	166	166	165	165	166	166	164

图像信号处理基础

○ 常见图像类型包括

● 24位彩色图像

- 每个像素用三个字节表示，其含义为该像素颜色的红、绿、蓝分量。



62	63	63	65	66	63	61	63			
63	61	59	64	63	60	61	64			
65	29	31	34	30	27	31	31	30		
63	31	31	32	30	28	31	31	32		
63	29	30	64	64	64	60	59	62	62	61
63	29	29	62	64	64	60	59	61	62	62
62	31	32	60	62	63	61	61	64	63	63
65	32	32	62	63	63	60	62	62	62	62
	30	30	62	63	62	61	61	60	60	61
	29	29	61	61	62	59	59	61	65	62
			63	61	60	59	61	63	66	62
			65	60	59	60	62	64	63	61

图像信号处理基础

- 常见图像类型包括
 - 索引（调色板）图
 - 每个像素存储一个调色板索引。
 - 调色板每条记录表示一个颜色。



108	116	9	0	0	0
112	112	1	0.0627	0.0627	0.0314
112	112	1	0.2902	0.0314	0
105	105	1	0	0	1.0000
112	105	1	0.2902	0.0627	0.0627
105	105	1	0.3882	0.0314	0.0941
105	112	1	0.4510	0.0627	0
105	105	1	0.2588	0.1608	0.0627

图像信号处理基础

- 常见图像类型包括
 - 黑白（二值）图像
 - 像素值为0或1，
 - 对应颜色为黑色，白色



1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

图像信号处理基础



人类视觉特点

○ 视觉范围

- 视觉范围是指人眼所能感觉的亮度范围。
- 客观上相同的亮度，当平均亮度不同时，主观感觉的亮度也不相同。

人类视觉特点

○ 视觉范围

- 视觉范围非常宽，但是人眼并不能同时感受这样宽的亮度范围。
- 当平均亮度比较适中时，能分辨的亮度的范围较大。
- 当平均亮度较低时，能分辨的亮度范围较小。

人类视觉特点

○ 视觉适应性

● 暗适应性

- 人由亮处走到暗处时的视觉适应过程，称为暗适应。
- 暗适应过程约需大约三十分钟。

● 亮适应性

- 人由暗处走到亮处时的视觉适应过程，称为亮适应。
- 亮适应过程约需大约一分钟。

人类视觉特点

○ 视觉惰性（视觉暂留）

- 人眼对亮度改变进行跟踪的滞后性质称为视觉惰性。
- 人眼所看到的影像消失后，人眼仍能继续保留其影像约**0.1-0.5**秒左右。

人类视觉特点

○ 视觉惰性（视觉暂留）

- 看纹理丰富的窗帘，然后再看墙壁，看墙时会发现窗帘的花纹重叠在墙上了。
- 看电视，电影是每分钟**20**几帧图片，通过暂留形成了流畅的动作。

图像质量评价——主观评价方法

○ 主观评价

● 平均意见分

- MOS (Mean Opinion Score)
- 观察者根据一些事先规定的评价尺度,
- 或自己的经验,
- 判断测试图像的视觉效果质量,
- 所有观察者给出的分数的加权平均。
- 五级评分

图像质量评价——主观评价方法

- 主观评价的优点：
 - 比较准确。
- 主观评价的缺点：
 - 受观察者的知识背景、情绪，
 - 和疲劳程度等因素的影响。
 - 可重复性较差。

图像质量评价——客观评价方法

○ 客观评价

- 处理和分析系统中输入和输出的图像信号，
- 提取图像特征参量作为研究分析对象，
- 比较特征参量，以此反映图像间的差别。

图像质量评价——客观评价方法

○ 常用客观评价方法

● 均方误差（MSE, Mean Square Error）

- $$MSE = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (x[i, j] - y[i, j])^2$$

● 平均绝对误差（MAE, Mean Absolute Error）

- $$MAE = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |x[i, j] - y[i, j]|$$

图像质量评价——客观评价方法

○ 常用客观评价方法

- 峰值信噪比（PSNR, Peak Signal to Noise Ratio）

- $PSNR = 10lg \left(\frac{x_{Max}^2[i,j]}{MSE} \right)$

图像质量评价——客观评价方法

- 客观评价优点：
 - 使用方便。
 - 可重复性强。
- 客观评价缺点：
 - 评价结果与主观感觉有时存在不一致。

图像的质量评价



- 为什么客观评价与主观评价不能很好吻合？
 - 人眼对所看到物体的理解，
 - 不仅与生理因素有关，
 - 还在相当大的程度上取决于心理因素，
 - 如“视而不见”，“听而不闻”。

图像的质量评价



- 为什么客观评价与主观评价不能很好吻合？
 - 对感兴趣的区域给予极大关注，对其它区域不在意。
 - 大脑对所接收的事务有一个过滤和取舍的过程，目前计算机还无法很好地模拟此过程。

例1



- 高斯噪声，椒盐噪声，PSNR相同

	
PSNR(dB)	PSNR(dB)
48.1183	50.9538



例1（续）

	
PSNR(dB)	PSNR(dB)
34.1389	35.1996

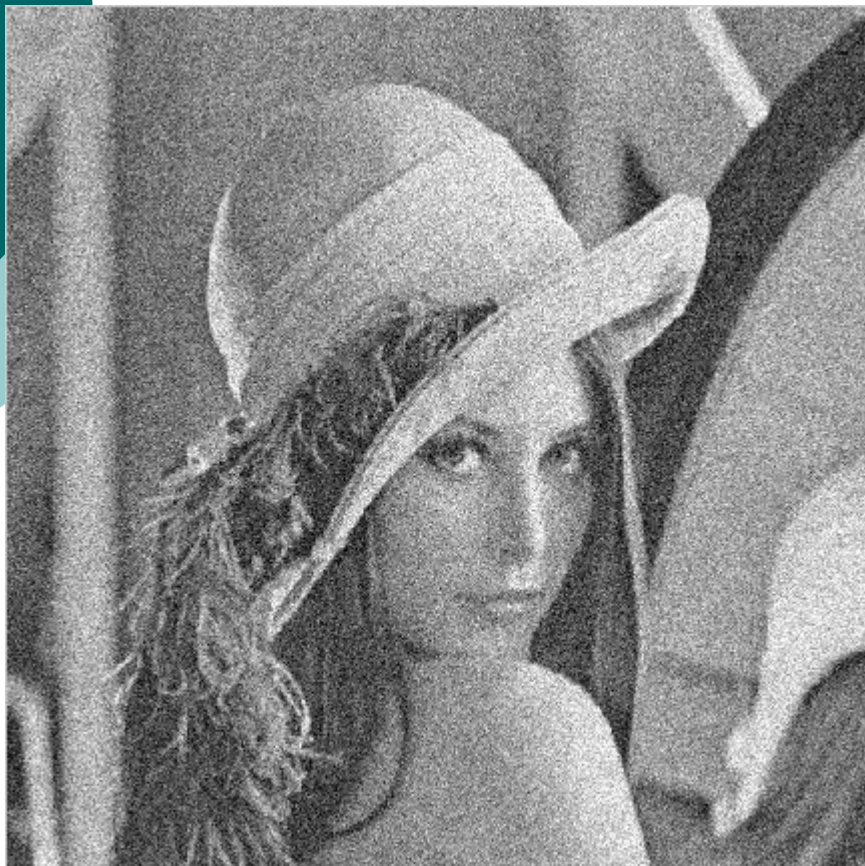
例1（续）

	
PSNR(dB)	PSNR(dB)
28.1183	29.0575



例1（续）

	
PSNR(dB)	PSNR(dB)
19.5175	19.0568

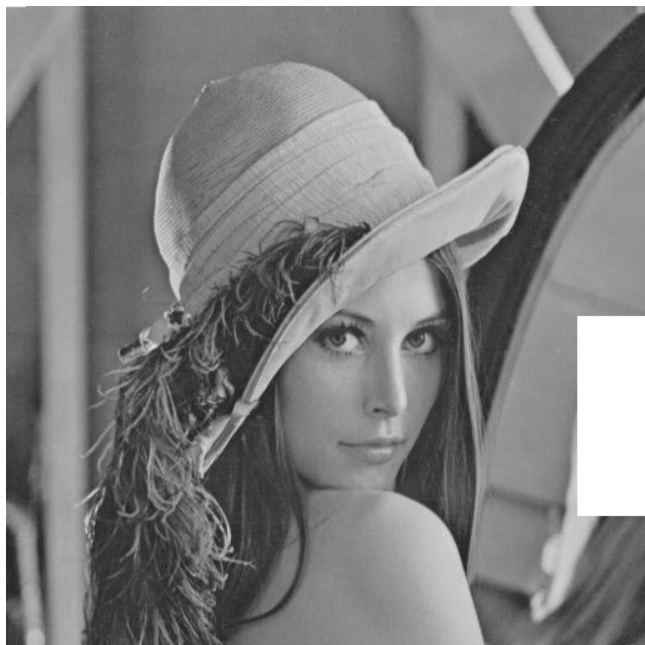

例1（续）





例2

	
PSNR(dB)	PSNR(dB)
23.1805	23.5164

例2（续）

	
PSNR(dB)	PSNR(dB)
23.5084	23.2079

例2（续）

	
PSNR(dB)	PSNR(dB)
33.8497	33.5907

常用图像信号变换

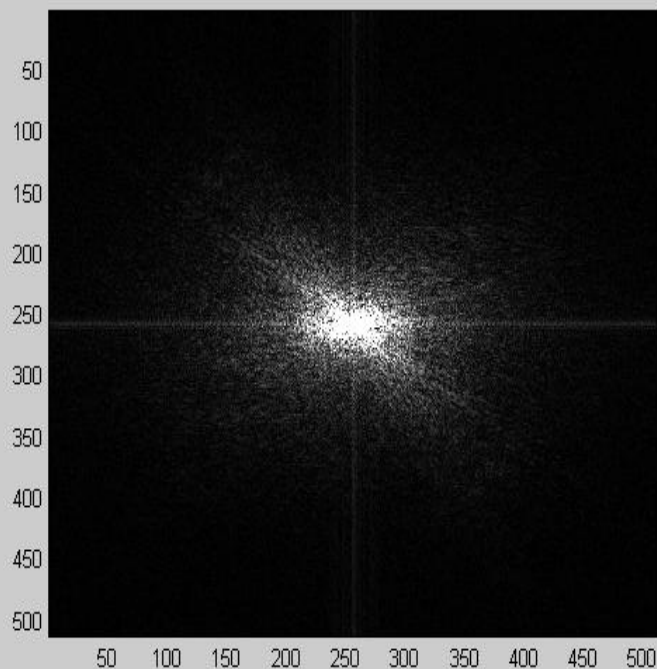
- 常见变换
 - DFT: 离散傅里叶变换
 - DCT: 离散余弦变换
 - DWT: 离散小波变换

常用图像信号变换

图像（空域）



图像（**DFT**）

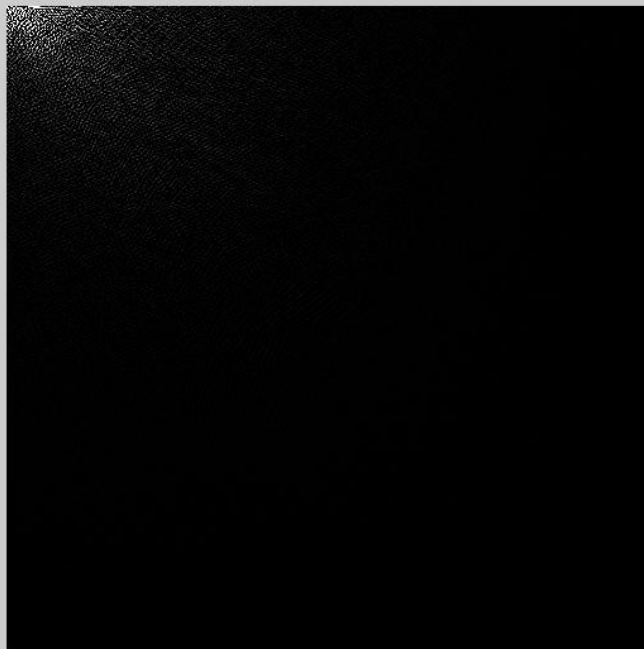


常用图像信号变换

图像（空域）



图像（**DCT**）



常用图像信号变换

图像（空域）



图像（**DWT**）

cA2	cH2	cH1
cV2	cD2	
cV1		cD1

常用图像信号变换

图像（空域）

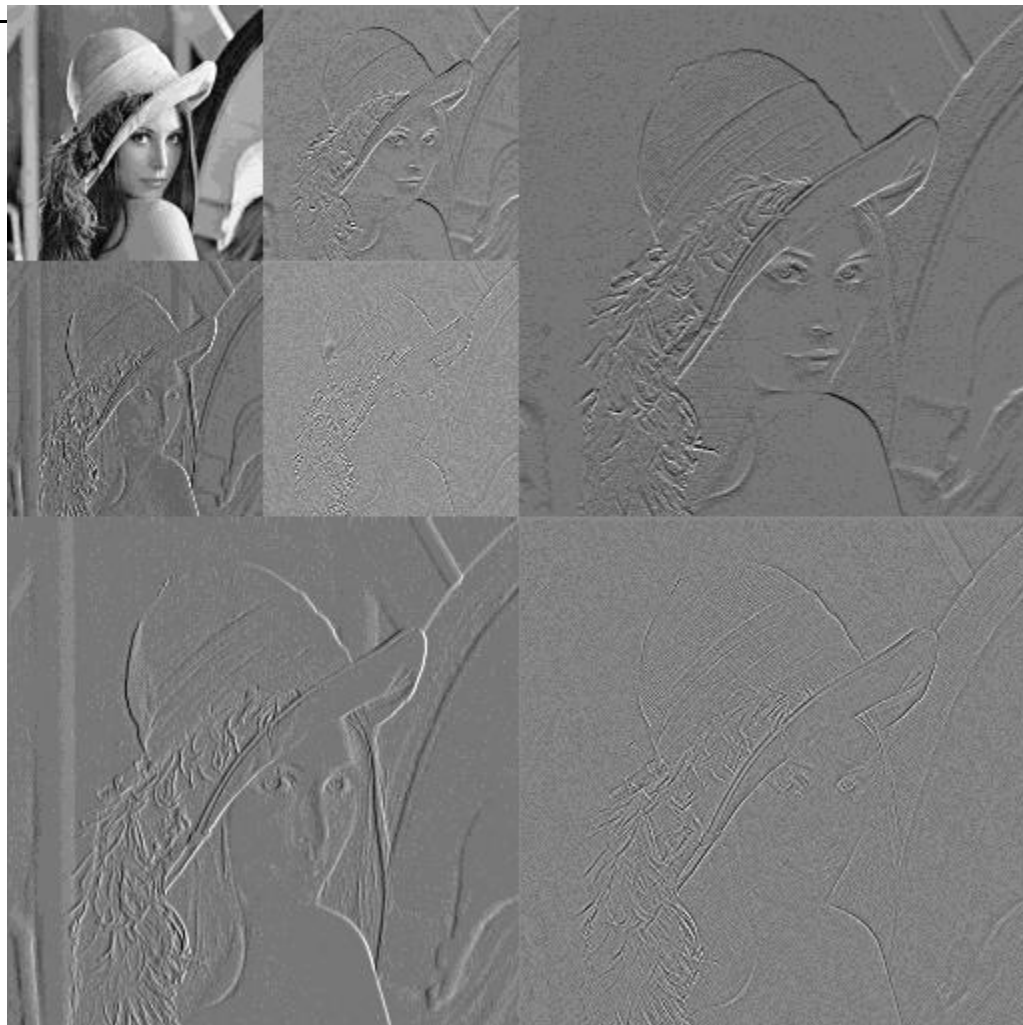


图像（**1级DWT**）



常用图像信号变换

○ 2级DWT



常用图像信号变换

○ 案例四：DCT系数特性研究

- DCT常用于图像信号处理操作，例如，JPEG图像压缩。图像不同位置的DCT系数大小，反映图像不同频率成份的强弱。位置与频率有对应关系，水平方向从左到右，频率逐渐升高；垂直方向从上到下，频率逐渐升高。
- 案例尝试修改图像信号DCT高频部份系数，观察修改前后图像质量变换情况。

常用图像信号变换

○ 案例四：DCT系数特性研究

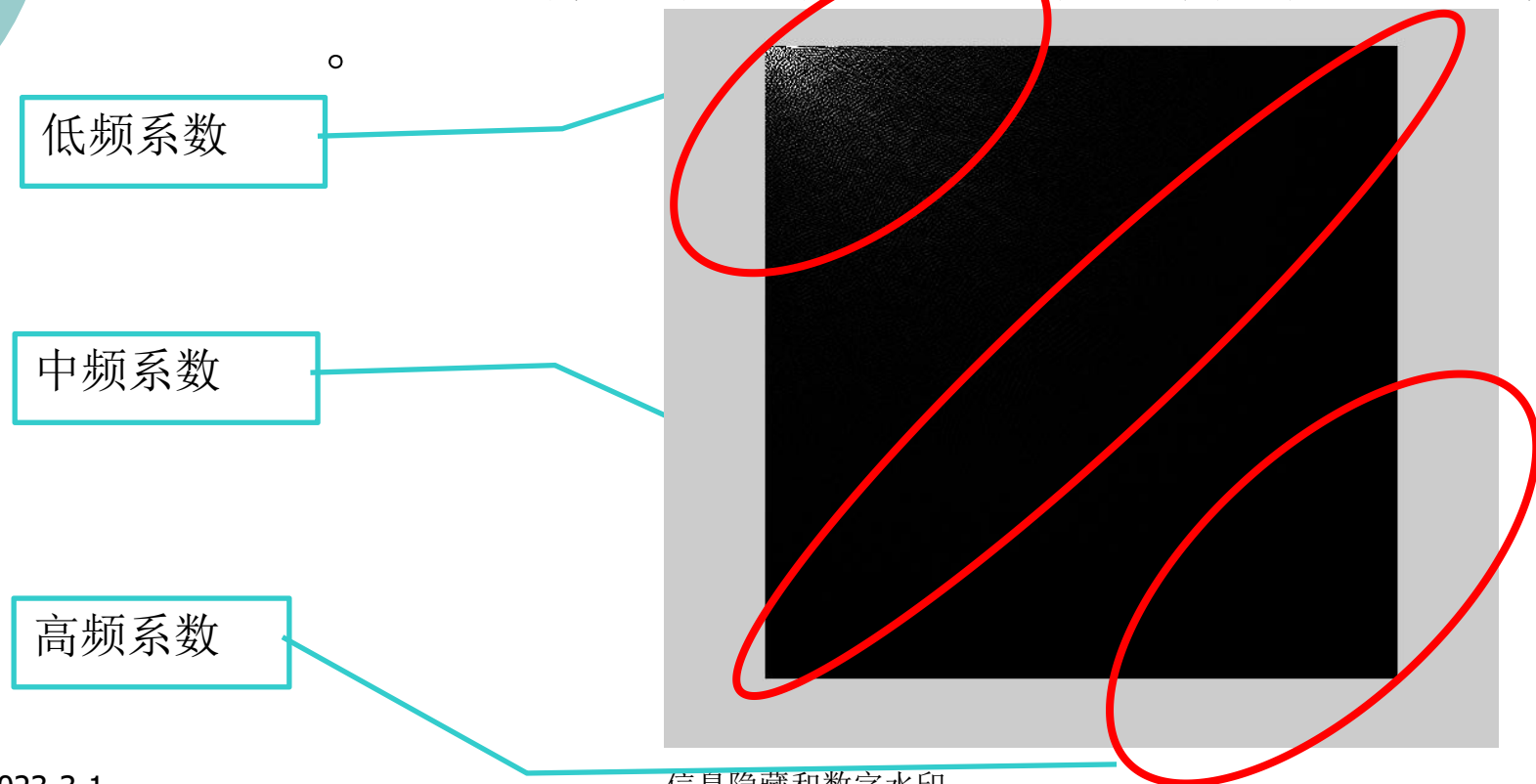
● 1. 准备图像信号



常用图像信号变换

○ 案例四：DCT系数特性研究

- 2. 对图像进行2维DCT变换，得到DCT系数



常用图像信号变换

○ 案例四：DCT系数特性研究

- 3. 分别用两组系数重构图像。第一组：原始系数（左图），第二组：部份高频系数置零（右图）



小结

- 数字音频是时间和样点值都离散的，计算机能处理的信号。
- 数字音频信息按一定的格式保存在磁盘文件中，通常包括采样率，编码方式，编码速率，单或双声道样点信息。
- 响度是人耳对声音强度的主观感受。人耳对不同频率的单音的敏感程度不同，物理强度相同的单音，可能因为频率不同，导致听起来响亮程度不同。

小结

- 人耳难以感知位于强信号附近的弱信号，这种声音心理学现象称为掩蔽。
- 强信号称之为掩蔽音，弱信号称之为被掩蔽音。掩蔽音和被掩蔽音同时存在所产生的掩蔽效应称为同时掩蔽或频域掩蔽，否则称之为异时掩蔽或时域掩蔽。

小结

- 数字图像是空间和样点值都离散的，计算机能处理的信号。
- 数字图像信息按一定的格式保存在磁盘文件中，通常包括分辨率，图像类型，调色板，样点值等信息。
- 常见图像包括二值、灰度、真彩色和调色板图像。
- 亮度是人眼对光强的主观感受。客观上相同的亮度，当平均亮度不同时，主观感觉的亮度也不相同。

小结

- 人由亮处走到暗处时的视觉适应过程，称为暗适应。人由暗处走到亮处时的视觉适应过程，称为亮适应。
- 人眼所看到的影像消失后，人眼仍能继续保留其影像约0.1-0.5秒左右。称为视觉惰性或暂留，电影和动画的制作利用了人眼的视觉惰性。

小结

- 质量评价包括主客观方法。MOS（Mean Opinion Score，平均意见分）为常用主观评价方法。
- SNR（Signal to Noise Ratio，信噪比）为音频质量常用客观评价方法。
- PSNR(Peak Signal to Noise Ratio，峰值信噪比)为图像质量常用客观评价方法。
- 客观评价与主观感受不一定相符。

小结

- 常用语音和图像信号处理操作包括DFT（Discrete Fourier Transform，离散傅里叶变换），DCT（Discrete Cosine Transform，离散余弦变换），DWT（Discrete Wavelet Transform，离散小波变换）。

小结

- DFT、DCT、DWT是一对可逆变换，正变换将信号从时域（音频样点）或空域（图像样点）转换为变换域（变换域系数），反变换则将变换域系数重构为音频或图像。
- 变换域系数的大小反映了信号不同频率分量能量的高低，变换域系数的位置与频率高低关联。通常修改高频系数所引起的失真不容易被察觉。

参考文献

1. 人类视觉与色度学_百度文库
<http://wenku.baidu.com/view/d04fa51614791711cc791758.html>
2. http://blog.sina.com.cn/s/blog_4b72aa9d01008mvl.html
3. <http://tieba.baidu.com/f?kz=82344802>
4. 鈕心忻主编，信息隐藏与数字水印
5. matlab图像处理宝典，电子工业出版社