# 《多媒体技术(含实验)课程》课程教学大纲

|  |  |
| --- | --- |
| 课程名称：多媒体技术（含实验） | 课程代码：ELIE2008 |
| 英文名称：Multimedia Technology |
| 课程性质：（通识、公共、大类等） | 学分/学时： |
| 开课学期：6 |  |
| 适用专业：信息工程、通信工程、电子信息工程 |
| 先修课程：高等数学、工程数学、信号与线性系统、数字信号处理 |
| 后续课程：信息系统综合实验 |
| 开课单位：电子信息学院 | 课程负责人：王加俊 |
| 大纲执笔人：王加俊、陈雪勤 | 大纲审核人：胡剑凌 |

## 课程性质和教学目标（在人才培养中的地位与性质及主要内容，指明学生需掌握知识与能力及其应达到的水平）

**课程性质：**多媒体技术是信息工程的专业必修课、以及通信工程和电子信息工程的专业选修课程，是信号与线性系统、数字信号处理等课程的后续课程。

**教学目标**：通过本课程的学习，使学生掌握语音、音频、图像以及视频等媒体的基本概念以及利用计算机对语音、音频、图像以及视频进行综合处理和分析的基本技术，了解多媒体信息处理技术的应用和发展概况，进一步培养学生解决多媒体信息的传输和处理中复杂问题的能力。本课程的具体教学目标如下：

1. 熟悉多媒体及多媒体技术的基本概念及特征；了解声音的基本特性及主观感觉；熟悉音频、图像、视频信号数字化的过程，掌握均匀量化的原理；
2. 掌握彩色三要素、三基色原理及混色方法等色度学基本知识；理解RGB、YUV、YIQ、YCbCr、HSI/HSV等颜色空间的表示及转换；熟悉ITU-R BT.601建议的主要内容；掌握Matlab中用于图像处理基本函数的用法。
3. 掌握数字图像增强的基本方法和技术；掌握二值图像的数学形态学处理的基本原理和常用方法；掌握图像的阈值化分割、边缘检测、区域分割等图像分割方法；掌握数字图像以及视频的无失真编码、预测编码、变换编码的基本原理、了解JPEG、MPEG等常用的图像及视频编码标准；掌握数字水印的基本概念、数字图像水印的嵌入与提取方法、了解数字视频水印的嵌入和提取方案。
4. 熟悉声学基本概念及度量参数；掌握人的三个主观听感要素：响度、音调、音色的概念；熟悉人的听觉特性和语音的发音机理；了解数字音频文件的多种格式，掌握WAV、RIFF格式的结构；掌握语音信号时频域声学特征提取方法。
5. 掌握数字音频压缩技术的原理；掌握线性预测编码、矢量量化、子带编码的原理，了解MPEG音频编码原理；了解基音频率与声调、说话人性别之间的关联；理解数字音频水印的原理，掌握基本的数字音频水印实现方法；理解声纹识别系统中的特征、分类模型的概念，掌握一种声纹识别系统的设计方法。

## 课程目标与毕业要求的对应关系（明确本课程知识与能力重点符合标准哪几条毕业要求指标点）

|  |  |  |
| --- | --- | --- |
| 毕业要求 | 指标点 | 课程目标 |
| 1、工程知识 | 1-2 掌握多媒体基础知识，具备对音频/语音、视频/图象进行信号分析与处理能力  | 教学目标1 |
| 2、问题分析 | 2-3 能正确运用多媒体技术原理分析复杂工程问题。 | 教学目标2 |
| 3、设计/开发解决方案 | 3-1 能够综合考虑，形成对复杂问题的设计方案。 | 教学目标3 |

## 课程教学内容及学时分配（含课程教学、自学、作业、讨论等内容和要求，指明重点内容和难点内容）（重点内容：★；难点内容：Δ）

**1、多媒体信息处理基础（7学时）**

1.1 多媒体的基本概念

1.1.1 多媒体的基本概念

1.1.2 多媒体与多媒体技术

1.2 图像信息处理基础

1.2.1 光的颜色与彩色三要素

1.2.2 三基色的基本原理

1.2.3 颜色空间模型及转换关系

1.2.4 图像信号的数字化

1.3 视频信号的数字化

1.4 Matlab在图像处理中的应用

1.4.1 Matlab简介

1.4.2 Matlab中图像文件的基本操作

1.5 声学基础知识

1.5.1 声波、声音与声学的基础概念

1.5.2 声音的参数与度量

1.5.3 室内声学基础

1.5.4 人耳的听觉特性

1.5.5 声音质量评价方法

1.6 音频信号数字化处理

1.6.1 音频信号的数字化

1.6.2 过采样与噪声整形

1.6.3 数字音频文件格式

1.6.4 数字音频编辑软件

* **目标及要求**

（1）熟悉多媒体及多媒体技术的基本概念及特征。

（2）了解声音的基本特性及主观感觉。

（3）熟悉音频、图像、视频信号数字化的过程，掌握均匀量化的原理。

（4）掌握彩色三要素、三基色原理及混色方法等色度学基本知识。

（5）理解RGB、YUV、YIQ、YCbCr、HSI/HSV等颜色空间的表示及转换。

（6）熟悉ITU-R BT.601建议的主要内容。

（7）了解MATLAB在图像处理和分析领域的应用。

（8）熟悉声学基本概念及度量参数。

（9）掌握人的三个主观听感要素：响度、音调、音色的概念。

 (10)了解影响室内声学环境的因素以及混响的作用。

（11）掌握人耳听觉范围和听觉掩蔽效应，了解听觉延时效应和双耳效应。

 (12) 熟悉评价声音质量的主客观指标。

（13）熟悉资源交换文件格式RIFF的结构和WAV文件格式的基本结构，了解mp3/midi/ra/rm/wma/ape/au等音频文件格式。

(14) 熟悉数字音频编辑软件如：cooledit/audition等的使用方法。

* **作业内容**

（1）媒体的概念与分类；

（2）多媒体技术的基本概念及特征；

（3）彩色三要素的物理含义，RGB、HIS如何对颜色进行描述?

（4）编写程序实现RGB与HIS之间的转换；

（5）复合数字编码与分量数字编码的基本原理及其优缺点；

（6）ITU-R BT.601建议的主要内容有哪些？有何实际意义？

（7）人耳的听觉范围及等响度曲线的特点？

（8）人耳区别不同人声及各类乐器依据的主要参量？

（9）描述听觉掩蔽效应，思考听觉掩蔽效应的应用。

（10）说明音频信号数字化的三个主要步骤及其功能。

（11）过采样以及噪声整形技术为什么可以提高A/D转换器精度？

**2、图像增强 (6学时)**

2.1 引言

2.2 图像的灰度变换

2.2.1 灰度的线性变换

2.2.2 灰度的非线性变换

2.2.3 直方图均衡化

2.2.4 直方图的规定化

2.3 图像平滑

2.3.1 模板操作和卷积运算

2.3.2 邻域平均法

2.3.3 中值滤波

2.3.4 频域低通滤波

2.4 图像锐化

2.4.1梯度算子

2.4.2 Sobel算子

2.4.3 Laplacian算子

2.4.4 频域高通滤波

2.5 图像的同态滤波

2.6 伪彩色增强

* **目标及要求**

（1）掌握数字图像增强的基本方法和技术；

（2）掌握数字图像灰度的线性与非线性变换的方法及应用；

（3）熟悉直方图均衡化、直方图规定化的步骤；

（4）掌握图像平滑的基本方法，如邻域平均法、中值滤波法、低通滤波；

（5）掌握图像锐化的基本方法，如梯度运算、Sobel算子、拉普拉斯算子、高通滤波；

（6）了解图像的同态滤波；

（7）了解伪彩色增强、假彩色增强的基本方法。

* **作业内容**

（1）图像增强的目的；

（2）利用线性变换实现灰度动态范围的拉伸或压缩；

（3）利用灰度直方图的均衡化和规定化技术实现图像增强；

（4）图像的平均值滤波、中值滤波的适用范围有何区别、分别利用这两种滤波器进行图像去噪处理；

（5）利用一阶微分算子对图像进行锐化处理；

（6）利用同态滤波对图像进行对比度增强处理；

（7）利用灰度分层法或灰度变换法将一幅灰度图像转换成彩色图像。

**3、形态学图像处理（6学时）**

3.1 引言

3.1.1 数学形态学的基本思想

3.1.2 集合论基础

3.1.3 数学形态学中的几个基本概念

3.2 二值形态学基本运算

3.2.1 腐蚀

3.2.2 膨胀

3.2.3 腐蚀与膨胀的对偶性

3.2.4 开运算

3.2.5 闭运算

3.3 二值图像的形态学处理

3.3.1 边缘提取

3.3.2 区域填充

3.3.3 骨架提取

3.3.4 细化

3.3.5 粗化

3.3.6 形态学滤波

* **目标及要求**

（1）了解数学形态学的基本思想；

（2）熟悉集合和子集的概念及表示方法，掌握集合间的关系及并、交、补、差等运算；

（3）理解数学形态学中结构元素的概念及作用；

（4）掌握膨胀、腐蚀、开、闭运算的物理含义，以及由上述各种基本运算导出的各种二值图像形态学处理算法。

* **作业内容**

（1）说明二值膨胀、腐蚀运算对图像处理的作用和特点，并通过编程实现之；

（2）说明二值开、闭运算对图像处理的作用和特点，并通过编程实现之；

（3）利用形态学处理实现二值图像的骨架化；

（4）利用形态学处理提取二值图像的内、外、形态学边缘；

（1）利用形态学滤波实现图像的细化、粗化、以及区域的填充处理。

**4、图像分割（6学时）**

4.1图像分割的概念及分类

4.1.1 图像分割的概念

4.1.2 图像分割的依据和方法分类

4.2 基于灰度阈值化的图像分割

4.2.1 阈值化分割的原理

4.2.2 全局阈值化分割

4.2.3 局部阈值分割法

4.3 基于边缘检测的图像分割

4.3.1 边缘检测的基本原理和步骤

4.3.2 梯度算子

4.3.3 Laplacian算子和 LOG算子

4.3.4 Canny 算子

4.4 基于区域的图像分割

4.4.1 区域生长法

4.4.2 区域的分裂与合并法

* **目标及要求**

（1）了解图像分割的依据和方法分类；

（2）掌握基于灰度阈值化的图像分割方法；

（3）掌握边缘检测的基本原理，熟悉Roberts、Sobel、Prewitt、LOG、Canny等边缘检测算子；

（4）了解轮廓跟踪的基本方法及步骤；

（5）掌握基于区域生长法、区域分裂与合并法的图像分割方法。

* **作业内容**

（1）简述图像分割的依据、分割方法的分类；

（2）分别利用迭代法、Otsu方法确定全局阈值，实现图像的分割；

（3）分别利用Sobel算子、LOG算子、Canny算子进行边缘提取并比较各自的优缺点；

**5、数字图像与视频压缩编码原理（6学时）**

5.1数字图像与视频压缩编码概述

5.1.1 数字图像与视频压缩的必要性和可能性

5.1.2 数字图像与视频压缩编码的主要方法及分类

5.2 无失真编码

5.2.1 游程编码

5.2.2 Huffman编码

5.2.3 算术编码

5.3 预测编码

5.3.1 图像差值信号的统计特性

5.3.2 帧内预测编码

5.3.3 帧间预测编码

5.4 变换编码

5.4.1 图像的频域统计特性

5.4.2 变换编码的基本原理

5.4.3 正交变换基的选择

5.4.4 DCT图像编码

* **目标及要求**

（1）熟悉数字图像与视频编码的基本原理及常用方法；

（2）重点掌握霍夫曼（Huffman）编码、算术编码、预测编码和DCT编码的基本原理；

（3）掌握运动估计和运动补偿预测编码的基本原理。

* **作业内容**

（1）简述图像压缩的依据、原理、方法及其分类；

（2）利用Huffman编码方法对单符号离散信源进行编码；

（3）利用算术编码方法离散信源符号序列进行无失真信源编码，并比较算术编码和Huffman编码的优缺点；

（4）说明预测编码的原理，并画出DPCM的原理框图；

（5）通过编程，将一幅512×512的256灰度级的图像分成4096个8×8的子块，并对每一块进行DCT变换，保留其中32个较大的变换系数实现2:1的压缩比；通过补零操作以及反变换恢复出512×512的图像，比较此图和原图的视觉质量。

**6、数字图像与视频压缩编码标准（6学时）**

6.1 静止图像的编码标准

6.1.1 JPEG标准概述

6.1.2 JPEG基本编码系统

6.1.3 基于DCT的渐进编码

6.1.4 分级编码

6.2 数字视频编码的标准进化过程

6.3 MPEG-1/MPEG-2视频编码标准

6.3.1 I帧、P帧、B帧

6.3.2 时频码流的分层结构

6.3.3 MPEG-1/MPEG-2视频编解码原理

* **目标及要求**

（1）掌握JPEG基本系统的编解码原理；

（2）理解MPEG-2视频编码中的“类”和“级”的含义；

（3）理解I帧、P帧、B帧图像的编码特点，熟悉MPEG-2视频码流的分层结构；

* **作业内容**

（1）JPEG标准采用了何种压缩算法？画出JPEG算法的原理框图，写出JPEG压缩算法的主要步骤；

（2）解释帧重排的原因；

（3）画出MPEG-1视频编码的原理框图，说明每一模块的功能；

**7、数字水印技术（3学时）**

7.1 数字水印概述

7.1.1 数字水印技术产生的背景和应用

7.1.2 数字水印的基本特征

7.1.3 数字水印系统的组成

7.1.4 数字水印的分类

7.2 数字图像水印算法

7.2.1 最低有效位方法

7.2.2 基于DCT域的方法

7.3 数字视频水印的嵌入和提取方案

7.3.1 基于未压缩的原始视频的水印方案

7.3.2 基于视频编码的水印方案

7.3.3 基于压缩视频码流的水印方案

* **目标及要求**

（1）了解数字水印的基本特征、分类和应用；

（2）掌握水印数字系统的组成，数字水印的嵌入和提取的基本原理；

（3）掌握最低有效位（LSB）法、基于DCT的数字图像水印嵌入和提取算法；

（4）了解数字视频水印的嵌入和提取方案。

* **作业内容**

（1）何为数字水印？可分成几类？

（2）简述数字水印的嵌入和提取过程；

（3）通过编程，实现数字图像DCT域水印的嵌入与提取。

**8、数字音频压缩编码（3学时）**

8.1 数字音频编码概述

8.1.1 音频信号的分类

8.1.2 数字音频压缩编码的原理

8.1.3 音频编解码器的性能指标

8.1.4 数字音频编码技术的分类

8.2 常用数字音频编码技术

8.2.1 线性预测编码

8.2.2 矢量量化原理

8.2.3 CELP编码

8.2.4 感知编码

8.2.5 子带编码

8.2.6 无损音频编码

8.3 MPEG-1/MPEG-2音频编码标准

8.3.1 MPEG-1音频编码算法的特点

8.3.2 MPEG-2音频编码标准

8.3.3 一些其他编码标准

* **目标及要求**

（1）掌握数字音频压缩技术的几类方法和压缩编码原理；

（2）掌握线性预测编码、矢量量化、子带编码的原理。

（3）了解MPEG-1，MPEG-2音频编码器的原理

* **作业内容**

（1）音频编码技术通常分为哪几类，各有什么优缺点？

（2）音频编解码器的主要性能指标有哪些？这些指标相互之间有何关系？

（3）子带编码的基本思想是什么？进行子带编码有什么好处？

（4）解释矢量量化与标量量化，举例说明；

（5）请说明LPC-10的2.4kbits传输率的组成？

**9、语音信号分析（5学时）**

9.1 语音信号数字化及预处理

9.1.1 预滤波及采样量化

9.1.2 预处理

9.2 时域分析

 9.3.1 短时能量及短时平均幅度

9.3.2 短时过零率

 9.3.3 短时相关分析

 9.3.4 短时平均幅度差

9.3 频域分析

 9.3.1 短时傅里叶变换

 9.3.2 临界频带特征

9.4 同态处理

 9.4.1 同态处理的基本原理

 9.4.2 复倒谱与倒谱

9.5 线性预测分析

 9.2.1 线性预测基本原理

 9.2.2 预测参数的求解

 9.2.3 LPC谱估计

9.6 基音周期估计

 9.6.1 自相关法

 9.6.2 平均幅度差函数法

 9.6.3 倒谱法

 9.6.4 一些优化方法

* **目标及要求**

（1）掌握短时能量、短时过零率、基音频率等特征计算方法；

（2）理解端点检测与基音检测之间的关系。

（3）理解同态滤波与倒谱之间的关系及倒谱计算方法；

（4）理解线性预测方法的原理；

（5）了解基音检测的改善方法。

* **作业内容**

（1）已知采样频率，那么基音周期、基音频率与样本点数之间的运算方法？

（2）对语音进行基音检测之前为什么要先做端点检测、加窗分帧处理？

（3）画出一帧语音信号采用短时自相关法进行基音检测的过程，用示意图画出中间状态。

（4）给出统计数据下，各年龄段、性别的人群基音频率的范围，分析产生该分布的原因。

**10、音频数字水印技术（3学时）**

10.1 音频数字水印概述

10.1.1 音频数字水印的应用背景

10.1.2 音频数字水印方法概述

10.2 音频数字水印基本原理

10.2.1 音频数字水印的嵌入

10.2.2 音频数字水印的提取

10.3 几种典型的音频数字水印方法

10.3.1 时域音频数字水印方法

10.3.2 变换域音频数字水印方法

10.3.3 压缩域音频数字水印方法

10.4 音频数字水印评价标准

10.4.1 感知透明性评价方法

10.4.2 鲁棒性评价

10.4.3 几种常见音频数字水印攻击

10.5 实验示例

10.5.1 水印的设计

10.5.2 实验方法

10.3.3 性能分析

* **目标及要求**

（1）掌握数字音频水印嵌入和提取的原理；

（2）了解数字音频水印的应用领域；

（3）掌握LSB的实现原理和方法；

（4）理解变换域和压缩域的方法；

（5）了解音频数字水印的攻击手段；

（6）熟悉音频数字水印的评价标准。

* **作业内容**

（1）画出数字音频水印嵌入和提取的原理框图。

（2）简述LSB方法实现音频数字水印的原理，以及该方法的优缺点。

（3）通常数字音频水印的主要性能指标包括哪几方面？指标间是否有关联？

（4）请列举音频数字水印的评价方法。

**11、声纹识别方法（3学时）**

11.1 声纹识别概述

11.1.1 什么是声纹识别

11.1.2 声纹识别发展现状

11.1.3 声纹识别应用

11.2 声纹识别基本原理

11.3 声纹识别中常用特征

11.3.1 语音声纹特征概述

11.3.2 线性预测系数（LPC）

11.3.3 线性预测倒谱系数（LPCC）

11.3.4 线性预测倒谱系数（LPCC）

11.3.5 Mel频率倒谱系数（MFCC）

11.4 高斯混合模型

11.4.1 高斯混合模型原理

11.4.2 基于GMM的声纹识别

* **目标及要求**

（1）掌握声纹识别的原理；

（2）了解声纹识别的应用领域；

（3）掌握Mel频率倒谱系数的计算方法；

（4）理解LPC、LPCC的计算方法；

（5）理解高斯混合模型的原理；

（6）掌握声纹识别的实现方案流程。

* **作业内容**

（1）什么是声纹识别，声纹识别可以怎样分类？

（2）声纹识别主要有哪些应用？

（3）画出MFCC特征的计算框图。

（4）画出基于高斯模型的声纹识别系统框图。

12、实验(18学时)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 项目名称 | 实验类型 | 学时分配 | 每组人数 | 必修/选修 |
| 1 | 图像增强和图像分割 | 综合性 | 9学时 | 1 | 必修 |
| 3 | 语音基音检测及应用 | 综合性 | 9学时 | 1 | 必修 |

**四、教学方法**

授课方式：a.理论课（讲授核心内容、总结、按顺序提示今后内容、答疑、公布习题和课外拓展学习等）；b.课后练习（按照理论内容进行）；c.实验环节（根据理论课教学内容，要求学生学会相应的Matlab命令和函数，编写相应的处理程序完成实验任务）；d.办公室时间（学生无需预约，可来教师办公室就课程内、外内容进行讨论）；e.答疑（全部理论课程和实验课程完成后安排1～2次集中答疑，答疑时间不包括在课程学时内，答疑内容包括讲授内容、习题、实验等）；g.期中考查和期末闭卷考试。

课程要求：a.理论课：在理论课讲授环节中，应注意概念讲清讲透，并贯彻理论联系实际的原则，注意学生逻辑思维能力、工程观点和分析与解决问题能力的培养。根据本课程的特点，必须严格要求学生独立完成一定数量的简单算法编程实现；b.实验环节：要求学生学会通过C++或MATLAB编程，完成综合性的多媒体信息处理实验，培养学生独立进行设计和分析问题的能力。

**五、考核及成绩评定方式**

**考核方式**：

期中考试：考查；

期末考试：闭卷笔试；

平时成绩：课堂表现及作业；

实验成绩：实验表现和实验结果。

**成绩评定方式**：期末成绩50%，期中成绩20%，平时成绩10%，实验成绩20%

**六、教材及参考书目**

**教材：**

卢官明，焦良葆. 多媒体信息处理，邮电出版社，2011

多媒体技术-音频部分讲义（自编，未出版，2016.2）

**参考书目：**

R.C.Gonzalez, R.E.Woods, Digital Image Processing (2nded) Prentice Hall, 2002.

R.C.Gonzalez, R.E.Woods,S.L. Eddins, Digital Image Processing Using MATLAB, 电子工业出版社，2004.

许录平. 数字图像处理，科学出版社，2007.

卢官明，宗昉. 数字音频原理及应用（第二版），机械工业出版社，2012