

卓越工程师教育培养计划配套教材

电气工程系列

MATLAB与音视频技术

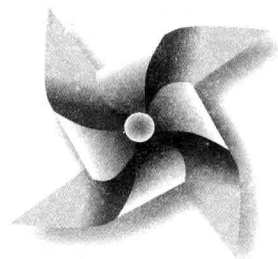


王永琦 编著

清华大学出版社

卓越工程师教育培养计划配套教材

电 气 工 程 系 列



MATLAB与音视频技术

王永琦 编著

清华大学出版社

内 容 简 介

本书系统地阐述了语音信号处理和图像处理的原理、分析方法、技术和应用。本书的内容主要分三篇：第1篇是“基础篇”，对 MATLAB 语音和图像视频处理的特点及基本知识进行了阐述，层层深入地介绍了 MATLAB 的二维和三维图像绘制功能，并介绍了科学计算可视化的相关知识。第2篇是“语音处理技术”，详细介绍了语音信号处理的基础知识和典型案例。第3篇是“图像处理技术”，详细介绍了图像处理工具箱的功能和相关函数，介绍了图像表示和描述，图像分割、增强、融合以及拼接等方面的知识，并给出图像处理应用的实例。

本书结构合理，实例丰富，注重理论和实际应用的结合，可读性强。本书可以作为高等院校相关专业的教材或者参考书，也可以作为图像和语音处理工程技术人员的参考书。

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

图书在版编目(CIP)数据

MATLAB 与音视频技术/王永琦编著. --北京：清华大学出版社，2013
卓越工程师教育培养计划配套教材. 电气工程系列
ISBN 978-7-302-33461-3

I. ①M… II. ①王… III. ①Matlab 软件—高等学校—教材 IV. ①TP317

中国版本图书馆 CIP 数据核字(2013)第 188539 号

责任编辑：庄红权
封面设计：常雪影
责任校对：王淑云
责任印制：沈 露

出版发行：清华大学出版社

网 址：<http://www.tup.com.cn>, <http://www.wqbook.com>

地 址：北京清华大学学研大厦 A 座 邮 编：100084

社总机：010-62770175 邮 购：010-62786544

投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈：010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者：北京富博印刷有限公司

装 订 者：北京市密云县京文制本装订厂

经 销：全国新华书店

开 本：185mm×260mm 印 张：21.75 字 数：523 千字

版 次：2013 年 11 月第 1 版 印 次：2013 年 11 月第 1 次印刷

印 数：1~2500

定 价：45.00 元



卓越工程师教育培养计划配套教材

总编委会名单

主任：丁晓东 汪 泓

副主任：陈力华 鲁嘉华

委员：（按姓氏笔画为序）

丁兴国 王岩松 王裕明 叶永青 刘晓民

匡江红 余 粟 吴训成 张子厚 张莉萍

李 毅 陆肖元 陈因达 徐宝钢 徐新成

徐滕岗 程武山 谢东来 魏 建

卓越工程师教育培养计划配套教材

——电气工程系列子编委会名单

主任：王裕明 李毅

副主任：陆肖元 史志才 张莉萍

委员：（按姓氏笔画为序）

孔勇 方易圆 王永琦 邓琛 余朝刚

张瑜 张颖 陈宇晨 陈益平 卓郑安

罗晓 高飞 黄润才



教育部于2010年开始实施的“卓越工程师教育培养计划”是要培养造就一大批创新能力强、适应经济社会发展需要的高质量工程技术人才,为国家走新型工业化发展道路、建设创新型国家和人才强国战略服务。为培养学生的工程意识、工程素质、工程实践能力、工程设计能力和工程创新能力,培养面向未来、高素质、具有国际竞争力的创新型卓越工程师,上海工程技术大学在办学过程中,始终以服务国家和地区经济建设为宗旨,坚持“学科链、专业链对接产业链”的办学模式。2010年,上海工程技术大学车辆工程等专业被列为教育部“卓越工程师教育培养计划”首批试点专业。2011年,电子信息工程(广电通信网络工程)等专业也被列为教育部“卓越工程师教育培养计划”试点专业。

针对目前工科大学生工程能力弱,理论水平与实践能力不匹配,相关教材的理论和实验与工程实际有一定距离,不能满足卓越工程师的培养目标要求等问题,上海工程技术大学电子电气工程学院组织有丰富教学经验和实践能力的骨干教师,联合业内专家,合作编写了“卓越工程师教育培养计划”电气工程系列教材。

本系列教材以社会需求为导向,以实际工程应用为背景,以工程技术为主线,着眼于提高学生的工程意识、工程素质和工程实践能力,按照理论与实践相结合的原则,参阅了大量的中、外文参考书籍和文献资料,吸收借鉴国内外同类教材的优点,参考电子信息产业的相关材料,综合各方面考虑,进行编写。全书坚持加强基础理论,并对基本概念、基础知识和基本技能进行详细阐述,同时强调企业和社会环境下的综合工程应用。

本系列教材注重基本概念、突出工程应用、内容编排新颖,具有基础性、系统性、应用性等特点,能够满足电子信息工程(广电通信网络工程)专业以及车辆工程等专业“卓越工程师教育培养计划”的电气工程类课程的教学目标和要求,体现了“面向工业界、面向世界、面向未来”的工程教育理念,凸显出上海工程技术大学“学科链、专业链对接产业链”和“面向生产一线,培养优秀工程技术人才”的办学特色。

朱仲英

上海交通大学电子信息与电气工程学院

2012年3月15日



随着 NGN、3G 及 3G 演进和 NGB 等对视频、多媒体业务与网络应用的需求飞速发展,作为媒体互动业务及应用的核心技术的多媒体音频视频信号的处理技术,越来越引起人们的关注,成为目前广播、视频与多媒体通信领域中的亮点与热点。

但在目前已出版的教材或者相关著作中,语音信号和视频图像的处理常常是分开撰写的,而广播电视工程专业的学生从事媒体互动业务开发时,需要掌握语音和视频处理的相关技术。编写本书的目的就是想要让学生通过对本书的学习,了解语音和视频技术的基本理论,学会用 MATLAB 语言处理实际的语音和视频信号的技术,培养解决实际问题的能力。本书的最大特色就是理论和实际相结合,融入了编者多年从事语音信号处理和视频处理的科研成果,在阐述基本理论的同时,辅以 MATLAB 源程序,加以详细注释,便于学生学习。

本书全面而细致地讲解了 MATLAB 语音和图像处理及界面编程功能,全书共有 14 章,分为 3 篇。第 1 篇是“基础篇”,包括第 1~3 章,对 MATLAB 语音和图像视频处理的特点及基本知识进行了阐述,层层深入地介绍了 MATLAB 的二维和三维图像绘制功能,并介绍了科学计算可视化的相关知识。第 2 篇是“语音处理技术”阐述,包括第 4~8 章。第 4 章对语音信号特性以及信号的采集和存储过程进行了介绍;第 5 章介绍了语音信号的短时域分析技术和方法;第 6 章介绍了语音信号的频域分析技术和方法;第 7 章介绍了语音信号线性预测分析技术和方法;第 8 章详细介绍了语音信号处理的一些典型案例实现。第 3 篇是“图像处理技术”,包括第 9~14 章。第 9 章详细介绍了图像处理工具箱的功能和相关函数;第 10 章介绍图像的正交变换,包括图像的小波变换、傅里叶变换和离散余弦变换等;第 11 章是图像增强,包括了频域增强、空域增强以及彩色增强等算法介绍;第 12 章是图像分割,包括阈值分割、区域分割和分水岭分割算法介绍;第 13 章是形态学处理,包括了膨胀与腐蚀,开操作与闭操作,击中与击不中变换以及形态学的应用;第 14 章详细介绍了图像处理中的一些典型案例。

本书力求内容丰富、图文并茂,使之成为 MATLAB 在音视频应用方面有价值的参考书。但是由于编者水平有限,错误和不足之处在所难免,敬请读者多提宝贵意见,以便继续完善。在此特别感谢家人对编者工作的大力支持,没有他们的支持,本书不可能那么快与读者见面。

编者

2013 年 7 月



CONTENTS

目录

第 1 篇 MATLAB 基础篇

第 1 章 MATLAB 程序设计基础	3
1.1 MATLAB 数据类型	3
1.2 脚本和函数	8
1.3 变量和语句	10
1.4 程序的调试	15
1.5 小结	18
第 2 章 MATLAB 数据可视化	19
2.1 二维绘图	19
2.2 三维绘图	23
2.3 图形处理	27
2.4 小结	32
第 3 章 MATLAB 图形用户界面设计	33
3.1 图形对象句柄	33
3.2 图形用户界面制作	37
3.3 GUIDE 图形界面开发	44
3.4 小结	51

第 2 篇 MATLAB 在语音处理中的应用

第 4 章 MATLAB 语音信号采集与处理	55
4.1 数字音频的基础	55
4.2 音频的数字化	58
4.3 声卡	61



4.4	基于 GUI 的音频采集处理系统	65
4.5	小结	74
第 5 章	MATLAB 语音信号时域分析应用	75
5.1	语音信号预处理	75
5.2	语音信号的时域分析	79
5.3	语音端点检测系统	91
5.4	基于短时自相关的说话人性别判决系统	103
5.5	小结	112
第 6 章	MATLAB 语音信号频域分析应用	113
6.1	语音信号处理的频域分析技术	113
6.2	基于频域的语音短时合成技术	117
6.3	语音信号频谱分析仪设计	120
6.4	基于 MATLAB 语谱图的显示与分析	125
6.5	小结	134
第 7 章	MATLAB 语音信号线性预测分析应用	135
7.1	语音线性预测原理	135
7.2	LPC 的自相关解法及 MATLAB 的实现	136
7.3	LPC 谱估计和 LPC 倒谱	139
7.4	线谱对 LSP 分析	143
7.5	线性预测分析技术应用	146
7.6	小结	152
第 8 章	MATLAB 语音处理的扩展编程	153
8.1	谱减法及改进算法语音增强系统	153
8.2	基于小波变换的语音水印系统	160
8.3	基于 DTW 的数字语音识别系统	167
8.4	BP 神经网络在说话人识别中的应用	176
8.5	小结	184

第 3 篇 MATLAB 在图像处理中的应用

第 9 章	图像处理基础	187
9.1	图像文件的读写及显示	187
9.2	图像的运算	196
9.3	邻域与块运算	212
9.4	小结	220



第 10 章 图像变换	221
10.1 二维离散傅里叶变换	221
10.2 二维离散小波变换	226
10.3 二维离散余弦变换	231
10.4 Radon 变换	235
10.5 Hough 变换	240
10.6 小结	243
第 11 章 图像增强	244
11.1 线性滤波增强	244
11.2 空域滤波增强	249
11.3 频域滤波增强	255
11.4 彩色增强	263
11.5 小结	267
第 12 章 边缘检测和图像分割	269
12.1 边缘检测和图像分割原理	269
12.2 边缘检测	271
12.3 边界跟踪	277
12.4 阈值分割	280
12.5 分水岭分割算法	283
12.6 区域分割	285
12.7 小结	289
第 13 章 图像的形态学处理	290
13.1 形态学基本运算	290
13.2 击中/击不中变换	296
13.3 形态学在图像处理中的应用	298
13.4 小结	303
第 14 章 MATLAB 图像处理扩展编程	304
14.1 基于 GUI 的小波图像融合系统	304
14.2 基于 DCT 的彩色图像水印	309
14.3 基于不变矩的图像检索	313
14.4 全景图像拼接	319
14.5 图像小波分析压缩	326
14.6 小结	332
参考文献	333



ARTICLE

第1篇 ●

MATLAB 基础篇



MATLAB程序设计基础

读者将在本章了解 MATLAB 的基本程序控制结构以及 MATLAB 中几种基本的数据类型,同时掌握 MATLAB 的脚本、函数两种 M 文件的编写方法,并能够掌握一些提高程序执行效率的编程方法。

1.1 MATLAB 数据类型

MATLAB 有 17 种基本的数据类型,每种类型的数据都以矩阵或数组的形式存在。其中常用的数据类型有数值型、字符和字符串、元胞数组和构架数组。下面具体介绍这 4 种数据类型。

1.1.1 数值型

数值型数据包括有符号和无符号整数、单精度和双精度的浮点数。在默认状态下, MATLAB 将所有的数据都看成是双精度的浮点数。但从节省空间的角度看,整数和单精度数组更节省内存空间。

所有的数值型数据都支持数组操作,如下标操作和尺寸的修改,除了 int64 和 uint64,都可以用于数学运算。

1. 整型

MATLAB 提供了 4 种带符号数和 4 种无符号数整型数据类型,整数的数据类型及其取值范围见表 1.1。

表 1.1 整数的数据类型及其取值范围

数据类型	取值范围	转换函数
单精度 8 位整数	$-2^7 \sim 2^7 - 1$	Int8
单精度 16 位整数	$-2^{15} \sim 2^{15} - 1$	Int16
单精度 32 位整数	$-2^{31} \sim 2^{31} - 1$	Int32
单精度 64 位整数	$-2^{63} \sim 2^{63} - 1$	Int64
无符号 8 位整数	$0 \sim 2^8 - 1$	Uint8
无符号 16 位整数	$0 \sim 2^{16} - 1$	Uint16
无符号 32 位整数	$0 \sim 2^{32} - 1$	Uint32
无符号 64 位整数	$0 \sim 2^{64} - 1$	Uint64



整数可以参与的运算有下面两种：

(1) 具有相同数据类型的整数或整数数组，运算结果的数据类型与操作数相同。

```
如： >> a = uint8(4)
      >> b = uint8(9)
      >> c = a * b
      c =
          36
```

(2) 整数或者整数数组与下标 double 型浮点数进行数值运算，在运算过程中保持浮点运算精度，但是最终结果为整数。

```
如： >> a = 0.3
      >> b = uint8(9)
      >> c = a * b
      c =
          3
```

2. 浮点数

浮点数有单精度(4 字节)和双精度(8 字节)两种，默认是双精度。两种格式可以进行强制转换。

(1) single：将其他类型的数据转换成单精度浮点数。

(2) double：将其他类型的数据转换成双精度浮点数。

例如：

```
>> a = uint8(4)
>> b = double(a)
>> class(a)
ans =
    uint8
>> class(b)
ans =
    double
```

class 函数返回的是对象类型，从上例可以看到，变量 a 原来是无符号 8 位整数，通过 double 函数将其变换成双精度浮点型数据。表 1.2 列出了常用的浮点数操作函数。

表 1.2 常用的浮点数操作函数

函 数 名	函 数 功 能
isfloat	检查输入是否为浮点数
realmax	返回本计算机能够表示的最大浮点数
realmin	返回本计算机能够表示的最小浮点数
eps	浮点相对精度
isreal	检查是否数组的所有元素为实数

1.1.2 字符和字符串

字符和字符串是 MATLAB 的重要组成部分。MATLAB 具有强大的字符串处理功能，在 MATLAB 中，字符串和字符数组基本上是等价的。



1. 字符

在 MATLAB 中,字符是以其 ASCII 码表示的,这样可直接在屏幕上显示字符或者在打印机上打印字符。输入字符数据时应用单引号括起来,例如输入:

```
>> name = '王'
```

这时采用 class 命令可以检查其类型。

```
>> class(name)
ans =
    char
```

这说明变量 name 的类型为字符型。

```
>> whos
Name      Size      Bytes  Class
name      1x1       2      char
```

通过 whos 函数查询得到,变量 name 含有 1 个汉字,占用了 2 个字节,这说明每个字符都采用 16 位的 ASCII 码存储,这大大方便了在 MATLAB 中使用双字节内码字符集,如汉字系统。利用 double 和 char 函数可在字符与其 ASCII 码之间进行转换。

2. 字符串

把字符放在一对单引号中,就定义了一个一维的字符数组,一维字符数组也称为字符串或字符向量。例如:

```
>> university = '上海工程技术大学'
university =
    上海工程技术大学
>> whos university
Name      Size      Bytes  Class
university 1x8       16     char
```

由于 MATLAB 语言是采用 C 语言开发的,因此它的字符串操作与 C 语言的相应操作基本相同。表 1.3 给出了 MATLAB 常用的字符串操作函数及其功能说明。

表 1.3 常用的字符串操作函数及功能说明

函数名	函数功能	函数名	函数功能
size	查看字符数组的维数	abs	查看一个字符的 ASCII 码
char	把数字按照 ASCII 码转换为字符串	strcat	字符串连接
strcmp	比较字符串	strrep	替换字符串
strcmpi	忽略大小写比较字符串	upper	转换为大写
strncmp	比较字符串前 N 个字符	lower	转换为小写
findstr	在一个字符串中查找另一个字符	strtok	返回字符串中第一个分隔符前的部分
strjust	对齐字符数组	blanks	产生空字符串
strmatch	查找匹配的字符串	deblank	删除字符串中的空格



下面通过一些实例介绍部分字符串常用函数的使用方法。

1) 函数 size()

返回字符串的类型和长度。

```
>> university = '上海工程技术大学'  
>> size(university)  
ans =  
      1      8  
% 返回这个字符串是一个一维数组,字符串长度是 8
```

2) 函数 strcat()

连接多个字符数组,可使用字符串连接函数 `strcat` 或者 `strvcat`,或者连接运算符 `[]`,这里重点讲解 `strcat` 函数的用法。

`strcat` 函数为水平连接字符串,调用格式为

```
T = strcat(s1,s2,s3,...)
```

`s1,s2,s3` 等经过水平连接得到的是一个更长的字符串。如:

```
>> s1 = 'hello'  
s2 = ' world!'  
T = strcat(s1,s2)  
T =  
hello world!
```

特别注意:当 `s1,s2,s3` 等为字符数组时,所有字符数组的行数必须相等,每行相连组成新的字符数组;如果 `s1,s2,s3` 中包含单个字符串,则将单个字符串扩展成与其他字符数组相同行数的字符数组后进行连接。

3) 函数 strcmp()

调用格式:

```
T = strcmp(S1,S2)
```

比较字符串 `S1,S2` 是否相同,如果相同,则返回值 `T=1`;反之,`T=0`。如:

```
>> a = 'hello'  
>> b = 'hello!'  
>> T = strcmp(a,b)  
T =  
0
```

4) 函数 findstr()

调用格式:

```
T = findstr(S1,S2)
```

自动比较输入的字符串 `S1,S2` 的长度,然后在长字符串中查找短的字符串,返回其所在位置,如:

```
>> s = 'How much wood would a woodchuck chuck?';  
T = findstr(s,'a')
```




```
T =
    21
>> s = 'How much wood would a woodchuck chuck?';
    T = findstr(s, 'H')
T =
    1
```

1.1.3 元胞数组

MATLAB设计的这种元胞数组(Cell array)允许存放并操作各种不同类型和不同大小的数据,但存放和操作复杂字符串是元胞数组的一个主要用途。元胞是元胞数组的基本组成部分,元胞数组是用下标来区分的,单元胞数组由元胞和元胞内容两个部分组成,元胞和元胞内容是两个不同范畴的东西。用圆括号表示元胞元素,用花括号表示元胞数组的内容。以二维元胞数组A为例,A(2,3)是指元胞数组A中的第2行第3列元胞元素;而A{2,3}是指元胞数组A中的第2行第3列所允许存或者取的内容。如:

```
>> A = cell(2)
>> A(1,1) = {'hello'}
>> A(1,2) = {'world'}
>> A(2,1) = {1+2i}
>> A(2,2) = {9}
A =
    'hello'          'world'
    [1.0000 + 2.0000i]  [    9]
```

1.1.4 构架数组

与元胞数组一样,构架数组(Structure array)也能在一个数组里存放各类数据,使用指针方式传递数值。构架数组由结构变量名和属性名组成,用指针操作符“.”连接结构变量名和属性名。从一定意义上讲,构架数组组织数据的能力比元胞数组更强、更富于变化。

构架(Structure)是构架数组的基本组成部分,数组中的每个构架是平等的,它们以下标区分。构架必须在划分“域”后才能使用。数据不能直接存放于构架,而只能存放在域中。构架的域可以存放任何类型、任何大小的数组(如任意维数值数组、字符串数组、符号对象等)。而且,不同构架的同名域中存放的内容可以不同。

(1) 直接创建一个构架数组。例如创建保存电子信息系统专业方向和学生人数的构架数组:

```
>> information.direction = 'Electronic and information engineering';
>> information.studentnumber = 60
information =
    direction: 'Electronic and information engineering'
    studentnumber: 60
```

(2) 利用 struct 函数创建一个构架数组。struct 函数的调用格式为

```
S = struct('field1',VALUES1,'field2',VALUES2,...)
```

其中,field1、field2 等为字段名;VALUES1、VALUES2 等为对应的字段数据,必须是