

《计算科学导论》课程总结报告

姓名：杨阳

学号：2007010325

院系：计算机科学与技术学院

专业班级：计算机科学与技术2003班

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 分项评价 | | | | | 整体评价 | | 总分 | 评阅教师 |
| 自我分析  10% | 环境分析  10% | 职业定位  15% | 实施方案  15% | 评估与调整  10% | 完整性  20% | 可行性  20% |
|  |  |  |  |  |  |  |  |  |

一、引言

转眼间，在中国石油大学（华东）已经度过了三个月的时间，在这三个月的时间里，认识了我的五个舍友，结识了很多同学，也认识了许多优秀的老师。怀着对未来的期待与对计算机的热爱，踏上了在中国石油大学（华东）计算机科学与技术专业四年的学习生活。

先说一下对计算机导论这门课的感受吧，给我们上课的其他两位老师都说这本书的理论知识过多而且难以理解。而我上这门课的感觉大概也是这样的感受。可能有一部分原因是上大学之前没怎么了解过这个专业，知识储备有些不足吧。但上课的过程中了解到不同的知识，也不断的加大了我对计算机的兴趣。其实这门课的目的就是把我们带入计算机的世界，而我感觉老师的目的已经达到了，虽然我们无法将课上的知识全部吸收，但的的确确让我们对计算机的世界产生了兴趣。

二、对计算机导论这门课的认识、体会

1.整体认识

孙运雷老师和其他两位给我们上过课的老师都说过计算机导论这门课就是为了让我们对计算机产生兴趣，把我们引进计算机世界。实际上老师们也是这么做的，虽然书上的内容对我们现在来说可能还太早了，但老师们并不是硬要把书本上的内容都让我们理解，而是在讲解书本内容的同时还会发散一下，给我们讲一些计算机有关的事情。这样的上课方式我觉得是很好的。

记得刚开学时孙运雷老师问过我们两个问题：“为什要学计算机？”“为什么要来中国石油大学（华东）学习计算机”。这是两个发人深省的问题，在认真思考后，我有了答案。原来我来学计算机是因为它前途光明，有钱赚。而来中国石油大学（华东）学计算机是因为我的分只够考到这里。这么一想之后其实能让我们对自身有更透彻的了解，也对以后的学习有很大的帮助。

作为一个计算机科学与技术专业的新生，希望可以通过本科四年的学习，为自己在计算机方面的知识建立出一个整体的思维架构和知识体系，为以后的深度学习和工作打下坚实的基础。也通过计算机科学导论这节课，能够指引自己找到适合自己的学习方法和学习方向，能够让我对整个学科有一个深入的、全面的了解和认识。

2.学到的知识内容

(1).二进制

二进制（binary）在数学和数字电路中指以2为基数的记数系统，以2为基数代表系统是二进位制的。这一系统中，通常用两个不同的符号0（代表零）和1（代表一）来表示。

在计算机被发明的时候，也许是图灵机读写带上只出现两个符号启发了研究者，在当时的技术条件下，从便于元器件的设计和制造考虑，计算机的研制很自然地选择了二进制。后来的实践也证明了这种选择具有极大的优点。

二进制和十进制一样，是一种数制，它用于表示数的符号（数字）由于在书写中的位置不同而具有不同的值。二进制表示的数的符号只有两个，即0或1，其数值与十进制中的0和1相同。此外，与十进制不同之处在于二进制数是逢二进1.

（2）存储程序式计算机的基本结构和工作原理

作为计算机科学与技术专业的一员，平时敲代码的时候免不了要和计算机的工作原理脱不了干系，在导论课上也让我对其有所了解。存储程序原理又称“冯·诺依曼原理”（1946年提出）。将程序像数据一样存储到计算机内部存储器中的一种设计原理。程序存入存储器后，计算机便可自动地从一条指令转到执行另一条指令。现代电子计算机均按此原理设计。

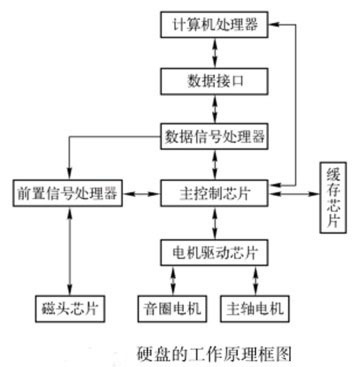
1、首先：把程序和数据通过输入输出设备送入内存。

一般的内存都是划分为很多存储单元，每个存储单元都有地址编号，这样按一定顺序把程序和数据存起来，而且还把内存分为若干个区域，比如有专门存放程序区和专门存放数据的数据区。

2、其次：执行程序，必须从第一条指令开始，以后一条一条地执行。

一般情况下按存放地址号的顺序，由小到大依次执行，当遇到条件转移指令时，才改变执行的顺序。每执行一条指令，都要经过三个步骤：第一步，把指令从内存中送往译码器，称为取指；第二步，译码器把指令分解成操作码和操作数，产生相应的各种控制信号送往各电器部件；第三步，执行相应的操作。这一过程是由电子路线来控制，从而实现自动连续的工作。

计算机必须具备五大基本组成部件：1.存储器2.运算器3.控制器4.输入设备5.输出设备。下面是电子数字计算机结构的一个简单框图：



（3）机器指令与汇编语言

汇编语言是任何一种用于电子计算机、微处理器、微控制器或其他可编程器件的低级语言，亦称为符号语言。在汇编语言中，用助记符代替机器指令的操作码，用地址符号或标号代替指令或操作数的地址。在不同的设备中，汇编语言对应着不同的机器语言指令集，通过汇编过程转换成机器指令。特定的汇编语言和特定的机器语言指令集是一一对应的，不同平台之间不可直接移植机器指令（Machine Instructions）是CPU能直接识别并执行的指令，它的表现形式是二进制编码。机器指令通常由操作码和操作数两部分组成，操作码指出该指令所要完成的操作，即指令的功能，操作数指出参与运算的对象，以及运算结果所存放的位置等。

机器指令一般可根据其功能划分为以下几类：

1. 控制指令；
2. 算术运算指令；
3. 逻辑运算指令；
4. 移位操作指令；
5. 传送操作指令；
6. 输入/输出指令；

汇编语言在可读性和编写程序时仍然是不能令人满意的，这导致进一步发展了高级程序设计语言。不过，由于高级语言在使用时通常还是要通过编译程序逐步将高级语言写的程序翻译成机器指令的程序，而这种翻译的结果往往不如机器指令或汇编语言写的程序效率高，所以，直到今天，不少工程师在系统软件的开发中还在使用机器指令和汇编语言。

三、进一步的思考

我觉得在计算机导论这门课中最能锻炼我的能力的就是分组演讲部分了。孙运雷老师给我们列出了1732道题目，其中包括安全，操作系统大数据，分布式计算，服务计算，互联网技术，机器学习，计算机视觉，计算机图形，计算机网络，计算理论，架

构，开发与运维，量子计算机，区块链，人工智能，人机交互，软件工程，数据安全，数据库，深度学习，搜索引擎，算法，网络，物联网，系统架构，信息安全，学习路线，用户界面，游戏设计，云计算等大类。看完这些题目，我觉得计算机专业是宇宙第一专业似乎一点也不夸张。，

我和我的搭档选择的题目是“声纹识别”，也称为“说话人识别”。

3.1什么是声纹识别

声纹识别（Voice Print Recognition，简称VPR），生物识别技术的一种，也称为说话人识别，包括说话人辨认（多选一）和说话人确认（一对一）。声纹识别就是把声信号转换成电信号，再用计算机进行识别。

多选一：即声纹辨认。一个声纹库包含了N个已收集的人员声纹特征，验证时，说一句话，即可找到库里的对应人员，比对时，验证语音和N个注册音频对比。家庭机器人如果要听声识人，就会用到声纹1:N。

一对一：即声纹确认。说话人事先录入过自己的声音后，验证时，只需要说一句话，即可验证自己的身份，比对时，验证语音只和一个注册音频的对比。微信和支付宝的声纹登录功能就属于声纹1:1。

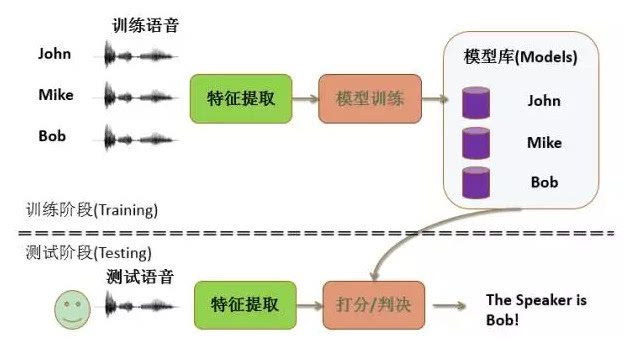
3.2声纹识别的发展历程

下面是从2019中国声纹识别产业发展白皮书中截取的声纹识别发展历程图：



从图中可以看到声纹识别的发展一共经过了五个阶段，声纹识别最早从17世纪60年代就开始萌芽，一直发展到现在，已经成为了一项相对成熟可利用的技术。

3.3声纹识别的步骤

下面是一张声纹识别的流程图：

一般的声纹识别过程是：首先提取语音特征，再把特征投入模型中训练，最后寻找分数最高或者最接近的结果。

目前针对各种特征而提出的模式匹配方法的研究越来越深入。这些方法大体可归为下述几类：

1. 概率统计

语音中说话人信息在短时内较为平稳，通过对稳态特征如基音、声门增益、低阶反射系数的统计分析，可以利用均值、方差等统计量和概率密度函数进行分类判决。其优点是不用对特征参量在时域上进行规整，比较适合文本无关的说话人识别

1. 动态时间规整

说话人信息不仅有稳定因素，而且有时变因素（语速、语调、重音和韵律）。将识别模板与参考模板进行时间对比，按照某种距离测定得出两模板间的相似程度。常用的方法是基于最近邻原则的动态时间规整DTW。

1. 最邻近方法

训练时保留所有特征矢量，识别时对每个矢量都找到训练矢量中最近的K个，据此进行识别，通常模型存储和相似计算的量都很大；

1. 矢量量化

矢量量化最早是基于聚类分析的数据压缩编码技术。矢量量化就是将若干个标量数据组构成一个矢量，然后在矢量空间给以整体量化，从而压缩了数据而不损失多少信息。Helms首次将其用于声纹识别，把每个人的特定文本编成码本，识别时将测试文本按此码本进行编码，以量化产生的失真度作为判决标准。这种方法的识别精度较高，且判断速度快。

1. VQ聚类方法

VQ聚类方法(如LBG，K-均值)：效果比较好，算法复杂度也不高，和HMM方法配合起来更可以收到更好的效果；

1. 隐马尔可夫模型

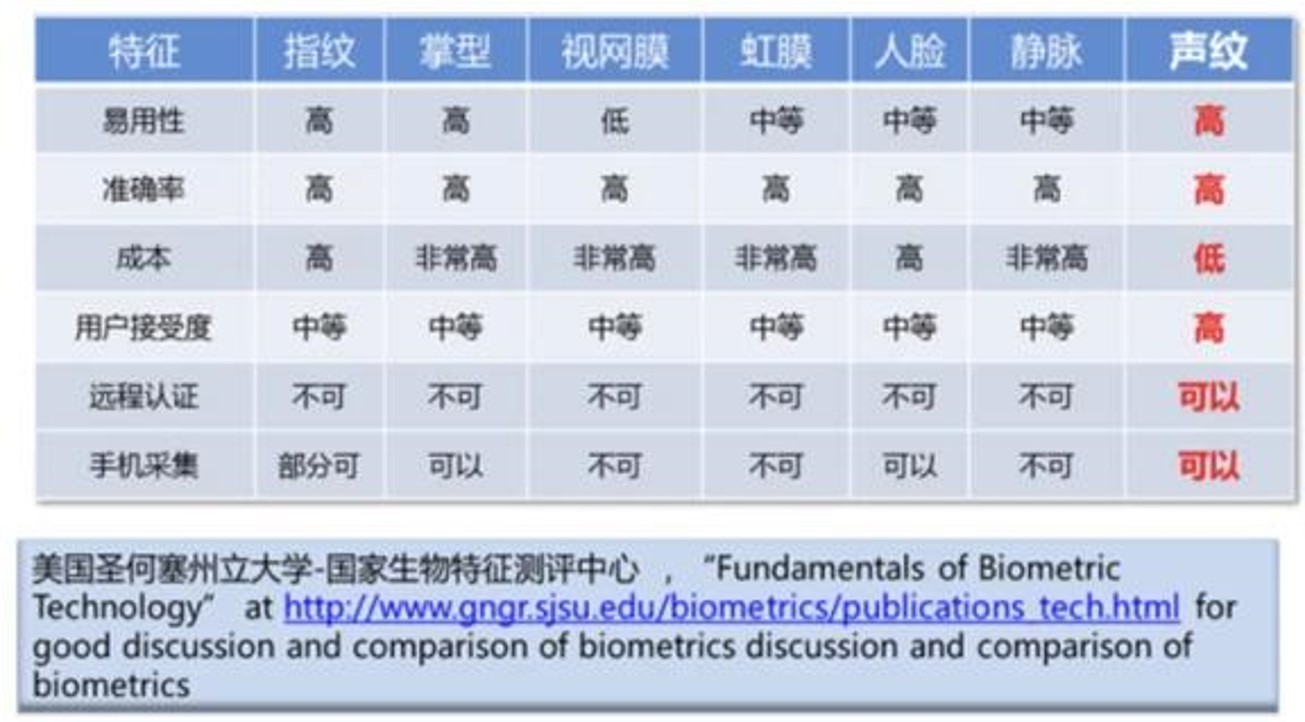
隐马尔可夫模型是一种基于转移概率和传输概率的随机模型，它把语音看成由可观察到的符号序列组成的随机过程，符号序列则是发声系统状态序列的输出。在使用HMM识别时，为每个说话人建立发声模型，通过训练得到状态转移概率矩阵和符号输出概率矩阵。HMM不需要时间规整，可节约判决时的计算时间和存储量，目前被广泛应用在工业领域，缺点是训练时计算量较大。

1. 人工神经网络

人工神经网络在某种程度上模拟生物的感知特性，它是一种分布式并行处理结构的网络模型，具有自组织和自学习能力、很强的复杂分类边界区分能力以及对不完全信息的鲁棒性，其性能近似理想的分类器。缺点是训练时间长，动态时间规整能力弱，网络规模随说话人数目增加时可能大到难以训练的程度。

3.4声纹识别的优劣

下面是一副比较声纹识别与其他识别技术优劣的图片：



从这张图片中可以看出，声纹识别在已知所有识别技术中其实是最好的一项技术了。在成本方面，声纹识别是所有技术中最低的。而在易用性，准确率以及用户接受度方面都是很高的，从这些可以看出，声纹识别这项技术有着很好的前景。

声纹识别的优点：

1. 准确率高：在理想情况下（环境安静、采集质量高、发音正常），声纹识别的准确率可以达98%以上；
2. 采集成本低：声纹采集对设备的要求不高，如今智能手机的普及，手机自带麦克风一般都能能满足采集要求，并且人在说话的时候就能无感采集，无附加操作成本；
3. 远程操作：只需要有麦克风，就可以远程采集声纹，并通过网络传输；
4. 不怕丢失： 生物特征的特殊优势，属于身体的一部分，不会像外部密码会丢失；

声纹识别的缺点：

1. 声音易变性：一个人在不同时间、不同地点下，说话音量、音调都会有所变化，这对算法要求比较高；
2. 环境噪音：噪音会干扰声纹的特征提取，使得识别准确率大大降低，所以一些声纹厂商会自研一套音频降噪处理算法；
3. 多人声音重叠：目前的算法技术还不能做到分离两个人重叠的声音，配合麦克风阵列，可以从源头就将不同发音方向的人声分离；
4. 说话时长限制：太短的语音提取不了足够的声纹特征信息，固定文本内容的最短有效时长至少是0.8秒，自由文本内容是2秒；

3.5声纹识别的应用与前景

（1）智能硬件：目前市场上常见的智能硬件有智能音响、机器人、智能车载等，具有查询（天气、机票等）、点播歌曲、设置闹钟、控制家电等功能。声纹识别可以让机器更智能，通过声音辨认发出语音的对象，结合其性别、年龄、历史偏好等信息，给出更加人性化和更智能的结果。

（2）公安司法：对于各种电话勒索、绑架、电话人身攻击等案件，声纹辨认技术可以在一段录音中查找出嫌疑人或缩小侦察范围；声纹确认技术还可以在法庭上提供身份确认（同一性鉴定）的旁证。在监狱亲情电话应用中，通过采集犯人家属的声纹信息，可有效鉴别家属身份的合法性。

（3）军队和国防：声纹辨认技术可以察觉电话交谈过程中是否有关键说话人出现，继而对交谈的内容进行跟踪(战场环境监听)；在通过电话发出军事指令时，可以对发出命令的人的身份进行确认(敌我指战员鉴别)。该技术在国外军事方面已经有所应用，据报道，迫降在我国海南机场的美军EP-3侦察机中就载有类似的声纹识别侦听模块。

（4）社保：声纹识别也可以在社保领域应用，利用声纹的远程采集验证的优势，可以在刷社保卡时，增加声纹验证，以确保一证一人。老人家每月领取养老金，需要进行生存验证，在很多农村地区，要求老人每年至少回老家一趟以证明还在世，对于一些已经跟儿女长居外地的老人来说，两地跑一趟，不仅花钱多，还劳累身心，验证也很不实时。如果声纹的身份验证可以普及开来，让老人家可以远程办理业务的同时，进行声纹注册和验证，对老人家们来说是极大的福音。

四、总结

通过计算机导论这门课的学习，我对计算机专业又有了新的认识。这门课为我打开了计算机世界的大门。计算机的未来是那么的丰富多彩。我已经等不及进入计算机的新世界了。

五、参考文献

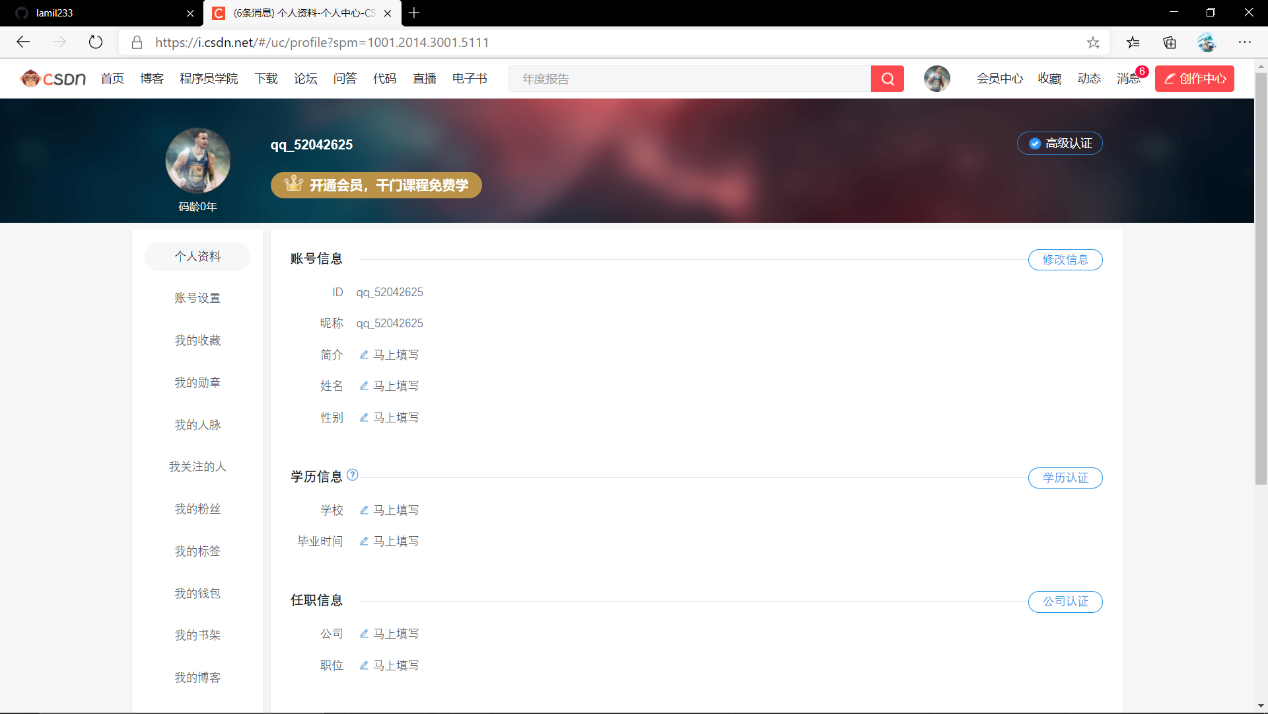
[1]简书Colleen\_oh声纹识别（说话人识别）

[2]2019声纹识别产业白皮书（百度文库）

[3]知乎 巧克力工厂的查理 声纹识别发展综述

[4]Kersta L G. Voiceprint identification[J]. Nature, 1962, 196(4861): 1253-1257.

[5]Pruzansky, S.. “Pattern‐Matching Procedure for Automatic Talker Recognition.”Jour

CSDN:

小木虫：



学习强国：



观察者：



博客园：



哔哩哔哩：

