



中國石油大學 (华东)  
CHINA UNIVERSITY OF PETROLEUM

## 《计算科学导论》课程总结报告

学生姓名：童思远

学 号：1907030124

专业班级：本研一体人工智能类 1901

学 院：计算机科学与技术学院

课程认识 30%	问题思考 30%	格式规范 20%	IT 工具 20%	Latex 附加 10%	总分	评阅老师

2019 年 12 月 25 日

# 《计算科学导论》课程总结报告

童思远

2019 年 12 月 25 日

## 1 引言

通过《计算科学导论》的学习，我对计算机科学这一门学科了解了很多，从比较科学的角度去认识和学习计算机科学，并且对如何将计算科学与人工智能的知识结合有了一定的思考。它既包含计算机科学与技术，即对计算机问题的一般研究；更是涵盖了包括计算机图形学与图像处理、数据库系统、人工智能和虚拟现实等更多更广的科学技术。并且通过对演讲主题 AR 增强现实技术的深入研究与学习，我对计算机视觉以及该方面的深度学习引起了极大的兴趣。

AR 技术如今火热，因此我想要通过进一步了解计算机视觉以及计算机神经网络，思考、学习和探讨如何将计算机视觉和计算机深度学习与 AR 技术相结合，以使 AR 技术得到进一步发展。

## 2 对计算科学这门课程的认识、体会

### 2.1 认识体会

该课程以计算机为基础，内容涵盖了计算科学的基本知识、计算科学的现实意义、内容和方法论。包括了计算模型、二进制算法、数字逻辑、机器指令与汇编语言，算法、过程与程序，高级语言以及程序设计，还有软件、计算机组织与体系结构、并行计算等。还简单介绍了计算机图形学、计算机视觉、图像处理和人工智能等概念。该书本所介绍的，不仅是计算机的原理、方法、用途、意义和前景，还蕴含着一种计算机学科所需要具备的，用以解决计算科学中遇到的实际问题的思维和方法论：我们需要具备一种科学的认识，一套科学的理论和方法，一个科学的程序。

因此，当我们面临一个程序问题时，会形成这样的思维，我们会思考：这个问题需要用到我们所学的那些知识？该问题的数学本质和数学模型是什么？如何制定一个能够写到程序中，合理、详细、思维缜密的程序结构？当我们遇到与我们自己的设想不符的情况时，也会查找是程序的语法错误呢，还是我们在编程方面二段函数用法与算法思想本身有问题。这就需要我们思维丰富，思想开阔，敢于推翻自己的错误，把握好编程语言的细节。而最重要的，还是要吃透计算机的理论，把握计算机的原理，在实践的过程中，一步一步提升自己的计算机水平。

## 2.2 深入学习

计算机，从最原始的完成计算任务，到现在能够对数字、文字、图形图像、音频等各种各样形式的数据，通过向量、矩阵等各种方法进行运算、处理。它通过硬件与软件的结合而具有自动高速的输入、处理、输出以及存储数据的强大能力。不论是数字计算机、模拟计算机、还是数字模拟计算机，各种类型的计算机虽然会在规模、用途的专长、性能、结构和预算处理原理方面有所差异，但它们运算速度快、运算精度高、具有记忆能力、具有强大的逻辑判断能力和存储程序的优势和特点都是大相径庭的。有赖于这些特点，计算机极大的增强了人类认识世界、改造世界的能力，在国民经济和社会生活的各个领域有着非常广泛的应用，对人类科学技术发展、生产生活等各个方面都产生了不可小视的深远广泛的影响。

然后来说一下我较为感兴趣的计算机图形学、图像处理与模式识别。计算机图形学处理的是由线条、点阵数据及数学模型生成的图形。数字图像处理是指将数据和几何对象的模型变成图像。第三个概念是计算几何，计算几何是指怎样研究建立几何形体的数学模型的学问，如三维立体造型、曲线曲面的构造与拼接等。

在上课的时候，老师讲到，计算机图形学是利用计算机研究图形的表示、生成、处理和显示的一门重要的计算学科，研究内容主要包括图形硬件、图形标椎、图形交互技术、光栅图形生成算法、曲线曲面造型、实体造型、真实感图形计算与显示的算法等。计算图形学在科学计算可视化、工业建模、游戏制作、生物信息、医药医疗、虚拟现实和增强现实、等各个行业都有着极其重要的作用。

熟悉图像处理的主要内容包括人类视觉原理、空间域图像增强、频率域图像增强、图像压缩编码、图像恢复、形态学图像处理、图形分割、图像描述、图像分析和图像理解等。运用于医疗、生物识别、工业自动控制、航空航天、卫星遥感、智慧城市、大数据处理等领域。最后说模式识别，模式识别就是通过计算机用数学技术方法来研究模式的自动处理和判读，把环境与客体统称为“模式”。模式识别以图像处理与计算机视觉、语音语言信息处理、脑网络组、类脑智能等为主要研究方向，研究人类模式识别的机理以及有效的计算方法。以上这些都是做好计算机图形学、图像处理的必要条件。

## 2.3 总结

总的来说，通过对这门课程的学习，大致了解了计算科学。正如书本前言部分中所说的：本书的内容生在引导学生怎么从科学、哲学的角度去认识和学习计算科学。书本知识对学生如何学好计算科学有着指导性的意义。在学习过程中，我明白了我现在所学有关其它计算机方面课程的重要性，并且具备了相关课程的一定知识基础。“一旦你选择了科学作为你终生为之奋斗的专业领域，就等于你选择了一条布满荆棘的路，一条充满艰辛的人生之路，一个有志于从事于计算科学研究与开发的学生，发须在大学的几年学习中打下坚实的基础，才有可能在将来学科的高速发展中，或在计算机产品开发和快速更新换代中有所作为。”因此，打好基础，开阔眼界，是我们的首要任务。并且在今后对于计算机，对于人工智能相关学科的学习中，我想将学着运用这门课程所教会我的科学的认识与学习方法。从事物的本质出发，从问题的根本挖掘，深刻思考，勤于实践操作，将问题从实际的角度解决。这，便是这门课程交给我的东西。

### 3 对演讲主题 AR 技术的研究

#### 3.1 AR 技术现存的技术问题

AR，增强现实技术使用相机和计算机屏幕结合虚拟和现实世界的图像技术 [5?]，现在人们已经熟悉。“1966 年，萨瑟兰，计算机图形和增强现实之父，第一次发明了增强现实设备，达摩克利斯之剑。[1]”可以说，它的发展历史已较为悠久。但现实是，如今的 AR 技术实际上远不及虚拟现实（VR）。因此，AR 技术发展的关键是找到技术瓶颈和突破口，也就是如何将计算机视觉和计算机神经网络与 AR 技术更好、更完善地结合的问题。

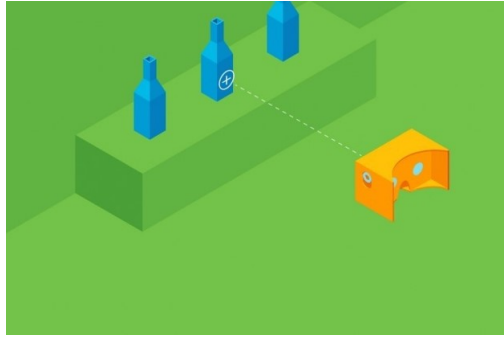


图 1: 识别物体

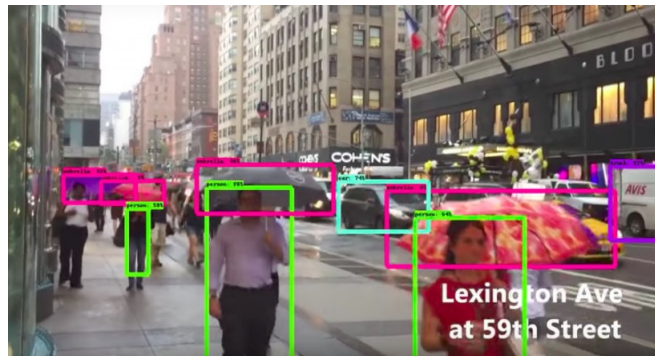


图 2: 识别与理解

不同于 VR 技术，VR 技术是在计算机中完全通过构建虚拟影像来，并运用这些影像模拟现实，达到与用户互动的目的。而 AR 技术是基于摄像头获取现实影像的，所以其中必须运用到计算机视觉的技术将现实中获取的影像进一步识别与理解。只有这样，当它在屏幕上附加上虚拟图像时，才能完成人机交互的功能。但是目前计算机用于 AR 中理解三维物体的能力还并不完善。

例如计算机无法明确区分地面，桌面与墙面的区别，在计算机“眼中”，这些都属于平面，甚至是可以无限延展的平面。因此，AR 通过计算机视觉达到对事物的理解是一大突破口。[4]

同样，与 VR 技术相比，AR 在显示器中能够呈现的视场角非常小。什么是视场角呢？就是图像能在用户眼中呈现的角度范围。视场角的大小与肉眼视野如果越相近，那么显示器的效果就越加优越。否则，使用 AR 设备，用户会认为他们使用的是望远镜，视野太有限。如果大



图 3: AR 设备产生的视场角



图 4: VR 设备产生的视场角

脑就无法通过直观的映射将 AR 世界看作真实世界的一部分，所谓的沉浸感也会化为乌有。由于 VR 完全是基于虚拟场景，所以，对于显示器呈现影像没有太大的视场角限制。但是 AR 是基于摄像头的，也就是说，要让 AR 显示器呈现图像的视场角增大，必须同时增大摄像头的视场角。并且也需要同时考虑 Tom Caudell 和 David Mizell 指出的，“要使得虚拟世界和真实世界更好地结合，对于增强现实的定位技术的要求在不断增强。[2]”因此，第二个问题就是扩大摄像头与显示器的视场角，并且增强 AR 的定位和角度测量能力。

其三，就是实现 AR 的自适应设计。AR 自适应设计的目的是使 AR 涵盖跨越三维空间的任意环境，尤其是快速适应一些先前未映射的场景，并进行虚拟图像的处理。有一个非常经典的例子，制作一款结合 AR 技术的射击游戏。要求游戏中的 NPC 从特定房间中钻出来，并且与用户进行实时互动。但是，如果 AR 系统不预先对环境进行映射处理，AR 系统连屋子里可能有另一个房间都不知道。因此，实现 AR 的自适应设计有助于增强 AR 系统的现实性，也是使其能够实现更加长效更加实用更加广泛深远影响的点睛之笔。

### 3.2 对于现存问题的分析

首先还是探讨一下 AR 系统对于物体的理解这一问题。如果单纯要让计算机通过摄像头识别和理解一个三维物体是非常困难的。又比方说，计算机可以通过摄像头获取一个啤酒瓶的图形信息，但是它难以探测到酒瓶内部还有大量空间，即无法理解酒瓶。更例如，计算机无法理

解酒瓶之中的啤酒平面无法高于瓶口所在的平面，树根不会长在天上，树叶不可能长在泥土里面。所以如果需要达到理解物体的目的，就需要较为庞大的理解算法体系，不仅对程序员是一种挑战，而且也无法达到“超越”人类智慧，真正达到人工智能的目的。那么，对于计算机理解物体这一步，就必须用到计算机的深度学习来构建神经网络。

AR 成像的关键设备是头戴式 AR 显示眼镜，它是通过将屏幕上现实的画面投射到用户眼中来实现信息输出的。但是往往该类设备体积与质量较大，不适于头戴，而如果缩小显示设备体积，那么其视场角也会随之减小。Oculus 此前表示，想实现真正的沉浸感，视场至少要达到 90 度，因此 AR 必须尽快翻过这座大山。不论是 AR 还是 VR，对于头戴式显示器的最大要求就是用户体验的沉浸感。就像如今影院的屏幕，越来越趋向大屏（即增大视场角）和 imax，用来增加观影人的沉浸式体验。但是如同我们所知道的，AR 显示器的成像是基于摄像头获取的图像信息，所以，扩大头戴设备视场角的一大关键也就是如何扩大其摄像头的“视野”。



图 5: AR 的应用

另外，就是计算机角度计算的问题。前些天我看到一个短视频，就是运用 AR 技术，将摄像头获取的人脸上附上猫脸的图，但是问题来了，当人的头低下俯视到一定程度时，计算机附上的猫脸就无法呈现了，它无法随着人脸角度的转变而转变图像。所以，我认为，这依旧需要使计算机对物体充分理解，让计算机明白，虽然人的脸俯下了，但依旧是人脸的一部分，依旧可以将虚拟图像即 AR 技术进行运用。

最后就是要分析 AR 自适应设计的问题了。就是说怎么样才能在现有设备的基础上，通过对计算机软件的改进，使其对先前未进行映射的物体快速做出反应。就还是比如说现实场景的射击游戏吧，首先是要靠计算机视觉和深度学习，对房间、墙面、门进行理解和快速反应，如果能做到看见即理解，理解即映射，那么，或许可以说是 AR 自适应设计的一个突破了。

### 3.3 对计算机视觉以及计算机神经网络的理解

计算机视觉是一种通过摄像头模拟人眼，并用计算机的算法模拟人类大脑并进行结合的生物视觉模拟技术。与 AR 类似，计算机视觉的最终目的是使计算机能像人和其他生物一样通过视觉观察理解世界，由此具有自己能够适应环境的能力。而 AR 技术，就是计算机视觉的一个应用和表现形式。[7?]

计算机神经网络，就是仿生人类大脑，从而构建起的具有判断、适应、决策能力，可以同时处理很多类型数据的算法体系。我们所说的计算机深度学习，就是构建计算机神经网络的大量神经元，以完善计算机处理事物的能力。先前所说的理解物体的能力，就是计算机神经网络的分支。[6?]

### 3.4 对于 AR 设备即技术改进的思考

以上发觉的问题可归为：

1、如何改进或者研发更加新型的 AR 设备，以使其在扩大视场角的同时不增加其体积和质量。

2、如何将计算机视觉与计算机神经网络和 AR 的自适应设计相连接并使之完善。

由于 AR 的虚拟影像的呈现是需要计算机的处理来实现的，但是又无法将整一台计算机戴在头上，所以需要研发一款体积小并且具有较强的图形处理系统的芯片。众所周知，如同笔记本一样，既要满足高性能，又要实现低散热是一件非常困难的事情，所以对于研发该芯片还要考虑散热问题。[3?] 而先前的较大体积 AR 设备主要问题并不是视场角不足，而是达到了可以实现沉浸式的设备处理器过大，体积过大。所以，研究一宽兼备散热小，体积小且性能高的芯片是成功推出便携，顺畅，优质的头戴 AR 设备的必经之路。

而说到将计算机视觉和神经网络与 AR 相结合这一问题，还是要训练出如同阿尔法狗这样的，可以对于棋盘上任意棋子摆放行为进行快速且最优化处理的神经网络。虽然棋子的摆放排列可以计算到无穷无尽的树木，但它终究是有局限的，而实地场景需要计算机考虑到的实物数据类型、计算思考方式要比下赢一盘棋复杂得多得多。

### 3.5 AR 的应用及市场方向

其实当今，像 Snapchat 和 Pokemon Go 等游戏与 app，在移动智能手机上已经掀起一波又一波热潮。尤其是在游戏一方面，对于年轻的游戏爱好者这一群体来说，不可谓没有市场。并且。2013 年以及谷歌眼镜的推出开启了专用 AR 设备市场。但是由于头戴式 AR 设备还存在体积质量与视场角性能的斗争，以及 AR 自适应设计的未完善，只能说 AR 有市场，但是还未有成熟的条件和恰当的时机。AR 头戴设备可以基于智能眼镜，其关键市场就在于能够完成对现实事物的实时模拟，在医疗手术、步行导航、以及室内设计家具摆放等方面有着非常大的应用。并且，如今娱乐、综艺行业发展得如火如荼，在表演、直播等行业也会有一定市场。尤其是直播行业，资金流水大，可以说如果要引入 AR 设备的话，是完全没有问题的。

## 参考文献

- [1] Clemens Arth, Raphael Grasset, Lukas Gruber, Tobias Langlotz, Alessandro Mulloni, and Daniel Wagner. The history of mobile augmented reality. 2015.
- [2] Tom Caudell and David Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. pages 660–661, 1992.
- [3] 包永刚. 理想的 ar 眼镜不会很快到来最大的阻碍是芯片设计. 2017.
- [4] 大壮旅. Ar 目前无法跨越的三座高山：视场角、理解物体和自适应设计. 2019.
- [5] 百度百科. 增强现实技术.
- [6] 百度百科. 计算机神经网络.



[7] 百度百科. 计算机视觉.

## 4 附录

- Github 网址: <https://github.com/TongSiyuan>
- CSDN 网址: <https://i.csdn.net//uc/profile>
- 博客园: <https://home.cnblogs.com/u/1869381/>
- 小木虫: <http://muchong.com/bbs/space.php?uid=19763821>

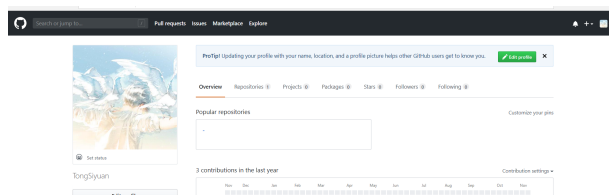


图 6: github



图 7: bilibili





图 8: 学习强国

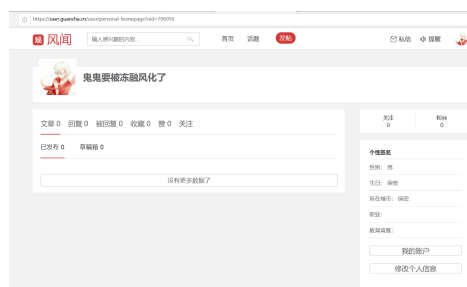


图 9: 观察者

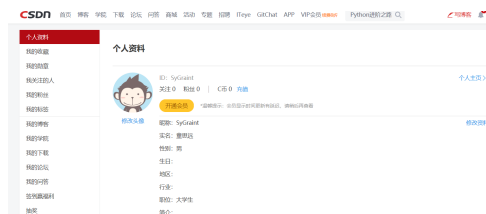


图 10: CSDN



图 11: 博客园



图 12: 小木虫