

山东大学暑假补习班学习总结

常琳

2016 年 7 月

这次能够参加山东大学暑期学校,是我绞碎了好多脑汁、写了一篇打动老师的报告之后,老师的特别恩典。虽说好多报告都听不懂,但是还算有所收货。

7 月 12 日

第一天并没有完全进入学习的状态,上午有一个讲仿真各种画风的报告,使用英文讲授的,可惜我英文不好,加之难以抑制的困意,理论方面并没有听进多少。另一个讲验证码的报告蛮有意思,倒不是说研究验证码识别很有意思,而是讲述人将一件枯燥的学术行为讲成了趣味实验探索。验证码识别的报告结束之后,有很多本科生提出了新的关于验证码识别方法的设想,虽然理论上不一定正确,但确实是跟着讲演者一起思考了。有人说真正的知识无法以有趣的方式讲出来,但是我觉得那一定是他自己不会,讲课中的“有趣”不一定是讲笑话,而是老师将知识讲活,最好是里面有自己的实验经历,这样讲起来能让大家跟着一起思考,明白其用途,试着想出新的解决方法。就像我老师说的,讲师在讲课之前,最好自己先去实践下,这样讲出来的东西才会深刻,也能够明白应该从哪个方面去指导学生。下午前两个是 3D 技术的介绍,我不是很懂,只知道之所以要重建 3D 环境,是为了实现人机交互,机器人实现各种功能的前提,是让它处于一个 3D 环境中。最后一个报告是用英文讲的,前一段是介绍了图像处理的硬件工具,后一段没怎么听懂。下午听 3 维重建时听到了两个与之有关的概念,一时好奇,大概的查了下,一个是 BRDF (Bidirectional Reflectance Distribution Function),它描述了入射光线经过某个表面反射后如何在各个出射方向上分布;一个是 SLAM (Simultaneous Localization and Mapping),它可以描述为机器人在未知环境中从一个未知位置开始移动,在移动过程中根据位置估计和地图进行自身定位,同时在自身定位的基础上建造增量式地图,实现机器人的自主定位和导航。

还有件事忘记说了,中午趁休息的时候去体验了一下 HTC 最新的 VR 头盔,好像是叫 HTC Vive,感觉棒极了,比半年前老师买的那台 Oculus 好太多,戴上玩了约十分钟,没有头晕,操作感还挺棒。在我读本科的时候,有段时间出了一款新的单机游戏,我着了迷似得从清晨一直玩到半夜,连每天出去锻炼的时间也从一个半小时压缩到了半个小时,那时我就能

不能让人进入一个虚拟的场景，玩家的视角就是原本玩家所操控的角色的视角，要想执行某个动作，必须要玩家先划出一个类似的动作，这样锻炼和娱乐两不误，多完美！不过闹心的是，我只是在想，人家都已经做出了产品。以前我总是被动地去了解新的技术和产品，跟着老师读研之后，一开始也是被逼着了解这个技术了解那门知识，后来变成自己会主动去查一些新技术，但是我的改变还远远不够达到质变。回想起自己这半年来根本就没有什么进步，而人家 VR 技术发展的却是这么快！果然是身处这个时代，原地踏步也会“被退步”的！

7 月 13 日

上午还是主要在讲 3D 方面的知识吧，和自己的方向似乎不太相关，上午听课时大脑也就没有高度集中地去，所能记住的就是老师强调的**数学在数字图像处理和计算机视觉中具有十分重要的作用**。这里也不难理解，从图像处理到机器学习，再到卷积网络，当中研究每取得一点微小的进展，都离不开相关算法的支撑。

中午进行了补觉之后，下午听得就比较带劲了，也可能是下午讲的比较有意思。首先是浙江大学的周昆老师介绍他们实验室这几年做的工作，他介绍了**计算机图形学在发型设计，室内场景扫描，面部表情与动画，3D 打印与物体表面非对称贴纸等四个方面的应用**。周老师的讲解通俗浅显，我这是第二次听他关于这几个方面的报告，第一遍的时候迷迷糊糊听了个大概，这次加深了一些理解。比如周老师讲的发型设计，亮点在于只是**从一幅图像中就能**

构建出一个完整的 3D 模型（图 1），如果以后能用这种方法构建其他物体，省时省力，简直太实用了。



图 1. 只使用一张图像建立的 3D 发型

而室内场景的扫描与构建，也是和 3D 发型的构建有相似之处的，都是只拍一张图就能重建 3 维场景。不过这里用到的方式是对扫描到的目标进行分割，将目标与已有数据库中的目标进行对比，选出数据库中最为相似的物体作为目标的 3 维图像（图 2），这里面应该是用到了[图像匹配](#)的知识。而发型设计中，首先要将头发分割成几部分，研究每一部分的发丝方向、深度等信息，其中发丝的深度信息又可以判断发丝在头皮上的位置来判断。不得不说想法确实简单而又实用！天天都梳头的我还真不一定能想到。当然，我觉得只凭一张图片就重建 3D 发型还是有点难度的，他们还是低估了发型的复杂度，尤其是女性的……

周老师还提到两个概念：[全自动](#)和[实时](#)，这也是我们老师所喜欢的，能够又好又快地完成一件事情，这本身就是世间最美妙的了。不过，有些东西就是因为手工才值钱的啊，像兵器这种东西，低端人士玩的都是机器锻造的，而大师手工打造的价值千万。[再一个就是数据集整理的问题，如果全部都可以自动完成，又哪来人工整理的可贵呢？](#)

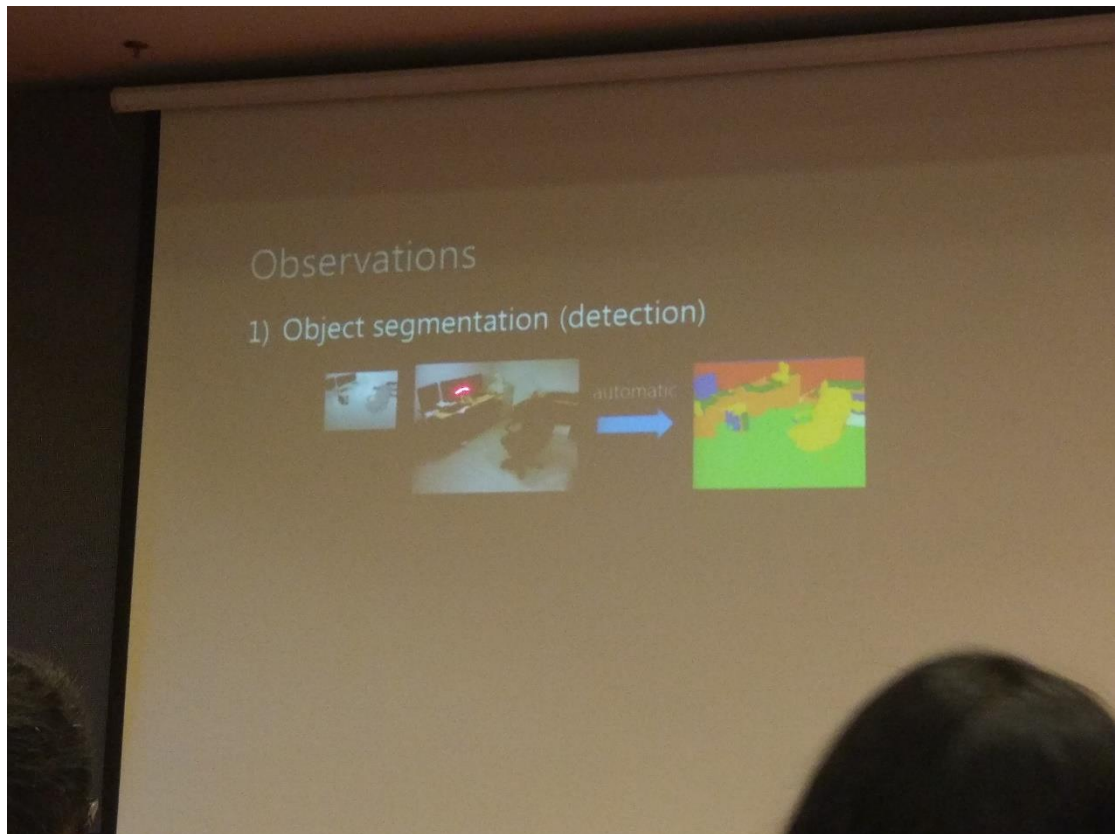


图 2. 三维场景的建立

还有位外国老师讲的对选中某块区域进行图像增强或者颜色修改等处理（图 3），PPT 中还给

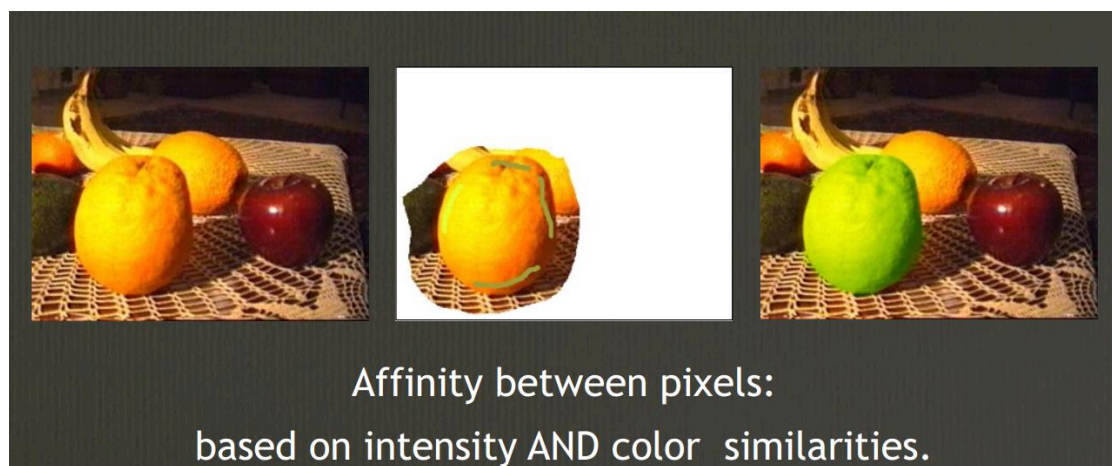


图 3. 局部图像处理

讲解了一下相关算法,虽然很想听懂,无奈英文水平不过关,最终只能对着人家的结果赞叹。大致描出一个区域,对该区域进行想要的处理,这个研究点应该是 Photoshop 等专注于图像处理的公司所喜欢的吧。另外两个报告是企业中的老师作的,内容偏工程一些,Pico 的那位老师上去作报告时不巧接到了财务处老师的骚扰电话,没有听太多。NVIDIA 的老师为我们讲解了创业公司初期的大致定位(to B),以及他们公司 GPU 的相关软件功能和技术支持(图 4)。以前总觉得硬件就是为软件或者技术提供一个运行的载体,听了老师的总结我才知道硬件研发的目的:降低开发门槛,使产品尽快落地。一切的研发都要围绕这两个目标进行。所以说,搞研究要有目标的,在实验之前,想要得到什么样的结果,或者说实验的目的是什么,自己要清楚,明白了目标,兴趣才会对接下来的工作有推动作用,不然的话,有限的精力会被兴趣分散。

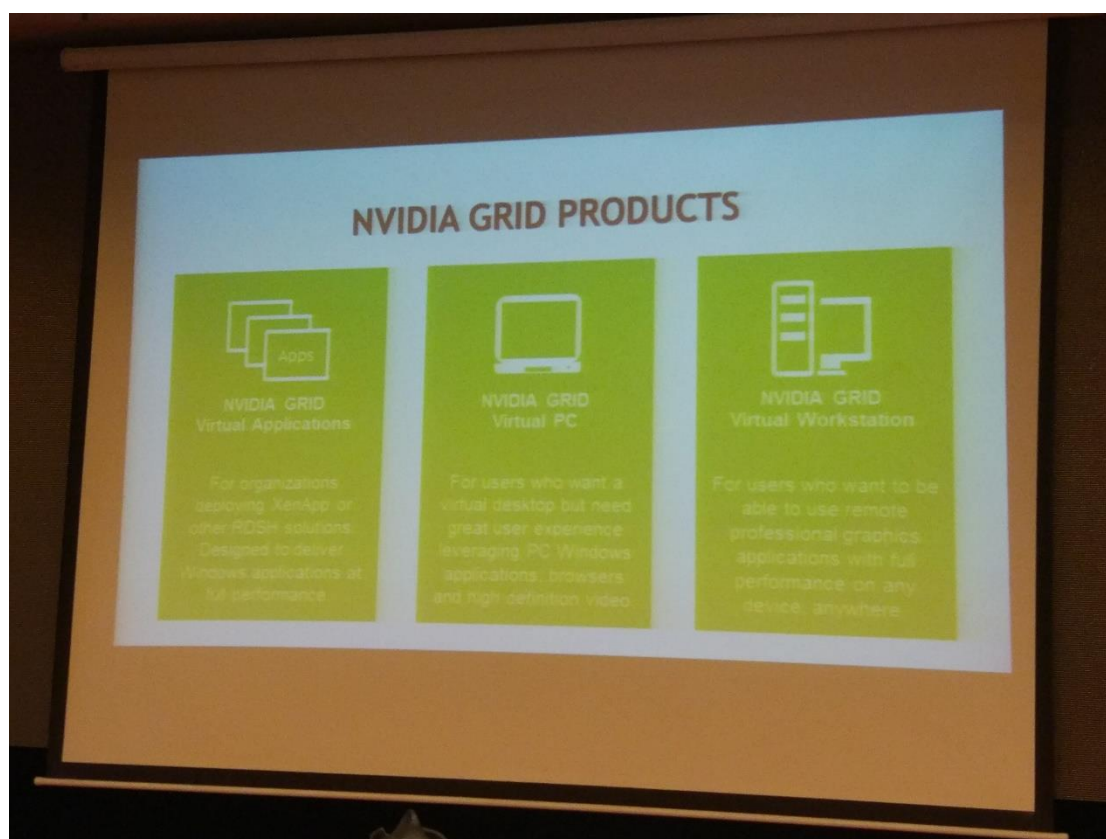


图 4. NVIDIA 相关产品应用

7 月 14 日

倒数第二天。

第一场报告的外国老师人很好，语速很慢，不过我也只是断断续续地听懂了一点知识。还有一个讲者是微软亚洲研究院的童欣老师，是关于 3 维模型重建的报告，这几天凡是做 3 维建模的报告中都会提到一个叫深度相机的设备，上网查了一下，这种相机可以获取目标的距离，也就是 depth，so 神奇。他这里讲 3 维重建时，尤其是物体在不同光照下的变化时，考虑到了物体的物理特性（有点像我们考虑的纹理特征），根据物理特性来判断物体在不同环境下（光照、亮度）其表面应有怎样的特点。还有发丝的动态三维建模，周昆老师的 3D 发型设计是根据一张图像做成的，而发丝的动态三维效果的建立首先需要多台摄像机从不同方位捕捉真实发丝的形态。他们的最终目的都是用手机扫描到图像后能够即时地显示 3D 信息，也就是面向普通用户的技术。此外，童欣老师还在报告中介绍了脸部 3 维模型重建在视频处理中的应用，一个有趣的例子就是给视频当中的人物做美颜：只要对第一帧图片中的人物做好美颜技术，余下视频中该人物的美颜效果都可以自动完成。这种技术也是先对视频中的面部建立三维模型，包括人物面部的皱纹也要捕捉到，美颜时将图片上的皱纹去掉，利用三维重建得到新视频里的目标人物将不会再有皱纹（好像是这样）。

下午贾佳亚老师那个报告很精彩，报告风格诙谐有趣很值得学习，内容也很摧残心灵碾压智商，贾老师的报告是按照年代讲的，虽然有些工作是在 2000~2010 年做的，但还是让人觉得难以超越。

贾佳亚老师的报告中提到了一种去模糊技术，是他的博士大约在十年前写出来的，我很想尝试，不知道用这种方法对许多张图片进行去模糊效果会是怎样的，好想知道这种技术能否对所有不清晰的图片都有效，这样就可以批量处理我们拍的不清晰的那些浮游植物的图像

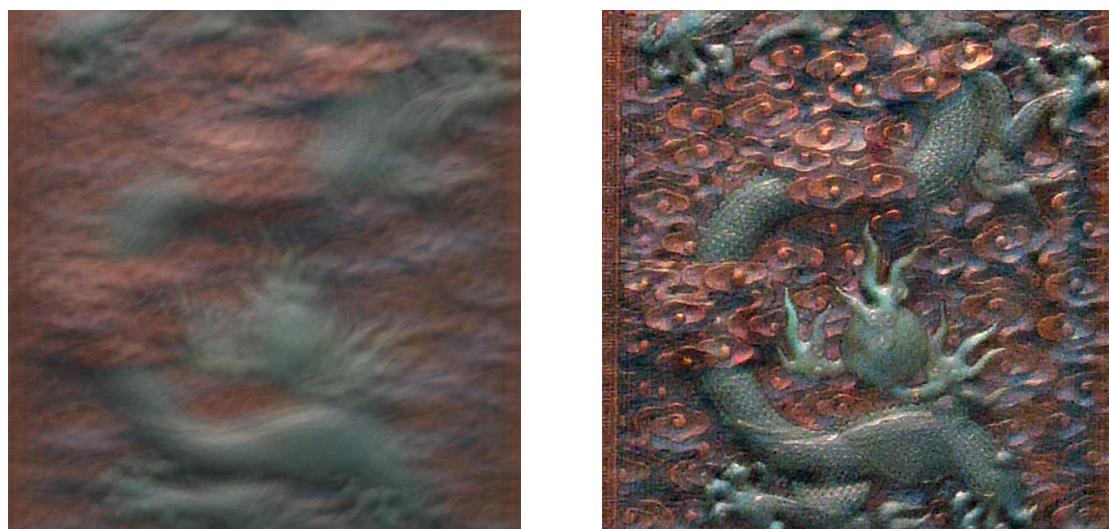


图 5. 图像去模糊

了。一想到我们现在的水平还不如人家十年前，真是连吃饭都觉得是在浪费粮食和时间。

还有一个不错的技术是图像融合，将某个目标与背景融合成一幅很自然的图像，我们的研究方向里暂时没有这个，不过作为闲时的消遣，写写这些代码也是蛮有意思的。

此外，说到培养学生的问题，贾老师不忘说一句，学生之所以能做这么好，“聪明是肯定的，再就是他们都很勤奋很努力”。好多导师提到让学生自己去探索，不过我觉得，能这么放心的前提也是学生够“聪明”，能自己探索的首要条件是懂得“举一反三”，“举一”是领会老师教的，“反三”即在此基础上去探索修正，发现未知。突然想起来老师居然收了我这么个笨学生，真是替老师心疼他自己。

另外我想说的是，**英语真的很重要**，英文的报告我就没一个能听懂的，也许和我抵触西方某些文化有关，几乎不看美剧英剧，听力就比别人差好多。以后一定一定要多看，哪怕是看动画片！

7月15日

暑假补习班的最后一天，今天的心情就相对比较轻松些，和师弟师妹们由不怎么说话变成了会交流几句。

今天的报告都是我喜欢的，像什么 AR 啦，游戏啦，自动驾驶啦。今天上午探讨了数据融合，特征融合以及降维，结合前两天提到的大数据变大数据问题，觉得自己的研究方向还是蛮有前途的，哈哈。

上午还观看了 Booming 科技的国产战争游戏战意的发布会，虽然画面感不如几年前发布的国内某款即时制武侠游戏，但贵在游戏引擎是自己研发的。明明与别人有很大差距，但是会以数倍甚至数十倍的速度去追赶，这种精神很值得学习。

下午是关于无人驾驶技术的讨论，技术细节倒是讨论的不多，但是从自动驾驶的整体需求去看，需要注意的事项确实很多，之前我曾想过做一个汽车识别的作业，就是扫描到一辆汽车后会显示汽车的品牌型号以及出厂价格，再做成一个手机上的 APP 或者车载软件，可以有效防止广大司机与豪车开斗气车。在无人驾驶领域，格林深瞳也考虑到了汽车识别，不知道其他公司是否也有考虑。以前我以为无人驾驶就是检测目标和障碍、然后进行躲避的问题，今天的讨论让我明白，无人驾驶是人工智能代替人驾驶车辆，为的是最大程度上地降低车辆发生事故的概率以及人员伤亡的数量，还需要考虑道路安全法以及道德等问题。从一种技术到一个成熟的应用，不仅是解决技术上的问题，还必须考虑法律、道德、安全性等等一系列问题。研究者对于自己所研究的事物应该始终持有高度负责的态度，本着实事求是的原则做实验，不能为了发文章而乱写乱编，也不能为了经济利益盲目夸大自己产品的功能，对于自己产品的使用应该充分考虑到消费者的身心健康以及公众安全问题。

最后，说一下今天各位 CEO、经理们讨论的个人培养问题，多数 IT 类公司看中的都是一个人的基本素质而非专业背景，从他们的报告中我总结了，基本素质包括良好的学习能力和身体素质，一颗平和心和耐心，当然还要有“爱”（热爱工作，热爱生活，热爱亲人），懂得团结协作。学生时代就要去多尝试，不要总担心失败或者被骂。