

## “可视计算”暑期学校参会总结

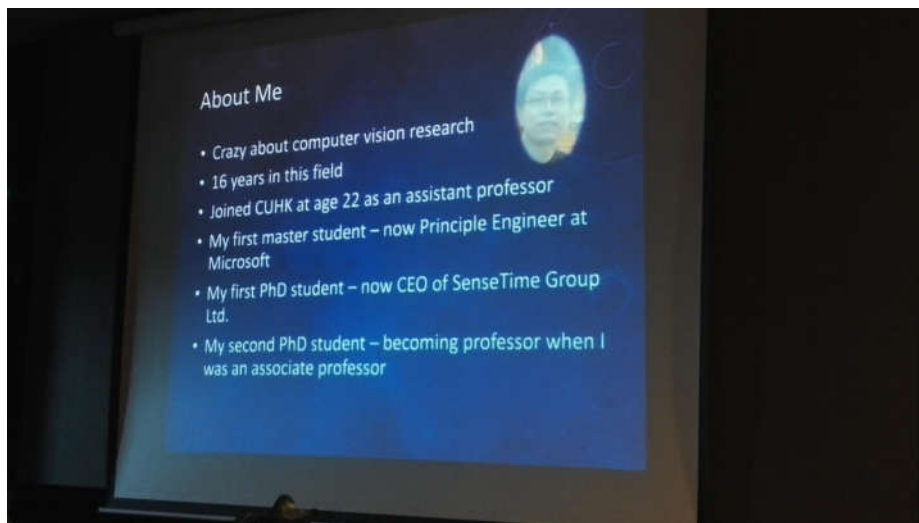
本次“可视计算”暑期学校围绕计算机图形学、视觉、人机交互、机器人、VR、AR 等进行展开。总共持续了 4 天，报告涉及内容较多，这里我只选择了几个来介绍。

### Jiaya Jia:

这次报告中，给我印象最深的就是来自香港中文大学的贾佳亚教授，他的报告让我对 Vision 这一领域有了新的认识和理解，报告的主题是“Two Goals of Computer Vision to Mimic and Surpass Human Ability”。

“计算机视觉的两大目标是模仿并超越人类的能力”，同时，贾教授说“视觉就是天生的智能”。这也就解释了计算机视觉为什么会和人工智能具有很大的关联性。

首先，贾教授简单地介绍了下自己（短短几行简历已让人五体投地）：



然后，贾教授简单地介绍了下什么是计算机视觉：计算机视觉就是要编写程序来解释理解图像，它是个看上去容易但是实质很难的任务。

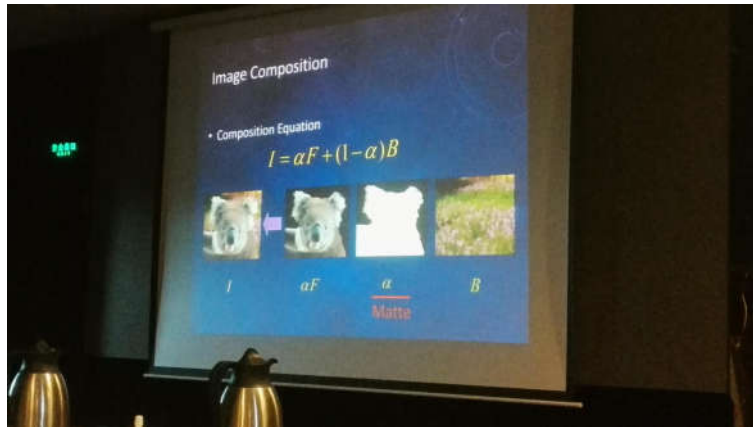
目前，计算机视觉的众多题目中一个新的任务就是使计算机能够模仿人的能力，使计算机超越人类。

之后，贾教授向我们展示了他们 vision 方面的相关研究工作。

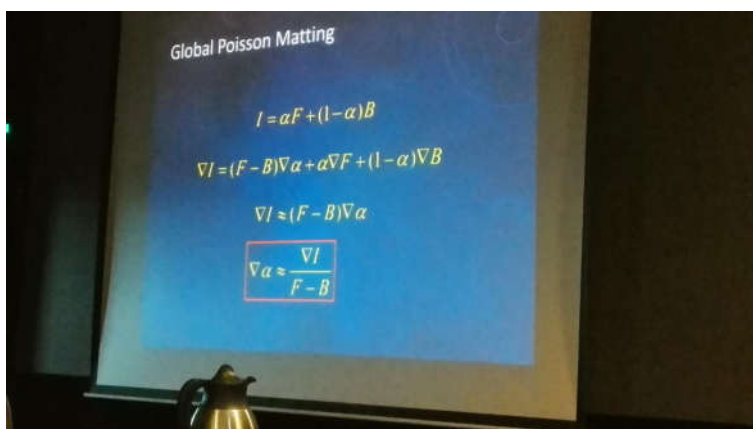
在 2009 年之前（贾教授读博期间），他做了一些图像修补工作（用于将图像中不想要的目标剔除掉并且修复剔除目标后区域获取新的图像）。虽然，在今天看来这项工作似乎已经不太具有挑战性了，但是在十几年前实现起来仍然非常困难。

还做了一些 Matting 的相关工作（已被微软收购，就是给图像中的目标换背景，首先获取它的蒙板，难点是发丝的部分）。

实现 Matting 的公式很简单，但是含有未知量：



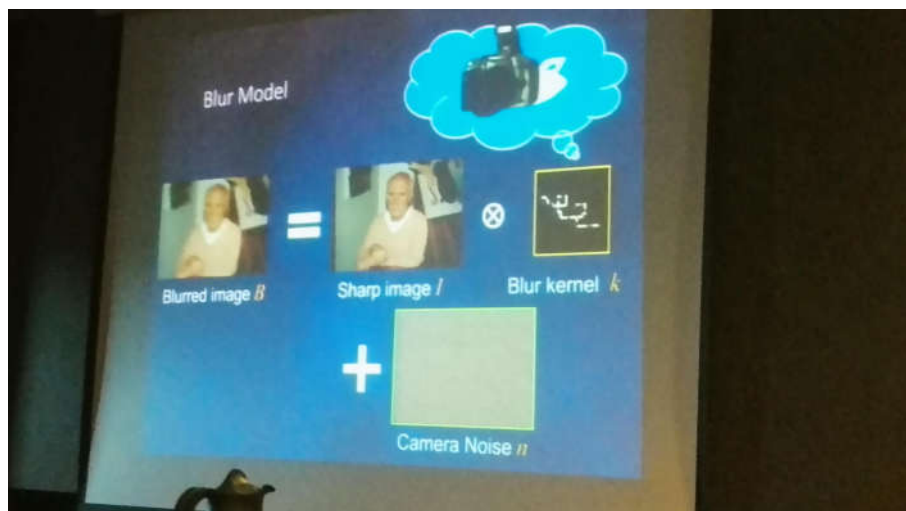
这里他们提出了梯度域的解法:



类似地，他还做了能使两张图像更好的融合以及图像增强的相关工作（和 TVB 合作过，可以放大视频，并保留更多的细节）。其次，中等视觉研究还做过通过输入图像和视频来获取深度信息。

贾教授最厉害的是在去模糊方面的成果。模糊问题来源于图像处理，但是 2007 年前的方法都是基于特定场景来实现的。

贾教授首先向我们展示了模糊模型：



然而，在实际情况中，存在原图像未知或者原图像和核函数都未知的情况。贾教授在此基础上，实现了去模糊，并且可以移除 size 非常大的模糊，甚至实现了快

速去模糊，平均处理一张模糊图像只需 4s。

还有用于解决一张图像中不同程度的模糊问题。

此外，贾教授提到了他们在美国旧金山开车时发现车在运动过程中看到物体模糊程度是呈发散状的，于是他们设计了算法用于解决这种运动产生的模糊（Vision 来源于生活，实用性较强，平时我们想到的一个点或者只是灵光一现的感觉都可能会和我们的课题紧密结合）。

还有一些看上去很完美但仍然存在着小模糊的图像。一张图像是否能达到我们所需的标准是由人为判断的。

12 年时，Deep Learning 开始一统江湖，贾教授开始用反卷积网络的方法来实现去模糊。

除此之外，贾教授还做了一些滤波的工作，通过保留并增强物体强边缘、弱化弱边缘来达到一种漫画的效果。

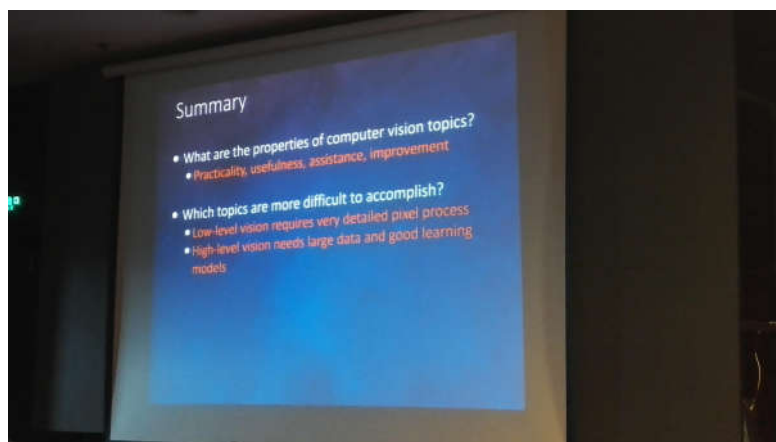
计算机视觉每天都会有新问题产生，贾教授的学生做了纹理移除的工作，并且还想出了一个应用就是解决漫画中的网点问题，实现代码也很简单，只需要一行（应该是封装好了），并且在 CVPR 上发表了。



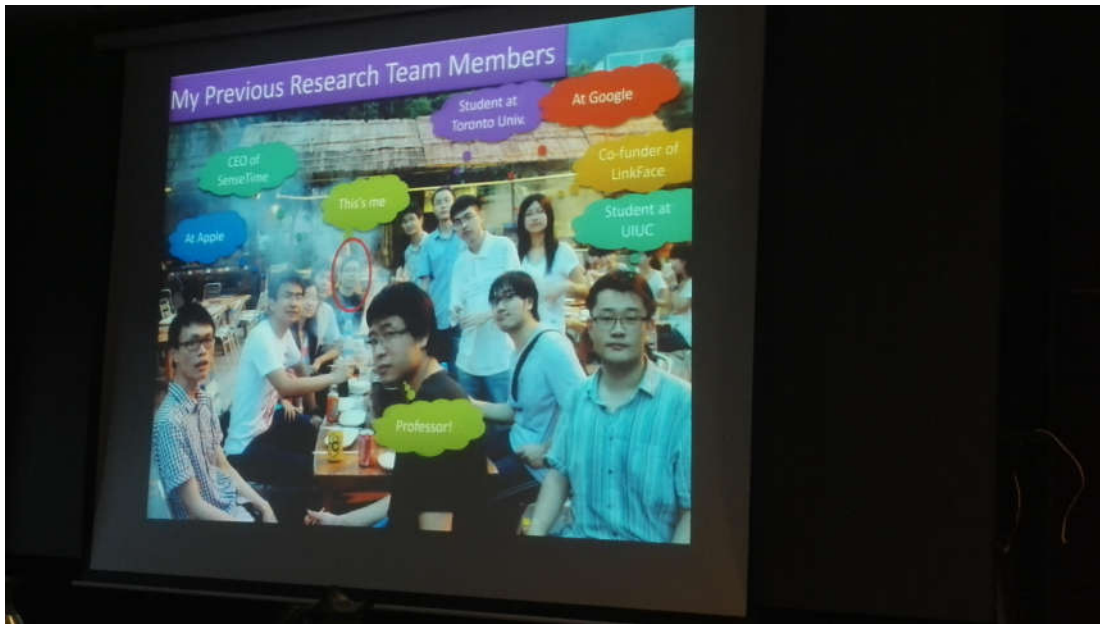
他用 DL 来训练 PS 结果图，最终结果和 PS 结果一样。

目前，贾教授在用 DL 做一些关于目标语义分割的工作。

最后，他对本次报告做了一个简要的总结：



他的团队，很厉害：



## Summary :

贾教授基本上列举了他的团队所做的所有工作，几乎囊括了大部分 Vision 领域。给我感触最大的就是他对研究工作的应用性和出发点的选择。针对于某一个特定的问题所展开一系列工作，另外就是他能从生活中发现研究点，明确什么需要做，可不可以做，做出来会有什么样的价值等问题，这是我们当下所欠缺的。虽然理论部分没有提到，但是报告确实精彩。除了感动，剩下更多的是佩服，佩服，还是佩服。

## Zhengyou Zhang:

下面介绍下来自微软亚洲研究院的张正有，他给我的印象也很深刻，因为他的报告在讲他几十年来一直研究的一个课题。

他报告的题目是 “Emotionally Intelligent Human-Computer Interaction”。

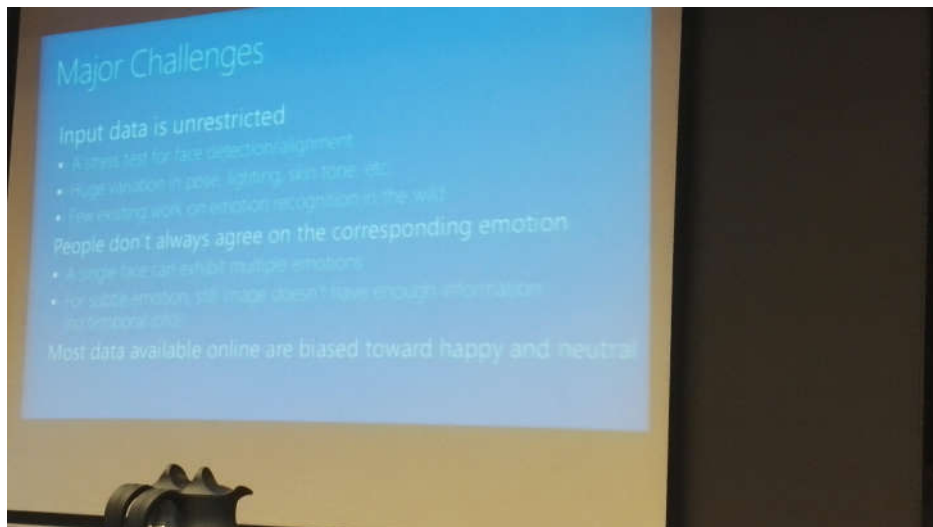
张教授首先介绍了目前微软在做的一个东西—Magic Mirror, 它有很多功能但最强大的功能就是可以读人的情绪。

然后，张教授介绍了他们团队在情绪判断方面所做的各项工作。

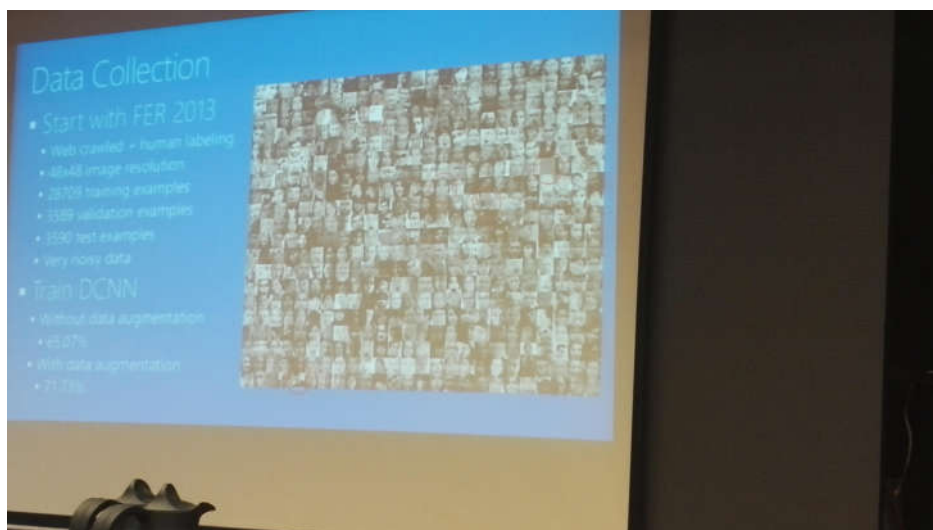
20 年前，他们采用了 2 层神经网络来判断人的情绪，这里用到了 Gabor Wavelets 的方法。



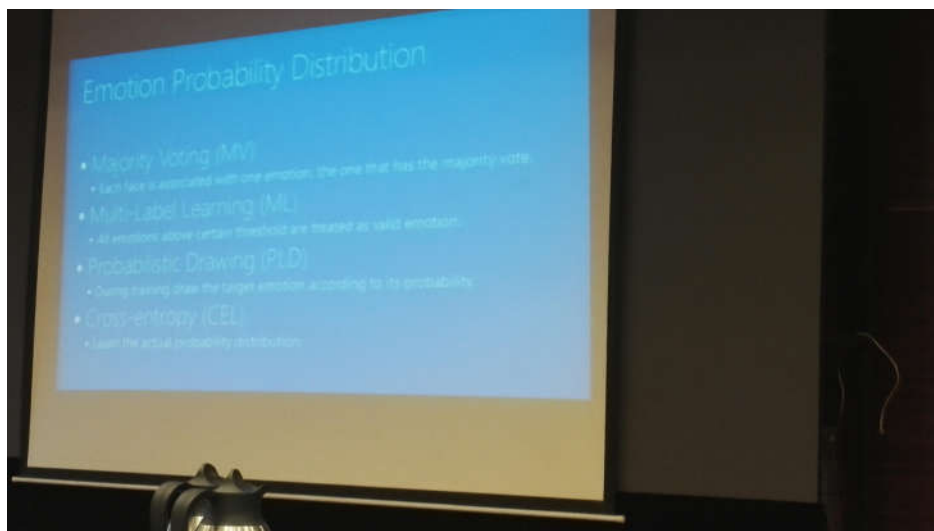
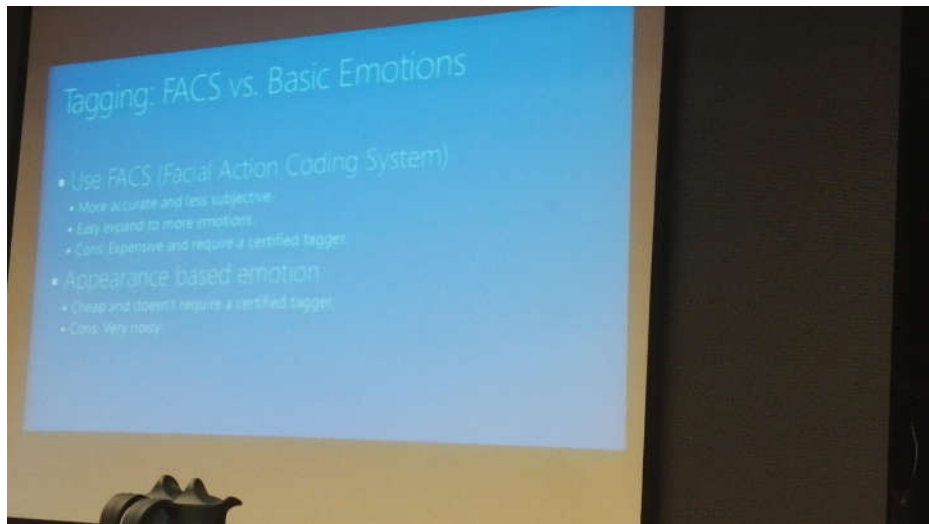
同时，他们还做了灵敏度分析的工作。  
现今，这项工作面临着众多挑战：



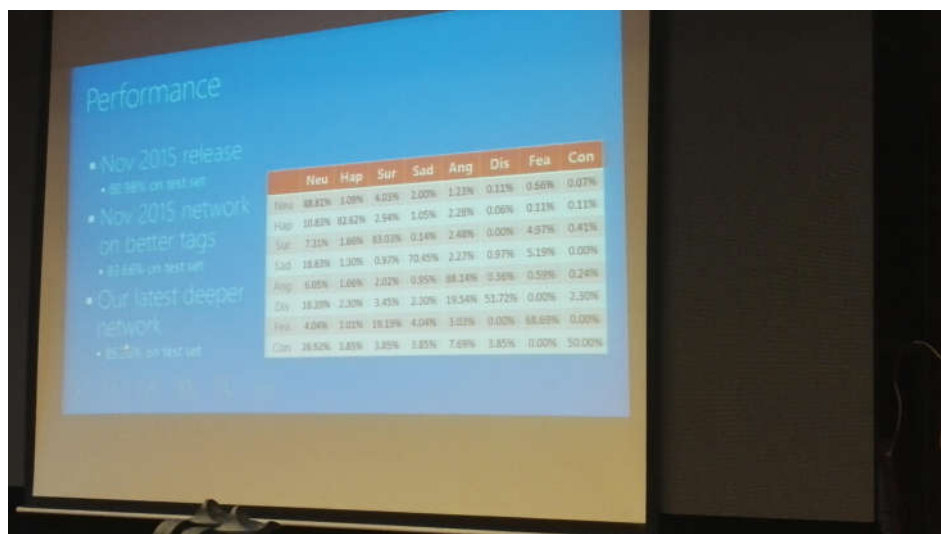
他们使用的一些方法（主要还是用 DL 的方法）：







最终得到的结果：6 个月内从 81%达到了 85%，不得不承认微软亚洲研究院的实力之强大。



Summary :

就我个人来说，张教授给我更多的是一种科研态度的鼓舞，几十年来始终坚持着同一个课题，不断的尝试、更改、精进。这种专注坚持绝非易事，而我们当下的年轻人，缺乏耐心、过于急躁，很多时候往往一失败就觉得这不行、那不行。我想，在未来的几年，我可能也会有这种想法，到时候再回来看看这份总结一定会有新的启示。

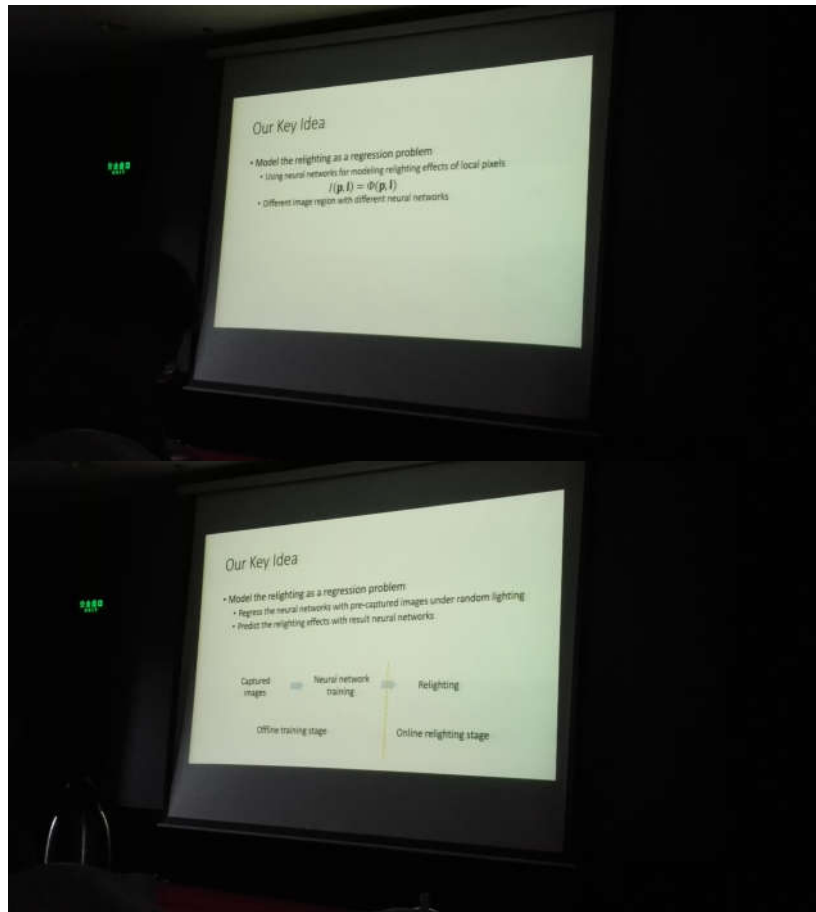
## Xin Tong:

微软亚洲研究院的童欣教授给我们带来了关于 3D 获取的精彩报告，他报告的主题是：“High Fidelity 3D Acquisition for Everyone”。

童教授首先介绍了传统 3D 获取方法的优势和劣势，然后列举了他们目前正在做的三种 3D 获取：获取环境、目标以及动态人物特性。

首先是重光照捕捉环境，童教授提到之前的方法都是在特定的光照条件下进行的，这里他介绍了他们的目标，既能使用数量较少的图片来得到一个令人满意的结果又能使所需设备简单化。

在经过仔细观察后，他们发现光源附近的像素点具有非线性特性，并且就此提出了解决方案：将重光照问题视为一个回归问题。



然后是捕获目标，目标的获取需要三维形状和表面信息，这里在未知几何形状和光源情况的条件下，通过输入视频序列来获取表面反射信息，而面临的一个挑战就是测量的平面可以由多种反射分布函数和光源产生：



他们采用了贾教授的方法，就是将这一问题视为一个去模糊的问题。

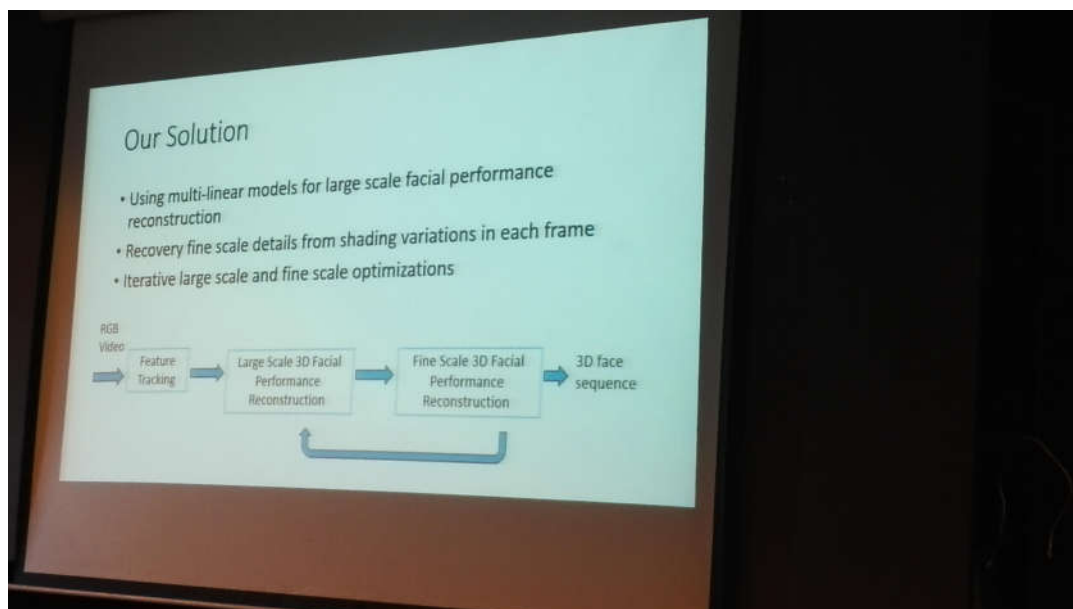
最终获取的结果可以把它放在一个虚拟的环境中进行渲染。

最后是捕捉人，捕捉人是一项高精度、高保真的任务。

之前的一些方法过于依赖设备并且需要用户输入。此外，对于 3D 头发的捕获缺少发丝运动的空间信息及时间连续性。

要想恢复高保真的面部信息，最大的挑战就是在未知光源等条件下获取运动、皱纹等信息实现完全自动的算法。

通过发现面部形状和表情都是在低秩空间中的，童教授他们提出了对应的解决方法：

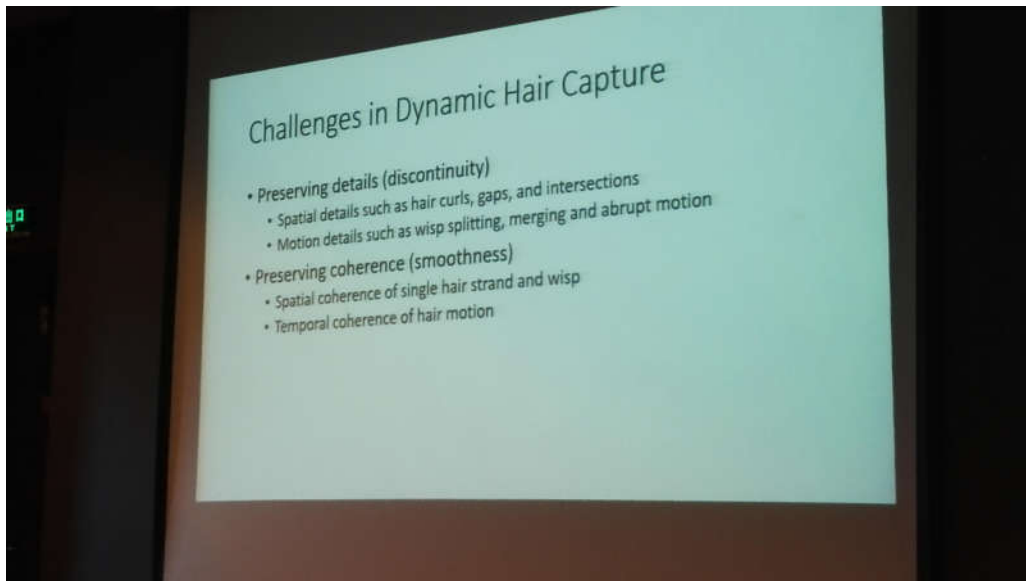


最终的结果可以用于给视频中的人脸加胡子、去皱纹等，并且只需要对第一帧进行操作，效果很逼真。

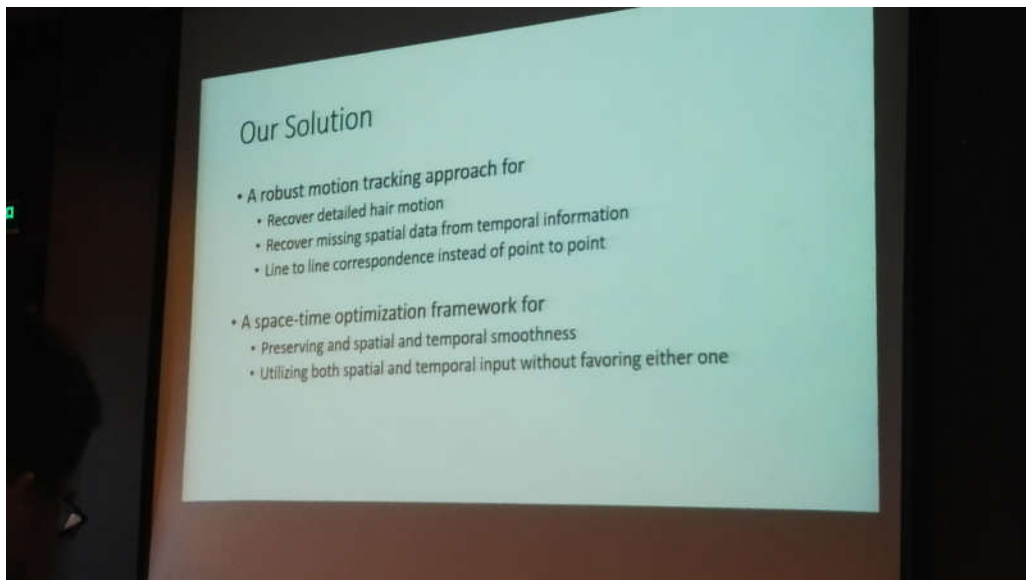
还有就是通过多视角视频序列对头发的运动进行 3D 重建，这里他们在假发模型四周放置了 21 台 GoPro（可以达到 120 帧 / s，有助于捕捉发丝运动序列）。

但是这项工作也面临着巨大的挑战：





对此，他们也提出了相应的方法：



## Summary :

童欣教授的报告着实精彩，但是更重要的是他的科研方法给了我一定的启示。我自己大致总结了一下他的科研方法：一是先明确需要解决的问题中有哪些难点；二是一个过渡段，这部分我不太好描述，我的理解就是找出问题的实质、根本，拿我的课题来举例，下一阶段我会总结下藻类特征，找出分类的关键点；三就是根据找到的实质制定相应的方法。当然这只是最基本的科研方法，我也在一步步摸索中。

## Yangyan Li:

来自山东大学的李扬彦给我们带来了有关验证码方面的报告，虽然他的报告给我的震撼力不是特别大，但是我觉得他的报告中解决问题方法的某个点很好，所以这里简单的介绍下。

李教授给我印象较深的就是他在提到一种验证码（把一张图像进行不同程度的扭曲，多

张图像组合成验证码，原始图像是最终正确的那个)，他说这种验证码有一个最简单的方法可以破解，就是检测直线。

我觉得这对以后的工作是一个很好的启示，就是把复杂问题简单化的这种思维方式，可能并不是所有问题都能够从简单思维的角度来思考，但是这未必不是一种尝试。

## My Feelings and Plan in the Future:

整体来说，暑期学校的报告都很精彩，这里我只选择了其中的一小部分简单介绍了下。除了 **Summary** 中谈到的对个别报告的一些体会之外，报告中还有一些其他的東西，我在这里简要的谈一下。首先就是 Panel 中提到的“考虑将来是什么”、“神经网络只是现在时，不是将来时”，我觉得这个需要我们好好考虑下，考虑以后要做什么样的工作、工作的实质是什么，而不是盲目选择一些当下流行的方法尝试。另外就是其中穿插了一个关于 PhD 的见解（“扩张人类的认知”），这就告诉我们要尽可能寻找还没解决的课题去做，并且要做好随时改变以往观念和思路的准备。其次，Panel 还谈到了“做什么并不重要，重点是要寻找自己喜欢的东西”，以及“这个行业需要耐心，随时可能陷入低谷，但是只要坚持，总会有成功的一天”。这些都给我一定的感触。

通过这次会议，我简单地制定了一下我的课题计划：首先我打算花一段时间整理下藻类特征，找出分类的关键点，然后我打算找找看是否有那种比较有针对性的相关文献可以参考，另外还想找找细粒度分类的方法是否可行，正好今年 CVPR 上有很多关于细粒度分类的方法，我计划回头看看。