# 视觉

2200013213 信息科学技术学院 柯宇斌

本文主要阐述本人对视觉的一些想法。

对视觉的研究起源于对人类的模拟，事实上，实现听觉、压力等传感信息的搜集往往是较为简单的，因为他们往往都能在一定程度上进行近似以至于达到可编码解码的地步。这也是为什么移动、声音识别等行为相对好实现。同时，在日常生活中，相对于崴脚、戴耳塞等这些让我们失去移动、听觉的场景，大多数人更惧怕的还是在黑灯瞎火的情况下移动，这也从侧面证明了，视觉对人的重要性。由此，想要实现机器人，视觉是必不可少的。

而视觉的困难点在于，它包含的信息很多，如果简单的使用摄像头，对每一帧的图像进行处理，最大的阻碍便是帧数和算力的矛盾。想要贴近人眼，就必须提高帧率，然而帧率的提高也意味着算力的大幅提升，这二者是难以同时满足的。而在课上，我了解到了仿照生物视网膜进行的高精度生物神经网络的模拟，这大大提高了计算的速率，也给我以深刻的感悟——科学是互联互通的。当一个问题难以解决的时候，可以去其他领域寻找机会。

然而，这节课最令我印象深刻的还是对拍摄技术的改进。时间丢失——帧与帧之间的信息完全消失，这种事情似乎已经成为人们所习惯乃至于认为理所当然的事情。甚至当我看到“图像和视频的重大缺陷：时间丢失”的标题时我还有点莫名其妙，这怎么能是缺陷呢。直到后来，我意识到，人类的视觉实际并不是离散的，那为什么照相机只能记录离散的呢。其后我便了解了以新光强定义为基础的新视觉。这也带给我一个启示：要时常探索自己认为理所应当的事情是否真的理所应当，突破口也许就在其中。

在我看来，视觉依然是计算机科学最艰深的一项，这是一个将大数量连续信息转换为离散有效可接受量级的信息的学科，具有极其深刻的意义和发展前景。