

## 报告总结

上次参加俞老师的老师的报告，由于我要去参加金工实习，所以只听了一场就走了，对于俞老师的老师 Minho Lee 所做的报告，我在离开之前做了这样一些笔记。

首先我的第一个关键词是 single easy optimal，通过对程序进行一点小小的优化，然后不断提升程序的性能，我记得 Lee 教授在中途提到了他们团队对于图像检索和识别进行了一点小小的优化算法使程序运行时间大大减少，但是由于我听力不是特别好，没能听出他所说的优化方法是什么。所以我的英语应该好好学啊！由此可以看出我们优化程序时可以侧重于单个部分和环节的优化，对程序的局部和算法进行优化，从而提高效率。

第二个关键词是 selective attention，对于一张图片，我们怎么能有选择性地关注图像的哪些部位，我上次听 valse 的网上视频时，曾经有个学者提出了这样的观点，这位学者针对图像检索出人的问题，他提出不管人的姿势怎么变，但是人的肩部的大致形状是一致的，肩部轮廓的特征具有更强的鲁棒性和泛化性，所以我们可以分析和提取出肩部的局部特征来检索出人的图像。在 Lee 教授做报告的时候，他放了一段视频，是一个机器人和人进行交互，机器人通过观察人的各种行为来判断人所要进行的各种活动和想法，写到这我的第二个关键词就要提出来了，机器人是通过什么观察并预测出人想要进行的活动的呢？我认为是 selective attention，机器人通过观察人的视线走向和手部指向的方向从而判断出人的行为方式的，换言之，如果此时人的这个动作是一个权重局部图的话，那么脸部、手部和眼睛的权重将远大于身体其他部位的权重，所以这引出了 selective attention 的概念，对于一张图片，我们可以关注于真正体现图像信息的局部区域，加强该局部区域对于整体的影响，减少其他区域对于整体的影响。还有个很有趣的事情是我原来试过微软的通过一张图片判断出照片中的各个人之间的关系，我觉得微软实现该功能的算法也是 selective attention，通过获取照片中的各个人之间的位置特征，姿势，表情和视线等局部特征做出判断。

第三个关键词是 perceptual memory, 知觉记忆，在这里我不是很理解这个概念，我有自己的一些理解，首先我写过 rnn，我对它的一些工作机制有一些了解，它是通过记忆一些特征来推测下一时间的行为和信息，简言之它是跟时间序列有关系的，在这里我认为 perceptual memory 可以用来描述图像和信息每个区域之间的联系和记忆，比如我通过记忆各个区域连接来记忆图片信息，通过感知提取图像的深层次特征来获取信息。

除了这主要的三个关键词之外，Lee 教授还提到了 MTRNN, SMTRNN 等比较实用的程序，俞老师自己写的 MTRNN 的程序已经发给我了，我已经看了一半了，其中也学习到了俞老师变成的一些技巧和提高效率的一些方式，比如使用指针来传递数据结构会大大加快程序的运行，还有一些就是一些很好的编程技巧。

关于 Lee 教授的这次报告，我的大致想法就是这一些。希望老师指正。

郑自强

2016年9月3日