**谈多媒体技术之现状**

学号：2016140256

姓名：魏嘉毅

要谈论的多媒体技术的发展，就首先定义一下“媒体”，这样也好为之后的论述铺垫。媒体本来就是传递信息的媒介，像是：文字、声音，图像，视频等等都是属于媒体的范畴。而我们要谈论的多媒体技术自然而然就是，对于我们上述的这些媒体进行处理的技术。而随着技术的发展，多媒体技术已经从以前简单的简单处理、压缩或是存储等，发展为了对于大规模的媒体数据进行处理，甚至是“教”会计算机对于媒体的处理方法，使得我们能够完成的任务更加复杂，甚至是伟大。例如：文字翻译，图片中语义信息提取，图像识别，图像检索，视频识别，视频检索等等。我们能够完成的任务越来越多样化，完成的效果也超乎想象，还有很多跨模态的任务也可以很好地被完成。

能让现在的多媒体技术在近两年间发展的如此迅猛，深度学习的帮助可以说是居功至伟。受制于之前的传统的机器学习方法的算法性能局限性，很多的领域里并没有办法将性能提升到接近于人类的能力。但是随着神经网络的出现，算法的性能得到了极大的提升，像是人脸识别的方法中，最高的准确率已经达到了99%，这已经是超过了人类的人脸识别能力。像跟更多的领域，都相比于传统的方法得到了非常高的提升，现在Google的在线翻译工具就已经开始使用RNN了，准确度已经翻译结果的近人程度都得到了巨大的提升，有兴趣的话可以去进行尝试，就连古诗都可以翻译成很好的英语，不得不令人佩服呢。

更多的我是想谈一下近两年神经网络的发展状况，毕竟现在使用NN的各种模型才是主流，所以我所了解的多媒体技术上的发展更多的都是利用神经网络得到的。虽然会有疏漏，但是我认为已经可以基本上囊括主要的技术发展了。其实神经网络的这个概念已经是很“古老”的概念了，而且当时的科技人员已经预见到了多层的深度网络对于拟合函数具有极强的表征能力，但是苦于没有优秀的计算能力并不能赋予实现。但是随着近些年英伟达这样的硬件厂商的革命性进展，当初的梦想照进了现实。刚开始时候的神经网络的是Hinton在2006年提出的一个7层的神经网络，这也就打开了神经网络的大繁荣。在神经网络的一开始出现就像是发现了新大陆一样，让科研工作者们纷纷发现了神经网络的巨大优势。在那时开始到现在的图像视频等相关的顶级会议中的论文都是关于神经网络的文章，而且也百花齐放的出现了很多的经典的模型结构或者是精彩理论，这些都对于神经网络的发展起到了巨大的推动作用。像是ReLU的激活函数，这个激活函数对于模型的训练有很与预见性的优势，之前的网络很大的问题就是梯度消失的现象，但是由于ReLU的导数具有很好的导数性质已经可以很好的传递导数，就很好的保证了反响传播的导数传递。再有就是残差网络的发明，由于残差网络的出现，极大地加深了网络的深度，毕竟由于网络参数的爆炸，如果没有残差的结构的话很难进行训练，但是残差的出现很机智的化解了这个问题，是我们可以大大的加深网络的层数，是我们的网络拥有更加出色的拟合能力，像是微软在ImageNet2016中所使用的网络结构已经达到了152层，正是得益于残差结构，这样就是的网络在获得很深的深度的情况下也可以方便于训莲。另外还有RCNN到Fast-RCNN再到Faster-RCNN这一系列的革新，简直就是引领了目标检测领域的走向，现在的最新的结构几乎都可以看到这一系列的结构的影子。对于文字方向的自然语言处理或是分词等技术我不是了解，单单就我的浅显了解来讲，这方面的技术也是增长迅猛并且许多的成果也是已经赋予应用，向我们之前提到的Google翻译其实就是一个非常好的例子了。

其实更想说的是各位科研工作者的对于多媒体技术方面的努力是如此不懈，他们在孜孜不倦的思考在想着革新，而不是简简单单的在想着为了竞赛的准确性提高一点点而在想办法的刷结果。他们是在踏踏实实的做实验，设计结构，在多媒体技术第一线的是这样的一批令人敬畏的工程师和科学家们，他们就是多媒体技术的“现状”。而更加令人，令我们欢欣鼓舞的是很多的接触的科学家和工程师是华人，我们华人在这个领域正在不断的贡献着不可忽视巨大力量。不单单在国内的巨头企业我们的人才在发挥着能量，即便是在世界巨头中，也是处处可见华人的身影。像是贾杨清，李菲菲这些名字都是华人之光也不为过吧。这一个个通过自己努力证明我们华人科学家工程师们的能力的人们，是我的榜样，我的梦想就是可以向他们一样，在自己的领域内可以有自己的见解和贡献。希望我自己也可以在未来的科研日子里成为多媒体技术这一领域里添砖加瓦的人，与自己共勉吧。