新生研讨课平时报告3

多媒体计算和其应用

——听魏骁勇老师关于《Multimedia Computing and Its Applications》之我的理解

一、对多媒体数据的了解和引入

魏老师刚开始从多年前国外的一个统计数据入题，Youtube上全网用户上传的视频都取出一帧，然后直线排列，最后的总长度与地球相比，地球在对比图中显得十分渺小。况且这数据还是十几年前的，还是一个网站的统计。由如今全世界的数据爆炸来看，如果再次统计，结果必然会更加震撼，必定远远超过人们的想象。不如更大胆地设想，也许我们已经无法统计这样庞大的数据。这正是对多媒体计算的必要性所在。通过先进高效的算法，大大节省人工检索信息的时间和劳动力，在信息数字化的同时实现信息管理数字化，推动信息价值的完全实现，使现代社会真正进入数字时代。

二、多媒体应用面临的现状和困境

非法信息，信息版权，人机交流……多媒体数据在计算机中的存储和处理方式与人类对多媒体信息的处理方式有着极大鸿沟。图片视频信息在我们眼中是平面的，多彩的，但在计算机中本质就是一堆0和1，如果人类要和计算机交流，仅用语言可能比较容易，但如果发出一个日常聊天中常用的表情包或动图，那么计算机就无法理解人的意思。还比如盗版内容检测，同样是盗版内容，图片信息、视频信息就比文字信息更难以监测。微信与QQ中的广告发布者通常使用图片和视频来发布广告信息，这样更不容易被拦截，原因就在于图像视频识别技术还不完善，相关软件无法准确高效地识别拦截。这都说明了多媒体计算还不太完善的困境。

要解决这个难题，就得把目光聚焦在人和计算机的理解方式上。魏教授说，人类在识别两张图片或多张图片的关联性，比如两张坦克和装甲车的照片，我们并不是浅层次地，只靠视觉上的相似性来判断推理。人有自己的经验和记忆，能储存和运用社会经验，所以人在做推理的时候是会联系早期经验来推理判断。就比如装甲车的图片上出现了军人，通过脑部的快速联系推理，我们很快就能判断这是军用汽车。

但计算机只能计算，只能通过图形图像的数据特征来判断，也就是说，计算机只能识别形状颜色这种表层信息，只能得到刻板结果，当图片中含有“语义信息”时，计算机无法像人一样与原先已有的印象产生联系，从而解读出其中的语义信息。甚至，在图形图像变换角度或改变形状时，计算机只能得出机械化的结果，而在人看来，图形图像几乎没有变化。

当然，要解决这种问题，最简单的就是将非文本信息利用文本进行标记，网络上大量图片、视频网站所运用的检索方式即使如此，可是，这种方式却不能保证所有的信息得到正确的标记和勘误。在我看来，多媒体技术并没有实现真正的突破。

三、计算机学习内容把握与方法借鉴

魏老师列出了我们不能忽视的一点——硬件工具的完成设计与制造。计算机科学不只是单单学习如何编程。很典型的例子就是《计算机导论》开始时的“洋葱”剖图，上面形象地展示了计算系统的各大层次，编程操作主要应用在软件层，但所有的一切都是建立在硬件层上的。这说明我们在学习计算机科学时，应该要打牢硬件基础，这样，在对计算机物理层次的理解就可以帮助我们在编写程序、解决实际问题时和计算机更高效地配合，还能根据硬件的不同，将一个程序调试，以保证在不同设备上都能运行良好，提高移植性和兼容性。

魏老师还结合自身经历给我们分享了学习计算机科学与技术的方法。他以任务驱动的模式启发我们，学习C语言应该多实践。实践才是学习最高效、最合理的方法。

所以我认为，首先应该对教材的框架有一个初步的了解，知道大体结构后再去结合实例，一个一个知识点地学习，而非按顺序啃书，记忆学习。在结合实例学习的过程中，我们不仅巩固了理论知识，还自学了很多新的知识、积累了许多项目开发经验！

四、多媒体计算的应用场景设想

（1）数据压缩方面的应用

多媒体计算机技术是面向三维图形、环绕立体声、彩色和全屏幕运动画面的处理技术。而数字计算机面临的是数值、文字、语言、音乐、图形、图像、动画视频等多种多媒体的问题，它承载着由模拟量转化成数字量信息的吞吐、存储和运输。

多媒体计算机技术是面向三维图形、环绕立体声、彩色和全屏幕运动画面的处理技术。而数字计算机面临的是数值、文字、语言、音乐、图形、图像、动画视频等多种多媒体的问题，它承载着由模拟量转化成数字量信息的吞吐、存储和运输。

在元宇宙实现中，对三维世界的建模压缩需要用到多媒体计算对图形图像的压缩，以便通过移动式设备来接入虚拟世界地图。

（2）监测非法图片视频方面的应用

多媒体技术实现理解图片含义，即使图片经过认为处理，模糊了其中含义。从而对非法内容进行拦截和监测，减轻人工筛查压力。

（3）美颜方面应用

如今，颜值焦虑盛行，通过多媒体计算技术实现对面部精细部分的分割识别和微调，做到不懂美容的人也能根据指导，对面部进行尽可能小的改变，以取得最美丽的照片效果。

（4）治安方面应用

多媒体计算技术大面积应用，通过天眼监控摄像头来搞清识别并自动比对通缉信息库中的人脸，实现无人追踪，提高社会治安

五、小结

我们身处一个数据爆炸的时代，一天接受的信息也许能相当于古人一生所接触的信息。正是数据时代，新的问题、新的矛盾也随即出现。文字到图像，音频到视频，太多数据等待被挖掘价值。而多媒体计算技术就是最好的洛阳铲——比人工更快更便利的选择。

所以，我们应该感谢多媒体技术领域的先辈，为时代发展做出不可磨灭的贡献！

最后，也应提醒自己，学无止境！现在应该打好基础，磨练出深入的数学思维，在实践中锻炼出应对实际问题的建模能力和创新意识。切不可一时自满，计算机技术的道路还很长，就像魏老师说的，人工智能发展了五六十年还没有成熟，但不可否认，人工智能的确取得了长足进步。

同时，也感谢魏骁勇老师的网上讲解，很遗憾没能与魏老师线下交流。