

# 国外的新技术新趋势

沈子鸣 1170301007

## 1. 移动设备上的面部识别技术

2017年9月13日, Apple公司发布了新一代iPhone手机iPhone X。除了卓越的屏幕设计以及强大的计算运行能力, iPhone X还一项吸引大众眼球的高新技术——面部识别技术。

人脸识别技术早在几年前就已经开始开发。在iPhone X之前的移动设备上的面部识别技术一直没有得到充分的开发利用, 一是因为移动设备需要高便捷性, 技术设备体积要足够小、功耗要足够低才能配置在移动设备上, 二是因为生物识别技术必须在安全性上有保证, 比如容错率在1/50000以下, 足以保护个人隐私和企业隐私, 三是生物识别要在短时间内完成, 如果“刷脸”解锁手机用了两三分钟, 这都是不满足用户需求的。

iPhone X上的人脸识别技术将传统的2D面部识别提升到3D面部识别。相比于传统的2D面部识别, 3D面部识别技术相当于重建人脸的3D模型, 不但有面部特征, 还增加了深度信息, 比如人脸的形状、头颅的形状、尺寸、长短等信息, 都可以用作面部识别的特征数据。

### 1.1 3D面部识别技术的三个解决方案

(1) 结构光。结构光是能够显示物体结构的光照算法, 技术原理为激光散斑编码。结构光通过将特定的光电信息投射到识别物体表面, 然后再由摄像头采集这些信息并记录处理, 通过物体表面光点信号的变化, 计算出物体的深度和位置等信息, 完成特征点的采集。结构光的优点在于分辨率较高, 并且在任何光照条件下使用效果都比较出色。光点越多, 采集出的结构模型越精确, 特征点越多, 但计算和识别时间会变长。现在的结构光已经由传统的单束光扫描采集发展到现在额多点投射扫描采集, 速度更快且更适合移动设备使用。

(2) 飞行光。飞行光(Time of light), 简称TOF, 技术原理为反射时差。这个原理相对更容易理解。使用TOF时, 使用光照射物体表面, 通过反射光和返回光时间的微小差异, 计算出物体和光源的距离, 从而得到物体表面的深度特征信息。显然, 飞行光对物体的精密识别不如结构光, 所以TOF得到的3D模型分辨率较低。但TOF的帧率较高, 可以快速识别变化较快的物体, 即使刷新信息, 它的算法也相对容易简单。

(3) 双目测距系统。双目测距系统的原理来自于双目匹配和三角测量。这中测距模式在生产生活中也有广泛应用。这项技术采用两个摄像头采集物体信息, 因为两个摄像头存在距离差, 采集信息后再采用三角函数, 就可以处理出物体的深度信息, 从而得到物体的3D模型。这种技术的分辨率和精度都足够高, 但由于需要双目计算, 在硬件成本上需要加入计算单元, 成本较高。而且这种测距需要场景中的自然光, 因此不适合在黑暗环境下使用。

### 1.2 iPhone X上的FaceID

iPhone X上的FaceID面部识别技术使用的就是结构光, 由结构光的点阵投影器、红外镜头以及泛光感应元件共同组成, 其中点阵投影器投射的红外光照射在人脸上, 再由红外镜头采集, 如果是黑暗、弱光场景的话, 泛光感应元件可以加强识别过程。不过苹果目前还没有透露太多有关FaceID的技术内容, 只是含糊地说在人脸上投射30000个红外点并建立特征库, 不受光照影响, 安全性要比

指纹识别技术高很多等。

## 2. 移动设备上的面部识别技术对生活的影响

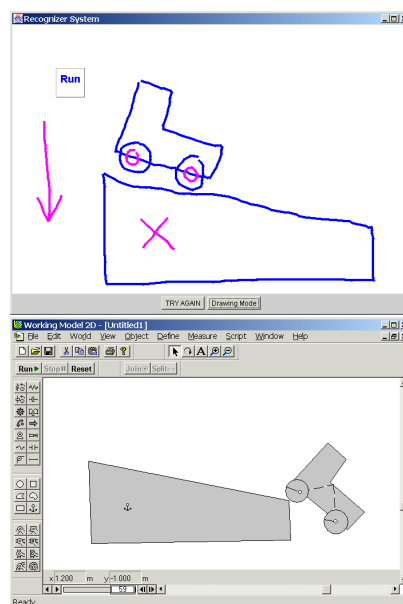
FaceID 所运用的结构光 3D 面部识别技术已然成为目前移动设备上最先进完整的最具可操作性的技术。在安全性上，FaceID 没有应用五官信息进行识别，而是只使用人脸的 3D 模型，这样即使用户化了妆、戴上帽子、口罩，仍然拥有一定的识别功能，另外，FaceID 加入了红外识别，这样很大程度上避免了无生命的人脸模型通过识别的可能性。

相比指纹识别技术，3D 面部识别技术的安全性更高。这种技术几乎没有学习成本，随看随用，是非常优秀的下一代移动设备生物识别技术的选择。这种技术使人们的日常生活更加便捷，使企业的隐私管理更加安全，并且为一众 APP 和软件的发展提供了机会，是人工智能应用在日常生活中的代表之一。

## 3. 国际一流大学的研究方向

### 3.1 Multi-modal Sketching of Human-Computer Interfaces in MIT

多模态描述。这个项目的目的是将人的语言整合到现有的被称为



“ASSIST”描述系统中。 ASSIST 是 Design

Rational 组织开发的一个高精度的描述理解和模拟工具。在机械工程设计当中，通常不会把合理设计记录下来，因为在实际过程中要去停下来记录每一步的设计决定是很不方便的。在电脑上直接设计可以有助于捕捉合理设计的过程。我们希望有一个基于电脑的设计工具可以记录下一个设计的全过程中的所有变化。遗憾的是，现有的基于电脑的机械工程设计工具都不是为了早期设计而制定的。大多数设计者会在开始用铅笔和草纸去设计，直到了设计近乎完成时才把这些放到 CAD 系统中。当一个工程师习惯了用纸笔去画图设计，而他也只会用手去画一些相对粗糙的形状时，让他再去使用一个精密的 CAD 软件，显得有点太难了。为此，Design Rational 公司开发了一款能让工程师就像手绘素描一样地画工程图的工程制图软件，并且能让人像工程机器一样和自己的设计互动，比如说通过看到工程师所画的模拟模型。该公司已经建成了一个该款

软件的早期前身——ASSIST，该软件能够使使用者在工程制图软件上简单地画图，并能看到自己所画的图的二维运动模型。

麻省理工学院的这项技术旨在创建一个多模态的环境，在这种环境中，达到输入模式之间相互的歧义消除，从而使使用者的目的更加明确。

在未来，智能空间中的基础设施可以加入其他的多模态，帮助明确使用者的目的意图。当人描述和谈论问题时，用户也会有时在描述时采用打手势的方式以明确自己的某部分描述，或者是表明自己的行为。智能房间可以提供这种数据，帮助用户明确意图，消除歧义。

### 3.2 An algorithm for your blind spot of CSAIL in MIT

麻省理工学院的计算机科学与人工智能实验室研发了一款新系统，这款系统的工作原理，是墙的边缘处的光线会受到其周围角落中物体反射的影响，在摄像机中可以看到这种影响变化。

麻省理工学院计算机科学与人工智能实验室（CSAIL）团队的成像系统，可以使用智能手机的摄像头，利用光线反射的信息来在隐蔽的环境中侦测物体或人，并且还能测算他们的速度和轨迹，这些都是实时的。

大多数观察障碍物周围的方法都设计特殊激光。具体来说，研究人员用相机去观察在可见和隐蔽的场景中都能可见的特定观察点上，然后测量光线返回所需的时间。然而，这些所谓的“飞行时间相机”是很贵的，而且很容易被周围的光线分散，尤其是在户外时。

相反，CSAIL 的技术不需要积极的投射光线进入空间，并且凭着现成的消费级的相机，就能在更大范围的室内和室外环境下工作。

在半影的视频当中，CornerCameras 会产生隐蔽场景的一维图像。单个图像不是特别有用，因为它包含了大量“嘈杂”数据。但是，通过观察几秒钟的场景并拼接几十幅不同的图像，系统可以区分运动中的不同物体，并确定它们的速度和轨迹。

该系统仍有一些局限性。显而易见的原因是，如果场景中没有灯光，它就不起作用，如果隐藏场景本身的光线很弱，就会产生问题。如果光线条件改变，它也会被绊倒，比如场景在室外，云在不停地移动。随着智能手机质量相机的信号也越来越弱，因为你远离角落。

### 3.3 Snake robot searches for Mexico City quake survivors

卡内基梅隆大学研发的蛇形机器人，在 9 月 19 日墨西哥城地震后倒塌的公寓楼里搜寻被困的幸存者。提供的多节点 Snakebot 救援人员用视频饲料从两个不同的穿过废墟，Matt Travers 说，但没有发现任何幸存者，系统科学家在 CMU 机器人研究所。

大学研究小组使用一种直径为 2 英寸，长度为 37 英寸的模块化蛇形机器人，并将其拴在控制和电力电缆上。它的主体由 16 个模块组成，每个模块有两个半连接在相邻模块上对应的半关节连接。机器人的头部包括灯光和摄像机。机器人有 16 个自由度，使其承担大量的配置和移动使用各种步态。一些类似蛇的自然起伏，但别人利用机器人的独特的机械特性，比如它有能力卷。

除了城市搜索与救援、Snakebots 可以通过和周围的管道，让他们在监督和核电厂和其他工业设施检查潜在的应用。这些机器人也进行了考古探测试验。

### 3.4 Carnegie Mellon Method Enables Telescoping Devices That Bend and Twist

卡内基梅隆大学的研究人员已经找到了一种设计伸缩式结构的方法，这种结构可以弯曲和扭曲，使各种形状的机器人能够折叠起来，以便运输或进入狭小的空间，使机器人的手臂和爪子能够越过或围绕大障碍物。研究人员设计了一种算法，可以将目标形状包括曲线或扭曲，并设计一个可伸缩的结构来匹配。他们还创建了一个设计工具，使初学者甚至能够创建复杂的、可折叠的程序集。设计的可能性范围从实用到快速部署的庇护所到奇异的创造，例如有腿、头和尾巴的伸缩蜥蜴，它们很容易缩回。

7月30日至8月3日，在洛杉矶，研究人员将在计算机图形和交互技术的SIGGRAPH会议提出他们的发现，是世界上最大和最有影响力的会议在计算机图形和交互技术。

研究小组发现球形、环形和螺旋形望远镜是可能的。一旦设计者为结构选择所需的曲线，他们的算法就可以设计出一种伸缩结构，这种结构可以在不碰撞的情况下扩展或收缩，并且在嵌套的部分之间没有浪费空间。他们还设计了将几个望远镜组合成更大的组件的连接器。

尽管嵌套的部分可以有多种截面，他们专注于那些具有圆形截面，就像海盗的望远镜。他们指出，一旦展开，圆形截面就可以使每一个曲线段旋转，增加三维扭曲，否则将是二维形状。

研究人员设计了几个原型，包括一个可折叠的帐篷状的掩体，一个可重构的手臂和几个分支关节，并用3D打印制作了原型。他们在模拟中探索了大量的设计，包括手、可重构的车辆和模仿蜥蜴和其他动物的几种有机形状。另一个是机器人手臂和爪子，可以从一个紧凑的圆柱体，并达到和越过障碍。

### 3.5 Automating Application Security for Cloud Services (Princeton University)

网络服务在用户在线体验中起着核心作用。这样做时，这些服务常常收集大量有价值的、特定于用户的、有时是隐私敏感的数据。不幸的是，在许多面向客户端的应用程序中，任何部分中的漏洞都可能危及整个应用程序。为了解决这个问题，我们设计并实现了过去，保护系统数据存储从意想不到的数据泄露和未经授权的写即使在应用妥协的脸。

通过自动分裂（以前的共享内存空间）应用到沙盘。这些成分和类型的访问每个组件可以共享存储之间的通信通过限制，如后台数据库。为了限制组件最小特权，通过采用动态分析对开发商提供的端到端的测试案例学习数据和控制流关系数据库查询和以前的查询结果之间，然后坚决执行这些关系。

我们是作为在 Django Web 框架原型置换一滴。通过运行十一未修改的，现成的应用程序，在帕塞，我们证明其提供强有力的安全保证是正确执行 96% 的应用程序的政策很少的额外开销的能力。此外，在原型的网站的特定设置，我们也减轻跨站脚本（XSS）跨组件的影响，结合浏览器的 HTML5 沙盒技术与我们的自动分离攻击。

## 4. 计算机科学前沿给我们的启示

从麻省理工学院、卡内基梅隆大学、普林斯顿大学的近期对计算机科学前沿的研究中可以看出，现在的计算机科学(CS)人工智能(AI)发展已经成为必然的大发展趋势，不论是CS中的哪个领域中都运用了AI的技术，甚至AI已经运用在生产生活，和其他各个科学领域之中。

机器可以解放人的生产力，这条规律自第一次工业革命就开始改变着世界的



生产生活。现如今人类正在享受着六次科技革命的幸福成果，我们横向对比现在科技发展，当然可以大胆地说，第七次科技革命，就将伴随着人工智能的大规模爆发。

计算机的深度学习，使机器越来越能够模拟人的思维，模拟人的学习能力，从而使社会自动化更进一步深入，更加解放全人类的生产力。

此外，除了人工智能这个大趋势，我们也可以看出有一部分研究正向着人工智能的“边缘领域”发展。比如人机交互，这个领域的研究成果可以更加有效地帮助 AI 的发展。AI 的实质是使机器像人，这本身就是人与机的联系。比如麻省理工学院的人机交互研究中的多模态识别，便可以使机器能够更加容易地理解人类的自然语言，达到“人与机沟通”的目的。

于是，我们期望未来的人类科技能够做到让人工智能全面爆发，为人类服务，那将是一个自动化达到极高水平的社会。为此，我们应该全身投入到计算机科学的学习当中，深刻领会人与机之间的关系，人脑与类脑之间的关系，为计算机科学向人工智能更进一步做出贡献。

## 5. 相关专业及作用

- (1) 计算机专业导论。“大家好，我们继续学习大学计算机。”这似乎已经成为了一句调侃的话语。但实际上，计算机科学的初识便是在计算机专业导论这门课中。作为计算机学院的学生，看似，我们需要掌握的是 C 语言等高级语言，这些是我们大部分学生将来毕业就业所凭借的，但此外，我们要涉足计算机科学这个领域的话，就不得不从计算机原理开始学起。计算机专业导论先从计算思维开篇，深度渐增，为学生展示了计算机的基本原理。这些对于今后对计算机原理的研究，对操作系统的开发和维护，对高级语言的开发和维护都起到了纲领性的作用。
- (2) 离散数学。这是一门大二才开始的数学学科。在大一，微积分的重要性不言而喻，而到了大二，主宰计算机科学的数学便是离散数学了。离散数学是传统的逻辑学、集合论、数论基础、算法设计、组合分析、离散概率、关系理论、抽象代数、布尔代数、计算模型等汇集起来的一门综合学科，在计算机科学与技术领域有着重要应用。学好此门课可以提高抽象思维能力和严格的逻辑推理能力。
- (3) 数据库。数据库顾名思义，与现在的大数据发展趋势有着紧密联系，在信息化社会，充分有效地管理和利用各类信息资源，是进行科学研究和决策管理的前提条件。严格来说，数据库是长期储存在计算机内、有组织的、可共享的数据集合。数据库中的数据指的是以一定的数据模型组织、描述和储存在一起、具有尽可能小的冗余度、较高的数据独立性和易扩展性的特点并可在一定范围内为多个用户共享。在大数据时代，我们需要高效的数据库来存储海量数据。
- (4) C 语言。没有什么东西是 C 语言不能描述的。C 语言极其重要，这是目前使用率最高，最受欢迎的高级程序语言，而且，这是我们计算机学院学生毕业后的饭碗。如果说，二十年后不能忘的有两样东西，除了计算机机器级程序的执行，再有就是 C 语言了！在某些领域上，C 语言相比 python、Java、C++ 都略显笨拙，但 C 语言是最浅显易懂也是最万能的语言，如果不会 C 语言，就不能算是程序员。

- (5) 英语。英语的学习重要性不用多说，虽然清华大学的计算机科学水平已经达到了世界顶尖水平，但大多数的前沿论文发表都是在海外，尤其是欧美地区，我国的图灵奖获得者也只有姚期智一人，所以，为了更加方便的查阅英文文献，我们应该学好英语，为以后的海外求学打下基础。

## 6. 参考文献

- [1] 《微型计算机》杂志，2017. 10. 15，总第 696 期，《认清你的生物特征——移动设备上的面部识别技术》。
- [2] 麻省理工学院官方网站，MIT-research-Agent-Based Intelligent Reactive. <http://aire.csail.mit.edu/projects.shtml#mmsketch>
- [3] 麻省理工学院官方网站，MIT-research-Computer Science and Artificial Intelligence Laboratory(CSAIL)-An algorithm for your blind spot [http://www.csail.mit.edu/corner\\_cameras](http://www.csail.mit.edu/corner_cameras)
- [4] 卡内基梅隆大学官方网站，CMU-Computer science- Snake robot searches for Mexico City quake survivors- <https://www.cs.cmu.edu/news/snake-robot-searches-mexico-city-quake-survivors>
- [5] 卡内基梅隆大学官方网站，CMU-Computer science-Carnegie Mellon Method Enables Telescoping Devices That Bend and Twist-  
<https://www.cs.cmu.edu/news/carnegie-mellon-method-enables-telescoping-devices-bend-and-twist>
- [6] 普林斯顿大学官方网站，Automating Application Security for Cloud Services, Aaron Blankstein 和 Michael J. Freedman. In Proc. IEEE Security and Privacy (SP ' 14), San Jose, CA, May 2014. <  
<http://sns.cs.princeton.edu/projects/passe/>>