多模态智能感知决策技术

多模态机器学习的定义：

1. 模态的概念：信息的来源或者形式，都可以称为一种模态。例如，人有触觉，听觉，视觉，嗅觉；信息的传播，有语音、视频、文字等；多种多样的传感器，如雷达、红外、加速度计等。以上的每一种都可以称为一种模态。
2. 多模态机器学习通过机器学习的方法实现处理和理解多源模态信息的能力。目前比较热门的研究方向是图像、视频、音频、语义之间的多模态学习。

举例：

1. 智能医疗：当前新冠肺炎的传播性非常强，因此智能医疗诊断被很多人认为是医疗下一代的发展方向，而大家都知道智能的医疗诊断需要同时“望、闻、问、切”，并综合许多不同项目的检测结果给出诊断结果，包括图像、数据等。这就涉及到一个多模态的数据处理过程。
2. 内容测试：例如，判断一个视频是否真实违规，需要把视频中的图像、文本和语音分别提取出来，送到三个模型中去分析，这样三者之间是孤立的，无法得到之间的关联信息，从而影响判定的结果。

因此，多模态处理能力是当前研究的热点。除此之外，我们还在多模态的处理前端，加入了快速感知技术，目标是实现多模态数据特征的有效且快速的提取。并且在多模态的后端，加入了综合决策，实现多模态复合型决策，从而构建一个多模态快速感知决策技术的闭环处理方式。

该技术的必要性：

1. 面向未来技术应用：人工智能技术的目标是构建具有通用处理能力的智能体，能让智能体像人一样工作。而人是具有多模态处理能力的，比如人与人交流，是同时考虑语音、神态、表情等等，而当前的手机助手或者智能客服往往仅通过语音去与人交流，智能化程度往往较低。
2. 当前的技术需要：目前看来，任何号称“AI”的设备，大多只能分析一种模态的数据，或者把不同模态的数据传输到处理中心，再利用不同的模型进行分析，AI模型本身没办法理解任何多模态数据。这种“伪AI多模态”在处理一些问题时是有效的，只是在处理实时化和复杂处理判断上会有心无力。多模态机器学习即为解决这一问题提出的技术。
3. 业界友商和学术界当前的研究热点：谷歌的AI掌门人在2019年的NIPS上也比较看好多模态机器学习的未来；微软也投资10亿美元给OPENAI构建下一代具有多模态处理能力的AGI平台；而国内的友商也在这一块发力，例如腾讯和阿里都在最近的AI会议上介绍了其在MMML上的进展，并且一些多模态的AI创业公司也逐渐增多，也证明该领域是下一代人工智能研发的重点。
4. 技术的创新性：我们将快速感知和综合决策和多模态机器学习相融合，构建一个闭环的技术，这是业界当前还未提出的。

该技术的应用：

1. 在终端产品上的应用：除了刚才介绍的智能医疗。还可以在AI教育领域有一定的应用前景，例如在针对儿童的表达及演讲能力的提升上，构建一个app，在儿童录制完自己的一段演讲视频后，通过学习海量的优秀演讲者，提示儿童在表情、动作及语言上的不足，完成快速矫正，使得儿童能快速学习，这是多模态机器学习在教育上的一个应用的范例。另外，还可以在检测这块有一定的应用。例如，最近近日，蚂蚁金服在IEEE成功立项“生物特征识别多模态融合”IEEE国际标准。由于金融支付需要在极低的金融级误识率（低至千万分之一误识率）下，实现用户的精准识别，基于传统单模态人脸识别的生物识别算法无法满足此性能要求。由此出发，蚂蚁金服提出了多模态融合人脸识别技术，并成功应用在支付宝刷脸支付的业务场景中。在性能指标上，即便是相比于业界传统单模态人脸识别的最优性能，蚂蚁金服多模态融合人脸识别误识率性能指标保持了两个数量级的优势；与此同时，能够支持的人脸识别1比N规模也提高了一个数量级。
2. 在IOT设备中的应用：一个是智能家居，比如采集时间、图像、语音各种数据，在家中无人的时候就自动进入休眠模式，当有人注释摄像头时会主动询问是否需要帮助等等。家中有独居老人，也能采集老人的健康信息以及家中环境信息，当出现数据异常时自动给出告警等。另一个应用是健身指导，比如人在家中健身时，电视摄像头能自动提取健身者的人体姿态，手环能提取人的心率等信息，送到数据处理中心，自动提示健身者的身体情况和姿态的正确性，完成自动矫正和健康提示。
3. 在构建复合决策的智能体中的应用：这块的实际应用比如无人驾驶、多模态机器人、游戏决策等等。例如，专为夜间自动驾驶研发的多模态传感系统。夜间的无人驾驶是相对困难的，因为激光雷达等提取的信息在夜间会受到一定影响。因此，苹果公司研发该专利，设计一种多模态传感系统，可在自动驾驶车辆夜间行驶时，用于检测和识别物体。比如在和平精英这个游戏中，声音也是很重要的特征。
4. 普惠AI，多模态机器学习可以提升常见AI问题准确性。

里程碑和规划：2020年踏实多模态数据的处理能力，补充语音、三维等素材的处理能力，重点分析上述几种素材的快速感知和特征提取能力。

2021年针对多模态机器学习进行重点研究，重点研究…

2022年对快速复合型决策技术进行重点研究…