

# 南洋理工大学开发高精度手势识别机器学习架构

VR/AR

来源：映维网 作者 颜昶华

— 2020 —

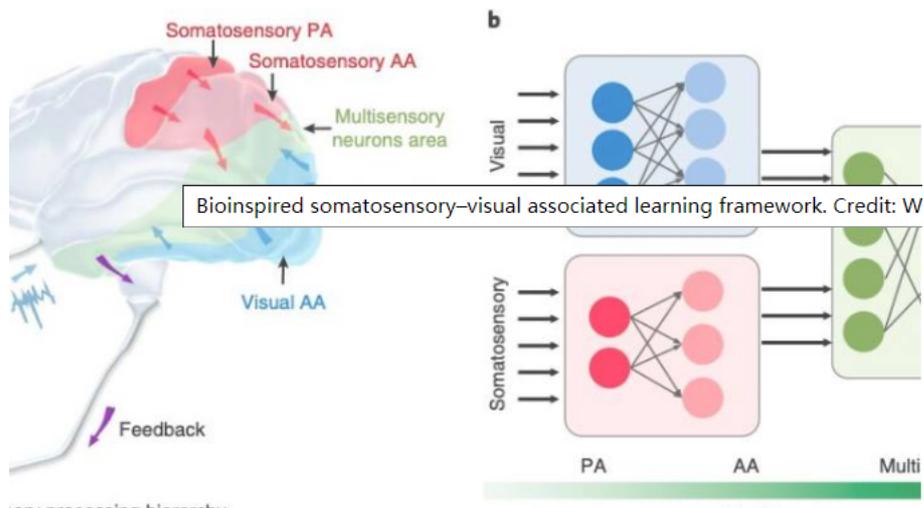
08/29

18:40

映维网

新加坡南洋理工大学和澳大利亚悉尼理工大学最近开发了一种旨在以高精度识别手势的机器学习架构，而方式是通过分析可伸缩应变传感器捕捉到的图像。研究人员已经将论文发表在《Nature Electronics (自然电子)》中，并表示新架构的灵感来自人类大脑。

南洋理工大学的陈晓东教授表示：“我们的项目概念源于人脑处理信息的方式。在人脑中，思维、计划、灵感等高感性活动不仅依赖于特定的感官信息，其同时属于不同感官的多种感官信息的综合整合。这启发了我们结合视觉信息和体感信息来实现高精度的手势识别。”



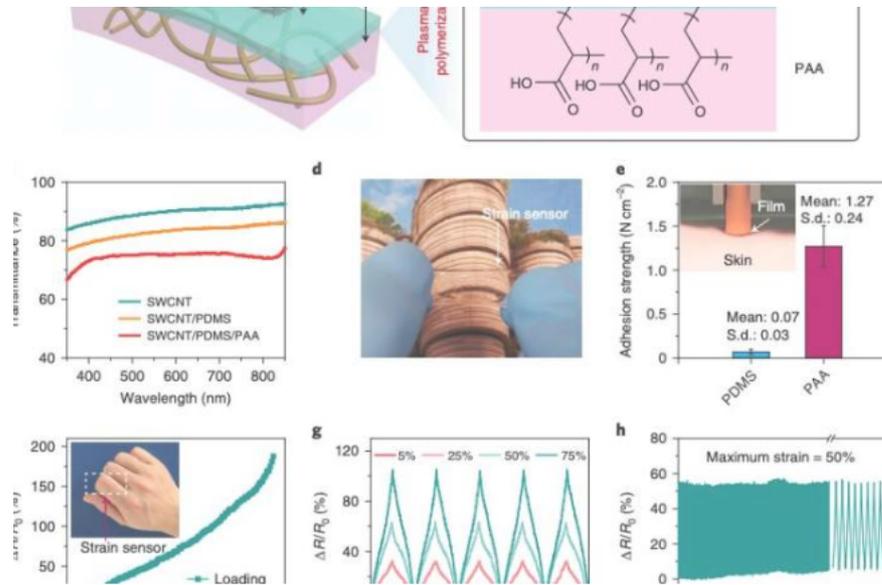
在解决实际任务时，人类通常会整合从周围环境收集的视觉信息和体感信息。这两种类型的信息彼此互补，而可以更好地帮助人类理解问题涉及的所有要素。

所以在开发手势识别技术时，陈晓东教授及同事确保它能够整合由多个传感器收集的不同类型信息。最终，团队的目标是建立一个能够以高精度识别人类手势的架构。

陈晓东教授解释道：“为了达到我们的目标，我们通过设计和制造可伸展的舒适传感器来改进传感器的数据质量。与现有的可穿戴传感器相比，这种传感器可以收集更精确的手势体感数据。另外，我们开发了一种生物启发式的体感视觉（Bioinspired Somatosensory-Visual; BSV）学习架构，它可以合理地融合视觉信息和体感信息，这类似于大脑中的体感-视觉融合结构。”

所述的BSV学习架构复刻了人脑是如何以多种方式融合体感信息和视觉信息，但方式是通过一个仿生神经网络结构。

另外，机器学习架构中的一些分段网络将如同大脑神经网络一样处理相同的模态感觉数据。例如，分段卷积神经网络（CNN）专门执行卷积运算，复制生物神经系统内局部感受场的功能，从而模拟人脑视觉处理部分发生的初始视觉信息处理。



最后，研究人员设计的架构使用了新开发的稀疏神经网络来融合特征。这个网络复刻了大脑多传感神经元是如何表示视觉信息和体感信息之间高效初始交互。

陈晓东教授表示：“我们开发的技术有三个独特的特点。首先，它可以处理视觉和体感信息的早期交互。其次，CNN的卷积运算类似于生物神经系统中的局部感受场的功能，它可以自动学习层次化的深空特征，并从原始图像中提取平移不变特征。最后，我们提出了一种基于弗罗贝尼乌斯定理条件数的疏剪策略来实现高效的稀疏神经网络。”

在一系列的初步评估中，这种BSV学习架构的结果优于单模态识别方法（即只单独处理视觉或体感数据，而不是同时考虑两者）。值得注意的是，与过去开发的三种多模式识别技术（加权平均融合（SV-V）、加权注意力融合（SV-T）和加权乘法融合（SV-M）架构相比，它能够更准确地识别人类手势。

**相关论文：** Gesture recognition using a bioinspired learning architecture that integrates visual data with somatosensory data from stretchable sensors

陈晓东教授说道：“与单模态识别方法和常见的多模态识别方法（SV-V、SV-T和SV-M）相比，我们的仿生学习架构可以达到最佳的识别精度。在图像噪点大、曝光不足或曝光过多的非理想情况下，它依然能够保持较高的识别精度。”

所述的大脑启发式架构最终可以支持一系列的用例，如能够读懂病人肢体语言的医疗机器人，帮助创造更先进的虚拟现实和增强现实系统等等。

他表示：“它独特的仿生特性使得我们的架构优于大多数现有的方法，我们的实验结果已经证实了这一点。我们的下一步计划是根据视觉数据和传感器数据的仿生融合构建一个VR和AR系统。”

---