机器人的视听觉信息处理

石 大 明

深圳大学 计算机与软件学院

履 历

哈尔滨工业大学
机电控制学博士

英国南安普顿大学
计算机科学博士

1997-2016 哈尔滨工业大学 讲师，教授，博导

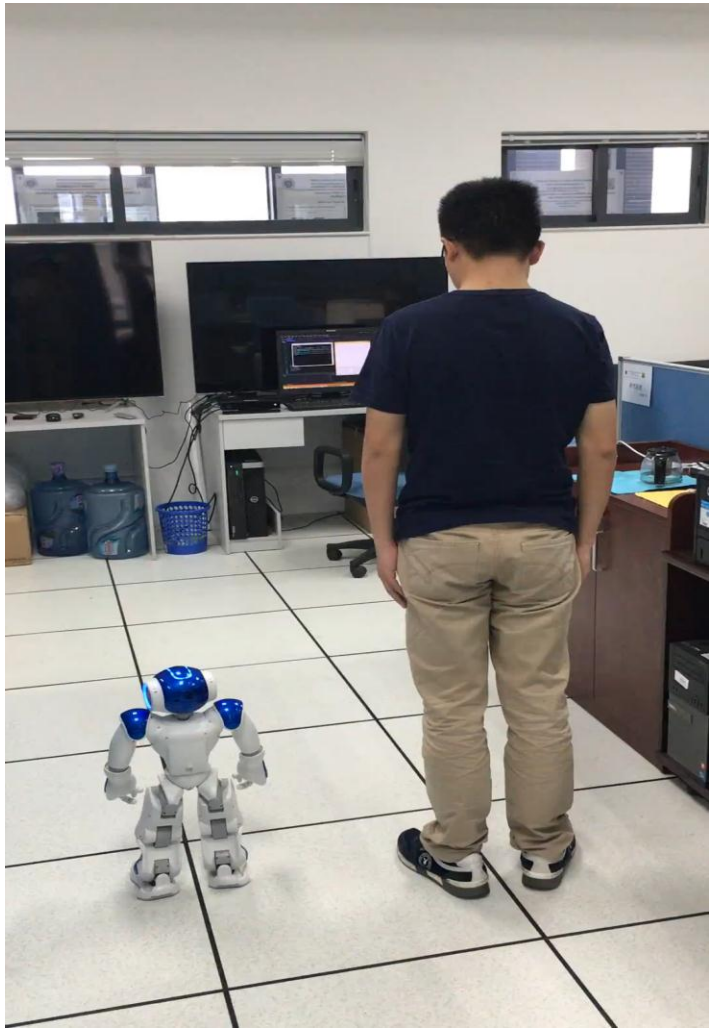
1998-1999 新加坡国立大学 博士后

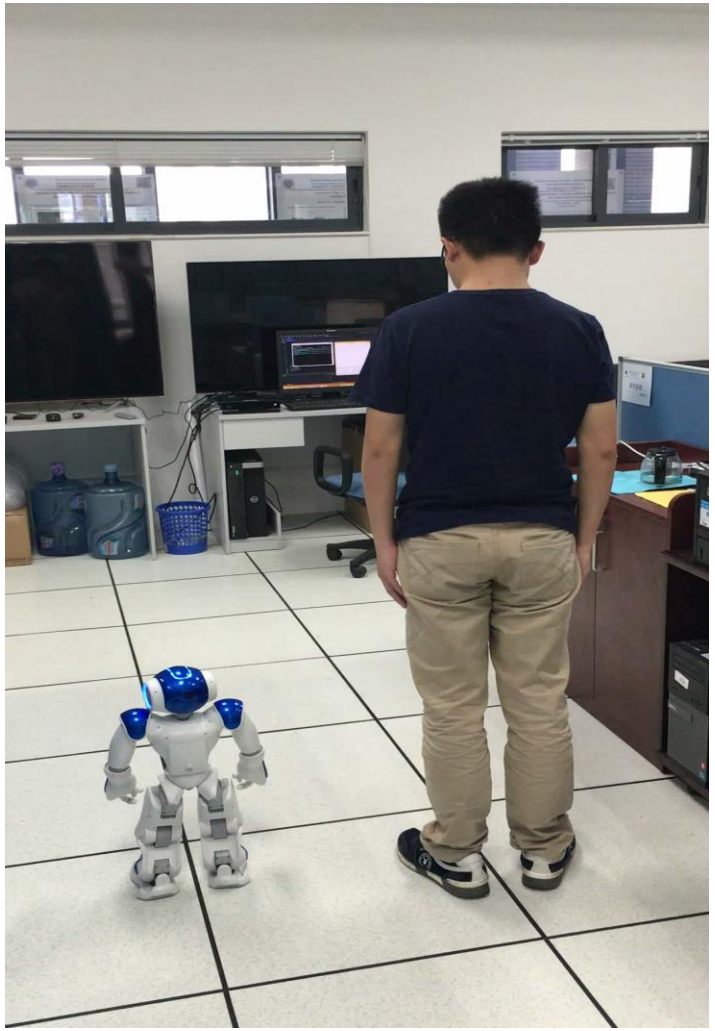
2002-2010 新加坡南洋理工大学 助理教授

2010-2016 英国密德萨斯大学 教授

2016-至今 深圳大学 特聘教授

机器人视听觉信息处理示例





演讲内容

一、机器人视听觉一览

二、机器人视觉导航实例

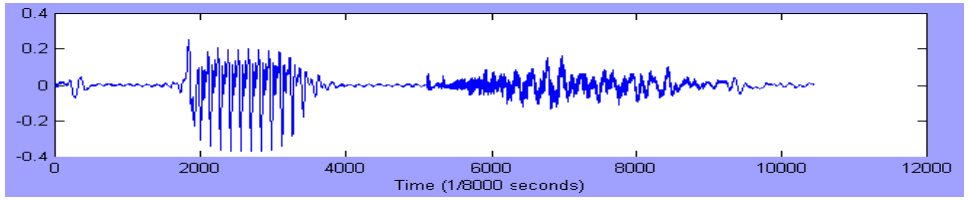
三、机器人的视听觉感知

四、机器人的视听觉认知

机器人视听觉一览

听觉

声音信号:

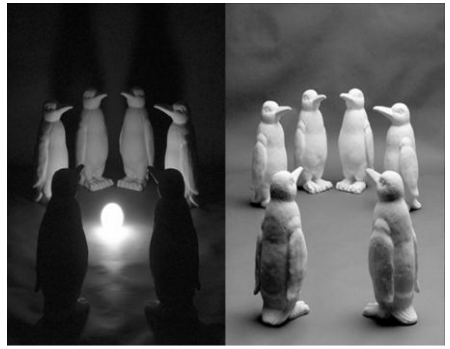


语音识别: 大 学 的 教 授

脑海猜测

	大达 …	雪学 …	里的 …	野兽教授 …

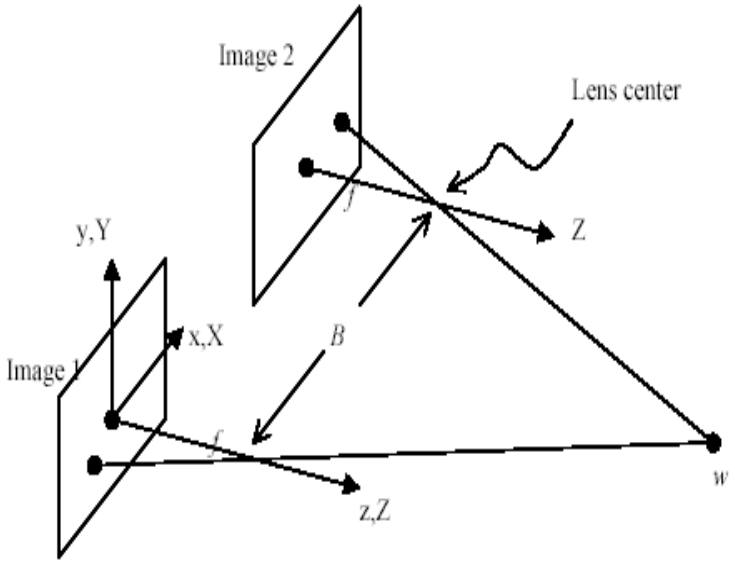
视觉



机器人视听觉：从信号处理到信息处理

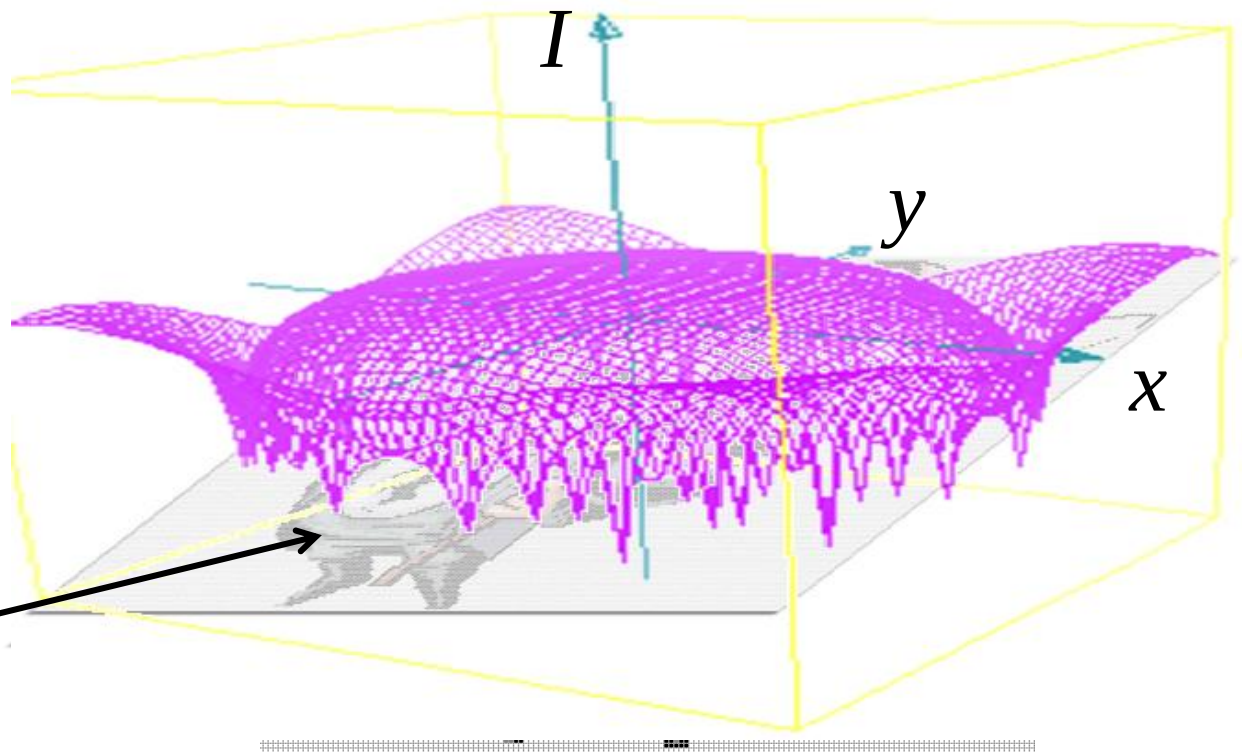
计算机视听觉一览

人类的80%信息来源于视觉



$$\left.\begin{aligned} \frac{x_1}{f} &= \frac{X_1}{f - Z_1} \\ \frac{x_2}{f} &= \frac{X_2}{f - Z_2} \\ X_2 &= X_1 + B \\ Z_1 &= Z_2 = Z \end{aligned}\right\} \Rightarrow Z = f\frac{f \cdot B}{x_2 - x_1}$$

视听觉数据 ＝ 离散函数



像素 (x,y)

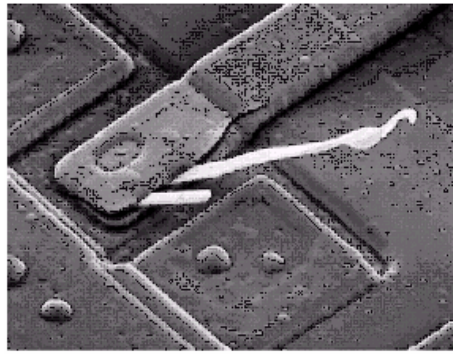
二值图像, if $I(x,y)\in[0,1]$

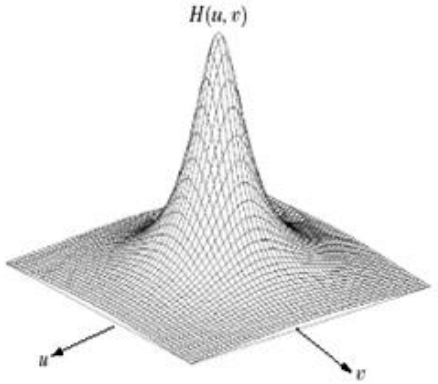
灰度图像, if $I(x,y)\in[0,255]$

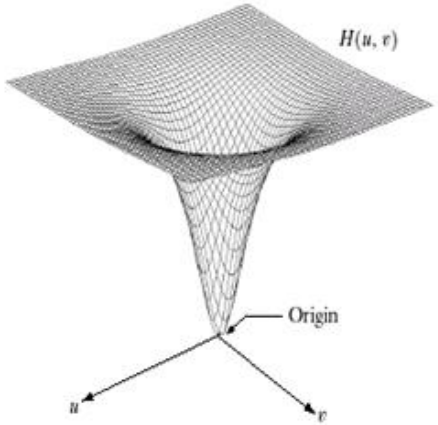
彩色图像, if $I(x,y)\in[(0,0,0),(255,255,255)]$

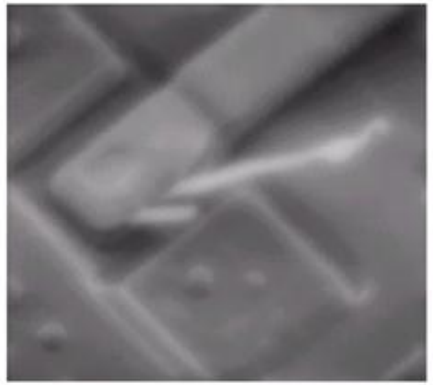
(红, 绿, 蓝三元色)

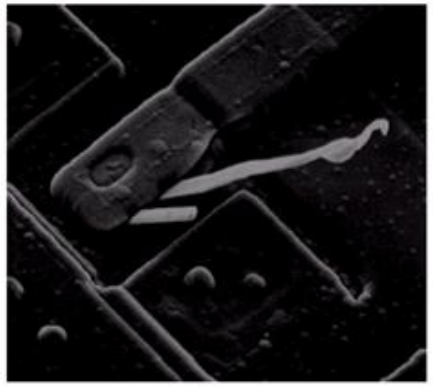
视听觉数据分析举例









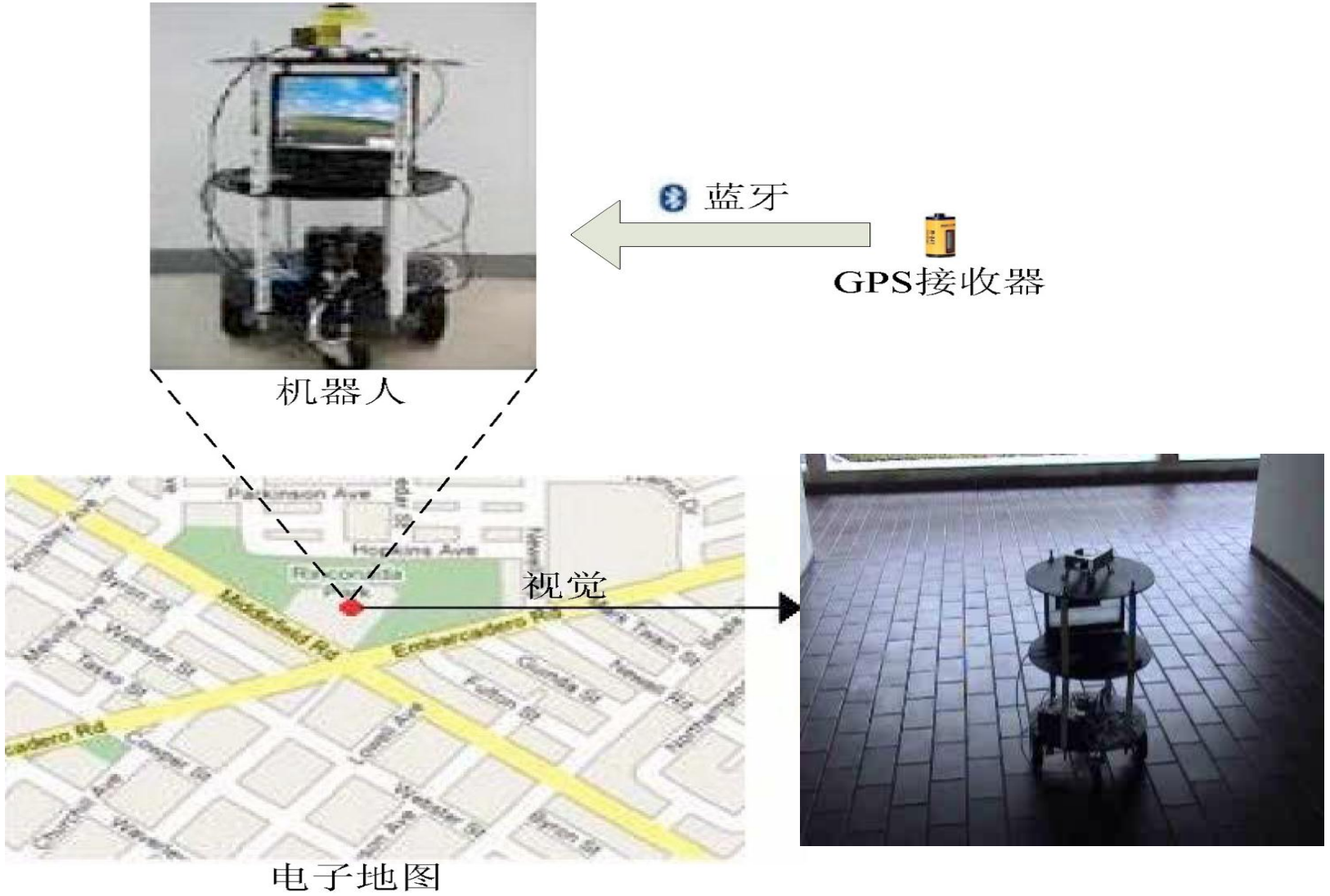


低通滤波：保持低频数据造成图像平滑

高通滤波：保持高频数据造成图像锐化

无人驾驶

无人驾驶是通过GPS和视觉信息定位车辆（机器人）在电子地图中的位置来实现的。



目标跟踪

目标跟踪的目的是定位目标在每帧视频图像中的位置， 产生目标运动轨迹。跟踪器能得到目标在每帧中的图像区域。

图像引导手术

图像引导手术就是根据图像提供的实时信息跟踪定位被手术部位及手术工具

过去： 皮开肉绽 ＋ 痛不欲生

麻醉药

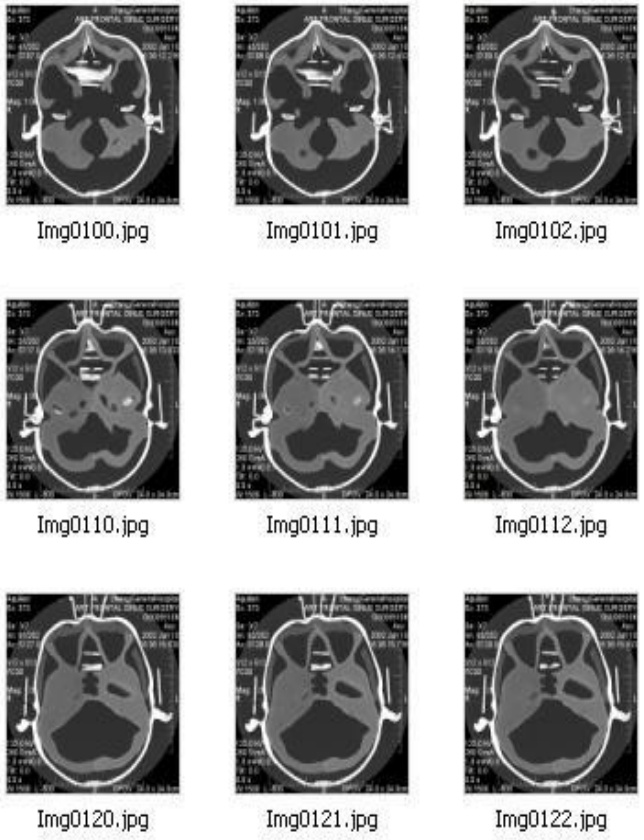
现在： 皮开肉绽 ＋ 微痛／无痛

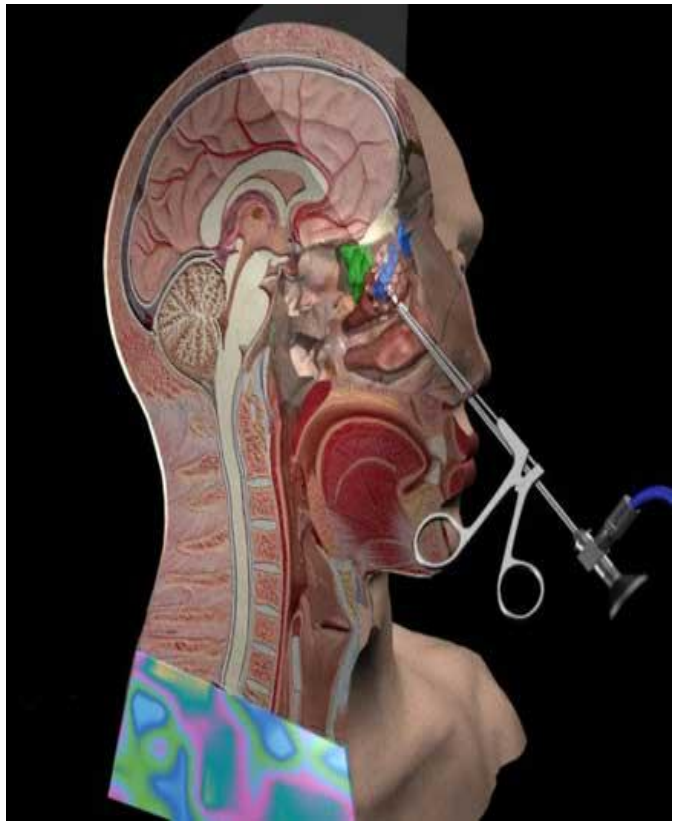
图像引导手术

未来： 微创／无创 ＋ 微痛／无痛

虚拟内窥镜

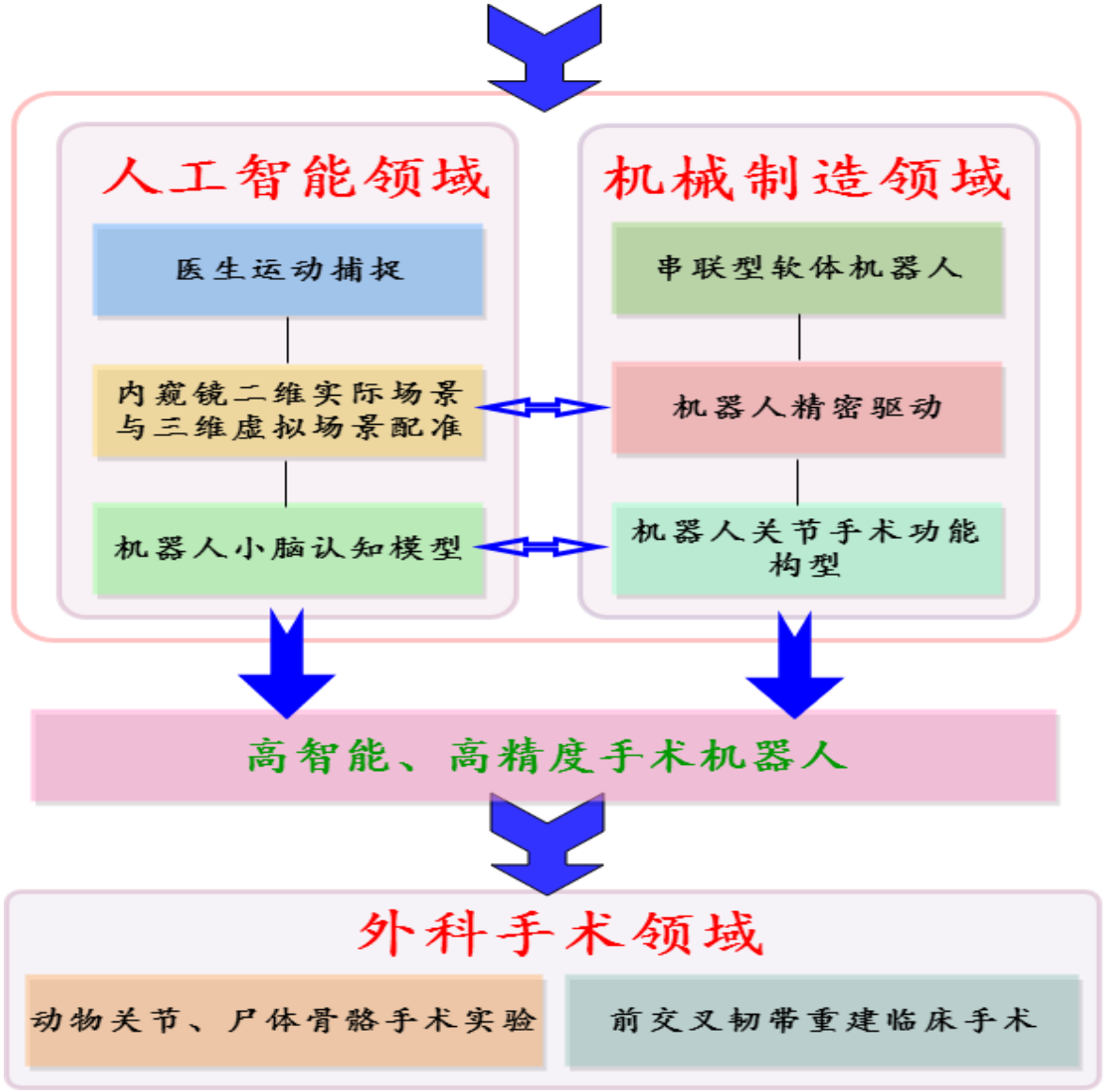
虚拟内窥镜是虚拟现实在医学无创、微创手术上的应用。它通过医学影像学与计算机后处理技术结合的三维影像，可以显示类似于真实内窥镜所见的人体腔道的解剖结构和病变。





手术机器人

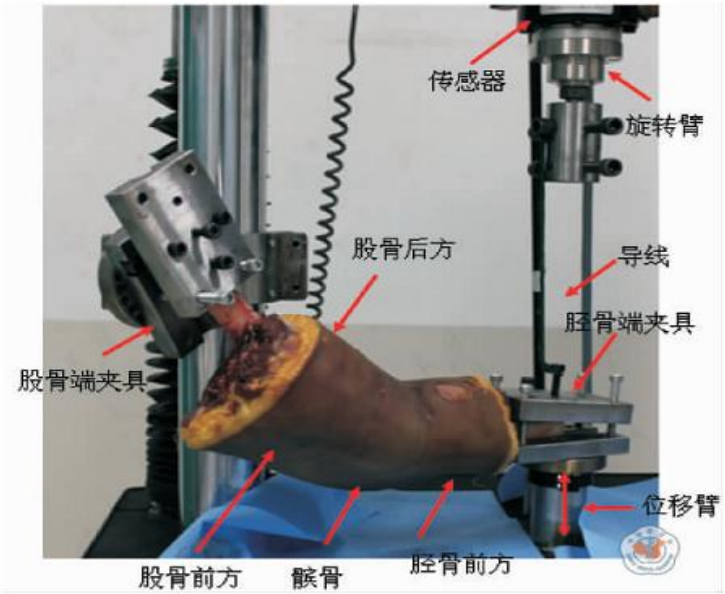
医生控制系统


人工智能领域
医生运动捕捉
内窥镜二维实际场景与三维虚拟场景配准
机器人小脑认知模型
机械制造领域
串联型软体机器人
机器人精密驱动
机器人关节手术功能构型
高智能、高精度手术机器人
外科手术领域
动物关节、尸体骨骼手术实验
前交叉韧带重建临床手术

机器人视觉导航实例

人工智能　　机械制造　　外科手术

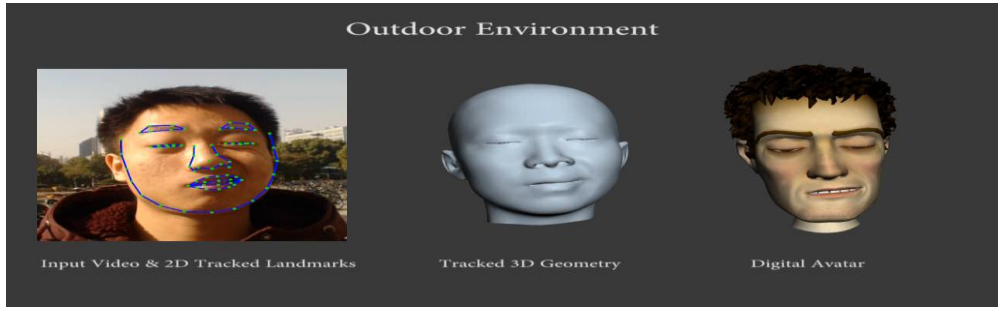
手术机器人



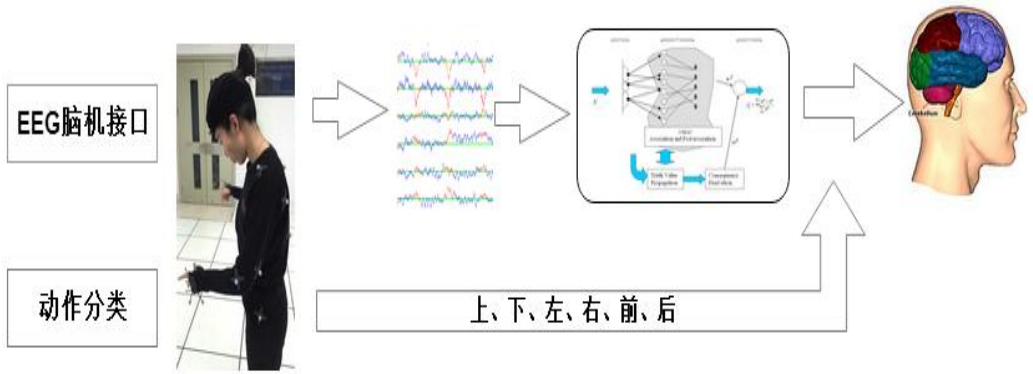
机器人的视听觉感知

感知是从感觉到知觉的过程。感觉是对事物个别属性和特性的认识，如感觉到颜色、声调、软硬等。而知觉是对事物的整体及其联系与关系的认识，如看到一面红旗、听到一阵嘈杂的人声、摸到一件轻柔的毛衣等。

第一步：
信息处理



第二步：
感知模型

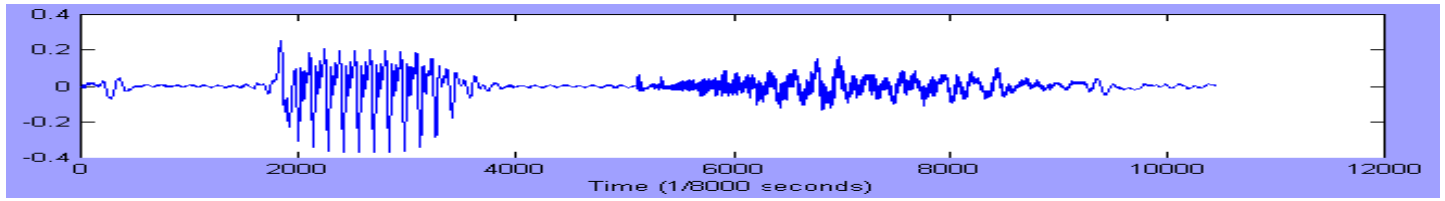


机器人的视听觉认知

认知是主动利用已有的知识和经验去概括地认识事物，揭露事物的本质及其内在的联系和规律，形成对事物的概念，进行推理和判断。

听觉

声音信号:



脑海猜测 大达… 雪学… 里的… 野兽教授…

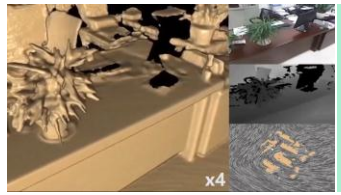
语音识别 大 学 的 教授

视觉

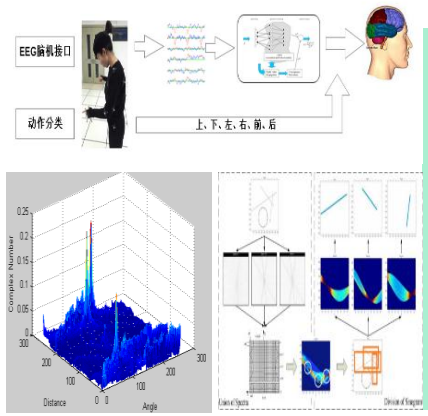
灾前场景

灾后场景

目标跟踪



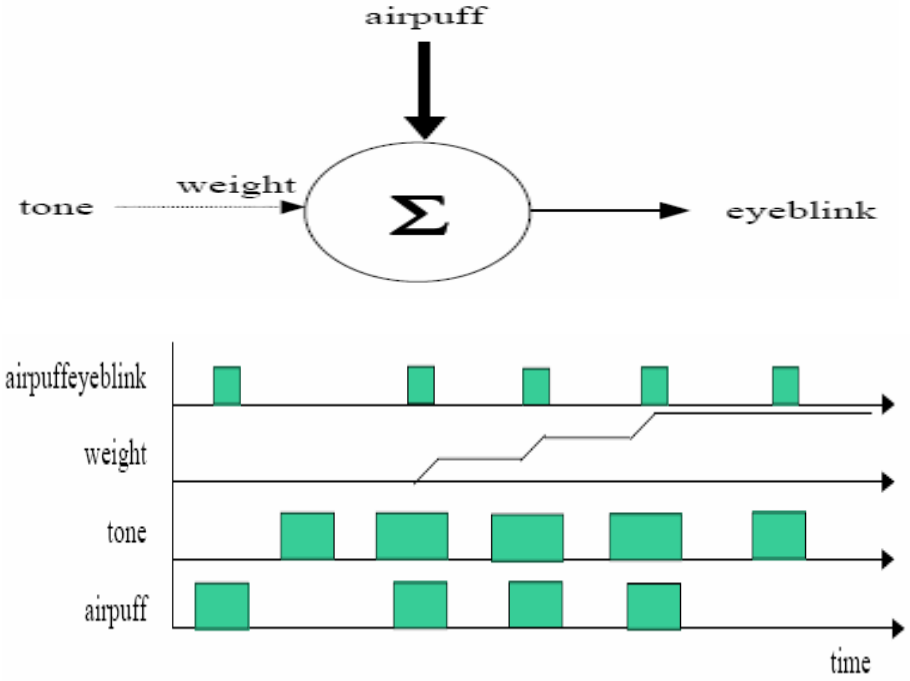
动态重建



视觉认知层面的未知场景定位与目标跟踪技术

联觉： 一种通道的刺激同时引起另一种通道的感觉。例如，看到红色会觉得温暖，看到蓝色会觉得清凉，听到节奏鲜明的音乐会觉得灯光也和音乐节奏一样在闪动。

利用联觉原理可以使得盲人通过声音“听图像”；使得失聪者“看声音”。



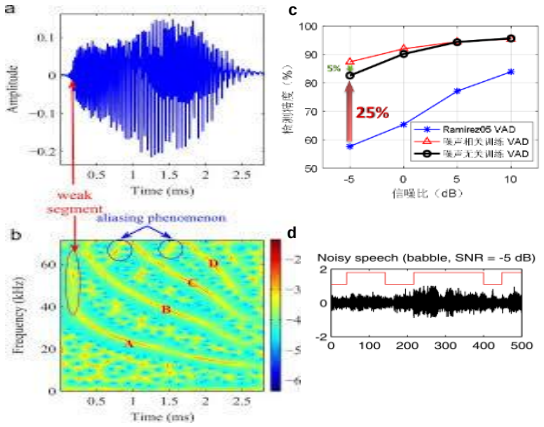
机器人的视听觉协同

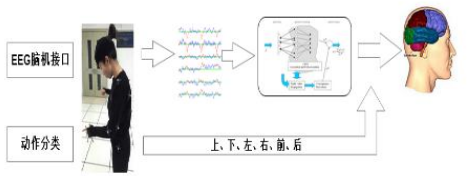
视听觉协同：利用视觉和听觉的信息互补，优化决策

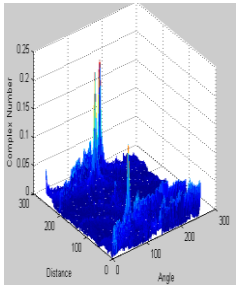
听觉

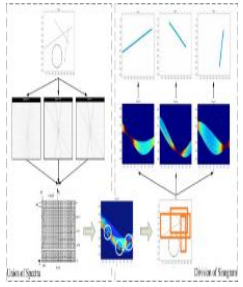
低质高噪环境下视听觉信息互补，实现协同认知

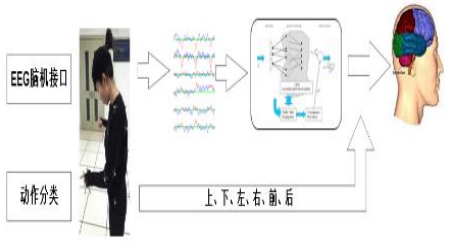
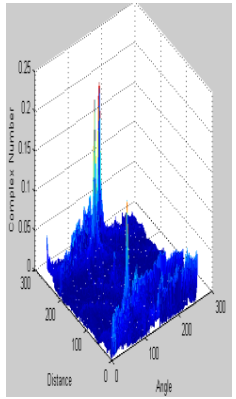
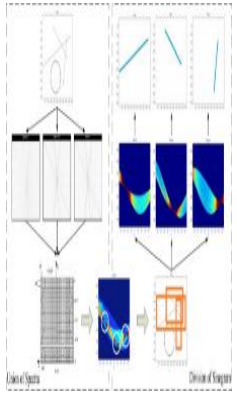
视觉









总结

- 机器人视听觉信息处理
- 信号→信息→感知→认知