

《人工智能通识》（科技素养）

# 第7讲 智能之躯——具身智能

主讲：丛润民

# 章节知识点概览

知识点1：让AI穿上机械战甲——具身智能

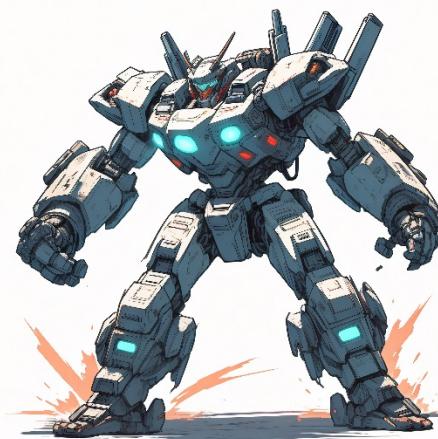
知识点2：从模仿中进化——模仿学习

知识点3：让理想照进现实——仿真到真实的迁移

知识点4：看智能机器人如何花式干活

# 知识点1： 让AI穿上机械战甲——具身智能

学  
习  
机  
械  
智  
能

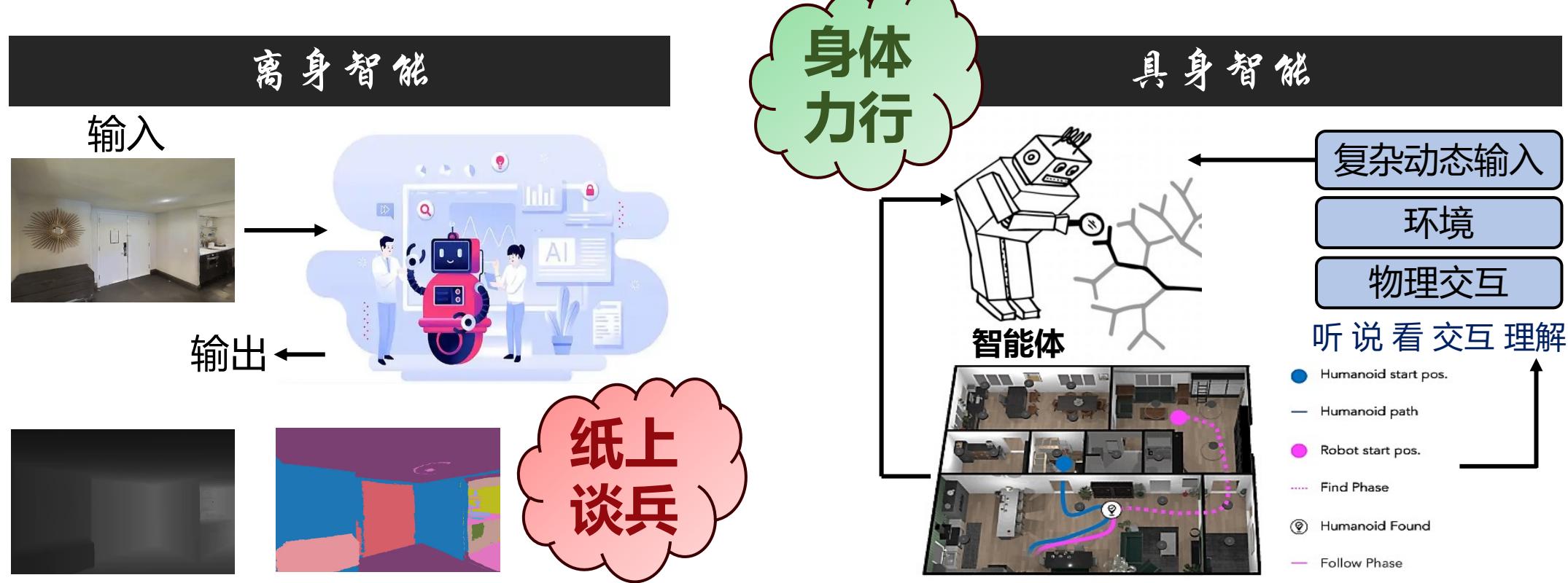


01 具身智能概述

02 具身智能的核心技术

**什么是具身智能？**

# 人工智能的机械战甲——具身智能

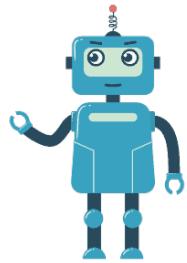


单一的符号智能往往与真实世界相脱节，  
认知与身体解耦

智能是具身化和情景化的，具身智能可通过  
与真实世界的交互完成任务

- 具身智能是以智能体作为本体支撑，不再局限于被动响应，而是能够像生物体一样，主动适应环境变化，应对噪声干扰，并适时调整自身行为。

# 人工智能的机械战甲——具身智能



## AI智能体

机器人等人工智能系统。



## 感知输入

搭载各类传感器收集环境信息，以支持智能系统分析与学习。



达到解决复杂问题的目标，  
例如运动规划、导航等。



## 交互学习

在与环境交互中学习，  
直至达成目标。



## 世界模型

能对世界时空维度的信息进行抽象表达和理解。

## ■ 具身智能的基本概念

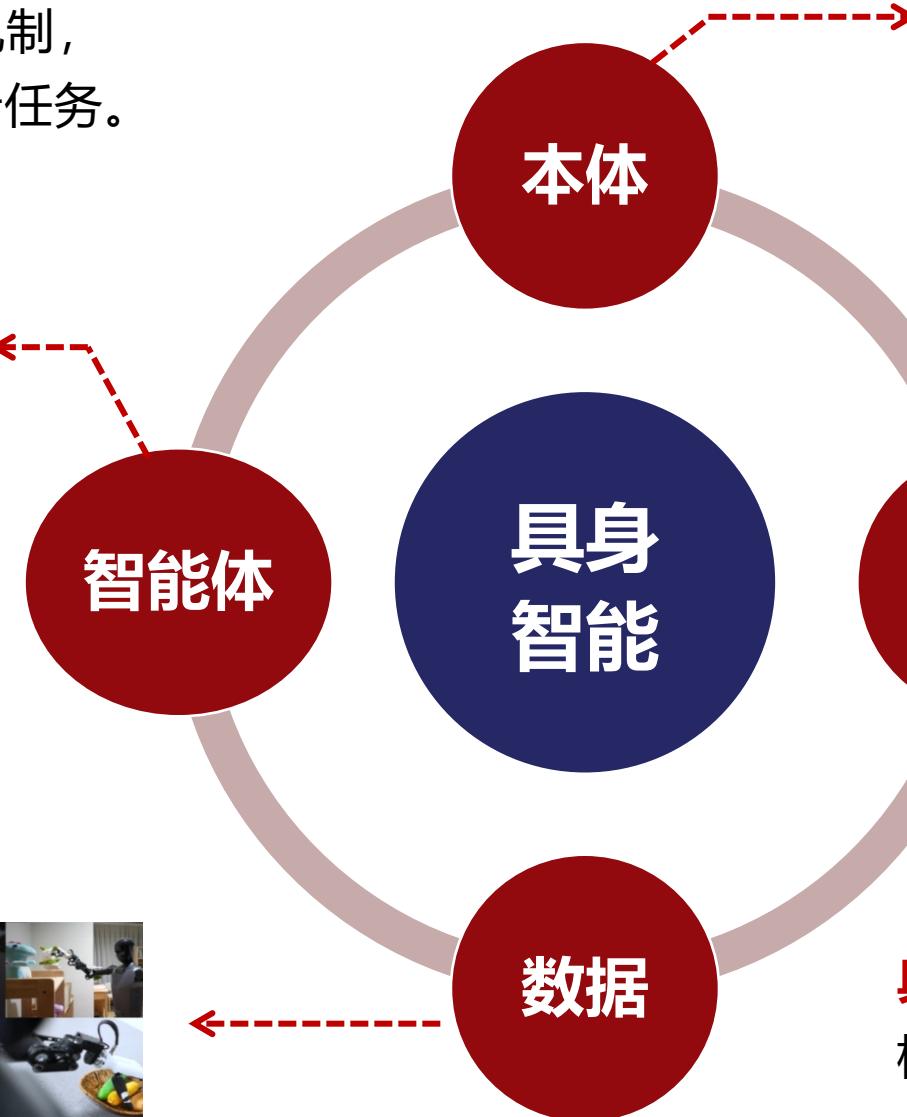
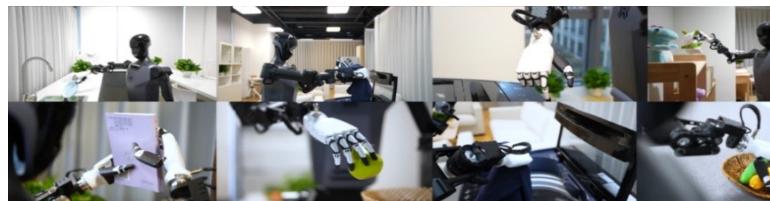
- **具身化 (Embodiment)** : 是一个跨学科概念，强调身体（本体）在认知、情感、社会互动中的核心作用，反对传统的身心二元论。
- **具身智能 (Embodied AI)** : 特指那些**拥有物理形态**，并能**直接参与物理世界交互**的智能系统，如服务型机器人、智能无人驾驶车辆等。它们通过“身体力行”的方式，展现出高度的环境适应性与任务执行能力。
- **具身任务**: 像人类一样通过观察、移动、对话以及与世界互动从而完成的一系列任务。
- **多模态**: 一个模型或系统**能够处理多种不同类型的输入数据**并融合它们生成输出，如文本、图像、音频和视频等。这种能力对于提升智能系统的环境感知与决策能力至关重要。
- **主动交互**: 机器人或智能体**与环境的实时交互**过程，从而提高智能体的学习、交流与处理问题的能力，是具身智能实现高效任务执行的关键。

# 人工智能的机械战甲——具身智能

**智能体**作为本体的智能核心，具备敏锐的感知能力和动态决策机制，能够解析复杂环境并高效执行任务。

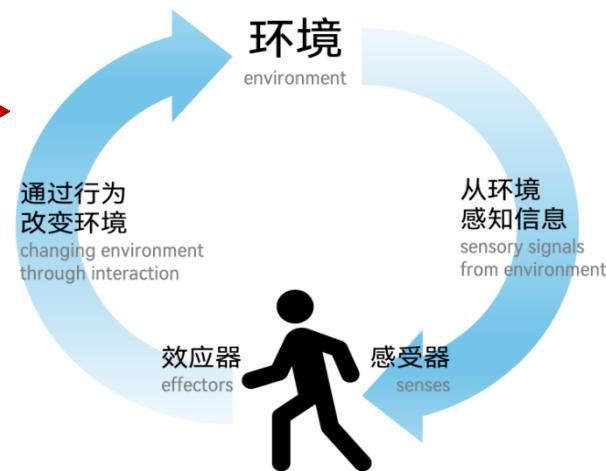


**数据**在机器学习与具身智能系统中至关重要，通过整合大量多样的具身数据，提升了智能体的任务执行成功率。



固定式机器人 自动驾驶汽车 四足机器人

**本体**在物理与虚拟空间中承担环境感知和任务执行。



**具身学习**通过智能体与环境及人类的互动，构建“感知-决策-行动(控制)”闭环。

# 案例引入1——人形机器人

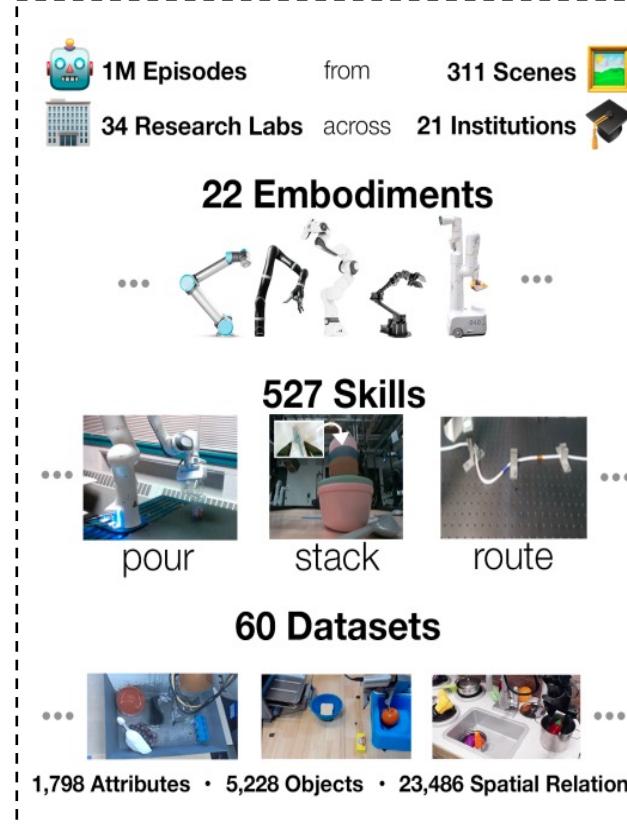




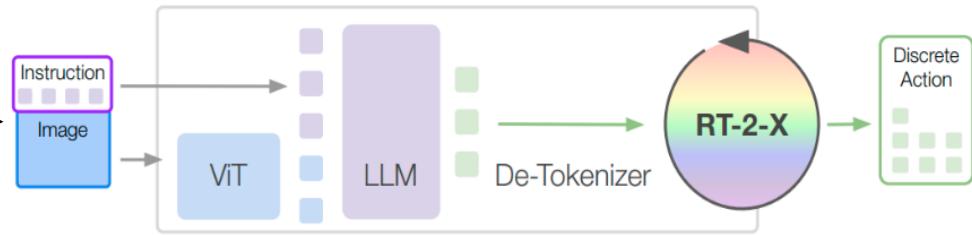
2023年7月，谷歌DeepMind发布具身智能视觉-语言-动作模型RT - 2  
(Robotics Transformer 2)

谷歌 DeepMind 提出的机器人模型 Robotics Transformer 2 (RT-2) 就是一个全新的具身智能模型，它从网络和机器人数据中学习，并将这些知识转化为机器人控制的通用指令。这样，在 web-scale 数据上进行预训练的视觉-语言模型 (VLM) 可以从 RT-1 的机器人数据中学习，以成为可以控制机器人的视觉-语言-动作 (VLA) 模型。

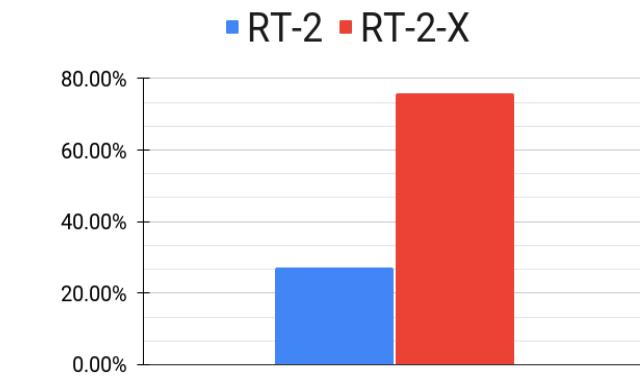
# 案例引入2——谷歌RT-X项目



- 互联网数据预训练，学习通用知识
- 具身数据微调，学习动作控制

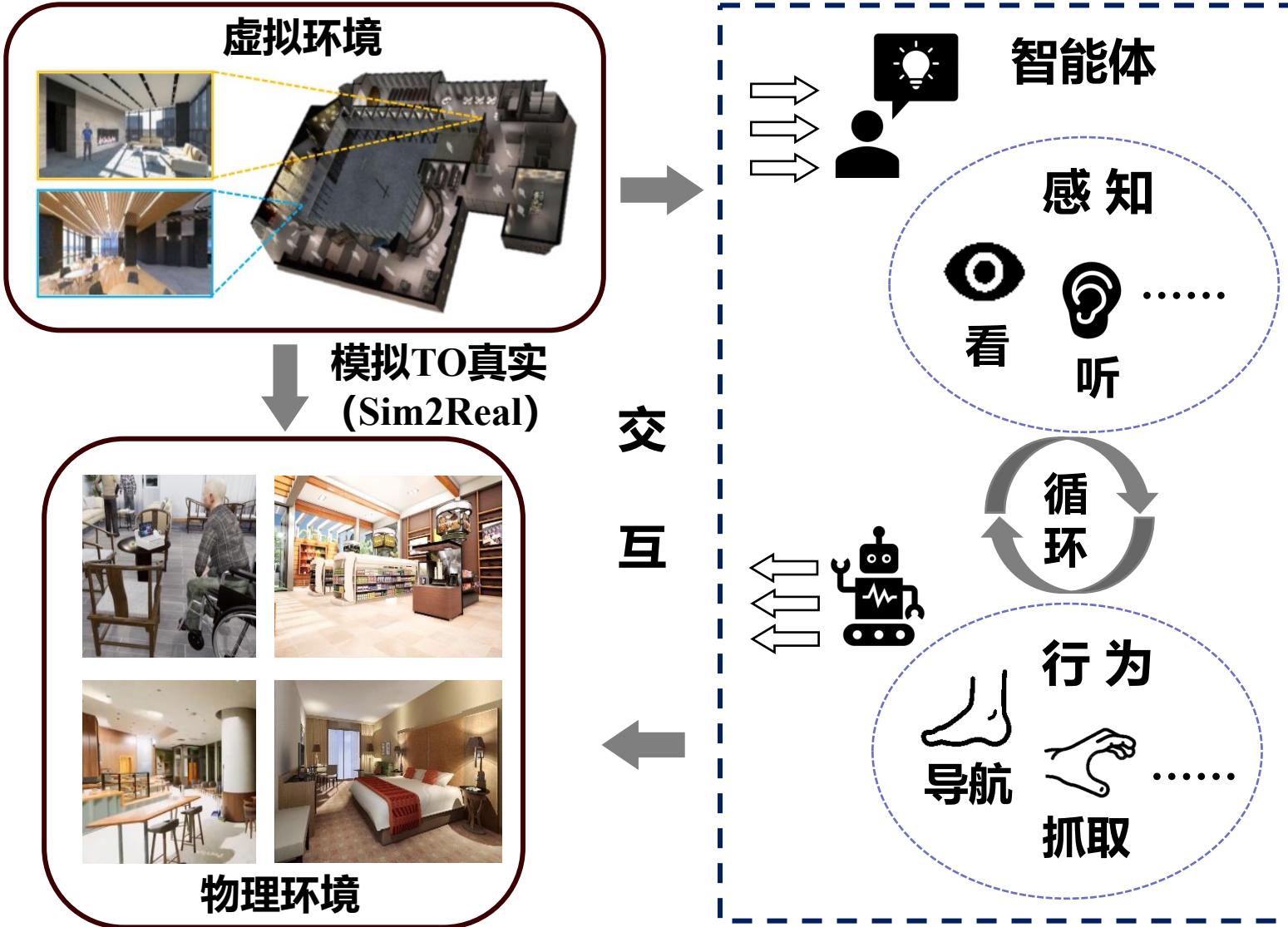


“将苹果移动至在罐子和橘子之间”



基于Open X-Embodiment 数据集训练，RT-2-X模型在众多具身任务上平均成功率提升近50%。

2023年10月谷歌发布的RT-X项目由RT-1-X和RT-2-X构成，可以用来构建通用具身数据集：涵盖多种机器人类型、任务和场景，整合了来自34家研究实验室的60个数据集，数据集总量惊人地达到1,402,930条记录。



## 核心技术：

- **具身感知:** 深度融入物理世界的智慧触角。
- **行为模块:** 复杂任务达成的执行者。
- **具身交互:** 构建人机协作的新生态。

## ➤ 核心技术1：具身感知

- ✓ **主动视觉感知：**智能体能够自主控制感知设备，如选择最佳视角和运用注意力机制。这种能力允许**智能体主动探索环境**，优化信息获取，从而提高任务执行效率。
- ✓ **三维视觉定位与物体感知：**智能体**需具备在三维空间中定位自身及周围物体的能力**，这对导航和物体操作至关重要。现代视觉编码器预训练技术增强了对物体类别、姿态和几何形状的精确估计，使智能体能在复杂动态环境中全面感知。
- ✓ **多模态感知融合：**除了视觉之外，触觉和听觉等感知模态同样重要，它们为智能体提供额外的环境信息。触觉传感帮助智能体感知物体的质地、重量和形状，支持精确的物体操作。**整合多模态感知数据**，能够显著提升智能体对环境的整体理解能力，使其在执行任务时更加灵活和高效。

## ➤ 核心技术1：具身感知的发展方向

### 感知大模型

- SAM: 视觉分割大模型
- DINO-X: 目标检测大模型



静态环境识别精度与人类相当

### 被动感知

第三人称

### 具身主动感知

通过主动获取图像，相比现有大模型的目标检测性能有显著提升



### 主动感知

第一人称



被门挡住了视线？

交互后，推开门看看是什么，能干什么？

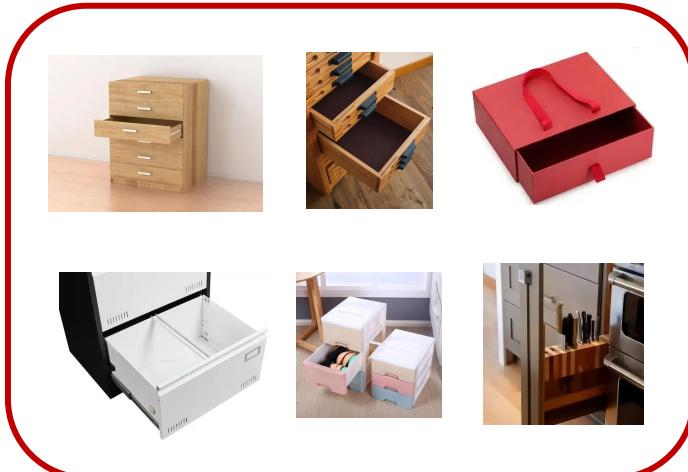
### 具身交互感知

第一人称  
行为交互+感知

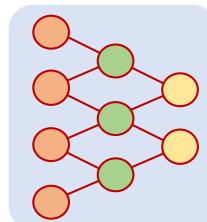
具身感知模式从被动感知向主动交互感知方向发展

## 核心技术1：具身感知与非具身感知的区别

### 第三人称智能



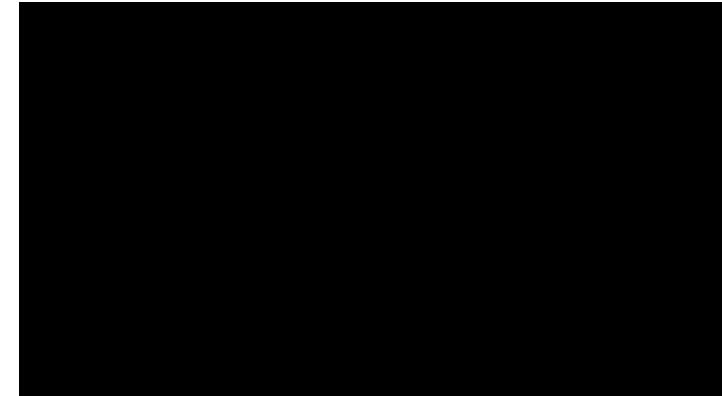
标签：抽屉



别人告诉我这就是抽屉

(a) 非具身感知

### 第一人称智能

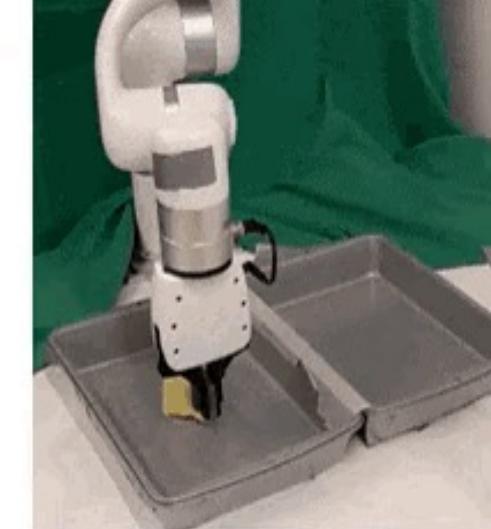


可以打开可以装东西  
我亲身体验抽屉是什么

(b) 具身感知

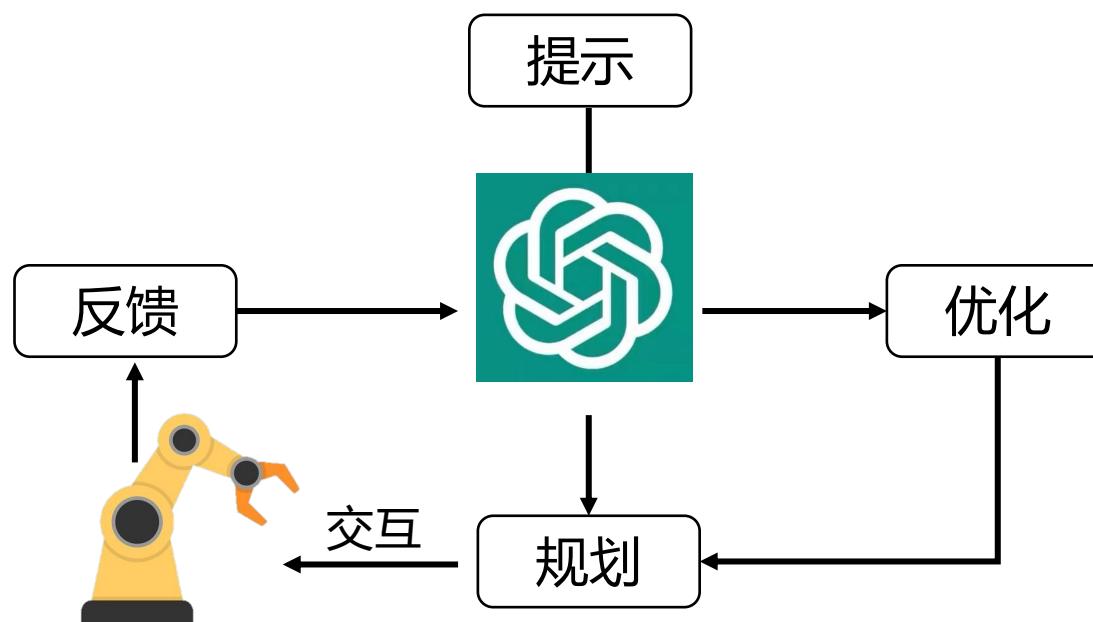
## ➤ 核心技术2：行为模块的应用示例

- ✓ 行为模块是**连接感知与行动的纽带**，它基于丰富的感知数据或人类指令，操纵智能体执行复杂的物体操作任务。
- ✓ 这一过程融合了语义理解、场景感知、决策制定与稳健的控制规划。

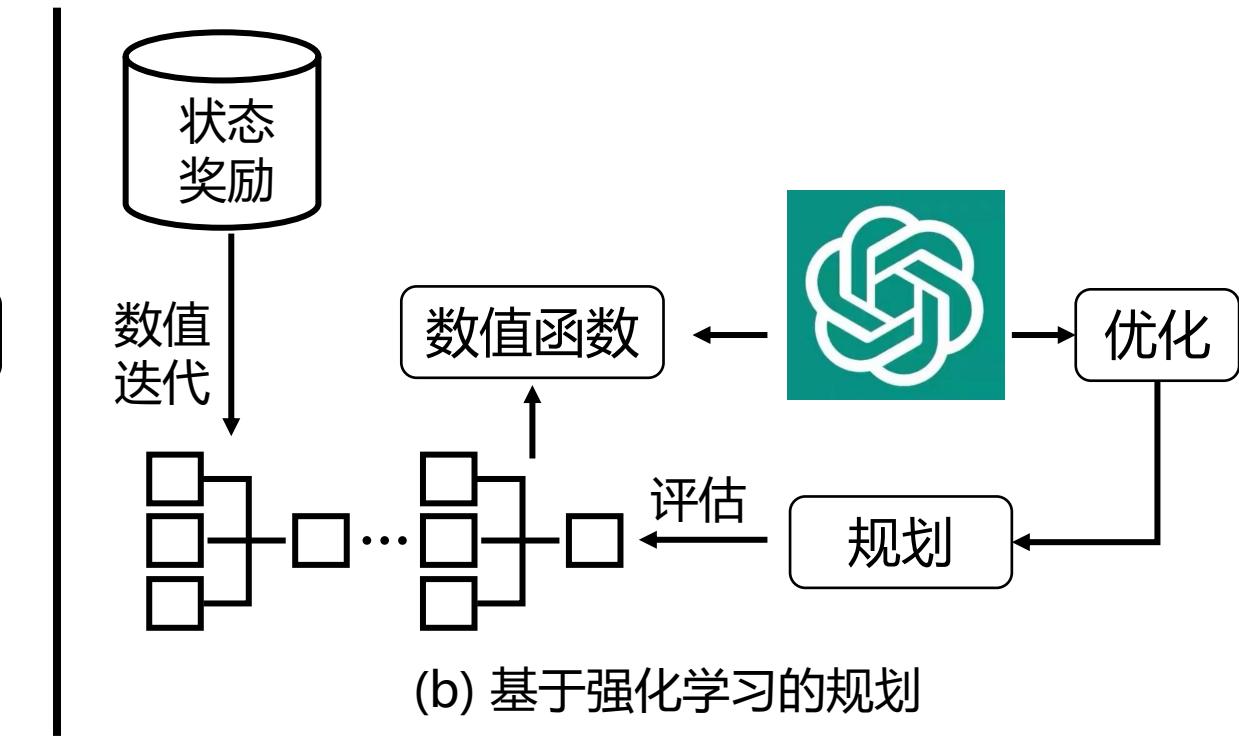


## ➤ 核心技术2：行为模块的规划策略

- ✓ 行为模块规划策略包括基于**物理反馈**和**强化学习**的规划。
- ✓ 前者通过人类提示和实时反馈确保任务准确性，适用于高精度控制；
- ✓ 后者通过交互和奖励机制优化策略，适应复杂动态环境，实现自主学习和快速适应。



(a) 基于物理反馈的规划

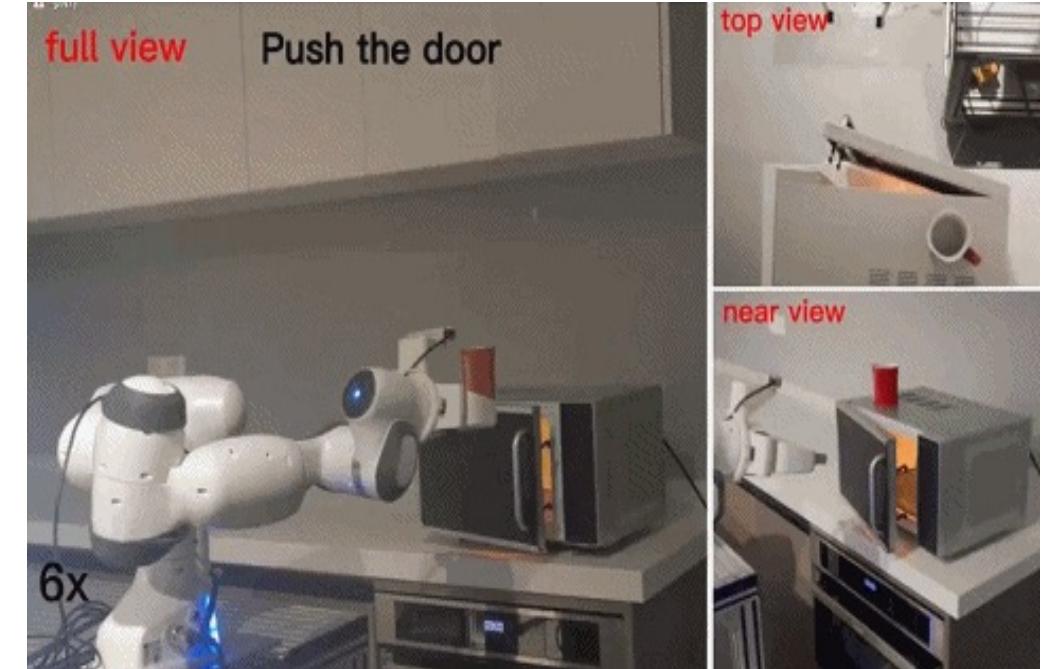
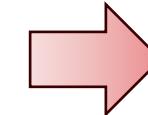
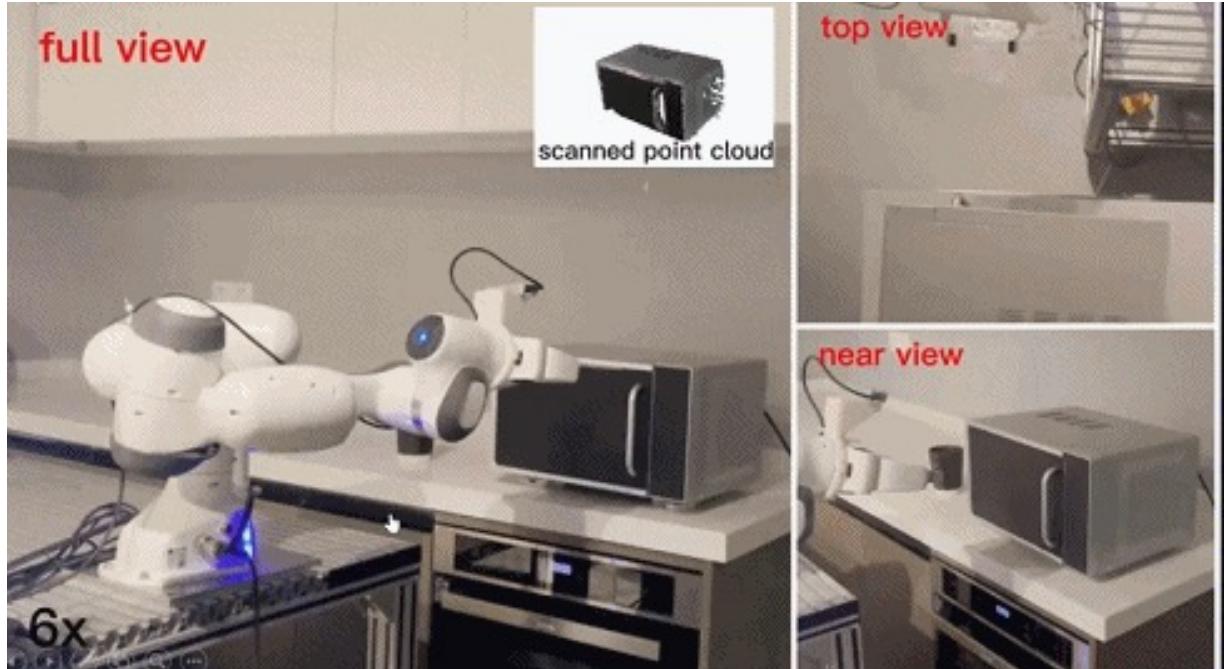


(b) 基于强化学习的规划

## ➤ 核心技术3：具身交互概念

- ✓ **人类监督与反馈的重要性：**人类在监督智能体行为轨迹的同时，确保其行动符合需求，并保障交互的安全、合法及道德边界。尤其在医学诊断等敏感领域，人类监督能有效弥补数据局限性与算法能力的不足。
- ✓ **从被动感知到主动交互的转变：**智能体通过在线互动实现模型发展与进化，从第三人称的被动感知转向第一人称的主动交互感知。如智能体能够通过行为交互主动适应实际场景，如“被门挡住视线”的情况。
- ✓ **人类与智能体交互的两种范式：**具身交互分为“不平等互动”模式，即“指导者-执行者”范式，人类发布指令，智能体辅助完成任务；以及“平等互动”模式，智能体与人类共同决策，预示更加协同的未来。

## ➤ 核心技术3：具身交互示例



初始状态下，机器人拉微波炉门的准确度较差

具身交互让机器人学会准确拉开微波炉的门，之后就可以在上面加技能了，比如把东西塞进去

# 知识点2：从模仿中进化——模仿学习

学  
习  
机  
制  
上  
进



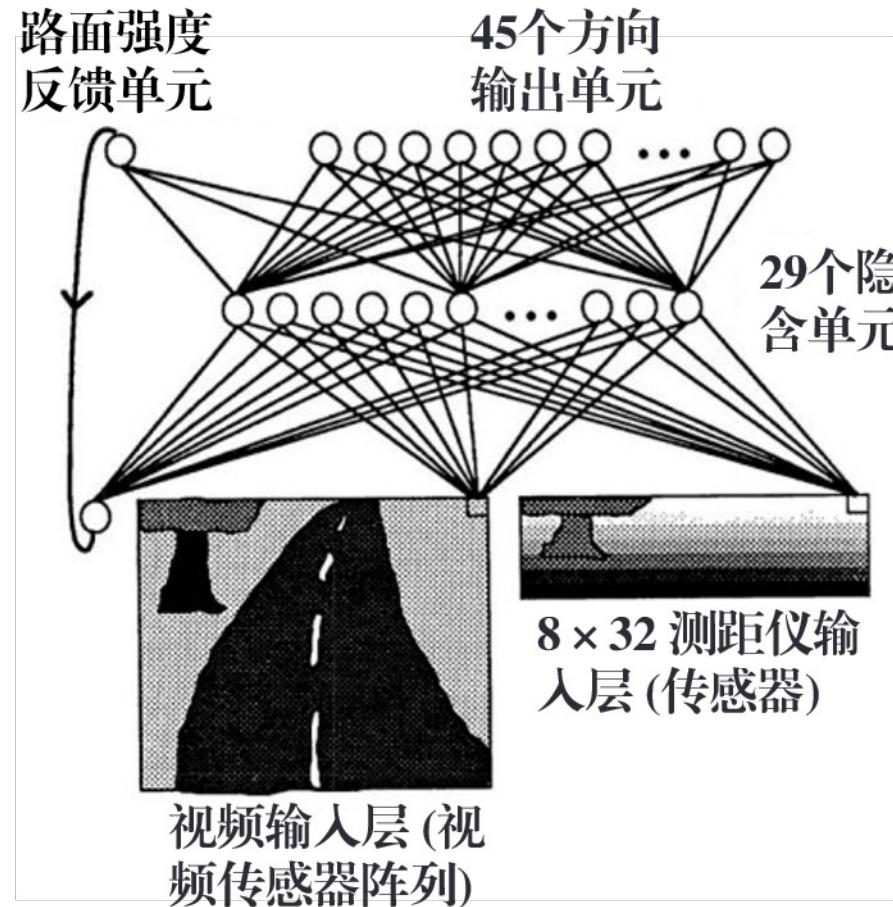
01 模仿学习概述

02 模仿学习的两种方法

# 生活中无处不在的模仿



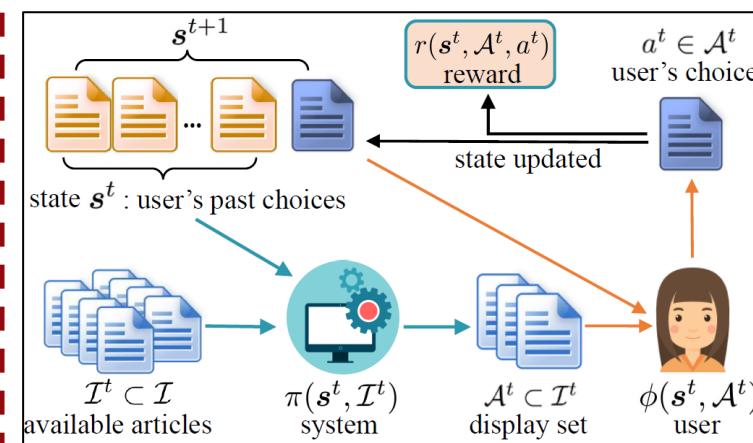
# 模仿学习体现在哪些实际场景?



20世纪80年代末，卡内基梅隆大学的研究人员已经利用人类专家示例成功训练一个浅层神经网络，其控制的无人车能够穿越北美洲。



2016年，AlphaGo成为第一个打败人类顶尖选手的机器人，也用到了模仿学习对人类棋谱进行学习。



互联网网约车派单、推荐系统等日常应用也用到了模仿学习，通过构建一个虚拟世界允许智能体试错与学习。

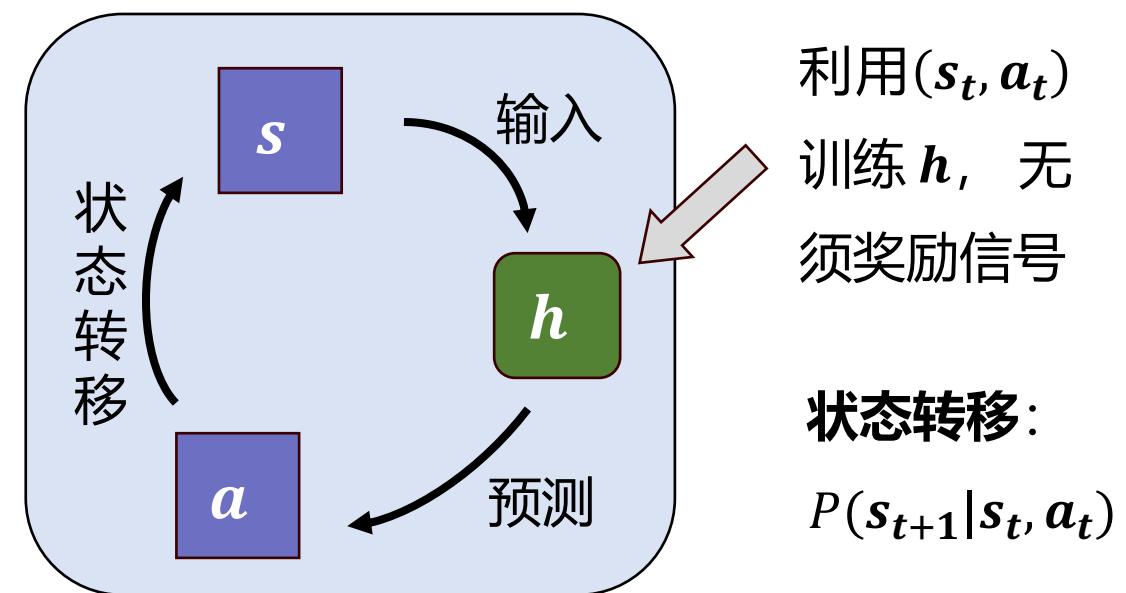
➤ **定义：**模仿学习可以理解为**从专家示例中学习**，是一种让智能体（如机器人）像人类专家能够进行智能决策实现进化的方法。在通向通用人工智能的路上，人们发现通过手工进行编程来教会智能体进行思考，涉及到大量的人工工程。相反，**人类**却能比较容易地完成“思考-决策”任务，并且**可以为智能体提供大量的示例行为**。

➤ **框架：**

**输入：**上下文或状态序列  $s$ 。

**输出：**动作序列  $a$ 。

**学习：**利用演示状态-动作对  $\{(s_t, a_t)\}$ ，训练一个智能体模型  $h$ 。

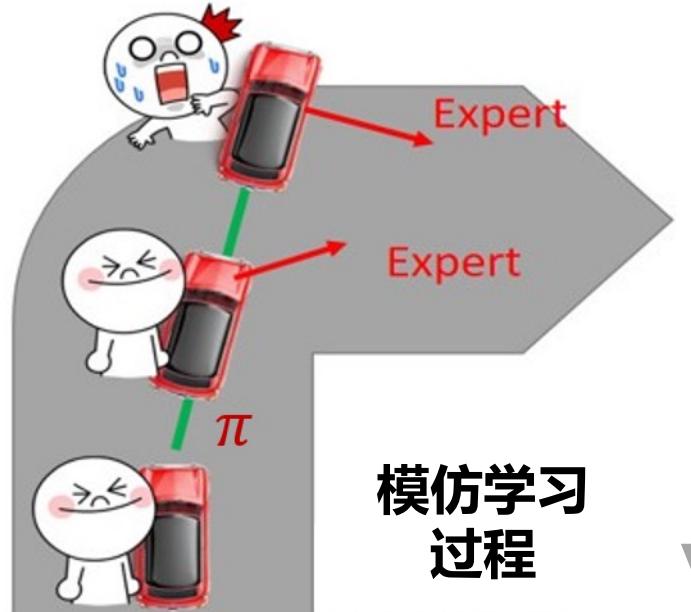
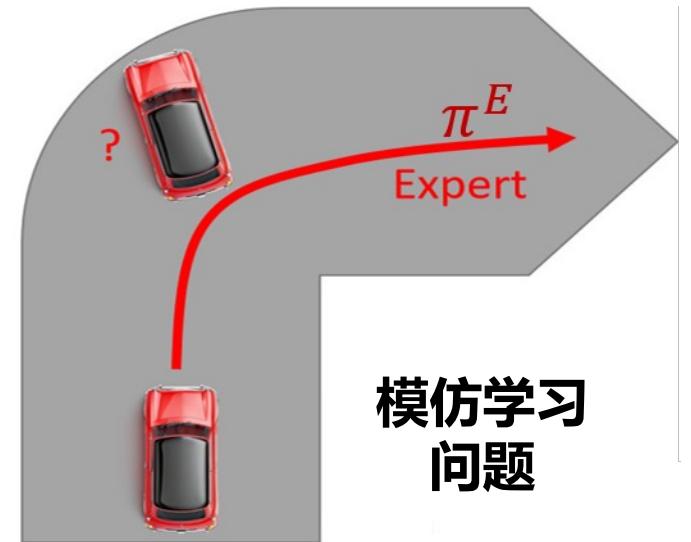


## ➤ 模型学习过程：

我们希望智能体能够模仿专家策略进行决策，即智能体与专家策略的**累计回报**比较接近。因此，模仿学习问题可以建模为如下的优化问题：

$$\min_{\pi} \{V(\pi^E) - V(\pi)\} \quad (1)$$

- (1) 最大化 $V(\pi)$ 的过程中是**不知道奖励函数**的。
- (2) 专家策略与环境交互产生**状态-动作对** $\{s, a\}$ ，称之为**专家示例**。
- (3) 专家示例作为训练数据集 $B$ ，由多条 $H$ 长的**轨迹** $\text{tr}$ 组成， $\text{tr} = \{s_1, a_1, s_2, a_2, \dots, s_H, a_H\}$ 。



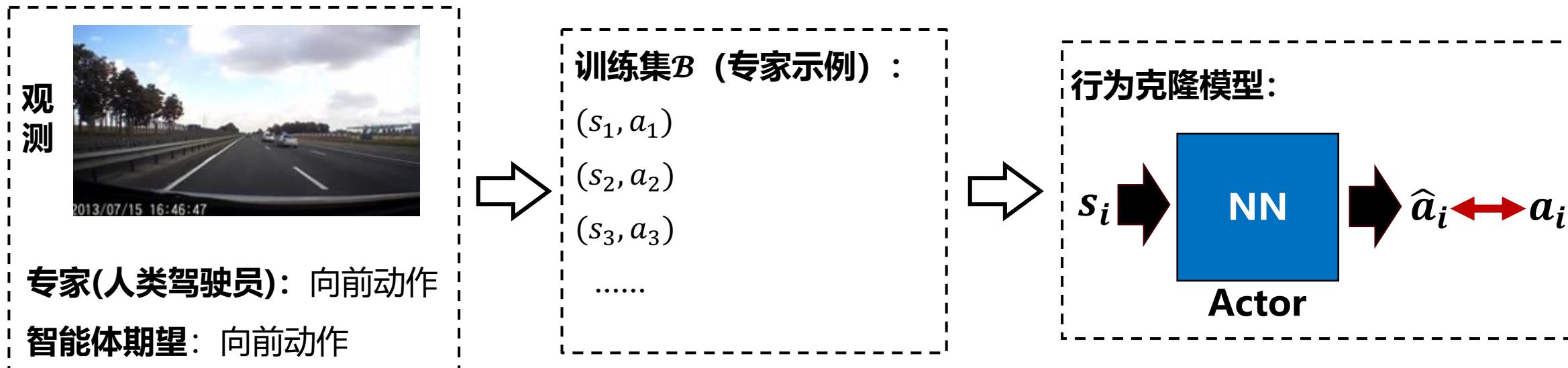
如何根据**专家示例**恢复**专家策略**？

# 模仿学习方法——行为克隆(Behavior Cloning)

➤ 定义：行为克隆采用直接的**有监督学习框架**，将专家示例 $\{(s_t, a_t)\}$ 中的状态 $s_t$ 作为样本输入，将动作 $a_t$ 视为标签。因此，其学习目标可以表示为：

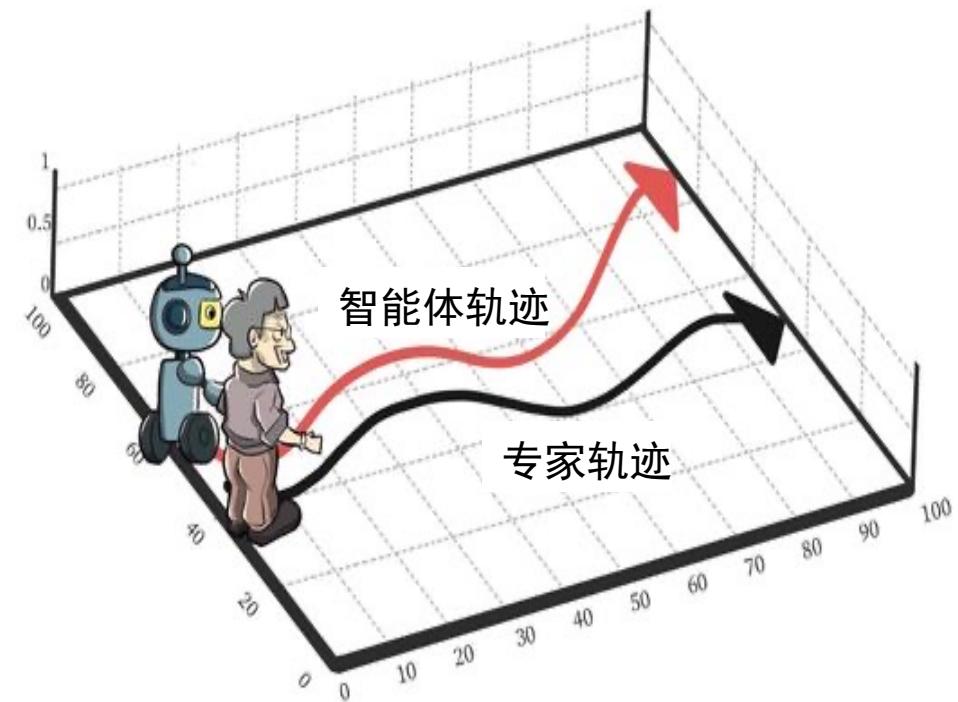
$$\theta^* = \arg \min_{\theta} \mathbb{E}_{(s,a) \sim \mathcal{B}} [\mathcal{L}(\pi_{\theta}(s), a)] \quad (2)$$

其中， $\mathcal{L}$ 为监督学习框架下的损失函数。如果动作 $a$ 呈现出**离散序列**的形式，损失函数可以采用**最大似然估计**来优化；如果动作 $a$ 是**连续序列**，则可以采用**均方误差函数**。



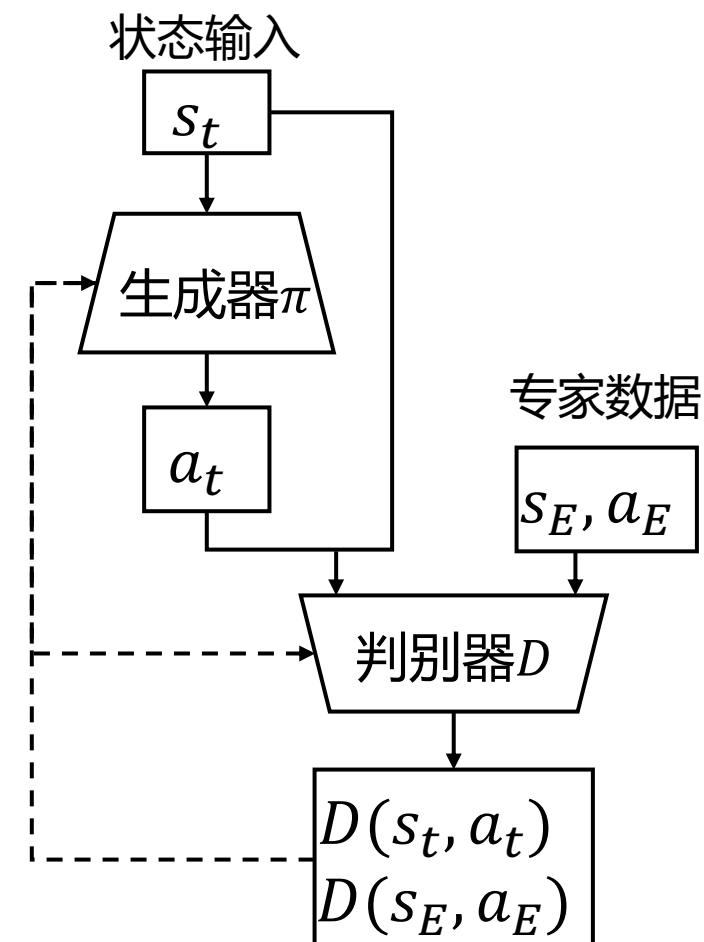
## ➤ 行为克隆中的复合误差问题

- 行为克隆算法**仅仅基于一小部分专家数据进行训练**，因此其策略仅能在这些专家数据的状态分布范围内做出准确预测。
- 然而，模仿学习涉及的是序贯决策问题，这意味着通过行为克隆学习到的策略在与环境进行交互时无法完全达到最优。一旦策略出现偏差，所遇到的下一个状态可能从未在专家数据中出现过，即**分布偏移问题**。



➤ 定义：借鉴生成对抗网络思想，使学习得到的策略所产生的状态-动作对分布尽可能接近专家策略的分布，即**智能体的占用度量 $\rho_\pi$ 尽量接近于专家的占用度量 $\rho_E$** 。

- (1) GAIL中的策略 $\pi$ （类似于GAN中的生成器）需要**与环境进行互动**，通过执行动作并观察结果来逐步调整自身。
- (2) 判别器 $D$ 的作用则是**评估状态-动作对 $(s, a)$ 是否源自专家**，用来估计状态-动作对 $(s, a)$ 来自学习策略而非专家的概率。
- (3) 判别器的目标是最大程度地区分专家数据与学习策略生成的数据。



- 判别器 $D$ 对应的目标函数定义为：

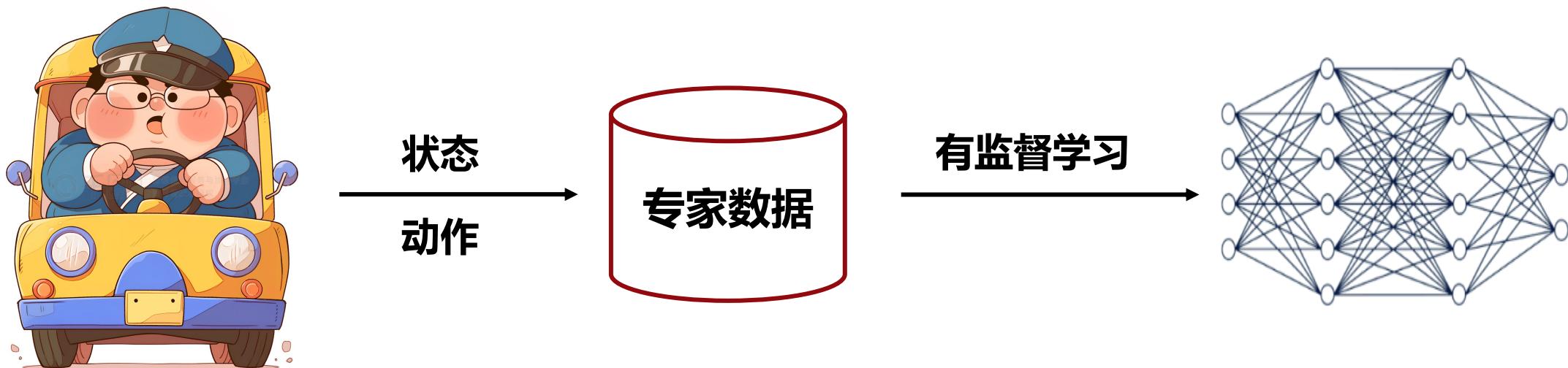
$$\mathcal{L}(\phi) = -\mathbb{E}_{\rho_\pi}[\log D_\phi(s, a)] - \mathbb{E}_{\rho_E}[\log D_\phi(s, a)] \quad (3)$$

其中，判别器 $D$ 的参数 $\phi$ 决定了其区分能力。

- 模仿者的优化目标是生成能够欺骗判别器的轨迹，使其难以分辨这些轨迹是否出自专家。为此，判别器的输出作为奖励信号用于训练模仿者策略，即**当模仿者在状态 $s$ 下执行动作 $a$ ，对应的状态-动作对 $(s, a)$ 被提交给判别器，其输出值作为奖励。**利用标准的强化学习算法，依据这些奖励优化模仿者策略，使其生成的数据分布逐渐逼近专家的真实数据分布，实现有效的模仿学习。

## ➤ 模仿学习的优点

- ✓ 通过不断与环境交互，采样最新的状态-动作对，具备**快速学习能力**。
- ✓ 由于直接利用专家的先验知识，**显著减少**智能体在试错过程中所需的时间和资源。
- ✓ **避免强化学习中常见的奖励稀疏或奖励欺骗问题**，提供更稳定的引导。



## ➤ 模仿学习的缺点

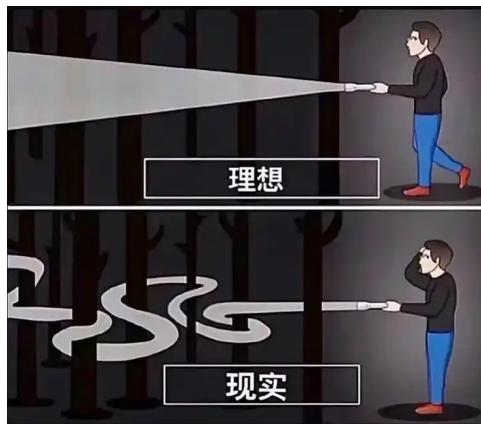
- 面临专家演示**数据局限性和偏见的问题**，可能导致学习到的**策略泛化能力不足**。
- 模仿学习通常**难以直接处理环境或任务变化带来的不确定性**，需要与其他方法结合以提高适应性。

## ➤ 模仿学习的发展方向：与强化学习相融合

- 首先通过**模仿学习快速构建基本行为模型**，再利用**强化学习进行微调和优化**，提高模型的泛化能力和鲁棒性。
- 例如，在机器人抓取任务中，先通过模仿学习掌握基本动作，再通过强化学习优化以提高成功率。

# 知识点3： 让理想照进现实：仿真到真实的迁移

学  
习  
有  
进  
步  
境

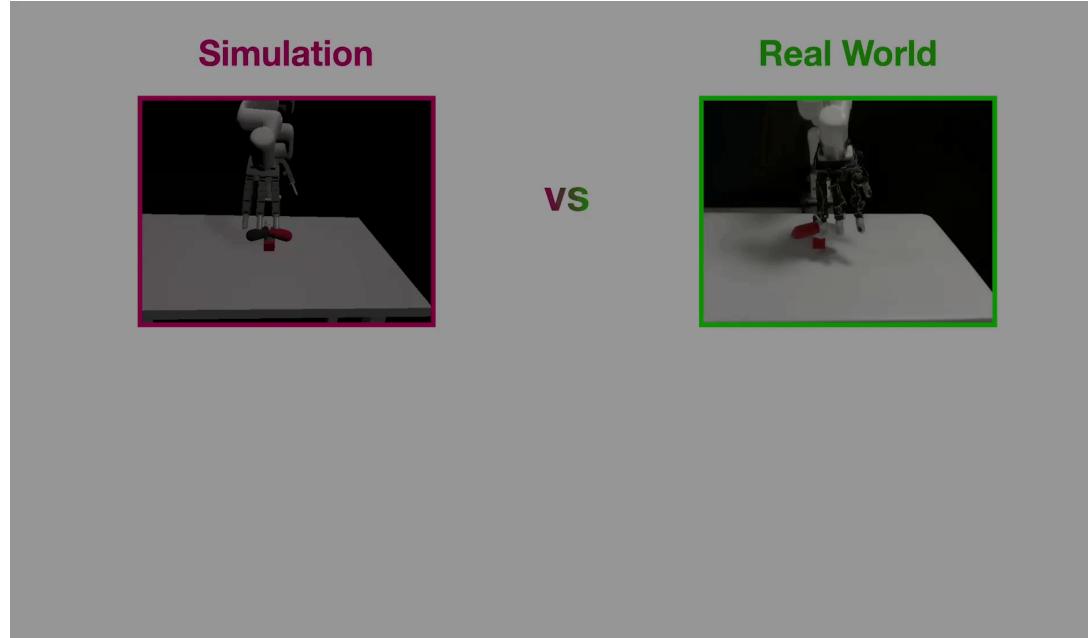


01 Sim2Real的基本概念

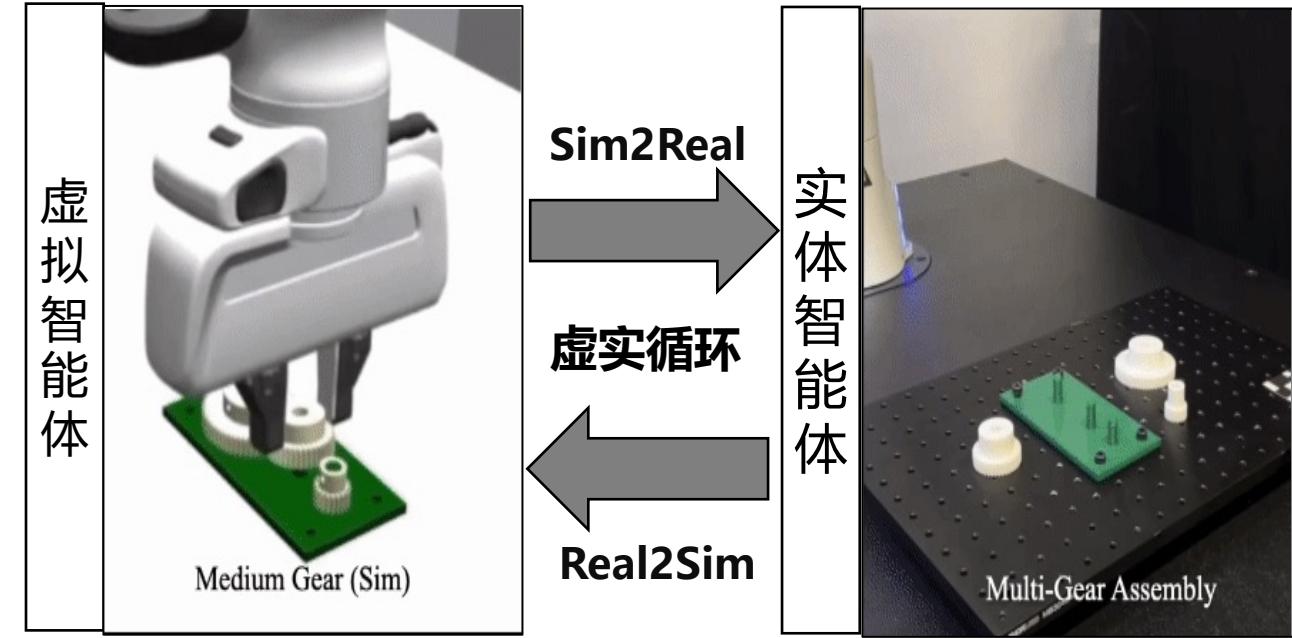
02 Sim2Real的实现方法



- **仿真到真实的迁移 (Simulation-to-Real, Sim2Real)** 关注于将在仿真环境 (如虚拟世界、模拟器等) 中训练得到的模型、算法或策略地迁移到现实世界中的物理实体 (如机器人、自动驾驶汽车等) 上，并确保其在实际应用中表现出良好的性能和稳定性。

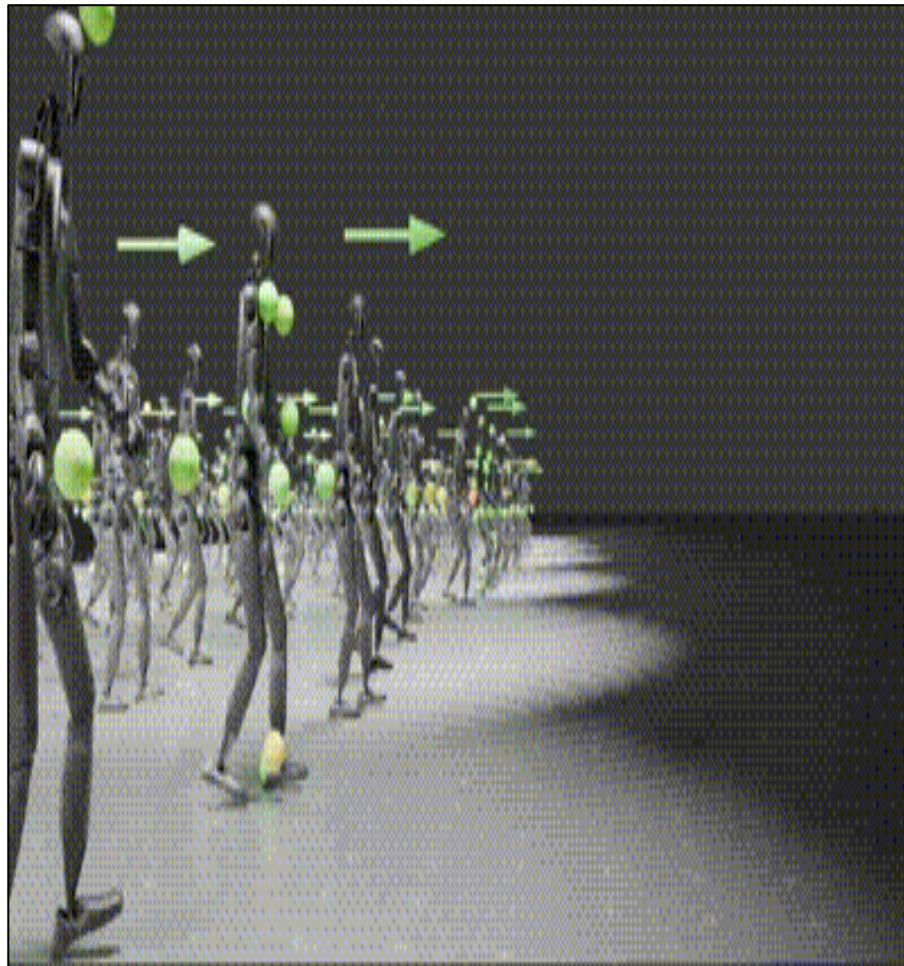


加州大学CyberDemo项目：机器人通过模拟人类的演示来快速完成真实世界的实际任务



工业机器人装配任务中的虚实循环

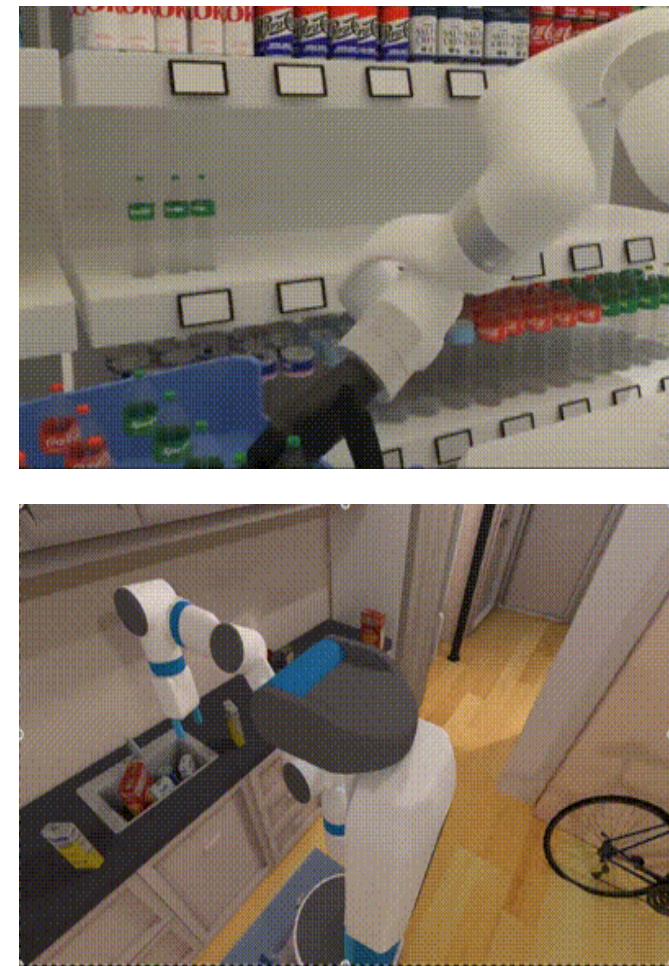
# Sim2Real在不同领域中的应用



人形机器人的运动与操作协调，实现多种控制模式的统一，打造一支步调一致的机器人军团。



多模态数据收集工具：3D对象生成，允许机器人在虚拟的物理环境中进行交互，培养其应对复杂情境的能力。



机械臂的抓取与摆放操作：机械臂在虚拟环境中进行大规模、多样化的操作学习。

## 1. 构建高精度、高逼真度的仿真环境

概述：构建高精度仿真环境是实现Sim2Real的关键步骤，通过**理解现实世界、选择合适工具、精细建模与校准**等方法，可以创建接近现实的仿真环境。**引入不确定性因素**并进行验证与优化，有助于提高模型在现实世界中的性能和稳定性。



3D表示的生成效果



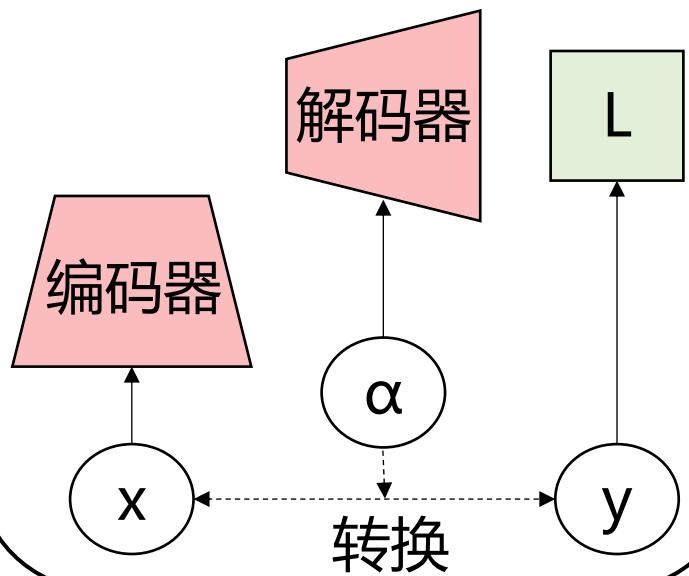
提示词：一片宁静的珊瑚礁

**生成式世界模型**通过模拟环境状态的变化与数据建模，成为一种构建和评估仿真环境的有效方法，实现对物体、场景、动作等要素的准确抽象与模拟。

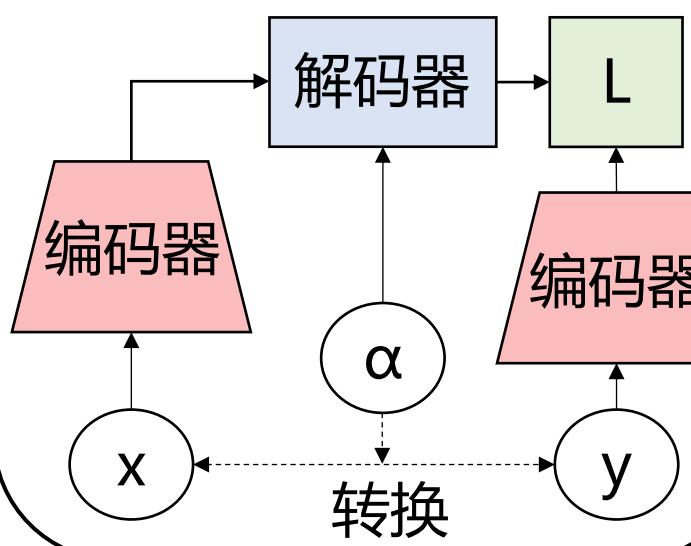
## 1. 构建高精度、高逼真度的仿真环境

生成式世界模型的设计方法分为：**基于生成的方法**、**基于预测的方法**以及**知识驱动的方法**。

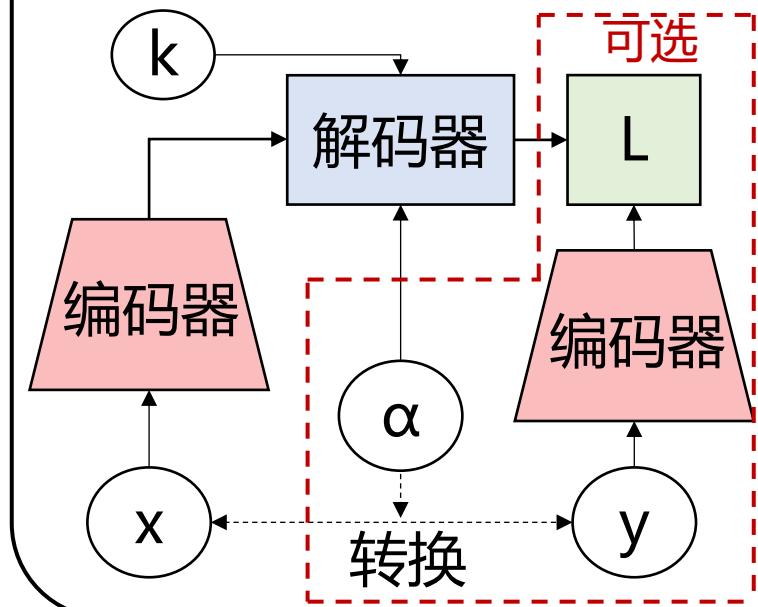
(a) 基于生成的方法



(b) 基于预测的方法



(c) 知识驱动的方法



## 2. 数据驱动的方法

机器人演示数据



打开机柜



将胡萝卜放在盘中

人体演示数据



切割甜椒



清洗玻璃

专家演示数据

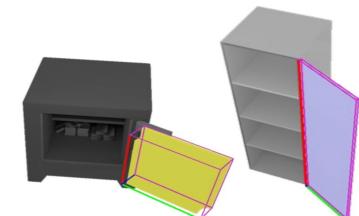


将蓝色块放到绿色碗中

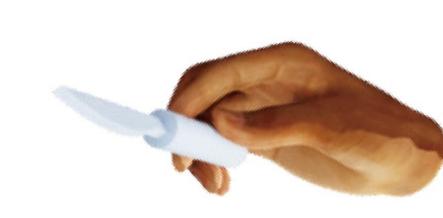


用微波炉加热汤

标注数据



打开门



只用两根手指握住刀柄



视频



点云



RGB+深度图



文本



声音



触觉



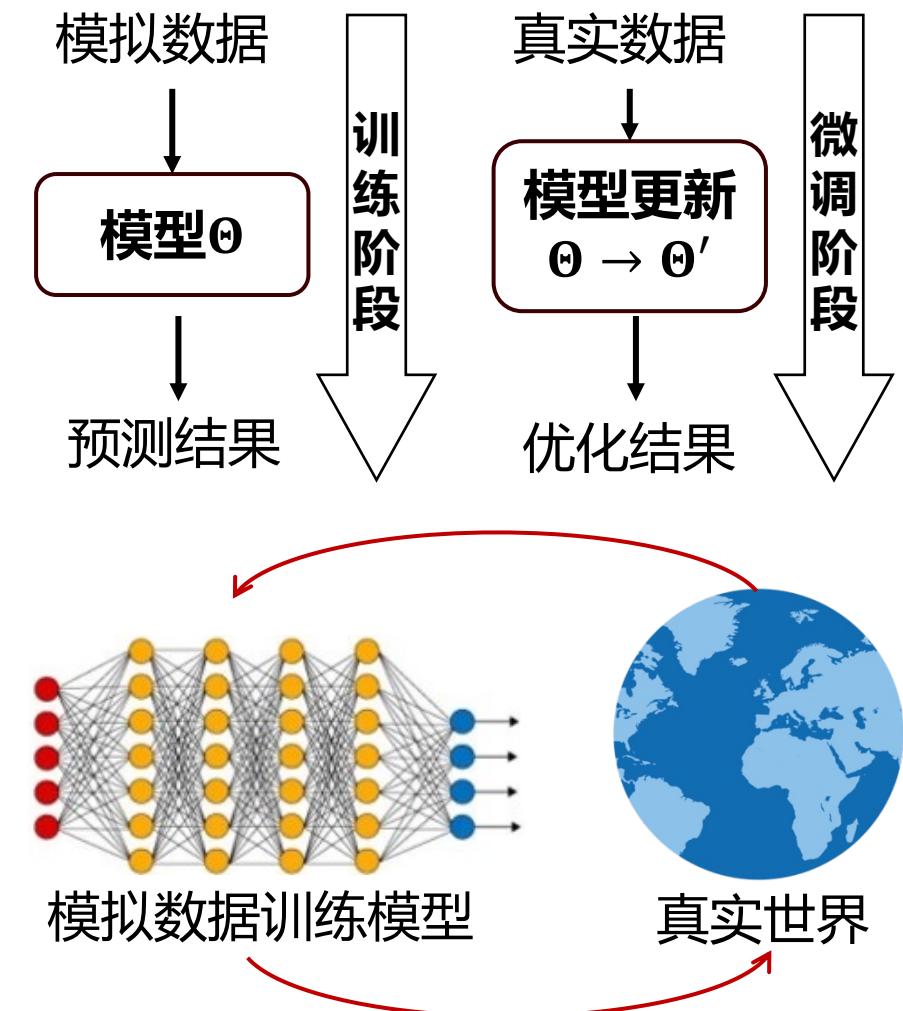
模拟

数据格式

## 2. 数据驱动的方法

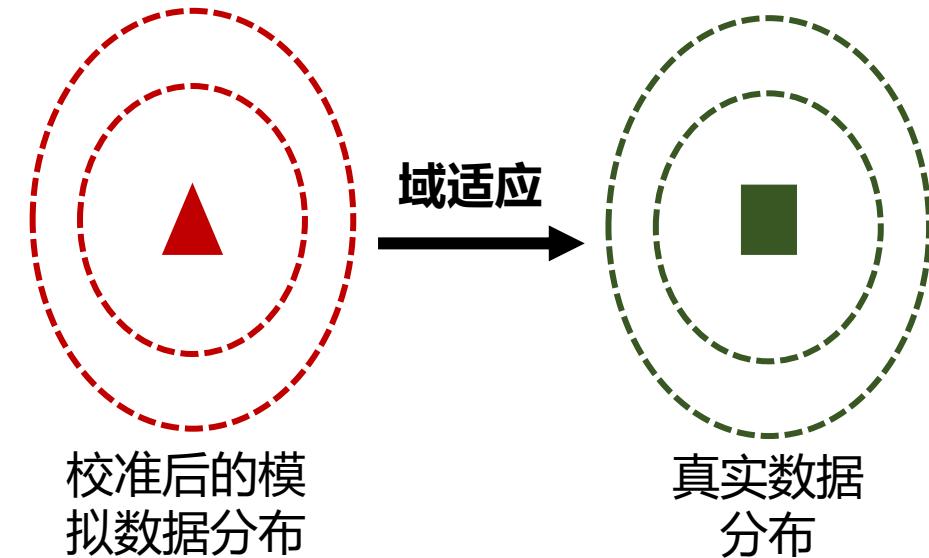
口 **结合模拟与现实数据优化模型**: 在模型训练阶段, **结合模拟数据与现实数据来优化模型参数**, 通过**预训练和微调策略**, 使模型在模拟环境中学习基本技能, 并通过现实数据反馈进行优化。这种方式增强了**模型适应现实环境的能力**。

口 **持续学习与迭代优化**: 通过不断收集新的现实数据并将其用于**模型再训练与更新**, 确保模型紧跟环境变迁与任务需求变化。这一过程提高了**模型在现实世界中的稳定性和可靠性**。



## 3. 域适应与域随机化

口 **域适应 (Domain Adaptation)** : 域适应使模型能够在不同但相关的数据分布上保持高性能，即从仿真环境迁移到现实环境中。通过**识别**仿真与现实环境的**主要差异**，并利用**特征对齐**等技术来缩小这些差异，帮助模型忽略特定噪声，关注有用特征。



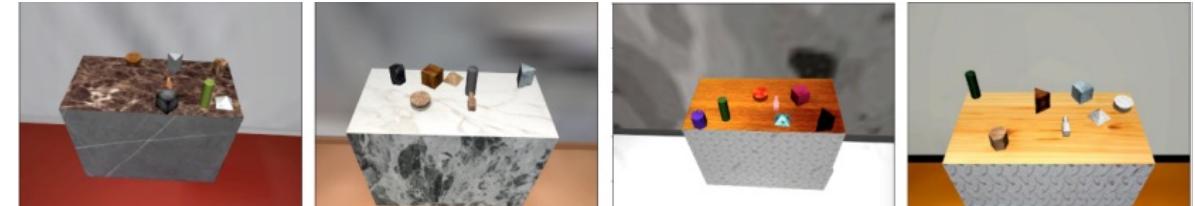
口 **特征对齐与生成对抗网络**: 特征对齐通过**学习共享特征空间**使仿真与现实数据表示接近，而**生成对抗网络**生成**具有现实特性的仿真数据**，作为补充训练集帮助模型适应现实环境。无监督或自监督方法**利用未标记的现实数据微调模型**，设计自监督算法捕捉现实世界的本质特征。

## 3. 域适应与域随机化

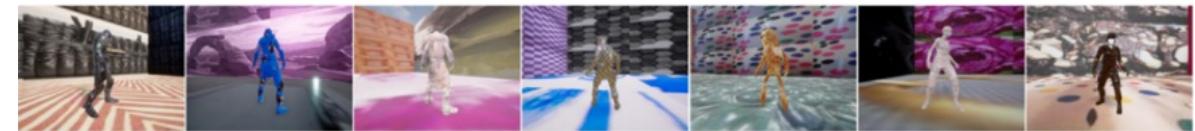
口 **域随机化 (Domain Randomization)** : 域随机化通过**增加仿真环境的复杂性和多样性**来提高模型训练的鲁棒性，通常在仿真训练阶段随机化环境参数。这种方法**不需要现实数据**，依赖模拟器和广泛的随机化策略，鼓励模型学习不依赖特定环境参数的特征，从而提高泛化能力，并**可作为域适应的预处理步骤**以加速适应过程。



(a) 仿真环境的纹理随机化前后对比

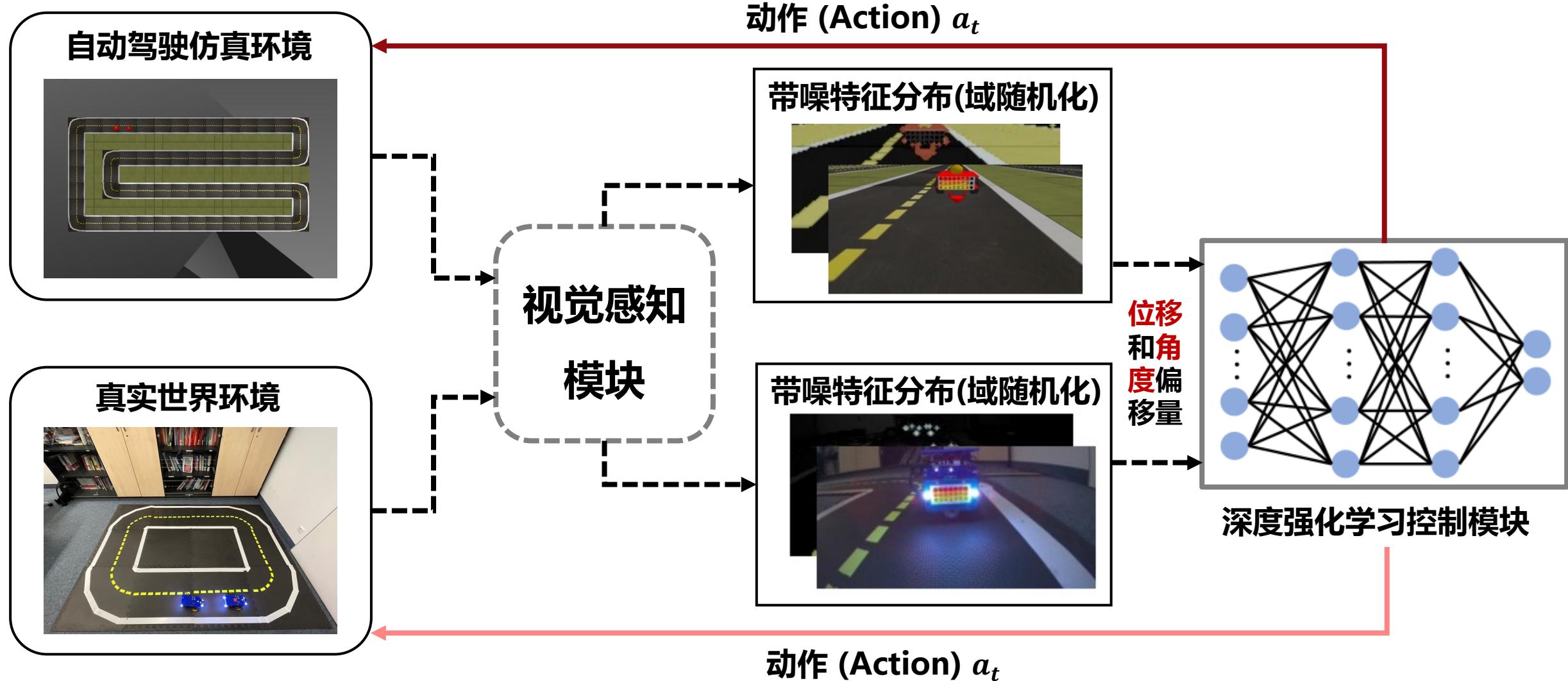


(b) 位姿估计任务中的域随机化



(c) 主动跟踪任务中的域随机化

# 案例总结：自动驾驶领域中的Sim2Real场景



# 知识点4：看智能机器人如何花式干活

学  
习  
机  
器  
人



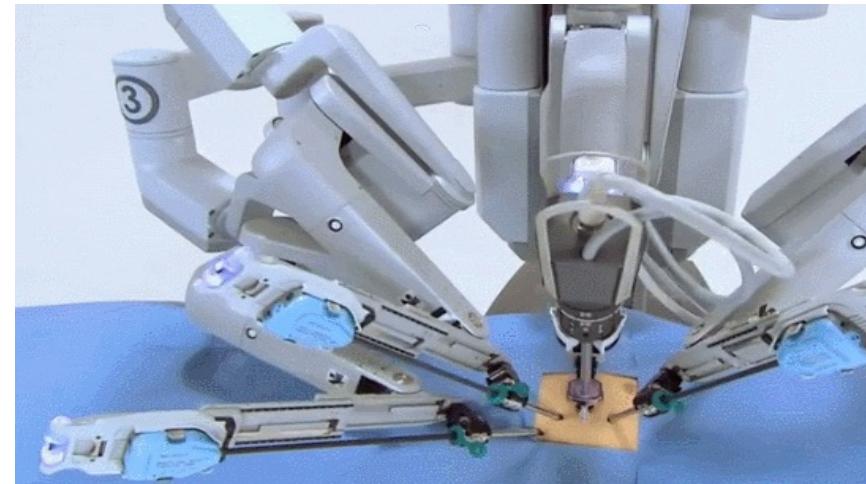
01 机器人能做什么？

02 机器人怎么干活？

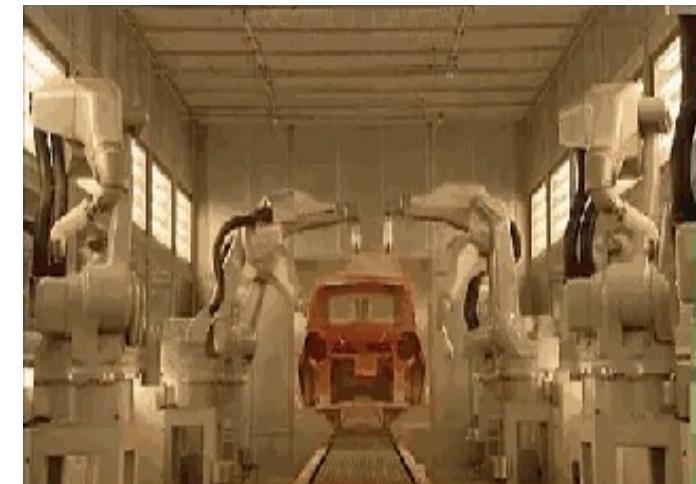
# 机器人能做什么？



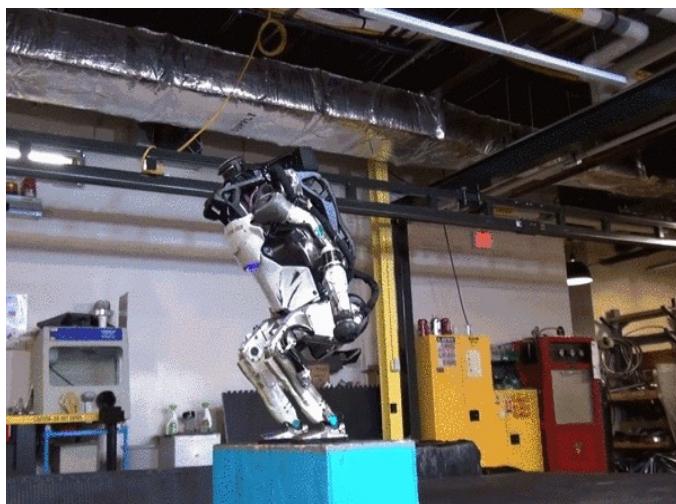
装配空调



手术机器人



工业喷涂



后空翻动作



家庭服务



机器狗巡检

CCTV 13  
新闻

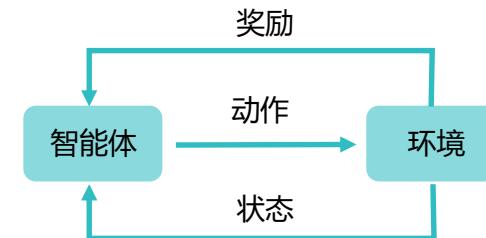
央视  
新闻



从人工智能的发展范式出发，具身系统的研究焦点在于**如何更有效地适应未知环境，特别是在机器人规划与导航等复杂任务中。**

## 具身智能

通过与环境交互得到的反馈进行学习：反馈-动作-反馈-动作.....



## 智能体

具身于本体之上的智能核心，负责感知、理解、决策、控制等核心工作。

## 智能机器人操作



机械臂通过与环境进行多模态交互，将物体移动到指定地点。

## 服务机器人导航



机器人通过与环境进行多模态交互，移动到指定地点。

■ 智能机器人操作是一个综合性的领域，它集成了视觉、语言等多模态输入处理能力，旨在输出精准的机器人动作以执行多样化的具身智能任务，如**物体抓取**任务。

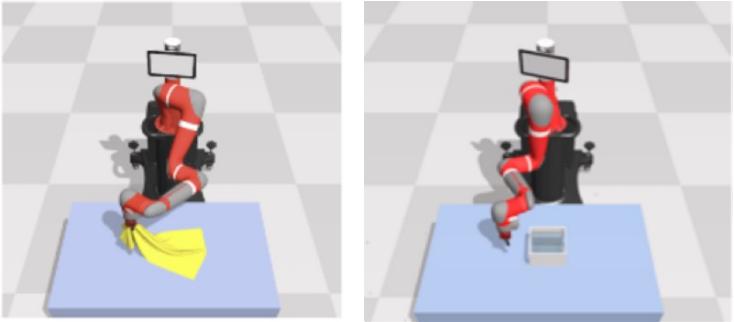


(a) 刚性物体操作-ManiSkill

提示：给定 我需要钉钉子，场景中有什么物件可能有用？

预测：石头。

行动：1 129 138 122 132  
135 106 127



(b) 柔性物体操作-SoftGym

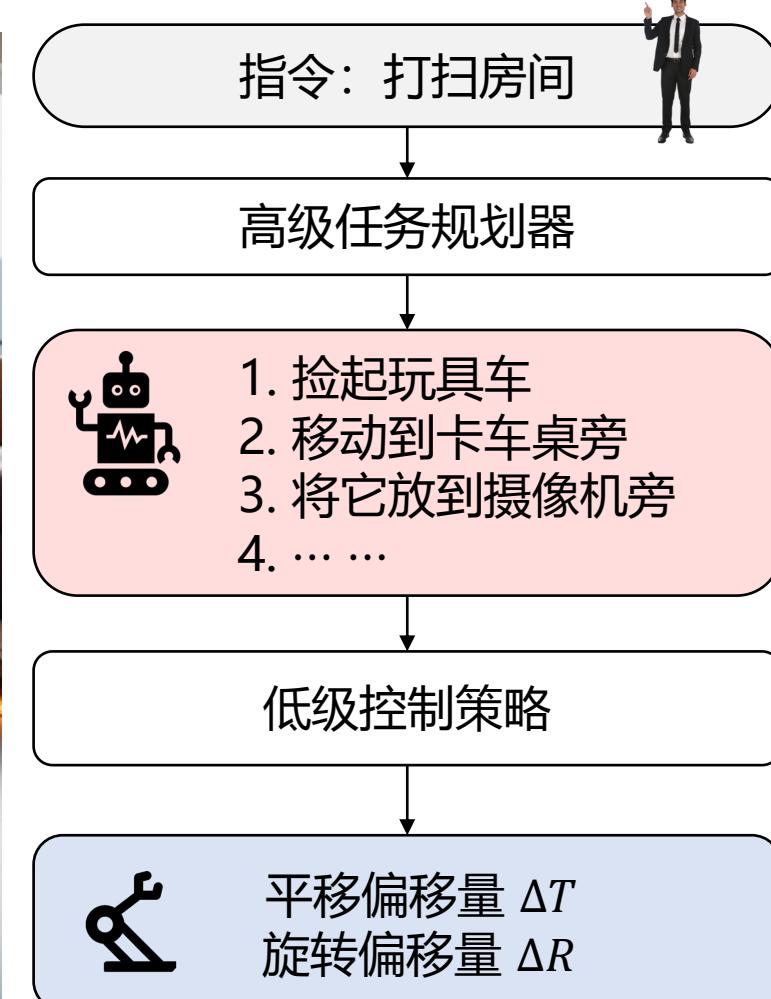
(c) 视觉语言操作

## ➤ 视觉-语言-动作 (Vision-Language-Action, VLA) 的基本概念与操作策略

- 定义：VLA是一种结合了视觉、语言与动作执行的更高级别任务处理框架，旨在结合视觉与语言信息，指导机器人或智能系统完成复杂任务（如清理桌面、拿取物品）。其核心在于其强大的多模态处理能力。
- VLA模型组成：视觉模块负责解析图像数据，语言模块则理解自然语言指令，动作模块据此生成动作指令并控制机器人执行相应的动作。三者之间通过深度协作与交互，使得模型不仅能理解复杂的场景与指令，还能灵活地执行任务，促进机器人综合能力的全面优化与提升。

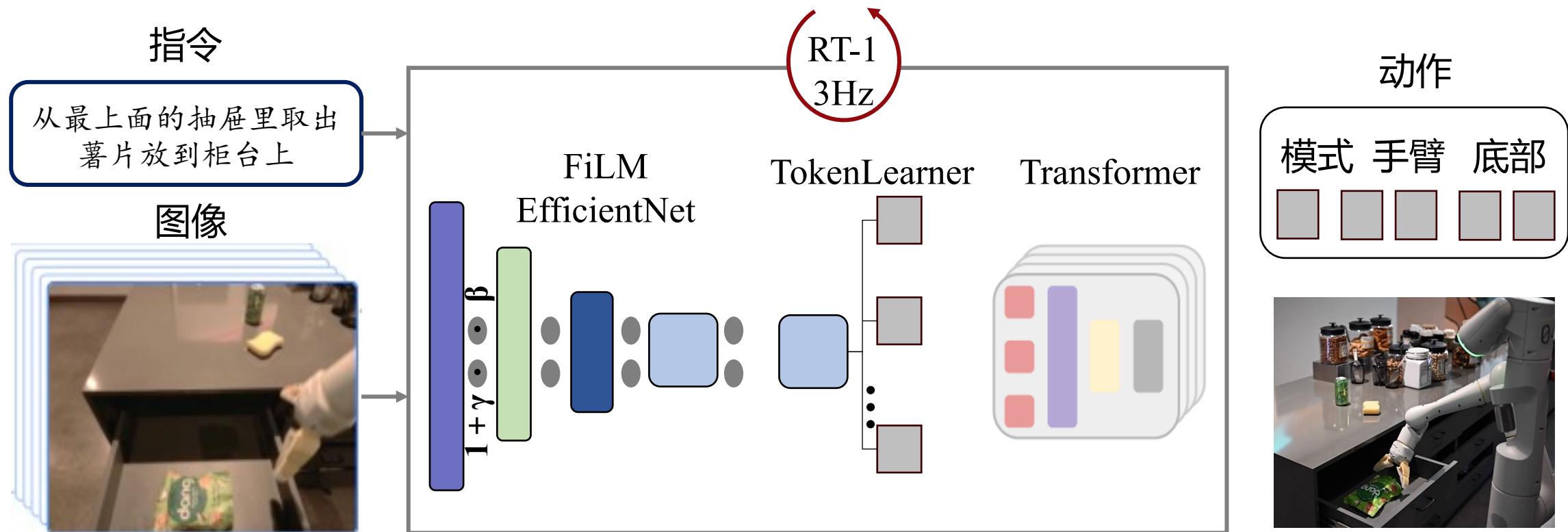
## ➤ 视觉-语言-动作 (VLA) 的基本概念与操作策略

分层机器人操作策略



## ➤ 视觉-语言-动作 (VLA) 的具体实现——Robotics Transformer系列模型

- 2022年12月，谷歌推出了名为Robotics Transformer 1 (RT-1) 的具身智能模型，这是一种多任务处理模型，能够将机器人的输入和输出动作转换为Token形式，从而提升实时控制。



## ➤ 视觉-语言-动作 (VLA) 的具体实现——Robotics Transformer系列模型

□ RT-2在模型设计上进行了重大创新，它将机器人的动作编码成一种独特的文本标记语言，这种创新性的表示方式使得RT-2能够利用互联网级别的庞大视觉-语言数据集进行训练。

互联网级别的视觉问答+机器人动作数据



问：这张照片的内容是什么?  
答：311 423 170 55 244

一只灰色的驴在街上行走



问：Que puis-je faire avec ces objets?(法语)  
答：311 423 170 55 244

Faire cuire un gâteau.



问：对于<任务>机器人应该如何操作?  
答：132 114 128 5 25 156

$\Delta$ 转移 = [0.1, -0.2, 0]  
 $\Delta$ 旋转 = [10°, 25°, -7°]

用于机器人控制的视觉-语言-动作模型

RT2

问：对于<任务>机器人应该如何操作? 答：...



ViT

大语言模型LLM

答：132 114 128 5 25 156

逆标记化

$\Delta$ 转移 = [0.1, -0.2, 0]  
 $\Delta$ 旋转 = [10°, 25°, -7°]

机器人动作描述

协同微调

部署

闭环机器人控制



把草莓放到正确位置



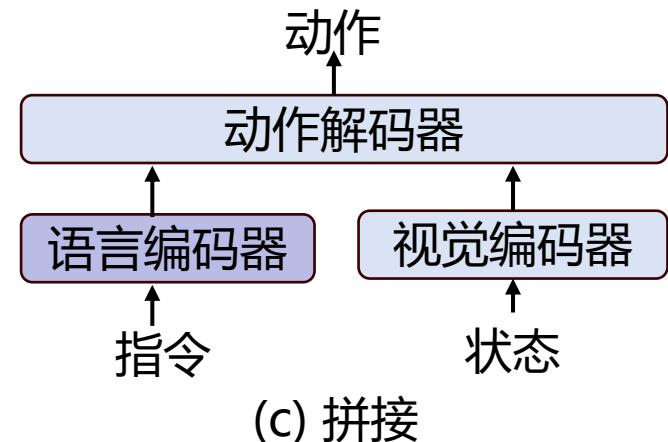
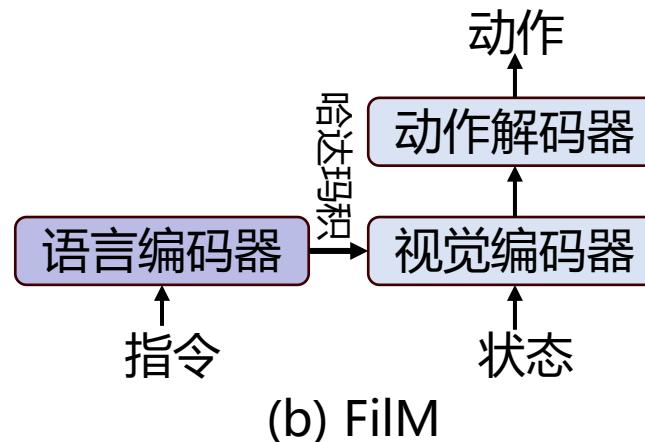
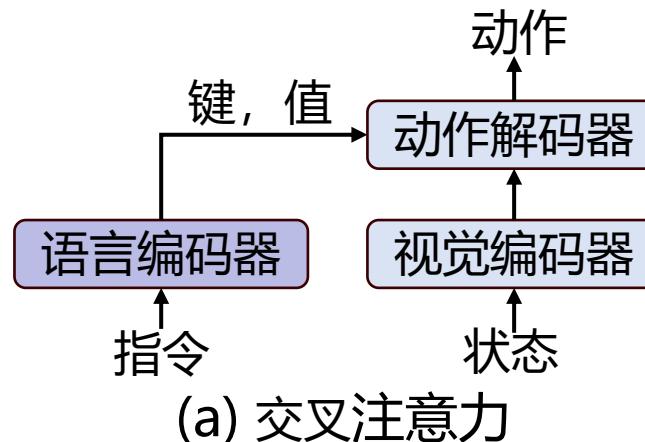
捡起快要掉下的袋子



捡起不一样的东西

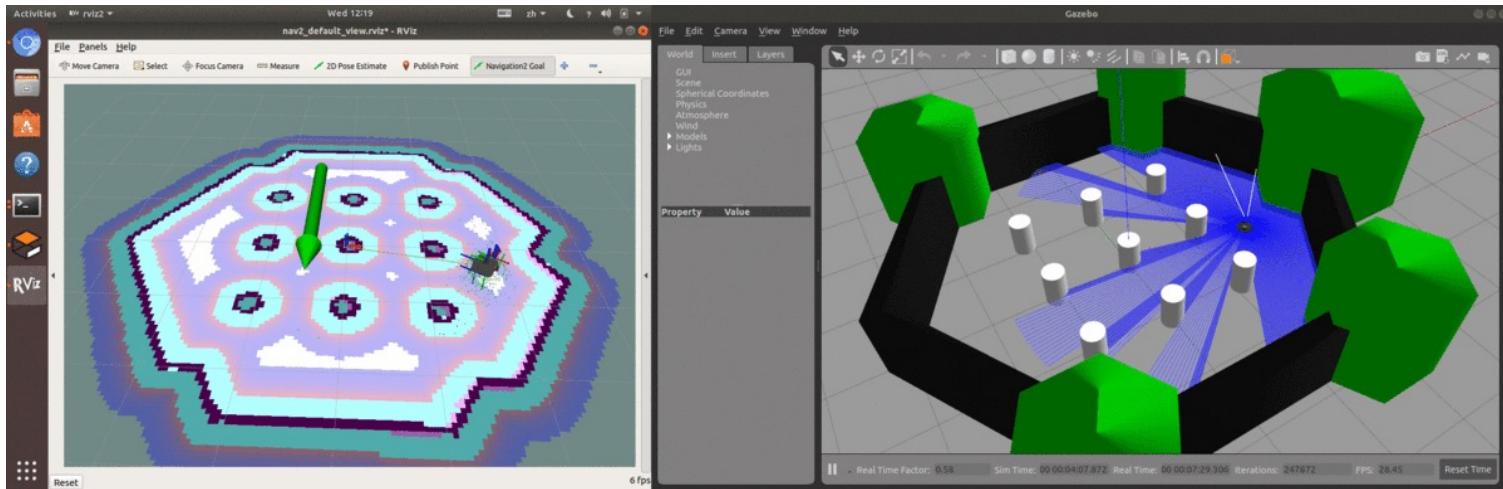
## ➤ 其他VLA的典型技术

- **预训练视觉编码器**: 通过大规模数据集的训练，获得能够捕捉复杂视觉特征并生成高质量视觉表示的模型。
- **环境动力学建模**: 包括利用前向动力学方程来预测物体在给定力作用下的运动轨迹，以及利用逆向动力学方程来推断产生特定运动所需的力或力矩。 **世界模型**
- **视觉-语言融合机制**:

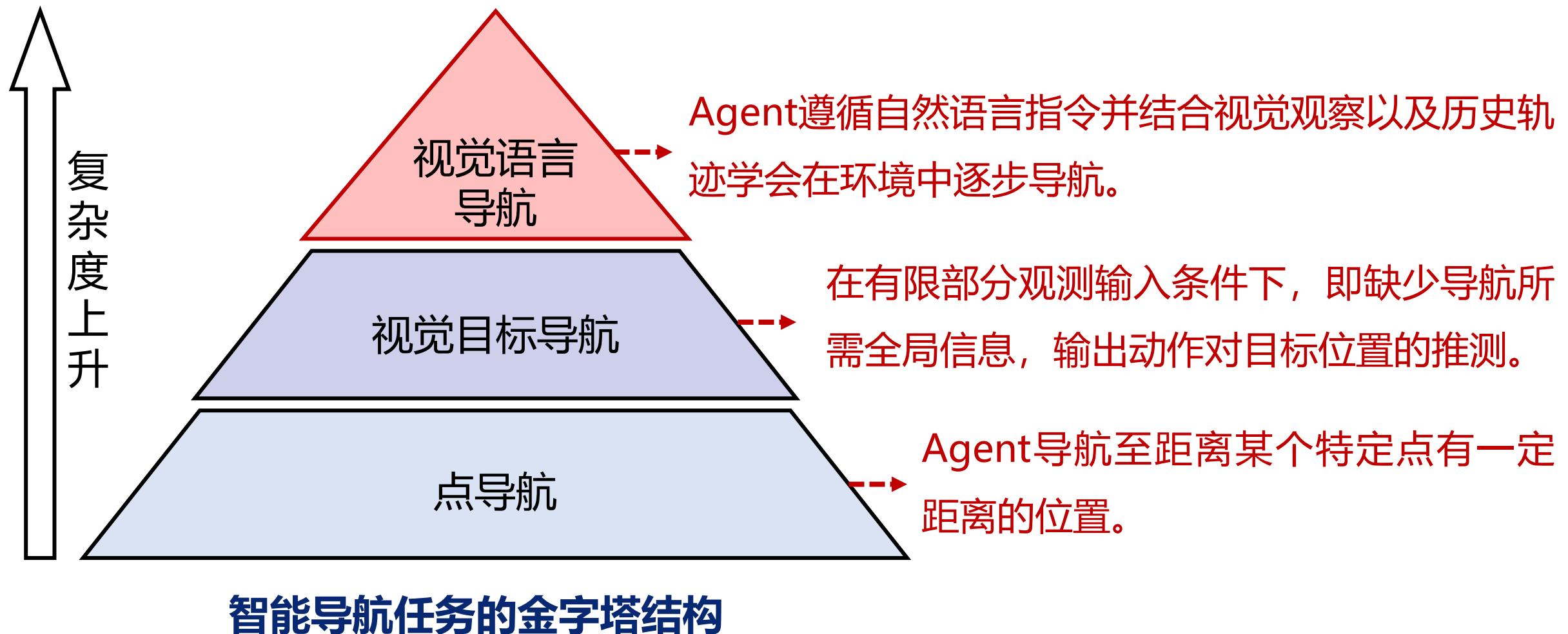




■ 服务机器人导航要求机器人在**未知且复杂的环境中**，  
**仅凭目标位置和多个视角的观测**（主要是视觉信  
息），通过集成的感知硬件与先进算法进行深度分  
析，并在与环境的持续交互与反馈中，高效且准确  
地在限定步数内抵达指定位置。







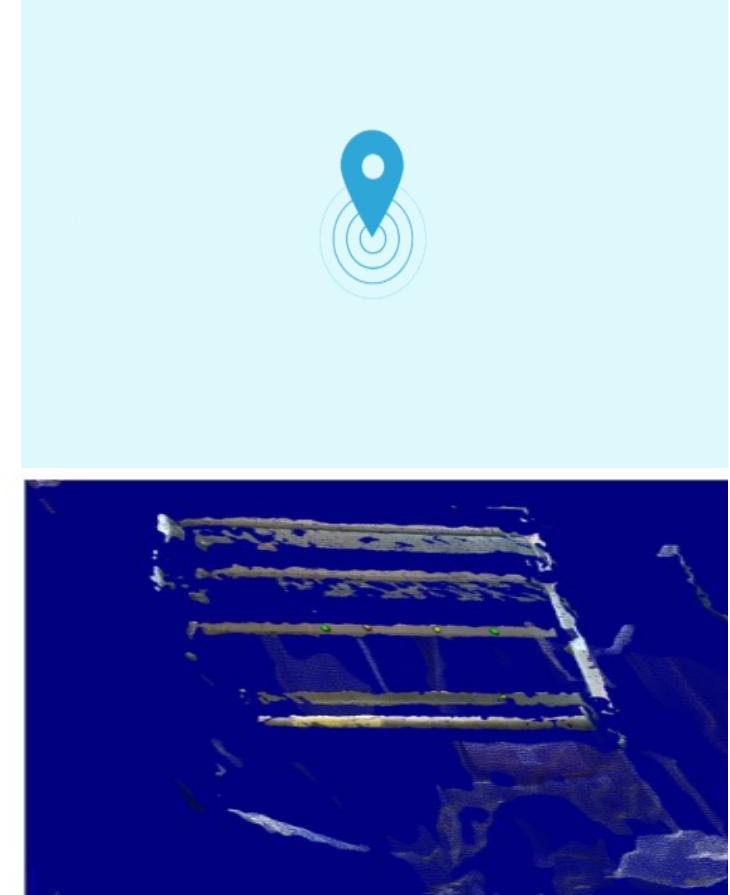
## ➤ 点导航任务

□ **初始化与目标定位：** Agent通常在环境的原点 (0,0,0)

初始化，**目标点通过相对原点的三维坐标指定。**

□ **导航硬件与位置感知：** 由于室内环境中定位的不准确性，目前视觉导航工作转向**基于RGB-D的在线定位**，不再依赖传统的GPS和指南针。

□ **学习型导航方法：** 基于学习的点导航方法探索端到端解决方案来处理未知环境中的导航，利用多种感官输入（如彩色图像、深度图及最近的观测动作），**无需真实地图或精确姿态信息。**



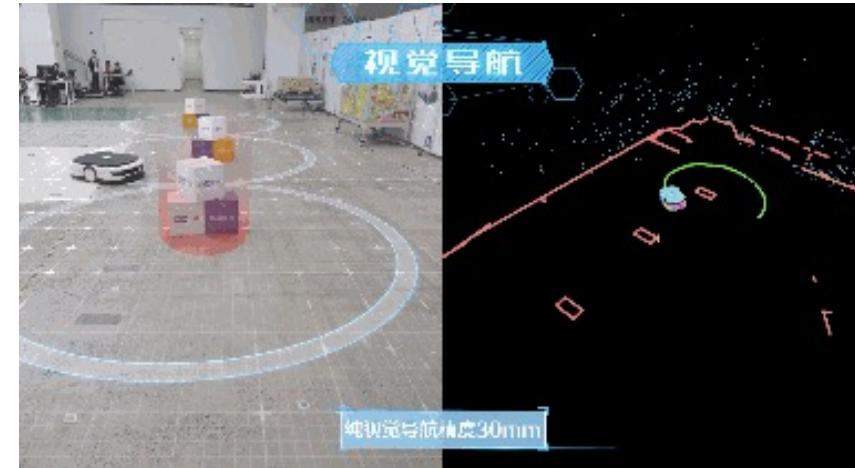
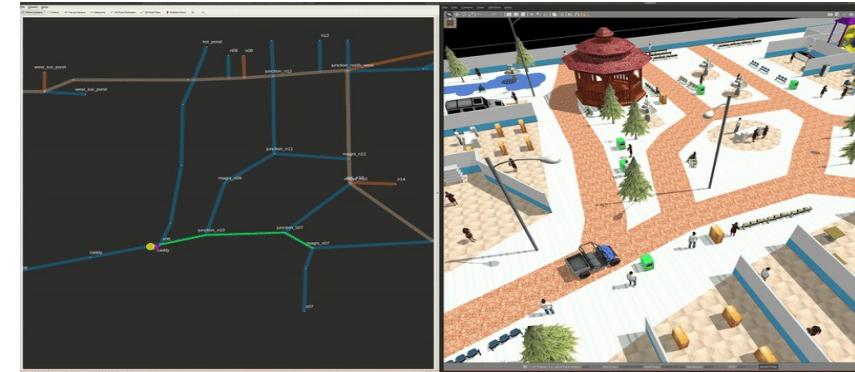
履带机器人观察到的楼梯RGB-D数据

## ➤ 视觉目标导航任务

□ **任务定义：** Agent仅依据第一视图的RGB图像导航至目标物体 $t$ ，该物体属于目标类别集合 $T$ ，并且**Agent事先并不了解环境。**

□ **数学描述：** 假设在每一次的导航过程中，Agent和目标物体的初始位置都是随机选取的，Agent可以通过策略学习网络 $\pi$ 根据当前的RGB图像 $I$ 和目标物体的特征向量 $w_t$ 采样动作 $a$ ，即 $a \sim \pi_\theta(I, w_t)$ 。

其中，动作 $a$ 属于集合 $A$ ，其中前行距离和旋转角度可以根据实际任务需求来定义，最后“Done”表示Agent已经找到目标时的状态，从而结束这一导航过程。



## ➤ 视觉-语言-导航 (VLN)

**框架：**视觉-语言-导航 (Visual Language Navigation, VLN) 任务旨在使得Agent遵循自然语言指令并结合视觉观察以及历史轨迹学会在环境中逐步导航。

转身，从门口进入房子。穿过客厅，经过椅子和沙发，穿过冰箱左侧的门口。向左急转弯，进入餐厅，在到达桌子之前，在门口前等待。

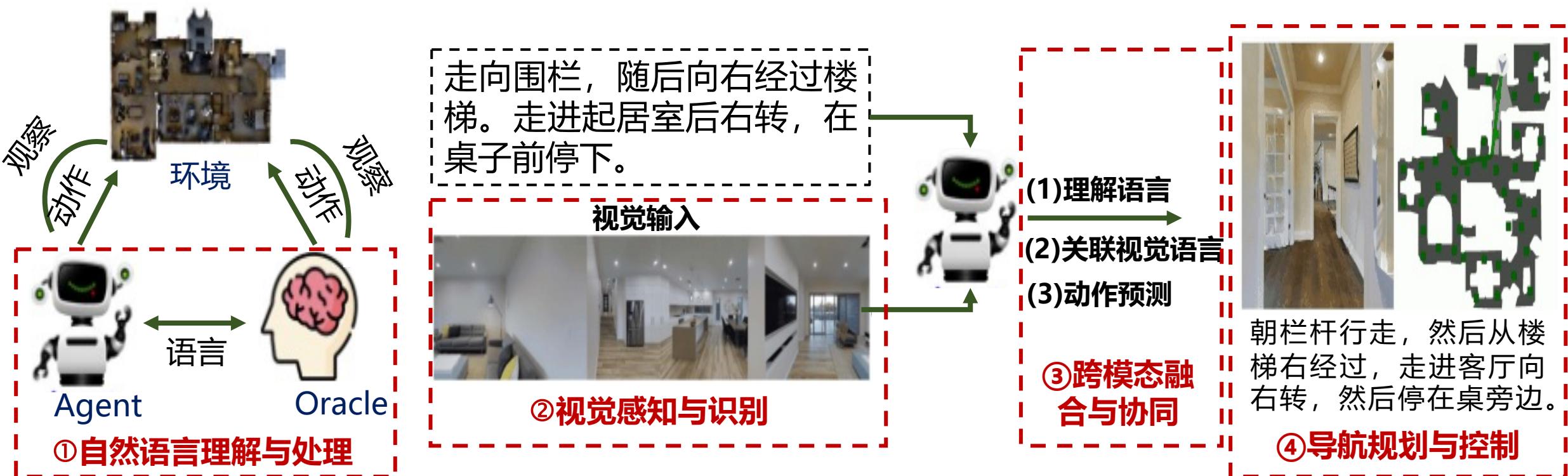


沿着走廊直走。走到走廊尽头后左转，穿过那扇门。稍微左转，穿过那个有四个椅子围着咖啡桌的房间。在大照片前停下，照片上有人在涂口红。



## ➤ 视觉-语言-导航 (VLN)

**方案：**借助于视觉观察、环境交互以及奖励机制，构建强化学习框架；利用语言指令指导Agent完成语言理解、视觉与语言关联以及动作预测，使得智能体移动到指定位置。





山东大学  
SHANDONG UNIVERSITY



中国图象图形学学会  
CHINA SOCIETY OF IMAGE AND GRAPHICS

- 山大AI (爱) 说唱 -

热爱改变世界

——宇树科技 王兴兴

作词 孙浩泽（核科学与能源动力学院2023级本科生）

丛润民（控制学院教授）

演唱 孙浩泽（核科学与能源动力学院2023级本科生）

策划 郑柔澄（校团委副书记）

丛润民（控制学院教授）

张伟（控制学院教授）

制作 郑彦童（新闻传播学院2023级本科生）

杜云飞（新闻传播学院2024级本科生）







山东大学  
SHANDONG UNIVERSITY

《人工智能通识》AI For Everyone

# 智能之躯——具身智能

学无止境 气有浩然

教育部-华为“智能基座”课程