

哈尔滨工业大学  
大学生创新创业训练计划项目  
中期检查报告

项 目 名 称：双足人形机器人 AlexBot

执 行 时 间：2023 年 9 月至 2024 年 9 月

负 责 人：黄一 Alexhuge

学院及专业：未来技术学院

# 1、项目简介

现拟打算设计一套以伺服电机为驱动器的十自由度的双足机器人下肢 AlexBot，以研究双足机器人运动规划与控制，首先对于机器人的结构以及控制系统进行设计以及拓扑优化；其次，进行以髋部坐标系为基坐标系的正运动学及逆运动学分析，并计算了机器人的实时质心位置以及机器人的质心相对于足端的雅可比矩阵；随后，说明了基于质心期望力的二维线性倒立摆理论，分别对单足支撑阶段以及双足支撑阶段的支撑脚与摆动脚进行规划，再将二维的情况推展到三维；之后，将线性倒立摆的行走数据当作参考在虚拟环境中对机器人进行强化学习训练，其中包括神经网络的结构的设计，强化学习算法超参数的设置，奖励函数、训练终止条件的设置以及将仿真迁移到现实机器人中需要进行的动态随机化策略；最后，为了验证方法的可行性，分别在仿真以及现实中实验。

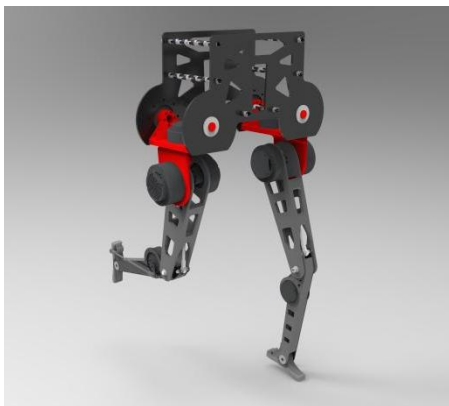


图 1-1 AlexBot 双足机器人设计仿真图

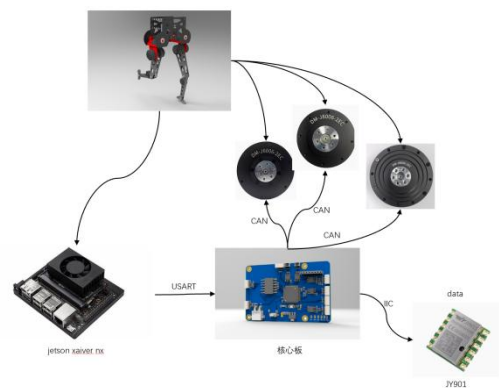


图 1-2 双足机器人架构图

## 2、立项背景

### 2.1 研究背景

在我国社会结构和经济发展的新阶段，人口老龄化问题已成为一个不容忽视的社会现象。根据最新的统计数据，65岁及以上的老年人口数量已经超过1.8亿，占总人口的比例达到12.6%，并且这一比例预计将在未来持续增长。这一趋势的出现，对我国的劳动力市场产生了显著影响。随着劳动力人口的减少，劳动力成本不断上升，这不仅对传统的经济发展模式提出了新的挑战，也对现有的社会保障体系带来了前所未有的压力。

在应对人口老龄化的挑战中，机器人产业的发展显示出了巨大的潜力和重要性。通过引入机器人技术，不仅可以有效缓解劳动力短缺的问题，还能显著提高生产效率和质量，从而推动经济的持续健康发展。



图 2-1 全球及主要国家 65 岁以上人口占有率

## 2.2 研究目的及意义

随着科技的不断进步，人们对机器人技术的需求和期待也在不断提高。在这一背景下，足式机器人，尤其是双足机器人的研究和开发成为了机器人领域的热点之一。双足机器人相较于其他类型的机器人，具有一系列独特的优势和潜力：它们不仅能够提高生产效率，保障人员安全，还能够在服务领域发挥重要作用，推动人机交互和人机协同技术的发展。



图 2-2 人形机器人应用于人类日常生活

## 2.3 研究现状

从上个世纪中叶开始，双足机器人的研究逐渐成为全球科技领域的热点，研究者们从对人类和各种动物的步行理论研究开始，随着科学技术领域的飞速发展，双足机器人的行走又融合了运动控制、人工智能等方式，双足机器人的行走理论被逐渐完善。美国和日本在这一领域的发展尤为突出，它们的研究成果不仅推动了机器人技术的进步，也引领了全球的研究趋势。

美国波士顿动力公司的 Atlas 机器人和日本本田公司的 ASIMO 机器人就是近十年来在学术界引起广泛关注的两大“网红”机器人。这些机器人不仅在技术上取得了飞速的发展，而且在运动规划和稳定性控制方面也取得了显著的成就。半个世纪以来，这两个国家涌现出了许多震惊学术界的先进机器人，如本田的 ASIMO 和波士顿动力的 Atlas，它们不仅在技术上达到了新的高度，而且在实际应用中也展现出了巨大的潜力。



(a) ASIMO 机器人



(b) ATLAS 机器人

图 2-3 美国和日本公司研发的部分人形机器人

特斯拉公司的人形机器人 Tesla Bot 在经历了多版本迭代之后，其研究进展远超预期。从

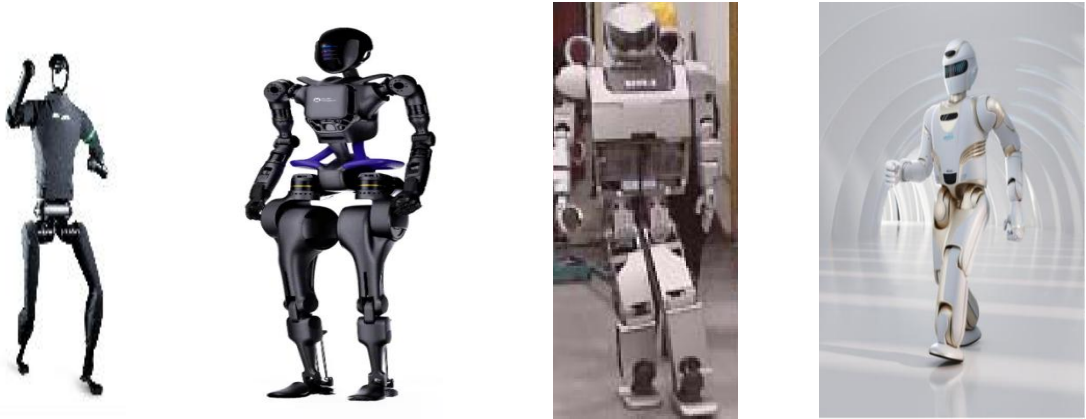
步态行走技术来看，Tesla Bot 从 2022 年完成第一次步态行走到 2023 年已经能够实现在工厂内稳步行走；从功能来看，Tesla Bot 从第一版仅具备抓取、搬运、上下料功能到具备双手处理复杂任务的能力，同时 Tesla Bot 关节电机转矩控制的精度已经可以达到在运动过程中不打碎鸡蛋的水平。与此同时，Tesla Bot 与特斯拉车辆共享核心自研的感知技术（FSD Computer 和 DOJO Training），并且其搭载的 Autopilot Camera 所具备的强大的视觉处理能力使得 Tesla Bot 在人形机器人研究领域处于领先地位。



(a) 特斯拉人形机器人 Optimus (b) Optimus 结构示意图

图 2-4 特斯拉公司研发的 Tesla Bot

中国在人形机器人产业的发展过程中，经历了从追赶到领先的转变，并逐步超越了早期投入此领域的日本、美国和韩国。其中创业企业典型代表包括优必选科技、追觅科技、智元、宇树、傅利叶智能等；科技企业典型代表则为小米、科大讯飞；高校领域则以哈尔滨工业大学研制的 GoRoBoT 机器人、北京理工大学研制的 BHR 系列双足机器人、清华大学研制的 Stepper-2D 机器人为代表。



(a) 宇树 H1 机器人 (b) 傅里叶 GR-1 机器人 (c) BHR-2 机器人 (d) 优必选 Walker 机器人

图 2-5 国内部分双足人形机器人



## 2.4 研究趋势

随着人形机器人产业奇点将至，市场将会呈现出更加明显的马太效应。在未来很长一段时间内，人形机器人的领域将围绕人类步态与行为的高度模仿、环境感知与交互能力的提升、轻量化和能效优化等一系列方向进一步深化发展。未来随着技术的不断成熟和创新，人形机器人尤其是双足机器人有望在更多领域展现其独特的价值和潜力。

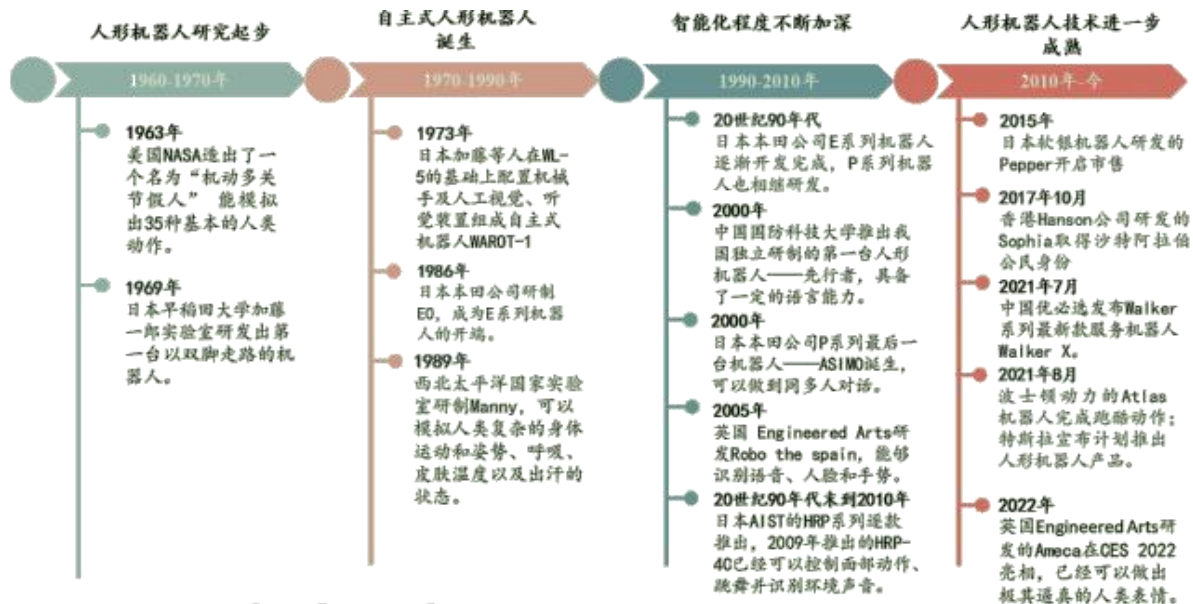


图 2-6 人形机器人发展历程及未来发展趋势

## 3、项目方案

本项目我们对双足机器人的行走进行研究，通过研究基于质心期望力的二维线性倒立摆拓展到三维线性倒立摆，并通过强化学习算法得到神经网络控制器，使得双足机器人可以执行行走任务，主要研究内容如下：

(1) 本项目想设计一套大型伺服双足机器人以供算法研究使用，同时严格控制成本并且编写教程开源全结构、硬件、算法，提供一套双足的完整的解决方案，产生一定的影响力。

(2) 研究基于二维线性倒立摆的步态规划器，通过线性倒立摆的动力学性质进行规划，分别对单足支撑期和双足支撑期进行规划，应用足底到质心的雅可比矩阵，将动力学简化模型应用到复杂机器人上，通过反馈重新分配腿部力矩，使得双足机器人的能量足以越过势能高点，形成稳态行走的步态，而后再将二维线性倒立摆拓展到三维线性倒立摆算法。

(3) 应用基于质心期望力的线性倒立摆所得到的规划步态形成参考轨迹，使用强化学习算法，在虚拟环境中迭代训练，运用动态随机化策略，提高网络的适应度，以便将仿真中训练的网络迁移到现实环境中。

(4) 将网络模型迁移到现实机器人身上，进行 Sim to Real 的行走实验，积累经验，提升机器人的鲁棒性。

#### 4、项目实施的进展情况及初步取得的成果

## 4.0 整体架构

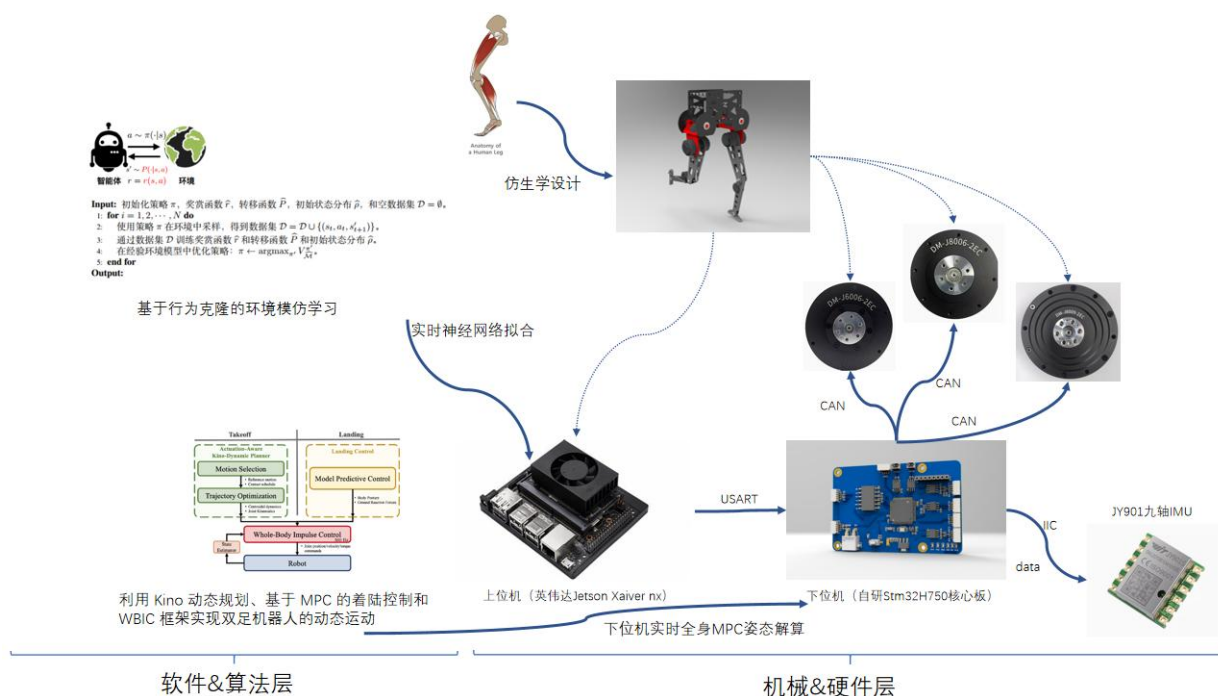


图 4-1 双足机器人架构图

本项目期望设计一套低成本，全开源，低入门门槛，完整系统的双足下肢机器人（之后会设计上肢），方便形成双足开源生态，增加影响力和为更多工程师提供帮助。（类比 MIT 的 Cheetah 四足机器人）



图 4-2(a) MIT Cheetah 四足机器人

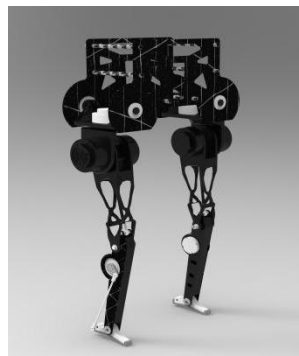


图 4-2(b) HIT Alexbot 双足机器人

本项目涉及两套双足机器人结构，一套为研究三维线性倒立摆所需要的十自由度桌面级机器人，其各个执行器都是在机器人矢状面内出力，配置以人类的下肢关节配置为参考，使得机器人得以在三维平面内行走。另一套为大型研究步态以及迭代优化双足架构的大型双足机器人。

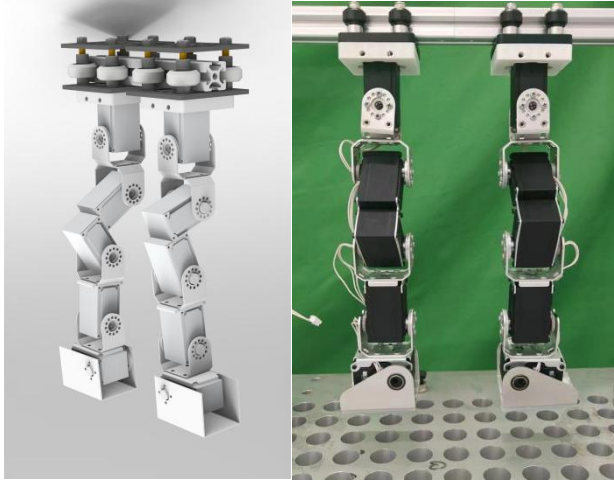


图 4-3(a) 研究性桌面级双足机器人

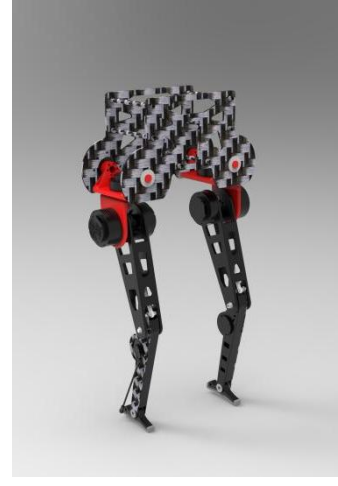


图 4-3(b) AlexBot 双足机器人下肢

我们先对两套双足机器人结构进行设计，搭建控制系统，此外为后续进行步态规划打下基础，对两套机器人分别进行正运动以及逆运动学分析，使得给定足端位姿可以求解各关节角度，给定各关节角度可以求解足端位姿。同时，通过给定关节角度以及髋部位姿，得出机器人实时的质心位置，求解足端到机器人质心的雅可比矩阵。

$$A_1 = \begin{bmatrix} C_1 & 0 & S_1 & 0 \\ S_1 & 0 & -C_1 & 0 \\ 0 & 1 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} C_2 & -S_2 & 0 & a_2 C_2 \\ S_2 & C_2 & 0 & a_2 S_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} C_3 & -S_3 & 0 & a_3 C_3 \\ S_3 & C_3 & 0 & a_3 S_3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_4 = \begin{bmatrix} C_4 & 0 & S_4 & 0 \\ S_4 & 0 & -C_4 & 0 \\ 0 & 1 & 0 & d_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} C_5 & 0 & -S_5 & 0 \\ S_5 & 0 & C_5 & 0 \\ 0 & -1 & 0 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_6 = \begin{bmatrix} C_6 & -S_1 & 0 & 0 \\ S_6 & C_6 & 0 & 0 \\ 0 & 0 & 1 & d_6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

图 4-4(a) D-H 法正解双足姿态

$$A_1^{-1}T = A_2A_3A_4A_5A_6$$

$$\begin{bmatrix} C_1 & S_1 & 0 & 0 \\ 0 & 0 & 1 & -d_1 \\ S_1 & -C_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

图 4-4(b) D-H 法逆解双足姿态

同时应用对于机器人的运动学计算结合线性倒立摆算法对支撑腿、摆动腿以及双足支撑阶段进行规划和控制，同时在仿真物理环境 Mujoco 中进行稳定行走，生成二进制数据，通过 agent 与环境的自主迭代互动，向着累计回报最高的方向进化，为了提高效率，同时为了避免强化学习学习的步态与人类行走步态相差过大，采用基于参考轨迹的强化学习方法，我们利用近端策略优化的强化学习算法训练神经网络控制器，在基于质心期望力的线性倒立摆的运动规划器所形成的参考轨迹的基础上进行学习训练，对于算法的流程、状态及动作空间以及奖励函数进行了设计，引入了动态随机化策略使得网络更容易迁移到现实机器人之中，进行强化学习步行控制器的设计。

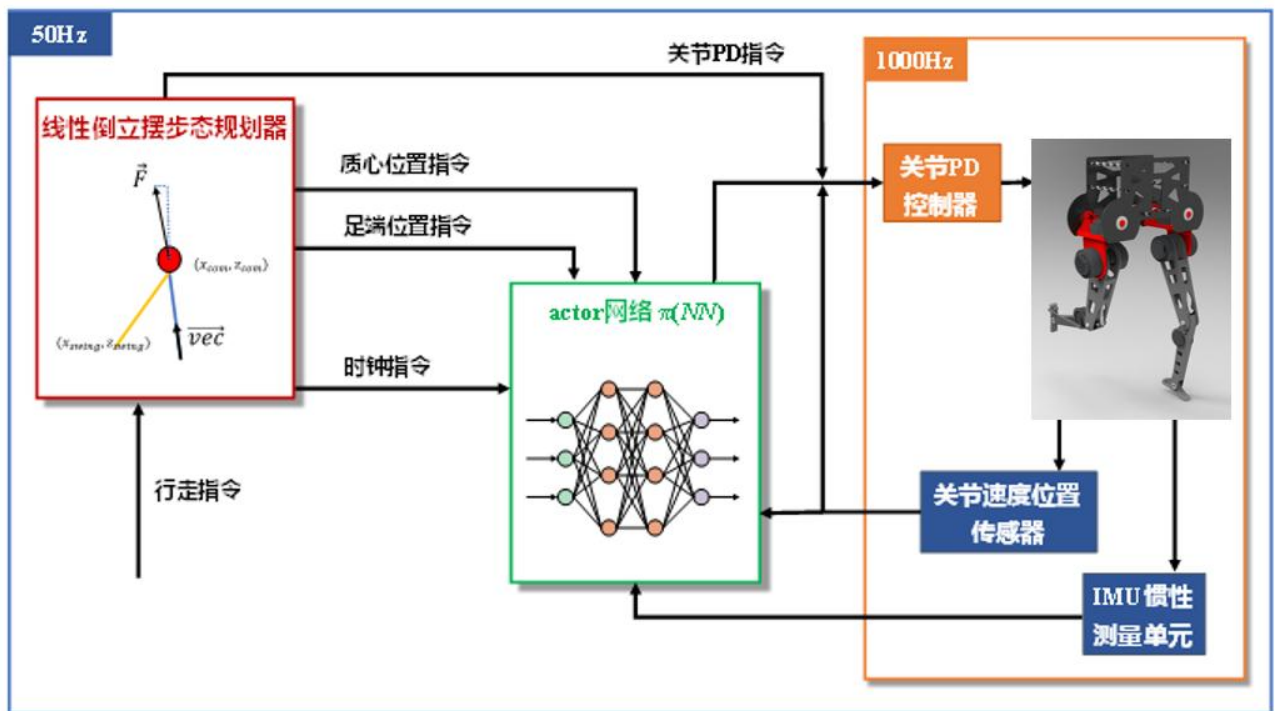


图 4-5 AlexBot 双足机器人基于倒立摆规划器的神经网络框图

为了验证强化学习的可行性，将从仿真和实物两个层面来进行算法验证。首先在仿真实验中分别进行机器人的行走实验，由于发现通过强化学习所得到的网络具有一定的鲁棒性，故额外进行了抗干扰力实验，并进行抗干扰训练提升机器人的抗干扰能力。同时将进行最为困难的 Sim To Real 迁移实验，实现双足机器人的行走和实现更多动作。

## 4.1 机械部分

### 4.1.0 机械总论

本项目目标是研制一款双足机器人，具有轻量化的结构以及优秀的性能和合理分配的惯量以及合适的运动性能。基于轻量化的考虑，本项目首先分析人体腿部结构以及肌肉分布和腿部惯量，使用一体闭环伺服关节电机以及有限元分析 Altair Inspire 分别代表人体的肌肉以及骨骼，使用最小体积质量比目标对机械腿的框架结构进行优化设计，并将拓扑优化结果导入到 SolidWorks 软件中，在考虑加工条件的前提下对零件进行修形，得到了满足强度要求且轻量化的机械腿的结构件。同时，机器人使用轻量化的结构件加工而成，如 3K 碳管，3K 碳纤维板，铝 6061 等材料，既兼顾材料重量也兼顾材料强度。

AlexBot 双足机器人包含单腿 5 自由度，其中包含三个髋关节自由度、一个膝关节自由度、一个踝关节自由度。其中为了整体惯量考虑，我们整体上移膝关节自由度并使用板簧结构传动到膝盖部分，对于踝关节自由度亦如此。通过惯量的计算以及参数测量，我们最终选定三款一体闭环驱动电机，分别为 DM-J8009、DM-J6006、DM-J8006。具有双编码器，输出轴单圈绝对位置，不惧掉电输出轴绝对位置丢失。同时电机和驱动器一体化设计，结构紧凑，集成度高。支持上位机可视化调试，支持固件升级。可通过 CAN 总线反馈电机速度、位置、



转矩、电机温度等信息。位置模式下支持梯形加减速。对于安全方面，具有双温度保护功能，保证电机稳定运行。

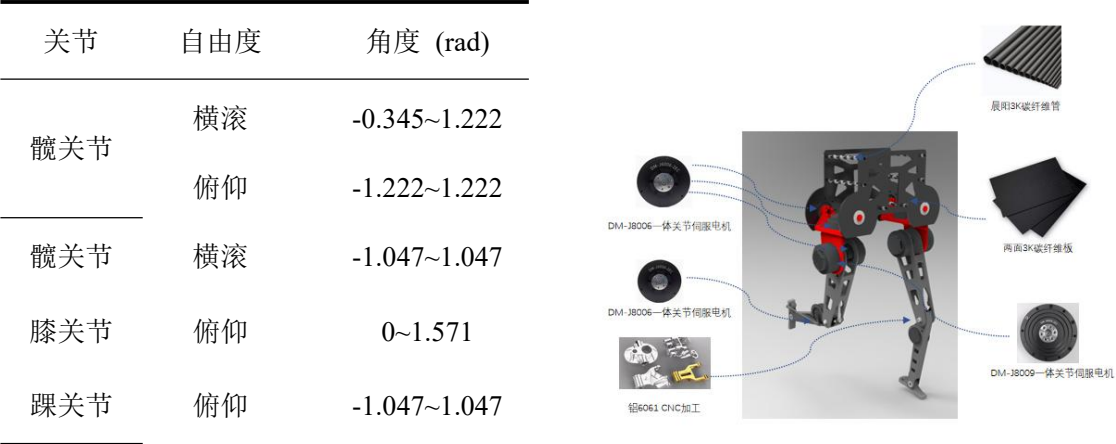


图 4-6 AlexBot 双足机器人机械结构以及关节自由度分配与限制角度

### 4. 1. 1 设计需求分析——腿部结构设计分析

在植物和生物的世界里，生物系统的基本结构往往是向着最优能源消耗的方向进化的，相比较于现有的机械结构而言，生物结构的重量更轻、寿命更长，并且具有一定的柔性。自然界生物进化的一个特点是“用进废退”，在需要承受较大载荷的地方会优先进化，在承受较小载荷的地方，则会发生退化。如长颈鹿的祖先的脖子并没有现在这么长，但是由于气候问题，较矮的灌木不复存在，为了吃到更高的乔木上的枝叶，长颈鹿的脖子渐渐往长的方向进化。人类作为灵长类动物，双足步行已有数百万年的历史，因此在机械腿设计中，可以参照人体运动学和构造来设计机械腿。

人体的腿部结构主要存在 6 个自由度，分别是髋关节的屈伸/伸展（俯仰、pitch）、内旋/外旋（偏航、yaw）、外摆/内收（翻滚、roll）三个自由度，膝关节的屈伸/伸展自由度，踝关节的屈伸/伸展、外摆/内收两个自由度。在本文的设计的柔顺机械腿中，主要研究的自由度为髋关节和膝关节的屈伸/伸展自由度，这是因为这两个自由度在人体运动中承受了最大的载荷，直接为人在行走、跳跃、奔跑等过程中重心的位移提供动力。

人体腿部整体外形曲线呈上粗下细的倒三角，如图 4-7 所示，这是因为人在运动过程中腿部存在支撑相和摆动相这两个状态，倒三角的结构有利于减小腿的转动惯量，提高腿在摆动相过程中的动态性能。在机械腿设计中，应该参照人体腿部结构，设计符合人体动力学的“上重下轻”的机械腿。此外，应该尽可能减少机械腿结构的重量，小腿的质量要小于大腿，机械腿结构重量的降低使得机械腿可以选择更小、更轻的驱动器，减少机械腿质量的同时可以调高其动力学性能。人体膝关节处有复杂的半月板-韧带结构，如图 4-7 所示，其中，半月板负责行走、跳跃、奔跑等过程中对外部冲击的缓冲以及膝关节运动的润滑，韧带负责连接肌肉与骨骼，提供动力。做过半月板切除手术的运动员膝盖往往容易再次受伤，其原因就是在切除了半月板之后，运动员股骨和胫骨连接处缺少缓冲和润滑，在进行跑跳等动作时膝盖会受

到刚性冲击。机器人的关节往往通过轴连接，通过轴承、衬套可以实现机器人膝关节的润滑，因此在机械腿设计中，需要重点考虑的是膝关节的被动柔顺性能。

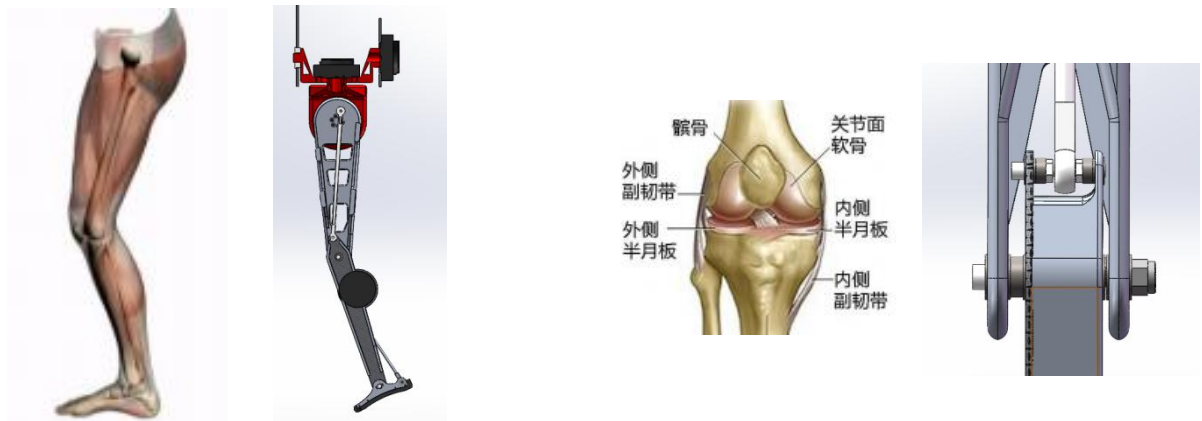


图 4-7(a) 人体下肢结构与机器人对应体 图 4-7(b) 人体膝关节解剖图与机器人对应体

### 4.1.2 设计需求分析——轻量化设计分析

腿部结构件的减重可以减少腿部质量和转动惯量。为此我们使用 Altair Inspire 对腿部主要结构件进行拓扑优化，其优化流程如图 4-8 所示，首先对原始模型进行填充，给出可优化区域。其次，根据机器人实际使用工况对结构件施加载荷和约束，进行拓扑优化。对拓扑后的模型进行后处理，使其具有可制造性。最后，对经后处理的模型再进行校核，如果满足强度和刚度条件，则完成优化，如果不满足，返回填充步骤，再次进行拓扑优化。

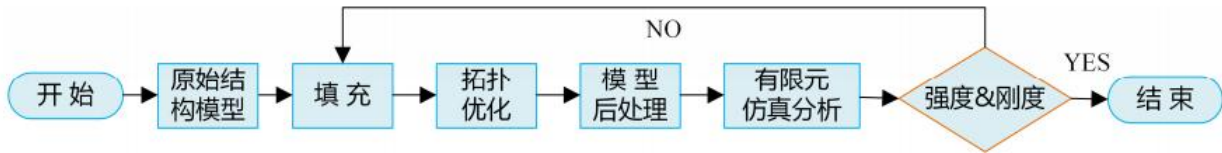


图 4-8 拓扑优化过程

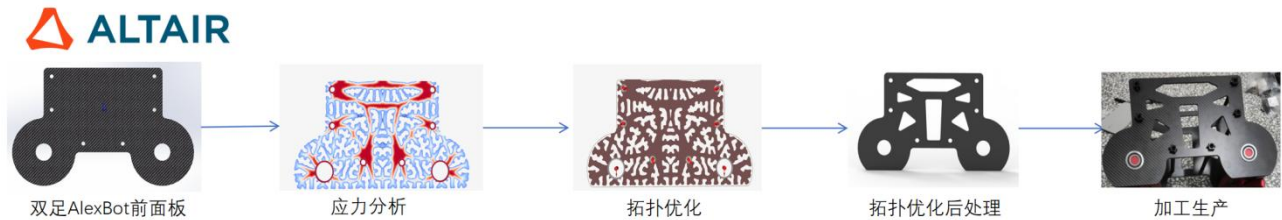


图 4-9 (a) 使用 Altair Inspire 进行二维拓扑优化全流程



图 4-9 (b) 使用 Altair Inspire 进行三维拓扑优化全流程

以双足前面板为例，AlexBot 采用前后双碳板通过碳管固定，通过在 Altair Inspire 添加载荷与固定点实现最小化质量与最大应力比进行 2 维拓扑优化，如图 4-9（a）所示。3 维拓扑优化同理。

### 4.1.3 运动周期分析

在机器人行走过程中，可以将一只脚着地到下一次着地这一段时间划分为一个运动周期。在本文的设计中，不考虑脚掌的设计，仅仅将其看成一个支点，因此可以将行走步态下机械腿的运动周期划分支撑相和摆动相两个状态。在行走过程中，机械腿需要为整个机器人质心的运动提供能量支持，在摆动腿落地瞬间，摆动腿切换为支撑腿，依靠机器人自身行走的惯性和膝关节的伸展 将质心从谷点推动到峰点，随后利用重力和膝关节的弯曲将质心从峰点降回至谷点，完成支撑相周期，期间髋关节的屈伸/伸展自由度需要配合膝关节的角度实时调整上半身的质心位置，防止机器人出现倾翻的状况；机器人双腿在摆动相和支撑相中相互切换，当其中一只腿处于支撑相时另一只腿则处于摆动相，准备下一次与地面接触。具体的步行中状态变化图如图 4-8 所示。从图中可以看出机器人在运动过程中，机械腿在摆动相和支撑相的切换瞬间，该腿与地面发生冲击，机器人垂直方向的能量瞬间被释放，在膝关节处的缓冲就显得尤为重要。虽然本文的线性/弹簧倒立摆步态已经考虑了落地时的冲击，但是在实际工作过程中，由于控制器与作动器的延迟，冲击依然存在，因此机械腿设计时需要考虑其被动柔顺性。此外，根据倒立摆模型的建模特性（机器人被看作质心和无重量的腿），在设计时，机械腿的重量需要尽可能轻，以保证机械腿的动态性能。从运动周期可以看出，机器人机械腿设计需要重点考虑柔顺性和轻量化，这也与前文的分析不谋而合。我们使用斯坦福大学的软件 OpenSim 实现人形结构的奔跑对应的肌肉与骨骼运动分析。

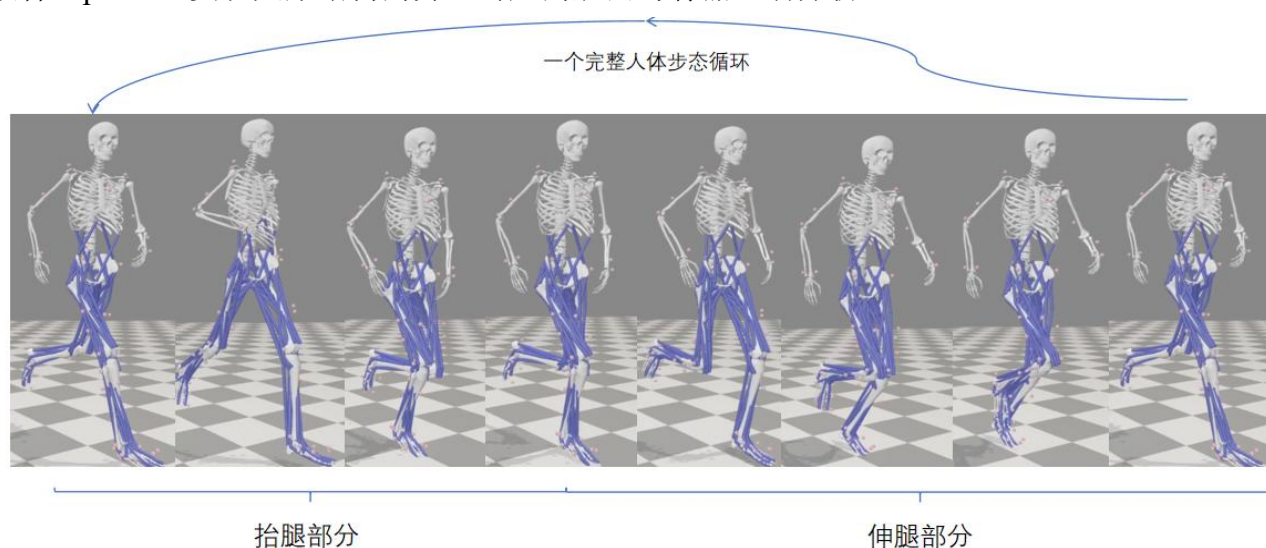


图 4-8 OpenSim 仿真人体一个完整的跑步姿态



#### 4.1.4 双足机器人实物



图 4-10 双足机器人实物建造流程图

## 4.2 硬件部分

#### 4.2.1 系统整体框架

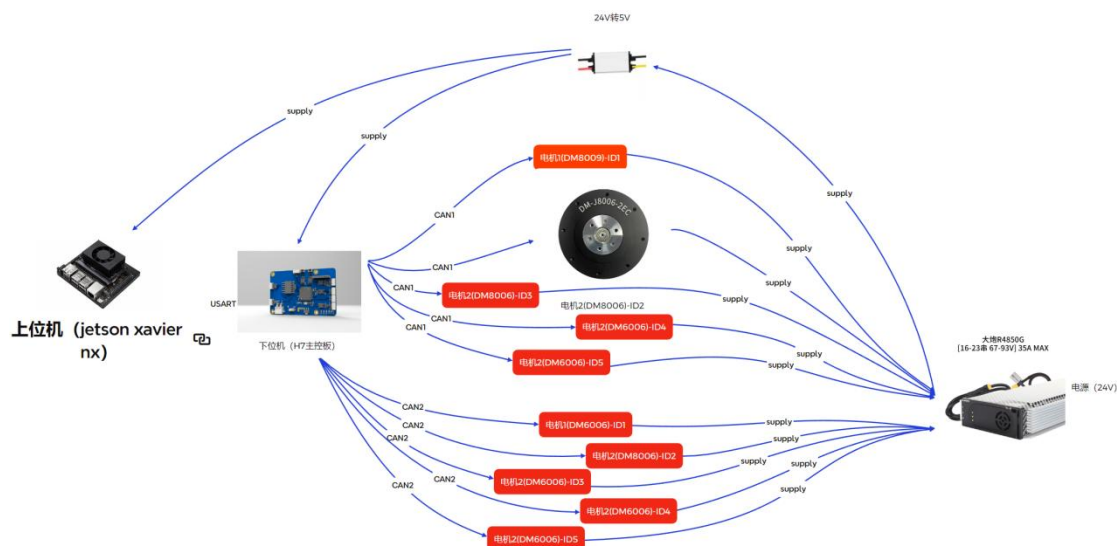


图 4-11 硬件架构整体框图

### 4.2.2 器件选型

电机品牌选用达妙公司 DM6006、DM8006 和 DM8009，具有出力大、质量轻等特点，该品牌的电机具有多种控制模式，如 MIT 控制模式、速度控制模式、位置速度控制模式、，根据不同的需求可以便捷的切换控制模式，功能强大，本项目使用了位置速度控制模式，；该



品牌电机都配备了绝对值编码器，同时响应速度快，满足整个系统对于快速性的要求。

其中三种电机 DM6006 的额定电压为 24V 但支持 48V 输入，额定电流为 4A，额定的扭矩为 4NM。DM8006 电机的额定电压为 24V 可支持 48V 输入，额定电流为 6.5A，额定的扭矩为 8NM。DM8009 电机额定电压为 48V，可支持 24V 输入，额定电流为 20A。考虑到整个功率以及供电的需求，我们这里统一采用 24V 对三种电机供电，根据不同关节所需的扭矩的不同，我们分别选取不同的电机。



图 4-12(a) DM-J6006



图 4-12(b) DM-J8006



图 4-12(c) DM-J8009

IMU 惯性测量单元选用维特智能公司的 WT-901C 型号，其具有精度高，速度快，误差小等优点，具有六轴和九轴的算法，可以根据需求选取模式，加速度计分辨率为 0.0005 g/LSB,RMS 噪声在 0.75 到 1 mg-LSB 之间；陀螺仪分辨率为,RMS 噪声在 0.0028 到之间；俯仰角及横滚角的分辨率为 0.0055°，其倾角精度为 0.2°，可以满足双足机器人使用的反馈精度。

在电源选取上，我们一共所需四个 DM6006 电机，四个 8006 电机以及两个 DM8009 电机，根据他们的额定电压和额定电流，这里我们假设他们的输入电压都为 24V，我们可得到总的功率为 1968W 的功率。所以我们这里选取了输出功率为 2KW 的 DC 直流源作为我们的供电电源。



图 4-13(a) JY901 九轴传感器



图 4-13(b) DC 直流源

### 4.2.3 硬件设计

整体电路所包含的电路板总共如下，包含两个分电板以及一个 STM32H7B0VBT6 为核心的控制板和上位机 jetson Xavier nx。由电源供电 24V 给到 4 个 DM6006，4 个 DM8006 和两个 DM8009。接入 24V 变压模块输出 5V 和 19V 电压分别给主控核心板和 jetson Xavier nx 供电，主控核心板主要处理矩阵变换以及和电机之间的 CAN 通讯问题，一条腿一共有五个电机，

分别是两个 DM6006,DM8006 和一个 DM8009,考虑到 CAN 总线带宽的占用问题,一个 STM32H7B0 核心板的一个 CAN 通道最多接入 5 个电机,也就是一条腿的所有电机。CAN1 接入左腿的五个电机,CAN2 接入右腿的两个电机。同时核心板皆有带有隔离的多路串口和 RS485 接口。可以顺利的完成和 jetson Xavier nx 的通讯。

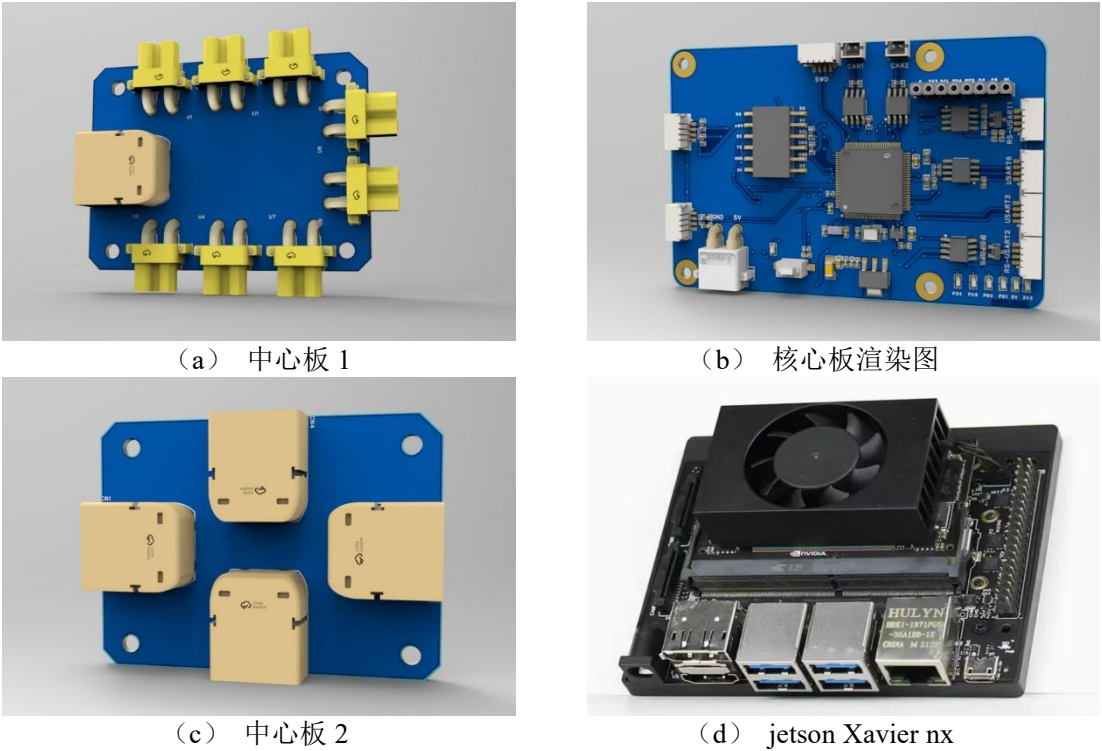
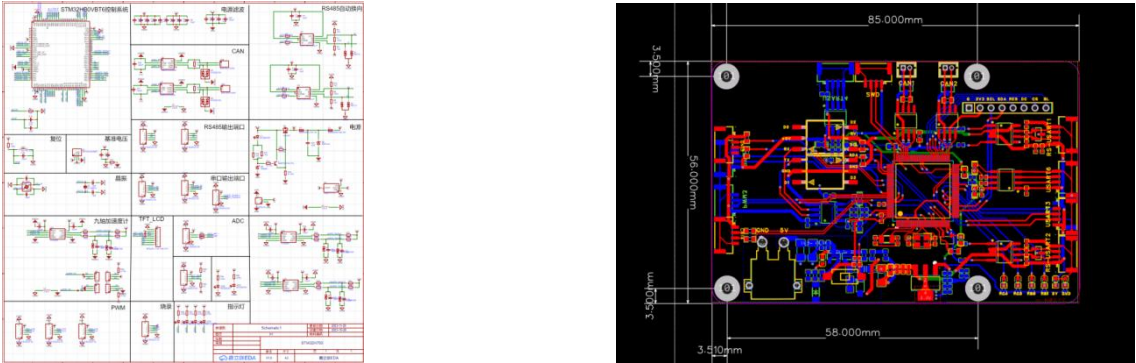


图 4-13 硬件部分板子

选用 STM32H7B0VBT6 作为核心芯片,引出两路 CAN 通道 FDCAN1 和 FDCAN2。电路设计包含输入的 5V 过压保护设计,通过一个 PNP 和一个 P-MOS 来完成电路过压时自动断开电路并直接接入 GND。同时内置 IMU,可以选择多种方式与之通讯,可以通过模拟 IIC 来获取九轴数据或者通过串口实现九轴的数据获取。同时板内内置两路 485 通讯接口,通过 MAX13487RS485 逻辑电平转换芯片,将 TTL 电平转换成 RS485 协议的电平,并通过外部电路的设计实现数据传输的自动换向,同时引出多路串口接口,为了防止高压通过串口直接损坏芯片,在每一路的串口输出中,接入了 II122M31 隔离芯片进行隔离。并引出几路定时器接口可以输出 PWM 波方便后续方案改进或者调试。



### 4.3 驱动部分

一条腿上一共有五个自由度，一个 DM8009 电机，两个 8006 电机和两个 6006 电机。DM8009，DM8006 和 DM6006 电机是 24V 供电。考虑到实际功率的需求，这里选用 XT60 作为我们电源的输入接口，其中 XT30 作为 24V 供电的接口。从模型上侧接入，一个 XT60 接入 24V 电源，并通过同板的另外两个 XT60 接入两个 DM8009,同时最后一个 XT60 接口接入另一个电源中心板，好给 DM8006 和 DM6006 电机进行供电。

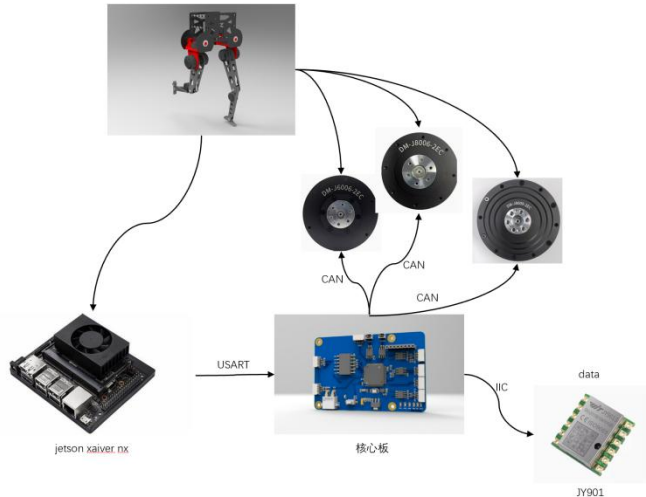


图 4-15 控制框图

代码编写我们选用 CUBEMX 软件进行配置编写，以下为代码的引脚配置情况整个引脚 FDCAN1 接入左腿的五个电机，FDCAN2 接入右腿的五个电机。DM 电机控制采用 CAN 通讯协议，这里我们通过对 DM 电机发送 CAN 指令来实现对电机的控制。以下为电机速度控制的协议：每次发送 8bit 的数据，要传输的速度数据为 float32 位类型，所以我们这里选择发送 4 次来完成一个速度指令的传输。加速度计的通讯方式有很多，有串口和 IIC，这里我们选用软件 IIC 获取九轴数据。以下为我们实际九轴数据获取的代码，如图 4-19 所示。关节控制就是对多个电机进行控制，这里我们先通过 DM 的上位机对各个关节的电机的 ID 进行修改好，我们对总的代码进行各个电机的控制，如图 4-20 所示。具体控制代码和效果如下：

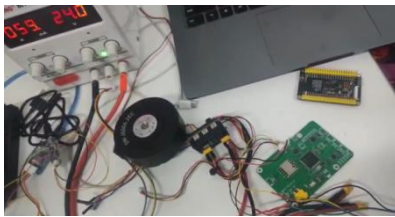


图 4-16 速度环控制

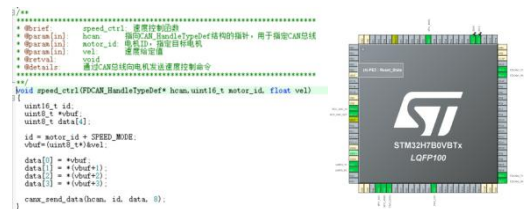


图 4-17 (a) 速度控制 CAN 协议代码与引脚配置

```
while (1)
{
    speed_ctrl(speed_val);
    // 加速度计数据
    // 获取加速度计数据
    for (i = 0; i < 3; i++)
    {
        accel[i] = readAccel() / 16384.0f * 16.0f;
        gyro[i] = readGyro() / 32768.0f * 180.0f;
    }
    // 更新加速度计数据
    printf("accel: %f %f %f\n", accel[0], accel[1], accel[2]);
    // 更新陀螺仪数据
    printf("gyro: %f %f %f\n", gyro[0], gyro[1], gyro[2]);
    // 更新角度数据
    printf("angle: %f %f %f\n", angle[0], angle[1], angle[2]);
    // 更新位置数据
    printf("pos: %f %f %f\n", pos[0], pos[1], pos[2]);
    // 更新速度数据
    printf("vel: %f %f %f\n", vel[0], vel[1], vel[2]);
}
```

图 4-17(b) 九轴加速度计代码

```
void speed_ctrl(FDCAN_HandleTypeDef* hcan, uint16_t motor_id, float vel)
{
    uint16_t id;
    uint8_t data[4];
    id = motor_id + SPEED_M08;
    data[0] = (uint8_t)vel;
    data[1] = (uint8_t)vel;
    data[2] = (uint8_t)vel;
    data[3] = (uint8_t)vel;
    can_send_data(hcan, id, data, 4);
}
```

图 4-17(c) 关节控制代码

## 4.3 算法部分

### 4.3.1 PPO 算法背景

让机器人像人类一样学习和做出人类的动作，一直是人工智能不断发展和完善的不竭动力和目标，深度学习和强化学习的不断进步和发展极大地促进这一目标的实现。深度学习的发展把从原始的数据中提取高水平特征变成现实。深度学习在感知能力表现强势，但是在决策能力上不尽人意；与此相反，强化学习在决策能力上表现上佳，但是对感知问题无特别优势。因此，将两者结合起来优势互补，为解决复杂系统的感知决策问题提供了方法。将具有感知能力的深度学习(deeplearning, 简称 DL)和具有决策能力的强化学习 (reinforcement learning, 简称 RL) 优势互补的结合在一起，构成深度强化学习(deep reinforcement learning, 简称 DRL) 算法。人类的进化是不停地通过学习的方式获得经验知识，这也是人类不同于其他生物的重要特点。因此在人工智能领域该如何教导机器人或者智能体学习人类的动作或者拥有人类的学习方式是一个十分重要的研究课题。机器学习是一门复杂而涉及众多学科的门类，强化学习(Reinforce Learning, 简称 RL)理论与算法和学习心理学与动态规划等至关重要。

策略梯度是一种常用的策略优化方法，它通过不断计算策略期望总奖赏关于策略参数的梯度来更新策略参数，最终收敛于最优策略。因此在解决深度强化学习问题时，可以采用参数为  $\theta$  的深度神经网络来进行参数化表示策略，并利用策略梯度方法来优化策略。值得注意的是，在求解深度强化学习问题时，往往第一选择是采取基于策略梯度的算法。原因是它能够直接优化策略的期望总奖赏，并以端对端的方式直接在策略空间中搜索最优策略，省去了繁琐的中间环节。因此与 DQN 及其改进模型相比，基于策略梯度的 DRL 方法适用范围更广，策略优化的效果也更好。

PPO(Proximal Policy Optimization Algorithms)是目前应用最广的强化学习算法，在对传统策略梯度算法引入使用 KL 散度决定参数，深度神经网络和 clip object 等方法之后，解决了在线强化学习(on-policy)容易更新过大，容易过冲的问题，且收敛快，效果好，被作为强化学习的范式使用。目前许多使用深度强化学习的机器人控制方案都是基于 PPO 算法的改进。

### 4.3.2 模仿学习

模仿学习 (Imitation Learning) ——从专家示例中学习——是一种让智能体（机器人）像人类专家一样能够进行智能决策的方法。在通往通用人工智能的路上，人们发现很难手工地进行编程来教会智能体进行思考，因为这么做涉及到大量的人工工程。比如，在教会车辆自动驾驶的过程中，需要有大量的约束进行考虑（安全驾驶而不发生事故、平稳驾驶而增加舒适度）等等，而针对这些约束设计特定的监督信息信号来引导智能体是一个比较困难的任务。相反之下，人类却能比较容易地完成这些任务，并且为智能体提供大量的示例行为。利用这些专家示例来教会智能体进行智能决策就是模仿学习主要解决的问题。通用地来讲，任何希望智能体能够像“专家”一样进行决策的系统都能从模仿学习的方法里受益。事实上，在 20



世纪 80 年代末，卡内基梅隆大学（Carnegie Mellon University）的 Pomerleau 教授已经利用人类专家示例成功地训练一个浅层神经网络，使得其控制的无人车能够穿越北美洲。前几年大放异彩的 AlphaGo，第一个打败人类顶尖选手的机器人，也用到了模仿学习方法对人类棋谱进行学习。除此之外，在现代的应用中，比如推荐系统，互联网网约车派单等系统中，也用到了模仿学习的方法来构建一个虚拟世界，从而允许智能体在其中自由地试错和学习。不断涌现的新的任务促使研究者们设计了各种各样的模仿学习算法。其中，普遍认为模仿学习有两大类算法：行为克隆（Behavioral Cloning）和对抗式模仿学习（Adversarial Imitation Learning）。其中，行为克隆算法尝试最小化智能体策略和专家策略的动作差异，把模仿学习任务归约到常见的回归或者分类任务。而对抗式模仿学习算法则是通过逆强化学习（Inverse Reinforcement Learning）来构建一个对抗的奖赏函数，然后最大化这个奖赏函数去模仿专家行为。基于这两种基本的算法设计思路，研究者们提出了各种各样的模仿学习算法变体。

机器人是模仿学习主要的研究领域，要解决如何让机器人像人一样行走，方式之一就是使用人类行走的数据集对智能体进行训练。我们适配的 Loco-mujoco 项目提供了大量人类行走的数据集，且规范了统一接口。

### 4.3.3 实施方案

本项目主要使用 PPO Agent 作为主要的强化学习方法，在基础的 PPO 算法上增加经验回放，使用 pytorch 库中的 torchrl 搭建而成。在 isaac gym 平台上，通过使用 api 创造不同材质和起伏的地面，使双足机器人具有适应各种地面的能力。适用于 Mujoco 的强化学习框架已经搭建并测试完成，并成功应用于 Loco-mujoco 的 benchmark 环境。

经验回放（Experience Replay）是在线强化学习中的一个技术，旨在改善学习的效率和稳定性。在实时与环境交互中获得的经验（状态、动作、奖励等）通常会被立即用于更新模型，但是实时更新的模型常常由于较少的数据而变得不稳定，经验回放通过存储这些经验到一个称为“经验回放缓冲区”的数据结构中，然后在训练过程中随机抽样以用于模型更新，从而解决了这一问题。

仿真环境的搭建是双足机器人强化学习中的重要一步，本项目先后使用 isaac gym 和 mujoco 作为训练的仿真环境，isaac gym 是英伟达开发的支持 Gpu 运行的并行化仿真环境，以解决强化学习普遍存在的仿真数据获取较慢的问题。Mujoco 是一个强大的物理引擎，对软物体，肌腱，生物肌肉等结构也有仿真能力，使用 Mujoco 平台极大提高了本项目的拓展性，可以在强化学习的基础上应用模仿学习。

## 5、特色与创新

### 5.1 极低的成本实现较强的运动性能

本双足机器人采用较易获取的材料和简易的 CNC 加工方式，争取将双足成本做到两万以内，使用达妙 DM-J8006\*6;DM-J8009\*2;DM-J6006\*2，额定扭矩分别为 8NM;20NM;4NM，通过轻量化设计与降低腿部惯量实现。



图 5-1 AlexBot 双足机器人机械结构以及部分订单价格

### 5.2 开源属性增加社群影响力

本双足机器人计划全部开源，包含机械结构设计，硬件设计，控制算法。目前开源机械结构设计，硬件设计，目前产生一定影响力，有部分社会人士愿意资助此项目落地，同时社群目前已有机器人相关硕博十余人密切关注，交流指导。

### 5.3 腿部关节以及计划足式仿生学设计

我们拟打算在结构设计上对于足踝关节进行仿生学设计。足踝关节是连接人类足部与腿部的生物机械结构，在人类行走中扮演着至关重要的角色。其主要功能包括负重、提供弹性以减轻地面冲击，并具备柔性以适应不平坦的地面。足踝关节由多个解剖结构组成，包括胫骨的下端关节面、内踝关节面、腓骨的外踝关节面，以及距骨滑车的上关节面和内外关节面。内踝和外踝是足踝关节的两个重要部分。在冠状面内，外踝相对于内踝较低，大约低 1 厘米左右；在矢状面内，外踝相对于内踝稍微向后倾斜 1 厘米；后踝相对于前踝向下延伸，从而限制了距骨向后移动。在正常的人类步态周期中，足踝关节发挥着重要作用。在触地前期，当胫骨绕足踝关节旋转时，跟腱被拉长，并储存能量。在触地后期，足踝关节向下蹬地，释放触地前期储存的能量。在摆动前期，足踝的跖屈肌协助加速胫骨向前运动，为摆动腿提供主要动力。足踝关节的正常运动范围通常为背伸  $26^{\circ}$  至  $27^{\circ}$  和跖屈  $40^{\circ}$  至  $45^{\circ}$ 。

目前，双足机器人一般采用不同类型的足部结构，包括平足、弧形足和球状足，以实

现步行和稳定性。举例如 ASIMO 系列双足机器人、McGeer 被动步行机、康奈尔完全被动机器人以及 RABBIT 平面双足机器人等。然而，与人类的高效、快速和稳定行走相比，现有的双足机器人在运动性能方面仍有改进的空间。为了提高机器人在不平坦地面上的行走稳定性并实现更自然的步态，研究人员致力于通过模仿人类足踝的功能来设计新型的仿生足踝结构。通常情况下，双足机器人的踝关节是刚性的结构，而人类的足踝具有刚柔耦合的结构。在人类行走时，内侧纵向足弓的弯曲和扭曲所产生的桁架机制和绞盘机制使足趾能够吸收地面冲击并执行蹬地行为。这种刚柔耦合的结构使得人类足踝在行走时能够适应不同地形和吸收地面的不平坦性，从而提供了更好的行走稳定性。因此，将类似的机制引入双足机器人的足踝结构，可以帮助提高机器人的运动性能，使其更有效地应对复杂的环境和地形。

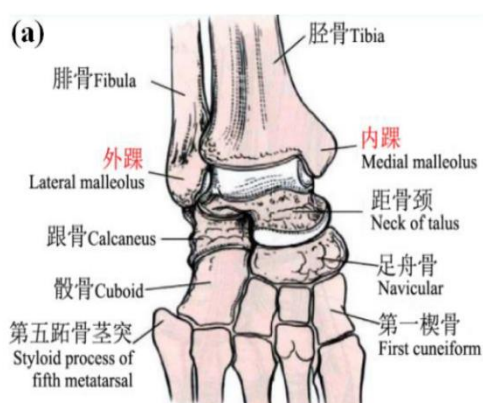


图 5-2 足踝骨骼组成 (a) 踝关节冠状面

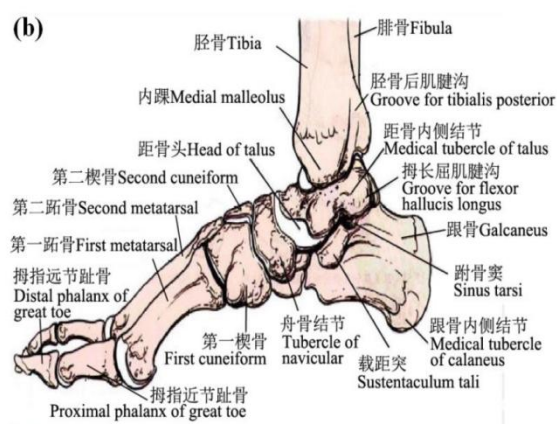


图 5-2 足踝骨骼组成 (b) 踝关节矢状面