

# 基于Vive VR的遥操作机械手

(申请清华大学工学硕士学位论文)

培养单位：天空工场

学 科：计算机科学与技术

研 究 生：薛 瑞 尼

指 导 教 师：郑 纬 民 教 授

副指导教师：陈 文 光 教 授

联合指导教师：某 某 某 教 授

二〇一八年三月



# **An Introduction to L<sup>A</sup>T<sub>E</sub>X Thesis Template of Tsinghua University v5.4.3**

Thesis Submitted to  
**Tsinghua University**  
in partial fulfillment of the requirement  
for the professional degree of  
**Doctor of Engineering**

by  
**Xue Ruini**  
**( Computer Science and Technology )**

Thesis Supervisor : Professor Zheng Weimin  
Associate Supervisor : Chen Wenguang

**March, 2018**



# 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容。

本人保证遵守上述规定。

（保密的论文在解密后应遵守此规定）

作者签名：\_\_\_\_\_

导师签名：\_\_\_\_\_

日 期：\_\_\_\_\_

日 期：\_\_\_\_\_



## 摘要

随着机器人技术的发展，机器人的应用场景越来越多。但是目前自动控制机器人尚不能执行多数复杂任务，特别是抢险救灾等需要随机应变的任务，这种情况下需要遥控操作机器人。

目前机械臂远程控制普遍采用手柄或键盘控制方式，且监控方式普遍为摄像头图像显示在监视器上，与现场操作差别很大。我们开发了一套用VR设备远程控制机器人的系统以及配套的三维实时场景采集及图传系统，使操作者能有身临其境的操作体验，同时大幅降低成本，减小了延迟。

本文的创新点主要有：

- 为使操作者看到具有立体感的实时画面，我们开发了由双目摄像头及VR显示系统组成的三维实时场景采集系统，此系统较图像拼接和场景重构实时性好，且对算力要求不高。
- 手持Vive追踪器操作机械臂末端符合我们日常使用手进行操作的习惯。较外骨骼和Optitrack运动捕捉系统廉价，且能够满足绝大部分需求。

**关键词：**机器人；虚拟现实；远程控制

## **Abstract**

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summarization of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

An abstract should be concise and to the point. It is a misunderstanding to make an abstract an outline of the dissertation and words “the first chapter”, “the second chapter” and the like should be avoided in the abstract.

Key words are terms used in a dissertation for indexing, reflecting core information of the dissertation. An abstract may contain a maximum of 5 key words, with semi-colons used in between to separate one another.

**Key words:** T<sub>E</sub>X; L<sup>A</sup>T<sub>E</sub>X; CJK; template; thesis

## 目 录

<b>第1章 项目背景 .....</b>	<b>1</b>
<b>第2章 现有系统问题 .....</b>	<b>2</b>
2.1 现有的三维图像采集系统 .....	2
2.2 现有的机器人遥操作系统 .....	2
2.2.1 计算机指令或手柄控制 .....	2
2.2.2 OptiTrack动作捕捉系统 .....	3
2.2.3 T-HR3人形机器人遥操作系统 .....	4
<b>第3章 系统简介 .....</b>	<b>5</b>
3.1 VR远程监视系统 .....	5
3.2 Vive追踪器遥操作系统.....	5
<b>第4章 项目展示 .....</b>	<b>7</b>
4.1 实时图像采集与显示系统 .....	7
4.2 机械臂遥控与执行机构（Vive+DOBOT机械臂） .....	7
4.3 机械臂遥控与执行机构（NAO人形机器人） .....	9
4.4 总结 .....	9
<b>第5章 技术细节 .....</b>	<b>10</b>
5.1 实时图像采集与显示系统 .....	10
5.2 机械臂控制系统.....	11
5.3 NAO机器人控制系统.....	13
<b>第6章 展望 .....</b>	<b>16</b>
6.1 项目改进 .....	16
6.2 未来应用 .....	16

## 主要符号对照表

VIVE	HTC公司开发的一套VR设备，包括头戴式显示器和手持追踪器
ZED	Stereolabs出品的双目摄像头，与人眼各参数类似
CUDA	由NVIDIA公司推出的一种通用并行计算架构，使GPU能够解决复杂的计算问题
DOBOT	越疆魔术师，由越疆科技研发的高精度机械臂，具有3个自由度
NAO	NAO高58厘米，是SoftBank Robotics推出的第一款类人机器人
TX1	NVIDIA Jetson TX1，AI超级计算机，采用 NVIDIA Maxwell架构，具有256颗NVIDIA CUDA核心和64位CPU
OptiTrack	NaturalPoint Inc出品的工业级动作捕捉系统，使用数十个高速红外相机确定物体上反光球的位置，较昂贵
Google Jump	Google的图像拼接系统，装备有16个独立的相机。可以创建360度全景，并用于虚拟现实设备。
全向轮	即Omni轮，在圆周上与转动方向垂直的具有小圆盘的轮子，其效果是，车轮可以纵向驱动，但也可以轻松横向滑动。
GoPro	运动摄像机，因其体积小，视野宽，分辨率高等特点，受到全景摄像解决方案青睐
T-HR3	丰田于2017年11月最新出的人形机器人，通过外骨骼和数据手套远程控制
MMS	远程控制(Master Maneuvering System)
Google Cardboard	Google所开发的折叠式纸板头盔，与智能手机配合使用的虚拟现实头戴式显示器。
CUDA	Compute Unified Device Architecture，统一计算架构，由NVIDIA所推出的一种集成技术

## 第1章 项目背景

随着机器人技术的发展，机器人的应用场景越来越多。但是目前自动控制机器人尚不能执行多数复杂任务，特别是抢险救灾等需要随机应变的任务，这种情况下需要遥操作机器人。目前机器人远程控制普遍采用手柄或键盘控制方式，且监控方式普遍为摄像头图像显示在监视器上，与现场操作差别很大。



图 1.1 控制面板

现有的基于动作捕捉的遥操作系统如OptiTrack和机械外骨骼等，虽然用户体验很好，但是成本极高，依赖性强，无法大规模应用；现有的实时VR显示大多基于图像拼接，图像采集设备昂贵，实时性和图像质量无法同时兼顾，且对算力要求很高。

本作品提出了一种廉价、低延时的遥操作系统，其双目摄像头及VR显示系统可使操作者看到具有立体感的实时画面，手持追踪器操作机械臂末端符合人类日常使用手进行操作的习惯，可以给予操作者以身临其境的操作体验。

## 第2章 现有系统问题

### 2.1 现有的三维图像采集系统

目前三维场景重构需要很高的算力，分辨率低且不能保证实时性，目前较成熟的实时三维场景采集系统基本上都是基于图像拼接技术：

就图像传输系统而言，要创建一个360度的视频，无论是使用一组特殊的摄像机同时记录360度的场景，还是将4个不同角度的GoPros图像拼接在一起。传入的360度视频文件是4K和更高，比特率可以超过50 Mb/s。而用于VR的3D 360度视频是其两倍，即每小时44 GB视频。如果要实时发送和接收这些图像信息，对带宽等网络环境要求非常高。另外无论采用立方体还是金字塔形拼接方式，拼接这些图像会消耗大量的算力，运算和传输延迟加起来都会达到2s以上，实时性和图像质量无法同时兼顾。

就图像采集设备而言，以Google Jump为例，Google Jump使用24个GoPro，环绕一个圆形分布，考虑到GoPro单个相机的价格（2000-3000人民币）上述系统造价非常昂贵。

### 2.2 现有的机器人遥操作系统

#### 2.2.1 计算机指令或手柄控制

许多需要人工远程操控的机器人，其操作方式基本都是向控制面板或计算机输入指令而实现。工作人员需要熟悉各种操作平台，并需要长时间紧盯电脑屏幕进行操作，其交互体验不是很理想。

另一些用户界面比较友好的机器人使用摇杆控制

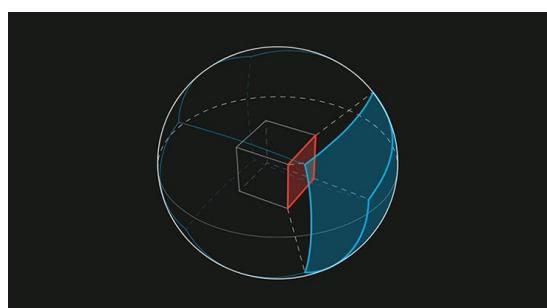


图 2.1 立方体拼接

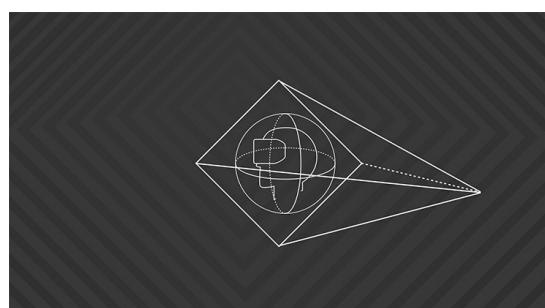


图 2.2 金字塔形拼接



图 2.3 Google Jump

### 2.2.2 OptiTrack动作捕捉系统

OptiTrack<sup>①</sup> 是基于一系列红外摄像机的动作捕捉系统解决方案，在操作者身上贴多个感光球，通过四周的十多个红外摄像机来计算每一个感光球的位置，从而获得人的动作数据。

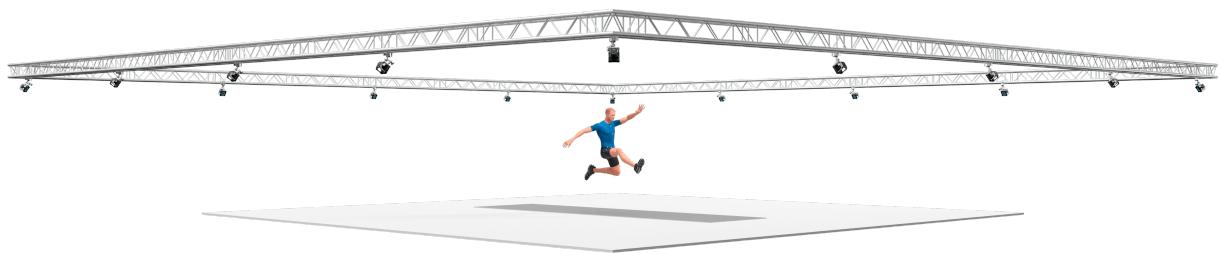


图 2.4 OptiTrack

由于此解决方案需要至少12个高速红外摄像机和配套的处理系统，而一个红外摄像机售价4000-5000美元，整个系统的成本非常高，用其做机器人控制并不现实。

---

① <http://optitrack.com/>

### 2.2.3 T-HR3人形机器人遥操作系统

T-HR3是一种远程机动系统，它通过外骨骼和数据手套将手，手臂和脚部动作映射到机器人。

就操作者近端方面而言，T-HR3采用机械外骨骼加上双手的数据手套。由于机械外骨骼以力信号作为控制端的输出信号，所以，机械外骨骼和操作者之间本身就存在一定的滞后性，而且，力信号易受外力干扰，所以存在信号不稳定的问题。

另外，T-HR3需要一套昂贵且笨重的外骨骼，其应用场景十分受限。



图 2.5 T-HR3

## 第3章 系统简介

该项目可分为两个子系统：图像显示系统及机器人控制系统。

### 3.1 VR远程监视系统

本项目旨在提供一个具有较强立体感的实时全景图像，并解决现有系统的问题以及最大化的降低造价，以使远程操控者获得身临其境的体验。我们用ZED双目摄像头来采集实时场景信息，两个目采集到的图像对应到VR眼镜的两个显示屏中，而搭载ZED的多自由度舵机支架可随操纵者头部的转动而实时旋转，以保证双目摄像头的朝向和操作者双眼朝向一致，从而使操作者可以实时地看到各方位的立体场景。

此系统硬件部分仅由一个双目摄像头，转台，虚拟现实头盔组成，成本低，且部件较常见。双目摄像头可以由两个独立的摄像头替代，只需保证两摄像头间距和瞳距相似即可，随着虚拟现实技术的发展，分辨率较高的Vive VR头盔(见图3.1)价格可以为普通人所接受，Google Cardboard(见图3.2)售价低至10RMB，和智能手机配合使用。

因为只需要传输两个摄像头的图像，且并不需要复杂的运算，这套机械式的实时全景系统延迟很小。

### 3.2 Vive追踪器遥操作系统

虚拟现实头盔和追踪器上的红外定位模块精确的确定了它们的绝对位置，由此可得它们的相对位置，我们可以使机械臂末端和双目摄像头的相对位置和其一致，从而实现操作者直接用手的位置来控制机械臂。



图 3.1 Vive VR头盔



图 3.2 Google Cardboard

此系统使得操作者能够利用自身的动作控制机器人的运动，符合人类的行为习惯，因而会带来良好的操作体验。

Vive虚拟现实套件<sup>①</sup>(见图 3.3)包括Vive追踪器，Vive VR头盔和Lighting基站，官网售价5488RMB，远低于OptiTrack或T-HR3外骨骼部分的售价。而这套系统完全可以胜任大多数任务。而且Vive VR套件对环境要求很低，可以随时搭建操作环境。



图 3.3 Vive虚拟现实套件

---

<sup>①</sup> <https://www.vive.com/cn/product/>

## 第4章 项目展示

下面以时间顺序分阶段介绍我们的成果：

### 4.1 实时图像采集与显示系统

实现了多自由度机械结构托举的双目摄像头，随着操作者头部的转动而同步转动，以采集实时图像信息。在这一阶段，我们的VR平台暂时使用谷歌折叠式纸板虚拟现实显示器。操作者头部姿态数据由VR头盔中的Android手机自身的陀螺仪获取，并发送到舵机的控制端。

同时，在这一阶段，我们在机械臂末端加上了吸盘执行机构，使用气泵和电磁阀组建了一个吸附装置(见图 4.2)。

同时我们改写了DOBOT机械臂底层及接口，解决了串口控制卡顿的问题。实现用遥控器-接收机控制系统，远程完成叠纸杯实验。

### 4.2 机械臂遥控与执行机构（Vive+DOBOT机械臂）

在这一阶段，我们完成了一个由全向轮底盘移动的机械臂(见图 4.1)，平台上搭载多自由度机械结构托举的双目摄像头，随着操作者头部的转动而同步转动，以采集实时图像信息。

我们可以实现读取追踪器和VR头盔绝对位置并通过串口发送。可以看到图(4.3)中输出了左手，右手和头盔的绝对坐标。我们改写了机械臂底层及接口，解决了串口控制卡顿的情况。

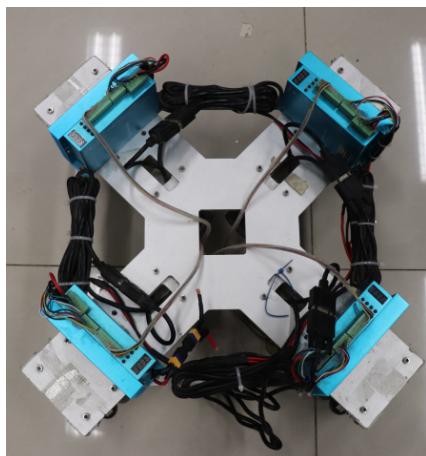


图 4.1 全向轮底盘

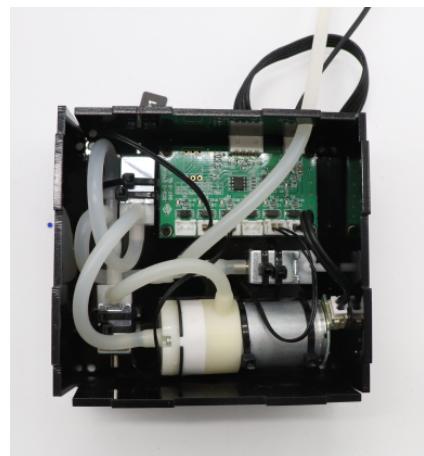


图 4.2 气泵和电磁阀组建的吸附装置

图 4.3 追踪器和VR头盔绝对位置

从而实现了用Vive追踪器远程控制机械臂，完成了叠纸杯实验。

### 4.3 机械臂遥控与执行机构（NAO人形机器人）

另外我们将此系统成功拓展到NAO人形机器人上，用追踪器控制机器人手臂的运动。<sup>①</sup>

成功完成Vive远程控制NAO的实验。

### 4.4 总结

经测试，我们这套系统远程控制实时性很好，视频传输延迟小于300ms，在VR眼镜中有很强的立体感，控制很方便，追踪器定位效果很好，而且具有很强的可迁移性。

---

<sup>①</sup> 由于NAO自身的缺陷（易失去平衡），我们现阶段并没有加入腿部动作，且将手臂动作对称处理

## 第5章 技术细节

此系统实现了远程操作的完整功能，整个系统数据流向如图 ??，硬件层面架构如图

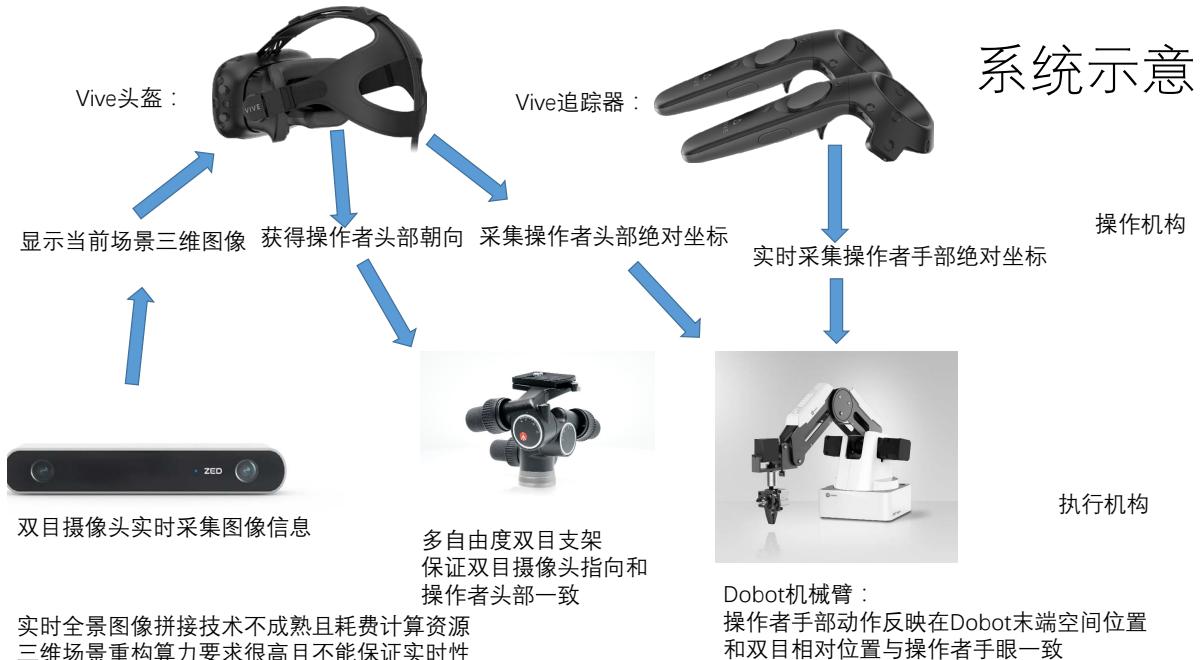


图 5.1 数据流向示意图

下面分三部分介绍技术细节。

### 5.1 实时图像采集与显示系统

ZED双目摄像头两个目采集到的图像分别投射到Vive头盔的两个显示屏中。ZED双目摄像头搭载在由两个数字舵机组成的双自由度平台上，以保证双目朝向与人眼朝向实时一致。

整个系统放置在4轮全向轮底盘上面，ZED双目摄像头采集到的图像，通过USB3.0发送给Nvidia Jetson TX1<sup>①</sup>，TX1上运行Ubuntu系统，图像在TX1上处理后，使用UV4L服务（User space Video4Linux）<sup>②</sup>发送到局域网。同时，Vive VR头盔的姿态数据通过HTC串流盒发送到计算机中进行处理，处理后通过蓝牙串口发送到全向轮车上面的下位机（Arduino Mega 2560），控制双自由度数字舵机做出

① <http://www.nvidia.cn/object/embedded-systems-dev-kits-modules-cn.html>

② <https://www.linux-projects.org/uv4l/>

与Vive VR头盔同步的转动，使安装在双自由度数字舵机平台上的ZED双目摄像头的指向与Vive头盔实时相同。

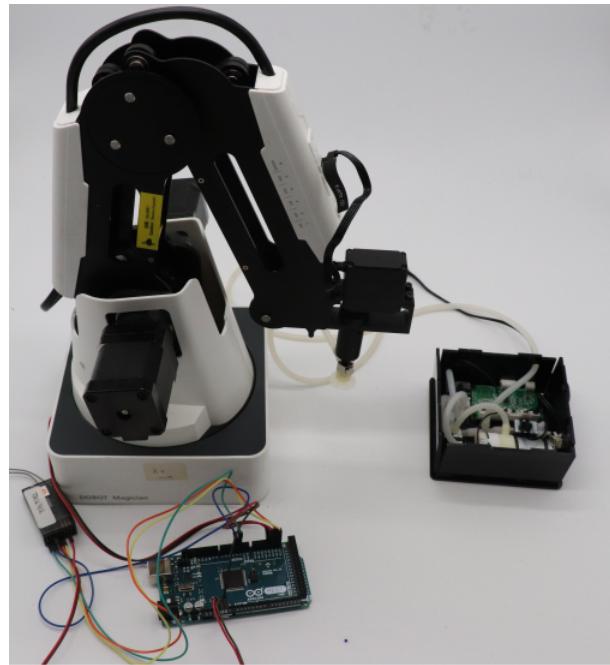


图 5.2 DOBOT机械臂及通讯系统

演示视频参见附件XXX

## 5.2 机械臂控制系统

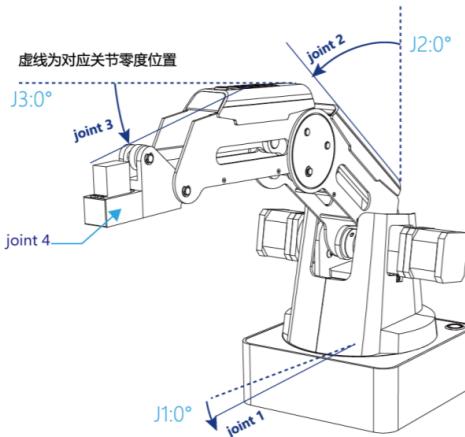


图 5.3 DOBOT机械臂各自由度的示意图

图 5.2为Dobot各自由度的示意图， 机械臂在三维空间中的移动需要用到joint1、2、3， 该三个自由度简化后的模型如下图所示。

其中a,b分别为两个臂的长度，根据官方资料， $a = 135mm, b = 147mm$ 。 $\alpha$ 、 $\beta$ 、 $\theta$ 为三个自由度对应的方位角。

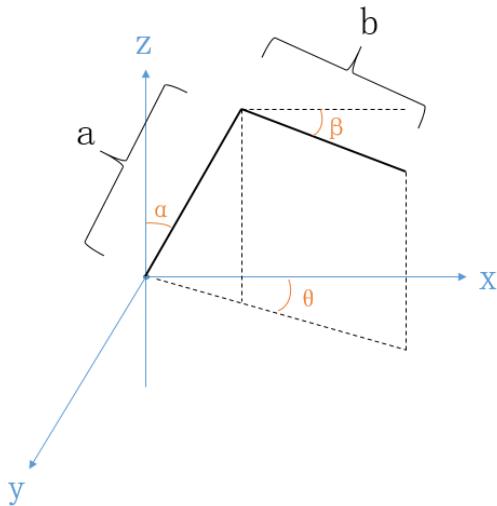


图 5.4 DOBOT机械臂末端坐标表示示意图

由图 5.3可得，末端坐标的位置( $x, y, z$ )可表示为：

$$x = (a \times \sin(\alpha) + b \times \cos(\beta)) \times \cos(\theta) \quad (5-1)$$

$$y = (a \times \sin(\alpha) + b \times \cos(\beta)) \times \sin(\theta) \quad (5-2)$$

$$z = a * \cos(\alpha) - b * \cos(\beta) \quad (5-3)$$

可见，任意一组给定的x,y,z坐标都有唯一的一组 $\alpha$ 、 $\beta$ 、 $\theta$ 与之对应，只要得到方位角与坐标之间的关系即可将机械臂移动到其可触及的任意位置。

1. 点对点移动：联立以上方程组，解得：

$$\alpha = \arcsin\left(\frac{x^2 + y^2 + z^2 - b^2}{2a\sqrt{x^2 + y^2 + z^2}}\right) - \arcsin\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)$$

$$\beta = -\arcsin\left(\frac{x^2 + y^2 + z^2 - a^2}{2b\sqrt{x^2 + y^2 + z^2}}\right) + \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)$$

由求出的 $\alpha$ 、 $\beta$ 可得到

$$\theta = \arcsin\left(\frac{y}{a \times \sin \alpha + b \times \cos \beta}\right) \quad (5-4)$$

由此可将任意一点的 $(x, y, z)$ 坐标用机械臂三个自由度的方位角( $\alpha$  fi  $\beta$  fi  $\theta$ )来表示，因而可通过各关节的转动而实现点对点运动。即初始坐标为 $(x_1, y_1, z_1)$ ，目的坐标为 $(x_2, y_2, z_2)$ ，则可用上述方法计算出相应的 $(\alpha_1$  fi  $\beta_1$  fi  $\theta_1)$ ,  $(\alpha_2$  fi  $\beta_2$  fi  $\theta_2)$ ，从而得知各关节需转动的角度。

2. 基于运动方向的移动机械臂在 $(x, y, z)$ 坐标系中各轴的速度分量也可由 $(\alpha$  fi  $\beta$  fi  $\theta)$ 坐标系中的角速度表示：

$$\begin{aligned} v_x &= \frac{dx}{dt} = \frac{\delta x}{\delta \alpha} \frac{\delta \alpha}{\delta t} + \frac{\delta x}{\delta \beta} \frac{\delta \beta}{\delta t} + \frac{\delta x}{\delta \theta} \frac{\delta \theta}{\delta t} \\ &= a \cos \alpha \cos \theta w_\alpha - b \sin \beta \cos \theta - (a \sin \alpha + b \cos \beta) \sin \theta w_\theta \end{aligned}$$

$$v_y = a \cos \alpha \sin \theta w_\alpha - b \sin \beta \sin \theta + (a \sin \alpha + b \cos \beta) \cos \theta w_\theta$$

$$v_z = -a \sin \alpha w_\alpha - b \cos \beta w_\beta$$

用矩阵形式表示，即：

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} a \cos \alpha \cos \theta & -b \sin \beta \cos \theta & -(a \sin \alpha + b \cos \beta) \sin \theta \\ a \cos \alpha \sin \theta & -b \sin \beta \sin \theta & (a \sin \alpha + b \cos \beta) \cos \theta \\ -a \sin \alpha & -b \cos \beta & 0 \end{bmatrix} \begin{bmatrix} w_\alpha \\ w_\beta \\ w_\theta \end{bmatrix} \quad (5-5)$$

简写为  $\vec{V} = A\vec{w}$

由于任意时刻机械臂都有确定的方位角，故 $A$ 为已知量，因而给定一组想要得到的 $(v_x, v_y, v_z)$ ，则可通过计算  $\vec{w} = A^{-1}\vec{V}$  得到对应的角度。

### 5.3 NAO机器人控制系统

项目中为了控制机器人的手臂，采用逆运动学的理论，通过机器人手臂的端点结合手臂的参数来计算关节必要的旋转值来进行动作。由于NAO机器人手臂（不包括手指部分）的自由度是5个，因此在仅要设置手指尖端的目标坐标点的情况下，会出现手臂姿态不唯一的情况。通过筛选使得手臂各个关节与端点的坐标实现一一对应。



图 5.5 NAO机器人

对于机器人的姿态设定、手臂末端的设定涉及到机器人的控制理论。每个关节有相应的标准坐标系来解读链接和位置以及决定关节转换的基本程序。从参考点开始到第一个关节，再最终到最后一个关节完成转换后可以得到机械臂的转换矩阵。利用这个转换矩阵可以根据手臂末端的目标位置得到每个关节的转动参数。

我们主要采用Denavit Hartenberg (DH)来进行运动学的计算。根据厂家提供的机器人的手臂参数（如长度等），得到每个关节n的转换矩阵 $A_n$

$$A_n = \begin{bmatrix} C_{\theta_n} & -S_{\theta_n}C_{\alpha_n} & S_{\theta_n}S_{\alpha_n} & a_nC_{\theta_n} \\ S_{\theta_n} & C_{\theta_n}C_{\alpha_n} & -C_{\theta_n}S_{\alpha_n} & a_nS_{\theta_n} \\ 0 & S_{\alpha_n} & C_{\alpha_n} & d_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-6)$$

其中  $C_{\theta_n} = \cos(\theta_n), S_{\theta_n} = \sin(\theta_n)$ ,  $\theta_n$  为官方给出的旋转角,  $a_n$  为两个关节参考坐标系中坐标轴交界口的距离,  $d_n$  为两个关节在同意旋转轴上的距离, 这些数据都可以在NAO的数据手册中查询到。

为了根据目的点的坐标分离方程, 以得到各个关节的转动值, 我们将变换矩阵"An"的逆矩阵相乘于关系方程的左边来取得计算角度值的元素。右手包含五个关节, 从上到下分别是肩膀关节(RShoulderPitch, RShoulderRoll 两个), 手肘关节(elbow RElbowRoll, RElbowYaw 两个), 和手腕关节(RWristYaw)。因此, 表示从肩膀到手腕各关节的转动矩阵为 ${}^w_s T$

$${}^w_s T = A_1 \times A_2 \times A_3 \times A_4 \times A_5 \quad (5-7)$$

$$A = A_1 \times A_2 \times A_3 \times A_4 \times A_5 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p_x \\ r_{21} & r_{22} & r_{23} & p_y \\ r_{31} & r_{32} & r_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5-8)$$

然后我们通过逐步乘以  $A_i^{-1}$  的方式得到各个关节的变换矩阵A, 进而可以通过以下公式得到每个关节的转动参数。

$$Pos\_w = A \times Pos\_s \quad (5-9)$$

由以上方程并结合NAO的链接关节变量数据 (在NAO官网上提供), 我们可以根据目的点得到每个关节点旋转的角度, 从而达到基于逆运动学控制机器人的机械臂。

## 第6章 展望

### 6.1 项目改进

给NAO编写行走程序

通过采集每次操作的数据，可以使用增强学习训练机器人自主地执行任务。

为了完善机器人的功能，可以加上手势识别元件实现机器人手部的动作。

### 6.2 未来应用

此技术可以用于拆弹，救援，远程交互等领域。

随着科技技术的不断发展，人类正向许多未知的领域前进，而这个过程注定危险重重。如深水、深地层、外层空间、强放射、高真空等新的生产和作业领域。

例如，随着航天事业的不断发展，如今人类已经不再局限于把飞行器送上天，还要求它们完成各种空间任务。在地球上尚且需要制造机器人来代替人完成各种工作，而在太空中只有几个宇航员，在这样危险的环境中，就只需采用本技术的Vive，就可以在空间舱中看到外面，进而遥控机器人协助或代替宇航员完成大量艰巨危险的任务。

而且，面对高强度工作，可通过本技术遥控机器人代替完成。比如，可解决大型装备的安装、拆卸等工作。

未来此技术可以用于拆弹，救援，远程交互等领域。