

# 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

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人类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

|                         |    |
|-------------------------|----|
| 第 1 章 项目背景 .....        | 1  |
| 第 2 章 现有系统问题 .....      | 2  |
| 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 .....   | 3  |
| 2.2.3 T-HR3 .....       | 4  |
| 第 3 章 系统简介 .....        | 5  |
| 3.1 VR 远程监视系统 .....     | 5  |
| 3.2 Vive 追踪器遥操作系统 ..... | 5  |
| 第 4 章 技术细节 .....        | 7  |
| 4.1 实时图像采集与显示系统 .....   | 7  |
| 4.2 机械臂控制系统 .....       | 7  |
| 4.3 NAO 机器人控制系统 .....   | 8  |
| 第 5 章 项目展示 .....        | 10 |
| 第 6 章 未来应用 .....        | 11 |

## 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。  |

## 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

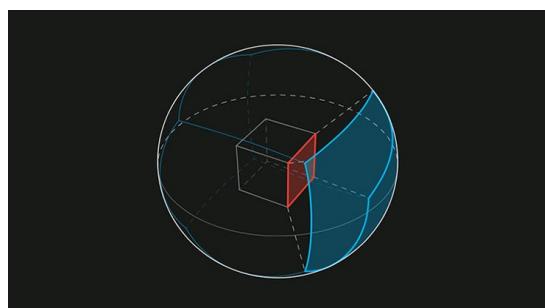


图 2.1 立方体拼接

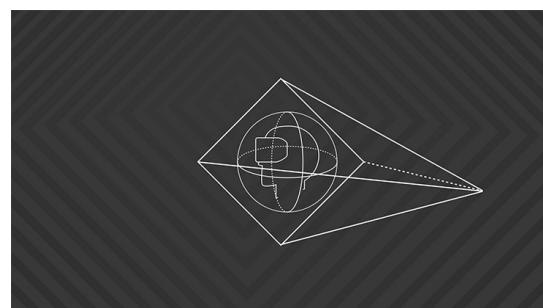


图 2.2 金字塔形拼接



图 2.3 Google Jump

### 2.2.2 OptiTrack

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

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 头盔价格可以为普通人所接受，Google Cardboard 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。

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。



图 3.1 Vive VR 头盔



图 3.2 Google Cardboard

Vive 虚拟现实套件<sup>①</sup>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 技术细节

### 4.1 实时图像采集与显示系统

项目用于和 Vive 头盔进行互动的是一个 ZED 双目摄像头。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，用于与用户交互。

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，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 Arduino Mega 2560，控制双自由度数字舵机做出与 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。

演示视频参见附件 XXX

### 4.2 机械臂控制系统

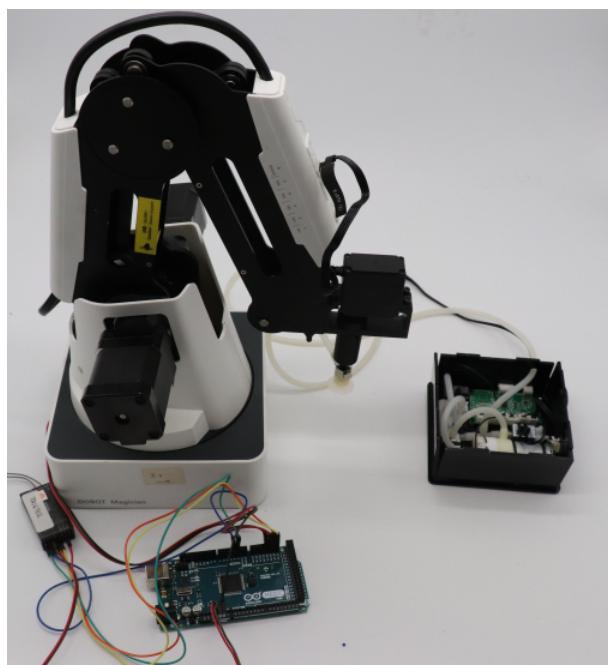


图 4.1 DOBOT 机械臂及通讯系统

① <http://www.nvidia.cn/object/embedded-systems-dev-kits-modules-cn.html>  
② <https://www.linux-project.org/uv4l/>

### 4.3 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 4.2 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  $A_n$

$$A_n = \begin{bmatrix} C_{\theta_n} & -S_{\theta_n}C_{\alpha_n} & S_{\theta_n}S_{\alpha_n} & a_nC_{\theta_n} \\ S_{\theta_n} & C_{\theta_n}C_{\alpha_n} & -C_{\theta_n}S_{\alpha_n} & a_nS_{\theta_n} \\ 0 & S_{\alpha_n} & C_{\alpha_n} & d_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4-1)$$

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  ${}^W_S T$

$${}^W_S T = A_1 \times A_2 \times A_3 \times A_4 \times A_5 \quad (4-2)$$

$$A = A_1 \times A_2 \times A_3 \times A_4 \times A_5 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p_x \\ r_{21} & r_{22} & r_{23} & p_y \\ r_{31} & r_{32} & r_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4-3)$$

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。

$$Pos\_w = A \times Pos\_s \quad (4-4)$$

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5章 项目展示

现阶段已完成一个由全向轮底盘移动的机械臂，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。可以看到图片中输出了左手，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。现在已实现用遥控器-接收机控制系统，远程完成叠纸杯实验。另外，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。四轮全向轮底盘也搭建完成。

经测试，我们这套系统远程控制实时性很好，视频传输感受不到延迟，在 VR 眼睛中有很强的立体感。此外，控制较传统方式更简便，追踪器较外骨骼造价低。

在之后的一段时间内，我们将调用追踪器位置信息，发送给下位机，实现追踪器控制机械臂的功能。

另外我们刚刚借到了 Now 机器人，准备在操作者手脚上放置四个追踪器，以此来控制机器人的运动，如站立，行走等。

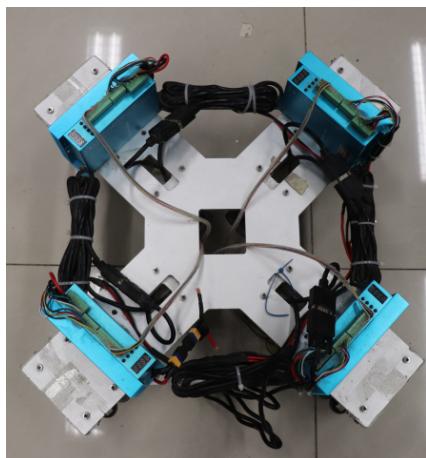


图 5.1 全向轮底盘

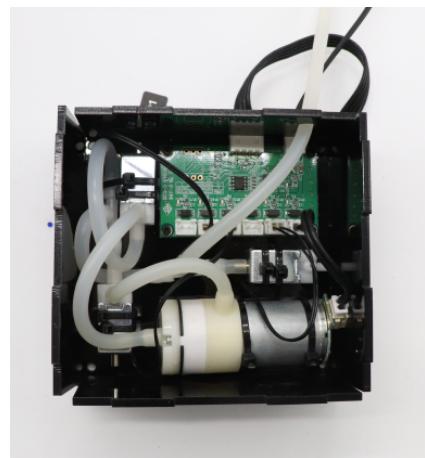


图 5.2 气泵和电磁阀组建的吸附装置

图 5.3 追踪器和 VR 头盔绝对位置

## 第6章 未来应用

未来此技术可以用于拆弹，救援，远程交互等领域。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。

此外，该系统中机器人可不断对操控者的行为进行学习。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。