附录1：（封面、封底用120克白色铜版纸打印）



**本科生毕业设计[论文]**

（华文中宋小初号加粗居中）

**基于图像的视觉伺服**（黑体2号加粗居中）

院 系\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

专业班级\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

姓 名\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

学 号\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

指导教师\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

年 月 日（华文中宋3号居中）

**学位论文原创性声明**

（黑体小2号加粗居中）

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

（宋体小4号）

作者签名： 年 月 日

**学位论文版权使用授权书**

（黑体小2号加粗居中）

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密囗，在 年解密后适用本授权书

2、不保密囗 。

（请在以上相应方框内打“√”）

（宋体小4号）

作者签名： 年 月 日

导师签名： 年 月 日

(注：此页内容装订在论文扉页)

**摘 要**

相比使用传统传感器来作为机器人获取周围信息的方式，由于视觉传感器具有成本低、信息丰富、可靠性高等优点而得到了广泛的应用[1]。视觉伺服，即基于视觉的控制系统，已成为机器人领域的重要发展方向之一。基于图像的视觉伺服不需要对系统参数进行精确的标定，直接从图像中获取信息。本文推导了在具有6个自由度的手眼情况下基于图像视觉伺服中基本的机器人动力学模型、交互矩阵、控制器的设计以及误差函数的设计[2]。并且本文在UR3双臂机器人系统上，基于ROS（机器人操作系统）构建了完整的算法验证平台。使用HOG特征加SVM识别手写体所在位置从而提取关键点来构建交互矩阵，求解正运动学与反运动学，设计位置控制器，实现机器人简单的位姿控制与手写体数字的跟踪。

**关键词：**基于图像视觉伺服；机器人；ROS；运动学；手写体识别；SVM；手眼；位置控制器设计

**Abstract**

Comparing to using the traditional sensor to gather the information around the robots, the vision sensor is widely used because of its low cost, rich information and high reliability[1]. Visual servo, which is a control system based on the vision data, has become one of the important development direction of the field of robotics. Image-Based Visual Servo can capture information directly from the image without accurate calibration for the system parameters. This paper deduces the basic robotics kinematic dynamics model、interaction matrix、design of the controller and the design of the error function based on Image-Based Visual Servo in the case of hand-in-eye with 6-DOF[2].In this paper,a complete algorithm validation platform has been built on the UR3 dual-arm robot system based on ROS (Robot Operating System).Using HOG feature and SVM to detect the hand written digit number’center point in image,constrcuting the interaction matrix,solving the forward kinematics and inverse kinematic,designing position controller, realizing simple postion-attitude control and object tracking of hand written digit number.

**Key Words：**Image-Based Visual Servo;Robotics;ROS;Kinematics;Hand written digit number recognition;SVM;Eye-in-hand;position controller

**1 绪论**

**1.1 前言**

近年来，机器人技术飞速发展，机器人在工业、农业、服务业以及军事等方向得到了广泛的应用。与此同时，对机器人执行复杂任务的需求也日益增加，具备感知能力的智能机器人被认为是机器人发展的一大方向。

为了使机器人能够在复杂环境下工作，必须提高它的对环境的探测能力与对自身运动控制的能力。使其能在复杂环境下完成自身定位、地图构建、自主搜索等任务。为此，必须为机器人本体装配各种传感器,使它们能够获取外部环境信息进一步完成定位、探查、搜索等功能。

传统的检测手段往往面临检测范围的局限性以及检测手段的单一性，而且可拓展性差，包含信息相对简单，成本过高等缺点。而视觉传感器具有低成本、高信息、非接触式测量、适用范围广和算法简单等优点，成为最重要的机器人传感器之一。采用视觉作为反馈的视觉伺服控制是机器人系统重要的控制手段，可运用如零件的自动检测、智能机器人控制以及生产线的监控等工业自动化生产过程中，同时视觉伺服在运动目标的自动跟踪与识别等方面也有广泛的应用。可以认为视觉伺服控制是真正实现机器人智能控制的关键。

目前大部分研究都建立在机器状态可解析性描述的基础上，并以某些特定的任务为目标，难以应用于复杂多变的实际应用中。因此如何使机器人真正具有与人一样的视觉能力，是一个需要长期研究的问题。

**1.2 研究现状**

视觉伺服控制涉及计算机视觉、机器人动力学、机器人运动学以及控制理论等多个领域的交叉学科，国内外学者在过去20年间进行了大量的研究。机器人视觉伺服主要是采用视觉传感器采集信息，通过姿态检测算法间接检测机器人的当前位姿或者相对位姿。国外Hutchinson等三篇论文[1-3]对视觉伺服控制研究做引导。国内，方勇纯[4]、薛定宇[5]等对视觉伺服的控制方法进行了综述，总结了经典的视觉伺服的控制算法，陶波[6]等对无标定视觉伺服控制研究进展进行了描述。

相比传统的在标定条件下使用视觉伺服控制系统，无标定视觉伺服控制系统具有更高的灵活性欲适应性，同时对模型的偏差以及图像噪声的等具有较强的鲁棒性，是机器人伺服控制系统未来重要的发展方向和研究热点[6]。虽然视觉伺服的研究已经有20年的历史，但是由于涉及学科众多以及这些学科的发展现状的限制，还有很多问题需要解决。视觉伺服未来研究的方向主要在图像特征的选择、无标定的视觉伺服的研究、视觉系统动态性能的研究、如何结合计算机视觉以及图像处理的最新研究算法、各种智能算法的结合以及多传感器的融合等[5]。

随着科技的不断进步，机器人作业任务将日趋复杂化、多样化。可以预见，视觉伺服控制技术将成为机器人控制的主要手段之一。机器人技术也将从专业技术领域逐步渗透到社会生活的各个方面。因此，在机器人技术的推广过程中，不需要专业标定环节的无标定视觉伺服方法凭借其更高的灵活性与适应性，势必会成为视觉伺服领域最为重要的分支之一。目前无标定视觉伺服技术虽然取得了较为丰硕的成果，但是远未达到成熟可靠的境界[6]。

**2 视觉伺服基础**

**2.1 视觉伺服基本概念**

视觉伺服控制是指使用计算机视觉数据控制机器人的运动。视觉数据可以是从直接安装在机器人上的相机获取操纵器或移动机器人上，在这种情况下的运动机器人引起相机运动，或者相机可以被固定在工作空间中，从而可以观察机器人的运动固定配置。

视觉伺服的控制方法的目的是最小化误差e(t)，一般定义如下：

这个公式具有一般性，他具有广泛的求解方法，将会在下面看到。式子(2-1)中的参数定义如下，向量m(t)是一组图像测量（例如，兴趣点的图像坐标或对象的质心的图像坐标）。这些图像的测量用于计算k个视觉特征的向量s（m（t），a），其中a是表示的一组参数关于系统的潜在附加知识（例如，粗相机固有参数或对象的3-D模型）。矢量s\* 包含所需要的特征的值。而视觉伺服方案的主要区别在于对目标函数s的选择与设计上面。

**2.2 基于图像的视觉伺服**

本文主要研究实现基于图像的视觉伺服控制的设计，基于图像的视觉伺服无需对系统参数进行精确的标定、信息可以直接从图像获取等优点得到广泛应用。本文在研究中，构建的目标函数所需要的图像的测量直接是感兴趣点所在图像坐标系中的坐标（例如多边形的定点），所以s(m(t),a)中m(t)可以直接由感兴趣点坐标直接表示，所以目标函数的构建可以直接完成。在本文中，m(t)由感兴趣矩形物体的四个顶点的图像坐标构成。

**2.3 摄像机系统**

由于采用的固定在6个自由度的机器手上面的手眼，手眼的运动会造成物体在摄像机成像平面上的运动，所以对摄像机系统的分析以及运动的模型建立十分重要。一般情况下，对于相机系中具有三维坐标点,最后投影在成像平面上的图像具有二维坐标点,我们有：

其中给出以像素为单位表示的图像点的坐标， (，，，)是相机固有参数集合：和是主点，是焦距，是像素尺寸的比率。式子(2-3)为相机的投影方程。

**2.4 交互矩阵推导**

为了实现视觉伺服的控制设计，s随时间变换和摄像机速度之间的关系。 让摄像机的空间速度表示为，其中是相机系原点的瞬时线速度，是相机系的瞬时角度。

和之间的关系由下式给出:

其中，叫做s相关的交互矩阵。在一些视觉伺服文献中也叫作雅克比特征矩阵。本文设计的是基于图像的视觉伺服，在2.2中提到直接采用图像坐标作为目标函数，所以直接选取,其中就是特征点或兴趣点在图像中的坐标。所以需要推导s随时间的导数与手眼相机6个自由度的速度的关系。

取投影方程(2-2)关于时间的导数，可以得到：

式子(2-5)得到了,还需要得到其与相机速度之间的关系。这里速度分为两种速度，一个平动（位移）速度以及一个转动速度，所以这里需要用到坐标系变换矩阵。坐标系变换矩阵可以写成,其中是表示旋转矩阵，表示平移矩阵。则坐标系的变换可以写成：

其中分别表示旋转后的在坐标系中的坐标以及旋转前的坐标。

本文使用的旋转矩阵为：

其中分别为绕轴转动的欧拉角。

假设在无穷小的时间内，角度的变化也很小,则有：

将(7)带入式子(6)中，可以得到;

而其中.

则有：

由于在无穷小的时间内，所以则有：

即：

但是所有的改变由手眼的运动而引起的，所以运动速度要改变相应符号。得出在由具有6自由度的手眼运动而产生图像中特征点在图像坐标系中运动的方程为：

**2.5 控制器设计**

一旦选择了s，控制方案的设计可以非常简单。 也许最直接的方法是设计一个速度控制器。 为了做到这一点，我们需要s随时间变换和摄像机速度之间的关系。 让摄像机的空间速度表示为，其中是相机系原点的瞬时线速度，是相机系的瞬时角度。

和之间的关系由下式给出:

其中，与s相关的交互矩阵。术语雅可比特征也在某种程度上可互换使用在视觉伺服文献中。

用式子（2-1）和式子（2-4），我们可以立即得到相机速度和误差的动态特性之间的关系：

其中，将设置成机器人控制器的输入，使误差值随时间而呈指数形式下降（即），我们通过（2-6）可以获得:

其中，被选择为的Moore-Penrose伪逆，就是当的秩是6满秩时，，

这个选择方法能够让和达到最小。当k=6时，且时，可以将求逆，然后可以给出控制

在现实的视觉伺服系统中，在实践中你不可能完全知道或者。因此，必须得到这两个矩阵之一的近似或估计。接下来，我们用符号来表示交互矩阵的近似的伪逆和交互矩阵的伪逆的近似。使用这个概念，控制律事实上就是:

**3 机器人运动学**

**3.1 机器人运动学基本概念**

机器人系统由机器人和作业对象与环境构成。机器人由机器人本体、控制器、软件三部分组成。为了获得作业对象信息与控制器输入，还需要传感器系统。机器人的本体主要有机械臂与移动机构组成，在许多应用中不需要移动机构。机械手由手臂与末端执行器构成。手臂通过多个关节的运动使末端执行器进行预设的运动到达预设的位置。每个手臂拥有一定自由度，自由度指机器人所具有的可进行独立坐标系运动的数目，一般情况下为了使机器人手部能在工作空间内以任意姿态到达任意位置，机器人至少需要六个自由度。

研究机器人运动学主要是要解决下面两个基本问题：

1. 对于一个给定的机器人，已知杆件的几何参数与关节参数，求末端执行器相对于给定坐标系的位置与姿态。给定坐标系一般为世界坐标系。
2. 对于一个给定的机器人，已知杆件的几何参数，给定末端执行器相对于机座坐标系位置与姿态，确定关节参数。

第一个问题常称为正运动学问题，第二个问题常称为反运动学问题。

机器人运动学运用几何学研究构成机器人系统结构的多个自由度运动学链的运动规律，几何强调机器人的连接部分被认为是刚体并且其中假设关节点只能进行转动或者平动。机器人运动学研究机器人系统中运动链的维度连通性与每个连接点的位置，速度和加速度之间的关系。为了规划以及控制运动路径并且输出给执行器的力和力矩，质量和固有内部属性，运动以及相关联的力和力矩需作为机器人动态系统的一部分来研究[wiki https://en.wikipedia.org/wiki/Robot\_kinematics]。

**3.2 运动学方程**

能够表示机器人运动链的运功学方程是机器人运动学中最基本的工具。运动学方程一般是一组非线性方程，描述成关节点相关参数与机器人配置系统之间的映射。运动学方程是机械系统的约束方程，例如机器人操纵器，其定义在一个或多个关节处的输入如何指定装置的配置，以到达任务位置或末端效应器位置。

Denavit-Hartenberg（D-H）方法在描述机器人动力模型中被广泛使用。D-H方法描述构成的串行机器人几何结构已经是研究机器人的基本准则。标准的D-H方法用来建立一个机器人的动力学模型。

在D-H方法表示的系统中，机械手由一串用转动或平移关节连接的刚体组成。每一对关节-杆件组成了一个自由度，每个杆件至多与两个杆件相连即不会构成闭环。

每个杆件可以由两个参数来表示。第一个参数为杆件的长度，即杆件上两个关节轴线的最短距离；第二个参数为杆件旋转角度，即两个相邻关节与的关节轴之间的夹角。

关节也用两个参数来表示。第一个参数是偏移量，即一个杆件到下一个杆件之间沿着关节的轴的距离；第二个参数是关节旋转角度，即一个杆件相对下一个杆件在关节轴上面的旋转角度。

通常这四个参数中三个参数是固定的，而一个叫做关节变量。对于一个旋转关节，关节变量就是，而对于一个移动关节，关节变量就是。D-H的标准相关定义在图XXX中描述。

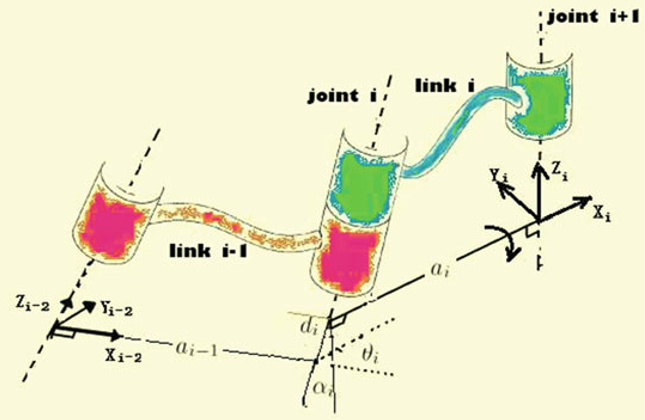


图 XXX□标准D-H杆件关节相关定义

为了描述相邻杆件之间平移与转动的关系，可使用D-H方法为每个关节处的杆件建立4X4齐次变换矩阵。这样通过可以将每个自由度的运动逐次传播至末端执行器。

对于坐标系与几何参数做如下定义：

1）轴为沿着关节的运动轴；

2）轴垂直于轴并指向离开轴的方向；

3）轴按右手坐标系建立；

4）为绕轴由轴向轴的关节角；

5）位从第坐标系原点到轴和轴的交点沿轴的距离；

6）为绕轴由轴转向轴的偏角；

7）为与两轴间的最小距离；

通过这些参数，连接部分其次变换矩阵将连接部分的坐标系转到坐标系可以由公式（3-1）表示。



其中缩写成,缩写成,表示绕着轴j的旋转，表示沿着轴j的平移。对于一个n杆件的机器臂，全部的手臂的变换运动叫做正运动学，所有独立的杆件的变换可以写成一个形式，如公式（3-2）所示。在笛卡尔坐标系中末端执行器的位置可以用公式（3-3）来计算。





其中 是笛卡尔坐标系中的一个增广向量。

**4 图像处理**

**4.1 图像处理目标与方法**

本文主要工作是基于图像的视觉伺服算法设计，所需要的信息与反馈需要直接从图像中获取，以此构建控制器。

本文将手写体数字作为识别目标分别为0到9的10个数字，使用图像的HOG特征，通过对MNIST[]的手写体数据库来SVM分类器作为识别方法，能够准确实时且在一定干扰下提取出手写数字信息，并获得各图像在摄像头中的像素坐标，控制机械臂按照目标一到目标三的顺序依次移动到目标正上方。

**4.3 HOG特征**

HOG(方向梯度直方图)是一种在计算机视觉以及图像处理中广泛应用于物体检测的特征描述符。HOG特征通过计算和统计图像中局部区域的梯度方向直方图来构成特征。HOG特征结合SVM分类器已经广泛应用于图像识别当中。关于HOG特征的一些介绍不再赘述。

**4.3 支持向量机**

支持向量机(SVM)是由Vapnik提出，针对回归于分类问题的一种统计学习理论。它是一种在高维特征空间中使用线性函数假设的学习方法，是具有准则性的最优化学习算法。

SVM的主要思想可以为以下两点：

1. SVM基于线性可分情况进行分析，若问题线性不可分，使用非线性[映射](http://baike.baidu.com/item/%E6%98%A0%E5%B0%84)算法将低维输入空间线性不可分的样本转化为[高维](http://baike.baidu.com/item/%E9%AB%98%E7%BB%B4)特征空间使其线性可分，从而将非线性特征分析转化为线性分析。
2. SVM基于结构风险最小化理论，通过在特征空间中构建最优分割超平面，使得问题得到全局最优解，并且在整个[样本空间](http://baike.baidu.com/item/%E6%A0%B7%E6%9C%AC%E7%A9%BA%E9%97%B4)的期望风险以某个概率满足一定[上界](http://baike.baidu.com/item/%E4%B8%8A%E7%95%8C)。

SVM在解决非线性以及高维模式识别中，表现出了克服“维数灾难”以及“过拟合”等传统统计学习方法所遇到困难的特点，目前广泛应用于人脸识别、手写体识别、手势识别等领域。

**4.4 手写体定位与识别**

本文的手写体数字识别为脱机手写体数字识别，该方法分为几个步骤，第一步为预处理，第二步为字体定位，第三部为手写体识别，下面针对这三个方面进行阐述。

**4.4.1 预处理**

为了提高手写体数字识别的精度，排除实际环境中的各种干扰，需要对数据库中的图像做一系列预处理。在预处理方面已经有非常深入且成熟的研究，本文的预处理包括图像灰度化、高斯模糊、二值化、腐蚀。

整个流程为图xxx所示。

图 XXX□图像处理流程

**5 ROS（机器人操作系统）**

**5.1 ROS介绍**

ROS (Robot Operating System, 机器人操作系统)是目前主流的机器人控制框架，它由斯坦福大学人工智能实验室和Willow Garage合作研发，极大的提高了开发效率。

ROS提供了操作系统应有的服务，包括硬件抽象，底层设备控制，常用函数的实现，进程间消息传递，以及包管理。它也提供用于获取、编译、编写、和跨计算机运行代码所需的工具和库函数。提供一系列程序库和工具以帮助软件开发者创建机器人应用软件。ROS遵循BSD开源许可协议。

节点（node）,话题（topic）,消息（message）是ROS的核心概念，三者一同构成了机器人中所有模块的交互方法。图xxx为ROS概念图。

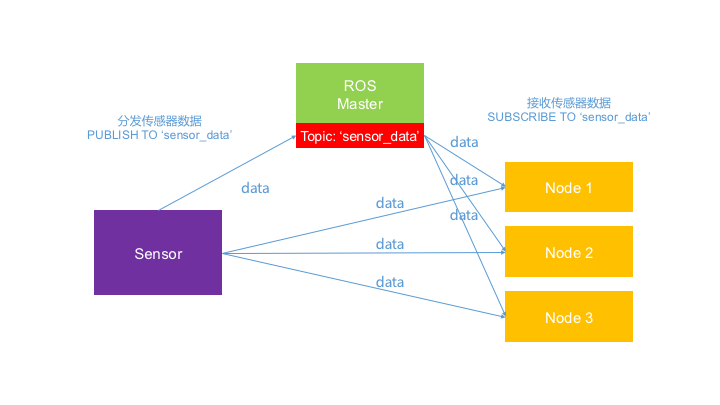


图3-1□ROS概念图

在ROS中，可以自由地定义节点，一个节点基本对应一个功能或实体，该节点可以与其他节点通信，具体方式是节点在整个系统中广播话题，话题下面包含各类型消息，其他的节点可以订阅消息，可以实现各节点之间通信以及数据的传输。

各节点之间关系是相互独立的，概念图中的节点能够订阅指定消息而不需要知道该消息是哪个节点发送的，这样的结构在项目建设中起到重要的作用，不会因为某些节点的异常而影响整个系统的运行。

**5.2 服务**

节点之间消息的发布/订阅模型是一个非常灵活的通信模式，但是经常需要在分布式系统中使用到的请求/回复的交互机制在多对多的单向传输是不适合的。所以服务相比消息的发布/订阅主要用来解决请求/回复的功能。请求/回复是通过服务来完成的，服务由一对消息定义：一个用于请求，一个用于回复。客户端通过发送请求消息并等待应答来调用服务。这样可以实现客户端与服务器端持久的连接。

**5.3 启动文件**

ROS中另外一个比较重要的概念就是启动文件。启动文件是一个xml格式的文本文件，ROS通过解析启动文件来完成启动节点，设置参数、传递参数和重映射名称等功能，同时能够启动其他的启动文件。这样为我们将多个节点、工程、启动文件以及一些参数的设置与传递提供了很大的方便。同时启动文件可以启动多个，个个节点都是独立运行。通过对启动文件的修改，可以轻松的将需要启动的节点以及一些参数与配置文件等结合起来形成一个完成的工程从而完成特定的任务与目标。

**6 算法平台构建**

**6.1 UR3机器臂驱动**

本文采用的算法验证平台的机器臂为Universal Robots公司推出的UR3机器臂。UR3具有6个自由度，最大工作范围19.7英寸，最大负载3公斤。采用的驱动的是Shaun Edwards等人开发的ROS的universal\_robot工具包，提供了操作机器臂各个关节的相关节点、服务、机器人模型以及参数等信息。通过该工具包可以实现对各个关节位置以及速度的控制从而实现对末端执行器的控制。

**6.2 KDL运动学求解库**

在6.1中提到通过universal\_robot工具包可以控制各个关机的位置以及速度，但是要控制末端执行器的位姿，就需要求解运动学方程。本文使用的是一个运动学求解库KDL(Kinematics and Dynamics Library)，是一个独立的框架可以建模和计算的运动链，如机器人，生物力学人体模型，计算机动画人物，机床等。它提供了各种几何对象（点，系，线，…）类库，各种运动链的集合（串行，仿人，并行，移动…）及其运动规范和插值操作。

所以对UR3机器臂的运动学方程的建立与求解通过KDL来完成、首先加载机器人模型、世界坐标系以及初始安装位置，同时获得当前机器人各个关节的位置，通过建立的正运动学方程计算出末端执行器在世界坐标系中的位置。控制末端执行器的运动就需要将期望的点的世界坐标系的坐标通过建立的反运动学方程求解出期望的各个关节的位置参数等，然后通过插值形成轨迹让机器臂的各个关节跟随轨迹从而实现末端执行器的运动控制。整个对末端执行器的控制流程图如图xxx所示。

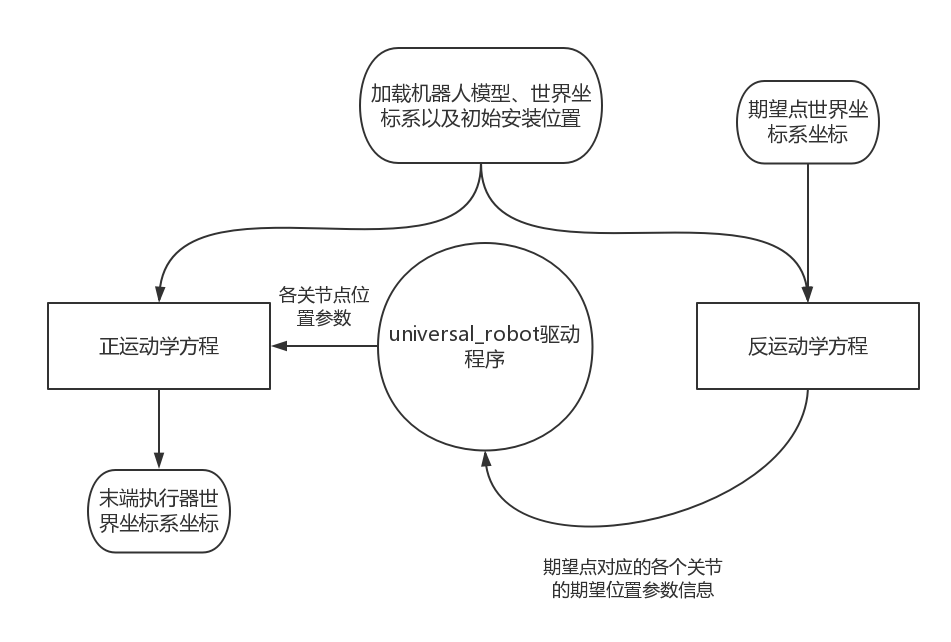


图3-1□控制流程

**6.3 手眼**

采集图片信息本文使用了通用USB摄像头作为图像捕获设备，将USB摄像头固定在机器臂的末端成为末端执行器。驱动采用ROS下的Benjamin Pitzer等人开发的usb\_cam工具包，将摄像头获取的图像作为话题以30HZ频率发布出去。以供图像处理节点订阅消息进行图像处理后获取所需信息。

**6.4 手写体识别**

图像处理采用OpenCV，具体流程在4.2中描述。采用的是ROS下的cv\_bridge工具包将获取到的图像消息转换成OpenCV中的图像数据类型（Mat或IplImage）后再运行图像检测的算法。

**6.5 整体架构**