

5.1 机器人视觉简述

机器人视觉(robot vision)是使机器人具有视觉感知功能的系统,是机器人系统组成的重要部分之一。机器人视觉可以通过视觉传感器获取环境的二维图像,并通过视觉处理器进行分析和解释,进而转换为符号,让机器人能够辨识物体,并确定其位置。机器人视觉广义上称为机器视觉,其基本原理与计算机视觉类似。计算机视觉研究视觉感知的通用理论,研究视觉过程的分层信息表示和视觉处理各功能模块的计算方法。机器视觉侧重于研究以应用为背景的专用视觉系统,只提供对执行某一特定任务相关的景物描述。机器人视觉硬件主要包括图像获取和视觉处理两部分,而图像获取由照明系统、视觉传感器、模拟/数字转换器和帧存储器等组成。根据功能不同,机器人视觉可分为视觉检验和视觉引导两种,它广泛应用于电子、汽车、机械等工业部门和医学、军事领域。

5.1.1 机器人视觉的相关定义

在机器人控制领域,视觉控制是当前的一个重要研究方向,也是目前的研究热点之一。机器人视觉控制与计算机视觉、控制理论等学科密切相关,但它采用的概念又有所不同。为便于理解,在此对机器人视觉控制中的部分概念予以简要介绍。

(1) 摄像机标定(camera calibration)。对摄像机的内部参数、外部参数进行求取的过程。摄像机的内部参数通常又称为内参数(intrinsic parameter),主要包括光轴中心点的图像坐标、成像平面坐标到图像坐标的放大系数、镜头畸变系数等;摄像机的外部参数又称为外参数(extrinsic parameter),是摄像机坐标系在参考坐标系中的表示,即摄像机坐标系与参考坐标系之间的变换矩阵。

(2) 视觉系统标定(vision system calibration)。对摄像机和机器人之间关系的确定称为视觉系统标定。例如,手眼系统的标定就是对摄像机坐标系与机器人坐标系之间关系的求取。

(3) 手眼系统(hand-eye system)。它是由摄像机和机械手构成的机器人视觉系统。摄像机安装在机械手末端并随机械手一起运动的视觉系统称为 eye-in hand 式手眼系统;摄像机不安装在机械手末端,且摄像机不随机械手运动的视觉系统称为 eye-to hand 式手眼

系统。

(4) 视觉测量(vision measure 或 visual measure)。根据摄像机获得的视觉信息对目标的位置和姿态进行的测量称为视觉测量。

(5) 视觉控制(vision control 或 visual control)。根据视觉测量获得目标的位置和姿态,将其作为给定或者反馈对机器人的位置和姿态进行的控制,称为视觉控制。简言之,视觉控制就是根据摄像机获得的视觉信息对机器人进行的控制。视觉信息除通常的位置和姿态之外,还包括对象的颜色、形状、尺寸等。

(6) 视觉伺服(visual servo 或 visual servoing)。利用视觉信息对机器人进行的伺服控制称为视觉伺服。视觉伺服是视觉控制的一种,视觉信息在视觉伺服控制中用于反馈信号。在关节空间的视觉伺服可直接对各个关节的力矩进行控制。

(7) 平面视觉(planar vision)。被测对象处在平面内,只对目标在平面内的信息进行测量的视觉测量与控制称为平面视觉。平面视觉可以测量目标的二维位置信息以及目标的一维姿态。平面视觉一般采用一台摄像机即可完成,而且摄像机的标定比较简单。

(8) 立体视觉(stereo vision)。对目标在三维笛卡儿空间(Cartesian space)内的信息进行测量的视觉测量与控制称为立体视觉。立体视觉可以测量目标的三维位置信息,以及目标的三维姿态。立体视觉一般采用两台摄像机,需要对摄像机的内外参数进行标定。

(9) 结构光视觉(structured light vision)。利用特定光源照射目标,形成人工特征,由摄像机采集这些特征进行测量,这样的视觉系统称为结构光视觉系统。由于光源的特性可以预先获得,光源在目标上形成的特征具有特定结构,所以这种光源被称为结构光。结构光视觉可以简化图像处理中的特征提取,也能大幅度提高图像处理速度,具有良好的实时性。结构光视觉属于立体视觉。

(10) 主动视觉(active vision)。对目标主动照明或者主动改变摄像机参数的视觉系统称为主动视觉系统。主动视觉可以分为结构光主动视觉和变参数主动视觉。

5.1.2 机器人视觉的基本原理

机器人视觉系统硬件主要包括图像获取和视觉处理两部分,而图像获取由照明系统、视觉传感器、A/D 转换器和帧数存储器等组成。根据功能不同,机器人视觉可分为视觉检验和视觉引导两种。

典型的机器人视觉系统一般包括光源、图像采集装置、图像处理单元(或图像采集卡)、图像分析处理软件、监视器、通信输入/输出单元等。实际上,图像的获取是将被测物体的可视化图像和内在特征转换成能被计算机处理的数据,这种转换的过程及其效果直接影响到视觉系统工作的稳定性及可靠性。被测物体图像的获取一般涉及光源、相机和图像处理单元(或图像捕获卡)。

光源是影响机器视觉系统输入的重要因素,因为它影响着输入数据的质量和至少 30% 的应用效果。由于没有通用的机器视觉照明设备,所以针对每个特定的应用实例,要选择相应的照明装置,以达到最佳的照明效果。许多工业用的机器视觉系统利用可见光作为光源,这主要是因为可见光容易获得,价格低廉,且便于操作。

常用的可见光源包括白炽灯、日光灯、水银灯和钠光灯。但是,这些光源的最大缺点是不能保持稳定。以日光灯为例,在使用的第一个 100 小时内,光能将下降 15%,随着使用时

间的增加,光能还将不断下降。因此,如何使光能在一定的程度上保持稳定,是实用性过程中急需解决的现实问题。另外,环境光将改变这些光源照射到物体上的总光能,使输出的图像数据存在噪声,一般采用加装防护屏的方法,减少环境光的影响。由于存在上述问题,在现今的工业应用中,对于某些要求严格的检测任务,常采用 X 射线、超声波等不可见光作为光源。

由光源构成的照明系统按其照射方式可分为背向照明、前向照明、结构光照明和频闪光照明等。其中,背向照明是将被测物放在光源和相机之间,其优点是能够获得高对比度的图像;前向照明是将光源和相机位于被测物的同侧,该方式的优点是便于安装;结构光照明是将光栅或线光源等投射到被测物上,根据它们产生的畸变,解调出被测物的三维信息;频闪光照明是将高频率的光脉冲照射到物体上,要求相机的扫描速度与光源的频闪速度同步。

对于机器人视觉系统来说,图像是唯一的信息来源,而图像的质量是由光学系统的合理选择所决定的。通常来说,由于图像质量低劣引起的误差不能用软件纠正。机器视觉技术把光学部件和成像电子结合在一起,并通过计算机控制系统来分辨、测量、分类和探测正在通过自动处理系统的部件。光学系统的主要参数与图像传感器的光敏面的格式有关,一般包括光圈、视场、焦距、 F 数等。

视觉传感器实际上是一个光电转换装置,即将图像传感器所接收到的光学图像转换为计算机所能处理的电信号。光电转换器件是相机的核心器件。目前,典型的光电转换器件包括真空电视摄像管、CCD、CMOS 图像传感器等。

5.1.3 机器人视觉的要求

应用于机器人导航的视觉算法有别于其他方面的应用,其具体要求主要体现在以下几方面。

1. 实时性要求

实时性要求即算法处理的速度要快,它不但直接决定了移动机器人能够行驶的最大速度,而且切实关系到整个导航系统的安全性与稳定性。举例来说,机器人的避障算法都需要提前知道障碍物的方位以便及时动作,这种信息获得的时间越早,系统就会有越多的时间对此做出正确的响应。视觉信息处理巨大的计算量对算法程序的压力很大,对室外移动机器人尤其如此。

2. 鲁棒性要求

由于移动机器人的行驶环境是复杂多样的,要求所采用的立体视觉算法能够在各种光照条件、各种环境下都尽可能保证其有效性。室内环境的机器人导航环境相对较好,但对于室外移动机器人或者是陆地自主车 ALV,不确定性因素增加了很多,如光照变化、边缘组织等,也不存在道路平坦假设。为此,视觉导航算法在各种环境下都要求保证其有效性。

3. 精确性要求

移动机器人的立体视觉算法满足匹配精确性要求,但这种精确性与虚拟现实或者三维建模所要求的精确性是有差别的,因为立体视觉算法对道路地形进行重建的最终目的是检测障碍物,而不是为了精确描绘出场景。对于移动机器人来说,有时候忽略细节可以提高整个系统的稳定性。



视频讲解

5.1.4 机器人视觉的应用

2015年5月,国务院发布《中国制造2025》,正式将机器视觉设备在生产中的应用推向了高潮。《中国制造2025》提出,加快机械、航空、船舶、汽车、轻工、纺织、食品、电子等行业生产设备的智能化改造,提高精准制造、敏捷制造能力。2015年11月23日至25日世界机器人大会在北京国家会议中心举办,这是中国首次举办关于机器人的国际性大会。中国国家主席习近平向会议发来贺电,指出以机器人科技为代表的智能产业蓬勃兴起,成为现代科技创新的一个重要标志。中国将机器人和智能制造纳入了国家科技创新的优先重点领域。机器视觉也成了人工智能发展最快的一个分支。

视觉是人类获取信息最丰富的手段。通常人类75%以上的信息来自眼睛,而对于驾驶员来说,超过90%的信息来自视觉。同样,视觉系统是移动机器人系统的重要组成部分之一,视觉传感器也是移动机器人获取周围信息的感知器件。近十年来,随着研究人员投入大量的研究工作,计算机视觉、机器视觉等理论不断发展与完善,移动机器人的视觉系统已经涉及图像采集、压缩编码及传输、图像增强、边缘检测、阈值分割、目标识别、三维重建等,几乎覆盖机器视觉的各个方面。目前,移动机器人视觉系统主要应用于以下3方面。

(1) 用视觉进行产品的检验,代替人的目检。包括形状检验,即检查和测量零件的几何尺寸、形状和位置;缺陷检验,即检查零件是否损坏划伤;齐全检验,即检查部件上的零件是否齐全。

(2) 对装配的零部件逐个进行识别,确定其空间位置和方向,引导机器人的手准确地抓取所需的零件,并放到指定位置,完成分类、搬运和装配任务。

(3) 为移动机器人进行导航,利用视觉系统为移动机器人提供它所在环境的外部信息,使机器人能自主地规划它的行进路线,回避障碍物,安全到达目的地并完成指定工作任务。

其中,前两者属于工业机器人的范畴,随着技术的发展,研究人员也提出了视觉伺服的概念。而视觉导航,不论是自主、半自主,还是最早期的遥控方式,都是为完成一定的任务执行者(主体)和环境(客体)的交互过程。尽管如此,移动机器人的视觉对于移动机器人来说还没有到达视觉相对于人类如此重要的地步,很大一部分原因是视觉信息处理理论与方法不够完善。摄像头能在极短的时间内扫描得到数百万上千万计的像素的环境图像,甚至超过了现在人类眼睛的信息采集能力,但在处理方法及处理速度上目前远不能和人类相比。不过可以相信,随着微电子技术的进步和计算机视觉的发展,移动机器人的视觉功能会越来越强大,同时机器视觉在移动机器人信息感知中所占的比重也会越来越大。

5.2 摄像机模型

摄像机是通过成像透镜将三维场景投影到它的二维像平面上,该过程投影可用于成像变换描述(摄像机成像模型)。

5.2.1 单目视觉模型

1. 单目视觉简介

单目视觉是指仅利用一台摄像机完成定位工作。因其仅需一台视觉传感器,所以该方

法的优点是结构简单、相机标定也简单,同时避免了立体视觉中的视场小、立体匹配困难的不足。其前提条件是必须已知物体的几何模型。

基于模型的单目视觉定位可以应用在多方面,包括机器人自主导航、陆地和空间移动机器人定位、视觉伺服、摄像机校正、目标跟踪、视觉监测、物体识别、零部件装配、摄影测量等。

基于模型的单目视觉定位问题所应用的几何特征可分为点、直线与高级几何特征等几类。相对来说,目前对基于点特征的单目视觉定位方法研究较多。直线特征具有抗遮挡能力强、图像处理简单的优点,所以有一部分学者致力于基于直线特征单目视觉定位方法的研究。而基于高级几何特征的单目视觉定位方法目前研究得还比较少。

本书根据基于模型的单目视觉定位方法所使用的定位特征类型把单目视觉定位方法分为基于点特征的定位方法、基于直线特征的定位方法、基于高级几何特征的定位方法,全面介绍了各种特征定位方法的研究现状。目的是方便读者了解各种特征定位方法的研究现状,为未来的研究奠定理论基础。

2. 单目视觉定位方法

1) 点特征定位

点特征定位又称为 PNP 问题,它是计算机视觉、摄影测量学乃至数学领域的一个经典问题。PNP 问题是在 1981 年首先由 Fischler 和 Bolles 提出的,即给定 N 个控制点的相对空间位置以及给定控制点与光心连线所形成的夹角,求出各个控制点到光心的距离。该问题主要被用来确定摄像机与目标物体之间的相对距离和姿态。

经典的问题从本质上来说是非线性的,而且具有多解性。目前对 PNP 问题的研究主要包括两方面:设计运算速度快、稳定的算法,以寻找 PNP 问题的所有解或部分解。对多解现象的研究,即找出在什么条件下有 1 个、2 个、3 个或者 4 个解。

PNP 问题的研究集中在对 P3P 问题、P4P 问题、P5P 问题的研究上。这是因为如果仅使用两个特征点即 P2P 问题有无限组解,其物理意义是仅有两个点不能确定两点在摄像机坐标系下的位置。而特征点的个数大于 5 时,PNP 问题就变成了经典的 DLT 问题,是可以线性求解的。目前,人们对 P3P、P4P 问题已研究得比较清楚,并有如下结论: P3P 问题最多有 4 个解,且可以达到解的上限,对于 P4P 问题,当 4 个控制点共面时,问题有唯一解,当 4 个控制点不共面时,问题最多可能有 5 个解,且可以达到解的上限。对于 P5P 问题,当 5 个控制点中任意 3 点不共线时,则 P5P 问题最多可能有两个解,且可以达到解的上限。

2) 直线特征定位

目前,基于模型单目视觉定位的模型特征分为点、直线与高级几何特征等几类。相对来说,目前对于基于点特征的单目视觉定位方法研究较多,对于基于直线特征的单目视觉定位方法的研究还比较少。在某些特定的环境中,采用直线特征进行定位比采用点特征进行定位具有一定的优势。直线特征的优势表现在以下几方面:首先,自然环境的图像包含很多的直线特征。其次,在图像上直线特征比点特征的提取精度更高。最后,直线特征抗遮挡能力比较强。

同时相对于更高级的几何特征,直线特征也具有优势,具体表现在以下几方面:首先,在周围自然环境的图像中,直线比其他的高级几何特征更常见,同时也更容易提取。其次,直线的数学表达式更简单,处理起来效率更高。因此综合来看,在某些方面采用直线特征进行视觉定位具有其他特征所不具有的一些优势,在实现高精度、实时自主定位方面有着广阔

的应用前景。

对于空间恢复,至少需要非共线的3个特征点来获得唯一解。如果使用直线,则需要3条直线,3条直线不同时平行且不和光心共面。目前,理论上研究最多的是利用三线定位的问题,即 Perspective Projection of Three Lines,以下简称 P3L 问题。

3) 高级几何特征定位

高级几何特征包括圆、椭圆、二次曲线等。对于基于模型的单目视觉定位问题,很多学者都做了这方面的研究工作。通常,他们使用点或直线的投影作为基元,由3个图像点或3条图像直线以及在物体坐标下点或直线之间的相对位置关系确定模型的姿态。有时,基于模型单目视觉定位问题的模型采用曲线表面的物体,所以使用曲线进行曲线表面物体的定位成为另外一个活跃的研究方向。

使用曲线定位的好处是:首先,自然界许多物体的表面上有曲线特征;其次,曲线包含三维物体的全局位姿信息;最后,对曲线的表示是对称矩阵,因此以数学方法处理起来很方便。在很多情况下,可以获得闭式解,从而避免了非线性搜索。相比于其他两种特征,不足的地方是自然界中还是点特征和直线特征更普遍存在,具有广泛的实用性。

对于曲线表面的物体,一些学者提出了使用曲线进行定位的方法。当用曲线进行姿态估计时,一定要对复杂的非线性系统进行求解。Forsyth 等对于共面曲线提出了一种定位方法,这种方法是两个四次多项式进行求解。Ma Songde 提出,对于两个非共面曲线,它的姿态可以对由6个二次多项式组成的非线性系统进行求解得到;当两个空间曲线共面时,可以得到物体姿态的闭式解。

5.2.2 双目视觉模型

双目立体视觉是计算机视觉的一个重要分支,即由不同位置的两台或者一台摄像机(CCD)经过移动或旋转拍摄同一幅场景,通过计算空间点在两幅图像中的视差,获得该点的三维坐标值。采用高精度的标定模板、完善的摄像机标定数学模型,并对靶特征点进行子像素检测,保证系统的标定精度。

该系统能够对视场范围内的标靶进行自动识别定位,可在复杂的背景环境下实现系统的现场标定,操作方便快捷。系统对运动体上特征点进行实时检测,通过左右图像中特征点的图像坐标,和双目测量原理,实现特征点的三维空间坐标的测量。通过对运动体上特征点的识别定位并对数据进行分析进一步获取运动体的位置三维坐标、姿态、特征点之间的相对距离。

20世纪80年代,美国麻省理工学院人工智能实验室的 Marr 提出了一种视觉计算理论并应用在双睛匹配上,使两张有视差的平面图产生有深度的立体图形,奠定了双目立体视觉发展理论基础。相比其他类的立体视觉测量、跟踪方法,如透镜板三维成像、投影式三维显示、全息照相术等,立体视觉测量、跟踪方法直接模拟人类双眼处理景物的方式,可靠简便,在许多领域均极具应用价值,如微操作系统的位姿检测与控制、机器人导航与航测、三维测量学及虚拟现实等。

双目立体视觉测量系统主要功能包括:现场系统标定;空间特征点距离三维测量;空间物体位置三维测量;空间运动体姿态的双目三维测量;特征点的自动识别定位。

系统的主要技术特点是:操作界面清晰明了,简单易行,只需简单设定即可自动执行测

量;测量软件及算法完全自主开发,系统针对性强;可灵活设置测量模板、测量范围;安装简单,结构紧凑,易于操作、维护和扩充;可靠性高,运行稳定,适合各种现场运行条件;基于PC平台,系统可扩充性强,基于EF-VS机器视觉软件平台可扩展其他功能。

双目标立体视觉技术的实现可分为以下步骤:图像获取、摄像机标定、特征提取、图像匹配和三维重建。

5.3 机器人图像处理技术

5.3.1 图像处理技术概述

视觉是最高级的人类感觉,因此,图像在人类感知中起着最重要的作用。然而,人类的视觉被限制在电磁波谱的可视波段,而成像机器几乎覆盖了全部电磁波段。它们还可以在人类不常涉及的图像源上产生的图像上进行展示,包括超声波、电子显微镜和计算机产生的图像。这样,数字图像处理就包含了很宽的应用领域。图像处理的步骤包括以下几点。

(1) 摄像机捕获图像。

(2) 图像数字化。

(3) 图像用一个二维函数 $X(x, y)$ 进行描述,其中, x 为矩阵的行, y 为矩阵的列, x, y 的范围由图像的最大分辨率决定。如果图像的大小为 $n \times m$, 此时 n 为行数, 而 m 为列数, 其中, $0 \leq y < m$, 且 $0 \leq x < n$ 。 x 和 y 为正整数或者零。

数字图像是由有限数量的元素组成的,每个元素都有一个特殊的位置和数值。这些元素被称为画素或像素。像素是广泛应用于定义数字图像元素的术语。

图像处理和计算机视觉之间并没有明显的界限,但我们可通过考虑3种等级的计算机化过程加以区分:低级、中级和高级处理。低级处理包括原始操作,如降低噪声的图像预处理、对比度增强和图像锐化,低级处理的特点是其输入与输出均为图像。图像的中级处理涉及诸如分割这样的行为,即把图像分为区域或对象,然后对对象进行描述,以便使它们简化为适合计算机处理的形式,并对对象进行分类。中级处理的特点是输入通常为图像,单输出则是从这些图像中提取的属性。高级处理通过执行与人类视觉相似的感知函数,对对象进行总体确认。

图像是人类获取和交换信息的主要来源,因此,图像处理的应用领域必然涉及人类生活和工作的方方面面,包括航天和航空技术、生物医学工程、通信工程、工业和工程、军事和公安、文化与艺术等方面。随着人类活动范围的不断扩大,图像处理的应用领域也将随之扩大。

5.3.2 图像的采集

图像是多媒体作品中使用频繁的素材,除通过图像软件的绘制、修改获取图像外,使用最多的还是直接获取图像,主要有以下几种方法。

(1) 利用扫描仪和数码相机获取。扫描仪主要用来取得印刷品以及照片的图像,还可以借助识别软件进行文字的识别。目前市场上的扫描仪种类繁多,在多媒体制作中可以选择中高档类型。而数码相机可以直接产生景物的数字化图像,通过接口装置和专用软件完成图像输入计算机的工作。在使用数码相机之前,应有一些摄影知识准备,如用光、构图、色



视频讲解

彩学等,这样就能更好地利用设备进行操作。

(2) 从现有图片库中获取。多媒体电子出版物中有大量的图像素材资源。这些图像主要包括山水木石、花鸟鱼虫、动物世界、风土人情、边框水纹、墙纸图案、城市风光、科幻世界等,几乎应有尽有。另外,还要养成收集图像的习惯,将自己使用过的图像分类保存,形成自己的图片库,以便以后使用。

(3) 在屏幕中截取。在多媒体制作中,有时候可以将计算机显示屏幕上的部分画面作为图像。从屏幕上截取部分画面的过程称为屏幕抓图。具体方法是:在 Windows 环境下,单击键盘功能键中的 Print Screen 键,然后进入 Windows 附件中的画图程序,用粘贴的方法将剪贴板上的图像复制到“画纸”上,最后保存。

(4) 用豪杰超级解霸捕获 VCD 画面。在多媒体制作中,常常需要影片中的某一个图像,这时可以借助“超级解霸”来完成。

(5) 从网络上下载图片。网络上有很多图像素材,可以很好地利用它们。但使用时,应考虑图像文件的格式、大小等因素。

当利用 Python 对以上已经得到的图像进行处理时,首先需要读取和保存图像,Python 读取和保存图像的方法有很多,比如 OpenCV 库的 `imread()` 和 `imwrite()` 函数;Image 包中的 `Image.open()` 和 `Image.save()` 函数等。

```
# 图像读写
import cv2
# 读取图像
im = cv2.imread('cat.jpg')
h,w = im.shape[:2]
print(h,w)
# 保存图像
cv2.imwrite('cat1.png',im)
```

利用 `im.shape` 函数可以获取图像的尺寸。

5.3.3 图像的转换

在实际操作中经常会需要对图像格式和颜色空间等进行转换。我们可以利用一些专门的软件达到多媒体制作要求,例如图像处理工具、ACDSee、Photoshop 等。利用 Python 语言同样可以实现一些简单的操作,并且使得在机器人应用中更加灵活、方便、容易移植。

要进行图像格式转换,只需要在保存图像时,将图像格式选择为需要的格式即可。通常生活中常用的颜色空间有灰度图、RGB、HSV、Lab 等,在 Python 的 OpenCV 图像处理库中有超过 150 种进行颜色空间转换的方法,比较常用的就是 RGB↔GRAY 和 RGB↔HSV。

这里要用到的函数是: `cv2.cvtColor(input_image,flag)`,其中,flag 就是转换类型。对于 BGR↔Gray 的转换,要使用的 flag 就是 `cv2.COLOR_BGR2GRAY`。同样,对于 BGR↔HSV 的转换,使用的 flag 就是 `cv2.COLOR_BGR2HSV`。

```
import cv2
import numpy as np

img = cv2.imread('img.jpg')
# 转换为灰度图
img_gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
```



```
# 转换到 HSV 空间
hsv=cv2.cvtColor(img,cv2.COLOR_BGR2HSV)

cv2.imshow('img', img)
cv2.imshow('HSV', hsv)
cv2.imshow('gray', img_gray), cv2.waitKey(0)
```

上述代码的运行结果如图 5-1 所示。



图 5-1 运行结果图

除了转换颜色空间外,还可以利用 Python 转换图像的颜色模式。这里要用到的是 image 工具包的 `convert()` 函数, `image.convert()` 有一个参数表示要转换的模式代码。

```
from PIL import Image
image = Image.open("img.jpg")
image=image.convert("1")
image.show()
```



图 5-2 二值图像

上述代码的运行结果如图 5-2 所示。

不同图像的模式表示方法如下。

(1) 模式 1: 二值图像,即非黑即白,像素只有 0 和 1 组成,图像由无数黑白格子组成。

(2) 模式 L: 灰度图像,它的每个像素用 8 个 bit 位表示,其中,0 表示黑,255 表示白,其他数字表示不同的灰度。

(3) 模式 p: 8 位彩色模式,它的每一个像素用 8 位表示,其对应的色彩是按照调色板查询出来的。

(4) 模式 RGBA: 32 位彩色模式,其中,24 位分别表示红色、绿色、蓝色 3 个通道,另外 8 位表示 alpha 通道,即透明通道。

(5) 模式 CMYK: 32 位彩色图像,它的每个像素用 32 位表示。模式 CMYK 就是印刷四分色模式,它是彩色印刷时采用的一种套色模式,利用色料的三原色混色原理,加上黑色油墨,共计 4 种颜色混合叠加,形成所谓“全彩印刷”。4 种标准颜色是: C(Cyan)表示青色,又称为“天蓝色”或是“湛蓝”; M(Magenta)表示品红色,又称为“洋红色”; Y(Yellow)表示黄色; K(Key Plate)表示定位套版色(黑色)。

(6) 模式 YCbCr: 24 位彩色图像,它的每个像素用 24 位表示。在 YCbCr 中,Y 是指亮度分量,Cb 指蓝色色度分量,而 Cr 指红色色度分量。人的肉眼对视频的 Y 分量更敏感,因此在通过对色度分量进行子采样来减少色度分量后,肉眼将察觉不到的图像质量的变化。

(7) 模式 I: 32 位整型灰色图像,它的每个像素用 32 位表示,0 表示黑,255 表示白,(0,

255)之间的数字表示不同的灰度。

(8) 模式 F: 浮点灰色图像,它的每个像素用 32 位表示,0 表示黑,255 表示白,(0,255)之间的数字表示不同的灰度。

5.3.4 图像处理软件简介

图像处理软件可以对图像进行编辑、加工、处理,使图像成为合乎需要的文件。下面简单介绍几种常用的图像处理软件。

(1) Windows 画图。“画图”是 Windows 下的一个小型绘图软件,可以用它创建简单的图,或用“画图”程序查看和编辑扫描好的照片;可以用“画图”程序处理图片,例如,JPEG、GIF 或 BMP 文件,可以将“画图”图片粘贴到其他已有文档中,也可以将其用作桌面背景。

(2) Photoshop。Photoshop 是目前最流行的平面图像设计软件,它是针对位图图形进行操作的图像处理程序。它的工作主要是进行图像处理,而不是图形绘制。Photoshop 处理位图图像时,可以优化微小细节,进行显著改动,以及增强效果。

(3) PhotoDraw。PhotoDraw 是 Microsoft 公司推出的图像处理软件,它具有丰富的功能及良好的易用性,且与 Office 及 Web 页面可以无缝地连接和嵌入。

(4) CorelDRAW。CorelDRAW 图像软件套装是一套屡获殊荣的图形、图像编辑软件,它包含两个绘图应用程序:一个用于矢量图及页面设计;另一个用于图像编辑。这套绘图软件组合带来了强大的交互式工具,操作简单并能保证高质量的输出性能。

根据三色原理,各种颜色的光都可以由红、绿、蓝 3 种基色光加权混合而成,因此,彩色空间是三维的线性空间,任何一种具有一定亮度的颜色光即为空间中的一个点。所以可以选择具有确定光通量的红绿蓝三种基色光作为该三维空间的基,这样组成的表色系统称为 RGB 表色系统。

人类的视觉处理也是通过对颜色的处理来达到的,但是在某些情况下人类是通过直接获取特别的色彩来解决问题。物体的颜色模型也能够用于计算机视觉中。这些方面的知识能够解决很多问题。例如,使用肤色模型来检测人脸的机器视觉应用。但是,这些应用也存在不确定性。如果背景中混杂的有与人类肤色相近的物体,那么就更容易检测失败。

5.3.5 图像处理技术的研究内容

图像处理技术是用计算机对图像信息进行处理的技术。主要包括图像数字化、图像增强和复原、图像数据编码、图像分割和图像识别等。

利用 Python 语言来进行图像处理,非常方便并且简单。接下来主要介绍图像增强和图像分割。

1. 图像增强

图像增强的目的是改善图像的视觉效果,它是各种技术的汇集,还没有形成一套通用的理论。常用的图像增强技术有对比度处理、直方图修正、噪声处理、边缘增强、变换处理和伪彩色等。在多媒体应用中,对各类图像主要进行图像增强处理,各类图像处理软件一般都支持图像增强技术。

图像增强主要解决由于图像的灰度级范围较小造成的对比度较低的问题,目的就是 will 输出图像的灰度级放大到指定的程度,使得图像中的细节看起来增加清晰。对比度增强有

几种常用的方法,如线性变换、分段线性变换、伽马变换、直方图正规化、直方图均衡化、局部自适应直方图均衡化等。

1) 线性变换

```
# 线性变换
import cv2 as cv
import numpy as np
import matplotlib.pyplot as plt
# 绘制直方图函数
def grayHist(img):
    h, w = img.shape[:2]
    pixelSequence = img.reshape([h * w, ])
    numberBins = 256
    histogram, bins, patch = plt.hist(pixelSequence, numberBins,
                                      facecolor='black', histtype='bar')

    plt.xlabel("gray label")
    plt.ylabel("number of pixels")
    plt.axis([0, 255, 0, np.max(histogram)])
    plt.show()

img = cv.imread("img.jpg", 0)
out = 2.0 * img
# 进行数据截断,大于 255 的值截断为 255
out[out > 255] = 255
# 数据类型转换
out = np.around(out)
out = out.astype(np.uint8)
# 分别绘制处理前后的直方图
# grayHist(img)
# grayHist(out)
cv.imshow("img", img)
cv.imshow("out", out)
cv.waitKey()
```

上述代码的运行结果如图 5-3 所示。

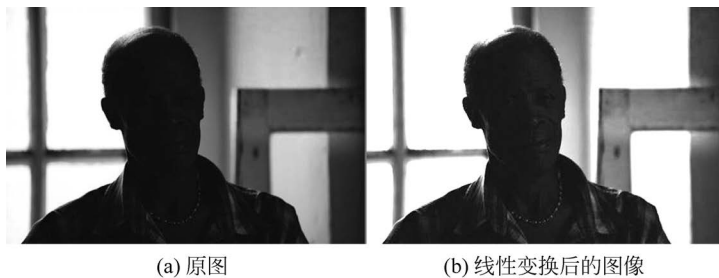


图 5-3 线性变换效果图

2) 伽马变换

```
# 伽马变换
img = cv.imread("1.jpg", 0)
# 图像归一化
fi = img / 255.0
# 伽马变换
gamma = 0.4
out = np.power(fi, gamma)
```

```
cv.imshow("img", img)
cv.imshow("out", out)
cv.waitKey()
```

上述代码的运行结果如图 5-4 所示。



图 5-4 伽马变换效果图

2. 图像分割

图像分割就是把图像分成若干特定的、具有独特性质的区域并提出感兴趣目标的技术和过程。它是由图像处理到图像分析的关键步骤。现有的图像分割方法主要分以下几类：基于阈值的分割方法、基于区域的分割方法、基于边缘的分割方法以及基于特定理论的分割方法等。从数学角度来看，图像分割是将数字图像划分成互不相交的区域的过程。图像分割的过程也是一个标记过程，即把属于同一区域的像素赋予相同的编号。

全阈值分割值是将灰度阈值大于 `thresh`(阈值) 的像素设置为白色，小于或等于阈值的像素设置为黑色；或者反过来，将大于阈值的像素设置为黑色，小于阈值的像素设置为白色，两者的区别只在于呈现的形式不同。

```
import cv2
image = cv2.imread('img.jpg', cv2.IMREAD_GRAYSCALE)
the = 140 # 设置阈值为 140
maxval = 255
dst, img = cv2.threshold(image, the, maxval, cv2.THRESH_BINARY)
cv2.imshow('image', image)
cv2.imshow('hand_thresh', img)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

上述代码的运行结果如图 5-5 所示。

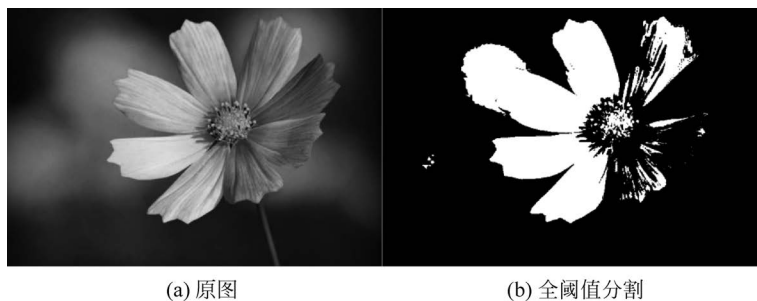


图 5-5 全阈值分割效果图

K-means 算法是最为经典的基于划分的聚类方法,是十大经典数据挖掘算法之一。K-means 算法的基本思想是:以空间中 k 个点为中心进行聚类,对最靠近它们的对象归类。通过迭代的方法,逐次更新各聚类中心的值,直至得到最好的聚类结果。

```
import cv2
import matplotlib.pyplot as plt
import numpy as np

def seg_kmeans_color():
    img = cv2.imread('cat.jpg', cv2.IMREAD_COLOR)
    # 变换一下图像通道 bgr->rgb, 否则很别扭啊
    b, g, r = cv2.split(img)
    img = cv2.merge([r, g, b])
    # 3 个通道展平
    img_flat = img.reshape((img.shape[0] * img.shape[1], 3))
    img_flat = np.float32(img_flat)
    # 迭代参数
    criteria = (cv2.TERM_CRITERIA_EPS + cv2.TermCriteria_MAX_ITER, 20, 0.5)
    flags = cv2.KMEANS_RANDOM_CENTERS
    # 聚类
    compactness, labels, centers = cv2.kmeans(img_flat, 2, None, criteria, 10, flags)
    # 显示结果
    img_output = labels.reshape((img.shape[0], img.shape[1]))
    plt.subplot(121), plt.imshow(img), plt.title('input')
    plt.subplot(122), plt.imshow(img_output, 'gray'), plt.title('kmeans')
    plt.show()

if __name__ == '__main__':
    seg_kmeans_color()
```

上述代码的运行结果如图 5-6 所示。

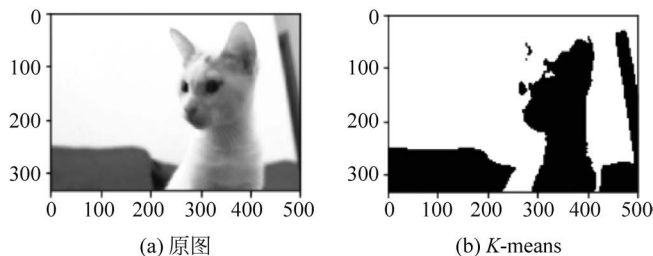


图 5-6 K-means 算法效果图

5.4 滤波算法概述

滤波是将信号中特定波段频率滤除掉的,是抑制和防止干扰的一种重要措施。滤波一词起源于通信理论,它是从含有干扰的接收信号中提取有用信号的一种技术。例如,用雷达跟踪飞机,在测得的飞机位置的数据中,含有测量误差及其他随机干扰,如何利用这些数据尽可能准确地估计出飞机在每一时刻的位置、速度、加速度等,并预测飞机未来的位置,就是一个滤波与预测问题。历史上最早考虑的是维纳滤波,后来 R. E. 卡尔曼和 R. S. 布西于 20 世纪 60 年代提出了卡尔曼滤波。现在对一般的非线性滤波问题的研究相当活跃。

5.4.1 常见的滤波方法

滤波电路常用于滤去整流输出电压中的纹波,一般由电抗元件组成,如在负载电阻两端并联电容 C ,或与负载串联电感 L ,以及由电容、电感组合而成的各种复式滤波电路。

滤波是信号处理中的一个重要概念。滤波分经典滤波和现代滤波。

经典滤波是根据傅里叶分析和变换提出的一个工程概念。根据高等数学理论,任何一个满足一定条件的信号,都可以被看成是由无限个正弦波叠加而成。换句话说,就是工程信号是不同频率的正弦波线性叠加而成的,组成信号的不同频率的正弦波叫作信号的频率成分或叫作谐波成分。只允许一定频率范围内的信号成分正常通过,而阻止另一部分频率成分通过的电路叫作经典滤波器或滤波电路。

5.4.2 滤波电路的分类

常用的滤波电路有无源滤波和有源滤波两大类。若滤波电路元件仅由无源元件(电阻、电容、电感)组成,则称为无源滤波电路。无源滤波的主要形式有电容滤波、电感滤波和复式滤波(包括倒 L 型、 LC 滤波、 $LC\pi$ 型滤波和 $RC\pi$ 型滤波等)。若滤波电路中不仅包括无源元件,还包括有源元件(双极型管、单极型管、集成运放),则称为有源滤波电路。有源滤波的主要形式是有源 RC 滤波,也被称作电子滤波器。

1. 无源滤波电路

无源滤波电路的结构简单,易于设计,但它的通带放大倍数及其截止频率都随负载而变化,因而不适用于信号处理要求高的场合。无源滤波电路通常用在功率电路中,比如直流电源整流后的滤波,或者大电流负载时采用 LC (电感、电容)电路滤波。

2. 有源滤波电路

有源滤波电路的负载不影响滤波特性,因此常用于信号处理要求高的场合。有源滤波电路一般由 RC 网络和集成运放组成,因而必须在有合适的直流电源供电的情况下才能使用,同时还可以进行放大。但电路的组成和设计也较复杂。有源滤波电路不适用于高电压大电流的场合,只适用于信号处理。

由滤波器的特点可知,它的电压放大倍数的幅频特性可以准确地描述该电路属于低通、高通、带通还是带阻滤波器,因而如果能定性分析出通带和阻带在哪个频段,就可以确定滤波器的类型。

识别滤波器的方法是:若信号频率趋于零时有确定的电压放大倍数,且信号频率趋于无穷大时电压放大倍数趋于零,则为低通滤波器;反之,若信号频率趋于无穷大时有确定的电压放大倍数,且信号频率趋于零时电压放大倍数趋于零,则为高通滤波器;若信号频率趋于零和无穷大时电压放大倍数均趋于零,则为带通滤波器;反之,若信号频率趋于零和无穷大时电压放大倍数具有相同的确定值,且在某一频率范围内电压放大倍数趋于零,则为带阻滤波器。

5.4.3 图像滤波处理

由于成像系统、传输介质和记录设备等的不完善,数字图像在其形成、传输和记录过程中往往会受到多种噪声的污染。另外,在图像处理的某些环节当输入的像对象并不如预想时也会在结果图像中引入噪声。

图像滤波是指在尽量保留图像细节特征的条件下对目标图像的噪声进行抑制。Python 中常用 Pillow 库提供的 ImageFilter 包对图像进行滤镜处理,其中包含模糊、平滑、锐化、边界增强等滤镜效果的处理。

常用的 Python 滤镜主要有 BLUR、CONTOUR、DETAIL、EDGE_ENHANCE、EDGE_ENHANCE_MORE、EMBOSS、FIND_EDGES、SMOOTH、SMOOTH_MORE、SHARPEN。

常用 Python 滤镜的滤波名称和含义如表 5-1 所示。

表 5-1 常用 Python 滤镜的滤波名称和含义

滤波名称	含 义
ImageFilter. BLUR	模糊滤波
ImageFilter. SHARPEN	锐化滤波
ImageFilter. SMOOTH	平滑滤波
ImageFilter. SMOOTH_MORE	平滑滤波(阈值更大)
ImageFilter. EMBOSS	浮雕滤波
ImageFilter. FIND_EDGES	边界滤波
ImageFilter. EDGE_ENHANCE	边界滤波
ImageFilter. EDGE_ENHANCE_MORE	边界加强(阈值更大)
ImageFilter. CONTOUR	轮廓

利用 ImageFilter 包中的 filter() 函数对图像进行滤波处理,该函数有一个参数,可以选择表 5-1 中的滤波名称来实现对应的效果。

```
from PIL import Image, ImageFilter
image = Image.open('img.jpg')
# 打开指定路径的 jpg 图像文件
image = image.filter(ImageFilter. BLUR)
# 模糊
image.show()
# 显示图片
```

上述代码的运行结果如图 5-7 所示。

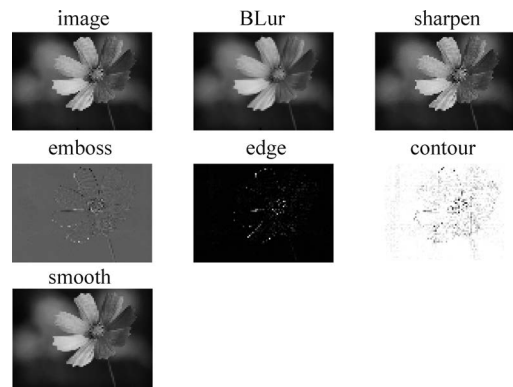


图 5-7 常用的 Python 滤镜

5.4.4 智能机器人复合滤波算法的实验研究实例

令 X_i 作为随机采样信号的统计量,共有 20 组统计量 $(X_k, k = 1, 2, \dots, n)$ 表示随机信号样本点集, S_k 为采样信号中的有用部分, W_k 为采样信号的随机误差, k 是每一组随机样

本点集包含的个数。剩余误差作为最后输出的判断基准。由于这组数已经按序排列,所以其中中位数值即为 $X_{n/2}$ 。下面分析其滤波结果。

(1) 求 n 次测量值 $X_1 \sim X_n$ 算术平均值 $\bar{X}$

$$X_k = S_k + W_k \quad (k = 1, 2, \dots, n) \quad (5-1)$$

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k \approx \frac{1}{n} \sum_{k=1}^n S_k \quad (k = 1, 2, \dots, n) \quad (5-2)$$

(2) 求各项的剩余误差

$$\Delta X_i = |X_k - \bar{X}|$$

$$\Delta X_i^2 = [(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_k - \bar{X})^2] \quad (k = 1, 2, \dots, n) \quad (5-3)$$

(3) 计算随机向量的标准偏差

$$\sigma = \sqrt{\frac{\left(\sum_{k=1}^n \Delta X_k^2\right)}{(n-1)}} \quad (5-4)$$

若本次采样结果为 X_k , 则本次复合滤波的结果由下式确定

$$\Delta X_i = |X_k - \bar{X}| \begin{cases} > 3\sigma, & X_i = \bar{X} \\ \leq 3\sigma, & X_i = X_{n/2} \end{cases} \quad (5-5)$$

机器人软件系统设计主要包括主程序设计, 该软件系统主程序的编写, 提高了传感器信号的判断和有效检出的准确程度。只有返回准确的数据, 机器人才能对障碍物做及时正确的反馈处理, 传感器的信息融合及应用才能得以实现。主程序如下:

```
// 去除最大最小极值后求平均
for(i = 0; i < 10; i++) d_sum += di;
average = d_sum / 10;
for(i = 0; i < 10; i++)
{
Standard deviation = standard
deviation + (di - average) * (di - average);
}
// 求该组数据的标准差
standard deviation /= 14;
// Serial.println(standard deviation);
if(the residual error <= 3 * standard deviation)
// 要求剩余误差小于一定门限, 否则返回均值
return (xi);
else
return(average);
}
```

实验是在静态无隔板和动态有隔板两种情况下交叉进行的。连接线路, 将智能机器人置于实验平台, 用一系列隔板模拟实验时的室内障碍物, 它的位置信息将障碍物归类成不同的状态, 隔板放置距离每次只做微调, 测试距离小于 50cm。根据已定的采样标准进行实验, 重复进行 3 组, 最后取 3 组测量数据的平均值作为测量结果, 如图 5-8 所示。

观察图 5-8 所示的具体实验效果可以看出, 算法中在系列 3 对于静态、动态隔板环境下

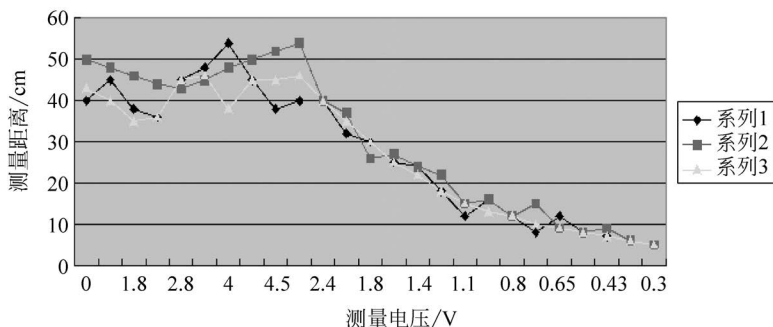


图 5-8 静态及动态隔板下电压与距离转换曲线图

距离测量的稳定性都比较好,为后续的障碍物目标躲避工作创造了条件。模拟室内动态与静态障碍物共存复杂环境中,在得到的电压-距离波形中系列 1 和系列 2 的测量距离值出现很大波动,甚至都超过设定的极限值 50cm,对环境描述精度偏低。

对比实验结果表明,算法在复杂环境下能从不同环境下准确获得数据,得到稳定的距离信号,送给控制器。控制器通过输入内部的算法,协调机器人工作,从而完成躲避障碍物动作。

5.5 基于摄像头路径识别的智能机器人控制系统设计

本节介绍智能机器人路径识别及速度控制系统的具体实现方法。在该系统中,由 CCD 摄像头实现路径识别,直流电机作为驱动,编码器检测速度,由 PID 算法实现小车的路径跟踪和速度的闭环控制。

5.5.1 硬件结构与方案设计

智能车控制系统以高性能的 16 位单片机 MC9S12DP256 为核心控制器,主要由电源管理、CCD 摄像头、图像采集模块、电机及其控制器、转向舵机及其控制器、上位机调试等功能模块组成,其中,上位机调试模块通过 RS232 串行接口与 PC 通信,结合基于 MATLAB 环境开发的应用软件实现在线综合调试、分析功能。系统总体结构如图 5-9 所示。

5.5.2 控制策略

CCD 摄像头每帧有 320×240 个像素点,每帧的前 20 行为场起始信号,后 10 行为场消隐信号,每行前 14 个数据为行起始信号,第 310~320 点为行消隐信号,摄像头每场信号有 320 行,其中,第 23~310 行为视频信号。从行起始信号到行消隐信号之间的时间间隔只有 $52\mu\text{s}$,在 24MHz 的总线时钟下,MC9S12DP256 最快的 AD 采样频率为 0.5MHz,那么每行最多只能采集到 26 个点,为了使采集的像素点尽量多,将单片机超频至 32MHz,这样每行最多可采集 34 个像素点。主要问题是现场光照环境不稳定,路径褪色以及路面光照反射能力的差别(路面不平整、路面上的污迹)造成路径图像灰度不均匀。本系统路径识别的目的是检测小车行驶方向的偏差。因此,路径识别的任务主要是将摄像头拍摄的灰度图像转换成二值图像,其中,路径标线以黑色表示,背景及其他图像内容用白色表示,以便下一步进行

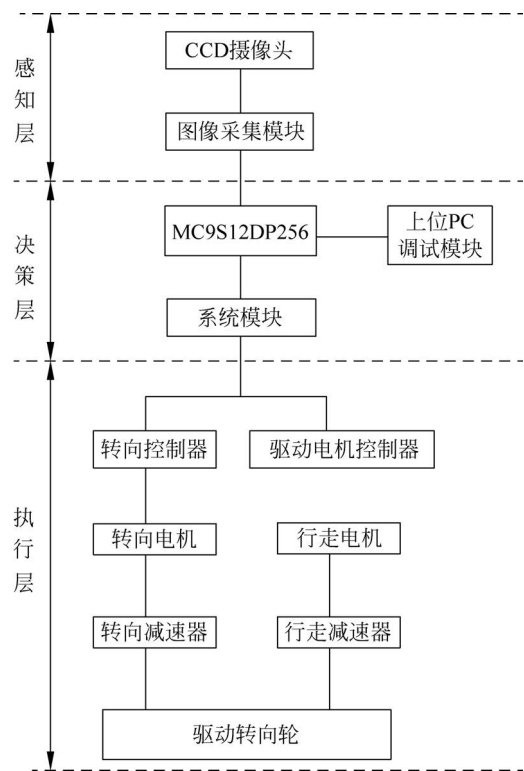


图 5-9 系统总体结构图

小车行驶测量,小车驱动电机电路如图 5-10 所示。

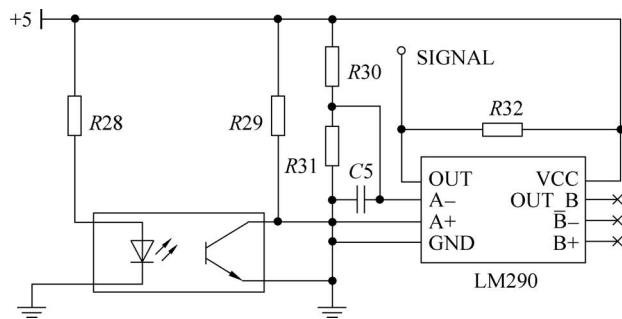


图 5-10 驱动电机电路

5.5.3 舵机转向控制分析以及驱动电机转速调节

智能车控制系统通过舵机输出转角驱动前轮转向,舵机自身为独立的位置闭环控制系统,在负载力矩小于最大输出力矩的情况下,舵机输出转角与控制信号脉宽成线性比例关系。在本控制系统中预先标定导引线位置与控制 PWM 信号的二维映射表。在电动小车自动行驶过程中,系统实时采集导引线的位置信息,通过查表给出当前 PWM 控制信号,调整舵机转向角度实现方向控制。为增强电动小车在不同目标速度下过弯时的稳定性和平顺性,可根据当前目标车速状态对输出的转向 PWM 控制信号进行调节。

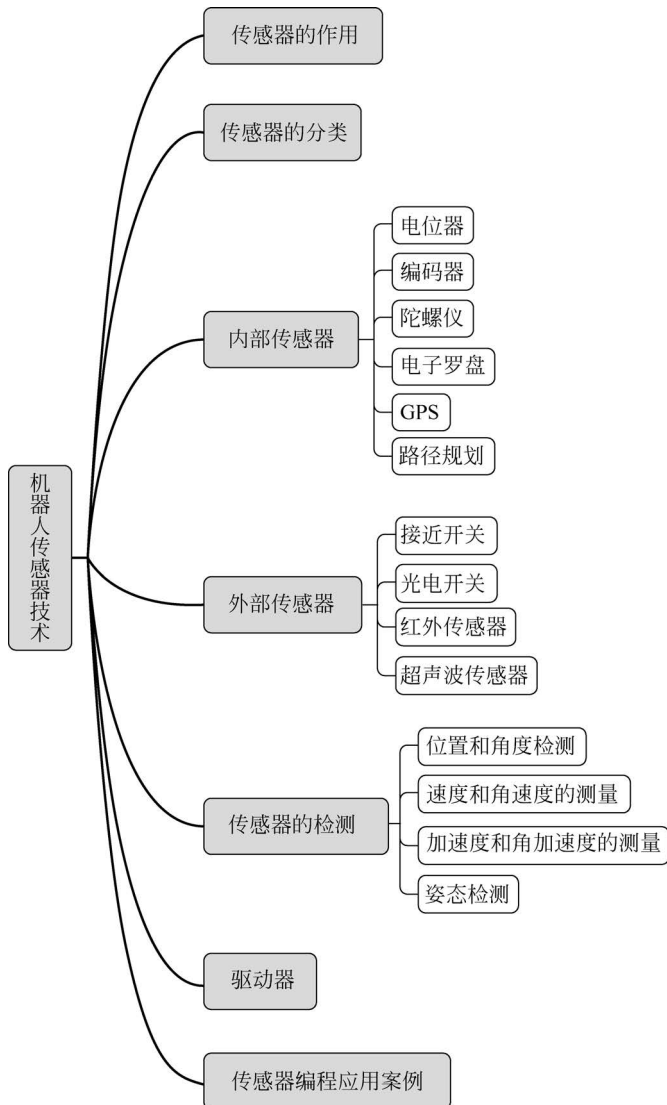
本控制系统中以驱动电机转速采样信息为反馈量,采用增量式数字 PID 控制算法,通

过输出 PWM 信号对电机实现闭环控制,为降低转速采样过程中的信号干扰,增强系统的稳定性,在控制软件中采用移动平均滤波法对转速采样信号进行了处理。

摄像头能获取的信息容量大、路径的设置和变更简单方便、系统柔性好,并具备性价比高、算法简便、实时性强等诸多优点,在路径信息分析过程中,采用阈值分割的优化算法,较好地解决了位置信息阶跃的问题,实现了连续路径识别功能。它具有现实应用的可能和广阔的应用前景,有望成为当前智能车辆导引技术研究的主流方向和发展趋势。

思考题

1. 什么是机器视觉定义?
2. Eye-in Hand 式手眼系统与 Eye-to Hand 式手眼系统的区别是什么?
3. 典型的机器人视觉系统一般包括哪些基本构成?
4. 机器人视觉要求包括: _____、_____、_____。
5. 图像处理技术包括哪些内容?



第 6 章思维导图