

双目立体视觉的目标识别与定位

尚倩^{1,2}, 阮秋琦^{1,2}, 李小利^{1,2}

(1. 北京交通大学 信息科学研究所, 北京 100044; 2. 北京交通大学 现代信息科学与网络技术北京市重点实验室, 北京 100044)

摘要: 双目立体视觉系统可以实现对目标的识别与定位. 此系统包含摄像机标定、图像分割、立体匹配和三维测距4个模块. 在摄像机标定部分, 提出了基于云台转角的外参数估计方法. 该方法可以精确完成摄像头旋转情况下外参数的估计, 增强了机器人的视觉功能. 并利用广茂达机器人系统, 基于改进的双目视觉系统进行目标识别与定位, 以此结果作为依据控制机器人的手臂进行相应运动, 最终实现了对目标物体的抓取, 验证了提出方法的可行性.

关键词: 双目立体视觉; 摄像机标定; 图像分割; 立体匹配; 三维测距

中图分类号: TP18 **文献标识码:** A **文章编号:** 1673-4785(2011)04-0303-09

Target recognition and location based on binocular stereo vision

SHANG Qian^{1,2}, RUAN Qiuqi^{1,2}, LI Xiaoli^{1,2}

(1. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China; 2. Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China)

Abstract: A binocular stereo vision system can be used for target recognition and location. This system consists of four modules, namely camera calibration, image segment, stereo matching, and binocular distance. furthermore, in the calibration section, a method of estimating the outer parameters based on the head angle was studied. In the experiment, this method could estimate the outer parameters accurately when the camera rotated, enhancing the vision function of the system. Then this improved binocular system was realized by the Grandar Robot platform. It was used for target recognition and location, and the results were used to control the robot motion. In the end, the target could be captured successfully, so the proposed method was proven to be feasible.

Keywords: binocular stereo vision; camera calibration; image segmentation; stereo matching; binocular distance

人类获取的信息 70% ~ 80% 来自视觉, 视觉是人类认识和改造世界的一个主要途径. 计算机视觉是使用计算机及相关设备对生物视觉的一种模拟, 它的主要任务就是通过对采集的图片或视频进行处理以获得相应场景的三维信息^[1].

根据视觉传感器的个数, 计算机视觉可以分为: 单目视觉系统、双目视觉系统和多目视觉系统^[1]. 单目视觉系统只通过一个视觉传感器来获取外界的视觉信息, 方式简单, 但是只能获得目标的二维空间信息. 而双目和多目视觉系统比较复杂, 一般具备获取三维空间信息的能力, 但实时性是双目和多目视觉系统的最大挑战.

双目立体视觉技术可以很好地模拟人眼的功能, 并且可以对三维世界进行立体感知, 因此, 在机器视觉的研究中发挥着越来越重要的作用. 文中着重研究双目立体视觉下目标的识别与定位技术, 双目立体视觉系统主要包括摄像机标定、图像分割、立体匹配和三维测距4个模块. 传统的摄像机标定技术一旦完成标定, 摄像机相对于世界坐标系的位置就不能发生变化, 否则, 就需要重新标定^[2]. 但从实际应用角度, 为了更好地进行目标的识别与定位, 摄像头需要进行相应偏转. 针对这个特点, 提出了一种基于云台转角的外参数估计方法, 经过误差分析, 此方法可以精确地完成摄像机旋转情况下对外参数的估计, 使机器人的云台在目标识别和定位过程中可以更加灵活地旋转, 大大增强了该机器人的视觉功能, 最后, 利用广茂达机器人系统作为实验平台, 基于改

收稿日期: 2010-12-09.

基金项目: 国家自然科学基金资助项目(60973060); 高等学校博士学科点专项科研基金资助项目(20080004001).

通信作者: 尚倩. E-mail: renxuefeier@gmail.com.

进后的双目立体视觉系统进行目标的识别与定位,并以此对目标识别定位的结果作为依据来控制机器人和手臂的相应运动,从而成功地实现了机器人对目标物体的准确抓取。

1 双目立体视觉系统

此系统的主要功能是完成对实际目标的识别与定位,主要分为摄像机标定、图像分割、立体匹配、三维重建4个模块。

1.1 摄像机标定

1.1.1 摄像机模型

摄像机采集到的图像是典型的二维信息,要想根据这个二维信息计算实际物体的三维信息,首先要建立图像坐标系与世界坐标系的映射模型。本文使用的摄像机模型为经典针孔模型,如图1所示^[34]。其中, $X_w Y_w Z_w$ 为世界坐标系, $X_c Y_c Z_c$ 为摄像机坐标系, xy 为图像物理坐标系, uv 为图像像素坐标系。

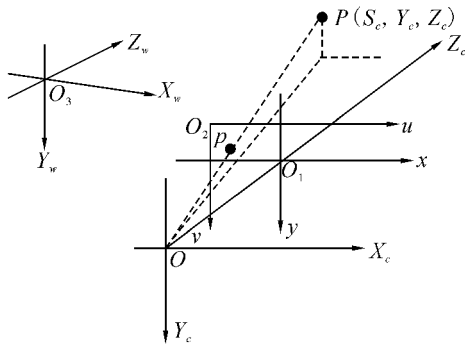


图1 摄像机针孔模型

Fig. 1 Camera pinhole model

上述模型存在如下的投影关系^[2,5]:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}.$$

式中: f_x, f_y, u_0, v_0 只与摄像机内部结构有关,所以 M_1 称为摄像机内部参数; R 与 t 完全由摄像机坐标系相对世界坐标系的方位决定,故称为摄像机外部参数。摄像机标定就是确定一部摄像机的内部参数和外部参数的过程^[2]。

1.1.2 摄像机标定

张正友等人提出了平面法来标定单摄像头^[2],本文利用此方法来标定双目立体视觉平台,标定物选择了标准棋盘格。为了简化计算,假定棋盘格所在的平

面为世界坐标系的 $Z_w=0$ 的平面。这样,对于棋盘格平面上的每一个点,上述投影关系可以简化为

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}. \quad (1)$$

令

$$M = [X_w \ Y_w \ 1]^T, m = [u \ v \ 1]^T,$$

$$H = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} = [h_1 \ h_2 \ h_3] =$$

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}.$$

则式(1)可以写成 $sm=HM$,其中 A 为摄像机内部参数, s 为尺度因子, H 称为单应性或投影矩阵,是模板平面上点和它的像点之间的一个映射,如果已知模板点的空间坐标和图像坐标,那么就已知 m 和 M ,可以求解单应性矩阵。利用这个矩阵就可以求解摄像机的内外参数。其中,内参可以通过Choleski分解得到,而利用矩阵 A 和 H ,就可以进一步求得外参矩阵^[2,5],公式如式(2):

$$\begin{cases} r_1 = \lambda A^{-1} h_1, r_2 = \lambda A^{-1} h_2, r_3 = r_1 \times r_2, \\ t = \lambda A^{-1} h_3, \lambda = \frac{1}{\|A^{-1} h_1\|} = \frac{1}{\|A^{-1} h_2\|}. \end{cases} \quad (2)$$

由于摄像机物理因素的影响,拍摄到的图片有一定的畸变,畸变会对摄像机的标定结果产生一定的影响。采用张正友的方法对畸变进行矫正^[2]。

1.1.3 算法改进

虽然张正友的方法可以精确地求得摄像机的内外参数,但是,参数一旦标定,摄像机坐标系相对于世界坐标系的位置关系也就静止。而在本实验平台中,摄像头会发生以下2种运动。

1) 摄像头会随着机器人的运动而运动。

2) 机器人不动时,摄像头也可以通过云台发生上下或左右的旋转。

上述2种运动都会导致摄像头的位置发生变化,理论上就需要进行重新标定,这在实验中不具备可行性。因此,摄像机的内参采用张正友的标定方法获得,而对于外参,针对本实验平台的特点,在其算法基础上进行了改进,使得当摄像头位置移动时,不用重新标定就可以计算出移动后的外参。改进方法归纳如下。

1) 世界坐标系转化为机器人坐标系。

在标定算法中,世界坐标系一旦选定了就是个绝对量,然而,当机器人移动时,如果仍然使用原

来标定的外参进行重建,就意味着认为世界坐标系也随机器人发生了完全相同的移动,而在这个过程中,二者的转换关系并没有发生变化.因此,可以选择一个相对不变的世界坐标系,即使机器人发生了移动,仍然可以通过固定的转换关系转换到这个坐标系,即机器人坐标系.实验过程中,机器人坐标系的原点选在2台摄像机云台的中点,三维测距后,将结果由世界坐标系转换到机器人坐标系,因为不关心目标物体在世界坐标系中的位置,只关心它与机器人的相对位置.

因此,首先需要知道原始标定时世界坐标系与机器人坐标系的转换关系.原始标定环境中,两摄像头平行地指向前方,即二者的云台旋转角度为0.在外参的标定结果中,假设左右摄像头的平移向量分别为 $t_1 = (t_{x1}, t_{y1}, t_{z1})$, $t_2 = (t_{x2}, t_{y2}, t_{z2})$,则左右摄像机坐标系原点在原世界坐标系中的坐标分别为 $(-t_{x1}, -t_{y1}, -t_{z1})$ 和 $(-t_{x2}, -t_{y2}, -t_{z2})$,所以机器人坐标系原点在原世界坐标系下的坐标为 $(-\frac{t_{x1} + t_{x2}}{2}, -\frac{t_{y1} + t_{y2}}{2}, -\frac{t_{z1} + t_{z2}}{2})$,由于标定时,标定物与机器人平面平行,所以旋转向量近似认为是0.

假设一个点在世界坐标系下的坐标为 (x_w, y_w, z_w) ,而其在机器人坐标系下的坐标为 (x_r, y_r, z_r) ,则二者的转换关系如下:

$$\begin{cases} x_r = x_w + \frac{t_{x1} + t_{x2}}{2}, \\ y_r = y_w + \frac{t_{y1} + t_{y2}}{2}, \\ z_r = z_w + \frac{t_{z1} + t_{z2}}{2}. \end{cases} \quad (3)$$

通过式(3),就把世界坐标系转换到了机器人坐标系.

2) 基于云台转角的外参数估计.

上述改进虽然解决了机器人运动时外参的使用问题,可是不允许摄像头有上下左右的旋转,这样就使机器人的等效视界大大减小,在一定程度上限制了机器人的视觉功能.由于智能机器人的主要思路是模仿人的眼睛,所以希望摄像头可以根据目标的位置,模拟人的双眼做相应地旋转,从而实现更加灵活的三维测距.

在三维空间中,旋转可以分解为绕各自坐标轴的二维旋转,实验用的云台有2个转动轴,可以进行旋转角度可控可知的左右转动和上下转动.设左右转动的角度为 θ_1 ,上下转动的角度为 θ_2 ,则左右旋转和上下旋转的旋转矩阵分别为:

$$\begin{aligned} R_1 &= \begin{bmatrix} \cos \theta_1 & 0 & -\sin \theta_1 \\ 0 & 1 & 0 \\ \sin \theta_1 & 0 & \cos \theta_1 \end{bmatrix}, \\ R_2 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_2 & \sin \theta_2 \\ 0 & -\sin \theta_2 & \cos \theta_2 \end{bmatrix}. \end{aligned}$$

绕2个轴均有转动的旋转矩阵为

$$R = R_1 R_2 = \begin{bmatrix} \cos \theta_1 & \sin \theta_1 \sin \theta_2 & -\sin \theta_1 \cos \theta_2 \\ 0 & \cos \theta_2 & \sin \theta_2 \\ \sin \theta_1 & -\cos \theta_1 \sin \theta_2 & \cos \theta_1 \cos \theta_2 \end{bmatrix}.$$

摄像机发生旋转后,设原来的摄像机坐标系为 $X_{c1}Y_{c1}Z_{c1}$,而旋转后的摄像机坐标系为 $X_{c2}Y_{c2}Z_{c2}$,则2个摄像机坐标系之间的关系为

$$\begin{bmatrix} X_{c2} \\ Y_{c2} \\ Z_{c2} \\ 1 \end{bmatrix} = \begin{bmatrix} R_c & T_c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{c1} \\ Y_{c1} \\ Z_{c1} \\ 1 \end{bmatrix}.$$

式中: R_c 为上面求得的旋转矩阵,由于摄像头只有旋转没有平移,故 $T_c = [0 \ 0 \ 0]^T$.

由于原来的摄像机坐标系与世界坐标系的关系为

$$\begin{bmatrix} X_{c1} \\ Y_{c1} \\ Z_{c1} \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}.$$

故

$$\begin{bmatrix} X_{c2} \\ Y_{c2} \\ Z_{c2} \\ 1 \end{bmatrix} = \begin{bmatrix} R_c & T_c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{c1} \\ Y_{c1} \\ Z_{c1} \\ 1 \end{bmatrix} =$$

$$\begin{bmatrix} R_c & T_c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (4)$$

从式(4)可推导出, $M = \begin{bmatrix} R_c & T_c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$,即为

摄像头旋转后的外参.这样,只要知道云台的转角,就可以求得任意旋转角度下的外参.这一方法成功地解决了摄像头任意旋转时外参的计算问题.

1.2 基于颜色相似度的图像分割

利用对目标图像的分割来显示目标识别的结果.基于颜色相似度的图像分割可以满足鲁棒性和实时性的要求,且具有良好的分割效果.

1.2.1 颜色相似度

从摄像机输入的彩色图像大多数是以 RGB 位图格式存储的,在 RGB 空间中,2 种颜色之间的距离与人眼对色彩差别的感觉有较大偏差.而在 HSV 空间中,颜色分量相应的欧式距离与人眼感知的颜色差吻合得较好. HSV 空间中的 2 种颜色 $C_1 = \{H_1, S_1, V_1\}$ 和 $C_2 = \{H_2, S_2, V_2\}$ 之间的距离^[1] 定义为

$$D(C_1, C_2) =$$

$$\sqrt{(S_1 \cos H_1 - S_2 \cos H_2)^2 + (S_1 \sin H_1 - S_2 \sin H_2)^2 + (V_1 - V_2)^2}.$$

两色彩之间的相似度^[8] 定义为

$$S(C_1, C_2) = 1 - \frac{D(C_1, C_2)}{\sqrt{5}}.$$

1.2.2 基于颜色相似度的图像分割

分割的目的是要从复杂的背景中提取出具有特定颜色特征的目标物体,具体实现步骤如下.

1) 颜色信息的统计及相似度计算.

设摄像头拍摄到的图片大小为 $m \times n$,首先计算图像中像素点颜色信息的均值,得到该图像的颜色模板 $C_1 = \{H_1, S_1, V_1\}$ ^[3,6-8]. 然后计算图像中每一像素点和标准颜色模板的颜色距离和相似度. 其中:

$$H_1 = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n h_{ij},$$

$$S_1 = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n s_{ij},$$

$$V_1 = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n v_{ij}.$$

2) 图像分割.

通过阈值的设置,可以完成对图像的分割. 计算如式(5)所示. 影响分割效果的主要因素是阈值的选取,找到一个合适的阈值是图像分割最关键的一步^[8]. 本实验中,阈值取 0.76 ~ 0.89 能够得到比较好的分割效果.

$$\begin{cases} P_{ij} = (0, 0, 0), & S(C_{ij}, C_i) \leq \text{threshold}; \\ P_{ij} = (255, 255, 255), & S(C_{ij}, C_i) > \text{threshold}. \end{cases}$$

(5)

式中: P_{ij} 表示像素点被分割后的 RGB 像素值, threshold 为相似度阈值.

3) 初步的分割结束之后,图像中可能存在小范围的噪声点或微小的孔洞,可以利用传统的腐蚀和膨胀处理,使分割效果更加平滑.

1.3 立体匹配

立体匹配是双目视觉中最关键的一步,立体像对之间的差异是由拍摄时观察点的不同引起的^[1].

本实验的目标物体形状较规则,故采用关键点匹配法,选择图像的重心作为关键点. 具体做法是利

用上一步骤中图像分割后获得的二值图像,计算其重心作为关键点实现立体匹配. 二值图像的重心坐标计算如式(6)所示^[8].

$$\begin{aligned} X_c &= \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n x_{ij} \times \alpha_{ij}, \\ Y_c &= \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n y_{ij} \times \alpha_{ij}. \end{aligned} \quad (6)$$

式中: (X_c, Y_c) 为重心坐标, (x_{ij}, y_{ij}) 为像素点坐标,

$$\alpha_{ij} = \begin{cases} 0, & \text{像素为白色,} \\ 1, & \text{像素为黑色.} \end{cases}$$

1.4 三维重建

立体视觉从左、右观测点观看对象时,可以看到对象的偏差. 这种偏差与观测点到被观测对象的距离有关. 三维空间内的点在投影图像上的位置是由三维空间中的某条直线决定的. 因此,如果一个对象从其他视点观测也有同样的三维投影图像,那么被观测对象的空间坐标就可以确定,2 条视线的交点就是它的位置^[8]. 这就是三维重建的过程.

1.4.1 基本模型

空间点三维重建的基本模型^[7] 如图 2 所示.

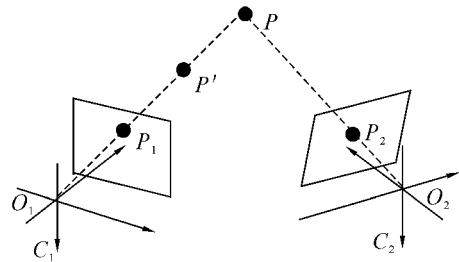


图2 空间点三维重建模型

Fig.2 3-D reconstruction model

对于空间物体表面的任意一点 P ,被摄像头 C_1 拍摄下来后,在图片上位于 P_1 ,被摄像头 C_2 拍摄下来后,在图片上位于 P_2 ,但是无法仅仅通过 P_1 或者 P_2 来得出 P 的三维位置,因为在直线 O_1P 或 O_2P 上的任意一点 P' ,其在图像上的位置都是 P_1 或 P_2 . 然而,现在利用 2 个摄像头来拍摄图片,这样就可以知道目标点位于 O_1P_1 和 O_2P_2 条直线的交点,即可以惟一确定目标点的三维位置.

1.4.2 利用最小二乘法计算三维坐标

假设 2 个摄像头的投影矩阵分别为 M_1 和 M_2 ,可以得到如下 2 个等式^[9-12]:

$$Z_{cl} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

$$Z_{c2} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

式中: (u_1, v_1) 和 (u_2, v_2) 分别为 P 点的成像点在 2 幅图像中的齐次图像坐标, Z_{c1} 和 Z_{c2} 分别为 P 点在 2 个摄像机坐标系下的 Z 坐标; (X, Y, Z) 为 P 点在世界坐标系下的齐次坐标;利用最小二乘法,就可以

求得 P 点的三维坐标^[9].

2 实验平台

以广茂达机器人作为实验平台,基于此改进的双目立体视觉系统进行目标的识别和定位. 并根据定位结果控制机器人对目标进行抓取. 图 3 为整个系统的流程图.

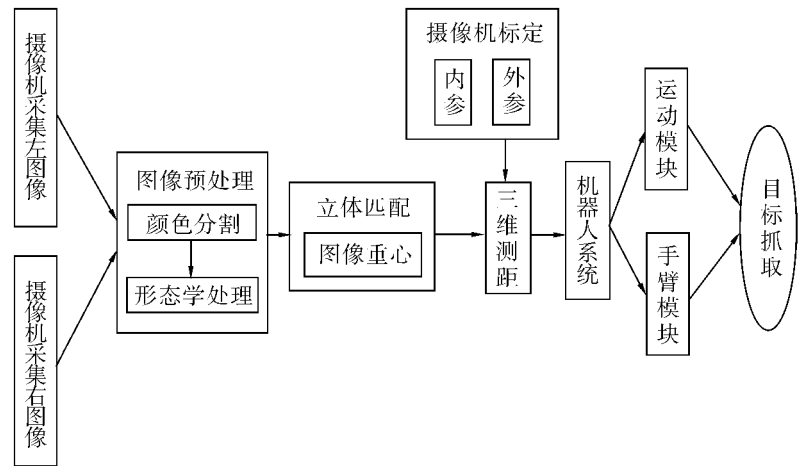


图 3 系统流程

Fig.3 Flow chart of the system

2.1 硬件平台

该机器人的视觉系统是由 2 个 Sony CCD 摄像机组成的. 每个摄像机包含 2 个自由度,即水平方向

转动范围为 $[-100^\circ, 100^\circ]$, 竖直方向的转动范围为 $[-25^\circ, 25^\circ]$, 同时具有 130 万像素和 30 帧/s 的最高帧率. 它的硬件系统如图 4 所示.

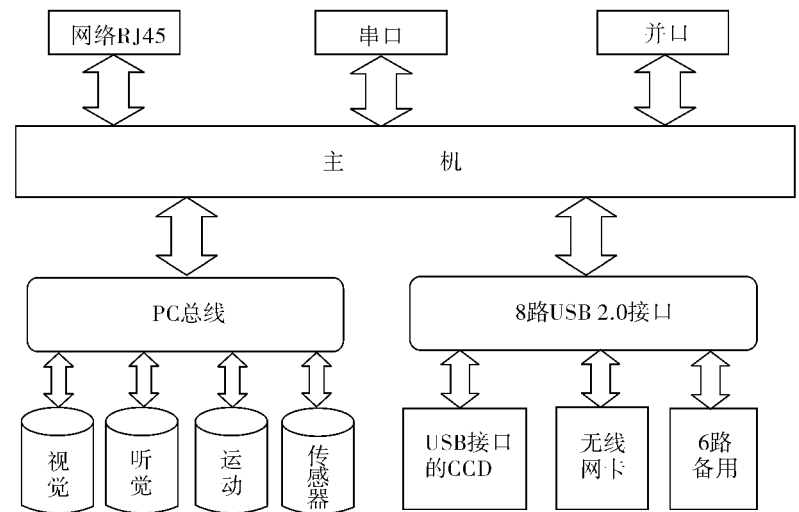


图 4 机器人硬件系统

Fig.4 Hardware of the robot

2.2 软件环境

机器人的核心是普通 PC, CPU 为 P4 3.0, RAM

为 256 MB, 可以根据需要对其配置进行升级. 开发环境为 VC++ 6.0, 在此基础上完成了完整的双目

立体视觉系统开发. 系统主界面如图5所示.

在此系统中, 可完成相应的视频采集、云台控制、三维测距、手臂控制和运动控制等功能.

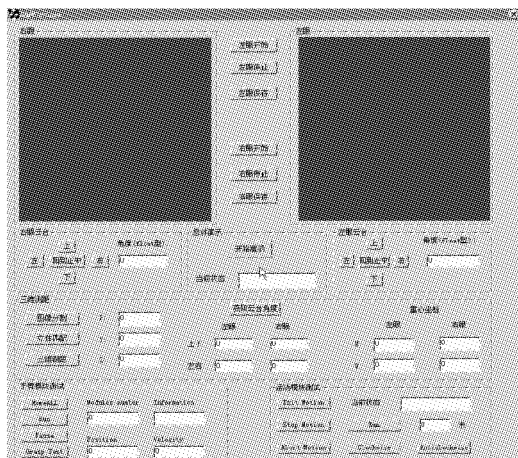


图5 软件系统主界面

Fig.5 The main interface of the software system

3 实验结果及分析

3.1 摄像机标定

利用上文介绍的标定方法, 对2个摄像头分别进行了标定. 标定物选择标准棋盘格, 标定过程中部分角点提取结果如图6所示, 左右摄像头参数的标定结果如表1所示.

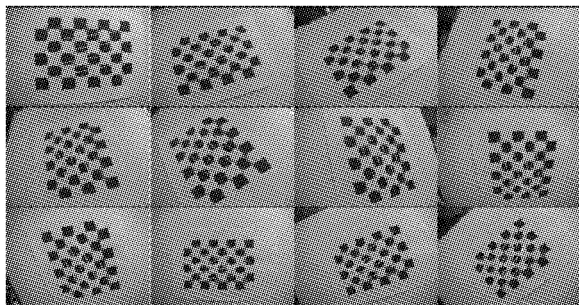


图6 标定过程部分角点提取结果

Fig.6 Part of the corner extraction results

表1 摄像头标定结果

Table 1 Camera calibration results

参数	左摄像头	右摄像头
内参矩阵	$\begin{bmatrix} 278.47 & 0 & 173.35 \\ 0 & 278.45 & 119.67 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 279.24 & 0 & 169.70 \\ 0 & 279.12 & 111.93 \\ 0 & 0 & 1 \end{bmatrix}$
外参旋转	$[2.20 \ 2.15 \ 0.04]$	$[-2.27 \ -2.24 \ 0.04]$
平移	$[-41.39 \ 73.46 \ 563.32]$	$[-226.02 \ 80.28 \ 573.51]$

从数据上看, 左右摄像头的焦距基本相同, 并且, 它们各自的 x 轴和 y 轴的等效焦距也相同. 摄像头拍摄得到的图像尺寸为 320×240 , 内参矩阵反映的光心点坐标也基本在 $(160, 120)$ 附近. 这些都符合标定的预期. 但是, 为了更加准确地衡量标定误差, 需要进行重投影误差计算. 假设 (u_1, v_1) 为角点检测得到的图像坐标, 而 (u_2, v_2) 为利用标定结果, 对实际三维坐标投影计算得到的图像坐标. 利用式(7)计算重投影误差, 得到左摄像头的重投影误差为 0.137 74, 右摄像头的重投影误差为 0.135 54. 误差级别低于一个像素, 说明此方法标定精度较高.

$$E_w = \frac{\sum_{i=1}^N \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}}{N}. \quad (7)$$

对于基于云台转角的外参数估计方法, 同样进行了误差分析. 假设 (t_{x1}, t_{y1}, t_{z1}) 为利用标定方法得到的平移向量, 而 (t_{x2}, t_{y2}, t_{z2}) 为利用改进的估计方法得到的平移向量. 利用式(7)计算总体平均误差, 计算结果如表2所示.

$$E_t = \frac{\sum_{i=1}^N \sqrt{(t_{x1} - t_{x2})^2 + (t_{y1} - t_{y2})^2 + (t_{z1} - t_{z2})^2}}{3N}. \quad (8)$$

表2 平移向量估计总体平均误差

Table 2 The overall average error of estimated translation vector cm

摄像头	上下移动	左右移动
左摄像头	1.294 9	1.532 9
右摄像头	1.395 3	1.652 9

从误差计算结果来看, 此外参估计方法可以比较精确地估计摄像头旋转之后的外参.

3.2 图像分割

图像分割采用的是上文提到基于颜色相似度的分割方法. 首先将图像从 RGB 空间转换到 HSV 空间, 进而计算颜色模板, 最后设置阈值进行图像分割. 实验结果如图7~9所示.

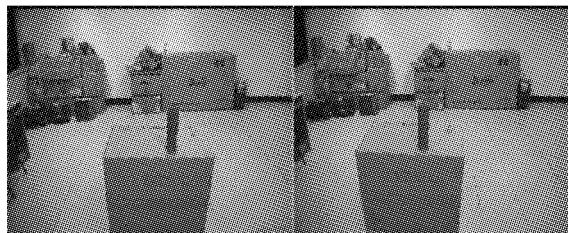


图7 原始图像

Fig.7 The original images

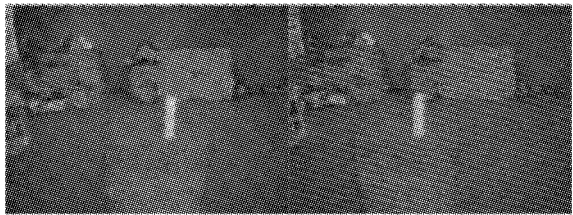


图 8 转化为 HSV 空间的图像
Fig.8 The images transferred to HSV space

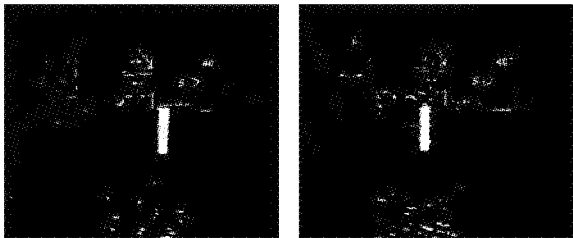


图 9 图像分割结果
Fig.9 Image segmenting results

直接颜色分割的图像含有一定的噪声,采用形态学的腐蚀和膨胀处理后,结果如图 10 所示.

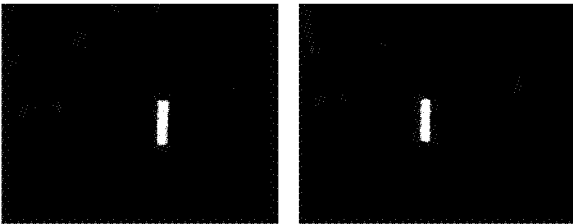


图 10 形态学处理后的结果
Fig.10 Morphological processing results

3.3 立体匹配

利用颜色分割得到二值化图像,计算目标物重心坐标.重心标记结果如图 11 所示.

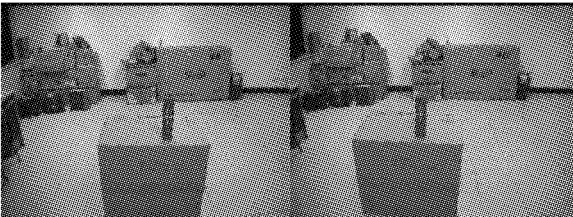


图 11 重心标记结果
Fig.11 The results of marking the center of gravity

3.4 三维重建

为了评测测距系统的功能和精度,对棋盘格中的角点进行了三维重建.实验中,使得棋盘尽量与机器人平行,以保证角点重建后的 Z 轴分量大致相同.拍摄得到的棋盘图如图 12 所示.棋盘格的尺寸为 28 mm,角点的尺寸为 7×5.利用 Harris 亚像素级角点检测方法得到图像中的角点坐标,然后对其进行三维重建.实验结果如表 3 所示.表 3 中,角点(1, 1)为棋盘图中左上角第一个角点,且所有角点按照从左到右,从上到下的顺序排列.同样,利用式(9)

计算测量平均绝对误差.计算结果为 2.356 1 mm.从结果上来看,此方法可以比较精确地实现目标三维测距.

$$E = \frac{\sum_{i=1}^N \sqrt{(X_{w1} - X_{w2})^2 + (Y_{w1} - Y_{w2})^2 + (Z_{w1} - Z_{w2})^2}}{2} \quad (9)$$

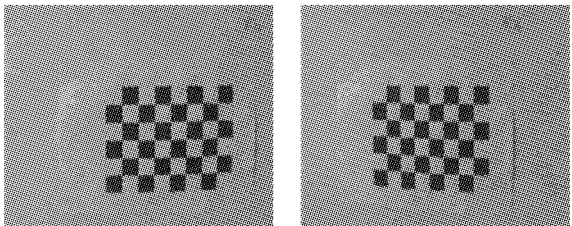


图 12 三维测距棋盘
Fig.12 Chessboard graph used for binocular distance

3.5 目标抓取

最后,利用广茂达机器人平台,实现了改进的双目立体视觉系统.并以此对目标识别定位的结果作为依据来控制机器人和手臂的相应运动,从而成功地实现了机器人对目标物体的准确抓取.

广茂达机器人抓取目标的流程为:首先识别目标物,然后根据目标在图像中的位置控制云台转动直至目标位于视野正中,进而进行立体匹配,并对目标物体实际位置进行定位,根据定位结果,首先控制机器人靠近目标物体,进而控制手臂模块进行相应的角度旋转,最终完成抓取动作.

实验中,机器人可成功地抓取目标物体,抓取过程部分截图如图 13~15 所示.

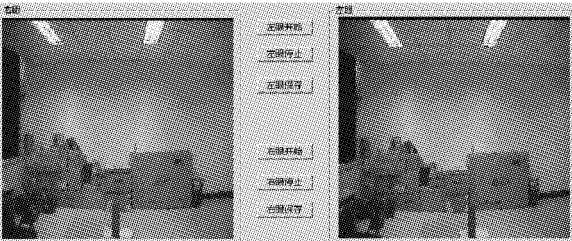


图 13 程序开始时刻
Fig.13 At the beginning of the process

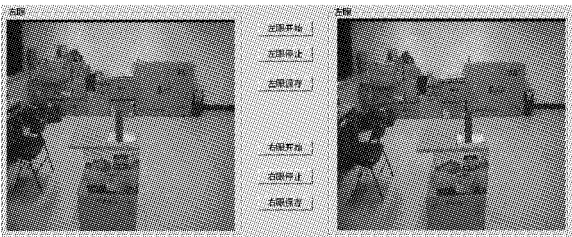


图 14 调整云台使目标位于视野中央
Fig.14 Adjust the head to let the target in the center of the vision



图15 实际抓取过程截图

Fig. 15 Actual capture process

表3 三维重建坐标与实际坐标比较

Table 3 Comparison of 3-D reconstruction coordinates and actual coordinates

角点	角点三维重建坐标/mm	实测三维坐标/mm
(1,1)	(-1.99, -0.979, -0.82)	(0,0,0)
(1,2)	(-1.57, 28.00, -0.48)	(0,28,0)
(1,3)	(-1.43, 56.76, 0.09)	(0,56,0)
(1,4)	(-1.67, 83.93, 0.96)	(0,84,0)
(1,5)	(-0.90, 111.00, 0.44)	(0,112,0)
(1,6)	(-1.42, 139.35, 0.33)	(0,140,0)
(1,7)	(-1.00, 166.00, 0.85)	(0,168,0)
(2,1)	(26.94, 0.06, 0.57)	(28,0,0)
(2,2)	(26.88, 28.85, 0.73)	(28,28,0)
(2,3)	(26.60, 56.04, 0.02)	(28,56,0)
(2,4)	(27.43, 84.22, -1.83)	(28,84,0)
(2,5)	(26.21, 111.10, -1.71)	(28,112,0)
(2,6)	(26.68, 138.60, 0.17)	(28,140,0)
(2,7)	(27.30, 167.16, 0.71)	(28,168,0)
(3,1)	(54.50, -0.60, 0.45)	(56,0,0)
(3,2)	(54.93, 28.12, 0.06)	(56,28,0)
(3,3)	(55.48, 56.58, -1.89)	(56,56,0)
(3,4)	(55.15, 83.52, -1.14)	(56,84,0)
(3,5)	(54.99, 111.50, -1.57)	(56,112,0)
(3,6)	(55.62, 138.41, 0.03)	(56,140,0)
(3,7)	(55.50, 166.38, 0.56)	(56,168,0)
(4,1)	(84.04, 0.66, 0.77)	(84,0,0)
(4,2)	(83.37, 27.85, 0.52)	(84,28,0)
(4,3)	(82.67, 56.58, -1.20)	(84,56,0)
(4,4)	(83.27, 83.72, -1.60)	(84,84,0)
(4,5)	(82.85, 110.77, -1.44)	(84,112,0)
(4,6)	(83.79, 139.02, -1.31)	(84,140,0)
(4,7)	(82.73, 165.68, -1.80)	(84,168,0)
(5,1)	(112.57, -0.07, 0.66)	(112,0,0)
(5,2)	(111.62, 29.14, -1.83)	(112,28,0)

续表3

角点	角点三维重建坐标/mm	实测三维坐标/mm
(5,3)	(111.61, 55.86, -1.06)	(112,56,0)
(5,4)	(110.86, 83.94, -1.44)	(112,84,0)
(5,5)	(111.40, 111.24, -1.31)	(112,112,0)
(5,6)	(110.86, 138.33, -1.77)	(112,140,0)
(5,7)	(111.48, 166.51, -1.67)	(112,168,0)

4 结束语

本文完成的是一个基于改进的双目立体视觉的目标识别与定位系统. 双目立体视觉系统主要包括摄像机标定、图像分割、立体匹配和三维测距4个模块. 其中,在摄像机标定部分,针对系统的运动特点,在原有算法的基础上提出了世界坐标系转化为机器人坐标系和基于云台转角的外参数估计2种方法,经过误差分析,基于云台转角的外参数估计方法可以精确地完成摄像机旋转情况下对外参的估计,使云台在实验过程中可以更加灵活的旋转,增强了该机器人的视觉功能,最后,利用广茂达机器人系统作为实验平台,基于改进后的双目立体视觉系统进行目标的识别与定位,并以识别定位的结果作为依据来控制机器人和手臂的相应运动,从而成功地实现了机器人对目标物体的准确抓取,此结果更充分验证了所提出方法的可行性.

参考文献:

- [1] 祝琨. 基于双目视觉的运动物体实时跟踪与测距[D]. 北京:北京交通大学, 2008: 1-103.
ZHU Kun. Real-time tracking and measuring of moving object based on binocular vision[D]. Beijing: Beijing Jiaotong University, 2008: 1-103.
- [2] ZHANG Zhengyou. Flexible camera calibration by viewing a plane from unknown orientations[C]//Proceedings of the IEEE International Conference on Computer Vision. Kerkyra, 1999: 666-673.
- [3] 吴亚鹏. 基于双目视觉的运动目标跟踪与三维测量[D]. 西安:西北大学, 2008: 1-58.
WU Yapeng. Moving objects tracking and 3-D measurement based on binocular vision[D]. Xi'an: Northwest University, 2008: 1-58.
- [4] CHEN Yuming. Research on binocular 3-D measurement model based on grating projection[C]//PEITS 2009—2009 2nd Conference on Power Electronics and Intelligent Transportation System. Shenzhen, China, 2009: 306-308.
- [5] ZHANG Zhengyou. A flexible new technique for camera calibration[J]. IEEE Transactions Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [6] 瞿中. 图像分割与评价及图像三维表面重建研究[D]. 重庆:重庆大学, 2009: 1-108.

QU Zhong. Research on image segmentation, evaluation and image three-dimensional surface reconstruction [D]. Chongqing: Chongqing University, 2009: 1-108.

[7] ZHAO Peng, NI Guoqiang. Simultaneous perimeter measurement for 3-D object with a binocular stereo vision measurement system [J]. Optics and Lasers in Engineering, 2010, 48(4): 505-511.

[8] 肖南峰. 智能机器人 [M]. 广州: 华南理工大学出版社, 2008: 10-98.

[9] 蔡健荣, 范军, 李玉良. 立体视觉系统标定及成熟果实定位 [J]. 农机化研究, 2007, 11: 3841.

CAI Jianrong, FAN Jun, LI Yuliang. Binocular system calibration and mature fruit location [J]. Agricultural Research, 2007, 11: 3841.

[10] 周子啸, 赵晓林, 胡峰, 等. 基于仿真机器人的双目视觉定位 [J]. 电视技术, 2010, 8: 106-109.

ZHOU Ziiao, ZHAO Xiaolin, HU Feng, et al. Binocular stereo location on emulator robot [J]. TV Technology, 2010, 8: 106-109.

[11] 郭巍, 张平, 朱良, 陈曦. 基于AdaBoost-ECOC的合成孔径雷达图像目标识别研究 [J]. 哈尔滨工程大学学报, 2010, 31(2): 232-236.

GUO Wei, ZHANG Ping, ZHU Liang, CHEN Xi. Research on synthetic aperture radar image target recognition based on AdaBoost-ECOC [J]. Journal of Harbin Engineering University, 2010, 31(2): 232-236.

[12] 王科俊, 魏娟. 基于共面圆的双目立体视觉分步标定法 [J]. 应用科技, 2010, 37(1): 36-40.

WANG Kejun, WEI Juan. Binocular stereo vision step by step calibration based on coplanar circle [J]. Applied Science and Technology, 2010, 37(1): 36-40.

作者简介:



尚倩, 女, 1987年生, 硕士研究生, 主要研究方向为计算机视觉、数字图像处理、信号与信息处理。



阮秋琦, 男, 1944年生, 教授, 博士生导师, 北京交通大学信息科学研究所所长, IEEE高级会员. 主要研究方向为图像处理、计算机视觉、多媒体信息处理、虚拟现实、计算机支持的协同工作. 曾获国家教委科技进步二等奖、铁道部科技进步二等奖和三等奖等. 发表学术论文150余篇, 出版著作3部, 获国家专利1项。



李小利, 女, 1986年生, 博士研究生, 主要研究方向为数字图像处理、计算机视觉和模式识别。

欢迎订阅《控制工程》期刊

跟踪国际自动化最新动态 发布自动化领域最新研究成果 展示自动化领域最新产品

《控制工程》(双月刊)

国际刊号: ISSN1671-7848 国内刊号: CN21-1476/TP

邮发代号: 8-216

国内定价: 15元/期(全年90元)

《控制工程》是由国家教育部主管、东北大学主办的国家级学术性期刊, 创刊于1994年. 《控制工程》系系北京大学中文核心期刊、中国科技核心期刊、中国科学引文数据库核心期刊(CSCD)、中国科技论文统计用刊、中国科学文献数据用刊、《中国学术期刊综合评价数据库》来源期刊, 俄罗斯《文摘杂志》(AJ)、美国《剑桥科学文摘》(CSA)、英国《科学文摘》(SA, INSPEC)三大著名检索机构全文检索期刊。

· 本刊是自动控制学科的学术期刊, 主要刊登反映自动化技术及产品的研究开发及应用方面最新成果的研究报告, 以及学术研讨和自动化高技术及其产品的介绍。

· 主要栏目: 综述与评论、过程控制技术的应用、智能控制技术的应用、优化控制技术的应用、计算机控制系统及软件、机电一体化、电气自动化、检测装置及仪表、企业资源与管理系统、综合自动化系统、故障诊断、典型自动化工程、模式识别及图像处理等。

地址: 沈阳东北大学310信箱

邮编: 11004

电话: 024-23883498/024-83687509

传真: 024-23883498

E-mail: kzgcbj@mail.neu.edu.cn

http://www.kzge.com.cn