

基于视觉的目标定位技术的研究进展

赵霞 袁家政 刘宏哲

(北京联合大学北京市信息服务工程重点实验室 北京 100101)

摘要 基于视觉的目标定位是目前计算机视觉领域的研究热点。就现阶段基于视觉的定位技术进行了综述,着重介绍了基于单目视觉定位、基于双目视觉定位和基于全景视觉定位 3 类定位技术的研究现状,并分析了各自的优缺点。最后就基于视觉定位方法的发展趋势做了简要分析,以期为以后基于视觉定位问题的研究提供参考。

关键词 目标定位,单目视觉,双目视觉,全景视觉

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.4.002

Advances in Vision-based Target Location Technology

ZHAO Xia YUAN Jia-zheng LIU Hong-zhe

(Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China)

Abstract The vision-based target location technology is the hotspot in the field of computer vision. The current state-of-the-art visual positioning technologies were summarized in this paper. This paper emphatically introduced research status, advantages and disadvantages of monocular vision-based positioning technology, binocular vision-based positioning technology and panorama vision-based positioning technology. Finally, the development trends of vision-based locating method were described briefly. This paper can provide reference for the research of visual positioning problem.

Keywords Object localization, Monocular vision, Stereo vision, Omni-directional vision

1 引言

近年来,随着计算机技术和人工智能的发展,计算机视觉在越来越多的行业得到广泛的应用。目标定位不仅是计算机视觉领域重要的研究内容,也是非常具有挑战性的研究课题,其在机器人视觉导航、公共场景监控、军事视觉制导、智能交通、博物馆智能展示等诸多领域有着十分广泛的应用^[1]。因此研究基于视觉的目标定位有着非常重要的意义。

传统的目标定位技术主要利用雷达(Radar)^[2]、激光(Laser)^[3]、红外线(infra-red)^[4]等传感器进行测距以达到目标定位的目的,但是用这些传感器存在以下缺点:一方面,它们很容易受到机器本身携带的其他传感器的影响,导致测距不准确^[5];另一方面,它们的最大缺点是不能识别目标物体。而随着视觉传感器的发展,近年来,利用视觉传感器获取景物特征实现定位的方法因其所需要的实验环境和硬件条件更加宽松、成本较低、在实际应用中更直观有效而受到了重视。根据使用视觉传感器的数目不同,视觉定位方法可分为基于单目视觉的定位、基于双目视觉的定位和基于全方位视觉的定位。

其中基于单目视觉的定位方法仅利用一个视觉传感器来完成定位工作,基于摄像机数学模型建立空间目标特征点与图像特征点之间的对应投影变换关系,从而确定目标特征点

位置信息。而基于双目视觉的定位方法仿照人类双眼感知周围环境空间深度的功能,利用两个视觉传感器从不同位置拍摄同一场景,对所拍摄的图像对进行匹配,并计算视差,然后利用三角测量原理实现距离测量。基于全方位视觉的定位方法利用全方位视觉传感器来完成目标定位工作。

2 基于单目视觉的定位方法

基于单目视觉的定位方法根据用于定位图像的数目,可以分为基于单幅、两幅或两幅以上图像的单目视觉定位。

2.1 基于单幅图像的定位方法

基于单幅图像的定位主要利用一幅图像的信息完成目标定位的工作,图 1 所示为基于单幅图像的目标定位模型。

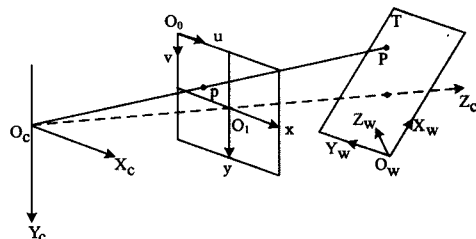


图 1 基于单幅图像的目标定位模型

图中 O_c 为摄像机光心, (O_c, X_c, Y_c, Z_c) 是摄像机坐标

到稿日期:2015-09-11 返修日期:2015-12-11 本文受国家科技支撑课题(2014BAK08B02),国家自然科学基金(61271369,61372148),北京联合大学“人才强校计划人才资助项目”(BPHR2014E02)资助。

赵霞(1990—),女,硕士生,主要研究方向为数字图像处理,E-mail: zhaoxia0835@163.com;袁家政(1971—),男,博士,教授,主要研究方向为图形图像处理、文物遗迹的数字化处理、数字博物馆、导航定位等,E-mail: xxtjiazheng@buu.edu.cn(通信作者);刘宏哲(1971—),女,博士,教授,主要研究方向为语义计算、数字图像处理、分布式系统、人工智能、数字博物馆等,E-mail: liuhongzhe@buu.edu.cn.

系, X_c, Y_c 分别与图像平面的 x, y 轴平行, Z_c 为摄像机光轴。 O_1 为光轴与图像平面的交点, 即图像平面的中心, O, O_1 为摄像机的焦距。世界坐标系 (O_w, X_w, Y_w, Z_w) 建立在平面 T 上。对于平面上的一点 $p(X, Y, Z)$, 在图像平面的投影 p 是点 P 与摄像机光心的连线 $O_c p$ 与图像平面的交点。

设图像中任意一点 p 在图像平面坐标系 (O_1, x, y) 中的坐标为 (x, y) , 在图像像素坐标系 (O_0, u, v) 中的坐标为 (u, v) , 则两者有如下关系:

$$u = \frac{x}{dx} + u_0 \quad (1)$$

$$v = \frac{y}{dy} + v_0$$

其中, (u_0, v_0) 为 O_1 在 u, v 坐标系中的坐标, dx, dy 为每一个像素在 x 轴、 y 轴方向上的物理尺寸, 用齐次坐标表示为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

设空间一点 P 在摄像机坐标系下的坐标为 $(X_c, Y_c, Z_c, 1)^T$, 在世界坐标系下的坐标为 $(X_w, Y_w, Z_w, 1)^T$, 则两者之间的关系为:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3)$$

其中, R 为 3×3 的正交单位矩阵, 表示从世界坐标到摄像机坐标的旋转矩阵; t 为 3×1 的向量, 表示从世界坐标到摄像机坐标的平移向量; M_1 为 4×4 矩阵。

空间任意一点 P 与其在图像中的投影 p 的比例关系如下:

$$x = f \frac{X_c}{Z_c} \quad (4)$$

$$y = f \frac{Y_c}{Z_c}$$

其中, (x, y) 为 p 点的图像坐标, (X_c, Y_c, Z_c) 为空间点 P 在摄像机坐标系下的坐标, f 为摄像机焦距。用齐次坐标关系式表示上述投影关系:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (5)$$

将式(2)、式(3)代入式(5), 得到世界坐标系中 P 点坐标与投影点 p 的坐标 (u, v) 的关系式:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (6)$$

其中, $f_x = \frac{f}{dx}, f_y = \frac{f}{dy}, f_x$ 和 f_y 分别称为 u 轴和 v 轴上的尺度因子, 也称为 u 轴和 v 轴上的归一化焦距; u_0, v_0 为图像平面中心。

令:

$$M_1 = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$$

则式(6)为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (7)$$

其中, $M = M_1 M_2, M$ 为 3×3 矩阵, 称为投影矩阵, M_1 为摄像机内参数, M_2 为摄像机外参数, 它们可通过摄像机标定获得。空间点的三维深度信息是指空间点在摄像机坐标系中的 Z 轴坐标, 即 Z_c 。

由式(7)得, 基于单幅图像进行目标定位, 其本身信息量较少, 仅仅依靠上述信息一般很难获取图像中目标的位姿信息。因此, 基于单幅图像的定位方法主要借助以下两种方式获取更多有效信息: 一种是设置人工标志, 例如特征光点、特征直线和特征曲线等, 在单幅图像中依靠这些特征的已知空间信息与检测的图像特征间的映射关系得到目标的位姿信息, 这种方式由于存在更多的已知信息, 因此求解相对简单; 另一种是无人工标志, 仅凭借目标物的自然属性, 例如平行、正交等几何特性, 再结合图像特征信息完成目标物的定位, 这类方法的求解则较为困难。基于单幅图像的定位根据已知目标特征(如点、直线、曲线、平面四边形等)在世界坐标系下的描述和这些特征在图像上的投影, 以及一些额外的特征参数, 来确定世界坐标系和摄像机坐标系之间的位姿关系, 以对目标进行定位。基于单幅图像的单目视觉定位的具体流程如图2所示。

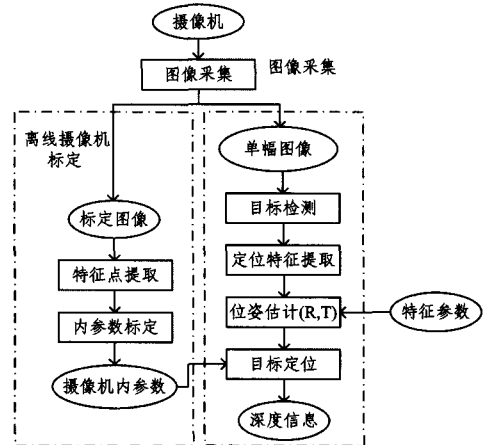


图2 基于单幅图像的目标定位的流程

2.1.1 设置人工标志的单幅图像定位

通过在目标物体上设置人工标志, 如特征点、特征直线、特征曲线等, 利用已标定的摄像机获取这些人工标志的图像, 对其进行图像处理以及相关的计算得到摄像机坐标系与人工标志坐标系的相对位姿信息。依据人工标志选取的不同, 定位主要分为: 基于特征点、特征直线以及特征曲线3类。

(1) 基于特征点的定位

基于特征点的定位方法又称 PnP 问题 (Perspective-n-Point Problem)。PnP 问题在 1981 年由 Fisher 等提出, 即利

用图像的 N 个特征点和其对应的空间位置的投影关系,来确定这些特征点的三维空间的位置坐标。目前研究者对 PnP 问题的求解主要有两种途径:一种是直接利用几何关系方法求解空间点到摄像机光心的距离;另一种是通过确定目标坐标系到摄像机坐标系之间的旋转矩阵 R 和平移向量 T ,即如式(3)所示,如果再已知某空间点在物体坐标系下的坐标,就可以求得其在摄像机坐标系下的坐标。第一种求解途径的困难在于约束方程组中有未知数的平方项,不易获得解析解。第二种求解途径中由于旋转矩阵 R 只有 3 个独立变量,导致元素间约束的实现较困难。多年来,研究者们主要对 P3P、P4P 和 P5P 问题进行了深入研究,同时取得了很多有价值的结论^[6,7]。在目标上设置点特征,有定位精度高、速度快等优点。

(2) 基于特征直线的定位

直线特征作为典型的图像特征,具备较强的抗遮挡能力,且提取方便,故利用直线特征进行单幅图像目标物定位也引起了众多学者的兴趣。Liu 等^[8,9]利用图像中的直线特征与摄像机光心形成的投影平面的法向量与目标物自身直线正交的几何特性,建立几何模型进行定位。此种定位方法对 3 条直线的空间位置关系有特殊要求:3 条直线不能同时平行并且不同时与光心共面,此时利用它们可以构造 3 个非线性方程,从而求取空间位姿参数。但是非线性方程组求解较为困难,且容易出现误差。于勇等^[10]在攻击目标可预知的情况下,提出了一种利用导引头图像进行飞机目标定位的方法,即利用飞机主轴及翼展线作为直线特征,借助特征直线成像的尺寸及其相互夹角来估计目标成像姿态的变化,使定位准确度高且不受目标姿态的影响。算法过程如图 2 所示。韩立伟等^[11]提出利用两条相交直线及其交点作为特征,利用直线的方向估计机器人的方位,再通过交点像素变化计算机器人的位置。

(3) 基于特征曲线的定位

基于特征曲线的定位方法一般需要对复杂的非线性方程组进行求解。秦丽娟等^[12]利用非共面曲线进行空间目标的定位研究,通过构造高次方程组进行计算,但是计算量较大,求解较困难。当两个空间曲线位于同一平面时,则能够获得非线性方程组位姿参数的封闭解。圆是常用的一种二次曲线特征,一般来说,经过投影变换后,圆会在图像上形成椭圆,二者之间的相关几何特性存在特定的函数关系,通过求解这种关系可以得到摄像机坐标系与物体坐标系的相对位姿关系。张磊等^[13]提出一种基于圆特征的单目视觉位姿测量改进算法,通过加入激光测距仪信息,解决了两组解的判定问题,在圆半径未知的情况下只需一幅图像就可以得到位姿信息。张广军等^[14]提出利用新型的双圆图像着陆平面标靶,通过建立机载摄像机的运动和投影模型,完成无人机自主着陆位姿参数估计。该方法抗噪声能力较强,可以满足实时性要求。

2.1.2 无人工标志的单幅图像定位

无人工标志的单幅图像定位不通过人工设置特征,而仅利用目标物本身的特征进行空间目标物定位,这种方法比包含人工标志的定位方法已知的信息更少,因此更具有挑战性。目前一些学者对此问题进行了研究,并取得了一定的成果。孙凤梅等^[15]讨论了基于单个平行四边形单幅图像的物体定

位方法,并证明了当仅知道单个平行四边形的单幅图像时,仅能确定物体坐标系相对摄像机坐标系的旋转矩阵和平移向量的方向;当进一步知道平行四边形的一条边长时,平移向量的大小同时可以确定,同时还表明,平行四边形邻边的边长比、邻边的夹角等信息由于不是射影变换下的不变量,不能对物体定位形成新的约束。王娟等^[16]提出了一种新的由二维投影图像来进行空间定位的方法,即采用目标上共面四线与其在二维图像上投影相匹配的测量来实现定位。王晓剑等人^[17]提出了利用单幅图像中目标的共面相交双平行线特征来进行空间定位。

2.2 基于多幅图像的定位方法

基于多幅图像的定位方法的基本思路是通过改变摄像机位置来获取两幅或者两幅以上的图像,然后利用提取的图像匹配特征点对并依据空间物点与所成像点的投影几何变换关系得到摄像机坐标系和物体坐标系的空间位姿参数。基于两幅图像的目标定位模型如图 3 所示,其中 O_1 和 O_2 是两次拍摄时摄像机的光心位置, I_1 和 I_2 是两次拍摄时的成像平面, P 为目标中选取的一个特征点, p_1 与 p_2 为该特征点在两幅图像上的投影,即 p_1, p_2 为对应匹配点。假设摄像机从甲处运动到乙处的旋转矩阵为 R ,平移向量为 T ,设 P 在甲处摄像机坐标系中的坐标为 $X_{c1} = (x_{c1}, y_{c1}, z_{c1})^T$,对应的图像像素坐标为 (u_1, v_1) ;在乙处摄像机坐标系中的坐标为 $X_{c2} = (x_{c2}, y_{c2}, z_{c2})^T$,对应的图像像素坐标为 (u_2, v_2) 。

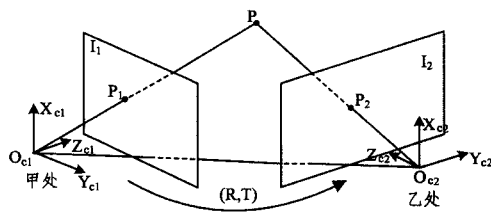


图 3 基于两幅图像的目标定位模型

由摄像机运动参数确定的两处摄像机坐标系的关系如下:

$$X_{c2} = RX_{c1} + T \quad (8)$$

将两坐标代入上式得:

$$\begin{bmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \end{bmatrix} = R \begin{bmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (9)$$

由式(2)、式(5)可得,图像像素坐标系与摄像机坐标系的转换关系为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = M_1 \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (10)$$

由式(11)、式(12)可推理得到:

$$Z_{c2} M_1^{-1} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = Z_{c1} R M_1^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (11)$$

变换可得:

$$\begin{bmatrix} -R M_1^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} & M_1^{-1} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} Z_{c1} \\ Z_{c2} \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (12)$$

$$\text{令 } C = \begin{bmatrix} -RM_1^{-1} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} & M_1^{-1} \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \end{bmatrix}, \text{ 则式(12)可表示}$$

为:

$$C[Z_{c1} \ Z_{c2}]^T = T \quad (13)$$

其中, M_1^{-1}, R, T 都是可求量, 图像平面点坐标值 (u_1, v_1) 与 (u_2, v_2) 可以由特征点的提取和匹配获得, 深度值 Z_{c1} 和 Z_{c2} 可以由最小二乘法计算得到, Z_{c2} 为当前摄像机光心到目标在摄像机坐标系 Z 轴方向上的距离值。

在基于两幅图像的单目视觉定位过程中, 求解两个视角下摄像机坐标系的位置关系是关键。文献[18]通过求本质矩阵来估计两次拍摄的摄像机坐标系之间的位置关系, 具体做法为通过两幅图像的匹配点对计算出基本矩阵, 然后结合摄像机内部参数矩阵求取本质矩阵, 再将本质矩阵分解得到两次拍摄的摄像机坐标系之间的位置关系, 进而实现目标定位, 主要流程如图4所示。

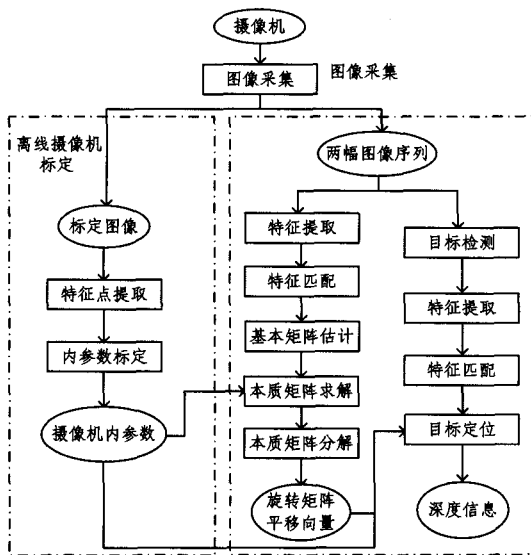


图4 基于两幅图像的单体视觉定位流程

还有一些学者在研究基于视觉信息的微型无人直升机位姿估计的问题时, 两个视角下摄像机坐标系的位置关系通过两次摄像机拍摄图像的相关单应矩阵分解得到。其中文献[19]采用基于尺度不变特征变换(Scale Invariant Feature Transform, SIFT)算法和基于随机抽样一致性(Random Sample Consensus Algorithm, RANSAC)的匹配策略实现航拍图像序列之间特征点的匹配, 来求解单应矩阵, 再将单应矩阵通过奇异值分解得到两次观测摄像机的相对运动关系, 即旋转矩阵 R 和平移向量 T , 进而实现微型无人直升机的位姿估计。但是由于航拍图像的采样频率低, 因此其不适合在摄像机快速移动的环境下运行。

3 基于双目视觉的定位方法

基于单目视觉的定位方法所使用的单个摄像头视野范围小、获取的深度信息少、定位精度低。近年来, 基于双目视觉的目标定位研究成为计算机视觉领域的研究热点。基于双目视觉的目标定位方法按使用的传感器的种类不同分为传统的双目视觉定位方法和基于 Kinect 视觉新工具的定位方法。传统的双目视觉定位系统利用两个彩色摄像头来进行视觉距离估计, 即通过获取同一时刻、同一场景且具有一定视差的两

幅彩色图像来进行视觉距离估计; 而基于 Kinect 视觉新工具的定位方法借助一个红外摄像头、一个红外投影仪以及一个彩色摄像头进行视觉距离的估计, 不仅能够获得彩色图像, 还能获得同步的深度图像。

3.1 传统的双目视觉定位方法

传统的双目视觉定位方法是模仿人类或生物系统的视觉距离估计和景物重建过程, 即由存在一定间距的两个彩色摄像头拍摄同一物体, 同一物体在两个摄像头中的成像像素点具有对应关系, 根据成像变换矩阵和对对应像素点在像空间中的位置就可以求出物体的空间位置。其基本原理如图5所示, 根据相似三角形原理, 物体距摄像机的距离 d 为:

$$d = \frac{bf}{x_l - x_r} \quad (14)$$

其中, b 为两摄像头之间的距离, f 为摄像机的焦距, x_l 和 x_r 分别表示物点在左右摄像机的投影点的水平位置。

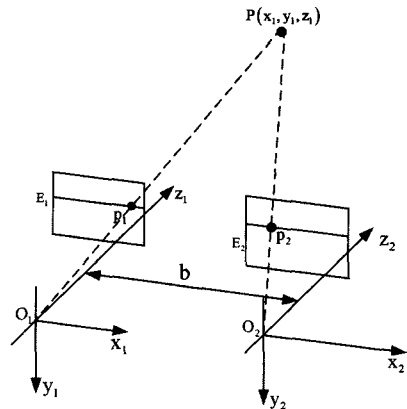


图5 双目视觉的基本原理

传统的双目视觉定位方法是在目标识别的基础上进行双目视觉定位的, 即将目标识别后, 再运用双目视觉的原理进行测距, 其基本流程如图6所示。

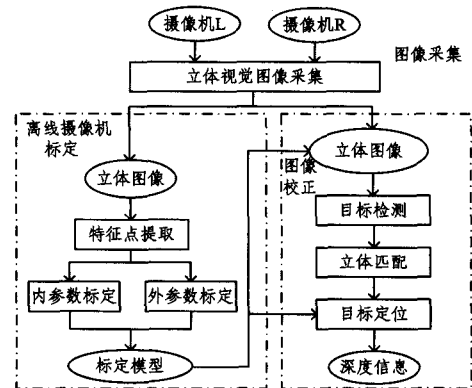


图6 传统双目视觉的目标定位流程

根据式(14), 要计算物体与摄像机的距离, 需要知道的参数有焦距 f 、摄像机之间的距离 b 和物点在左右摄像机的投影点的横坐标的差 $x_l - x_r$, 即视差。 f 和 b 可以通过摄像机标定获得, 视差成为计算物体与摄像机的距离的关键。在背景环境中正确获取目标后, 如何在左、右摄像机采集的图像中选取具有空间位置一致性的目标点一直是双目视觉技术的难点和关键, 立体匹配的结果直接影响到目标定位的精度。近年来, 有许多学者对基于双目视觉的目标定位做了大量研究, 依据其在立体匹配中选用方法的不同, 将基于双目视觉的目标定位研究分为基于特征点的双目视觉定位方法、基于区域的

双目视觉定位方法以及基于特征和区域联合的双目视觉定位方法。

3.1.1 基于特征点的双目视觉定位方法

基于特征点的双目视觉定位方法主要通过提取左右眼图像中的一些局部特征点来实现图像的匹配与定位,如 Harris 角点特征、FAST(Features from Accelerated Segment Test)、SIFT、SURF(Speeded Up Robust Features)、ORB(Oriented FAST and Rotated Brief)等。其中 Harris 角点特征因算法编程灵活、稳定性好等而被广泛应用于立体视觉图像匹配中。但当立体图像匹配对之间几何尺寸、平移以及旋转角度发生较大变化时,利用 Harris 特征进行图像匹配时得到的效果并不理想,因此有大量学者对上述问题进行了改进,如罗桂娥等人将尺度不变性融入了 Harris 特征。尽管改进的 Harris 特征对图像立体匹配可以得到较好的匹配精度,但是其对于场景中角点不明显的图像检测的失败率较高。SIFT 特征^[28,29]因对尺度、旋转、视角等变化具有良好的鲁棒性而被广泛应用于双目视觉立体匹配中^[27],但是其计算复杂度较高,实时性较差。因此有学者^[30,32]提出了基于 SURF^[31]特征匹配的立体视觉定位方法,其克服了 SIFT 算法实时性差的缺点。

还有许多研究者根据立体视觉定位所针对的特定目标进行特殊特征点的提取,这种方法不仅能有效地解决实际问题,而且准确率更高。例如,明祖衡^[34]将手指尖作为特征匹配点,研究了一种基于立体视觉原理测量空间手指距离的算法。郑晓东等人^[35]和项荣等人^[36]利用番茄的形心特征进行立体匹配,实现番茄的深度信息的获取,满足番茄采摘机器人视觉系统在大多数采摘作业环境下的工作要求。

3.1.2 基于区域的双目视觉定位方法

特征匹配具有较好的匹配精度和鲁棒性,基本不受外界光照变化的影响,而且特征匹配方法将匹配的搜索范围限制在一系列比较稀少的、较为符号化的特征上,因此计算量较小,匹配速度较快。然而,特征在图像中的稀疏性决定了特征匹配只能得到稀疏的视差场,不能呈现细节部分^[37,38]。基于区域的立体视觉定位方法以区域匹配算法作为基础,能够获得稠密、均匀的图像视差图,并具有一定的匹配精度;但是,其对于图像旋转和外界光照的变化较为敏感,鲁棒性较差,不能够处理图像中纹理相似区域或遮挡图像区域^[38,39]。所谓区域立体匹配的算法,即给定一幅图像上的某一点,选取该像素点邻域内的一个子窗口,在另一幅图像中的一个区域内根据某种相似性测度算法,寻找与子窗口图像最为相似的子图,而其匹配的子图中对应的像素点就为该像素的匹配点,基本原理如图 7 所示。常用的相似性测度算法如表 1 所列,其中 p_i 和 p_j 分别为待匹配点窗口和候选匹配点窗口中各像素的灰度值。目前也有一些学者对基于区域的视觉定位方法进行了研究,如蒋焕煜等人^[40]根据立体图像对中像素点的相关性来对图像进行立体匹配,采用像素差的平方和(Sum of Squared Differences, SSD)作为相似测度,计算像素点的位置信息而获得一幅深度图像,最后对照番茄形心在深度图像中的位置,获取番茄的三维位置信息。实验结果表明:立体摄像机与番茄的距离小于 1000mm 时,番茄的距离误差为 ±20mm,基本满足番茄采摘机器人的工作要求。Li 等人^[41]利用绝对差之和(Sum of Absolute Differences, SAD)进行立体匹配,得到深度图,计算目标的位置,利用此方法有效实现了机器人抓取可口可乐瓶。

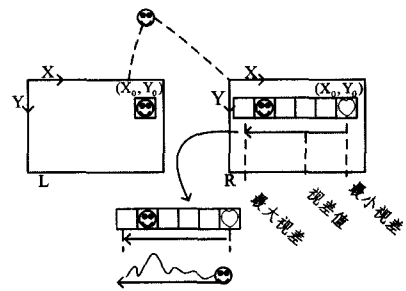


图 7 基于区域的立体匹配原理

表 1 常用相似性测度算法

相似性测度算法	数学表达式
像素差平方和(SSD)	$\sum (p_i - p_j)^2$
像素差绝对值和(SAD)	$\sum p_i - p_j $
归一化交叉相关(NCC)	$\frac{\sum p_i p_j}{\sqrt{\sum p_i^2} \sqrt{\sum p_j^2}}$
零均值归一化交叉相关(ZNCC)	$\frac{\sum (p_i - \bar{p}_i)(p_j - \bar{p}_j)}{\sqrt{\sum (p_i^2 - np_i^2)} \sqrt{\sum (p_j^2 - np_j^2)}}$

3.1.3 基于特征和区域联合的双目视觉定位方法

针对区域匹配与特征匹配两种算法的优势与不足,采用互补策略,充分利用区域匹配视差图的致密性以及特征匹配视差图的鲁棒性^[22]将两种算法结合起来,提出了一种基于特征和区域联合的双目视觉定位方法。罗桂娥等^[22]针对特征匹配得到的是稀疏视差图的问题,提出了改进 Harris 角点匹配与区域匹配相结合的立体匹配算法。实验结果表明,所提出的联合的匹配算法可得到致密的视差图,并在边缘等视差不连续区域可获得较好的匹配精度,验证了该算法的有效性和可行性。项荣等人^[36]为了提高番茄采摘机器人视觉系统的定位精度,提出一种基于组合匹配方法和深度校正模型的双目立体视觉测距方法。该方法在形心匹配的基础上,将形心匹配得到的视差值作为区域匹配时设定视差范围的参考值,这样可以减小区域匹配计算量误匹配的概率。结果表明,当工作距离小于 1050mm 时,其可满足番茄采摘机器人在大多数采摘作业环境下的工作要求。

3.2 基于 Kinect 视觉新工具的定位方法

随着微软新体感工具 Kinect 的推出,基于视觉的目标定位变得方便简单,通过 Kinect 工具不仅可以获得目标场景的高质量彩色图像,而且还可以得到同步的高质量深度图像。目前国内外众多研究机构或者学者利用 Kinect 代替传统的双目摄像头实现目标的定位,利用 Kinect 所提供的深度数据,可以更准确地识别与跟踪目标^[42]。例如,董建民等人^[43]结合 Kinect 获得的彩色图像与深度图像信息实现西红柿的定位,用于辅助真空吸盘机械手进行西红柿的自动采摘。与其相类似的研究还有, Figueroa 等人^[44]提出用 Kinect 进行目标识别与定位,其还能用于机器人手臂抓取物体。Zhang 等人^[45]将 Kinect 用于跟随机器人,能够准确地定位移动机器人与跟随人脊椎点的距离以指导机器人向前向后移动,实验证明了所提方法的有效性。除了机器人方面的应用, Kinect 也常用于室内、外的定位, NAKANO 等人^[46]将 Kinect 用于室内定位,不仅硬件简单、成本较低,而且定位准确度超过了目前其他的室内定位系统。Wang 等人^[47]提出了一种新颖的基于 RGB-D 图像检测与识别楼梯、人行道和交通信号灯的方法,以为盲人或者有视觉障碍的人群提供帮助。虽然 Kinect

在视觉定位技术上取得了一些进展,但因其识别的最大距离约为4m,因此阻碍了其对外场景中的目标和远距离目标的检测、定位与跟踪^[48]。

4 基于全方位视觉的定位方法

鉴于传统的摄像机视野范围存在局限,不能同时获取周围所有目标的位置,日本东京大学 Yagi 等人于1990年^[49,50]首次将双曲反射镜面的全方位视觉系统应用于移动机器人导航,随后基于全方位视觉系统的研究在世界范围内掀起了一股热潮,并取得了较好的成果。而国内于2003年才开始有关于全方位视觉技术的研究,起步较晚,目前仍处于基础研究阶段。目前全方位视觉传感器主要分为两类,一类为传统视觉传感器组成的全方位视觉传感器,另一类为反射镜面型全方位视觉传感器。前者一般采用图像拼接技术将各个传统视觉传感器获得的图像拼接成一个完整的全方位图像,然后进行相关的操作和处理。此类全方位视觉传感器需要对大量的视觉信息进行处理,计算量大,难以达到实时性。反射镜面型全方位视觉传感器由一个或者几个曲面镜和一个传统的视觉传感器构成,通过对曲面形状的设计来控制全方位视觉传感器的视野范围。对于这类全方位视觉传感器,由于反光镜的镜面形状会造成全方位图像的分辨率分布不均匀且有一定的畸变,加上曲面反光镜的水平截面为圆形,造成全方位图像也为圆形,很难直接处理。虽然这种全方位视觉传感器需要解决图像分辨率低和图像畸变等问题,但其计算量小,不需要解决前者全方位视觉传感器的图像融合问题,因此被广泛应用于目标检测和视觉跟踪。如霍振华^[51]考虑到全方位图像几何特征不利于图像的分析与处理,提出了一种结合全方位图像展开和全方位图像鸟瞰图的分析方法,从全方位鸟瞰图中分割障碍物区域和地面区域,从全方位图像展开图像中进行人体检测,结合这两个处理结果得到人体目标在全方位鸟瞰图中的位置。最后通过大量的实验验证了所提方法的可靠性。

结束语 目标定位一直是计算机视觉领域的研究热点和难点,近年来国内外研究机构或学者对此进行了大量的研究,已取得了一定的发展,并提出了一些有意义的算法;胡占义等人通过对单幅图像设置人工标志的特征点来对目标进行定位;Liu 等人通过利用在单幅图像中人工标志的直线特征与摄像机光心形成的投影平面的法向量和目标物自身直线正交的几何特性,建立几何模型进行定位;鉴于胡占义等人提出的方法需要设置人工标志,制约了视觉定位的应用场合;王娟等提出了一种新的由二维投影图像来进行空间定位的方法,即利用目标上共面四线与其在二维图像上投影相匹配的测量来实现定位;张治国也提出一种基于两幅图像的单目视觉定位方法;考虑到单目视觉视野范围小,获取的深度信息少,导致定位精度低,Mammeri 等人提出了基于特征点的立体视觉定位方法;由于基于特征点的立体视觉定位方法不适用于图像特征贫瘠的场合,蒋焕煜等人提出基于区域立体匹配的双目视觉定位方法;项荣等人考虑到特征点匹配和区域匹配的特点,提出了一种基于组合匹配方法的双目视觉定位方法,用以指导番茄采摘机器人精准地采摘;随着微软推出新体感工具 Kinect, Wang 等人利用 Kinect 同步获得的彩色图像和深度图像识别楼梯、人行道和交通信号灯;鉴于传统摄像机视角狭小,不能同时得到附近所有目标的位置, Yagi 等人将双曲反射镜面的全方位视觉系统应用于移动机器人导航。上述方法

各有优劣,总体来说,基于单目视觉的定位方法精度较低,但结构简单,运动灵活,易于标定且实时性较好;基于传统的双目视觉的定位方法精度较高,但其运算量较大且匹配对应像素点时较困难;而基于全方位视觉的定位方法虽然视野范围广,但测量深度较复杂;随着微软推出视觉新工具 Kinect,基于 Kinect 工具的目标定位虽然能够获得高质量的深度图,但目前其深度图最大的识别距离大约为4m,因而阻碍了其对外场景中的目标或远距离目标的检测、定位与跟踪。

本文通过阅读有关基于视觉目标定位方面的重要文献,分析了目前基于视觉目标定位的有关方法,总结得出:实现两幅或多幅图像之间对应特征元素的匹配是基于视觉的目标定位算法的关键,同时更深入地研究人类视觉的立体融合和形成机制或者根据具体的应用设定一些约束条件和匹配模型是提高基于视觉的目标定位的精度和实时性的关键,也是基于视觉的目标定位进一步研究的方向。

参 考 文 献

- [1] Yun D. Machine vision[M]. Beijing: Science Press, 2000 (in Chinese)
云得. 机器视觉[M]. 北京: 科学出版社, 2000
- [2] Mobus R, Joos A, Kolbe U. Multi-target multi-object radar tracking[C] // Proceedings of Intelligent Vehicles Symposium, 2003. IEEE, 2003: 489-494
- [3] Wang J, Gao F, Shi S, et al. A New Method for Distance and Relative Velocity Measurement in Vehicle Collision Warning System[C] // Sixth International Conference on Intelligent Systems Design and Applications, 2006 (ISDA'06). IEEE, 2006, 1: 1142-1147
- [4] Chen Y, Wang C. Vehicle Safety Distance Warning System: A Novel Algorithm for Vehicle Safety Distance Calculating Between Moving Cars[C] // IEEE 65th Vehicular Technology Conference, 2007 (VTC2007-Spring). IEEE, 2007: 2570-2574
- [5] Lim Y, Lee C, Kwon S, et al. Distance estimation algorithm for both long and short ranges based on stereo vision system[C] // Intelligent Vehicles Symposium, 2008. IEEE, 2008: 841-846
- [6] Fishler M, Bollers R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395
- [7] Hu Z Y, Lei C, Wu F C. A short note on p4p problem[J]. Acta Automatica Sinica, 2001, 27(6): 770-776 (in Chinese)
胡占义, 雷成, 吴福朝. 关于 P4P 问题的一点讨论[J]. 自动化学报, 2001, 27(6): 770-776
- [8] Liu Y, Huang T, Faugeras O. Determination of camera location from 2-D to 3-D line and point correspondences[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(1): 28-37
- [9] Chen H. Pose determination from line to plane correspondences: existence solution and closed form solutions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 13(6): 530-541
- [10] Yu Y, Guo L. Target passive location method based on line feature[J]. Opto-Electronic Engineering, 2009, 36(1): 41-46 (in Chinese)
于勇, 郭雷. 基于特征直线的目标被动定位方法[J]. 光电工程, 2009, 36(1): 41-46

- [11] Han L W, Xu D. A Visual Dead Reckoning Positioning Method Based on Lines and Single Feature Point for Mobile Robot [J]. Robot, 2008 (1): 79-84 (in Chinese)
韩立伟, 徐德. 基于直线和单特征点的移动机器人视觉推算定位法[J]. 机器人, 2008 (1): 79-84
- [12] Qin L J, Hu Y L, Wei Y Z, et al. A algorithm for vision location based on plane quadrangle[J]. Transactions of Shenyang Ligong University, 2009, 28(2): 66-69 (in Chinese)
秦丽娟, 胡玉兰, 魏英姿, 等. 一种基于平面四边形的视觉定位算法[J]. 沈阳理工大学学报, 2009, 28(2): 66-69
- [13] Zhang L, Xu K J, Zhao R, et al. Improvement of position and orientation measurement algorithm of monocular vision based on circle features [J]. Journal of Hefei University of Technology (Natural Science), 2009, 32(11): 1669-1673 (in Chinese)
张磊, 徐科军, 赵锐, 等. 一种基于圆特征的单目视觉位姿测量算法的改进[J]. 合肥工业大学学报(自然科学版), 2009, 32(11): 1669-1673
- [14] Zhang G J, Zhou F Q. Position and Orientation Estimation Method for Landing of Unmanned Aerial Vehicle with Two Circle Based Computer Vision[J]. Acta Aeronautica Et Astronautica Sinica, 2005, 26(3): 344-347 (in Chinese)
张广军, 周富强. 基于双圆特征的无人机着陆位置姿态视觉测量方法[J]. 航空学报, 2005, 26(3): 344-347
- [15] Sun F M, Wang W N. Pose determination from a single image of a single parallelogram[J]. Acta Automatica Sinica, 2006, 32(5): 746-752 (in Chinese)
孙凤梅, 王卫宁. 基于单个平行四边形单幅图像的物体定位[J]. 自动化学报, 2006, 32(5): 746-752
- [16] Wang J, Zhao Y. Target spatialization algorithm based on single image[J]. Journal of Hebei University of Science and Technology, 2011, 32(4): 347-350 (in Chinese)
王娟, 赵艳. 基于单幅图像目标空间定位的算法研究[J]. 河北科技大学学报, 2011, 32(4): 347-350
- [17] Wang X J, Pan S L, Qiu L W, et al. Analytic algorithm of pose estimation based on two pairs of parallel lines[J]. Chinese Journal of Scientific Instrument, 2008, 29(3): 600-604 (in Chinese)
王晓剑, 潘顺良, 邱力为, 等. 基于双平行线特征的位姿估计解析算法[J]. 仪器仪表学报, 2008, 29(3): 600-604
- [18] Zheng Q X. Object recognition and localization based on monocular vision[D]. Jinan: Shandong University, 2013 (in Chinese)
郑全新. 基于单目视觉的目标识别与定位[D]. 济南: 山东大学, 2013
- [19] Ren Q Y, Li P, Han B. Vision-based mini unmanned helicopter pose and position estimation[J]. Journal of Zhejiang University (Engineering Science), 2009, 43(1): 18-22 (in Chinese)
任沁源, 李平, 韩波. 基于视觉信息的微型无人直升机位姿估计[J]. 浙江大学学报(工学版), 2009, 43(1): 18-22
- [20] Harris C, Stephens M. A combined corner and edge detector [C] // Alvey Vision Conference. 1988, 15: 50
- [21] Jiang Y, Xu Y, Liu Y. Performance evaluation of feature detection and matching in stereo visual odometry[J]. Neurocomputing, 2013, 120: 380-390
- [22] Luo G E. Some issues of depth perception and three dimension reconstruction from binocular stereo vision[D]. Changsha: Central South University, 2012 (in Chinese)
罗桂娥. 双目立体视觉深度感知与三维重建若干问题研究[D]. 长沙: 中南大学, 2012
- [23] Lindeberg T. Feature detection with automatic scale selection [J]. International Journal of Computer Vision, 1998, 30(2): 79-116
- [24] Lindeberg T. Edge detection and ridge detection with automatic scale selection [J]. International Journal of Computer Vision, 1998, 30(2): 117-156
- [25] Mikolajczyk K, Schmid C. Scale and affine invariant interestpoint detectors[J]. International Journal of Computer Vision, 2004, 60(1): 63-66
- [26] Cheng B S, Tang X W. Study of Harris scale invariant keypoint detector [J]. Journal of Zhejiang University (Engineering Science), 2009, 43(5): 855-859 (in Chinese)
程邦胜, 唐孝威. Harris 尺度不变性关键点检测子的研究[J]. 浙江大学学报(工学版), 2009, 43(5): 855-859
- [27] Meng H, Cheng K. Object location technique for binocular stereo vision based on scale invariant feature transform feature points [J]. Journal of Harbin Engineering University, 2009 (6): 649-652 (in Chinese)
孟浩, 程康. 基于 SIFT 特征点的双目视觉定位[J]. 哈尔滨工程大学学报, 2009(6): 649-652
- [28] Lowe D. Object recognition from local scale-invariant features [C] // The proceedings of the seventh IEEE International Conference on Computer Vision, 1999. IEEE, 1999, 2: 1150-1157
- [29] Lowe D. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110
- [30] Wang C, Sun Y. A new method of depth measurement with binocular vision based on surf [C] // Second International Workshop on Computer Science and Engineering, 2009 (WCSE'09). IEEE, 2009, 1: 568-571
- [31] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features [M] // Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417
- [32] Mammeri A, Boukerche A, Zhao M. Keypoint-based binocular distance measurement for pedestrian detection system [C] // Proceedings of the fourth ACM International Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications. ACM, 2014: 9-15
- [33] Bhowmick B, Bhadra S, Sinharay A. Stereo vision based pedestrians detection and distance measurement for automotive application [C] // 2011 Second International Conference on Intelligent Systems, Modelling and Simulation (ISMS). IEEE, 2011: 25-29
- [34] Ming Z H. Research on binocular stereo vision ranging algorithms [D]. Beijing: Beijing Institute of Technology, 2008 (in Chinese)
明祖衡. 双目立体视觉测距算法研究[D]. 北京: 北京理工大学, 2008
- [35] Zheng X D, Zhao J W, Liu M H. Tomatoes recognition and location from nature background based on binocular stereo vision [J]. Computer Engineering, 2005, 30(22): 155-156 (in Chinese)
郑小东, 赵杰文, 刘木华. 基于双目立体视觉的番茄识别与定位技术[J]. 计算机工程, 2005, 30(22): 155-156
- [36] Xiang R, Ying Y B, Jiang H Y, et al. Localization of tomatoes based on binocular stereo vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(5): 161-167 (in Chinese)
项荣, 应义斌, 蒋焕煜, 等. 基于双目立体视觉的番茄定位[J]. 农业工程学报, 2012, 28(5): 161-167

特别是对利用非关系数据库 HBase 的存储方式进行了深入研究,提出了基于 HBase 的本地存储模型 HBase-OntSM,其用一种稀疏多维排序的方式将数据存储到 HBase 数据库中,以少量的存储空间换取高效的查询效率;以西藏文化本体为例解释该存储模型及其存储过程;最后通过设计实验与其他模型进行对比来突出本文所提模型的优势,特别是对于大规模的本体,其可以将它分为多个图进行存储,这更能体现出性能优势。

参 考 文 献

- [1] Lv Yan-hui. Storage of Fuzzy Ontologies Based on Relational Databases[J]. Computer Science, 2011, 38(6): 217-222, 245 (in Chinese)
吕艳辉. 基于关系数据库的模糊本体的存储方法[J]. 计算机科学, 2011, 38(6): 217-222, 245
- [2] Abraham J, Brazier P, Chebotko A, et al. Distributed storage and querying techniques for a Semantic Web of scientific workflow provenance[C]// 2010 IEEE International Conference on Services Computing (SCC). IEEE, 2010: 178-185
- [3] Chebotko A, Abraham J, Brazier P. Storing, Indexing and Querying Large Provenance Data Sets as RDF Graphs in Apache

(上接第 16 页)

- [37] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J]. International Journal of Artificial Intelligence, 1995, 78(1/2): 87-119
- [38] Yu Q X. Research on mobile localization techniques for wheeled restaurant service robots[D]. Shanghai, Shanghai Jiaotong University, 2013 (in Chinese)
于清晓. 轮式餐厅服务机器人移动定位技术研究[D]. 上海: 上海交通大学, 2013
- [39] Kanade T, Kano H, Kimura S, et al. Development of a video-rate stereo machine[C]// 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems 95. IEEE, 1995, 3: 95-100
- [40] Jiang H Y, Peng Y S, Ying Y B. Binocular stereo vision applied to harvesting robots[J]. Journal of Jiangsu University (Natural Science Edition), 2008, 29(5): 377-380 (in Chinese)
蒋焕煜, 彭永石, 应义斌. 双目立体视觉技术在果蔬采摘机器人中的应用[J]. 江苏大学学报(自然科学版), 2008, 29(5): 377-380
- [41] Li H, Chen Y L, Chang T, et al. Binocular vision positioning for robot grasping[C]// 2011 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2011: 1522-1527
- [42] Guo B, Sun J, Wei Y, et al. Kinect identity: Technology and experience[J]. Computer, 2011, 44(4): 94-96
- [43] Dong J M, Chen W H, Yue H S, et al. Automatic recognition and location of tomatoes based on Kinect vision system[J]. Chinese Journal of Agricultural Mechanization, 2014, 35(4): 169-173 (in Chinese)
董建民, 陈伟海, 岳昊嵩, 等. 基于 Kinect 视觉系统的西红柿自动识别与定位[J]. 中国农机化学报, 2014, 35(4): 169-173
- [44] Figueroa J, Contreras L, Pacheco A, et al. Development of an

Hbase[C]// 2013 IEEE Ninth World Congress on Services, 2013: 1-8

- [4] Sun J, Jin Q. Scalable rdf store based on hbase and mapreduce [C]// 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). IEEE, 2010; V1-633-V1-636
- [5] Papailiou N, Konstantinou I, Tsoumakos D, et al. H2RDF: adaptive query processing on RDF data in the cloud[C]// Proceedings of the 21st International Conference Companion on World Wide Web. ACM, 2012: 397-400
- [6] Qiang Yan, Lu Jun-zuo, Liu Tao, et al. HBase Based Parallel BFS Method[J]. Computer Science, 2013, 40(3): 228-231 (in Chinese)
强彦, 卢军佐, 刘涛, 等. 基于 HBase 的并行 BFS 方法[J]. 计算机科学, 2013, 40(3): 228-231
- [7] protege[EB/OL]. <http://protege.stanford.edu>
- [8] Khadilkar V, Kantarcioglu M, Thuraisingham B. Jena-HBase: A Distributed, Scalable and Efficient RDF Triple Store [C] // CEUR Workshop Proceedings. 2012, 914: 85-88
- [9] Guo Y, Pan Z, Heflin J. LUBM: A benchmark for OWL knowledge base systems[J]. Web Semantics: Science, Services and Agents on the World Wide Web, 2005, 3(2): 158-182
- Object Recognition and Location System Using the Microsoft Kinect™ Sensor[M]// RoboCup 2011: Robot Soccer World Cup XV. 2012: 440-449
- [45] Zhang H, Yan R J, Zhou W S, et al. Binocular Vision Sensor (Kinect)-Based Pedestrian Following Mobile Robot[J]. Applied Mechanics and Materials, 2014, 670: 1326-1329
- [46] Nakano Y, Izutsu K, Tajitsu K, et al. Kinect Positioning System (KPS) and its potential applications[C]// International Conference on Indoor Positioning and Indoor Navigation. 2012
- [47] Wang S, Pan H, Zhang C, et al. RGB-D image-based detection of stairs, pedestrian crosswalks and traffic signs[J]. Journal of Visual Communication and Image Representation, 2014, 25(2): 263-272
- [48] Premevida C, Carreira J, Batista J, et al. Pedestrian detection combining RGB and dense LIDAR data[C]// 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014). IEEE, 2014: 4112-4117
- [49] Yagi Y, Yachida M. Real-time generation of environmental map and obstacle avoidance using omnidirectional image sensor with conic mirror[C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. IEEE, 1991: 160-165
- [50] Yamazawa K, Yagi Y, Yachida M. Obstacle detection with omnidirectional image sensor hyperomni vision[C]// Proceedings of 1995 IEEE International Conference on Robotics and Automation, 1995. IEEE, 1995, 1: 1062-1067
- [51] Huo Z H. Omnidirectional vision based human detection and localization in complex environments[D]. Beijing: Peking University, 2010 (in Chinese)
霍振华. 基于全方位视觉的复杂环境下人体目标检测与定位[D]. 北京: 北京大学, 2010