基于 CAD 模型的随机工件视觉识别和定位方法

王中任[1],路清彦[2],刘海生[1,3]

(1. 湖北文理学院 机械与汽车工程学院,湖北 襄阳 441053;
2. 武汉科技大学 汽车与交通工程学院,湖北 武汉 430081;
3. 汽车零部件制造装备数字化湖北省协同创新中心,湖北 襄阳 441053)

摘 要:获取工件目标位置和姿态的信息是实现工业机器人准确地抓取工件的关键。提出了一种用单目相机对随机堆放工件的目标进行识别和定位的方法。采用基于物体三维 CAD 模型的几何信息进行模板训练,对轮廓分明的工件可以实现高效的识别。结合高效的金字塔模型分层搜索策略,以及具有归一化相似性度量的匹配方式,保证了全局域范围内能训练出最合适的模板。测试目标对象是轮廓明显、具有反光特性的金属保持架工件,在随机堆放工件中进行了识别定位的实验。实验结果表明,用该方法可以准确识别随机堆放工件中的目标,目标工件在 X,Y,Z 轴方向上的平均定位误差为 0.948 mm、1.078 mm、2.175 mm。

关键词:识别; 定位; CAD 模型; 模板训练; 金字塔搜索
中图分类号:TP242.2 **文献标志码**:A **文章编号**:1007-2276(2015)S-0230-06

Visual recognition and location method of randomly placed parts based on CAD model

Wang Zhongren[1], Lu Qingyan[2], Liu Haisheng[1,3]

(1. School of Mechanical and Automotive Engineering, Hubei University of Arts and Science, Xiangyang 441053, China;
2. School of Automobile and Traffic Engineering, Wuhan University of Science and Technology, Wuhan 430081, China;
3. The Co-Innovation Center of Hubei Province of Auto Parts Manufacture Equipment Digitizing, Xiangyang 441053, China)

Abstract: Accurate recognition and location of randomly placed industrial parts is the key technology to realized intelligent bin-picking of industrial robot. An approach to recognition and location of target object in randomly placed industrial parts using monocular camera. Efficient recognition of industrial parts with distinct profile could be realized when the geometric information based on 3D CAD model is used to train the temple. The most appropriate temple was obtained after the combination of a pyramid approach with hierarchical model views and the matching mode characterize by normalized similarity measure. The tested objective is a metal couple cage with distinct profile and reflective feature. Some experiments were carried out for randomly placed parts. Experimental results show that the target can be recognized accurately in randomly placed parts. The average location errors of target parts in X, Y and Z direction are 0.948 mm, 1.078 mm and 2.175 mm.

Key words: recognition; location; CAD model; temple training; pyramid search

收稿日期:2015-10-15; **修订日期**:2015-11-26
基金项目:湖北省科技支撑计划(2015BAA034)
作者简介:王中任(1974-),男,副教授,博士,主要从事机器视觉技术方面的研究。Email:xfu_wangzhongren@126.com

0 引　言

随着智能装备的不断发展和成熟，机器人技术已成为工业领域智能化生产的主要生产力。传统的机器人分拣方式采用对机器人进行简单试教，对机器人和工件的姿态及摆放位置有严格要求，一旦姿态及位置发生改变,会极大地影响生产效率,甚至会导致生产损失。为了能有效解决这一问题,将机器视觉与机器人技术相结合，通过视觉使机器人作业更加柔性化、智能化,对于降低劳动成本、保证产品质量、提高生产效率具有重要意义[1-3]。

机器视觉技术作为“眼睛”引入到机器人技术中,能够使机器人具有分析和判断的能力,是实现“机器人换人”和“智能制造”的关键技术。国外,以 ABB 和 Fanuc 为代表的机器人公司,将机器视觉技术和工业机器人的结合已经日益成熟。ABB 公司将 Bubblebee 双目视觉系统集成到关节机器人上[1],能够实现对随机摆放工件的三维识别和抓取。参考文献[2]在机器人系统中搭建了一套 DLP 结构光视觉系统,通过投影格雷码和解码计算获取工件的三维信息，提取 3D 边缘用于模型匹配。日本 FANUC 公司的iRVision 视觉系统将机器人控制器与视觉系统相结合，不仅可以通过目标的 3D 姿态对料框中的工件进行定位,还可以有效地识别被不完全遮挡的工件[3]。国内杜平等[4]在研究飞机部位识别算法中提到了一种利用三维扫描物体获取三维模型、用 OpenGL 类库对三维模型进行处理后获得飞机各姿态下的二维投影的方法，但该方法只能针对局部信息且不能获取具体的姿态信息。李星云等[5]基于单目视觉系统,提出了一种在嵌入式平台上的单目视觉定位方法。在数据采集的过程中，分别让采集模块在不同高度对目标图像进行采集,以判断不同高度下的系统误差。在对目标姿态的测量过程中，选择了四种不同的角度进行定位测量,但只针对单一目标进行定位,没有涉及到多种工件混杂堆放的识别后的定位问题，难以适用于工业实际。

文中提出了一种基于三维 CAD 模型作为训练模板的方法，以单目相机下目标工件的平面信息为参考对象,匹配三维模型各个面的投影,这种方法不依赖其表面特征及光照对工件的影响，对工业生产中具有反光特性的金属零件,可以准确、快速识别并且确定目标工件在随机工件中任意位置和姿态下的三维坐标。

1 三维 CAD 模型模板匹配及定位原理

三维 CAD 模型匹配是一种基于投影的方法,对物体的三维模型在多视角及不同距离条件下进行投影,利用单目相机从工作平面上获取物体的图像,并将图像中目标物体的平面信息与三维模型各视角及不同距离下的投影进行相似性匹配,如图 1 所示。通过相机的标定获取相机内外参数,确定物体坐标系、相机坐标系和世界坐标系三者之间的位姿关系,通过坐标转换物体在世界坐标系下的三维姿态，其工作流程图如图 2 所示。

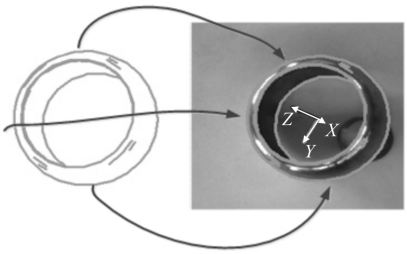


图 1 从三维 CAD 线框模型到目标工件的投影

Fig.1 Projection of CAD frame model to the target object

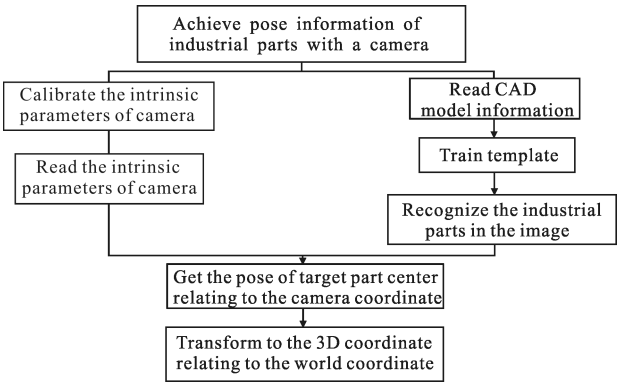


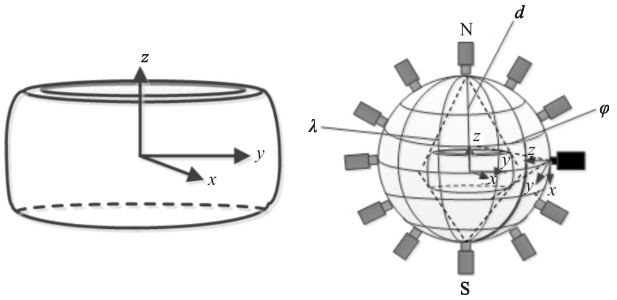
图 2 三维模型模板匹配及定位流程图

Fig.2 Flow chart of CAD model matching and location

1.1 利用物体三维模型训练模板

物体的三维模型可以由三维制图软件构建线框模型并以 DXF、OFF、STL 或者 PLY 格式储存,物体的三维轮廓(如图 3(a)所示)是由物体在三维空间中各姿态不同视角下的平面信息所构成的。这些二维

视图的获取可以设想成在三维空间中通过放置无数虚拟的相机以球体的形式围绕在物体三维模型周围(如图 3(b)所示),将三维模型投影到不同位置的虚拟相机中。各虚拟相机中生成的二维轮廓模型储存在形状模板中，通过在形状模板中的数据搜索匹配真实相机下物体的成像信息，进而识别图像中物体在某个视角下的三维姿态[5]。假设三维模型在一个被定义了地理坐标系的球体的中心，将用来获取物体二维视图的虚拟相机以指向球心的方式围绕物体，虚拟相机覆盖的范围可以被限定在球体表面的某一区域，这一区域的范围由地理坐标参数 λ (球面经度)、φ(球面纬度)及 d(三维模型原点到球体表面的距离)所决定。



(a) 物体的三维 CAD 模型 (b) 虚拟相机围绕球体的模型

(a) 3D CAD model of the object (b) Model of virtual camera around the sphere

图 3 投影模型

Fig.3 Projection model

1.2 确定模板搜索策略

搜索策略是指在搜索空间中寻找最大相似性以获得匹配位置的方法[6]。用三维模型训练的模板搜索图像中目标物体时,好的搜索策略可以提高整个图像匹配过程中的速率和增强匹配的鲁棒性[7]。图像是以金字塔形式排列的分辨率逐步降低的图像集合，金字塔底部是待处理图像的高分辨率表示,而顶部是低分辨率的近似。其原理是在图像金字塔的低分辨率图像上用模板匹配的方法寻找初始化匹配点，随着金字塔层数逐渐降低再逐层匹配,每层匹配点的搜索区域都以上层的匹配结果来设定搜索范围,依次循环直到获得较准确的匹配点位置。图 4 所示为球体表面上某一区域生成的模型视图，每个金字塔图层中，匹配点以黑色小三角形的形式表示。

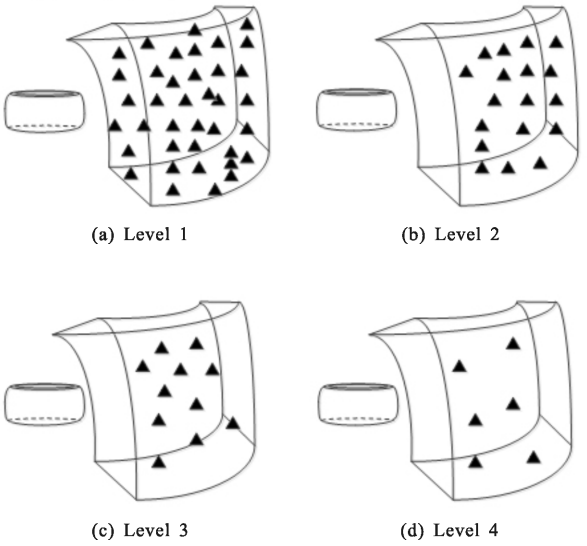


图 4 金字塔搜索模型

Fig.4 Pyramid search model

1.3 选定相似性度量

在图像中利用金字塔模型进行模板匹配时,相似性度量就是匹配点对之间的相似性比较的重要的衡量依据[8],相似度比较有以下三种常用的方式:平均平方差度量、相关匹配度量和归一化相关匹配。这三种常用的比较方式均具有其归一化的形式，在进行度量系数归一化的同时既减小了模板，又减小了噪声和图像光线变化的影响。减少噪声影响的原因是，即使在图像中进行模板搜索的过程中某一特征丢失,噪声虽然将导致一个随机的方向向量,但是这些向量不会对模板搜索的整个过程造成重要影响。对于各向同性像素点均可以设置其方向向量为零向量,仅保留各向异性的像素点,并计算其归一化方向向量，这样遮挡和混乱构成的特征就不会对综合造成影响。使其归一化的另一优势是减小了光线的影响,方向向量的模取决于图像的亮度。将这个相似度量归一化,所有的方向向量都为 1,这样任意光照变化影响就不会影响相似度量。

1.4 三维物体定位

用物体的三维模型创建模板可以仅用一台相机识别物体在三维空间中相对于世界坐标系的位姿。三维模型是按实际比例 1:1 创建的，利用图像金字塔搜索策略及具有归一化形式的相似性度量算法识别出与图像中物体相似度高的模板。在图像中,模板的三维轮廓投影到物体轮廓后用 Levenberg –

Marquardt[9]非线性优化算法进行多次增强鲁棒性的迭代,得到模板轮廓(线框模型)与物体轮廓完全啮合,啮合后通过三维模型和物体在图像中的几何关系,可以知道物体在图像中的三维姿态。利用标定相机所得的内参、模板实际尺寸和在图像中缩放的比例以及三维模型在图像中的三维姿态,求取物体原点与光轴中心之间的对应关系,获得物体原点相对相机坐标系 6 个自由度的坐标,标定相机内参的公式如下:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & k_s & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1/z_1 \\ y_1/z_1 \\ 1 \end{bmatrix} = M_{\text{in}} \begin{bmatrix} x_1/z_1 \\ y_1/z_1 \\ 1 \end{bmatrix} \tag{1}$$

式中:M_{in} 为相机的内参;k_s 为图像的径向畸变系数;k_x、k_y 为每个像元的物理尺寸的长、宽;u_0,v_0 为在图像中心的像素坐标;x_1,y_1,z_1 分别为图像中(u,v)点相对相机坐标系的实际坐标,z_1 等于相机的焦距 f。

图像中物体原点的坐标为(u_0,v_0),将三维模型训练出的模板投影在图像平面的目标工件上,通过标定的内参所获得的参数 M_{in},获得成像平面上工件中心相对于相机坐标系的坐标(x_0,y_0,z_0)。根据模板投影在图像平面上的三维模型,可以求得物体实际中心和成像平面上物体中心的位姿关系,通过空间坐标齐次变换公式,求得物体相对相机坐标系的坐标(x,y,z)。

2 视觉定位系统中各坐标系的关系

对各种姿态的物体进行识别定位时,得到物体的姿态是相对于以相机坐标系 CCD 芯片中心为原点的姿态信息,但是在机器人抓取物体的过程中,由于 CCD 芯片安装在相机内部,以芯片中心为原点建立参考坐标系外不易获取,所以除了相机坐标系、物体坐标系外还需要建立一个以工作平面上某点为参考位姿的世界坐标系,方便机器人根据视觉定位的信息进行精确抓取。

2.1 确立世界坐标系相对相机坐标系的位置

要得到目标相对于世界坐标系的坐标,首先需要对相机进行标定,获得相机相对工作平面的内、外参数。确定世界坐标系坐标原点的方法:设置平面上某一固定参考物为坐标原点,如图 5 所示。

确定世界坐标系后,以标定板坐标系的原点为参考点和图像上该原点的坐标,确立图像坐标系和相机坐标系之间的变换矩阵 ${}^{c}H_{\text{cp}}$。因为在相机所拍摄的图像中,世界坐标系和标定板的坐标系在相机坐标系下位置关系已知,可以求取两者之间的变换矩阵 ${}^{w}H_{\text{cp}}$。再利用矩阵之间的变换关系如公式(2),求取世界坐标系 $O-X_wY_wZ_w$ 在相机坐标系 $O-X_cY_cZ_c$ 下的位姿关系 ${}^{c}H_{\text{w}}$。

$${}^{c}H_{\text{w}} = {}^{c}H_{\text{cp}} \cdot {}^{\text{cp}}H_{\text{w}} = {}^{c}H_{\text{cp}} \cdot ({}^{\text{w}}H_{\text{cp}})^{-1} \tag{2}$$

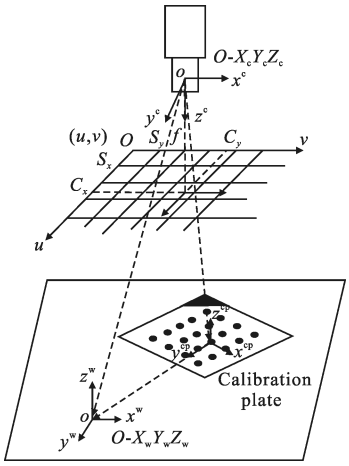


图 5 确定世界坐标系的位姿

Fig.5 Determination of the pose of world coordinate system

2.2 确定物体坐标系相对于世界坐标系的位置关系

获取物体中心坐标系在相机坐标系下的坐标(x_c,y_c,z_c)后,需要将相机坐标系下物体的中心坐标(x_c,y_c,z_c)转换成世界坐标系下的坐标(x_w,y_w,z_w),利用齐次矩阵 ${}^{c}H_{\text{w}}$ 求取物体坐标系在世界坐标系 $O-X_wY_wZ_w$ 下的变换矩阵 ${}^{w}H_{\text{o}}$,如图 6 所示。

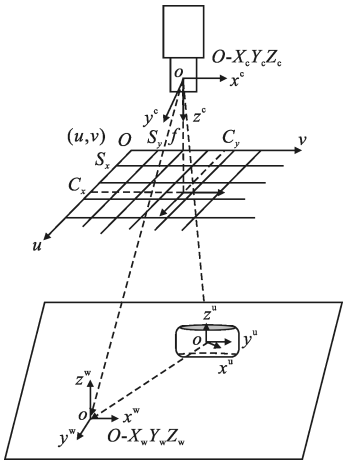


图 6 相机坐标系、物体坐标系及世界坐标系之间的关系

Fig.6 Relationship of camera coordinates, object coordinates and world coordinates

世界坐标系 $O-X_wY_wZ_w$ 和物体坐标系 $O-X_oY_oZ_o$ 相对于相机坐标系 $O-X_cY_cZ_c$ 的关系 ${}^cH_w$、${}^cH_o$ 已知，可以通过齐次矩阵变换将物体的中心坐标转换到世界坐标系 $O-X_wY_wZ_w$ 下的位置关系 ${}^wH_o$，如公式(3)：

$$ {}^wH_o={}^wH_c\cdot{}^cH_o=({}^cH_w)^{-1}\cdot{}^cH_o \tag{3} $$

3 散乱堆放工件中识别物体定位实验

实验选用的相机型号为Basler acA1300-60gm，分辨率 1 280 pixel×1 024 pixel。镜头为 PENTAX TV LENS 镜头，焦距 12.5 mm。选用的目标工件是汽车前轮中轴承的保持架部件，它是一个外表面形状为球形曲面的环状金属件，表面光滑无纹理，具有一定的反光特性。为了识别工件轮廓，工作平面选择以白色为底，可以与工件金属色泽明显区分开来，相机的视野范围为 440 mm×355 mm，可以摆放 6~8 个以不同姿态摆放的目标工件，世界坐标系的参考点是以工作平面右下角黑色矩形的中心为原点建立的坐标系，世界坐标系的 X_wY_w 平面与相机坐标系 X_cY_c 平面平行，实验现场如图 7 所示。

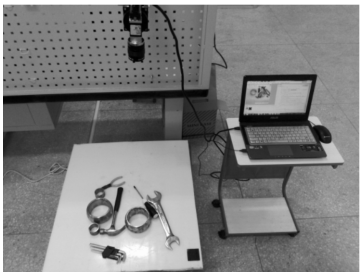


图 7 实验场景

Fig.7 Experimental scene

为了测试识别工件在复杂环境下的稳定性，在单一的环境中加入一些其他的干扰，如：物体堆叠摆放、加入多个相同目标物体、添加其他工具作为干扰因素，定位测试的结果如图 8 所示。

物体最终的位姿是物体中心相对于世界坐标系的位姿关系，下列 6 组数据(表 1)是以图像上工作平面上与相机光轴交点为原点建立的世界坐标系，测量物体中心与世界坐标系之间 6 个方向的位姿。

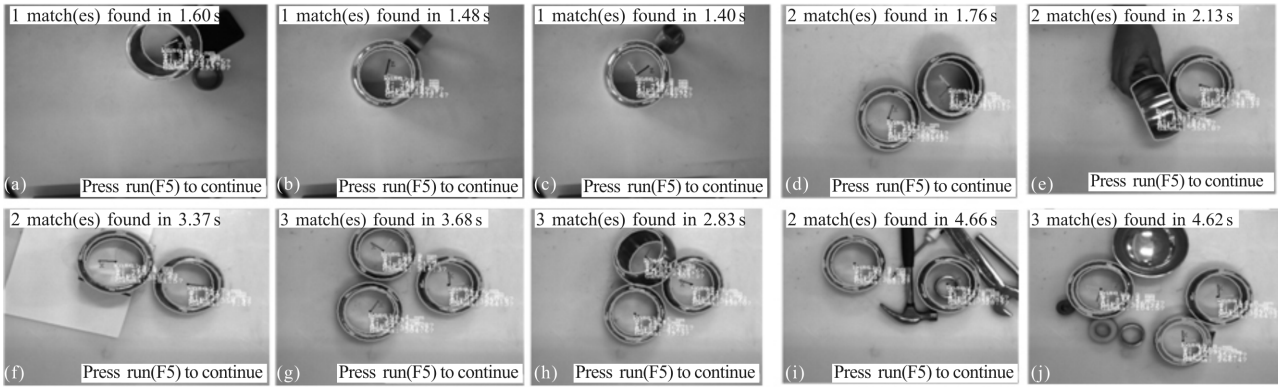


图 8 随机摆放工件的目标识别

Fig.8 Object recognition in randomly placed industrial parts

表 1 测量数据与实际数据的对比(单位：mm)

Tab.1 Comparison of measured data and actual data (Unit：mm)

No.	X		Y		Z	
	Measured value	Real value	Measured value	Real value	Measured value	Real value
1	−2.13835	−3.00000	−4.13984	−3.00000	25.47030	27.71000
2	5.45442	4.00000	−3.32152	−4.00000	24.24070	27.51000
3	1.01735	1.40000	−4.01354	−4.50000	67.12420	70.12400
4	4.21687	4.75000	−7.57261	−6.35000	76.36920	75.32000
5	2.84718	4.21000	−9.18935	−7.81000	82.55390	80.15000
6		2.43000	−4.61349	−3.05000	81.36030	80.27000

将工件随机摆放 6 个不同姿态，将视觉定位数据(如图 9 所示)与物体相对世界坐标系的实际位姿(表 1)进行比较，并列入表 1。

经计算，工件在 X 轴方向上最大定位误差为 1.454 mm，平均定位误差为 0.948 mm。工件在 Y 轴方向上的最大定位误差为 1.563 mm，平均定位误差为 1.078 mm。工件在 Z 轴方向上的最大定位误差为 3.269mm，平均定位误差为 2.175 mm。可见，工件在 Z 轴方向上的定位精度明显低于在 X、Y 轴方向。由公式(1)可知，相机内参焦距 f 与工件相对相机坐标系下 Z 轴方向上的数值有关，所以相机内参中焦距 f

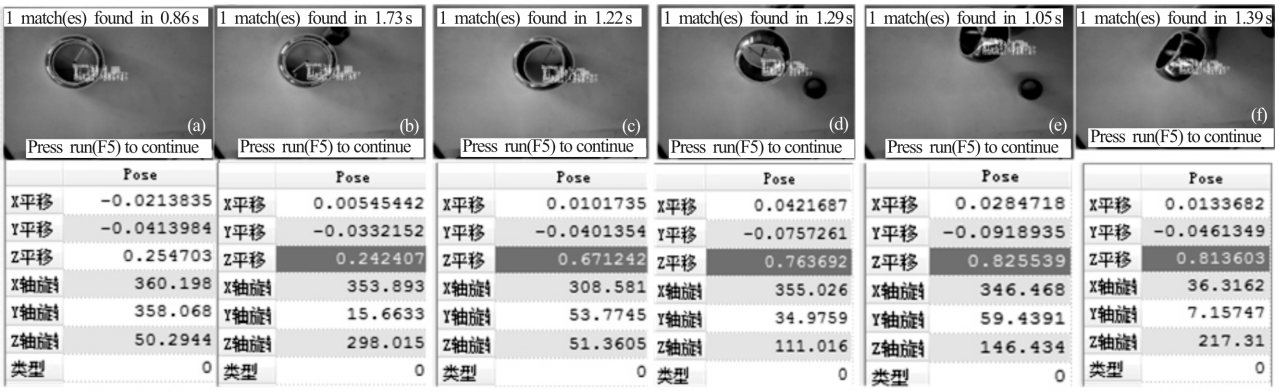


图 9 六种位姿的物体定位

Fig.9 Location of the object under six postures

标定的准确性直接影响了工件 Z 轴方向上的定位精度。因为工件最后的位姿是相对于世界坐标系下的姿态，所以 Z 轴精度和工作平面上世界坐标系的位姿也有关系，由公式(2)~(8)可知，世界坐标系相对于相机坐标系下的位姿影响到工件在世界坐标系下的姿态，所以对世界坐标系参考点的选取也影响了 Z 轴方向上的定位精度。

4 结 论

文中提出了一种用单目相机对随机堆放工件进行识别和定位的方法。实验结果表明，采用的 CAD 三维模型创建模板进行匹配的方法在物体任意姿态、任意位姿下，能够准确地对图像中的物体进行匹配，并可以在随机堆放及复杂环境中对目标物体进行准确识别，具有较高的稳定性和可靠性。目前在识别精度上存在一定的偏差，可能与图像质量、算法优化及标定结果有关，在优化这些因素后可以减小定位的误差范围，调高定位精度。

参考文献：

[1] Carlos Martinez, Remus Boca, Biao Zhang, et al. Automated bin picking system for randomly located industrial parts[C]// IEEE International Conference on Technologies for Practical Robot Applications(TePRA), 2015.

[2] Jong-Kyu Oh, Kyeong Keun Baek, Daesik Kim, et al. Development of structured light based bin picking system using primitive models[C]//IEEE International Symposium on Assembly and Manufactunng, 2009.

[3] Shanghai-FANUC Robotics Co.Ltd. FANUC intelligent robot loading and unloading for machine tool [J]. ***Servo Control***, 2013(2): 94-97. (in Chinese)

[4] Du Ping, Xu Dawei, Liu Chongqing. Research of simulation of aircraft status [J]. ***Infrared and Laser Engineering***, 2000, 29(5): 1-4. (in Chinese)

[5] Li Xingyun, Li Zhongli, Liao Xiaobo. Design on industrial robot positioning system based on monocular vision [J]. ***Machine Tool & Hydraulics***, 2015, 43(9): 35-38. (in Chinese)

[6] Zhang Kaixing, Zhang Shushen, Liu Xianxi. Current research and future development of 3-D CAD model retrieval [J]. ***Transactions of the Chinese Society for Agricultural Machinery***, 2013, 44(7): 256-263. (in Chinese)

[7] Christian Wiedemann, Markus Ulrich, Carsten Steger. Recognition and tracking of 3D objects [C]//Proceeding of Pattern Recognition, 30th DAGM Symposium, 2008.

[8] Chen Weibing. Comparison of matching capabilities in similarity measurements[J]. ***Journal of Computer Applications***, 2010, 30(1): 98-100, 110. (in Chinese)

[9] Li Chenglong, Zhong Fan, Qin Xueying. Model-based irregular object localization from a single view image [J]. ***Journal of Computer-Aided Design & Computer Graphics***, 2015, 27(1): 69-75. (in Chinese)