

DOI:10.19651/j.cnki.emt.2211745

基于双目视觉的棒材端面中心点定位^{*}

刘 昶 姜 楠

(沈阳理工大学信息科学与工程学院 沈阳 110159)

摘要: 在利用标签焊接机器人为捆装棒材焊接标签时,需要向其提供一组棒材端面中心点的三维坐标,针对从捆装棒材中选择并定位棒材端面中心点的问题,提出一种基于双目立体视觉的方法。该方法利用双目视觉系统的虚拟像平面模型,通过将平面标定板摆放在平行于捆装棒材端平面的位置进行两台相机的外参标定,根据标定结果生成目标在该标定平面上的虚拟图像;采用 SVM 和连通区的方法提取虚拟图像中棒材端面中心点特征;采用极线约束和中心点共面约束的方法进行特征匹配,向机器人提供一组端面中心点三维坐标。仿真实验结果显示提供给机器人的推荐点均来自正确匹配点对,说明本文提出的特征匹配算法有效;真实实验中推荐点深度位移最大误差为 0.20 mm,平均误差为 0.09 mm,说明本文提出的棒材中心点特征提取方法有效。

关键词: 双目视觉;相机模型;支持向量机;极线约束;特征匹配

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Bar bottom center localization based on binocular vision

Liu Chang Jiang Nan

(School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract: An automatic tagging robot needs to be provided the 3D coordinates of a bar's bottom center to weld a tag on the bundle. A method based on binocular vision was proposed to select and localize bars' bottom centers. Virtual image camera model was adopted in the binocular vision system. The extrinsic parameters of two cameras were calibrated from a planar calibration pattern which was put parallel to the common end plane of bars. The virtual images of two cameras were created according to the calibrated results. A method using SVM and connected region was adopted to extract the center point features of bars in both virtual images from two cameras. A group of candidate features pairs were selected using epipolar constraint to the features and coplanar constraint to the recovered physical points. The 3D coordinates of corresponding physical points were recommended to the robot to try to weld the tag. Simulation results showed all recommended points to the robot were from correct matched pairs. It demonstrated effectiveness of the features matching method presented. In real experiment, the maximum depth displacement error of the recommended points was 0.20 mm, the average error was 0.09 mm. It demonstrated the effectiveness of bar bottom center extraction method presented.

Keywords: binocular vision; camera model; SVM; epipolar constraint; feature matching

0 引 言

钢厂等企业生产的棒材在入库或出厂前,需要对捆装棒材挂标签,目前多数钢厂采用人工焊接标签的形式,具有效率低、危险性高、劳动强度大、成本高等问题,采用机器人搭载焊接设备实现标签的自动焊接可有效解决这些问题,为使这种标签焊接机器人能将标签准确地焊接到棒材的端面,需要从捆装棒材中选择若干棒材并将它们的端面的中

心点在基准坐标系下的三维坐标提供给机器人,引导机器人焊接。考虑到计算机视觉技术在目标测量和定位方面具有精度高、非接触等优势,因此利用计算机视觉技术实现棒材端面中心点的定位是一种有效的手段。

一般的视觉定位方法主要包括基于单目视觉和基于双目视觉的方法,其中单目视觉的方法一般需要合作目标配合,不适合捆装棒材这种缺乏明确合作目标的定位问题;双目视觉方法是通过两台相机分别拍摄目标图像,通过特征

收稿日期:2022-10-19

^{*} 基金项目:辽宁省教育厅高等学校基本科研重点项目(LJKZ0243)资助

提取和匹配实现空间点的三维测量^[1-2]。双目视觉系统中常用的特征提取和描述的方法有尺度不变特征转换(scale-invariant feature transform, SIFT)^[3]、加速稳健特征(speed up robust feature, SURF)^[4]、加速特征检测算法(features from accelerated segment test, FAST)、BRIEF(binary robust independent elementary features)^[5]、ORB(oriented fast and rotated brief)^[6-7]等,此类特征通用性强,但对于捆装棒材这种具有大量重复性目标的图像缺乏针对性。目前针对棒材端面中心点特征提取的主要方法有基于连通区的方法、Hough 变换的方法、模板匹配方法等。Yan 等^[8]提出采用连通区的方法提取棒材端面特征,并应用于棒材计数问题;刘晓玉等^[9]采用形态学的腐蚀方法,将棒材端面连通区域不断缩减实现棒材分离,最后将分离的端面连通区处理为单个像素,记为中心点,但这种方法不适用于当棒材端面存在相互遮挡的场景;Zhao 等^[10]采用边缘检测提取棒材端面的边缘点,采用 Hough 变换的思想,以边缘点为圆心构建圆的方式向棒材端面圆心投票提取棒材中心点,这类方法依赖棒材端面边缘提取的结果,当棒材端面的边缘信息提取不好时,容易导致棒材中心的提取误差变大;侯维岩等结合形态学、连通区的方法对棒材图像进行预处理,再采用十字形模板确定可能的棒材中心^[11],模板匹配通常采用二值图与模板进行比较,对于存在棒材遮挡、照明差异等情况下,图像的二值化可能损失有效信息,会导致棒材中心的提取误差变大。

双目视觉系统中应用的特征点匹配算法主要有基于灰度的特征点匹配方法、基于特征描述的匹配方法、基于几何关系的匹配算法等。灰度特征点匹配方法主要有像素差绝对值匹配(sum of absolute differences, SAD)、互相关匹配(normalized cross correlation, NCC)等算法,张一飞等^[12]提出融合边缘特征的 SAD 匹配算法,用边缘特征图进行立体匹配,该方法对边缘较多的图像效果好;吴鹏等^[13]提出结合小波金字塔的快速 NCC 图像匹配算法,利用分层匹配提高图像的匹配效率,但这类基于灰度的匹配方法不适用于具有相似弱纹理棒材端面中心特征匹配。基于特征描述的匹配方法主要基于 SIFT、SURF、ORB 等特征进行匹配,由于棒材定位问题主要依赖棒材中心点特征,而此类特征提取结果中往往包括大量无关特征,因此并不适用于捆装棒材图像。Elashry 等^[14]提出基于图像的几何约束进行特征点的匹配算法,利用特征点间的角度和距离等几何特征对提取的特征点进行匹配,提高了匹配的准确率;Shen 等^[15]提出利用仿射协方差进行特征点匹配,实验表明具有良好的鲁棒性,但从测量的角度考虑,双目相机基线距离大更有利于减少深度测量误差,这时双目视角的差异将导致立体目标上特征点间的几何不变性假设不易满足。

本文采用虚拟像平面相机模型的双目视觉系统,通过特别选择的虚拟像平面,由左、右相机的原始图像生成在虚拟像平面上的虚拟图像,通过支持向量机(support vector

machine, SVM)结合连通区的方法提取虚拟图像中棒材端面中心点特征,利用极线约束和棒材端面中心点近似共面的约束进行特征点匹配,获取最佳的匹配对,将对应的空间点的三维坐标提供给机器人使用。

1 双目视觉系统的虚拟像平面模型

1.1 虚拟像平面相机模型

如图 1 所示,设空间中物理点 P 在像平面上的无畸变投影点为 q , 受镜头畸变的影响,其真实有畸变投影点为 q' ; 在一个标定平面 π 上建立世界坐标系 $O-xyz$, 光心 O_c 和像点 q 的连线(即投影射线)与平面 π 交于点 P' , 根据透视投影关系可知, P 、 P' 、 q 和 O_c 共线;若将平面 π 看作是相机的虚拟像平面, P' 看作是 P 的虚拟像点,则目标在虚拟像平面上的投影也会形成一幅图像,称之为虚拟图像;虚拟图像与光心构成相机的虚拟像平面相机模型^[16], 该模型是一个无畸变的针孔模型,由光心 O_c 和虚拟像点 P' 的连线确定 P 的投影射线,光心 O_c 到虚拟像平面的垂直投影点 S 为虚拟主点,光心 O_c 到虚拟像平面的距离为虚拟焦距。

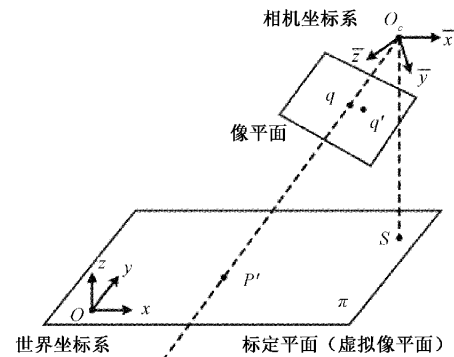


图 1 虚拟像平面相机模型

在相机参数和镜头畸变参数已标定的条件下,虚拟图像可由原始有畸变图像生成。记世界坐标系与相机坐标系 $O_c-\bar{x}\bar{y}\bar{z}$ 间的变换关系为:

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \mathbf{T} \quad (1)$$

其中, \mathbf{R} 和 \mathbf{T} 是坐标变换的旋转矩阵和平移向量。记 P' 点的世界坐标为 (x, y) , 无畸变像点 q 的图像坐标为 (u, v) , 则它们满足:

$$t \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{K} \begin{pmatrix} r_1 & r_2 & \mathbf{T} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2)$$

其中, \mathbf{H} 是虚拟像平面与像平面间的单应矩阵, \mathbf{K} 为相机内参矩阵, r_1 和 r_2 为 \mathbf{R} 的左两列, t 为比例因子。虚拟图像的生成过程是在虚拟像平面上划分等间距网格,以网格点作为虚拟图像的像素点,网格点图像坐标 (\bar{u}, \bar{v}) 与

对应的世界坐标 (x, y) 间的关系为:

$$\begin{cases} x = x_0 + \tilde{u} \cdot \delta \\ y = y_0 + \tilde{v} \cdot \delta \end{cases} \text{ 或 } \begin{cases} \tilde{u} = (x - x_0) / \delta \\ \tilde{v} = (y - y_0) / \delta \end{cases} \quad (3)$$

其中, (x_0, y_0) 为虚拟图像左上角点对应的世界坐标, δ 为网格在两个方向的间距, 为虚拟图像的像素距。式(3)建立了虚拟图像的像素点 (\tilde{u}, \tilde{v}) 与世界坐标 (x, y) 的对应关系。为了获取 (\tilde{u}, \tilde{v}) 的灰度值, 首先利用式(3)计算虚拟像点对应的世界坐标, 再利用式(2)计算对应的原始图像的无畸变像点坐标, 再利用镜头畸变模型和参数计算对应的有畸变像点坐标, 最后利用双线性插值方法获取该点灰度的估计值作为 (\tilde{u}, \tilde{v}) 点的灰度值。在获取了每个格点的灰度值后, 虚拟图像生成完成。

从以上过程可以看出, 虚拟图像同时实现了对原始图像上的镜头畸变校正和几何校正, 能真实地反映出目标在虚拟像平面上的投影情况; 相机的虚拟像平面模型在虚拟图像的像点坐标与世界坐标间建立了更直接的联系, 更有利于问题的分析。该模型在实际应用时需事先完成对相机内、外参及镜头畸变参数的标定, 一般可用张氏标定法等常规方法实现。

1.2 虚拟像平面双目视觉模型

对于双目视觉系统, 若针对每台相机均采用虚拟像平面模型, 并为它们选择同一平面作为虚拟像平面, 则构成虚拟像平面双目视觉模型。如图 2 所示, 设左、右相机光心分别为 O_L 和 O_R , 对应虚拟主点分别为 S_L 和 S_R , 物点 P 在公共的虚拟像平面 π 的投影点分别为 q_L 和 q_R 。虚拟像平面双目视觉模型下的投影关系具有以下性质:

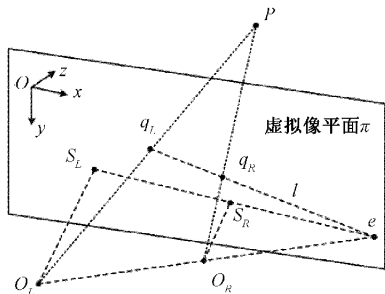


图 2 虚拟像平面双目视觉模型

性质 1) 若左、右相机的虚拟焦距不同, 则光心 O_L 和 O_R 的连线与虚拟像平面 π 的交点 e 为有限点, 为双目视觉系统的虚拟极点。

性质 2) 若虚拟极点 e 为有限点, 则 e 与两相机的主点 S_L 和 S_R 共线, 且任意物点 P 的两个虚拟像点 q_L 和 q_R 与 e 共线。

性质 3) 若左、右相机的虚拟焦距相同, 虚拟极点 e 为 $S_L S_R$ 方向的无穷远点, 任意物点 P 的两个虚拟像点 q_L 和 q_R 的连线与两个虚拟主点 S_L 和 S_R 的连线平行。

性质 4) 物点 P 必位于左、右相机光心与对应的虚拟像点确定的投影射线上, 即 P 为 $O_L q_L$ 和 $O_R q_R$ 的交点。

以上性质说明虚拟像平面双目视觉模型下仅存在一个极点; 性质 2) 体现了在有限极点条件下的极线约束性质; 性质 3) 体现了在无穷远极点条件下的极线约束性质; 性质 4) 用于根据虚拟像点匹配对计算物点的三维坐标, 考虑到实际应用时存在误差, 导致两条投影射线可能不相交, 一般选两条投影射线的公垂线段的中点作为物点的估计。

1.3 虚拟像平面双目视觉系统标定

本文针对捆装棒材中棒材端面中心点的定位问题, 采用双目视觉系统的方案, 其中两台相机分别采用传统的张氏标定法进行标定, 标定板选择平面棋盘格标定板。考虑到棒材在捆装前都进行过端面对齐的操作, 因此可以认为, 成捆棒材中所有棒材的端面中心点近似共面, 称该平面为棒材的端平面。但由于棒材的对齐过程不够精准, 棒材运输过程中有滑动等因素, 可能导致真实的棒材端面中心点并不严格在端平面上, 本文假设棒材中心点到端平面的距离服从正态分布。在进行标定时, 特别选取平行于棒材端平面的一个位置摆放标定板, 该平面一般与地面垂直, 与运输棒材的轨道平行, 将此时的标定板平面选作虚拟像平面。根据标定板上棋盘格角点的位置定义世界坐标系, 以标定板平面上左上方角点作为世界坐标系原点, 沿水平方向向右为 x 轴, 垂直方向向下为 y 轴, 再按右手规则确定 z 轴。

在确定了两台相机外参后, 由式(1), 左相机光心在世界坐标系下的坐标为:

$$(x_L, y_L, z_L)^T = -R_L^{-1} T_L \quad (4)$$

其中, R_L 和 T_L 是左侧相机坐标系与世界坐标系变换的旋转矩阵和平移向量。同理计算出右侧相机在世界坐标系下的坐标 (x_R, y_R, z_R) 。若两相机的虚拟焦距不相等, 世界坐标系下虚拟极点的坐标 $e(x_e, y_e, z_e)$ 为:

$$\begin{cases} x_e = (x_L z_R - x_R z_L) / (z_R - z_L) \\ y_e = (y_L z_R - y_R z_L) / (z_R - z_L) \\ z_e = 0 \end{cases} \quad (5)$$

2 棒材端面中心点特征提取与匹配

2.1 基于 SVM 的棒材端面中心点特征提取

本文仅针对端面为圆形的棒材进行处理, 采用基于方向梯度直方图 (histogram of oriented gradients, HOG) 特征^[17]和 SVM 的方法识别虚拟图像中棒材端面的中心区域, 再利用连通区方法提取端面中心点坐标, 具体步骤如下:

1) 棒材端面的规格化

若相机的真实像平面不与棒材的端平面平行, 则拍摄的图像中棒材圆形端面的像表现为椭圆, 但通过本文设定的与端平面平行的虚拟像平面, 不论两台相机的摆放角度如何, 生成的虚拟图像中的棒材端面都被恢复成圆形。考虑到虽然同一捆棒材的直径相同, 但不同批次的棒材的直径可能不同, 因此虚拟图像中棒材端面的尺寸也会不同。为便于应用机器学习方法进行特征提取, 先利用边缘提取

方法提取虚拟图像中的边缘点,再利用 Hough 变换提取圆特征,估计出图像中该捆棒材的直径 r ,再通过图像缩放,使缩放后的棒材直径为 r_s ,实现棒材端面尺寸的规格化,本文取 $r_s = 36 \text{ pixel}$ 。

2) 提取棒材端面 HOG 特征

针对缩放后的虚拟图像,提取图像的 HOG 特征。提取过程中,检测窗口的尺寸设为 $40 \text{ pixel} \times 40 \text{ pixel}$;每个 Cell 的尺寸为 $4 \text{ pixel} \times 4 \text{ pixel}$,将梯度方向划分 9 个区间,区间的角度间隔为 20° ,这样每个 Cell 对应一个 9 维向量描述的方向梯度直方图;将窗口中上下左右相邻的 4 个 Cell 组合为 1 个 Block,将 Block 中所有 Cell 的特征向量依次连接起来,构成 36 维的 Block 特征向量;再将窗口中所有 25 个 Block 的标准特征向量依次连接起来,得到窗口图像的 900 维的 HOG 描述子。

3) 构建 SVM 分类器模型

通过大量采集棒材图像,经步骤 1) 规格化后,通过人为标记棒材中心点的方式,以标记点为窗口中心,构造 40×40 像素大小的棒材端面样本图像,再根据步骤 2) 提取其 HOG 特征作为正样本,本文构建的正样本集共包含 3 500 个样本。

从规格化后的虚拟图像中随机选择像素点,保证其到已标注棒材中心点的距离大于 $0.7r_s$,以这些点作为中心点建立 40×40 像素的窗口,提取 HOG 特征作为负样本,本文构建的负样本集共包含 6 500 个样本。

本文使用具有高斯核函数的 SVM 分类器模型,应用棒材端面样本的 HOG 特征向量进行训练,构建棒材端面中心点 SVM 分类器,高斯核函数的形式为:

$$K(\mathbf{h}_1, \mathbf{h}_2) = \exp\left(-\frac{\|\mathbf{h}_1 - \mathbf{h}_2\|^2}{2\sigma^2}\right) \quad (6)$$

其中, \mathbf{h}_1 和 \mathbf{h}_2 为样本的 HOG 特征向量,本文取超参数 $\sigma = 3.8$ 。利用构造的样本集对 SVM 分类器进行训练,所得分类器用于判别检测窗口是否包含棒材。

4) 棒材端面中心点特征提取

针对待检测的虚拟图像,先对其规格化,再采用滑动窗口的方法检测棒材中心点,检测中窗口尺寸为 40×40 像素,步长为 2,应用训练好的 SVM 分类器对每个待检测窗口进行分类,根据分类结果对窗口中心的 4 个点进行标记,如分类结果为正,则标记为 1,否则标记为 0;待整幅图像检测完成后,图像中标记为 1 的像素将形成一些连通区,这些连通区大多位于棒材端面的中心部分;再应用连通区检测方法获取所有连通区,设置超参数 T_s ,对于每个连通区,若其面积小于 T_s 则认为其为干扰或误识别点不做处理,否则计算该连通区的重心作为一个端面中心点;对所有连通区处理结束后,再对图像进行规格化中缩放的逆操作,恢复所提取的中心点在规格化前图像中的坐标值,作为棒材端面中心点特征提取的结果,最后由式(3)计算出该点的世界坐标。

2.2 基于极线约束的特征点预匹配

受相机视角和棒材位置影响,相机拍摄的捆装棒材图像中可能存在棒材的遮挡,导致在利用 2.1 节的方法提取棒材端面中心点时,有可能出现部分棒材的中心点出现一定程度偏差;另外由于可能出现漏检,从左、右两幅虚拟图像中提取出的中心点数量也可能不同。为计算某根棒材中心点的三维坐标,需要对从两幅图像中提取的中心点特征进行匹配,本文首先利用极线约束对这些特征点进行预匹配。记从左、右两幅虚拟图像中提取出的棒材中心点特征点集分别为 $U_L = \{q_i^L \mid i = 1, 2, \dots, N_L\}$ 和 $U_R = \{q_j^R \mid j = 1, 2, \dots, N_R\}$,其中 N_L 和 N_R 分别是点集中特征点的数量。记这些特征点的世界坐标分别是 $q_i^L(x_i^L, y_i^L, 0)$ 和 $q_j^R(x_j^R, y_j^R, 0)$,连接 q_i^L 与 q_j^R 的直线 $l_{i,j}$ 的直线方程为:

$$\begin{cases} (y_i^L - y_j^R)x - (x_i^L - x_j^R)y + y_i^L x_j^R - y_j^R x_i^L = 0 \\ z = 0 \end{cases} \quad (7)$$

计算虚拟极点 $e(x_e, y_e, 0)$ 到直线 $l_{i,j}$ 的距离 $d_{i,j}$,有:

$$d_{i,j} = \frac{|(y_i^L - y_j^R)x_e - (x_i^L - x_j^R)y_e + y_i^L x_j^R - y_j^R x_i^L|}{\sqrt{(y_i^L - y_j^R)^2 + (x_i^L - x_j^R)^2}} \quad (8)$$

设定距离阈值 T_D ,当 $d_{i,j} < T_D$ 时,认为 q_i^L 与 q_j^R 是一个预匹配点对,记为三元组 $\langle q_i^L, q_j^R; d_{i,j} \rangle$,否则认为不满足极线约束,不构成匹配对。记所有满足以上条件的预匹配对的集合为 $U_0 = \{\langle q_k^L, q_k^R; d_k \rangle \mid k = 1, 2, \dots, N_0\}$,为便于表述,这里将预匹配点对对三元组中特征点的下标重新按点对序号 k 进行标记, d_k 表示虚拟极点到直线 $q_k^L q_k^R$ 的距离, N_0 代表预匹配对的数量。

2.3 基于中心点共面约束的匹配对过滤

由于从捆装棒材图像中可能提取出大量的中心点特征,考虑到棒材的密集性,根据 2.2 节预匹配获得的结果可能并不一定是真实的匹配点对,因此需要进一步的过滤。

针对 U_0 中的每个匹配对 $\langle q_k^L, q_k^R; d_k \rangle, k = 1, 2, \dots, N_0$,根据性质 4 计算对应的空间物点的三维坐标 $P_k(x_k, y_k, z_k)$,利用六元组 $\langle q_k^L, q_k^R; d_k, x_k, y_k, z_k \rangle_k$ 表示匹配对。若该匹配对是正确匹配,则 P_k 是对应的真实棒材中心点的估计,否则称 P_k 是一个假点。由于已经假设真实的棒材端面中心点到端平面的距离满足正态分布,可以认为若 P_k 距离端平面 π_E 更近, $\langle q_k^L, q_k^R; d_k, x_k, y_k, z_k \rangle_k$ 为正确匹配的可能性越大,反之可能性越小。如图 3 所示, q_1^L 与 q_1^R, q_1^L 与 q_2^R 是两对预匹配对,在极线约束下无法区分它们的正确性; P_1 和 P_2 分别是与两个匹配对相对应的物点。由于两个匹配对具有公共的左视图特征点 q_1^L ,则 P_1 与 P_2 的深度必然不同,图中显示 P_1 较 P_2 距离 π_E 更近,则两个匹配对相比, q_1^L 与 q_1^R 为正确匹配的可能性更大,即 P_2 为假点的可能性更大。

根据以上思想,首先对端平面进行估计,由于标定时将虚拟像平面所在的标定板摆放到与端平面平行的位置,因

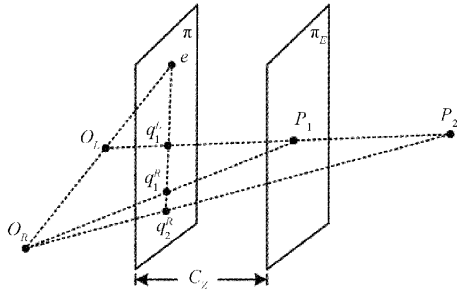


图 3 端平面估计与假点过滤原理示意图

此端平面的方程可表示为：

$$z = C_z \tag{9}$$

其中, C_z 为端平面到虚拟像平面的距离。利用 U_0 中的匹配对计算：

$$\begin{cases} \mu_1 = \frac{1}{N_0} \sum_{k=1}^{N_0} z_k \\ \sigma_1 = \sqrt{\frac{1}{N_0 - 1} \sum_{k=1}^{N_0} (z_k - \mu_1)^2} \end{cases} \tag{10}$$

以 μ_1 作为 C_z 的估计, 根据对应的物点到端平面的距离对 U_0 中的预匹配对进行过滤, 若 $|z_k - \mu_1| > \sigma_1$ 则将其从 U_0 中删除, $k = 1, 2, \dots, N_0$, 记过滤后的匹配对集合为 $U_1 = \{ \langle q_k^L, q_k^R; d_k, x_k, y_k, z_k \rangle \mid k = 1, 2, \dots, N_1 \}$, 这里 N_1 为过滤后匹配对的数量, 为便于表述, 在不影响理解的前提下, 这里仍用 k 标记匹配对的下标。再重新估计端平面的位置, 利用 U_1 中的匹配对计算：

$$\mu_2 = \frac{1}{N_1} \sum_{k=1}^{N_1} z_k \tag{11}$$

以 μ_2 作为 C_z 的新的估计。根据 $|z_k - \mu_2|$ ($k = 1, 2, \dots, N_1$) 的值对 U_1 中的匹配对排序, 选择最小的 M_1 个匹配对作为候选匹配对, 本文选取 $M_1 = 10$ 。在不影响理解的前提下, 记此时的匹配对集合为 $U_2 = \{ \langle q_k^L, q_k^R; d_k, x_k, y_k, z_k \rangle \mid k = 1, 2, \dots, M_1 \}$ 。

再根据 U_2 中匹配对 d_k 的值, 对 U_2 中的候选匹配对进行排序, 选最小的 M_2 个作为推荐的匹配点对, 将对应的物点坐标依次提供给机器人供焊接, 本文中选取 $M_2 = 3$ 。向机器人提供 M_2 个候选焊接点的原因是为机器人提供更多的选择, 标签焊接机器人并不严格要求在某一根特定的棒材上焊接标签, 若由于定位不够精确导致某次焊接操作失败, 可以依据后面推荐点的坐标继续尝试焊接, 由此可提高焊接的成功率。

3 实验研究

3.1 仿真实验

为检验本文提出的棒材端面中心点特征匹配和定位算法的效果, 进行以下仿真实验。本实验中物理坐标和距离均以 mm 为单位, 设左、右相机光心在世界坐标系下坐标

分别为 $(-80, 53, -295)$ 和 $(130, 58, -293)$; 目标棒材的直径为 10 mm, 50 根成捆包装; 随机生成 50 根棒材端面中心点三维坐标作为准确值, 其中 z 坐标服从高斯分布 $N(10, 4)$, 为避免出现极端情况, 去掉了区间 $[4, 16]$ 以外的取值; 假设从虚拟图像中提取的特征在 x 和 y 方向上的误差都服从 $U[-\lambda, \lambda]$ 的均匀分布, 利用本文的特征匹配和过滤方法选择最后的 M_2 个推荐匹配点对, 考察推荐点对的匹配正确性及三维定位的精度。

分别取 λ 为 0.1、0.2、0.4、0.8 四种情况, 实验中设定虚拟图像的尺寸为 1500×1000 像素, 像素距为 0.1 mm, 则每根棒材端面在虚拟图像中的直径约为 96 个像素, 以上四种情况对应的最大特征提取误差分别为 1、2、4、8 像素。针对每个 λ 的取值, 重复以上过程各 1 000 次。图 4(a) 和 (b) 展示了 $\lambda = 0.8$ 时其中一次仿真实验中从左、右两幅虚拟图像中进行特征提取的模拟结果, 图中圆圈表示棒材的端面圆的真实位置, 圆内的点表示模拟的中心点特征提取位置。

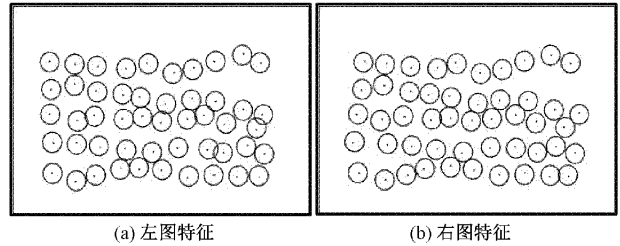


图 4 两幅图像的中心点特征提取结果样例

算法的评价指标选用匹配准确率 A_M 、棒材中心点深度误差 E_z 、棒材中心点端平面误差 E_{xy} , 计算方法为：

$$A_M = \frac{n_M}{M_2} \times 100\% \tag{12}$$

$$E_z = |z_m - z_r| \tag{13}$$

$$E_{xy} = \sqrt{(x_m - x_r)^2 + (y_m - y_r)^2} \tag{14}$$

其中, n_M 表示推荐的匹配对中正确匹配的个数; (x_r, y_r, z_r) 和 (x_m, y_m, z_m) 分别表示真实的端面中心点坐标和计算的结果。

表 1 展示了 λ 取不同值时算法获得的 M_2 个推荐点的正确率和最大误差, 在统计 E_z 最大误差时, 先统计每次仿真推荐的 M_2 个点的 E_z 最小值, 再对 1 000 次仿真的最小值取最大; 在统计 E_{xy} 最大误差时, 先统计每次仿真中 E_z 取最小值时对应的 E_{xy} , 再对 1 000 次仿真的结果取最大。可见所有仿真实验中向机器人推荐的 M_2 个点均来自正确匹配, 说明本文匹配算法是有效的; 深度误差和端平面误差受特征提取精度的影响比较大, 并随着提取误差的增加而增加; 在本实验设置的误差水平下, 端平面中心误差均小于棒材的半径, 能保证机器人将标签焊接到棒材端面范围以内。

本文选取 $M_2 = 3$, 1 000 次仿真实验共产生 3 000 个推荐结果, 图 5 和 6 分别显示了 $\lambda = 0.8$ 时这 3 000 个点

表1 仿真实验结果

λ	A_M 最小/%	E_Z 最大/mm	E_{XY} 最大/mm
0.1	100	0.33	0.21
0.2	100	0.45	0.33
0.4	100	0.83	0.61
0.8	100	1.80	1.14

的 E_Z 和 E_{XY} 的分布情况。可见总体上深度误差超过 2.0 mm 或中心误差超过 1.0 mm 的点占比非常少,因此推荐的 3 个目标点误差均较大的可能性极低。

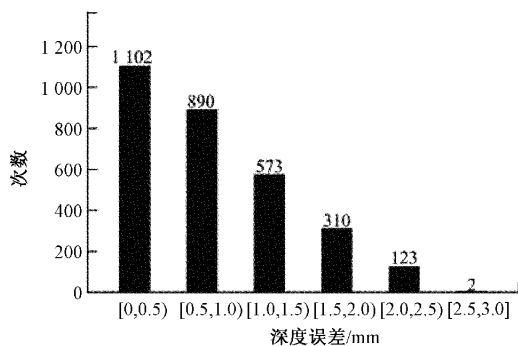


图5 推荐点深度误差分布图

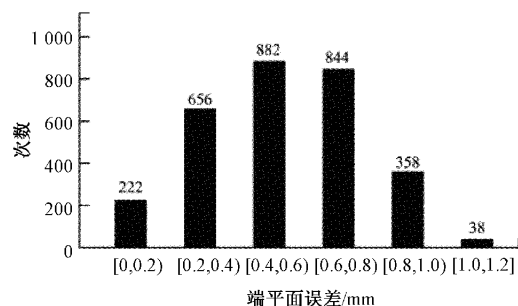
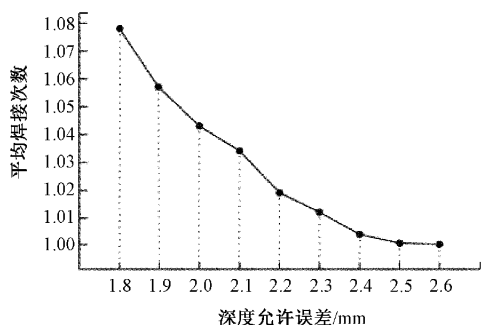


图6 推荐点端面误差分布图

标签焊接机器人一般存在一个深度误差允许值 D_M , 即当 $E_Z < D_M$ 时都能保证焊接成功,图7显示了 $\lambda = 0.8$ 时机器人根据 M_2 个推荐点坐标成功进行焊接时尝试的次数与 D_M 的关系曲线,从图7中可知,多数情况下机器人能1次焊接成功,说明本文算法推荐的棒材中心点可靠性较好。

图7 机器人平均焊接次数与深度允许误差 D_M 的关系曲线

3.2 消融实验

为了验证基于中心点共面约束过滤算法的有效性,对本文算法进行消融实验。仍然采用 3.1 节的仿真方法,将本文完整的算法与仅通过极线约束进行特征点对匹配的方法进行对比,实验中取 $\lambda = 0.8$, 两种方法各进行 1 000 次仿真实验,每次实验中两种方法模拟的特征点位置均相同。由于取 $M_2 = 3$, 算法结果的匹配准确率 A_M 只可能有 0%、33.3%、66.7%、100% 四种取值,表2显示了两种方法 1 000 次实验中与 A_M 值相对应的次数。可见本文提出的完整的算法的结果远优于仅基于极线约束的匹配方法,说明基于中心点共面约束的过滤方法能有效提高匹配的正确率。

表2 消融实验结果

匹配正确率/ %	次数	
	仅极线约束方法	本文完整方法
0	323	0
33.3	455	0
66.7	189	0
100	33	1 000

3.3 真实实验

在实验室搭建双目视觉定位实验环境,实验装置如图8所示。实验中使用 1 捆 29 根圆棒作为被测对象,圆棒端面直径约为 8 mm;将两台相机和直线运动平台固定在 T 形铝框上,实验中保证两台相机的相对位置不变;两台相机型号均为 Basler acA1600-20gm,配 25 mm 定焦镜头,拍摄的图像为灰度图,尺寸为 1 600×1 200 像素;直线运动平台的运动精度为 0.01 mm;棒材放置在直线运动平台上,方向与平台运动方向基本一致。

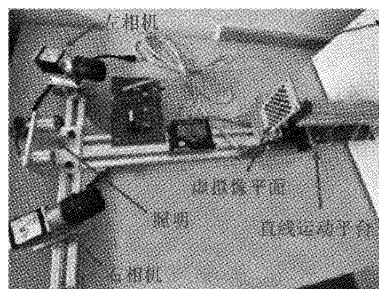


图8 实验装置图

将标定板放置在垂直于 T 形框架平面、与平台运动方向垂直的位置设置虚拟像平面;左、右相机拍摄的标定板图像分别如图 9(a)和(b)所示,实验中利用相机标定参数生成的虚拟图像的尺寸均为 1 000×800 像素,像素距为 0.1 mm。

图 10(a)和(b)分别展示了棒材在初始位置时的双眼的虚拟图像,它们由左、右两相机拍摄的原始图像生成;图 11(a)和(b)分别展示了采用本文的 SVM 方法识别的棒

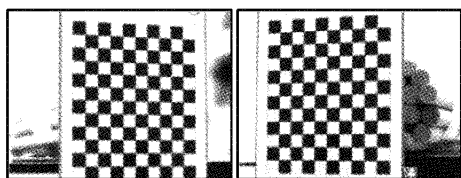


图 9 两相机拍摄的标定板图像

材端面中心点,图中棒材端面内的黑色像素点是被标记出的棒材端面中心;图 12(a)和(b)分别展示了通过连通区方法提取棒材端面中心点的结果,图中棒材端面内的黑色像素点为最终提取出的棒材中心点位置;图 13 展示了最终筛选出的提供给机器人的匹配对。

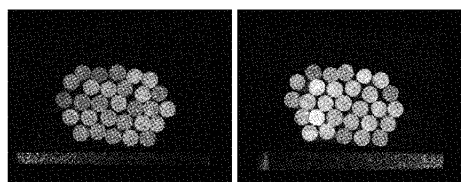


图 10 双眼虚拟图像

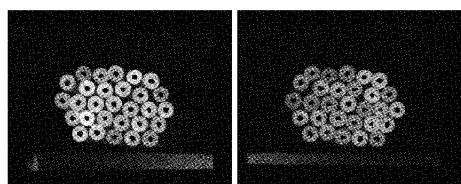


图 11 SVM 方法标记棒材端面中心点

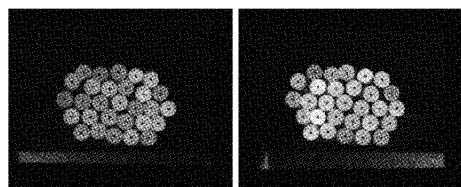


图 12 连通区方法标记的端面中心点

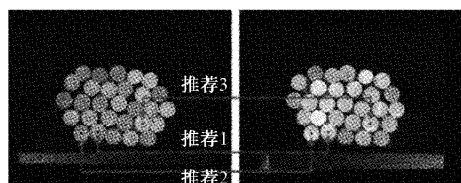


图 13 推荐的棒材端面中心匹配对

由于无法确定棒材中心点坐标的准确值,本文采用直线平台带动棒材移动一段距离,按本文方法计算初始位置推荐的 3 个端面中心点三维坐标后,再计算它们的深度变化用于评价定位的精度。由于直线运动平台的精度较高,认为运动平台的位移值为深度变化的准确值。控制平台向后运动,位移分别为 10、20、30 mm,计算 3 个中心点深度的位移,并与准确值进行对比。考虑到深度误差主要来自特征点提取误差,实验中还使用文献[9]的基于形态学的方法和文献[10]的基于 Hough 变换的方法提取棒材端面中心点坐标,并与本文方法进行对比。表 3 显示了实验的结果。

从实验结果可以看出,本文方法深度位移误差最大为 0.20 mm,平均误差为 0.09 mm,具有较高精度,说明本文提出的棒材端面中心点特征提取方法是有效的;本文方法

表 3 推荐点深度位移计算结果

真实位移/ mm	推荐序号	计算结果/mm			误差/mm		
		本文	文献[9]	文献[10]	本文	文献[9]	文献[10]
10.00	1	10.08	10.07	9.39	0.08	0.07	0.61
10.00	2	10.07	9.61	10.20	0.07	0.39	0.20
10.00	3	10.04	9.95	10.80	0.04	0.05	0.80
20.00	1	20.05	19.99	20.56	0.05	0.01	0.56
20.00	2	19.83	19.88	22.11	0.17	0.12	2.11
20.00	3	20.13	20.10	19.12	0.13	0.10	0.88
30.00	1	30.10	29.89	29.82	0.10	0.11	0.18
30.00	2	30.00	29.82	31.19	0.00	0.18	1.19
30.00	3	30.20	29.89	31.72	0.20	0.11	1.72
	最大误差				0.20	0.39	2.11
	平均误差				0.09	0.13	0.92

在最大误差和平均误差水平上均优于文献[9]和文献[10]的方法,说明本文方法具有一定的优越性。

4 结 论

针对向标签焊接机器人提供捆装棒材端面中心点三

维坐标问题,提出一种基于双目视觉的方法,采用虚拟像平面模型处理双目视觉系统,分析了该模型下双目视觉极线约束的特点,对生成的虚拟图像采用 SVM 和连通区方法提取棒材的中心点特征,通过极线约束和中心点共面约束进行特征点匹配和过滤。虚拟像平面模型的应用,不仅

简化了视觉系统的分析,也使中心点共面约束便于应用。实验结果表明本文方法获得的特征匹配点对可靠性较好,棒材端面中心点三维定位精度较高,对实际应用中设计标签焊接机器人引导系统有指导意义。

参考文献

- [1] 董国梁,张雷,辛山. 基于深度学习的钢筋绑扎机器人目标识别定位[J]. 电子测量技术, 2022, 45(11): 35-44.
- [2] 庄苏峰,屠大维,张旭,等. 水下双目立体视觉对应点匹配与三维重建方法研究[J]. 仪器仪表学报, 2022, 43(5):147-154.
- [3] 张学峰,马行,穆春阳. 基于四叉树的 SIFT 与 K-D 树融合的图像匹配研究[J]. 电子测量技术, 2021, 44(22): 121-127.
- [4] 崔建国,孙长库,李玉鹏,等. 基于 SURF 的快速图像匹配改进算法[J]. 仪器仪表学报, 2022, 43(8): 47-53.
- [5] WANG T, WANG Z, CAO Y, et al. A multi-BRIEF-descriptor stereo matching algorithm for binocular visual sensing of fillet welds with indistinct features[J]. Journal of Manufacturing Processes, 2021, 66(3):636-650.
- [6] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]. Proceedings of IEEE International Conference on Computer Vision, 2011, 2564-2571, DOI: 10.1109/ICCV.2011.6126544.
- [7] 廖国庆,吴文海,曾鑫鹏. YOLOv4 与 ORB 深度融合的绝缘子识别定位研究[J]. 电子测量与仪器学报, 2022, 36(2):131-138.
- [8] YAN X, CHEN X. Research on the counting algorithm of bundled steel bars based on the features matching of connected regions[C]. Proceedings of 3rd IEEE International Conference on Image, Vision and Computing, 2018: 11-15, DOI: 10.1109/ICIVC.2018.8492784.
- [9] 刘晓玉,吴冰,苏志祁,等. 基于数学形态学的棒材自动计数方法[J]. 控制工程, 2013, 20(2): 286-288,294.
- [10] ZHAO J, XIA X, WANG H, et al. Design of real-time steel bars recognition system based on machine vision[C]. Proceedings of 8th International Conference on Intelligent Human-Machine Systems and Cybernetics, 2016: 505-509, DOI: 10.1109/IHMSC.2016.75.
- [11] 侯维岩,张利伟,党麟,等. 一种基于图像处理的棒材计数测量系统的设计与实现[J]. 仪器仪表学报, 2013, 34(5):1100-1106.
- [12] 张一飞,李新福,田学东. 融合边缘特征的 SAD 立体匹配算法[J]. 计算机工程, 2020, 46(4): 236-240,246.
- [13] 吴鹏,徐洪玲,宋文龙. 结合小波金字塔的快速 NCC 图像匹配算法[J]. 哈尔滨工程大学学报, 2017, 38(5):791-796.
- [14] ELASHRY A, SLUIS B, TOTTH C. Improving RANSAC feature matching based on geometric relation [C]. Proceedings of 24th International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Congress, 2021, 43:321-327, DOI: 10.5194/isprs-archives-XLIII-B2-2021-321-2021.
- [15] SHEN L, ZHU J H, FAN C Y, et al. A novel affine covariant feature mismatch removal for feature matching[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-13, DOI: 10.1109/TGRS.2021.3104146.
- [16] 刘昶,张敏,CHEN C. 应用虚拟像平面方法的圆柱表面单目重建[J]. 仪器仪表学报, 2017, 38(10): 2581-2588.
- [17] LLORCA D F, ARROYO R, SOTELO M A. Vehicle logo recognition in traffic images using HOG features and SVM[C]. Proceedings of 16th International IEEE Conference on Intelligent Transportation Systems, 2013: 2229-2234, DOI: 10.1109/ITSC.2013.6728559.

作者简介

刘昶,博士,副教授,主要研究方向为计算机视觉、嵌入式系统。

E-mail:syliuch@126.com

姜楠,硕士研究生,主要研究方向为计算机视觉。

E-mail:1587525777@qq.com