

文章编号: 2095-2163(2023)12-0062-07

中图分类号: TP241

文献标志码: A

# 基于双目视觉引导的纸箱切割系统研究

刘乐<sup>1</sup>, 邹光明<sup>1,2</sup>, 侯宇<sup>1,2</sup>, 刘源洞<sup>1,2</sup>, 皇甫小峰<sup>1</sup>

(1 武汉科技大学 机械传动与制造工程湖北省重点实验室, 武汉 430081;

2 武汉科技大学 冶金装备及其控制教育部重点实验室, 武汉 430081)

**摘要:** 针对传统模式机器人在仓储出库作业时缺乏智能性的问题, 本文研究一种基于双目视觉引导机械臂切割系统, 搭建了双目立体视觉系统并进行双目标定, 通过机械臂手眼标定获得纸箱坐标系, 相当于机械臂基坐标系的坐标变换, 提出一种基于标准尺寸纸箱作为模板的灰度匹配方法, 获得纸箱规格信息进而引导机械臂完成切割。实验结果表明, 3类纸箱定位误差绝对值分别在 1.3 mm、1.5 mm 和 2.2 mm 以内, 满足工业生产要求, 具有较好通用性。

**关键词:** 视觉引导; 手眼标定; 灰度匹配; 切割系统

## Research on carton cutting system based on binocular vision guidance

LIU Le<sup>1</sup>, ZOU Guangming<sup>1,2</sup>, HOU Yu<sup>1,2</sup>, LIU Yuanjiong<sup>1,2</sup>, HUANGFU Xiaofeng<sup>1</sup>

(1 Key Laboratory of Mechanical Transmission and Manufacturing Engineering of Hubei Province, Wuhan University of Science and Technology, Wuhan 430081, China; 2 Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** Aiming at the lack of intelligence of the traditional mode robot in warehouse operation, this paper researches a binocular vision guidance robotic arm cutting system based on binocular vision, builds a binocular stereoscopic vision system, and performs binocular calibration, obtains the carton coordinate system equivalent to the coordinate transformation of the base coordinate system of the robotic arm through the robotic arm hand-eye calibration, and puts forward a grayscale matching method based on the standard-size carton as a template to obtain the carton specification information. And then guide the robotic arm to complete the cutting. The experimental results show that the absolute values of the positioning errors of three types of cartons are within 1.3 mm, 1.5 mm, and 2.2mm respectively. The vision-guided robotic arm-cutting system meets the requirements of industrial production and has good versatility.

**Key words:** vision guidance; hand-eye calibration; gray-scale matching; cutting system

## 0 引言

目前工业生产过程中, 机器人通常是利用示教器编程或者利用离线仿真软件, 控制到达预定的位置<sup>[1]</sup>。但是这种模式在智能化生产过程中存在较大的局限性, 若外界环境或者工件状态发生改变, 机器人无法做出及时判断, 导致任务中断或失败。随着视觉引导技术日益成熟, 将视觉引导技术和机器人技术相关联, 按照视觉所采集的图像信息, 实现目标识别与定位, 能够完成多目标、多类型工件分类作业, 提高生产效率。

在工业机器人中引入视觉引导技术实现目标的

识别、上下料、焊接、装配、切割等<sup>[2]</sup>。I. Giouroudi等<sup>[3]</sup>将机器视觉应用于微装配领域, 通过相机获取光纤图像, 识别定位光纤的位姿, 并引导夹钳完成直径 150  $\mu\text{m}$  光纤的微装配; J. Miura等<sup>[4]</sup>通过视觉检测零件表面的特征信息, 进而求出零件的安装位置和角度; Li. Y. F等<sup>[5]</sup>研制的基于视觉引导的食品封装机器人, 实现对传输带上任意位姿的食品精确拾取和包装。中国对视觉引导技术与机器人的研究相比国外起步晚, 王君臣等<sup>[6]</sup>提出结合机器学习与双目视觉, 实现精确位置测量的星载设备智能装配, 整个过程体现了智能化和精准化; 杨树明等<sup>[7]</sup>提出一种结合精度补偿的机器人手眼关系同步标定方法, 弥补了机器人视觉测量定位的精度不足的缺点。

**基金项目:** 湖北省技术创新专项重大项目(2018AAA027)。

**作者简介:** 刘乐(1997-), 男, 硕士研究生, 主要研究方向: 机器视觉与结构设计。

**通讯作者:** 邹光明(1970-), 男, 博士, 教授, 硕士生导师, 主要研究方向: 机器视觉、机器人。Email: zoug Guangming@wust.edu.cn

收稿日期: 2022-12-02

视觉引导的纸箱切割系统, 实现以智能化仓储出库瓦楞包装纸箱为目标物体的机械臂智能切割。利用双目相机采集纸箱图像, 上位机将处理的纸箱图像特征转化为目标定位信息, 引导机械臂完成智能切割实验。

## 1 切割系统平台构成

本文以协作机械臂为基础, 搭建了如图 1 所示的视觉引导切割系统。视觉引导切割系统主要包括: 双目视觉系统、纸箱切割夹具、切割刀具、协作机械臂和上位机。

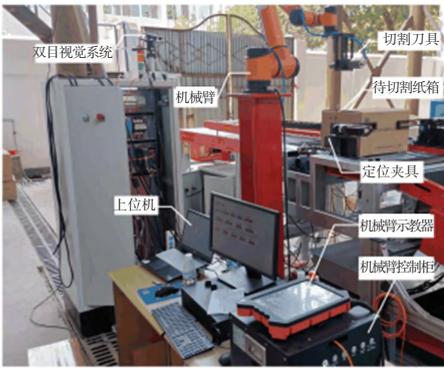


图 1 视觉引导切割系统

Fig. 1 Vision-guided cutting systems

首先, 双目相机对工位中的纸箱进行图像采集, 并采用双边图像滤波及边缘特征提取方法对图像进行处理; 其次, 通过模板匹配识别出当前要切割的纸箱并计算纸箱 4 个顶点在图像中的位姿; 最后, 根据构建的切割系统参数化模型将图像中位姿转化为相对于机械臂基坐标的位姿信息。引导机械臂依次经过 A、B、C、D 点后, 再回到 A 点, 完成纸箱上盖切割, 其切割方向示意图如图 2 所示。

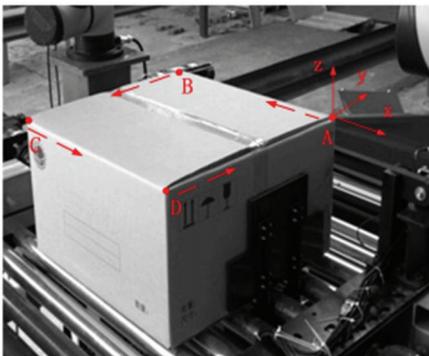


图 2 视觉引导切割方向示意图

Fig. 2 Schematic diagram of visually guided cutting direction

## 2 双目标定与手眼标定

视觉引导系统借助双目相机采集待切割纸箱的图像实施, 通过特征处理得到所需的特征信息, 最后将纸箱特征信息转化为在机械臂的基坐标系下的位姿信息。双目标定通过求得相机内外参数进而获得物体世界坐标系与像素坐标系之间的关系; 手眼标定求得像素坐标系与机械臂基坐标的联系, 进而求得目标纸箱特征点在机械臂的基坐标下的位姿信息。

### 2.1 双目标定

双目立体成像的过程实质上是四大坐标系之间的坐标转换, 其相互位置关系如图 3 所示。这 4 个坐标系依次是世界坐标系 ( $O_w - X_w Y_w Z_w$ )、相机坐标系 ( $O_c - X_c Y_c Z_c$ )、图像坐标系 ( $O_1 - xy$ )、像素坐标系 ( $O_0 - uv$ ), 其中  $O_0$  与  $O_1$  之间距离为相机焦距  $f$ , 点  $P(X_w, Y_w, Z_w)$  是世界坐标中真实存在的一点, 其在图像中的成像点为点  $p(x, y)$ 。

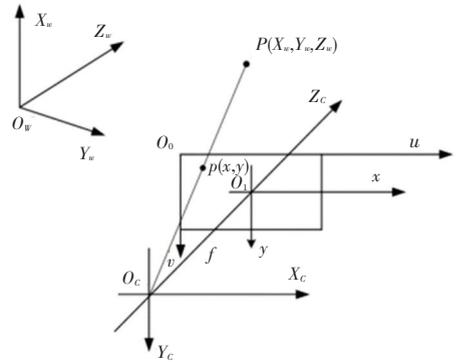


图 3 4 种坐标系的位置关系

Fig. 3 Position relation of the four coordinate systems

像素坐标系与世界坐标系之间的转化关系式 (1):

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

其中,  $\mathbf{R}$  为旋转矩阵;  $\mathbf{T}$  为平移矩阵; 图像中心点坐标为  $(u_0, v_0)$ ;  $f_x, f_y$  分别为  $X$  和  $Y$  方向广义焦距。

根据式 (1) 可以采集不同位置的标定板图像, 求解双目相机的内参和外参。本文基于张正友标定法对搭建的视觉系统进行双目标定, 双目视觉系统采集 16 组不同姿态标定板图像, 如图 4 所示。

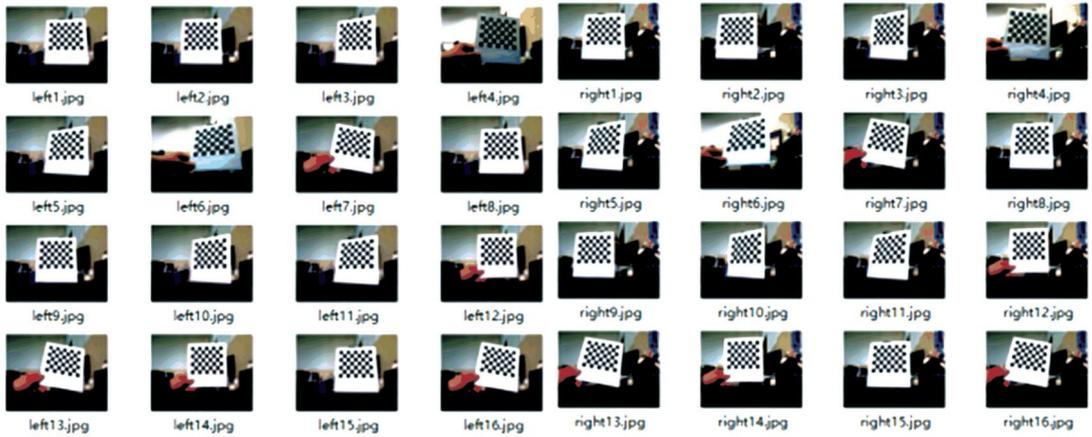


图4 采集到的16组双目相机标定板图像

Fig. 4 16 groups of calibration plate images of binocular camera collected

标定的总体平均误差结果符合要求,左右相机的内部参数标定结果见表1。

表1 左右相机内部参数

Table 1 Left and right camera internal parameters

内部参数	左相机	右相机
焦距 $f_x$	541.064 000 0	540.650 000
焦距 $f_y$	608.510 000 0	607.630 000
中心点坐标 $u_0$	316.952 000 0	321.799 000
中心点坐标 $v_0$	245.180 000 0	238.361 000
径向畸变 $k_1$	-0.060 700 0	-0.084 200
径向畸变 $k_2$	-0.088 700 0	0.261 200
切向畸变 $p_1$	0.000 074 5	-0.000 715
切向畸变 $p_2$	0.002 088 9	-0.001 181

标定可得到左右相机外部参数的旋转矩阵和平移矩阵。

## 2.2 手眼标定

手眼标定是为了创建像素坐标系与机械臂基坐标系转换关系  ${}^{base}H_{cam}$ 。眼在手外的手眼模式具有大视野、响应速度快及系统更加简洁等特点,适用于仓储纸箱出库的大视野定位场合,故选用眼在手外的手眼模式进行手眼标定,眼在手外手眼标定模型如图5所示。

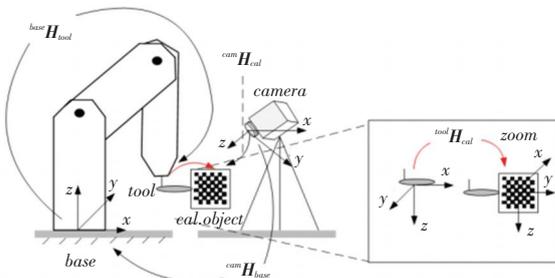


图5 眼在手外手眼标定模型

Fig. 5 Eye to hand calibration model

根据机器人运动学关系,可通过机械臂示教器得到机械臂末端工具坐标系到机械臂基坐标系的变换矩阵  ${}^{base}H_{tool}$ ,通过双目标定也可以得到标定板到相机的变换矩阵  ${}^{cam}H_{cal}$ ,联立可得转换关系式(2):

$${}^{base}H_{cam} * {}^{cam}H_{cal} = {}^{base}H_{tool} * {}^{tool}H_{cam} \quad (2)$$

其中,  ${}^{base}H_{cam}$  为相机到基坐标系变换矩阵,  ${}^{tool}H_{cam}$  为相机到末端工具坐标系变换矩阵。

标定板固定在机械臂末端,移动机械臂运动到不同位置,采集对应时刻的标定板图像,并记录示教器上机械臂末端位姿数据,代入式(2)整理可得关系式(3):

$${}^{tool}H_{base}^{-1} * {}^{tool}H_{base}^0 * {}^{base}H_{cam} = {}^{base}H_{cam} * {}^{cam}H_{cal}^{-1} * {}^{cam}H_{cal}^0 \quad (3)$$

设  $A = {}^{tool}H_{base}^{-1} * {}^{tool}H_{base}^0$ ,  $B = {}^{cam}H_{cal}^{-1} * {}^{cam}H_{cal}^0$ ,  $X = {}^{base}H_{cam}$ , 则式等价于求解  $AX = XB$  方程。

本文在手眼标定时,共采集13组机械臂末端标定板图像并记录示教器上机械臂末端位姿信息,将这些数据写入OPENCV手眼标定程序中进行求解,最终得到手眼标定转换矩阵。

## 3 目标纸箱识别与匹配

为了得到目标纸箱位置信息,本文在切割过程中搭建视觉引导系统,仓储出库的纸箱规格尺寸分别为  $500 \times 290 \times 300$ 、 $500 \times 495 \times 300$  和  $600 \times 390 \times 400$  (单位 mm)3类,类型依次记为A、B、C类型。为了完成纸箱定位,要从目标纸箱图像中提取特征点,系统以纸箱图像作为输入,基于像素处理,经过图像预处理、纸箱边缘轮廓提取以及模板匹配,得到更易提取特征的图像。

### 3.1 图像预处理

图像在经过采集到传输的过程中会受到噪声干

扰,图像中一些重要细节会丢失,会对后续图像的分析与判断产生影响,本文采用双边滤波对纸箱图像去噪操作。双边滤波在考虑像素间空间距离和灰度相似性的同时,又对图像平滑操作过程进行了边缘信息的保留操作,是非线性滤波的一种形式<sup>[8]</sup>。双边滤波优化了高斯滤波器只顾及像素间欧式距离,使用的模板系数大小与窗口中心距离成反比。双边滤波基于一维高斯函数生成值域模板,基于二维高斯函数生成距离模板。

生成距离模板系数的关系式(4):

$$d(i, j, k, l) = \exp\left(\frac{-(i-k)^2 - (j-l)^2}{2\sigma_d^2}\right) \quad (4)$$

其中,  $(i, j)$ 、 $(k, l)$  分别是模板窗口的其他系数坐标点和中心坐标点,  $\sigma_d$  是高斯函数的标准差。

生成值域模板系数的关系式(5):

$$r(i, j, k, l) = \exp\left(-\frac{\|f(i, j) - f(k, l)\|^2}{2\sigma_r^2}\right) \quad (5)$$

其中,函数  $f(i, j)$  表示在点  $(i, j)$  对应的像素值;  $f(k, l)$  表示图像在  $(k, l)$  的像素值;  $\sigma_r$  是高斯函数的标准差。

以上模板相乘得到双边滤波模板函数,式(6):

$$\omega(i, j, k, l) = \exp\left(\frac{-(i-k)^2 - (j-l)^2}{2\sigma_d^2} - \frac{\|f(i, j) - f(k, l)\|^2}{2\sigma_r^2}\right) \quad (6)$$

用这个模板对图像进行卷积,就得到双边滤波之后的图像。对双目相机获得的 B 类规格纸箱(尺寸  $500 \times 495 \times 300$ ) 图像进行预处理前后效果如图 6 所示。



(a) 灰度原图 (b) 双边滤波图

图 6 纸箱预处理前后效果图

Fig. 6 Effect picture of carton before and after pretreatment

### 3.2 边缘检测

对纸箱图像预处理操作后,获得了清晰的目标纸箱图像。提取图像的边缘信息能够划分出感兴趣的目标区域,同时能减少计算数据量,最大程度保留图像的重要信息是实现图像识别定位的关键操作<sup>[9]</sup>。

本文采用改进的 Canny 算子边缘检测算法,该

算法在平滑处理时通过双边滤波来取代高斯滤波,图像的边缘检测采用 Otsu 算法进行自适应的阈值设定。

改进的 Canny 算法检测步骤:

(1) 双边滤波,采用双边滤波器后在滤除噪声的同时极大地保存了图像边缘信息,得到更好的边缘检测效果。

(2) 计算梯度幅值和方向。传统 Canny 算法梯度计算时只考虑了像素在  $x$  和  $y$  方向梯度值,得到部分像素点丢失的图像;改进的 Canny 算法在 Sobel 算子的基础上,添加  $45^\circ$ 、 $135^\circ$  两个方向梯度模板,如图 7 所示。

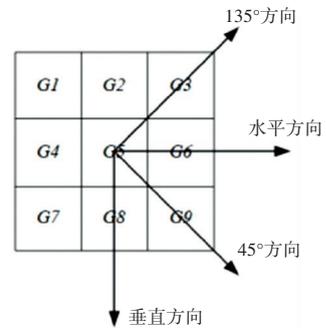


图 7 4 个方向梯度模板

Fig. 7 4-direction gradient template

图 7 中,4 个方向的梯度模板分别为:

$$\left\{ H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, H_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \right\} \quad (7)$$

$$\left\{ H_{xy} = \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}, H_{yx} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} \right\} \quad (8)$$

其中,  $H_x$  为水平方向梯度模板;  $H_y$  为垂直方向梯度模板;  $H_{xy}$  为  $45^\circ$  方向梯度模板;  $H_{yx}$  为  $135^\circ$  方向梯度模板。

4 个梯度模板卷积得到 4 个梯度分量  $Q_x$ 、 $Q_y$ 、 $Q_{xy}$ 、 $Q_{yx}$ , 对应该像素点的梯度幅值为

$$E = \sqrt{Q_x^2 + Q_y^2 + Q_{xy}^2 + Q_{yx}^2} \quad (9)$$

(3) Otsu 算法确定高低阈值,是一种对非极大值抑制后的梯度图像进行二值化分割阈值的算法<sup>[10]</sup>。Otsu 算法确定自适应的阈值,进而避免了认为确定高低阈值存在的局限性<sup>[11]</sup>。

图像总平均梯度表示为式(10):

$$\mu = \omega_0 \times \mu_0 + \omega_1 \times \mu_1 \quad (10)$$

类间方差表示为式(11):

$$L = \omega_0 (\mu_0 - \mu)^2 + \omega_1 (\mu_1 - \mu)^2 \quad (11)$$

其中,  $\omega_0$  为目标像素点占整个图像的比例;  $\mu_0$  为目标像素平均梯度;  $\omega_1$  为背景像素点占整个图像的比例;  $\mu_1$  是背景像素平均梯度。

本文对 B 类规格纸箱 (500×495×300) 灰度原图提取感兴趣区域后的图像,用改进后 Canny 算法进行边缘检测,其检测效果如图 8 所示。

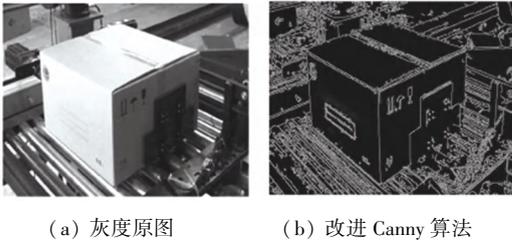


图 8 改进的 Canny 算法检测效果

Fig. 8 Improved Canny algorithm to detect the effect

由图 8 可知,改进的 Canny 算法在噪声的去除和细节边缘处理方面更加出色,纸箱上盖折痕部位更加明显,箱子的轮廓更加完整。同时改进的 Canny 算法采用 Otsu 算法进行了高低阈值自适应确定,具有更好的自适应性。

### 3.3 基于自适应尺寸的模板匹配

针对环境中特定物体的识别与匹配,模板匹配是一种快捷、高效的方法。纸箱图像的边缘特征提取和模板匹配是机械臂实现对纸箱定位的重要步骤,常用的模板匹配方法主要有基于灰度特征和基于形状特征两种。本文的目标工件是由专用夹具夹紧定位,工作台传输辊道和专用夹具在几何形状特征匹配下产生干扰较多,所以采用基于灰度的模板匹配算法。

基于灰度的模板匹配是对同一位置的灰度值进行对比,差值最小或者达到最小阈值时表明此次匹配成功。为了防止光线的线性变化对灰度值的匹配产生干扰,本文采用归一化互相关匹配作为相似性度量准则,式(12):

$$NCC(r, c) = \frac{1}{n} \sum_{(u, v) \in T} \frac{(t(r, c) - m_t) \times (f(r + u, c + v) - m_f(r, c))}{\sqrt{s_t^2} \times \sqrt{s_f^2(r, c)}} \quad (12)$$

其中,  $m_t, s_t^2$  分别是模板图像所有灰度值的平均值和方差,  $m_f(r, c)$  和  $s_f^2(r, c)$  分别是匹配的当前位置图像所有灰度值的平均值和方差。

基于灰度的模板匹配算法相对成熟,但是匹配时间较长。为了提高箱装物资自动化出库过程视觉

处理实时性,优化视觉引导切割系统响应速度,为此采用图像金字塔分层搜索策略来减少匹配过程的运算量。

图像金字塔分层搜索策略首先从最顶层低分辨率的图片中进行初步搜索匹配,此时是从所有位置上进行全搜索。因为分辨率最低,搜索之后很快就可以发现多个粗匹配位置;将获得的粗匹配位置结果回溯到分辨率较高的下一层图像中,下一层中只从上一个产生粗匹配区域附近执行相关搜索;以此一层层往下回溯,在高分辨率待匹配图片中能够发现一些小范围领域,再通过相似性度量函数实现最后精准匹配,最终模板图像在底层待匹配图像中找到对应匹配位置。金字塔分层搜索策略优化了搜索次数,极大减少了运算量。

智能化仓储中标准化包装纸箱有多种尺寸规格,为了确定每次出库纸箱规格类型进而完成特定纸箱的匹配与定位,本文采用标准化包装纸箱作为模板的灰度模板匹配方法。首先,将 3 类已知规格纸箱提取感兴趣区域制作成模板图像;其次,对出库的 3 类未知纸箱及 3 类模板图像分别经过 3 次高斯平滑处理和亚采样处理,得到 6 组不同分辨率和尺寸的图像;按照分辨率不同摆放成金字塔结构,分辨率最高的原始图像在金字塔的最底部,往上的图像分辨率逐渐变低。以 B 类纸箱为例采样后的金字塔待匹配图像如图 9 所示,其中图 (a) 是原始图像,每张图像对应着采样后分辨率数值。

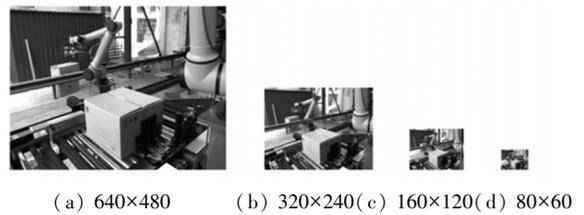


图 9 待匹配图像金字塔示意图

Fig. 9 Pyramid diagram of images to be matched

本文结合金字塔分层搜索策略,采用归一化互相关匹配依次用 3 类模板图像对出库过程中尺寸未知的纸箱进行匹配计算。若匹配成功待匹配纸箱则属于对应模板图像的尺寸规格,完成了对目标纸箱的切割点进行定位。由于纸箱模板与待匹配图像的局部特征相似,模板匹配过程容易出现误匹配情况,本文采用自适应尺寸的模板匹配来剔除误匹配情况。令载入的图像每次循环缩放为一定的比例,再进行匹配,将得到的感兴趣区域与原始模板图片进行相似度对比,选取最高相似度区域作为最佳的模板匹配区域。

最终得到 3 类出库纸箱模板匹配结果如图 10 所示,绿色框表示待匹配图片与模板重合的区域。传统灰度匹配由于图像分辨率较大,匹配耗时较长,金字塔分层搜索,匹配耗时大幅降低。

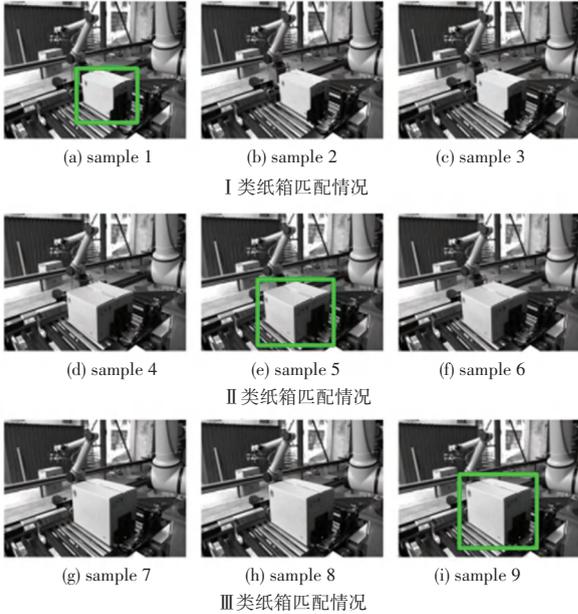


图 10 模板匹配结果

Fig. 10 Template matching results

### 4 实验

搭建基于协作机械臂的双目视觉引导切割系统来进行切割实验。引导机械臂依次对 3 类目标纸箱经过 A、B、C、D 点后,再回到 A 点,实现视觉引导切割。以 B 类规格的纸箱 (尺寸 500×495×300) 为例,切割实验过程如图 11 所示。

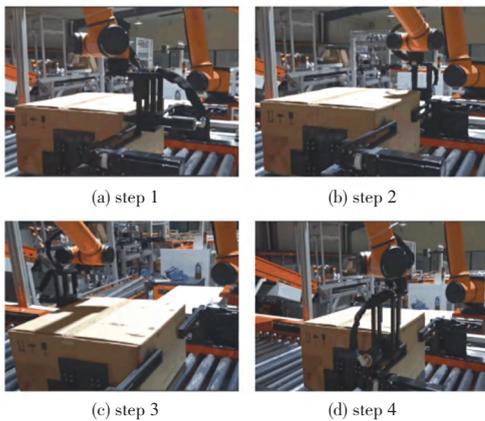


图 11 机械臂切割过程

Fig. 11 Cutting process of mechanical arm

利用构建的机械臂手眼模型,对于 B 类纸箱换算求得其在机械臂基坐标系下的位姿,将此与机械臂单独走到纸箱 4 个特征点时读取的实际位姿相比较,其对比结果见表 2。

表 2 视觉定位结果与实际位置对比

Table 2 Comparison between visual positioning results and actual position

点位	位置	视觉定位结果	实际坐标	误差
A	X/mm	-516.677	-518.018	-1.341
	Y/mm	230.066	229.322	-0.744
	Z/mm	21.961	20.693	-1.268
B	X/mm	-1 113.208	-1 111.590	1.618
	Y/mm	227.744	229.322	1.578
	Z/mm	22.628	20.693	-1.935
C	X/mm	-1 112.534	-1 111.590	0.944
	Y/mm	-235.739	-233.596	2.143
	Z/mm	19.470	20.693	1.223
D	X/mm	-645.625	-647.926	-2.301
	Y/mm	-234.901	-233.596	1.305
	Z/mm	22.515	20.693	-1.822

为了更直观反应 3 类纸箱切割之后误差情况,对 A、C 两类纸箱进行引导切割,接着将 3 类纸箱切割之后视觉定位结果和实际坐标之间误差转化为图表进行对比,3 种纸箱切割误差情况如图 12 所示。

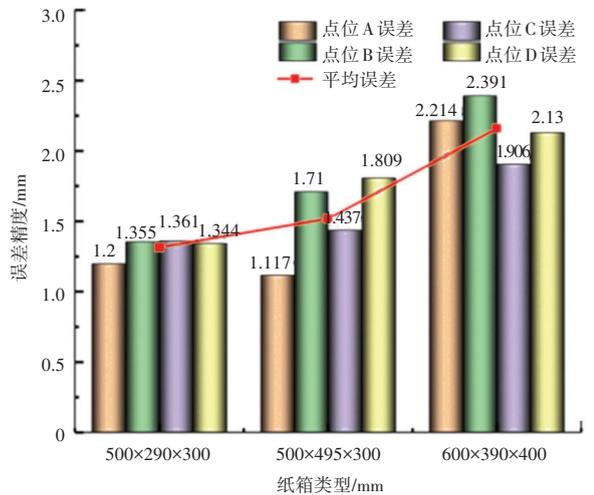


图 12 3 种纸箱切割误差情况

Fig. 12 Three kinds of carton cutting error

对表 2 和图 11 分析可知,纸箱经过视觉定位结果和实际末端位姿结果之间误差主要来源于双目相机标定、手眼模型建立以及立体匹配的精度,最终视觉定位结果。切割 A 类纸箱其 4 个点位置平均精度误差绝对值小于 1.3 mm, B 类纸箱平均误差小于 1.5 mm, C 类纸箱平均误差小于 2.2 mm,该模板匹配识别精度满足机械臂切割纸箱任务时对位置精度的定位要求。

### 5 结束语

本文搭建了基于双目视觉的机械臂切割平台, (下转第 74 页)