

# 基于 Kinect 的机器人织物抓取系统<sup>①</sup>

王晓华, 戴庭宣, 李 珣, 洪 良, 张 蕾

(西安工程大学 电子信息学院 陕西省纺织印染自动化工程技术研究中心, 西安 710048)

**摘 要:** 现有纺织行业自动化水平较低、劳动力成本高, 机器人代替人工必将成为必然. 本文研制了基于 Kinect 的机器人织物抓取系统. 利用 Kinect 传感器获得织物图像, 应用 CamShift 算法找出目标织物, 获得该目标织物的质心, 应用机器人运动学将此质心坐标转换为机器人运动参数, 实现机器人对织物的抓取. 实验结果表明, 该方法能够快速准确获得目标织物质心的三维坐标, 较双目视觉机器人系统具有优势.

**关键词:** Kinect; 织物; CamShift; 机器人

## Robot Fabric Scraping System Based on the Kinect

WANG Xiao-Hua, DAI Ting-Xuan, LI Xun, HONG Liang, ZHANG Lei

(Shaanxi Research Center of Automation of Textile Printing & Dyeing, College of Electronic Information, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** With the lower level of automation and high labor costs in the present fabric industry, it is inevitable that the robot will replace human to work. The article presents the robot fabric scraping system based on the Kinect. First of all, target fabric image is obtained by making use of Kinect. Then, target fabric can be found out by CamShift algorithm and the center of mass is gotten. Finally, the coordinates can be transformed into motion coefficient with robot kinematics to make robot scraping fabric come true. The experiment results imply the method which can get three-dimensional coordinates of target fabric quickly has more advantages than visual robot.

**Key words:** Kinect; fabric; CamShift; robot

随着机器人技术的迅猛发展, 各种各样的机器人已经非常广泛的应用于工业、医疗、农业、教育以及人们日常生活等众多领域. 目前在纺织行业中, 传统织物生产后的整理工作仍然是依靠人工进行, 一方面跟不上纺织业智能化发展的要求, 另一方面其生产效率也不高, 同时, 生产车间的高温、湿气等因素也容易对人体健康产生不良影响. 将机器人应用到织物整理中来将成为今后的发展趋势. 而且目前的机器视觉理论<sup>[1,2]</sup>技术发展已经很成熟, 将机器人和机器视觉结合起来应用于纺织业不仅能够提高纺织业生产效率, 扩大利润, 也能够减少环境对工人产生的不良影响. 对于目标物体的三维信息获取的方法主要有双目视觉、单目视觉等方法, 然而这些方法都有各自的缺陷, 比如单目视觉的测量速度比较慢, 而双目视觉计算和匹

配过程又很复杂, 微软公司研制的 Kinect 作为本机器人系统的视觉传感器, 它能够同时提供二维图像及深度信息, 并且价格低廉, 被国内外研究者应用到机器人的导航与定位<sup>[3]</sup>、目标的跟踪<sup>[4]</sup>、障碍物的检测<sup>[5]</sup>等方面, 本文应用 Kinect 作为视觉传感器, 将其安装于四自由度机器人上, 结合机器视觉技术、Camshift 跟踪算法以及机器人运动学方法, 实现机器人对不规则形状织物的自主抓取.

## 1 机器人织物抓取系统组成

实验室内视觉机器人织物抓取系统如图 1 所示, 主要包括下面几个部分: Kinect 传感器, GBR-400 四自由度工业机器人, 计算机、GBR-400-SR 运动控制卡、松下 MinasA4 AC 伺服驱动器和电机、织物传送带. 计

<sup>①</sup> 基金项目:国家自然科学基金(61301276);陕西省自然科学基金基础研究计划(2016JQ5106,2015JQ6258)

收稿时间:2016-07-26;收到修改稿时间:2016-09-23 [doi:10.15888/j.cnki.csa.005739]

算是系统的核心,承担着图像处理、操作界面的显示和控制以及系统各部分的通讯等重要任务;织物放置在传送带上,传送带织物传送一定位置并停止,因为织物在传送带上放置位置的随意性以及形状的不规则性,将 Kinect 摄像头安装在传送带正上方以获得织物图像信息。

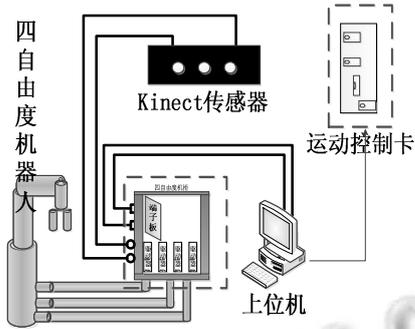


图1 视觉机器人系统框图

首先对 Kinect 进行标定以及将深度图像与彩色图像进行融合配准,将采集到的图像传至计算机,然后对采集的图像进行预处理,接下来就需要进行图像分割以及对织物的定位,最后根据图像坐标系与机器人坐标系的转换关系通过机器人轨迹规划来实现对目标织物的抓取,具体系统工作流程如图 2 所示。

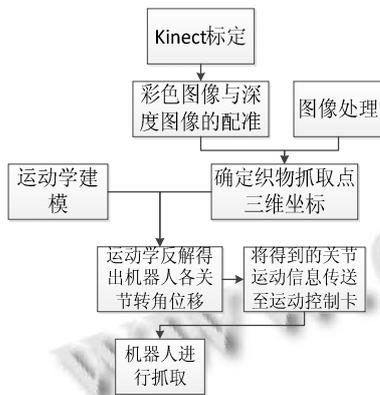


图2 系统工作流程

## 2 基于Kinect的织物识别

### 2.1 Kinect 标定

Kinect 是由一个 RGB 摄像机、一个红外发射器以及与之配合的 CMOS 红外传感器和一组麦克风阵列所组成,其中 RGB 摄像头来产生彩色图像,而红外 CMOS 和红外发射器则是通过 LightCoding 光编码技

术来产生深度图像.由于 Kinect 生成的彩色图像和深度图像并不是完全重合的,所以要将 Kinect 获得的 RGB 图像和深度图像进行匹配,使获得的质心位置在 RGB 图像中与在深度图像中的像素位置对应起来,从而得到该质心在深度图像中的索引位置,获得该点在深度图像中的像素值,也就是该点的深度值。



图3 RGB 图像与深度图像以及融合后的图像

本系统的抓取对象为具有一定形状的织物,抓取范围较大,因而对标定精度的要求不是非常严格,对于标定方法的选择上以简单快速为主,采用文献[6]的标定方法,摄像机模型应用针孔模型,其定义如下<sup>[7,8]</sup>:

$$sm = A[R | t]M$$

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

式中的(u,v)为点投影在图像平面坐标系下的坐标,A则是摄像机的内参数矩阵,[R|t]是摄像机的外参矩阵,(X<sub>w</sub>,Y<sub>w</sub>,Z<sub>w</sub>)是点在世界坐标系下的坐标.式中 X<sub>w</sub>、Y<sub>w</sub>是彩色图像中的对应的是世界坐标,Z<sub>w</sub>为深度图像中相同像素点的深度值。

应用此模型与 OpenNI 自带函数可将图像坐标系的点转换到真实世界坐标系中.利用 ConvertProjectToRealWorld 函数将图像三维坐标点转换到真实世界坐标系中<sup>[9]</sup>的具体实现方法如下<sup>[10]</sup>:

1) 获得 Kinect 的视角值,也即 Kinect 在垂直方向和水平方向的视角。

2) 计算深度映像图的宽度值 W、高度值 H 与距离 d 的比值,如图 4 所示为真实世界坐标系,该坐标系表示了现实世界中 Kinect 的视场范围和深度值获取的原理,坐标原点为红外 CMOS 的中心,X-Y 平面即 Kinect 所在平面,Z 轴方向代表了 Kinect 镜头所对方向,α 表示水平视角角度值 β 表示垂直视角角度值,由此得到图像中的一点(x,y)与距离 d 的单位换算比例为:

$$xTz = 2 \tan(\frac{\alpha}{2}) \quad yTz = 2 \tan(\frac{\beta}{2}) \quad (1)$$

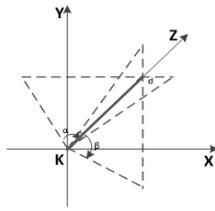


图 4 以 Kinect 位置为原点的真实世界坐标系

3) 从图像坐标系转换到摄像头坐标系, 实际上只是对(x,y)的值进行了转换, z 的值并不会改变, 则真实世界坐标(X<sub>w</sub>, Y<sub>w</sub>, Z<sub>w</sub>)的计算公式<sup>[10]</sup>如下:

$$\begin{aligned}
 X_w &= \left( \frac{x}{\text{outputMode.n}X_c \text{ Res}} - 0.5 \right) * z * xTz \\
 Y_w &= \left( 0.5 - \frac{y}{\text{outputMode.n}Y_c \text{ Res}} \right) * z * yTz \\
 Z_w &= z
 \end{aligned} \tag{2}$$

应用上述方法求得图像点(x,y)对应的真实世界中的点(X<sub>w</sub>, Y<sub>w</sub>, Z<sub>w</sub>)。本文标定的部分图片如图 5 所示。



图 5 部分标定模板图片

彩色相机的参数标定结果:

$$A_1 = \begin{bmatrix} 622.267 & 0 & 379.133 \\ 0 & 616.283 & 273.816 \\ 0 & 0 & 1 \end{bmatrix}$$

深度相机的参数标定结果:

$$A_2 = \begin{bmatrix} 620.891 & 0 & 380.066 \\ 0 & 614.528 & 279.686 \\ 0 & 0 & 1 \end{bmatrix}$$

相较于传统的双目视觉获取目标点深度信息的情况, Kinect 传感器具有天然优势。

## 2.2 目标识别

本文采用 Camshift 方法对柔性织物进行识别和定位。Camshift 提取目标的颜色特征而不受目标形状约束, 被广泛用于目标跟踪<sup>[11]</sup>, 分为背投(Back Projection)计算、均值漂移(Mean Shift)<sup>[12]</sup>和 camshift 算法三步骤。

Back Projection 利用模板的颜色空间分布, 改变源图像的像素值, 进而突出源图像中与模板类似的图像特征, 利用 backproject 的结果可用于对目标织物的

分割检测。由于 RGB 颜色空间对光照的亮度变化比较敏感, 需要将 RGB 图像转换到 HSV 色彩空间中, 然后对其中的 H 分量作直方图, 在直方图中代表了不同 H 分量值的概率或者像素个数, 即得到了颜色概率查找表。将织物图像中每个像素的值用其颜色出现的概率替换, 就得到了颜色概率分布图。

MeanShift 通过迭代寻优找到概率分布的极值来定位目标。从每个像素开始, 首先估计有相似颜色的邻近像素点的密度(局部密度)的梯度, 而后利用迭代算法求出局部密度的峰值(即重心点), 把能够聚类到同一个峰值点的所有像素点划分成一个区域。该算法最终得到的通用符号表示的核密度估计式<sup>[13]</sup>:

$$\hat{f}_{h,k}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\| \frac{x-x_i}{h} \right\|^2\right)$$

将 meanshift 算法扩展到连续图像, 就形成了 Camshift 算法。将机器人视场内获取的视频的所有帧做 meanshift 运算, 并将上一帧的结果(即搜索窗口的大小和中心), 作为下一帧 meanshift 算法的搜索窗的初始值, 就可以对织物进行实时的定位。

Camshift 算法的具体计算步骤如下<sup>[14]</sup>:

1) 读入图像, 将 RGB 色彩空间的值转化到 HSV 空间, 计算出 H 分量的特征直方图。

2) 根据获得的特征直方图将最初获得的 RGB 图像转化成色彩概率分布图像。

3) 初始化搜索窗口的大小和位置(x<sub>c</sub>, y<sub>c</sub>)。计算此时的搜索窗口内质心位置。假设(x<sub>c</sub>, y<sub>c</sub>)为搜索窗口中像素的位置, I(x<sub>c</sub>, y<sub>c</sub>)为投影图中(x, y)处的像素值, 定义搜索窗口的零阶矩 M<sub>00</sub> 和一阶矩 M<sub>01</sub> 和 M<sub>10</sub> 分别如下:

$$M_{00} = \sum_x \sum_y I(x, y)$$

$$M_{10} = \sum_x \sum_y x I(x, y)$$

$$M_{01} = \sum_x \sum_y y I(x, y)$$

由此可以计算出搜索窗口的质心:

$$(x_c, y_c) = \left( \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right)$$

4) 移动搜索窗的中心到质心位置, 判断是否收敛(窗口中心移动到质心的距离小于预设的固定阈值, 或者达到设定的迭代次数)。如果不满足收敛条件, 则重复 3)直至收敛为止; 若满足收敛条件, 则返回搜索结果。

当目标织物进入视场后开始跟踪,直到静止,质心坐标不再发生变化,质心位置即为最终的目标位置.



图6 通过 Camshift 算法跟踪毛巾到最终位置

### 3 抓取实验

织物质心位置就是机器人末端执行器要到达的位置.根据机器人运动学逆解分析,获得机器人四个轴运动控制指令,驱动机器人末端执行器工作,具体实验系统如图7所示.



图7 实验系统

#### 3.1 实验结果

本文对各种颜色的纯色毛巾进行了实验,机器人能够识别绿色、红色、紫色、黄色等等颜色的织物.图8所示为紫色织物处理结果,检测抓取点坐标为(-0.587, -0.260, 0.642),经测量目标抓取点的实际坐标为(-0.592, -0.254, 0.649),通过 Kinect 进行目标检测的误差在 1cm 内.



(a)RGB 图像 (b)深度图像

图8 实验中的采集到的图

图像采集处理后机械臂实施抓取的过程如图9所示.实验中用多种不同颜色毛巾进行识别抓取,结果均能快速成功识别抓取,说明本文算法快速稳定对于不同颜色的毛巾也能够快速的识别.

实验结果如表1所示,在三个坐标方向的最大误差在 1cm 内.尤其是 Z 方向上的误差均在 0.2cm 左右.



图9 机械臂实施抓取过程

表1 实验结果

	目标坐标(cm)	偏差(cm)
1	(-0.384, -0.370, 0.645)	(-0.3, -0.3, 0.1)
2	(-0.432, -0.253, 0.642)	(-0.4, -0.2, 0)
3	(-0.528, -0.276, 0.647)	(0.3, 0.2, 0.2)
4	(0.171, -0.364, 0.643)	(-0.5, 0.3, 0.1)
5	(0.258, -0.296, 0.648)	(0.5, -0.1, 0.1)

为了说明本系统的性能,用双目视觉传感器替代 Kinect 传感器,就构成了双目视觉机器人系统,在其余配置及方法不变的情况下,将本系统的所有算法应用于应用双目视觉视系统,并应用视差法获得抓取点三维坐标,将本系统实验结果与双目视觉机器人抓取系统的实验结果做了对比,如表2所示.

表2 两种方法结果耗时对比

	平均耗时
双目视觉机器人	0.88s
Kinect 机器人	0.43s

可以看出,应用 Kinect 作为视觉传感器在获取目标深度信息时准确度较高.在得到的抓取点位置的精度和速度上, Kinect 比双目视觉要更加精确且快速.

### 4 结论

在本系统中,将 Kinect 作为视觉传感器,通过将视觉定位技术应用于织物抓取机器人,代替了人工劳动力提高了生产效率,提高了机器人的智能化水平和工作灵活性,利于扩大其在工业生产中的应用范围.实现了对目标物体的快速自主抓取,初步的研究结果表明,应用上述技术和方法能够较精确地实现机器人对目标物体的抓取和定位,本系统实现了对特征较为明显的单目标的抓取,而基于颜色的跟踪也存在它自身的缺陷,如果背景跟运动目标的颜色分布近似,可能会有检测失败的情况.于是在接下来对于特征不是

很明显或多目标的情况的研究中,我考虑先使用帧差法来检测出运动目标,可以只对去掉背景的图像使用 Camshift 算法来进行颜色匹配,这样就有效消除了背景的干扰,提高匹配准确率.对多目标的识别和选择性抓取及提高系统的快速性还有待进一步研究.

### 参考文献

- 1 吴俊君,管贻生,张宏,等.仿人机器人视觉导航中的实时性运动模糊探测器设计.自动化学报,2014,40(2):267-276.
- 2 王麟琨,徐德,谭民.机器人视觉伺服研究进展.机器人,2004,26(3):277-282.
- 3 Ganganath N, Leung H. Mobile robot localization using odometry and Kinect sensor. 2010 IEEE International Conference on Emerging Signal Processing Applications (ESPA). NJ, USA. 2010. 91-94.
- 4 庄佳园,徐玉如,万磊,等.基于雷达图像的水面无人艇目标检测技术.哈尔滨工程大学学报,2012,33(2):129-135.
- 5 Ravari ARN, Taghirad HD, Tamjidi AH. Vision-based fuzzy navigation of mobile robots in grassland environments. Proc. of International Conference on Advanced Intelligent Mechatronics. IEEE. ASME Press. 2009.
- 6 Zhang ZY. Flexible camera calibration viewing a plane from unknown orientations. Proc. of International Conference on Computer Vision (ICCV'99). Corfu, Greece. 1999,1. 666-673.
- 7 高文娟,李健.基于 OpenCV 的摄像机标定问题研究.计算机与数字工程,2008,36(12):128-130.
- 8 刘睿祯,于仕琪.OpenCV 教程.北京:北京航空航天大学出版社,2009.
- 9 Heresy. Kinect+OpenNI 的深度值 .http://viml.nchc.org.tw/blog/paper\_info.php?CLASS\_ID=1&PAPER\_ID=278. [2012-09-27].
- 10 Opensource. Kinect 运用 OpenNI 产生点云.http://blog.csdn.net/opensource07/article/details/7804246.[2012-09-27].
- 11 侯志强,韩崇昭.视觉跟踪技术综述.自动化学报,2006,32(4):603-617.
- 12 Comaniciu D, Ramesh V, Meep P. Real-time tracking of non-rigid objects using mean shift. Proc. of IEEE Conference on Computer Vision and Pattern Recognition. New York. IEEE Press. 2000. 142-149.
- 13 刘冬冬.基于双目视觉和 CamShift 算法的目标检测与跟踪[硕士学位论文].济南:山东大学,2006
- 14 邬大鹏,程卫平,于盛林.基于帧间差分和运动估计的 CamShift 目标跟踪算法.光电工程,2010,37(1):55.