

DOI: 10.13652/j.issn.1003-5788.2020.02.018

基于 RGB-D 弹性可形变物体跟踪识别控制 披萨厨师机器人方法研究

Research on pizza chef robot control method based on RGB-D
sensor elastic deformable object tracking

李凌乐 李瑞华

LI Ling-le LI Rui-hua

(许昌学院工程技术中心, 河南 许昌 461000)

(Engineering Training & Technology Center, Xuchang University, Xuchang, Henan 461000, China)

摘要:以披萨厨师机器人为研究对象,提出了一种基于 RGB-D 传感器所提供的实时点云数据对弹性形变的 3D 无纹理对象进行跟踪的方法。着重分析了视觉分割的时间一致性和实时性问题;通过有限元分析的方法建立了带有轮廓加权的同向旋转 FEM 模型;论述了注册点云数据方法,提出了跟踪合成数据与披萨面坯实际图像的算法,采用模拟开放框架架构模拟器,验证了算法的有效性。

关键词:机器人;RGB-D 物体识别;Grabcut 分割;弹性形变建模

Abstract: Aiming at the tracking problem of elastic deformable objects, a control method based on RGB-D sensor of pizza chef robot was proposed in this study. An in-depth study on the difficult problem of dough tracking and recognition in pizza making. A method to track 3D textured objects with elastic deformation based on real-time point cloud data provided by RGB-D sensor was also proposed. Firstly, the Grabcut in the visual segmentation was described in detail, and the temporal consistency and real time of the visual segmentation were emphatically analyzed. Secondly, the FEM model with contour weighting was established by finite element analysis. Finally, the method of registration point cloud data was discussed in detail to realize the tracking and recognition of deformable objects. In order to evaluate the proposed method, the validity of the proposed algorithm was verified by using the simulated open frame architecture (SOFA) simulator. The research results showed a

基金项目:河南省科技厅科技攻关重点项目(编号:172102210094)

作者简介:李凌乐,女,许昌学院讲师,硕士。

通信作者:李瑞华(1978—),男,许昌学院副教授,博士。

E-mail: lilingle2008@163.com

收稿日期:2019-09-23

positive significance for the quick response of pizza chef robot in food cooking process.

Keywords: robot; RGB-D object recognition; Grabcut segmentation; elastic deformation modeling

基于视觉的非刚性物体跟踪技术在计算机视觉、计算机图形学和机器人社区中逐渐成为研究的热点方向^[1]。在诸如增强现实、医学成像、机器人操纵等领域,可通过这项技术处理各种各样的物体,如纸张、橡胶、黏性流体、电缆、食物等,具有潜在的广泛应用场景^[2]。近年来,机器人学术界及工程界越来越多地关注涉及食品操作任务的机器人,这有助于提高食品制作的效率、品控,同时可保障食品的卫生^[3]。图 1 为披萨厨师机器人的操作场景示意图,通过对人形双臂的操作,完成披萨的制作,需要机器人支持可形变物体的复杂动态操纵任务。



图 1 机器人制作披萨

Figure 1 The robot makes pizza

人形披萨厨师机器人控制中的核心问题即机器人的感知,通过传感器来指引机器人可形变物体的跟踪,力求良好的准确性、稳定性以及实时性。处理感知系统中的形变问题需考率诸多影响因素,如对所选材料特性进行建模,以及使该模型适配视觉和距离数据^[4]。此外,还需考虑实时问题,这是机器人动态操作的硬性要求^[5]。中

国对于可形变物体跟踪识别技术的研究尚处于起步阶段,尽管许多研究提出了用有效的实时技术来处理经历等距或微小弹性形变的 3D 表面(如纸张、衣服等),但未能解决更大的弹性形变问题。对于无纹理的弹性物体的建模通常需要依赖于有限元方法(Finite Element Method, FEM),但典型的有限元方法对于待识别物体发生较大形变以及快速形变后的识别效果仍不够理想。因此,试验拟提出一种实时跟踪方法,使用由 RGB-D 传感器提供的视觉和距离数据,通过跟踪大幅度形变和快速刚性运动来处理无纹理的弹性物体,在应用有限元方法的基础上对模型进行改进,建立一种带有轮廓加权的同向旋转 FEM 模型,实现可形变物体的跟踪识别,对提出跟踪合成数据和实际的披萨面坯实际图像应用模拟开放框架架构(SOFA)模拟器,验证文中所述算法的有效性。

1 视觉分割

逐帧跟踪系统架构如图 2 所示。为达到可形变对象跟踪的目的,对所考虑对象的视觉分割,使用基于图像分割的方法以确保时间一致性。使用生成的分段点云,执行刚性迭代最近点(ICP)以估计从点云到网格的刚性变换。使用上一步得到的分段点云,计算从点云到网格,施加在网格顶点上的外部线性弹性力,在相反方向则参考最接近点的对应关系。应用意义在于提高了数值分辨率,自动发现目标。捕捉到清晰的目标特征,以利于实时的判断和事后对照取证。

1.1 Grabcut 分割

Grabcut算法是一种由微软公司提出的图像分割手

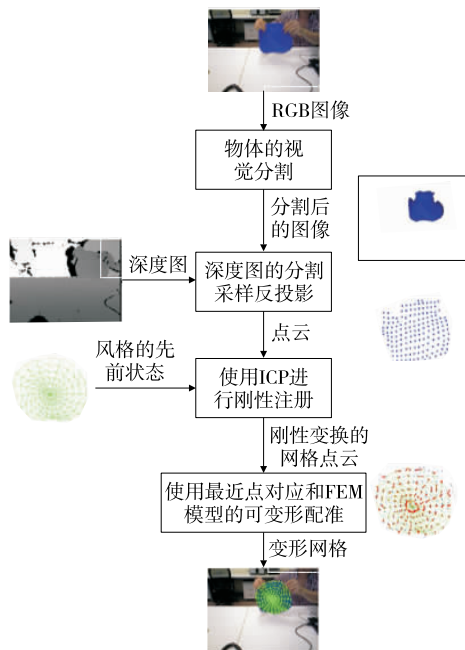


图 2 可形变对象跟踪方法

Figure 2 Deformable object tracking method

段,通过在图像中圈定一个方框,框外部分均为背景,而框内部分则可能是前景或背景^[6]。该算法被广泛应用于图像分割中,具有高效的特征。Grabcut 算法基于前景(对象)和背景的统计模型将视觉双层分割任务转化为能量最小化问题^[7]。

对于输入图像 I ,用 $\alpha = \{\alpha_i\}_{i=1}^N$ 表示该组像素的未知二进制标记的集合(背景像素 $\alpha_i = 0$,前景 $\alpha_i = 1$)。估计标签的 α 可表示为基于能量的马尔可夫随机场目标函数 $E(\alpha)$ 相对于 α 的最小化^[8]。

$$\hat{\alpha} = \arg \min_{\alpha} E(\alpha), \tag{1}$$

$$E(\alpha) = E_{\text{data}}(\alpha) + \gamma E_{\text{smooth}}(\alpha), \tag{2}$$

$$E_{\text{data}}(\alpha) = \sum_i U_i(\alpha_i), \tag{3}$$

式中:

E_{data} ——数据能量项,定义像素 p_i 属于前景或背景的观察概率为 $p(p_i | \alpha_i)$, $U_i(\alpha_i) = -\log[p(p_i | \alpha_i)]$ 为像素上观察到的图像数据,如密度、颜色、位置等;

E_{smooth} ——平滑能量项,其目标是支持像素内的平滑度或空间相干性。

为了计算能量最小化问题的最优解并确定 $\hat{\alpha}$,采用图切割最小化算法,提供分段帧 I^s ,用于数据能量函数的统计模型是基于颜色分布的高斯混合模型(Gaussian Mixture Model, GMM),其针对前景层和背景层学习,由用户通过在初始图像上的前景周围手动定义的边界框来确定。此外,此边界框外的像素明确分配给背景图层 [$U_i(\alpha_i = 0) = \infty$],而其内部未知,因此能量最小化仅在边界框内有效。

1.2 时间一致性和实时问题

一旦通过用户交互来分割初始图像,则类似地处理以下帧,该操作并不是分割整个帧,而是围绕分割区域的轮廓周线逐帧更新有效分割区域,据此提供分割过程中时间的一致性。如图 3 所示,提取先前分割的前景轮廓周线,并在其上计算距离变换,从而向这些轮廓提供带符号的距离图 d (负的外部,正的内部)。

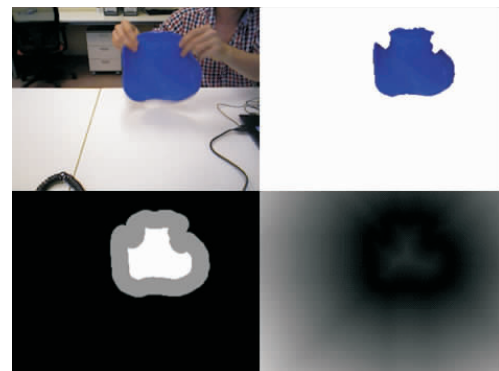


图 3 分割的时间一致性

Figure 3 Time consistency of segmentation

根据距离图上的固定阈值 d_t , 在轮廓周围定义一条窄条 ($|d_i| < d_t$, 灰色区域, 其中像素 p_i 标签未知), 通过设置 $U_i(\alpha_i = 1) = \infty$ 将该区域分配给条带内侧的前景 ($d_i < -d_t$, 白色区域), 否则设定 $U_i(\alpha_i = 0) = \infty$ 分配给背景 ($d_i > d_t$, 黑色区域)。以这种方式, 确保了时间一致性, 由于能量最小化仅在该条带内, 在先前分割边界附近有效, 避免外部或内部异常值, 并可大幅度减少计算, 使得该分割过程对于假定前景和背景颜色分布恒定的序列是有效的。

2 弹性形变的建模

试验主要处理的是可能发生大幅度弹性形变的物体, 因此主要问题在于相关物理模型的定义。相较于质量弹簧系统的有限差分, FEM 方法提供了一种真实的物理模型, 依靠连续介质力学能更好地表达相关的特性^[9]。考虑建模精度问题, 采用具有四面体单元的体积线性 FEM 方法, 其性能优于其他拓扑结构, 同时兼顾了计算效率, 并且便于使用拓扑结构进行网格划分。

2.1 形变场的基函数

将元素 e 上的形变场 u_e 近似为其 4 个顶点的位移矢量 \hat{u}_e 的连续插值, 通过 3×12 矩阵 $N_e(x)$ 描述转换关系, 其包含元素的多项式基函数为:

$$u_e(x) = N_e(x)\hat{u}_e, \quad (4)$$

其中 x 为元素中的一个点, $\hat{u}_e = x_e - x_{e,0}$, x_e 、 $x_{e,0}$ 分别为 e 的 4 个顶点的形变后和初始的世界坐标。

2.2 弹性形变建模

为对弹性形变进行建模, 根据无穷小应变理论计算四面体内的柯西线性应变张量 ϵ_e , 其与 \hat{u}_e 为线性关系:

$$\epsilon_e = L_e \hat{u}_e, \quad (5)$$

其中 L_e 为一个 6×12 常矩阵。

为将应力与应变联系起来, 通过胡克定律的线性弹性理论可知, 对于连续的各向同性材料, 无穷小的应力张量 σ_e 可表示为:

$$\sigma_e = C_e \epsilon_e, \quad (6)$$

其中 C_e 为 6 阶对称矩阵, 取决于材料的两个弹性参数——杨氏模量 E 和泊松比 ν 。

通过式(5)、(6)可得:

$$\sigma_e = C_e L_e \hat{u}_e. \quad (7)$$

e 中的应变能量可通过 σ_e 、 ϵ_e 和元素的体积 V_e 来计算, 进而可以获得施加在 e 的 4 个顶点上的内部弹性力 f_e 。根据式(5)、(7), f_e 与顶点的位移 \hat{u}_e 线性相关。

$$f_e = K_e \hat{u}_e = V_e L_e^T C_e. \quad (8)$$

$K_e = V_e L_e^T C_e L_e$ 为 e 的 12×12 刚度矩阵, 对平移变换不敏感, 但在对元素进行大幅度旋转时, 建模是不准确的, 是因为非线性效应导致合力非零。在顶点的当前和初始状态之间的变换矩阵上使用极性分解来提取刚性旋

转矩阵 R_e 。则刚度矩阵 K_e 可以相对于该旋转进行调整, 以便适应旋转变换:

$$f_e = R_e K_e \hat{u}_e^r = R_e K_e (R_e^{-1} x_e - x_{e,0}). \quad (9)$$

当 $\hat{u}_e^r = R_e^{-1} x_e - x_{e,0}$ 时, $R_e^{-1} x_e$ 将 e 顶点的后旋转变换坐标转换为未旋转帧, 通过乘以 R_e 将力 $K_e \hat{u}_e^r$ 旋转到当前形变的元素。通过这种方式, 整个网格上合力为零, 与非线性 FEM 方法相比, 可预先计算 K_e , 确保计算效率。

3 注册点云数据

形变配准问题在于将 RGB-D 传感器提供的点云数据与已知的四面体网格适配, 基本思想是导出点云对网格施加的外力, 并将外力与物理模型计算的内力整合至求解所得机械方程的数值求解中。这些外力是基于点云和网格之间的点对点对应计算的, 放宽了具有纹理对象或具有粗糙表面对象的限制, 可以提取和匹配其中 2D 或 3D 关键点。可以考虑通过离线自动重建和网格化技术构建网格, 通过对自旋图像或局部 3D 特征的一些学习和识别解决初始化问题。此外, 假设所选材料杨氏模量和泊松比为已知的。

3.1 分段和采样点云

使用获取的 RGB 图像序列从其背景和遮挡中分割所需对象, 由于不依赖于一些独特的视觉特征, 为避免模糊, 深度传感器提供的点云仅限于所考虑对象, 在与背景匹配过程中或具有遮挡形状, 并能处理从输入点云到网格的对应关系。

在帧 k 处, 使用分割图像 I^s , 通过将原始输入深度图 D 与 I^s 中的分割区域对齐及交叉来获得分割深度图 D^s 。然后通过传感器框架中反投影 D^s , 获得所需分段点云 $Y = \{y_j\}_{j=1}^n$, 其中 y_j 为传感器帧中 3D 点。考虑到计算问题, 通过预先在图像平面中的规则网格上采样 D^s 来限制 Y 的大小。

3.2 最近点刚性迭代

根据刚性平移和旋转变换记录观察到的分段点云 Y , 预先将对象的网格设定为刚性。在 3.1 的计算状态中, 对于帧 $k-1$, 首先定义 $X = \{X_i\}_{i=1}^n$ 为网格的顶点集合。对于 RGB-D 数据变换, 在 Y 和网格可见表面 X_V 的顶点间应用刚性 ICP 算法。对于帧 $k-1$, 通过对对象的 3D 投影网格渲染深度图执行可见性测试来确定 X_V 。通过这种快速收敛过程, 可以跟踪快速刚性运动, 并可实现非刚性过程的初始化。

3.3 可形变的迭代最近点

以非刚性方式向网格注册分段点云, 采用类似 ICP 程序, 假设通过 3.2 中估计的刚性变换更新集合 X 、 X_V 。

3.3.1 最近邻关系 采用 Kd 树搜索, 通过分段点云到网格可见表面确定最近邻对应关系。此过程提供了最近邻居集合 $N_{X_V} = \{nm^Y(X_i)\}_{i=1}^n$ 、 $N_Y = \{nm^{X_V}(y_j)\}_{j=1}^n$ 。

由于依赖于唯一的几何接近度可能导致使用单点对点匹配的不一致,因此采用两组对应关系。如图 4 所示,从分段点云到网格,对应性使得能够跟踪,如在拉伸力下的膨胀形变,对于该拉伸力,观察到分段点云 Y 将在网格 X_V 可见表面上扩展, Y 相对于 X_V 扩展区域可与 X_V 外部区域匹配(右侧红色箭头)。这些对应关系还能处理遮挡和分割错误,对象的相应未观察区域不会影响 X_V 底层区域。相反,从 X_V 到 Y ,对应物更适合跟踪压缩动作下的收缩形变, X_V 外部区域与被压缩物体的观察点云 Y 外部区域相互匹配(左侧绿色箭头)。未观察到区域将影响与 X_V 的最近区域匹配的下面区域 X_V 。

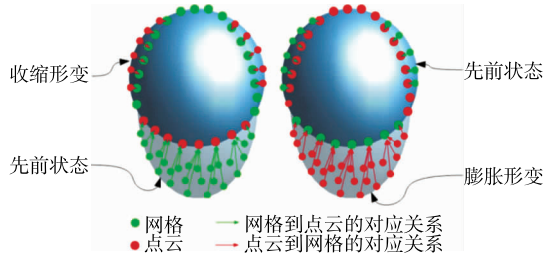


图 4 网格和点云间的对应关系

Figure 4 Correspondence between grids and point clouds

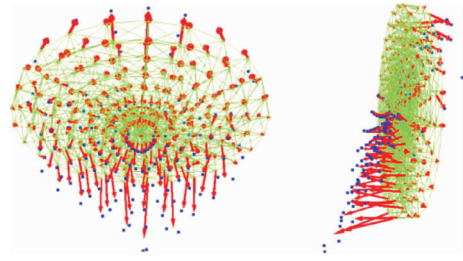
3.3.2 计算外力 基于由 N_{X_V} 、 N_Y 给出的两组网格和点云的对应关系,在 X_V 中对每个 x_i 施加的外部弹力 f_{ext} 可按式(10)计算。

$$f_{ext}(x_i) = k_{ext}(x_i - y_i^f), \quad (10)$$

$$y_i^f = \begin{cases} \lambda n n^Y(X_i) + (1 - \lambda) \frac{1}{n_{N_i}} \sum_{y_j \in N_Y} y_j, & n_{N_i} > 0; \\ \lambda n n^Y(x_i) + (1 - \lambda)x_i, & n_{N_i} = 0, \end{cases} \quad (11)$$

其中 $N_Y = \{y_j \in Y \mid nm^{X_V}(y_j) = x_i\}$ 为 Y 中一组点,其最近邻居为 x_i ; y_i^f 充当吸引 x_i 的虚拟点,由多点云到网格对应产生的 y_i^f 分量按 $1/n_{N_i}$ 缩放,其中 n_{N_i} 的维度等同于 N_Y ,与单个网格到点对应的分量是均匀的;若 N_Y 为空,则缺少点云到网格的对应关系,被顶点 x_i 的自回归所取代,迫使其保持在当前位置; k_{ext} 为外部弹簧弹力刚度,可使用质量弹簧模型和线性有限元模型间的等效性来相应地设置杨氏模量。通过这种方式建模,使用等效弹簧的视觉数据仅影响可见顶点的外力。

如图 5 所示,展示了向量 $x_i - y_i^f$ 。点云中异常值可引起异常对应,导致在某些顶点上施加异常力。为解决异常对应,丢弃点云中的点,由于其与网格中最近邻的点对点距离相对于整个点对点的平均值和标准偏差高于某个阈值。此情况下, $f_{ext}(x_i) = 0$ 。而关于 X 中不可见的点,设置 $f_{ext}(x_i) = 0$ 。整套力最终在大小为 n_X 的向量文本中连接起来。



$k_{ext} = 1 \text{ N/m}, \lambda = 0.7$

图 5 外力施加在网格顶点上

Figure 5 External forces are applied to the vertices of the grid

3.3.3 使用轮廓的加权力 基于 3D 几何形状建立对应关系,位于网格遮挡轮廓上的顶点被吸引到点云中的扩展区域。通过加权网格可见表面顶点来增强吸引力,给定其与投影网格的遮挡轮廓距离。基于投影网格的深度图 d^M 计算网格遮挡轮廓距离图,进而计算顶点 x_i 权重 w_i 。

$$w_i \propto e^{-\frac{d_i^M}{\sigma}}, \quad (12)$$

式中:

d_i^M ——从 x_i 到投影网格的最近轮廓距离,mm;

σ ——根据经验设定的参数。

对 w_i 归一化,以便得到观察概率,按式(13)计算力。

$$f_{ext}(x_i) = w_i k_{ext}(x_i - y_i^f). \quad (13)$$

3.3.4 用于计算变形的数值解算器 估计网格形变包括求解基于拉格朗日动力学方程,涉及内部和外部的常微分方程:

$$M \ddot{x} + C \dot{x} + f = f_{ext} \quad (14)$$

$$f = K'x + f_0, \quad (15)$$

式中:

x ——包含 X 、 M 中顶点位置的 n_X 阶向量;

C —— $n_X \times n_X$ 质量和阻尼矩阵;

K' —— $n_X \times n_X$ 全局刚度矩阵。

使用得到的网格顶点的估计位置 x 可以更新 X 。

4 试验

为了评估所提出的方法,测试各种物体形变和条件。对于非刚性注册阶段,采用模拟开放框架架构(SOFA)模拟器,该模拟器能处理各种物理模型并实时演化模拟。

4.1 跟踪合成数据结果

依靠 SOFA 框架,首先生成一个涉及圆柱形弹性物体形变序列,由 FEM 同向旋转方法建模,具有 $E = 800 \text{ Pa}$ 的杨氏模量和 $\nu = 0.3$ 的泊松比。基于手工设计的半径 \times 高度为 $0.11 \text{ m} \times 0.02 \text{ m}$ 的圆柱面三角网,使用 3D Delaunay 三角剖分,通过 CGAL 库生成体积四面体网格,得到网格由 1 369 个元素和 497 个顶点组成,如图 6 所示。

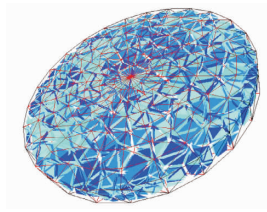


图 6 处理后的网格

Figure 6 The processed grid

模拟形变并生成序列,在 Z 方向上施加弹性拉伸力如图 7 所示。对 4 种建模方法的效果进行对比,分别获取第 15,50,55,150 帧的形变识别情况,表明试验所提出的方法效果最好。

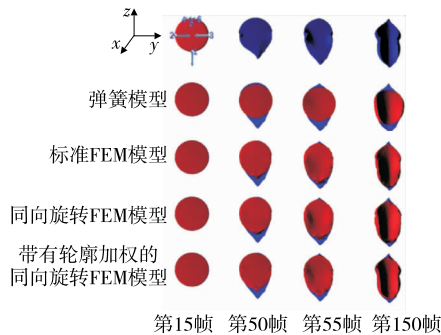


图 7 形变跟踪结果

Figure 7 Deformation tracking result

4.2 实际数据结果

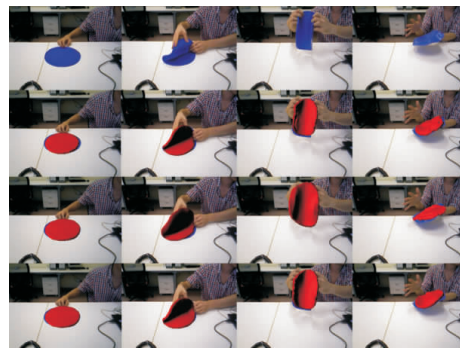
为了对真实数据进行试验,从校准的 RGB-D 相机华硕 Xtion,320×240 RGB 和正在处理的深度图像中获取被调查场景的点云,使用带有 NVIDIA GTX 1060 显卡的标准笔记本电脑以及 2.4 GHz Intel Core i7 CPU。分段过程涉及循环,由于需要快速的实时性能,主要依赖于 CUDA 实现。试验对象涉及用硅制无纹理光滑弹性物体,网格尺寸 0.12 m×0.01 m,包括 574 个顶点和 1 675 个元素。

在第一个特征序列中,物体经大幅度刚性运动、各种等长和弹性形变,设定 $k_{ext}=1\text{ N/m}$, $E=300\text{ Pa}$, $\nu=0.3$,跟踪识别结果如图 8 所示,使用带有轮廓加权的同向旋转 FEM 方法与其他模型方法进行比较。

由表 1 可知,试验所提出的带有轮廓加权的同向旋转 FEM 模型在时间开销上并没有劣化,不会影响控制系统的实时性。

5 结论

试验通过 RGB-D 传感器,采用视觉分割和有限元分析方法建立了弹性形变物体的模型,并实现了点云数据的注册。试验提出的带有轮廓加权的同向旋转 FEM 模型可较好地实现对无纹理形变物体的跟踪,在保障识别



第 1 行显示输入 RGB 图像;第 2 行显示用质量弹簧模型跟踪对象的 3D 网格;第 3 行用标准 FEM 模型跟踪;第 4 行用提出的方法

图 8 披萨面团的跟踪效果

Figure 8 Pizza dough tracking effect

表 1 4 种模型算法各个阶段的计算时间

Table 1 The calculation time of each stage of each of the four model algorithms

阶段	质量弹簧模型	标准 FEM 模型	同向旋转 FEM 模型	带有轮廓加权的同向旋转 FEM 模型
分割	10.7	10.5	10.7	10.7
刚性 CP	3.0	2.5	2.7	2.6
可视性	8.1	8.2	7.6	7.4
外部力处理	3.4	3.5	3.5	4.0

效果的前提下,并未由于模型复杂度的提升而影响运算的时间开销。试验缺少对更大的弹性形变的多种目标进行研究分析来验证算法的有效性,后续可增加试验的多目标验证来完善算法的正确性。

参考文献

- [1] 王婧林, 黄剑. 基于动态减背景图像处理算法的可变形线性物体识别[J]. 天津科技大学学报, 2013, 2(6): 65-70.
- [2] 关印, 王向军, 阴雷, 等. 基于物体表面形貌的单机视觉位姿测量方法[J]. 光电工程, 2018, 4(1): 43-50.
- [3] 黄海松, 张松松, 姚立国, 等. 基于机器视觉的物体阶跃表面检测算法研究与实现[J]. 激光杂志, 2019, 40(3): 45-48.
- [4] 向程瑜, 王冬丽, 周彦, 等. 基于 RGB-D 融合特征的图像分类[J]. 计算机工程与应用, 2018, 54(8): 178-182.
- [5] 卢良锋, 何加铭, 谢志军, 等. 基于深度学习的 RGB-D 物体识别算法[J]. 移动通信, 2015(10): 52-56.
- [6] 周良芬, 何建农. 基于 GrabCut 改进的图像分割算法[J]. 计算机应用, 2013, 33(1): 49-52.
- [7] 王万耀, 段先华, 徐丹, 等. 基于显著性的 Grabcut 图像分割方法[J]. 计算机工程, 2018, 44(7): 230-236.
- [8] 辛月兰. 基于超像素的 Grabcut 彩色图像分割[J]. 计算机技术与发展, 2013, 23(7): 48-51.
- [9] 李元元, 赵志敏, 潘崎, 等. 基于有限元仿真的板状物体的变形分析[J]. 理化检验: 物理分册, 2012, 48(1): 7-10.