

广义Hough变换在遮挡图像识别中的应用

姚立健, 丁为民^{*}, 赵三琴, 杨玲玲

(南京农业大学工学院, 南京 210031)

摘要: 介绍了一种基于广义 Hough 变换的空间有部分遮挡的茄子目标识别方法。从生物学的角度, 用广义柱近似等效样本茄子形状; 通过坐标转换获得了不同位姿的立体目标在平面上的投影边缘, 并建立茄子外形的 4 维参数索引表; 采用一种广义 Hough 变换计算茄子目标的潜在位置, 并通过比较各旋转角度下的“面积比”, 筛选出目标实际的位姿。试验表明: 这种广义 Hough 变换检测目标位置和姿态的相对误差分别约在 1.3% 和 5.5% 以下, 说明其对空间不同位姿、部分遮挡情况下目标的识别具有良好的效果。

关键词: 茄子, 轮廓计算, 遮挡, 广义 Hough 变换

中图分类号: TP391.41

文献标识码: A

文章编号: 1002-6819(2008)-12-0097-05

姚立健, 丁为民, 赵三琴, 等. 广义Hough变换在遮挡图像中的应用[J]. 农业工程学报, 2008, 24(12): 97-101.

Yao Lijian, Ding Weimin, Zhao Sanqin, et al. Applications of the generalized Hough transform in recognizing occluded image[J]. Transactions of the CSAE, 2008, 24(12): 97-101.(in Chinese with English abstract)

0 引言

在果蔬收获机器人的研究中, 采摘目标被茎、叶遮挡或者目标之间相互重叠时, 会造成对采摘目标位置的误判, 因而影响目标定位的精度。陈振羽^[1]等人通过测量目标和景物轮廓的局部极大张角从含有遮挡目标的景物图像中识别出目标。申金媛^[2]、Cozar^[3]等人介绍了两种识别外表面是平面的三维立体的方法。

Hough变换(Hough Transform, HT)^[4]及其拓展-广义Hough变换(Generalized Hough Transform, GHT)^[5]自提出以来, 因其良好的全局判别能力在识别遮挡和被噪音干扰的目标问题上得到广泛应用。谢志勇^[6]、蔡键荣^[7]等人采用Hough变换检测出圆形草莓和柑橘。饶洪辉^[8]、侯学贵^[9]等人检测出农作物行中心线以便于施药和导航; 程义民^[10]、李智磊^[11]等人采用广义Hough变换分别检测出工作台上堆集的有部分被遮挡的平面工具和破碎的平面图形。

上述研究对象主要是被部分遮挡的平面目标或多面体目标。自然生长的果蔬目标既不可能是平面物体也不可能是平面立体, 且在空间中具有不同的位置和姿态。本文以茄子为例, 在汲取前人研究成果的基础上, 首先采用广义柱(generalized-cylinders)^[12]描述三维茄子形状, 并提取外表的二维边缘轮廓^[13], 然后采用GHT, 对不同视角和不同遮挡程度的茄子位姿的判别进行了研究。GHT将变量空间物体位姿的计算问题, 变换为参数空间

的累加问题, 该方法不必得到物体全部特征, 为空间中部分遮挡物体的识别, 提供了一个良好的途径。

1 茄子三维表达方法和二维投影轮廓的计算

1.1 用广义柱描述茄子 3D 形状

GHT 首先要求将样本茄子的平面轮廓建成参数表, 这就需要用一或一组已知的数学函数曲线或曲面来近似表示复杂的茄子外形。Cozar 等人^[3,14-16]用了不同的方法对果蔬的形状进行了描述。从生物学角度看, 一般果蔬的形状都与圆球形有关, 这是因为: 同样的体积, 球形的表面积最小, 这样保水和保温性能比较好, 害虫的立足之处也少; 另外球形能承受较大的应力、重力作用, 能经受风雨的吹打。为简化研究, 通过观察, 本文假定: 1) 同品种、同等级茄子外观都是接近的; 2) 茄子可以看成是一个半径呈按一定规律变化的圆沿着与圆平面垂直的弯曲的中心轴平移所得到的广义柱体; 3) 该中心轴的弯曲程度与茄子的弯曲程度一致, 茄子的中心轴为平面上的一条曲线。

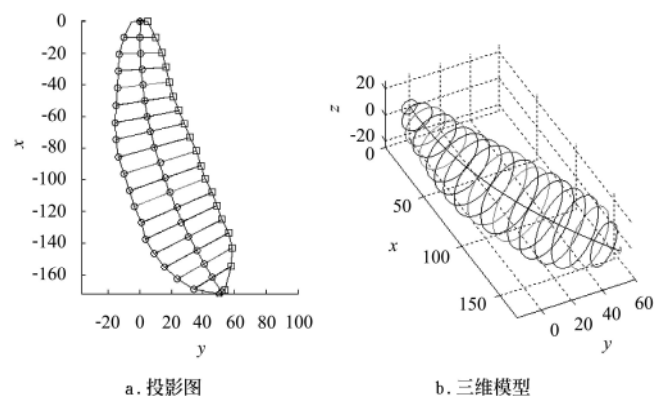


图1 样本茄子的三维模型

Fig.1 Three-dimensional model of sample eggplant

收稿日期: 2007-11-04 修订日期: 2008-05-27

作者简介: 姚立健(1974—), 男, 江苏盐城人, 博士生, 主要从事农业机器人、计算机图像处理和计算机控制等研究。南京 南京农业大学工学院, 210031。Email: yao_njau@163.com

*通讯作者: 丁为民(1957—), 男, 安徽合肥人, 教授, 博士生导师, 主要从事农机现代设计理论与方法、设施农业及其环境控制研究。南京 南京农业大学工学院, 210031。Email: wmding@njau.edu.cn

基于以上假定,参考标准 NY/T 581-2002,选择一个一等品“种都墨茄”为样本,以茄子中心轴所在的面为透视投影面(本文称为 xoy 面,如图 1a,下同),获取茄子的投影图像,利用数学形态学的方法提取骨架,骨架即为茄子中心轴;用二次函数拟合中心轴,使其平滑且无拐点;将拟合后的中心轴等分为若干点,过这些点分别作曲线的垂线,与边缘两交点的距离即为过该点茄子截面圆的直径;最后根据比例关系换算成实际坐标和实际直径。以茄子中心轴所在平面为 xoy 平面,果梗为原点的三维茄子广义柱模型如图 1b 所示。

1.2 用 2D 边缘描述 3D 任意形态的茄子

由三维变换矩阵^[17]:

$$T_{3D} = [R_{\alpha,\beta,\gamma}, p_{x,y,z}] \quad (1)$$

可得,描述茄子在空间中的位置和姿态,至少需要 6 个参数。式中, $R_{\alpha,\beta,\gamma}$ 表示旋转矩阵,共有 3 个旋转角参数, $p_{x,y,z}$ 表示位置矢量,也有 3 个坐标参数。因为茄子一般都按固定行距种植,因此在行间移动的收获机器人摄像头距离目标的距离可以近似认为不变;又因为茄子在无阻碍的情况下,一般都长成自然下垂状。因此为简化研究,本文忽略了目标的深度位置信息 z 和其绕 y 、 z 轴的旋转姿态信息 β 和 γ , 仅仅保留了平面位置信息 x 、 y 和绕 x 轴的旋转信息 α 。具体坐标系的描述可参看图 1。

如图 2,当茄子的中心轴与 xoy 平面重合时,旋转角设为 0° ,中心轴在三维空间的参数方程 $[x, y, z]^T$ 为:

$$[x, y, z]^T = \begin{bmatrix} x \\ ax^2 + bx + c \\ 0 \end{bmatrix} \quad (2)$$

则中心轴在 xoy 平面投影方程为: $y = ax^2 + bx + c$;茄子在 x_k 处与中心轴垂直的截面圆为 O_k ,此状态下,圆 O_k 在 xoy 平面上的投影为一线段 ab ;当茄子绕 x 轴旋转 α 角度后,中心轴在三维空间的参数方程 $[x', y', z']^T$ 为:

$$\begin{aligned} [x', y', z']^T &= Rot(x, \alpha) \times [x, y, z]^T \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ ax^2 + bx + c \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x \\ (ax^2 + bx + c) \cos \alpha \\ (ax^2 + bx + c) \sin \alpha \end{bmatrix} \end{aligned} \quad (3)$$

此时中心轴在 xoy 平面上的投影方程为: $y' = (ax^2 + bx + c) \cos \alpha$, 圆 O_k 旋转到了 O'_k 位置,此时该圆在 xoy 平面上的投影为一椭圆 O'_k , 椭圆 O'_k 长轴 cd 长度等于圆 O_k 直径,且长轴 cd 垂直于二次曲线 y' 在 O'_k 处的切线。

该椭圆长轴 cd 上的两顶点 c 、 d 即为茄子旋转 α 角度后其在 xoy 平面上投影的一对边缘点,以此类推,就可

以得到不同旋转角度下,茄子在 xoy 面上的投影边缘。对这些边缘点进行曲线拟合,就可以得到旋转 α 角度后在 xoy 平面上投影的封闭的边缘曲线。如图 3,分别为绕 x 轴旋转 0° 、 30° 、 60° 、 90° 、 120° 、 150° 、 180° 时,样本在 xoy 面上投影的边缘。

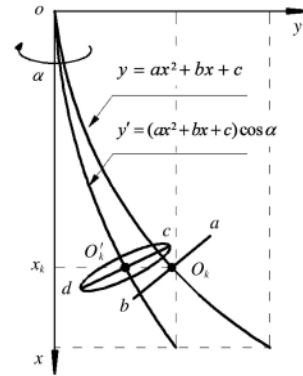


图 2 中心轴和截面圆的旋转示意图

Fig.2 Rotation schematic diagram of central axis and circle section

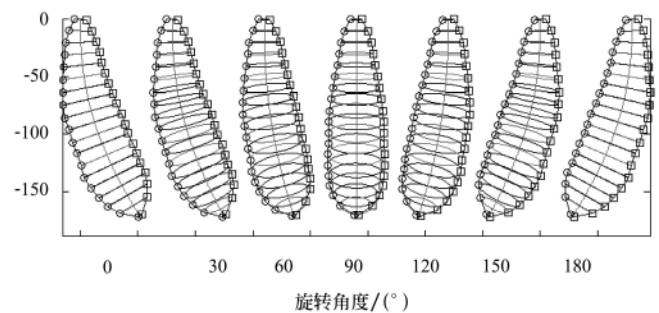


图 3 不同旋转角度下茄子 2D 边缘投影

Fig.3 Two-dimensional eggplant edge projection of different rotation angles

2 茄子形状参数表的制作

为空间样本茄子建立 4 维参数表:第 1 维存放绕 x 轴旋转的各角度索引;第 2 维存放茄子边缘梯度角索引;第 3、4 维分别存放在各索引下的位置特征向量 r 和 φ 。

2.1 参考点 (x_c, y_c) 的选取

本文以图像的中心为参考点 (x_c, y_c) ,具体算法为:

$$\begin{cases} x_c = \frac{1}{n} \sum_{i=1}^n x_i \\ y_c = \frac{1}{n} \sum_{i=1}^n y_i \end{cases} \quad (4)$$

式中 n ——图像边缘点个数; (x_i, y_i) ——边缘点的坐标。

2.2 旋转角度索引 α_k 的确定

绕 x 轴旋转一周的角度为 360° ,但由图 3 知, $180^\circ \sim 360^\circ$ 之间的投影与 $0 \sim 180^\circ$ 之间投影是镜像对称的,因此只需计算 $0 \sim 180^\circ$ 即可。将茄子绕 x 轴旋转的角度在 $0 \sim 180^\circ$ 之间量化成 m 个的等级,在每个旋转角度 α_k ($k = 1, 2, \dots, m$) 下记录茄子在 xoy 平面上的离散的投影

边缘图像。

2.3 边缘梯度角索引 θ_i 的确定

李智磊^[11]和王鑫^[18]等人针对离散的数字图像边缘提出了两种索引项的算法。本文为获取更为精确的索引，对这些离散的边缘点按照顺时针或逆时针进行排序，对排好序的边缘点进行三次样条插值，目的是使目标的边缘光滑、连续、可导。若第 i 个和第 $i+1$ 个边缘点之间的拟合函数为： $y_i = a_i x^3 + b_i x^2 + c_i x + d_i$ ，则待测目标的任意一点 i 的梯度角索引 θ_i 为计算方式为：

$\theta_i = \arctan[-1/(3a_i x_k^2 + 2b_i x_k + c_i)]$ ，并根据需要，将 θ_i 量化成 n 个等级。

2.4 特征向量 r_i 和 φ_i 的计算

边缘特征点与参考点之间的距离为 r ， $r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$ ， r_i 与 x 轴之间的夹角为 φ_i 。可见 r_i 与目标平移无关， φ_i 与平移和缩放都无关。

3 广义 Hough 变换

样本参数表建好之后，就可以通过 GHT 对在不同位置和姿态、不同遮挡程度的待识别目标进行位姿的识别，具体步骤如下：

Step1: 在每个旋转角度 α_k 下建立 2 维累加数组 (a_k, b_k) ，一共建立 m 个累加数组。

Step2: 读入含有待识别目标的图片，为提高检测速度，避免对无效点进行检测，对图像预处理，提取待识别图像的边缘点。对离散的边缘点进行三次样条插值，计算方法同上，求取每个边缘点 (x_i, y_i) 的梯度索引 θ_i ；在 m 个旋转角度 α_k 下通过该点的梯度索引，逐个查找样本 4 维参数表，找到其对应的一组 $r_{k,i}$ 与 $\varphi_{k,i}$ 值；代入下式：

$$\begin{cases} a_{k,i} = x_i + r_{k,i}(\theta_i) \cos(\varphi_{k,i}(\theta_i)) \\ b_{k,i} = y_i + r_{k,i}(\theta_i) \sin(\varphi_{k,i}(\theta_i)) \end{cases} \quad (5)$$

反算 $a_{k,i}$ 和 $b_{k,i}$ ，并在其相应的累加器里加上 1。此时 m 个二维累加器数组 (a_k, b_k) 中，每个累加器中最大值所对应的坐标 $(a_{k-\max}, b_{k-\max})$ 就是在旋转角度 α_k 下对应潜在参考点坐标。

Step3: 筛选待测目标的参考点。计算 m 个参考点所对应的目标轮廓中所包含的待测目标的面积的比例 R_k ：

$$R_k = \frac{S_k \cap S_{object}}{S_k}; \quad (k=1, 2, \dots, m) \quad (6)$$

式中 S_k ——各旋转角度下潜在参考点 $(a_{k-\max}, b_{k-\max})$ 所对应的轮廓线包含的面积； S_{object} ——Step2 中得到的待测目标的面积； $S_k \cap S_{object}$ —— S_k 和 S_{object} 的公共面积。显然， m 个 R_k 中，最大的 $R_{k-\max}$ 所对应的旋转角度即为待测目标的空间姿态 $\alpha_{k-\max}$ ，该角度下的累加器的最大值所对应的坐标即为待测目标的位置坐标。

4 试验结果分析

试验采用的图片大小为 400×300 ；图像处理软件采用 MATLAB7.0；处理设备为 PC 机。样本参数表的旋转角索引 α_k 根据实际情况按 5° 量化成 37 个等级；梯度角索引按 1° 等级量化成 181 个等级。累加数组按 1 个像素点分别量化成 400 和 300 个等级。

试验 1: 模拟合成了样本茄子目标的不同旋转角度、不同的位置和不同遮挡程度的情况，如图 4 上面一行。验证 GHT 对目标位姿的识别能力。为直观显示 GHT 效果，图 4 下面一行给出了 GHT 后求出的参考点所对应的边缘与待测目标的叠加。图 4e 出现目标被遮挡成两部分，在 GHT 前，首先对该二值图像进行连通区域标注，在不同的连通区域内对各自边缘分别进行排序，然后各自进行 GHT，最后将两个累加器叠加，即为该图的最终 GHT 结果。表 1 给出了试验 1 的试验结果。

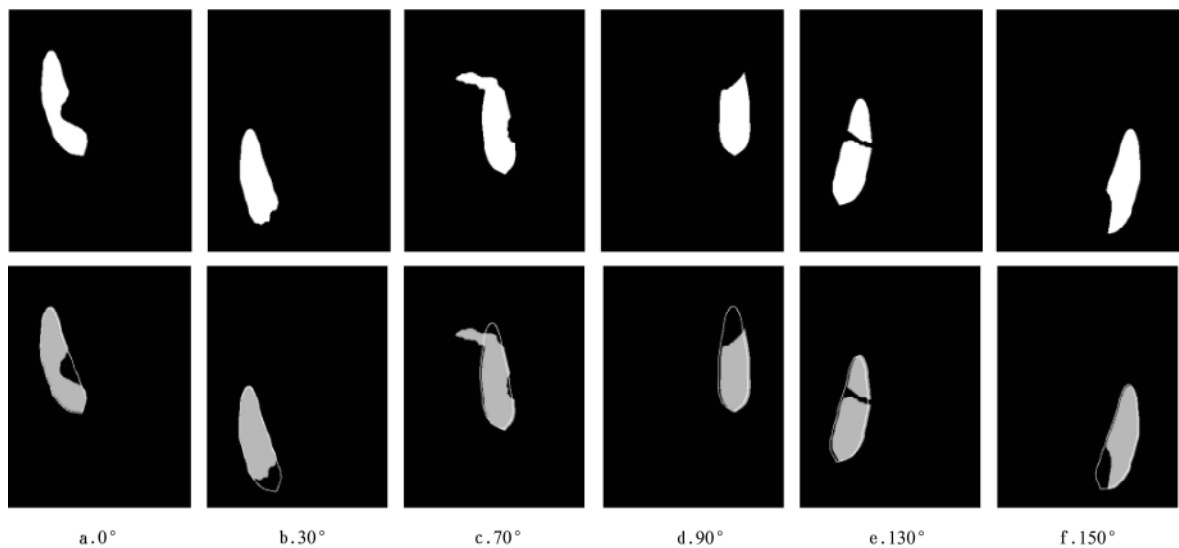


图 4 不同位姿和不同遮挡程度下 GHT 识别效果

Fig.4 Recognition effects of different poses and various occlusions

表 1 不同位姿和不同遮挡程度下 GHT 识别结果统计表
Table 1 Statistical results of different poses and various occlusions

| | 旋转角度误差统计 | | | 位置误差统计 | | | | | 送检边缘点 个数 (pixel) | 检测 时间 /s |
|---|--------------|----------------|--------------|-----------------|-------------------|-------------------|-----------------|-------------------|------------------------|----------------|
| | 模拟角度 /(°) | GHT 角度 /(°) | 角度误差 /(°) | 模拟位置 (pixel) | 实测遮挡位置 (pixel) | GHT 位置 (pixel) | 实测误差 (pixel) | GHT 误差 (pixel) | | |
| 1 | 0 | 0 | 0 | (158,87) | (160,83) | (159,86) | (2,4) | (1,1) | 426 | 8.4 |
| 2 | 30 | 30 | 0 | (288,84) | (278,79) | (289,84) | (10,5) | (1,0) | 399 | 11.3 |
| 3 | 70 | 65 | 5 | (188,156) | (184,152) | (186,153) | (4,4) | (2,3) | 462 | 22.0 |
| 4 | 90 | 100 | 10 | (156,220) | (162,222) | (154,216) | (6,2) | (2,4) | 376 | 13.1 |
| 5 | 130 | 130 | 0 | (238,89) | (256,77) | (240,85) | (18,12) | (2,4) | 443 | 19.8 |
| 6 | 150 | 150 | 0 | (289,206) | (285,209) | (287,204) | (4,3) | (2,2) | 409 | 9.3 |

注：模拟角度—模拟目标的标准旋转角度；GHT 角度—经过 GHT 检测的目标旋转角度；模拟位置—模拟目标的标准中心坐标；实测遮挡位置—遮挡后目标的实际中心，即图 4 上边一行的二值图像中心坐标；GHT 位置—经过 GHT 后检测出的目标位置。

由表 1 可得，GHT 在检测旋转角度方面误差较小，只是在旋转到 90°附近出现最大误差角为 10°，相对误差为 $10/180=5.5\%$ ，这是因为在 90°左右，目标投影的边缘梯度差异很小的原因。在检测位置方面，实测的水平方向和垂直方向的最大绝对误差分别为 12 和 18 pixels，其相对误差为 $12/300=4\%$ 、 $18/400=4.5\%$ ；经 GHT 后，最大误差分别为 4 和 2 pixels，其相对误差为 1.3%、0.5%。可见，通过 GHT 检测出来的目标位置较实测位置明显更接近模拟位置。检测时间从图像预处理开始，到筛选出实际参考点止，检测时间与待检测的像素点个数 i 以及这些像素点所能索引到的特征向量 r_i 和 ϕ_i 的数量有关，还与算法设计及精度要求有关。若在确保一定精度的情况下，适当降低参数表的量化等级，可明显提高检测速度。

试验 2：设计了多果遮挡下，目标的识别。图 5a 中模拟合成两个目标，旋转角度都为 0°，图 5b 也是两个目标，旋转角度分别为 50°和 170°。

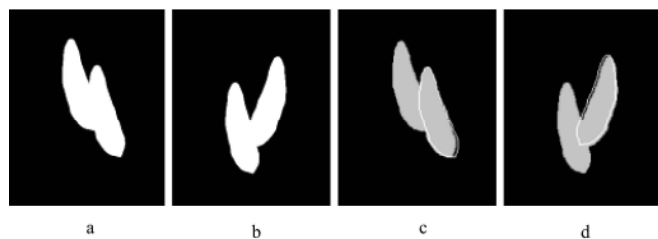


图 5 多果遮挡时的识别效果

Fig.5 Recognition effect of occlusion multiple target

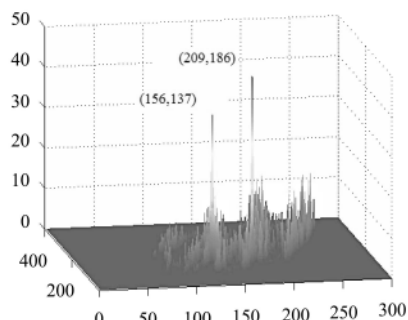


图 6 累加器的拓扑图

Fig.6 Topology chart of accumulator

图 6 显示了图 5a 在旋转角为 0°的情况下的累加器数

组累加结果的拓扑地形图，图 6 中的两个峰值分别表示两个目标的位置。图 5c 显示了较大的一个峰值对应的边缘与原图叠加。

图 7 显示了图 5b 在 19 个旋转等级上的面积比 R_k （此处显示了 10°为一个等级）。由图 7 可见，50°和 170°处的面积比明显要大于其他角度等级下的面积比。图 5d 显示了较大的一个面积比对应的边缘与原图叠加。试验 2 表明，GHT 在检测不同位姿的多果遮挡效果良好。

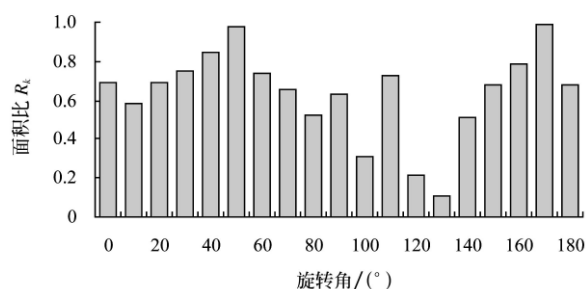


图 7 各潜在参考点下的面积比

Fig.7 Area-ratios of potential reference points

5 结论与讨论

1) 从生物学的角度，用广义柱近似抽象了茄子的外形，并通过坐标的旋转和平移，用 2D 边缘描述 3D 目标，将 GHT 由平面边缘的识别拓展到空间，为果蔬形状的表述提供了新思路。

2) 试验表明，GHT 具有很强的抗噪能力，在遮挡不严重的情况下，对目标位姿判别效果良好。

3) 对待识别图像进行适当的预处理，提取边缘点作为送检特征点，可明显降低计算量，提高计算速度。通过对边缘点的三次多项式拟合，可提高梯度索引的获取精度。

在后续的工作中，需通过添加更多的特征来更全面地描述茄子外形；可通过优化算法和降低不必要的量化等级来减小 GHT 检测时间；另外用广义柱来近似茄子形状，可能会造成描述的多义性，即形状不同的目标，其 2D 投影可能是相同的。尽管不会影响对目标位置的判断，但可能会造成对目标姿态的误判，需进一步研究。本文所述的方法还适用于在工业生产中，对可数学描述的三

维刚体的识别。

【参 考 文 献】

- [1] 陈振羽, 李德华, 周 焰, 等. 利用多种特征和 Hopfield 神经网络的有遮挡的目标识别[J]. 中国图像图形学报, 2000, 5(12): 1034—1038.
- [2] 申金媛, 李现国, 常胜江, 等. 相位特征在三维物体识别中的应用[J]. 物理学报, 2005, 54(11): 5157—5162.
- [3] Cozar J R, Guil N, Zapata E L. Detection of Arbitrary Planar Shapes with 3D Pose[J]. Image and Vision Computing, 2001, 19(14): 1057—1070.
- [4] Hough P V C. Method and means of recognizing complex patterns[P]. U S Patent 3069654, 1962—12.
- [5] Ballard D H. Generalizing the Hough transform to detect arbitrary shapes[J]. Pattern Recognition, 1981, 13(2): 111—122.
- [6] 谢志勇, 张铁中, 赵金英. 基于 Hough 变换的成熟草莓识别技术[J]. 农业机械学报, 2007, 38(3): 106—109.
- [7] 蔡健荣, 周小军, 李玉良, 等. 基于机器视觉自然场景下成熟柑橘识别[J]. 农业工程学报, 2008, 24(1): 175—178.
- [8] 饶洪辉, 姬长英. 基于标记信息的 Hough 变换检测早期的作物行中心线[J]. 农业工程学报, 2007, 23(3): 146—150.
- [9] 侯学贵, 陈 勇, 郭伟斌. 除草机器人田间机器视觉导航[J]. 农业机械学报, 2008, 39(3): 106—108, 112.
- [10] 程义民, 王以孝, 张冬青, 等. 基于广义 Hough 变换的部分遮挡物体识别[J]. 中国科学技术大学学报, 1997, 27(4): 411—418.
- [11] 李智磊, 翟宏琛, 王明伟. 一种可识别破碎图形的特殊广义 Hough 变换方法[J]. 物理学报, 2007, 56(6): 3234—3239.
- [12] Ponce J, Chelberg D, Mann W B. Invariant properties of straight homogeneous generalized cylinders and their contours[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(9): 951—966.
- [13] Chen J L, Stockman G C. 3D free-form object recognition using indexing by contour features[J]. Computer Vision and Image Understanding, 1998, 71(3): 334—355.
- [14] 蔡健荣, 许月明. 基于主动形状模型的苹果果形分级研究[J]. 农业工程学报, 2006, 22(6): 123—126.
- [15] 林开颜, 吴军辉, 徐立鸿. 基于计算机视觉技术的水果形状分级方法[J]. 农业机械学报, 2005, 36(6): 71—74.
- [16] 应义斌, 桂江生, 饶秀勤. 基于矩的水果形状分类[J]. 江苏大学学报(自然科学版), 2007, 28(1): 1—3, 67.
- [17] 蔡自兴. 机器人学 [M]. 北京: 清华大学出版社, 2000: 29—43.
- [18] 王 鑫, 荆 晶, 葛庆平. 检测不规则图形的改进广义 Hough 变换[J]. 计算机工程, 2007, 33(8): 178—179, 184.

Applications of the generalized Hough transform in recognizing occluded image

Yao Lijian, Ding Weimin^{*}, Zhao Sanqin, Yang Lingling

(College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

Abstract: This paper introduced a method of recognizing the partially occluded eggplant based on generalized Hough transform in three-dimensional space. The shape of eggplant was approximately described by using generalized-cylinders from the view of biology. Different edges of projection were obtained through coordinate transformation and 4-D parameter index tables were established. Generalized Hough Transform was adopted to count the potential pose of object and the real pose was screened by comparing the area-ratio of different rotation angles. The experiment demonstrates that the relative errors of position and pose detecting for target are lower than 1.3% and 5.5% by using Generalized Hough Transform, and this method is feasible and effective to recognize different poses and partially occluded objects.

Key words: eggplant, contour measurement, occlusion, Generalized Hough Transform