

基于双目立体视觉技术的成熟番茄识别与定位

蒋焕煜, 彭永石, 申川, 应义斌

(浙江大学生物系统工程与食品科学学院, 杭州 310029)

摘要: 研究了基于双目立体视觉技术的成熟番茄的识别与定位方法, 获取了成熟番茄的位置信息, 用于指导温室内成熟番茄的自动化采摘作业。该方法利用成熟番茄与背景之间颜色特征的差异进行图像分割来识别成熟番茄; 根据图像分割的结果, 用形心匹配方法获取番茄中心的位置; 然后根据相邻区域像素点灰度的相关性, 利用区域匹配方法计算番茄表面各点的深度信息。使用限制候选区域和两次阈值分割的方法减少计算量, 提高了计算精度。实验结果表明, 工作距离小于 550 mm 时, 番茄深度值的误差约为 ± 15 mm。利用该测量方法可以实现成熟番茄位置信息的获取, 测量精度较高。

关键词: 立体视觉; 定位; 形心匹配; 区域匹配; 番茄

中图分类号: TP242.62; TP391.41

文献标识码: A

文章编号: 1002-6819(2008)-8-0279-05

蒋焕煜, 彭永石, 申川, 等. 基于双目立体视觉技术的成熟番茄识别与定位[J]. 农业工程学报, 2008, 24(8): 279—283.
Jiang Huanyu, Peng Yongshi, Shen Chuan, et al. Recognizing and locating ripe tomatoes based on binocular stereo vision technology[J]. Transactions of the CSAE, 2008, 24(8): 279—283.(in Chinese with English abstract)

0 引言

为了节省劳动力, 目前国内外研究人员已经广泛地开展了果蔬采摘机器人的研究。果蔬采摘机器人能准确地识别与定位成熟的果蔬, 是成功实现自动化采摘作业的关键。双目立体视觉技术将二维图像推广到三维景物, 增加了目标对象的深度信息, 能精确地获取目标对象的位置信息。国外比较早就开始了这方面的研究, Kassay 等^[1]将彩色 CCD 立体摄像机安装在机器人末端执行器上, 并采用霍夫变换进行图像处理获取水果的中心位置。Kondo 等^[2]用彩色摄像机作为视觉传感器, 基于番茄的颜色特征寻找和识别成熟的果实, 采用立体视觉技术获取番茄的位置信息。Hayashi 等^[3]根据茄子的颜色特征利用彩色图像分割的方法对不同光照条件下的茄子进行识别。Takahashi 等^[4-7]根据苹果在立体图像对中的视差将给定范围的三维空间分割成若干个等距离的区间, 将两幅图像合成一幅中心图像, 然后获取苹果的深度信息。Kondo 等^[8]研制了一种草莓采摘机器人, 机器视觉系统主要由三个彩色摄像机和四个光源组成, 其中两个彩色摄像机组成立体视觉系统, 识别并定位目标对象, 第三个摄像机装在末端执行器上, 判断草莓花萼的位置。Hannan 等^[9]提出了基于图像分割、区域标记、直径提取和边缘检测的立体视觉算法, 对不同光照条件或遮挡下的柑橘进行了识别与定位。国内, 张瑞合等^[10]对红色番茄图像进行灰度变换, 提取背景区域边缘, 用拟合曲线实现图像分割, 将番茄从背景中分离出来, 并运用立体视觉成像原理获取目标的三维坐标。郑小东等^[11]提出了基于形心

特征匹配的双目立体视觉技术对番茄进行识别与定位, 该方法在计算过程中仅依靠形心特征获取目标对象的位置信息, 对图像质量的要求很高, 对噪声的影响很敏感。赵杰文等^[12]将计算机视觉系统获取的番茄 RGB 图像转换成 HIS 图像, 根据 H 分量的灰度分布用阈值法分割出成熟番茄区域, 实现对田间成熟番茄的识别。徐惠荣等^[13]根据柑橘的颜色特征利用图像色差对树上柑橘进行识别, 并分别对顺光条件和逆光条件下的柑橘进行了识别研究。张铁中等^[14]研制了一种草莓采摘机器人, 其立体视觉系统的一个摄像机获取图像中收获区域内草莓的重心位置, 另一个摄像机提取目标草莓的重心和采摘位置。尹建军等^[15]使用 Otsu 算法对番茄彩色图像的色差灰度图像进行动态阈值分割, 实现自然光照条件下多个番茄的识别。通过以上分析可看出, 近年来国内外研究者在果蔬的识别以及定位方面开展了一些研究, 然而由于果蔬采摘机器人的作业环境比较复杂, 目前依然存在视觉系统的识别和定位精度比较低, 工作效率不高等问题, 还不能完全应用于实际的生产作业。目前国内的研究尚处于初步阶段, 还需要做大量的基础研究及实际应用工作。

本文以成熟番茄为研究对象, 根据成熟番茄的颜色特征, 选用适当的阈值对成熟番茄图像进行图像分割, 将目标对象从背景中识别出来, 并将形心匹配和区域匹配相结合来计算目标番茄的位置信息。该方法可以快速地识别并定位成熟的番茄, 用于指导机器人的采摘作业。

1 系统组成

双目立体视觉系统主要有两个相同型号的摄像机(TMC7DSP, Pulnix)、两个相同型号的图像采集卡(Matrox Meteor II/Standard)和一台研华 IPC-610 工控机组成。左右两摄像机的光轴平行且高度相同, 触发器控制两个摄像机同时拍摄同一个场景的图像。采集图像的分辨率为 640 像素 \times 480 像素。

收稿日期: 2008-02-29 修订日期: 2008-05-14

项目基金: 国家 863 计划(2006AA10Z257); 国家 863 计划(2006AA10A304)

作者简介: 蒋焕煜 (1972—), 男, 浙江宁海人, 副教授, 主要从事生物图像与机器视觉研究。杭州 浙江大学生物系统工程与食品科学学院, 310029。

Email: hyjiang@zju.edu.cn

2 标定

立体视觉系统的标定是使用立体视觉系统采集图像前不可缺少的一步。立体视觉标定主要是获取摄像机的内部参数（如焦距、镜头失真系数、不确定性图像因子）、外部参数（如旋转矩阵和平移矢量）和两个摄像机之间的位置关系^[16]。本文采用 Small Vision System (SVS) 立体视觉软件进行摄像机的标定，预先制作了标准的 7×9 黑白方格图，每个方格的大小为 $54 \text{ mm} \times 54 \text{ mm}$ 。校正过程如下：先用立体摄像机拍摄黑白方格图至少 5 个不同位置的图像，如正放、左右倾斜、上下倾斜等，图像对的数量必须不少于 5 对；然后读入对应的立体图像并设置焦距（mm）、单位像素的尺寸（mm）、图像的分辨率等先验参数；最后该软件会自动检测所有黑白方格的交点，校正图像并计算两摄像机的位置关系等参数。

3 识别

实时准确地识别果蔬图像中的目标对象，是采摘机器人视觉系统的关键，而目标识别的实质是图像分割。大部分果蔬处于采摘期时，果实表面颜色与背景颜色存在较大差异，在色彩空间存在着不同的分布特性，利用图像分割可以将目标果蔬从背景中提取出来。在图像分割之前，要将彩色图像进行灰度变换，本文采用公式（1）将彩色图像变换为灰度图像。

$$I = 0.3R + 0.6G + 0.1B \quad (1)$$

式中 I ——灰度值； R 、 G 和 B ——RGB 颜色空间中像素的 3 个分量。

图 1 是用双目立体视觉系统拍摄的成熟番茄立体图像对。在图像上成熟番茄呈现红色，背景大部分是绿色的枝叶，还有少部分是呈现介于黄色和红色之间颜色的枯萎枝叶。当像素呈现红色时，其 R 值大于 G 值；呈现绿色时，其 G 值大于 R 值；呈现介于黄色和红色之间的颜色时， R 值大于 G 值，但二者的差值远小于呈现红色时 R 值与 G 值的差值。根据这个特征，对图像中的每个像素进行了色差处理，以增强背景和目标的反差，图像中每一个像素的色差表示为 $C=R-I$ ，即：

$$C = 0.7R - 0.6G - 0.1B \quad (2)$$

根据灰度图像的直方图，采用合适的阈值进行图像分割，设 T 是所取阈值，阈值分割理论如下公式（3）所示。当图像中像素的色差值小于所设定的阈值时，将像素的灰度值设置为 0；当图像中像素的色差值大于或等于所设定的阈值时，将像素的灰度值设置为 255。

$$\begin{cases} C < T, & C = 0 \\ C \geq T, & C = 255 \end{cases} \quad (3)$$

通过计算番茄图像中各像素点的色差值可知，背景的色差值主要集中在 $0 \sim 40$ 之间，而目标对象的色差值主要集中在 $140 \sim 160$ 之间，目标对象与背景之间的色差值存在较大的差异。当阈值 T 为 105 时，可以得到较好的分割效果。经过分割后的图像中仍有许多像素以噪声的形式存在，这些噪声主要是背景区域的像素被认为是目标对象或是番茄部分被误判为背景而形成的。利用形

态学方法对分割后的图像进行处理，成功地消除了噪声的影响，最终的识别结果如图 2 所示。

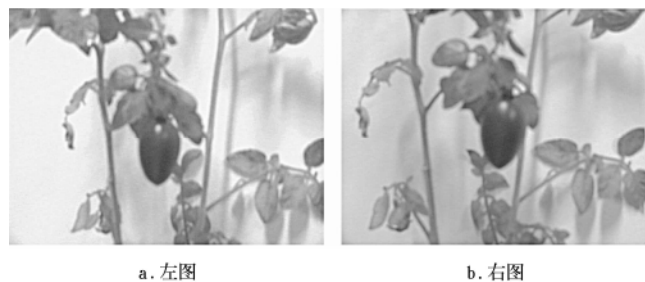


图 1 原始立体图像对

Fig.1 A pair of original stereo images

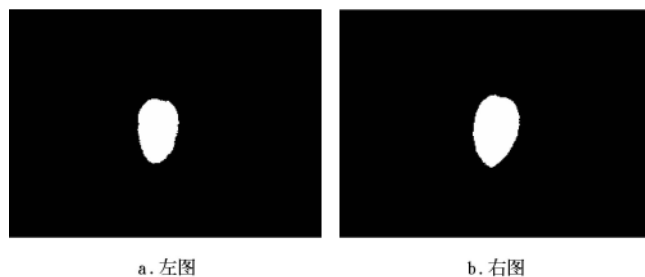


图 2 识别结果

Fig.2 Result of recognition

4 定位

双目立体视觉计算目标点深度值的模型如图 3 所示， f 为 CCD 摄像机的焦距， b 为两摄像机的中心距离， R 为目标点的深度值， X_L 和 X_R 分别是目标点在左右图中的位置。目标点在立体图像对中的视差为其在两图中水平方向上的距离，即视差 $D=|X_L-X_R|$ 。根据三角测量原理，可获得深度值 R 与视差 D 的关系：

$$R = \frac{b \cdot f}{D} \quad (4)$$

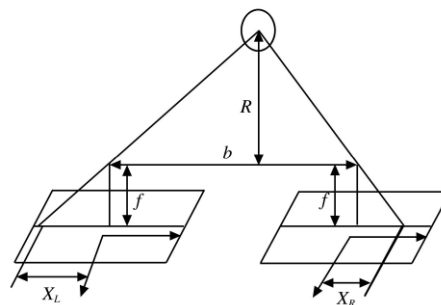


图 3 双目立体视觉模型

Fig.3 Model of binocular stereo vision

4.1 形心匹配

标准球体在二维图像中的形心是该球体质心在摄像机成像平面上的投影，所以球体在左右图像中的形心是一对天然的匹配点，进行形心匹配可计算其深度值。番茄的形状接近于球体，可以采用形心匹配方法进行目标对象的初步定位。通过对左右两图像进行图像分割，获

取目标对象在立体图像对中的形心坐标，由公式(4)计算得到目标对象的初步深度值。

4.2 区域匹配

通过形心匹配方法可以获得目标番茄的初步深度信息，然而仅仅依靠目标对象的形心获得深度信息对噪声非常敏感。在形心匹配的基础上结合区域匹配算法进一步确定番茄表面点的位置，能更加准确地获得目标对象的位置信息，提高算法的鲁棒性。在图像处理中提取感兴趣区域(ROI)可以减小计算量，屏蔽不必要的信息，提高处理速度和精度。本研究中的ROI就是成熟番茄所在的区域，该区域由成熟番茄的最小外接矩形确定，以左右两图中最小外接矩形的长中的最大者作为长，宽中最大者作为宽。

4.2.1 匹配基元

本文以区域作为匹配基元，把一幅图像中某一点的灰度邻域作为匹配模版，在另一幅图像中搜索具有相同(或相似)灰度值分布的对应点邻域。由于原始图像是RGB图像，需要对三个分量进行线性变换计算其灰度值。Ohta等^[17]通过对大楼、海滨等不同类型的彩色图像进行方差计算，归纳出3个正交的彩色特征：

$$\begin{cases} I_1 = (R+B+G)/3 \\ I_2 = R-G \text{ 或 } I_2 = G-R \\ I_3 = (2 \times R - G - B)/4 \end{cases} \quad (5)$$

经实验发现， I_3 可以获得适合番茄表面的匹配。 I_2 所包含的信息量太少， I_1 包含了大量的信息，然而在番茄的ROI中 I_3 比 I_1 具有更大的区分能力。所以采用 I_3 的色彩空间对原始图像进行线性变换。

4.2.2 匹配窗口

区域匹配的实质是一个在立体图像对中不断寻找最大相似程度的过程，因此计算量非常巨大，利用ROI作为搜索空间在很大程度上减少了计算量。为了更大程度上减少计算量，提高匹配的精度，采取如下办法来约束匹配区域：

1) 限定立体图像对的左图中作为匹配模版的点仅是属于番茄表面的点，同时对左图上番茄表面点采用隔点采样的方式以减少计算时间。

2) 限定匹配时立体图像对中右图的搜索范围，把右图的搜索范围限定在一个较小的区域内。

4.2.3 匹配算法

区域匹配中常用的相似测度算法有像素差平方和(SSD)、像素差绝对值(SAD)、图像的互相关性(NCC)等。本研究选取SSD作为相似测度，计算匹配窗口间的相关性。SSD算法的数学表达式如下：

$$SSD(x, y, d) = \sum_{i=1}^n \sum_{j=1}^m [L(x+i, y+j) - R(x+i+d, y+j)]^2 \quad (6)$$

式中 $L(x, y)$ 和 $R(x, y)$ ——左右两幅图像的像素； d ——视差。

直接用SSD计算有较大的冗余度，而采用归一化相

关量进行相关计算，可以减少计算量，不会受到左右图像亮度差异的影响，而且归一化相关量 $R(x, y)$ 是点 (x, y) 处视差值可信性的一个自然衡量依据。

归一化方法处理过程如下，将式(6)展开则有：

$$SSD_{(x,y,d)} = \sum_{i=1}^n \sum_{j=1}^m [L(x+i, y+j)]^2 - 2 \sum_{i=1}^n \sum_{j=1}^m [L(x+i, y+j)R(x+i+d, y+j)] + \sum_{i=1}^n \sum_{j=1}^m [R(x+i+d, y+j)]^2 \quad (7)$$

式中的第一项是立体图像对中左图匹配模版的总能量，第二项是立体图像对中的右图窗口与左图匹配模版的自相关量，第三项是在左图匹配模版覆盖下右图中窗口的能量。当第二项取最大值时，右图窗口和左图模版相匹配，于是可以得到归一化的相关量：

$$R(x, y, d) = \frac{\sum_{i=1}^n \sum_{j=1}^m [L(x+i, y+j) \times R(x+i+d, y+j)]}{\sqrt{\sum_{i=1}^n \sum_{j=1}^m [R(x+i, y+j)]^2} \sqrt{\sum_{i=1}^n \sum_{j=1}^m [R(x+i+d, y+j)]^2}} \quad (8)$$

根据施瓦茨不等式可知，式(8)中 $0 < R_{(x,y,d)} < 1$ 。立体匹配过程中，在立体图像对的右图中寻找与左图匹配模板窗口之间的相关量 R 达到最大值的窗口，且当该窗口的相关量 R 大于预先设置的阈值时，则将该窗口作为匹配成功的窗口。

番茄容易受到非完全漫反射的影响，在相关性上表现为相关量 R 出现很大的跳变，造成错误匹配。通过限制相关量 R 的波动范围，对相关量 R 采用两次阈值分割方法，可以排除非完全漫反射的干扰。

第一次阈值分割在立体图像对右图搜索区域的同一行像素点中进行，阈值分割的公式如下：

$$\frac{R_{\max R} - R_{\sec R}}{R_{\max R} - R_{\min R}} > T_1 \quad (9)$$

式中 $R_{\max R}$ ——该行像素点中相关量的最大值； $R_{\sec R}$ ——该行像素点中相关量的次大值； $R_{\min R}$ ——该行像素点中相关量的最小值； $R_{\max R} - R_{\min R}$ ——波动的幅度； T_1 ——阈值。

第二次阈值分割则针对整个匹配区域，阈值分割的公式如下：

$$R_{\max} > T_2 \quad (10)$$

式中 R_{\max} ——所有剩余 $R_{\max R}$ 中的最大值； T_2 ——阈值。这样就减少了错误匹配，提高了匹配精度。

经过番茄表面区域相关量 R 的计算可得，当阈值 T_1 为 0.25，阈值 T_2 为 0.85 时，两次阈值分割的效果较好，保证了相关量的最大值与次大值之间有足够的差距，提高了匹配的精度，同时也限制了相关量 R 的波动幅度，屏蔽了非完全漫反射的干扰。

5 结果与分析

本文中图像的阈值分割、噪声去除、立体匹配以及目标对象深度值的计算是在 Visual C++6.0 环境下进行程序编写实现的。

5.1 形心匹配

测定了不同位置的两个番茄深度信息的实际值, 并利用形心匹配方法计算目标对象的预测值。利用形心匹配方法完成一次目标对象的定位测量所需时间约为 0.07 s。番茄 A 和番茄 B 的测量值与预测值之间的差异如图 4 所示, 其中纵坐标为番茄深度信息的测量值与预测值的偏差, 横坐标为形心的测量值。由图可见, 当工作距离小于 700 mm 时, 除了个别奇异点外, 番茄深度信息的误差基本上在 ± 15 mm 以内。

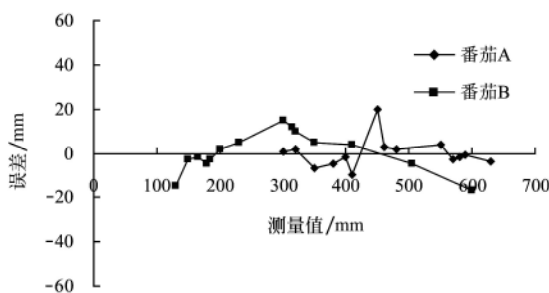


图 4 番茄 A 和 B 的误差与测量值

Fig.4 Errors and measured depth information values of tomato A and B

5.2 区域匹配

为了定量地说明匹配的准确度, 采用手工测量方法对区域匹配算法进行评估, 将番茄表面分为 3×3 共 9 个区域如图 5 所示。测量番茄表面的每个区域到摄像机平面的距离作为该区域的实际测量值。将每幅图像中成熟番茄区域的像素点按 x, y 坐标分成相应的 9 个区域, 取每个区域中各点的平均深度作为该区域深度的预测值。利用区域匹配方法完成一次目标对象的定位测量所需时间约为 0.35 s。番茄各区域的实际测量值与预测值之间的

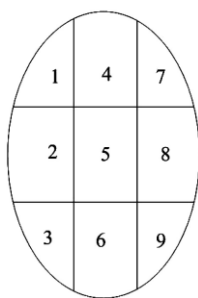


图 5 番茄表面各区域分配

Fig.5 Division of tomato surface

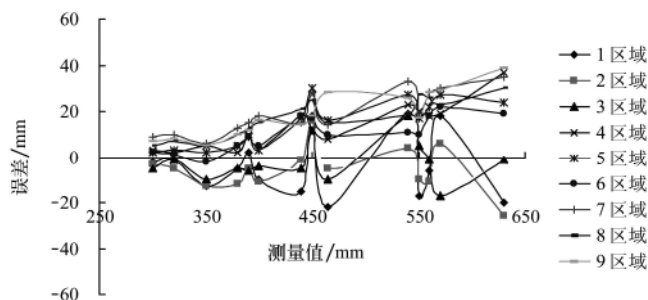


图 6 9 个区域的误差与测量值

Fig.6 Errors and measured values of the nine divisions of tomato surface

差异如图 6 所示。从图中可以看出, 在 450 mm 以内, 9 个区域的误差均在 ± 15 mm 之间, 其中在 350 mm 以内时, 各区域的误差接近于零, 这一误差范围满足果蔬采摘定位精度的要求。当距离大于 450 mm 小于 550 mm 时, 9 个区域的最大误差范围放大到 ± 20 mm, 而 2、3、4、6、8 五个边缘区域的误差在 ± 15 mm 之间。当距离大于 550 mm 时, 各区域的误差较大。

6 结 论

1) 在 RGB 颜色空间下成熟番茄与背景之间的颜色特征差异较大, 通过对图像中的像素点进行色差处理, 设定合适的阈值可以得到较好地目标对象分割出来。

2) 利用形心匹配与区域匹配相结合的方法, 计算目标番茄的深度值, 能够较准确地获得番茄的空间位置信息。选用了合适的匹配模板窗口和匹配区域, 使用由粗到精的匹配策略, 大大减小了立体匹配的计算量, 同时对相关量 R 采用了两次阈值分割方法, 排除了番茄表面受非完全漫反射的干扰, 提高了匹配精度。实验结果表明, 利用该测量方法可以有效地获取目标对象的位置信息, 而且测量精度较高。

[参 考 文 献]

- [1] Kassay L, Slaughter D C, Molnar S. Hungarian robotic apple harvester[A]. St. Joseph, Mich.: ASAE, 1992, ASAE Paper No.922047.
- [2] Kondo N, Nishitsuji Y, Ling P P. Visual feedback guided robotic cherry tomato harvesting[J]. Transactions of the ASAE, 1996, 39(6): 2331-2338.
- [3] Haysshi S, Ganno K, Ishii Y, et al. Robotic harvesting system for eggplants[J]. Japan Agricultural Research Quarterly, 2002, 36(3): 163-168.
- [4] Takahashi T, Zhang S, Sun M, et al. New method of image processing for distance measurement by a passive stereo vision[A]. St. Joseph, Mich: ASAE, 1998, ASAE Paper No. 983031.
- [5] Takahashi T, Zhang S, Fukuchi H, et al. Binocular stereo vision system for measuring distance of apples in orchard (Part 1) - Method due to composition of left and right images[J]. Journal of the Japanese Society of Agricultural Machinery, 2000, 62(1): 89-99.
- [6] Takahashi T, Zhang S, Fukuchi H, et al. Binocular stereo vision system for measuring distance of apples in orchard (Part 2) - Analysis of and solution to the correspondence problem[J]. Journal of the Japanese Society of Agricultural Machinery, 2000, 62(3): 94-102.
- [7] Takahashi T, Zhang Shuhuai, Fukuchi H. Measurement of 3-D locations of fruit by binocular stereo vision for apple harvesting in an orchard[A]. St. Joseph, Mich.: ASAE, 2002, ASAE Paper No.021102.
- [8] Kondo N, Ninomiya K, Hayashi S, et al. A new challenge of robot for harvesting strawberry grown on table top culture[A]. St. Joseph, Mich.: ASAE, 2005, ASAE Paper No.053138.
- [9] Hannan M W, Burks T F, Bulanon D M. A real-time machine vision algorithm for robotic citrus harvesting[A]. St. Joseph, Mich.: ASAE, 2007, ASAE Paper No. 073125.

- [10] 张瑞合, 姬长英, 沈明霞, 等. 计算机视觉技术在番茄收获中的应用[J]. 农业机械学报, 2001, 32(5): 50—52.
- [11] 郑小东, 赵杰文, 刘木华. 基于双目立体视觉的番茄识别与定位技术[J]. 计算机工程, 2004, 30(22): 171—173.
- [12] 赵杰文, 刘木华, 杨国彬. 基于 HIS 颜色特征的田间成熟番茄识别技术[J]. 农业机械学报, 2004, 35(9): 122—124, 135.
- [13] 徐惠荣, 叶尊忠, 应义斌. 基于彩色信息的树上柑橘识别研究[J]. 农业工程学报, 2005, 21(5): 98—101.
- [14] 张铁中, 陈利兵, 宋 健. 草莓采摘机器人的研究: 基于图像的草莓重心位置和采摘点的确定[J]. 中国农业大学学报, 2005, 10(1): 48—51.
- [15] 尹建军, 毛罕平, 王新忠, 等. 不同生长状态下多目标番茄图像的自动分割方法[J]. 农业工程学报, 2006, 22(10): 149—153.
- [16] 隋 婧, 金伟其. 双目立体视觉技术的实现及其进展[J]. 电子技术应用, 2004, 10: 4—6.
- [17] Ohta Y, Kanade T, Sakai T. Color information for region segment[J]. Computer Graphics and Image Processing, 1980, 13(3): 222—242.

Recognizing and locating ripe tomatoes based on binocular stereovision technology

Jiang Huanyu, Peng Yongshi, Shen Chuan, Ying Yibin

(College of Biosystems Engineering and Food science, Zhejiang University, Hangzhou 310029, China)

Abstract: A method for recognizing and locating ripe tomatoes based on binocular stereovision technology was developed for tomato harvesting in greenhouse. The color features were used to recognize ripe tomatoes by segmenting target objects from background. According to the results, the position of the fruit centroid was acquired with centroid-based matching. Then in terms of grey correlation of neighborhood regions, depths of the tomato surface points were calculated with area-based matching. And the method of limited candidate region and twice thresholds was used to reduce computation load and improve precision. The error ranged within ± 15 mm when the distance was less than 550 mm. It's feasible to obtain position information of ripe tomatoes with this measurement, and the measurement precision is high.

Key words: stereovision; location; centroid-based matching; area-based matching; tomato