灵巧手主从操作中的人手运动跟踪与识别

李继婷 任大伟 张玉茹

(北京航空航天大学 机械工程及自动化学院，北京 100083)

摘 要：在灵巧手主从抓持规划系统中，一项重要的研究内容是获取人手的运动信息.利用计算机双目立体视觉系统作为灵巧手主从操作的人机接口，研究人手指尖运动实时跟踪和识别的方法，提出了一种人手指尖位置的动态识别算法及其实验系统.利用本实验系统，完成左右两像面 4 个目标点所用的最长识别时间为 0.4 ms，且人手在正常运动速度时，识别率达 100%，表明本算法同时满足实时性和识别率的要求.

关 键 词：灵巧手；计算机视觉；目标识别；识别率

中图分类号：TP 242.6^{+}2

文献标识码：A **文 章 编 号**：1001-5965(2007)05-0618-04

Motion tracking and recognition of human hand in master-slave manipulation for dexterous hands

Li Jiting Ren Dawei Zhang Yuru

(School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Acquiring the motion information of human hands is a very important research problem in the master-slave grasping system for dexterous hands. Used to be the human-computer interface for the master-slave manipulation system for dexterous hands, a binocular stereo computer vision system was investigated to track and recognize the dynamic moving fingertips of human hands. A corresponding identifying algorithm and the experimental system were proposed. Experiment results show that the longest time is 0.4 ms to simultaneously recognize the four targets in two images of left and right ones, and the recognition rate reaches 100 percent when the human hand moves in usually used speed, which illustrate that the presented method meets both the real-time and accurate requirements.

Key words: dexterous hand; computer vision; target recognition; recognition rate

在灵巧手主从操作系统中，人手作为主操作手来控制作为从操作手的灵巧手，这需要借助有效的人机接口装置准确、实时地测量人手的运动.其中，数据手套应用较为广泛，但其精度较低，因此需要探索更为有效的人机接口装置.比较而言，计算机视觉测量精度高，但存在信息处理速度慢的缺点.另外，当其用来识别人手指运动时，还需要具有非常高的正确识别率，这需要高效的视觉识别算法来保证.目前计算机视觉在灵巧手抓持规划系统中多用来记录操作环境和被操作物体的信息，或用于数据手套标定的辅助装置，较少用来直接作为人手运动的测量装置[1-3].

计算机视觉硬件设备的性能不断提高，而价格却在不断下降，这使得计算机视觉具有优越的性价比；同时硬件技术的飞速发展及计算机视觉应用领域的不断扩大，也极大地促进了其软件技术的研究和发展：这两方面保证了计算机视觉具有无可比拟的发展潜力.

本文探索用计算机视觉作为人机接口设备，直接测量人手指尖点的位置，获得人手的运动信息.在原有研究工作的基础上[4]，针对人手手指的运动无特定运动规律、但经过路径重复性等特点，

收稿日期：2006-06-19

基金项目：国家教委博士点基金资助项目(2000000605)

作者简介：李继婷(1967—)，女，辽宁抚顺人，副教授，lijiting@buaa.edu.cn.

提出了一种利用双目立体视觉来实时获取人手指尖三维坐标的方法,在提高测量精度的同时,保证系统识别准确、快速;并通过实验验证了该方法的有效性.

1　人手运动识别系统

如图 1 所示,人手运动识别系统的硬件包括 2 个摄像机、1 块图像采集卡、1 台计算机和人手. 人手掌固定不动,手指运动时,由摄像机实时采集其运动图像,通过图像采集卡转换为计算机能够处理的数字信号,并传递给计算机. 系统软件即人手运动识别算法,用来识别硬件系统记录的人手指尖的位置信息,通过左右像面中的人手指尖所成像的像面坐标,计算出人手各指尖的空间坐标,再通过映射算法把人手的空间坐标转化为灵巧手的运动控制信号. 为尽量简化算法,保证识别的实时性,人手佩戴白色手套,并在人手拇指、食指、中指和无名指指尖分别粘贴相同大小的黑色小球作为识别目标点.

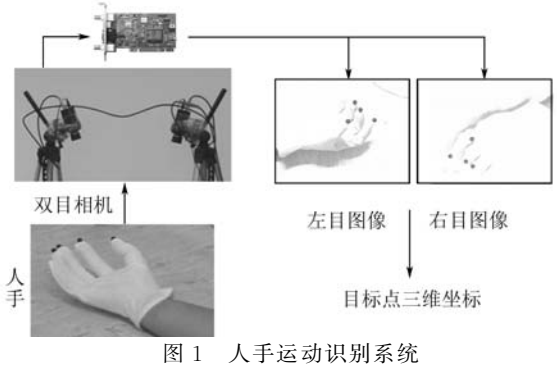


图 1　人手运动识别系统

2　人手运动识别算法

2.1　目标点匹配初始化

整个人手运动识别算法的流程如图 2 所示.

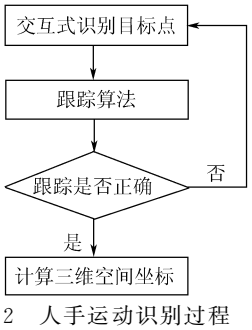


图 2　人手运动识别过程

首先在识别开始时通过手动方式依次匹配左右像面上的 4 个目标点,图 3 为匹配完毕的左右 2 幅图像,相同颜色的方框包围的目标点为同一个手指指尖点. 以下称该方框为搜索框,其边长为 L_S,比目标点的直径稍大,可根据具体的实验条件进行调整. 在本实验系统中,目标点像面直径小于 30 个像素,L_S 取 36 个像素. 初始匹配完成之后,程序进入自动跟踪模块.

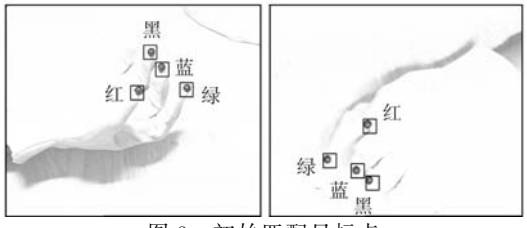


图 3　初始匹配目标点

2.2　自动跟踪算法模块

2 幅图像素个数为 80 多万,需要处理大量的数据,又因实时性的要求,给程序后续的处理带来很大的压力.

在灵巧手主从操作系统中,人需要根据灵巧手的运动来实时调整人手的运动,以保证灵巧手达到预期位置,因而人手通常保持运动平稳. 基于此,特征点在像面上的运动可以近似看作匀速运动. 所以,可以用前一时刻目标点的速度来预测下一时刻目标点可能出现的位置,在预测位置附近搜索目标点. 这样可极大地减少数据处理量,从而减少计算时间,提高跟踪和识别速度. 该自动跟踪算法分目标点预测、目标点形心计算和目标点速度修正 3 个步骤.

2.2.1　目标点预测

如图 4 所示,设在 k 时刻,某一目标点在左(或右)像面内的位置和速度分别为 (p_k^x, p_k^y) 和 (v_k^x, v_k^y),在经过图像采集周期 T 后的 $k+1$ 时刻该目标点的位置和速度由式(1)预测,设预测位置和速度分别为 $(\tilde{p}_{k+1}^x, \tilde{p}_{k+1}^y)$ 和 $(\tilde{v}_{k+1}^x, \tilde{v}_{k+1}^y)$,则

$$\begin{bmatrix} \tilde{p}_{k+1}^x \\ \tilde{p}_{k+1}^y \\ \tilde{v}_{k+1}^x \\ \tilde{v}_{k+1}^y \end{bmatrix} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_k^x \\ p_k^y \\ v_k^x \\ v_k^y \end{bmatrix} \tag{1}$$

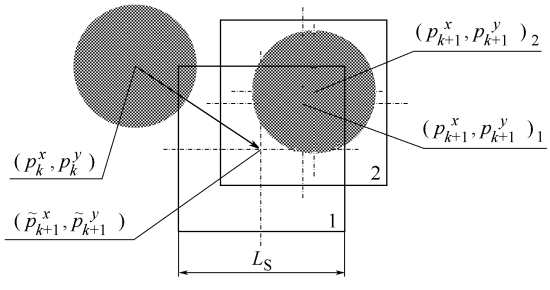


图 4　目标点预测和形心计算

2.2.2 目标点形心计算

在 $k+1$ 时刻，若目标点在以 $(\tilde{p}^x_{k+1},\tilde{p}^y_{k+1})$ 为中心、边长为 L_S 的搜索框内，则该目标点的形心为[5]

$$\left.\begin{aligned} p^x_{k+1}&=\frac{\sum_{n=\tilde{p}^y_{k+1}-\frac{L_S}{2}}^{\tilde{p}^y_{k+1}+\frac{L_S}{2}}\sum_{m=\tilde{p}^x_{k+1}-\frac{L_S}{2}}^{\tilde{p}^x_{k+1}+\frac{L_S}{2}}mf(m,n)}{\sum_{n=\tilde{p}^y_{k+1}-\frac{L_S}{2}}^{\tilde{p}^y_{k+1}+\frac{L_S}{2}}\sum_{m=\tilde{p}^x_{k+1}-\frac{L_S}{2}}^{\tilde{p}^x_{k+1}+\frac{L_S}{2}}f(m,n)}\\ p^y_{k+1}&=\frac{\sum_{n=\tilde{p}^y_{k+1}-\frac{L_S}{2}}^{\tilde{p}^y_{k+1}+\frac{L_S}{2}}\sum_{m=\tilde{p}^x_{k+1}-\frac{L_S}{2}}^{\tilde{p}^x_{k+1}+\frac{L_S}{2}}nf(m,n)}{\sum_{n=\tilde{p}^y_{k+1}-\frac{L_S}{2}}^{\tilde{p}^y_{k+1}+\frac{L_S}{2}}\sum_{m=\tilde{p}^x_{k+1}-\frac{L_S}{2}}^{\tilde{p}^x_{k+1}+\frac{L_S}{2}}f(m,n)}\end{aligned}\right\}\quad(2)$$

式中 $f(m,n)=\begin{cases}1 & I(m,n)\leqslant Z\\ 0 & I(m,n)>Z\end{cases}$

$I(m,n)$ 为第 m 行 n 列的像素灰度值；Z 为灰度阈值，其选取根据经验来确定，在本实验中取 160. 若目标点不是完全位于搜索框内，如图 4 所示搜索框 1，则先由式(2)计算以预测点 $(\tilde{p}^x_{k+1},\tilde{p}^y_{k+1})$ 为中心的搜索框内的图像的形心坐标，设为 $(p^x_{k+1},p^y_{k+1})_1$，下标表示搜索框计算目标点形心的次数，再计算以 $(p^x_{k+1},p^y_{k+1})_1$ 为中心的搜索框内的图像的形心 $(p^x_{k+1},p^y_{k+1})_2$. 显然，在图 4 中有 $(p^x_{k+1},p^y_{k+1})=(p^x_{k+1},p^y_{k+1})_2$. 在得到预测点后，搜索框计算次数的多少与目标点和搜索框的大小有关，本实验中，最多搜索 3 次就可以找到目标点的形心.

当手指运动速度在相隔的 2 个时刻变化过大时，那么以预测点为中心的搜索框将包围不到目标点，即目标点位于搜索框外，如图 5 所示.

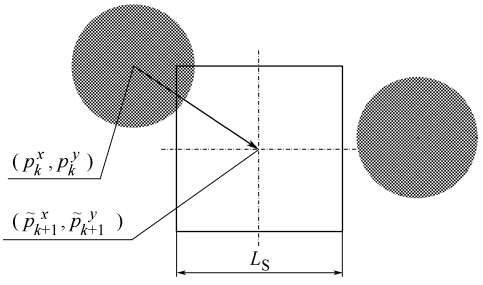


图 5 目标点位于搜索框外

由于在整个识别过程中，手掌固定不动，人手反复张合运动过程中指尖点的运动范围有限，指尖点的运动轨迹很接近，重复性较大. 因此在跟踪和识别的同时，把目标点的距离相隔 L_S 的历史轨迹点记录下来，如图 6 所示. 一旦出现图 5 所示的情况，就到所记录的历史轨迹点中去查找，直到查找到该目标点为止.



图 6 指尖点路径

2.2.3 目标点速度修正

当指尖点正确识别后，即可对 $k+1$ 时刻目标点的速度进行修正，如式(3)所示.

$$\left.\begin{aligned} v^x_{k+1}&=\frac{p^x_{k+1}-p^x_k}{T}\\ v^y_{k+1}&=\frac{p^y_{k+1}-p^y_k}{T}\end{aligned}\right\}\quad(3)$$

至此完成了一次目标点的识别. 在识别开始时，$v^x_0=v^y_0=0$，p^x_0，p^y_0 通过手动匹配 2 幅图像目标点得到.

同理，可以得到右(或左)图像的像面位置坐标，于是就可以通过目标点的左像面位置坐标和右像面位置坐标实时地计算出目标点的空间位置坐标 (p^x_k,p^y_k,p^z_k).

3 实验及结果分析

实验场景如图 7 所示，人手戴一只粘贴标识点的手套，摄像机是黑白面阵 CCD；采集卡分辨率是 768 像素 × 576 像素，单路采集频率是 25 帧/s；内存 512 MB，赛扬 2.5 GHz. 实验时，人手掌固定不动.



图 7 实验场景图

3.1 识别速度

由于系统要求实时地识别目标点，所以目标点跟踪与识别算法的速度是评价系统性能重要的指标. 使用该实验系统进行了大量实验，结果表明完成识别左右两像面目标点所用的最长识别时间为 0.4 ms，图像采集周期为 40 ms，显然，系统的

实时性可以保证.

3.2　目标识别率

在一定时间内,视觉正确识别目标点的次数与总实验次数的比值称为识别率,用以衡量算法效果的优劣.选择了人手常用的抓持动作作为实验动作,此动作包括手指的张开和抓合 2 个过程.

实验结果如表 1 所示,人手每分钟张合 45 次,运动速度正常;每分钟张合 61 次,即人手在 1 s的时间完成一次张合运动,此时人手运动速度较快;在这 2 种运动速度下,识别率均为 100%;当人手每分钟张合次数大于 72 次时,人手速度已非常快,人手操作时几乎不会使用这样的速度.所以本算法可以保证在人手的正常速度下对人手运动进行跟踪识别.

表 1　视觉识别率

实验次数/min^{-1}	成功次数/min^{-1}	识别率/%
45	45	100
61	61	100
72	69	96
81	75	92

4　结 束 语

本文提出了一种利用计算机双目立体视觉实时跟踪、识别人手指尖点的方法.实验结果表明,在人手手指较快的运动速度下,此方法可以保证识别率 100%;并且算法所用时间也远远小于图像采集周期的 40 ms,充分满足实时性要求.该算法不但可以用来识别人手运动,也可以应用于对其它匀速运动目标进行实时跟踪.

参考文献(References)

[1] Turner M L, Findley R P, Griffin W B, et al. Development and testing of a telemanipulation system with arm and hand motion[C]//Dynamic Systems and Controls. Okamura: ASME,2000:1－8

[2] Kawasaki, Nakayama H, Mouri K, et al. Virtual teaching based on hand manipulability for multi-fingered robots[C]// Proc IEEE Int Conf Rob Autom. Piscataway, NJ: IEEE, 2001,2:1388－1393

[3] Kang S B, Ikeuchi K. Toward automatic robot instruction from perception-mapping human grasps to manipulator grasps[J]. Robotics and Automation,1997,13(1):81－95

[4] Li Jiting, Zhang Yuru, Su Wenkui, et al. Grasp planning and implementation for dexterous hands by human demonstration[C]//Huang Tian. The 11th World Congress in Mechanism and Machine Science. Tianjin: China Machine Press,2004,4:1－4

[5] 贾云得. 机器视觉[M]. 北京:科学出版社,2000
Jia Yunde. Computer vision[M]. Beijing: Science Press, 2000(in Chinese)

(责任编辑:文丽芳)

(上接第 567 页)

参考文献(References)

[1] Lamport L. Password authentication with insecure communication [J]. Communication of ACM, 1981, 24:770－772

[2] Hwang M S, Li L H. A new remote user authentication scheme using smart cards[J]. IEEE Transactions on Consumer Electronics, 2000, 46(1):28－30

[3] Ku W C, Chen S M. Weakness and improvements of an efficient password based user authentication scheme using smart cards [J]. IEEE Trans Consumer Electronic, 2004, 50(1): 204－207

[4] Yoon E J, Ryu E K, Yoo K Y. Further improvement of an efficient password based remote user authentication scheme using smart cards [J]. IEEE Trans Consumer Electronic, 2004, 50(2):612－614

[5] Manik Lal Das, Ashutosh Saxena, Ved P Gulati. A dynamic ID-based remote user authentication scheme [J]. IEEE Transactions on Consumer Electronics, 2004, 50(2):629－631

[6] Liao I-En, Lee Cheng-Chi, Hwang Min-Shiang. Security enhancement for a dynamic ID-based remote user authentication scheme[C]// Next Generation Web Services Pracitces. Seoul, KOREA: IEEE, 2005

(责任编辑:娄　嘉)