

第一届全国计算机视觉学术会议

论文集

1st National Symposium on
Computer Vision Proceedings



(下)

中国计算机学会人工智能与模式识别专业委员会
上海市计算机学会人工智能专业委员会

41-53

1989

目录

<一> 立体视觉

1.1	基于知识的双摄像头三维视觉系统	竺卫东	李允明	1
1.2	用单个透视投影识别空间多边形	郭雷		7
1.3	基于模型的三维物体识别方法的研究	陆乃将	李衍达	13
1.4	用二幅图象和边长信息唯一地恢复平面块三维坐标	宣国荣	姚键超	16
1.5	立体视觉多通道结构匹配法	顾创	吴立德	22
1.6	关于三维形态估计的一个问题	巍湘曙		27
1.7	GRIMSON-LOZANO 三维物体识别算法的扩展试验	袁太刚	潘士先	31
1.8	成像几何的若干问题	夏宇闻		
1.9	基于特征的立体视觉并行匹配算法	吴渝彦	徐建华	35
1.10	三相机立体视觉	杨敬安	张奠成	39
		胡荣祥	董铁军	45
1.11	用计算机立体视觉原理重建三维数字表面	毛素银		
1.12	一个重建二维图象的 AI 方法	李荣兴		47
1.13	摄像机模型矩阵的分解方法	费耀平	谷士文	56
1.14	体视法提取景物深度信息	宣雷	徐建华	60
1.15	用于变焦视觉的弹性匹配方法	厉万庆	张森	64
1.16	从灰度及其变化恢复表面形状的研究	郭杆		67
1.17	序列图像中提取刚体运动参数的最小二乘算法	陈臻祺		69
		宣国荣	姚键超	73

<二> 运动序列分析

2.1	基于序列图象估计物体三维运动参数一种新算法	何明浩	刘武	76
2.2	用于目标跟踪的一种新的符号匹配方法	熊衍仁		
2.3	一种有效的动态图象序列平滑处理方法	曾迎生	张桂林	79
2.4	用运动一致性准则约束点匹配	李江		81
		魏国庆	何振亚	86
2.5	用 N 对观察点来确定运动物体是否为平面的有效方法	马颂德		
2.6	基于层次基元结构的运动图象匹配算法	邢小良	庄新华	90
		王晓军	傅洁	96
2.7	运动图象分析的连续处理方法	吴立德		
2.8	多目标实时跟踪测量系统	傅洁	吴立德	101
		吕维雪	胡大可	106
2.9	用多摄像机求解三维运动刚体的位移和运动朝向	顾列平		
2.10	平面型目标在三维空间运动的估计	王志凌	罗秀芳	110
2.11	不用点对应的三维运动参数估计	陈刚	方廷键	114
2.12	从视觉流对应恢复物体的三维空间运动	林敏杰	鞠德航	117
		陆庆	周新	119

<三> 基于知识的视觉系统

3.1	飞机目标的特征选择和提取	戴震	毕都彦	126
		李响		
3.2	遥感图象分类专家系统的黑板控制模型	李爱生	柳健	128
		黄铁侠	郭建文	

3.3	ALV 自学习道路跟踪	李树祥	黄晓德	133
3.4	两维重迭物体的识别	张鸿宾		135
3.5	关系数据结构的一种不精确匹配	冯建生		140
3.6	对于 KBVS 的现状及其发展的探讨	李允明	张晓军	142
3.7	用知识模式网理解室外自然景物的智能视觉系统	王志凌	罗秀芳	146
		路祖铭	程琪青	
3.8	匹配误差的形态学度量	龚 炜	石青云	151
3.9	计算机视觉系统中的分割区域知识表达形式及算法	吴春明	朱森良	155
3.10	三维物体理解中的知识结构初析	过耀中		159
3.11	地图知识库系统	孙建伶	吴春明	161
		朱森良		

<四> 形状表示与分析

4.1	形状的沃尔什分析	陆乃将	李衍达	166
4.2	一种八元树的快速旋转算法	沈新权	谷士文	168
4.3	利用线性四元树中轴变换重建线性四元树	张田文	李仲荣	172
4.4	标准线性四元树在模式识别中的应用	张田文	李仲荣	178
4.5	一种基于峰值的景物形状分析方法	雷向康	周曼丽	184
4.6	用链码模板对区域进行描述与匹配	李玉山	李凯峰	186
4.7	多面体的 HOUGH 空间特征识别方法的研究	王华迪	施鹏飞	189

<五> 图像理解系统

5.1	荧光粉涂层混色自动检测系统	詹 宏	丁明跃	191
5.2	一种用于工件识别的机器人视觉系统	彭嘉雄		
5.3	静态图象采集板 SIPS-II 的研制	白 焱	彭嘉雄	193
5.4	PIPS 体育录像解析系统	毕都彦	戴 震	197
		吴敏金	蔡文康	199
5.5	低价图象输入接口系统	龚铭新		
5.6	一种简易实用的图象处理系统	杨兆选	国澄明	201
		彭 强	唐 强	203
5.7	一种简便的图象处理系统	姚卫东		
		渠 速	聂大同	206
5.8	机械扫描式图像 / 图形生成及处理系统	高鹏飞	毛 刚	
5.9	医用超声成像系统的新型数字化扫描变换器	乐光启	于 雷	208
5.10	CID-100 生物图象数据微机处理系统图象接口电路分析	乐光启	胡修泰	210
5.11	PC/AT 微机图象处理系统	张耀远	吴立丹	212
		夏良正	顾宗恂	215
5.12	高效率的帧存储器	朱佩英	丁则信	
5.13	带钢头部形状检测及优化剪切系统	蔡燕波	莫玉龙	218
5.14	图象板技术在信函分拣机中的应用	孙一康	王力飞	221
5.15	机器人视觉实验系统	许海华	夏海良	223
5.16	基于规则的图像理解专家系统	李敦玲	梁震源	225
		陶文祥	曾湘燕	229

<六> 图形生成与仿真

6.1	线框图模型和相应简单光照模型三维显示消隐算法	王 柏	宣国荣	233
6.2	一种新的角度表示法——泛角及相应的绕坐标旋转算法	刘鸿键	袁 琛	237
6.3	图象图形合成法实现视觉模拟初探	彭建华	黄振华	239
6.4	机器人图形仿真的研究	李春梅	王棣棠	242

<七> 机器人视觉

7.1	基于特征的作业面空间姿态和旋转变换参数计算方法	康南生	吴翼娜	246
7.2	论机器人视觉	胡庆茂	陈锦江	251
7.3	基于摄像机和超声传感器的机器人视觉闭环控制系统	詹跃东	曹柏林	259
7.4	一种用双相机检测障碍物的新方法	陈辉堂		
		朱 枫	吴继显	263
7.5	行走机器人探测路面 3-D 形状的视觉技术	徐心平		
7.6	利用改进的 B-P 神经模型实现多体汉字的学习与识别	洪金木		267
7.7	一种基于彩色视觉自治车视觉导航的道路分割方法	黄晓非		271
7.8	机器人视觉系统的研究	朱铁山	吴继显	274
		李金伴		276

<八> 编码与压缩

8.1	五邻域单边界差分链码的图形压缩	郭新平	楚天鹏	278
8.2	一种新的多灰度图象的压缩编码方案	袁占亭	郭椿标	282
8.3	用变换域中的自适应矢量量化方法实现图象信息的压缩	许金生	李长河	
8.4	视觉在图象编码中的作用	张又又	董志信	285
8.5	矢量量化在静态图象编码中应用	孙山东		290
8.6	图象信号离散余弦变换域的矢量量化编码方法	杨有为	赵若平	292
		何任杰	张总成	294
8.7	图像等值线编码存储及重建的改进	沈雪勤	刘锡安	
		宁 飞	赵 红	298
		林 虹		

<九> 神经网络

9.1	利用神经网络对手写数字识别的探讨	罗四维	杜 矛	302
9.2	基于 HOPFIELD 神经网络模型的启发式学习算法	袁宏文		
	及其在数字模式处理中的应用	姚毓林	张逸敏	304
9.3	用于视觉模式识别的神经网络模型	洪 进		
9.4	识别手画的一般几何图形	张胜持		308
9.5	用于文字识别的自适应神经元网络系统	季寒冰	戚飞虎	312
		何 煜	叶以正	316
9.6	随机点立体图对体视匹配神经网络的研究	李仲荣		
9.7	基于人工神经网络的低限制手写数字识别	贺朋令	蔡元龙	318
9.8	单层逻辑神经网络	林慎机		320
		杨国庆	陈松涛	323
		徐 涛		

<十> 汉字识别

10.1	手写体汉字的 PLS 识别法	吴敏金	李 蓓	327
10.2	基于数学形态学并具有启发式搜索功能的	唐 松	郑南宁	331
	印刷体汉字分类方法	郭椿标		
10.3	一种基于笔划元和笔划元对的限制性手写汉字识别方法	刘 永		335
10.4	基于形状分析法的行驶车辆牌照数字识别	胡 超	戴 莹	337
		郑南宁		
10.5	自由手写字符预处理的专家系统	张心叶		341

<十一> 视觉应用

11.1	计算机物体识别与定位系统	毕红葵	彭嘉雄	343
11.2	微循环血流中白细胞的识别	徐尚平	骆力	347
11.3	干涉条纹图的计算机自动分析及其应用	李树祥	李允明	351
11.4	指纹模式匹配的“中心 - 近邻”方法	赖建苏		
11.5	多层印刷电路板 PCB X 线照片的自动识别	朱剑平		355
11.6	印刷电路板缺陷的视觉检验	李军		361
11.7	用保护连通性算法检查印刷线路板表面缺陷	王创新	罗仲广	363
11.8	手描逻辑电路图的计算机自动识别系统	李杰	宣国荣	366
11.9	肺癌细胞图像的计算机识别系统	王正安	宣国荣	371
		张文蔚	严洪范	375
		严洪范	刘敏	379
11.10	保持连通性收缩的印制电路板检测方法	唐立吾	施鹏飞	
		陈文琦		

<十二> 图像处理

12.1	图象处理中形态学变换的软件实现	戴震	郭明杰	384
12.2	灰度图象放大的平滑处理	毕都彦		388
12.3	图象平滑在红外成像系统中的应用	李名世		390
12.4	一种用于线图两值图象的细化算法	黄沛然		397
12.5	考虑灰度倾斜方向的点阵打印机浓淡图象表示	王吉仁		401
12.6	一种文本图象快速平滑细化的算法	金天荣		403
		唐松	郭椿标	
12.7	图像分析中的边界链码应用	郑南宁		407
12.8	子像素对准研究综述	胡泽民		412
12.9	一种用于图象识别和畸变参数估计的快速方法	郑兆瑞	田琦	423
12.10	数学形态图象边缘检测	李克勤	宋文强	431
12.11	数学形态学加权运算及其在二值图象处理中的应用	丁润涛	路坤	433
12.12	积木世界物体透视线图的三维理解	谭兆信		436
		姚键超	宣国荣	

<十三> 其它

13.1	模式信息处理的进展	宣国荣	黄泰翼	438
13.2	自动寻的制导与计算机的人工智能技术需求	张国友		445
13.3	场域视觉信息时空转译——计算机图象处理和航空航天遥感技术实用于风景景观工程的方法研究	刘滨谊		447
13.4	计算机图象与视觉设计	赵淑华	渠速	453
13.5	一种适合模式线状分布的快速分类法	聂大同	高鹏飞	456
13.6	从图象处理到计算机视觉	刘建庄		460
13.7	人类视觉和机器视觉中的群集	程虎		462
13.8	一类快速匹配算法的实验与比较	方驰		473
		丁明跃	王宇军	
13.9	手写汉字识别的自然层次映射法及其模拟实验系统	彭嘉雄		475
		郑重	陈玲	
		陈学德		

附录

- A1. VICOM/VME 系列图像与图形处理工作站
南科电子有限公司
- A2. 翔裕系统工程公司 (FORTEK SYSTEMS COMPANY)

基于特征的作业面空间姿态和旋转变换 参数计算方法

康南生

吴真娜

(中国科学院合肥智能机械研究所) (安徽大学电子工程系)

【摘要】假定已知目标物体的几何结构及特征数据,用其经视觉中心投影所得2—D图象数据来计算其3—D空间的姿态参数和旋转变换参数,以期实现满足实时性要求的单眼视觉单幅图象进行机械手运动引导的目的。本方法基于下述命题,即目标物体作业面上的三个特征点组成的特征三角形与其中心投影所成之2—D象是一连续映照关系,其勃劳维尔不动点即为该特征三角形的内心。

一、引言

在机械手执行拾取,装配等任务作业时,需要估计目标物体上与此任务有关表面的3—D空间位姿参数,以引导机械手运动至适当的作业空间。我们称这任务有关表面为目标物体的作业面,如在拧螺丝任务中,目标物体上螺孔所在面。在智能机器人系统中,目标物体位姿参数的获取和运动轨迹规划,主要是依赖于机器人视觉,即摄像机得到的观察数据。

本文提出的计算方法使用手上单眼视觉系统。其优点是:① 较之双眼视觉,单眼视觉提供进行图象预处理(如特征点抽取和匹配等)的数据量较少,有利于满足系统实时性的要求;② 由于摄像机离目标物体较近,能给出较高的空间采样分辨率,可提高视觉定位精度;③ 通过控制机械手的运动,可避免目标物体作业面被遮挡的可能。其缺点是:① 缺乏直观有效的深度信息;② 对各种算法的实时性有更苛刻的要求。

降低计算时间耗费的基本措施之一是要确定,在满足特定环境和约束条件下,求解问题所需的最少数据量,即最少的图象幅数和特征数。R. Horaud [1] 用立方体的一项点及共此点的三条边的2—D投影象去匹配物理世界中的真实物体,并给出其空间位姿参数的估计值Roach和Aggarwal [2] 用单眼视觉两幅图象,五个特征匹配点估计作连续,平滑运动的刚性物体的运动参数。本方法用目标物体作业面上的三个特征点的单幅2—D投影象来估计该作业面的3—D空间姿态参数。用2—D图象数据估计目标物体3—D空间6个自由度位姿参数,即3个位移参数和3个转动参数,难点在于缺乏深度信息,以前的研究[3]和工作(如[4]),都假定目标物体是放置在工作台或托架上,而放置平面在机械手基本笛卡尔空间位姿参数为已知,这样只要观察到目标物体在已知平面上的位置即可求出其深度参数。本方法无此约束,作者在[5]中给出了位移参数的算法,本文则局限于讨论三维空间中姿态和变换矩阵参数的算法。

二、目标物体作业面的投影模型

1、作业面,特征三角形及变换矩阵表示法

若目标物体作业面是一平面,则该作业面可由其上3个不共线的特征点组成的平面完全定义。若作业面是一曲面,作业面上3个特征点构成的平面与作业面的关系可由 4×4 齐次变换矩阵 T 表示,其脚标表明由特征点平面到作业面的变换矩阵,由于目标物体的几何结

构已知, T 为一常数矩阵, 因而本文将用特征点平面上的特征三角形替代作业面进行讨论。

机械手手爪坐标系与安装在手爪上的摄像机构成的投影坐标系之间的关系可由一已知常数齐次变换矩阵 T 定义。本文讨论的特征三角形在投影坐标系中的参数可由 T 和 T 变换为作业面与机械手手爪之间的关系。在机器人学中 [3], 齐次变换矩阵 T 定义为:

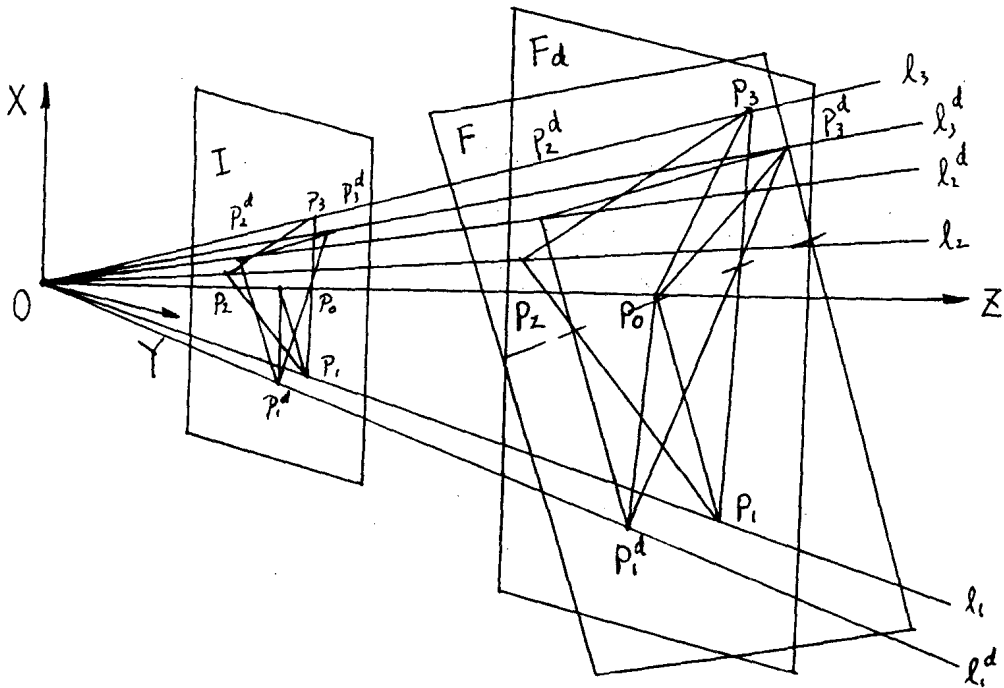
$$T = \begin{bmatrix} R & P \\ 0 & 1 \end{bmatrix} \quad (1)$$

式中 R 为 3×3 旋转变换矩阵, P 为 1×3 位移变换列向量, 本文局限于讨论作业面特征三角形在投影坐标系中的旋转变换矩阵 R 的参数。

2、坐标系的定义(参见附图)

(1) 投影坐标系 $OXYZ$, 原点 O 与摄像机透镜中心重合, Z 轴与光轴重合, 方向为机械手手爪 $\vec{a}$ 矢量 [3], X 轴和 Y 轴分别与手爪的 $\vec{s}$ 和 $\vec{r}$ 矢量方向相同。

(2) 图象平面坐标系 oxy , 为分析方便起见, 设图象平面与焦平面重合, 且焦距为单位长度 1, 坐标系原点在投影坐标系中的坐标为 $(0, 0, 1)$, 其 x 轴和 y 轴分别平行于 X 轴和 Y 轴, 且具有同样的标度。



附图: 特征点平面和特征三角形

(3) 理想作业面特征三角形 F , 坐标系 $OXYZ$, F 为作业面 F 经机械手相对运动后在投影坐标系中具有理想的姿态参数, 设 F 平面平行于图象平面, 原点 O 在投影坐标系中坐标为 $(0, 0, Z)$, 且位于 F 特征三角形内心。 Z 轴方向为负 Z 轴方向, 且与 F 平面法向量相同, 而 X 轴, Y 轴则分别与 X 轴、 Y 轴方向相同, 且具有相同的标度。

设 F 三特征点及其坐标分别为: $P_1', (X_1', Y_1', Z); P_2', (X_2', Y_2', Z); P_3', (X_3', Y_3', Z)$, 其在图象平面上的对应投影象 I 仍为三角形, 其顶点分别为 $P_1^d, (x_1^d, y_1^d); P_2^d, (x_2^d, y_2^d); P_3^d, (x_3^d, y_3^d)$;

和 p_i' , (x_i', y_i') , 根据投影定理有:

$$[x_i' \ y_i']' = \frac{1}{Z} [x_i \ y_i]', \quad i=1, 2, 3 \quad (2)$$

已知的特征参量还有3特征点到内心的距离, 分别设为 E_1, E_2, E_3 , 其在图象平面上的对应特征为 e_1', e_2', e_3' , 有:

$$e_i' = E_i / Z, \quad i=1, 2, 3 \quad (3)$$

(4)在[5]中作者提出的方法将作业面 F 特征三角形的内心经位移变换至与 F_i 特征三角形内心重合, F 的3特征点 P_1, P_2 和 P_3 在投影坐标系中的坐标分别为 $(X_1, Y_1, Z_1), (X_2, Y_2, Z_2)$ 和 (X_3, Y_3, Z_3) , 其在 F_i 坐标系中的坐标分别为 $(X_1', Y_1', Z_1'), (X_2', Y_2', Z_2')$ 和 (X_3', Y_3', Z_3') , 其中:

$$Z_i' = Z_i - Z, \quad i=1, 2, 3 \quad (4)$$

F 在图象平面上的投影象三角形 I 对应顶点及坐标分别为 $p_1(x_1, y_1), p_2(x_2, y_2)$ 和 $p_3(x_3, y_3)$, I 内心到三项点特征线长度分别为 e_1, e_2 和 e_3 , 有:

$$[x_i \ y_i]' = \frac{1}{Z_i} [X_i \ Y_i]', \quad i=1, 2, 3 \quad (5)$$

三、特征三角形的映照性质

命题I, 作业面特征三角形到图象平面上的中心投影象是一连续映照, 其反投影也是一连续映照。

该命题的证明可从附图中简单而直观地得到, 我们更感兴趣的是它们之间的其它某些映照关系和性质。(1)三维空间的作业面到二维图象平面是一降维变换。在平行投影下投影变换核的零度为1, 即投影方向上的矢量(深度信息)被映射成零向量。而在中心投影下被映射成零向量的是空间基向量的某种组合, 从而有结论为: 在具有一定先验知识条件下, 可从2—D中心投影象恢复出原象的3—D空间参数。(2)从图中可看出二维象 I 到三维空间的反投影象可以是分别位于中心投影射线 l_1, l_2 和 l_3 上任意三点构成的无穷多的三角形, 即反投影是一对多的。但在原象几何参数一定的条件约束下, 其反投影是一对一的, 如作业面特征三角形的特征已知量, 则:

$$E_i = (X_i^2 + Y_i^2 + (Z_i - Z)^2)^{1/2}, \quad i=1, 2, 3 \quad (6)$$

由(5)(6)组成的方程组可以求解出 F 平面3特征点的空间坐标参数, 进而确定作业面的位姿参数。但不幸的是对于非线性方程组还没有一种系统的求解理论和方法, 若采用数值迭代法求近似解, 那么初始估计值选择的好坏直接影响着收敛的性质和速度, 或者可能收敛至错误的答案导致失败, 显然在实时性很强的系统中是不合适的。

命题II, 具有任意空间姿态的作业面特征三角形与其中心投影所成之三角形象, 有一连续映照不动点即三角形的内心。

从附图可得该命题的直观证明叙述如下: ①设在投影空间中, 平面 ZOl_1 为平面 l_1Ol_2 和平面 l_1Ol_3 的分角平面, 可看出顶点位于 l_1 射线上, 两边分别位于平面 l_1Ol_2 和 l_2 和 l_3 上任意角的平分线为该角所成平面与平面 ZOl_1 的交线, 进一步可证得顶点分别位于 l_1, l_2 和 l_3 上的任意三角形, 其三个角的平分线为该三角形组成的平面分别与三分角平面的交线, 而该三分角平面交于 Z 轴, 即顶点分别位于 l_1, l_2, l_3 上的任意三角形内心必位于 Z 轴上。即内心位于 Z 轴上的任意三角形在图象平面上所成象三角形的内心仍然在 Z 轴上, 反之也成立, 即图象平面上的三角形内心位于 Z 轴上, 则其反投影到三维空间的任意三角形的内心仍位于 Z 轴上。②设 F 坐标系与 F_i 坐标系的原点重合且位于 F 和 F_i 特征三角形的内心, 由①可知 F 与其在图象平面

上的中心投影象 I 有一连续映照不动点位于 Z 轴上, 同样 F_i 与其中心投影象 I_i 也有一连续映照不动点位于 Z 轴上。即三维空间平面特征三角形旋转变换前后的中心投影象有一变换映照不动点即三角形内心。

四、参数计算

1、目标物体作业面特征点空间坐标参数

由于 F_i 特征三角形在投影坐标系中的坐标为已知参数, 由公式(2)可在图象平面上生成其的投影象 I_i , 若将 F 绕 Z 轴旋转 ϕ 角度, 则使得在图象平面上 e_i 与 e_i' 重合, 在投影空间中则有平面 l_iOZ 与平面 $l_i'OZ$ 重合。 ϕ 的值可直接由图象平面上的特征矢量值得得:

$$\phi = \cos^{-1} \left(\frac{\vec{e}_i \cdot \vec{e}_i'}{e_i \cdot e_i'} \right) \quad (7)$$

式中, $\vec{e}_i$ 和 $\vec{e}_i'$ 分别为图象平面坐标系中的矢量 $\vec{op}_i$ 和 $\vec{op}_i'$,

经 $R(Z, \phi)$ 旋转变换后, E_i' 与 E_i 位于同一平面 $l_i'OZ$ 上, 其夹角 θ 可由特征量表示如下,

$$\theta = \cos^{-1} \left(\frac{e_i}{e_i'} \sin(tg^{-1}e_i) \right) - tg^{-1}e_i \quad (8)$$

而 $Z_i = Z + \text{SIGN}(e_i - e_i') e_i' Z \sin(tg^{-1}e_i), \quad i=1, 2, 3 \quad (9)$

式(5)(9)完全定义了作业面特征点的空间坐标参数。

2、旋转变换矩阵和运动学方程求解。

当作业面特征点空间坐标确定后, 可以直接求出 F 与 F_i 坐标系的旋转变换矩阵参数。

$$\begin{bmatrix} X_i' \\ Y_i' \\ Z_i' \end{bmatrix} = \begin{bmatrix} n_i & o_i & a_i \\ n_i & o_i & a_i \\ n_i & o_i & a_i \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i' \end{bmatrix}, \quad i=1, 2, 3 \quad (10)$$

由旋转变换矩阵表示的机械手笛卡尔空间姿态到关节空间的变换求解过程即为运动学方程问题, 以求表示旋转运动参数横滚(Roll), 俯仰(Pitch), 偏转(Yaw)角为例

$$\begin{bmatrix} n_i & o_i & a_i \\ n_i & o_i & a_i \\ n_i & o_i & a_i \end{bmatrix} = \begin{bmatrix} C\phi C\theta & C\phi S\theta S\psi - S\phi C\psi & C\phi S\theta C\psi + S\phi S\psi \\ S\phi S\theta & S\phi S\theta S\psi + C\phi C\psi & S\phi S\theta C\psi - C\phi S\psi \\ -S\theta & C\theta S\psi & C\theta C\psi \end{bmatrix} \quad (11)$$

式中 C 表示函数 $\cos(\quad)$, S 表示函数 $\sin(\quad)$ 。求解过程参见[3]。

3、基于特征的RPY(ϕ, θ, ψ)矩阵参数求解

为方便起见而又不失一般性, 设 X_i 轴与 E_i' 重合, 在图象平面上 I_i 中则有 e_i' 与 x 轴重合。

(1)横滚转动 $R(Z, \phi)$, 即 F 绕 F_i 坐标系 Z_i 轴旋转 ϕ 角度, 从而使 e_i 与 e_i' 在图象平面上重合, 而在投影空间中则有平面 l_iOZ 绕 Z 轴旋转至与 $l_i'OZ$ 平面重合, ϕ 角由式(7)决定。

(2)俯仰转动 $R(Y, \theta)$, 该旋转变换将使三维空间中的 P_i 点与 P_i' 点重合, 在图象平面上使 p_i 得与 p_i' 点重合。此时 F 与 F_i 将相交于 X_i 轴。 θ 角由式(8)确定。

(3)偏转转动 $R(X, \psi)$, 设 P_i 点经两次旋转变换至 P_i' 点, 其坐标为,

$$[X_i' \ Y_i' \ Z_i']^T = R(Y, \theta)R(Z, \phi)[X_i \ Y_i \ Z_i]^T \quad (12)$$

由(5)式可得 P_i' 点在图象平面上的坐标为 (x_i', y_i') , 此可写出图象平面上连结 p_i' 和 p_i' 点的直线

方程并得到它在y轴上的截距, 设为d, 则

$$d = y'_1 - x'_1 (y'_2 - y'_1) / (x'_2 - x'_1) \quad (13)$$

在 F_1 坐标系中得连结 P'_1 与 P'_2 点的直线参数方程为,

$$x = (x'_1 i + y'_1 j) + t(X'_2 i + Y'_2 j + Z'_2 k) \quad (14)$$

设该直线与 $Z_0 Y_0$ 平面交点到原点的距离为D, 则

$$D = [(Y'_1 - X'_1/X'_2)^2 + (X'_1 \cdot Z'_2/X'_2)^2]^{1/2} \quad (15)$$

$$\psi = \cos^{-1} \left(\frac{D}{d \cdot Z} \sin(\text{tg}^{-1} d) \right) - \text{tg}^{-1} d \quad (16)$$

(4)求得RPY (ϕ 、 θ 、 ψ) 矩阵为:

$$\text{RPY}(\phi, \theta, \psi) = R(X, \psi) R(Y, \theta) R(Z, \phi)$$

五、结论

本方法假定已知机械手操作目标的几何特征参数, 这种前提具有实际应用的意义, 因为目标物体几何参数可离线测量, 且可用离线编程或CAD法输入系统。依据决定一平面的最少特征点数, 即3个特征点构成的特征三角形与其中心投影象之间的连续映照关系和连续映照不动点即特征三角形内心的性质, 由其二维图象的特征数据, 计算出目标物体的三维空间参数和旋转变换参数。该方法使用单眼单幅图象不仅恢复了空间物体的深度信息, 而且给出了尽可能满足实时系统要求的算法。本方法只适用于静止的目标物体, 但它不象其它一些视觉方法对目标物体有一定的约束条件, 因而为实现目标物体运动参数估计和动态视觉引导研究打下了基础, 这是下一步要开展的工作。

参考文献:

- [1] R. Horaud, "New Methods for Matching 3—D Objects With Single Perspective Views," IEEE Trans, Pattern Anal. Machine Intell., Vol. PAMI—9, No. 3, pp. 401—412, May 1987.
- [2] J. W. Roach, J. K. Aggarwal, "Determining the Movement of Objects from a Sequence of Images," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI—2, No. 6, pp 554—562, Nov. 1982.
- [3] R. P. Paul, "Robot Manipulators, Mathematics, Programming, and Control," Cambridge, Mass., MIT Press, 1981.
- [4] A. G. Maktin, "Stability and Sensitivity of Servo Vision Systems," Proc. 5th Int. Conf. on Robot Vision and Sensory Controls. (Amsterdam, Oct. 1985) RoViSeC5, pp 79—89.
- [5] _____, "Feature Based an Approach to Vision Guide," will be published.

论 机 器 人 视 觉

胡庆茂 陈锦江

华中理工大学自控系

【摘要】 本文分析了视觉在机器人中的应用特点，把用于导引的视觉定义为机器人视觉，而把用于检验的视觉划入计算机视觉之列，并给出了机器人视觉的定义及成功的标准，从基本概念、基本方法及实现方面论述了机器人视觉与计算机视觉的差异，并着重讨论了机器人视觉研究的一些特殊问题：机器人视觉传感器、场景的照明、二值图像与灰度图像、机器人视觉传感器在空间上的配置策略。

1. 引言

在机器人的各种传感器控制中，相对于触觉、听觉而言，视觉的研究最多。然而，机器人视觉通常与用于自动检验中的视觉甚至是场景分析中的智能相混淆。首次进入机器人视觉领域的用户和研究人员会感到棘手，因为目前他能得到的是一大堆计算机视觉方面的书，而机器人视觉还没分离出来，只是其中的一部分。但是，机器人视觉不论是其基本概念、基本方法还是其实现都与计算机视觉的其他领域有明显的不同。机器人视觉的研究目前还处于感性认识的阶段，没有一套系统的理论，甚至还不曾有过“机器人视觉”的定义。因此，很有必要探讨机器人视觉学科的基本概念，以及机器人视觉领域的一些特殊问题。

2. 成功的机器人视觉的评判标准

典型的视觉系统由摄像机、计算机及接口、软件组成，这种系统能

胜任很多工作。然而，从“看”的常规意义上讲，还远不能说计算机或机器人能“看”，因为不论是生物视觉还是人工视觉都还没得到很好的了解，对它们进行描述也是困难的。因此，将从视觉系统的外部特性如速度、精度、成本来评价它们。

成功的机器人视觉要求如下：低成本、可靠操作、快速图像处理、场景照明简易。

上述要求通常与研究人员发表的研究成果相抵触，高分辨的灰度图像通常能得到精度较高的结果，但代价是处理器的结构复杂或需更多的处理时间。专用机（如并行处理器）能以较高的速度获取精度较高的结果，但这通常需要多个领域的工程人员相互协作来解决图像处理的技术问题，使图像处理复杂化了，同时成本又太高。

场景的照明控制是一个难题，但通常又被忽视。早期成功的机器人视觉系统例子如 CONSIGHT，充分显示了结构光照明的优越性。只有通过合适的视觉传感器、良好的环境照明及灵活的软件共同配合才能形成成功的机器人视觉。早期的研究成功实例多局限在场景照明控制能够实现的场合。对于场景照明不好控制的环境，今后将考虑采用专用机并行处理结构，这一方面是由于电子技术的飞速发展使得其价格大幅度下降，另一方面在于这种场合要想获得理想的效果，必须使用复杂的计算机结构或专用芯片。

3. 机器人视觉的一些基本概念辨析

下面分析一下视觉在机器人中的应用特点。对于工业制造过程，视觉在机器人中的应用可分为两类〔1〕，即视觉导引与视觉检验。这两者存在着很大的差异，用于导引的视觉通常需要机器人以主动的方式完成其功能（如抓取、传送、运动、放置、跟踪、切割、装配等），速度是极其重要的因素；而用于检验的视觉通常不需要机器人对环境的作用，且由于检验通常追求的是可靠，速度要求放宽，分辨率要求

高，这与计算机视觉的其它应用领域（如航空图像的资源分析等）相似。因此，我们把用于机器人检验的视觉划入计算机视觉之列，而把用于机器人导引的视觉叫机器人视觉。

根据Marr教授的观点〔3〕，视觉是一个处理过程，它能从外部世界的图像中得到一个既有用又不受无关信息干扰的描述。

计算机视觉也叫机器视觉，它是将用于视觉感知的各种处理和表示综合在一起，并使其实现自动化的一种技术〔3〕。它包含许多本身已较成熟的技术，如图像处理（图像的变换、编码及传输）、模式识别、几何建模以及认识处理的一些手段。

纯计算机视觉研究常常涉及相对来说与应用领域无关的一些问题。其基本研究方法是模仿生物视觉，这种方法在机器人视觉导引中是欠妥的，这正是将机器人视觉从计算机视觉中分离出来的原因之一。当然还有其它因素，这就是在4中要讨论的机器人视觉中的一些特殊问题。

视觉是计算机视觉中研究得最多的一个问题，这也是计算机视觉模仿人类眼睛的例证。在计算机视觉中，在传感器的空间配置上，一般也是“双目”位于空间的上方，这种特定的空间配置使得计算机视觉给出的描述只能是物体的位置、姿态、类别，一般地讲不可能给出物体的抓取位置，这种描述在机器人视觉中又显得不够。为充分中解这一点，只需看一反例。作为面向机器人操作的机器人视觉，它给出的描述必须能使机器人对对象进行操作（主动式），如果物体外形比较复杂，并不是在任何位置都可抓取，则必须给出抓取位置才能保证机器人手爪将物体可靠地抓取和操作，原因是机器人的手爪不可能做得很复杂（比如象人的手那样灵活复杂）。早期的机器人视觉研究没有提出抓取位置这一指标，是由于物体形状简单，抓取位置易于在模型中得到。对于复杂物体，必须以某种方式得到抓取位置这一特征，这在机器人的传感器空间配置时必须充分考虑。

机器人视觉中通常对速度、可靠性的要求较高,有时精度倒并不是十分关键的因素。也就是说,成功的机器人视觉并不是去追求摄像机的高分辨率,基本的考虑应是选择执行操作所需的最优分辨率。利用 50×50 的低分辨率也取得过成功的应用〔4〕。在装配中,运用视觉将位置与姿态的误差检测出来,形成主动伺服反馈,这种校正亦是粗糙的;如需细微的校正,可用顺应腕进行被动式协调。在机器人视觉中,精度的提高亦不是靠提高分辨率得到的,使摄像机离物体更近是最有效的手段,这在机器人的传感器配置策略中还要讨论。

当然,机器人视觉与计算机视觉有许多相似之处。例如,很难讲出机器人视觉所要求的图像处理技术与计算机视觉所要求的图像处理技术的差异;对于复杂的视觉计算问题,二者都要解决提高速度的问题。在计算机视觉中,一方面研究并行处理器来提高速度,另一方面还在利用通用机来完成一些实时性要求不很高的工作。在机器人视觉中,由于速度要求甚高,于是处理只有两类。第一类,回避复杂的处理,即用结构光来形成二值图像进行处理,早期的CONSIGHT、PUMA/VS-100属此列;第二类,采用专用硬件,或使用并行处理芯片、并行处理器进行处理,如AIS-5000并行处理器〔5〕、OLIPTA图像处理器〔6〕OMS-视觉系统〔7〕。有人〔8〕甚至对视觉系统进行了这样的分类,即第一代视觉系统没有利用并行处理,处理器是进行串行处理的冯·诺曼处理器,第二代视觉系统使用了并行处理计算机,算法亦是并行的。

基于如上的分析,对机器人视觉给出如下定义。

机器人视觉是这样的一种过程,其输入为一幅或多幅图像,而其输出为满足下列条件的描述:

- a: 这种描述必须是待感知场景内物体有关特性的某种描述;
- b: 这种描述必须具有为使机器人完成某一特定任务所需的信息。

一般地,这种描述为物体的位置、姿态、类别、抓取位置;

○:这种描述必须满足机器人操作的实时性要求。

按照上面给出的定义,只有机器人的视觉伺服控制与图像处理及理解同时进行的视觉系统才是机器人视觉系统。然而,由于速度限制,目前大多数视觉伺服实验均是静态看——动〔9〕,即机器人的控制及图像处理不同步,这种情况在生产过程中不可能出现。

4. 机器人视觉研究的一些特殊问题

4.1. 机器人视觉传感器

研制机器人视觉传感器的目的就是要充分利用现有的信息处理系统来有效地获取为机器人视觉所需的图像信息。电视摄像机是目前用得最多的、具有广泛来源的图像传感器,但它却不是合适的机器人视觉传感器,理由是:第一,三维世界通过电视摄像机的光学系统变成图像平面的二维光强分布时,三维物体对二维光强的变换是模糊的,因为这种变换依赖于光学、照明及物体表面特性。人由于具有大量的先验知识,因此能补偿这种变换的模糊性。然而,要让机器人具备这些先验知识至少在目前是不可能的。这样,采用电视摄像机作为机器人视觉传感器时,成像就带来了极大的模糊性,使得后续处理复杂,难于满足实时性要求。第二,电视摄像机本为工业电视设计,并未考虑在机器人视觉中应用的特殊问题。电视摄像机用作机器人视觉传感器时太重、尺寸太大,无法安装在机器人的手爪上。正因如此,机器人视觉传感器的研制一直为机器人视觉领域所关注,在能得到的关于机器人视觉较权威的最近 *Robotics* 文献中(1986),视觉传感器的文章占了一大半。

视觉传感器的研制表现在如下几个方面。第一,是研制能直接获取三维信息的三维传感器,在〔10〕中利用PSD(位置敏感器件)与LED(发光二极管)已研制出了样机,这可使得上述的变换模糊性不复存在。第二,研制适宜于安装于机器人手爪上的视觉传感器。在〔11〕

中有报道。这时传感器尽管是二维的，但由于摄像机离物体很近，因而分辨率可以低些。这样信息相对来说变少了；同时坐标变换变得简易，终端执行器至目标的距离只需进行相对测量，减少了机械位置的校准，以上这些对提高速度、降低成本都是十分有效的。第三，研制能直接提取有用信息的摄像机：传统的电视信号含有大量的信息，机器人视觉的图像处理任务只需这些数据中的一小部分来作出决策。轮廓摄像机O I D，从一幅图像中产生经过大量压缩了的数据，不失为一种研究方向。值得一提的是，虽然光纤不是视觉传感器，它却在机器人视觉中起着非常重要的作用。在手——眼系统中，利用光纤可得到分布式视觉：安装于手爪附近的光纤，利用相干光束将视场内的物体光强值传送至远处的摄像机，再由摄像机将这种光强转换成图像信号〔12〕，这对解决手爪的负荷能力有限与摄像机太重、尺寸太大的矛盾十分有用。

4.2. 场景的照明

场景的照明是研究图像恢复时容易忽视的问题。通常，由图像提取特征的算法之所以过于复杂，正是由于图像缺乏适当的对比度。尽量考虑照明设计来配合快速算法的实现是机器人视觉研究的又一个鲜明特点。

设计场景照明时，需要考虑选择何种照明器以及照明器的光线至视觉传感器的路径问题。

光线由照明器至传感器的方式有多种，它们是正面光(*front light*)、逆光(*rear light, back light*)、偏置光(*offset light*)、镜面光(*specular light*)。

通常正面光用来照射感兴趣的区域，此时光束可直接到达物体的表面，很大一部分结构光属于此类。

逆光用来检测由遮挡所致的阴影，这在相对测量中非常有用〔13〕

偏置照明使得入射光有一定的朝向，这可使诸如孔之类的特征产生阴影，从而确定特征的位置。

镜面光可由多种方法得到，目的是将光源的方向变成传感器光轴的方向。这样一方面可看见哪些无法直接照明的场景部分，另一方面又可检测出某些缺陷。

适当的场景照明设计是一项技巧性很高的工作。只要有可能，机器人视觉的应用都必须考虑照明的设计。

4.3. 二值图像与灰度图像

机器人视觉系统的输入是连续电信号，经过图像输入装置变成数字信号。此数字信号的比特数为多少时更便于后续的处理呢？

问题的关键在于现在的通用计算机处理速度有限，而机器人视觉的实时性要求甚高。当比特数较大时（比如说 8 比特），要想获取实时性，必须用专用硬件或并行处理器，但此时成本很高，不利于在生产实践中实现。二值图像便于处理，但却丢失了大量的信息，为获得满意的效果，通常采用结构光、增加对比度来补偿，这时可用一般的通用计算机完成处理。

实际情况可分成如下三类。第一类，场景照明能够控制，场景不复杂（例如物体较少，彼此不重叠），可考虑用二值图像，配合适当的照明。第二类，场景照明能够控制，但场景较复杂（例如物体较多，彼此可能重叠）；或场景照明不能控制，场景不复杂。前面要考虑用适当的照明配合，二者都用比特数较少（如 4 比特）的图像，添加一个辅助处理器进行处理。第三类，场景照明不能控制，场景较复杂，此时需采用比特数较大的灰度图像，用专用硬件或并行处理器进行处理。对复杂程度可变化的场景，利用〔1〕中介绍的工具箱进行多次感知，可很好地解决每一场景表述的优化比特数。

4.4. 机器人视觉传感器在空间的配置策略

这里，有两个问题决定了传感器在空间的配置问题。第一，获取的物体信息在空间上怎样分布时更有效；第二，要构造一个基于传感器控