

模糊化多视觉信息融合的视觉跟踪策略

张国亮, 谢宗武, 蒋再男, 王捷, 刘宏

(哈尔滨工业大学机器人技术与系统国家重点实验室, 150080, 哈尔滨)

摘要: 为了提高视觉跟踪系统在空间环境对外界变化的适应能力, 提出一种模糊化多视觉信息融合的视觉跟踪策略. 该策略综合了多个反映目标特征的视觉信息, 通过对选择的每一项子特征集进行模糊化处理, 提高了识别的智能化程度. 针对运动物体分割阈值受噪声干扰的问题, 提出一种结合人口统计与区域生长的区域分割方法, 实现了运动区域特征的稳定聚类. 相对于传统的局部特征识别方法, 这种多特征融合技术适合应用于复杂的动态环境, 冗余信息间的互相补充可确保特征在不确定环境中的识别. 试验结果表明, 该策略能够实现鲁棒的特征提取和跟踪, 尤其当出现被跟踪物体部分遮挡以及外界发生光线变化时, 视觉跟踪系统仍能够实现稳定、快速的识别.

关键词: 多视觉信息融合; 空间机器人; 视觉跟踪

中图分类号: TP242.6 **文献标志码:** A **文章编号:** 0253-987X(2009)08-0033-05

Visual Tracking Strategy Based on Fuzzy Multi Visual Cue Integration

ZHANG Guoliang, XIE Zongwu, JIANG Zainan, WANG Jie, LIU Hong

(State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150080, China)

Abstract: To improve the adaptability of visual tracking system for various circumstances in space, a visual tracking strategy based on fuzzy multi visual cue integration is proposed, which synthesizes multi visual cue together to represent the features of object. The recognition of object is intelligentized by integration of each fuzzy cue. To deal with the effects of noise on segmentation threshold of moving object, a new region-based segmentation is chosen to realize stable clustering of moving region by integrating population statistics and region growing. Compared with conventional local feature recognition, this strategy is adaptive to complex and dynamic circumstances. Redundant information complements each other for ensuring feature recognition and tracking in uncertain environment. Experimental results demonstrate the robust feature recognition and tracking with this strategy, especially in part occlusion and unstable lighting.

Keywords: multi visual cue integration; space robot; visual tracking

随着人类征服太空脚步的加快, 空间机器人的发展和应用受到有关国家的高度重视, 由此而产生的空间机器人相关技术也成为机器人研究领域中最具吸引力的课题. 与传统的遥操作方法相比, 视觉自主跟踪及伺服方法由于能够切断遥操作通讯时延, 使空间机器人具有自主操作的能力, 因此受到了国内外学者的广泛关注. 目前, 许多空间机器人的任务都是由视

觉控制完成并实现的, 如 Canadarm2、ETS7^[1]等.

针对空间环境对于视觉跟踪的要求, 本文设计了一种模糊化多视觉信息融合的视觉跟踪策略. 与传统的单一特征集识别跟踪方法不同, 此策略综合了多个反映目标特征的视觉信息, 通过对这些特征子集进行一定的加权, 实现了多视觉特征的融合, 从而获取到期望的视觉特征输出.

1 多视觉信息融合的视觉跟踪

高效、快速的图像处理算法是实施视觉跟踪的前提,而图像特征的选择对于视觉跟踪系统的性能具有重要的影响,不合适的特征选择由于不能提供充分的目标信息而可能导致视觉伺服失败。目前,视觉跟踪研究所用到的图像特征主要有两类:一类是采用简单几何形状的局部特征,如点、直线、圆、直线斜率、椭圆的轴长等^[2];另一类是采用全局特征,如基于目标轮廓的特征选择和基于区域的特征选择^[3]。然而,无论采用何种特征,如何提高视觉跟踪算法的鲁棒性和可靠性一直是视觉跟踪算法面临的难点^[4]。当跟踪运动目标发生遮挡或运动目标所处环境的光照发生变化时,仅仅依靠单一特征集进行特征识别,往往会由于特征的暂时失效而导致视觉跟踪的失败。例如采用基于局部几何特征的识别,当需要处理摄像机运动的视觉跟踪时,容易造成图像特征在视场范围内的丢失,而基于区域的特征识别方法虽然不要求特殊的图像结构,对图像的畸变具有较强的鲁棒性,但对背景变化敏感,当环境光照发生变化时,以区域或全局特征作为跟踪基础的视觉跟踪算法一般都会失效。

在空间环境中,光线的变化以及由于目标的运动而引起的部分遮挡是经常发生的,这就要求视觉跟踪算法对于这种环境的变化具有足够的鲁棒性。针对这一问题,本文设计了一种模糊化多视觉信息融合的视觉跟踪策略。该策略并不是简单地将图像的局部特征或全局特征进行叠加,而是采用多个图像特征类别的组合,例如颜色、边缘、纹理、图像的运动等,然后通过模糊逻辑和加权的方法得出结论,这样可以保证在更大的尺度和范围内来表示目标的信息,从而提高视觉跟踪系统对于环境变化的鲁棒性。图1表示了这种多视觉信息融合的结构。

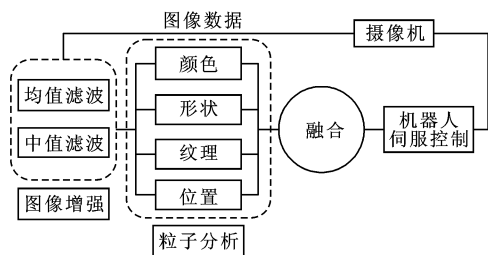


图1 多视觉信息融合的结构

2 视觉处理算法

2.1 图像增强与颜色识别

在机器人的彩色视觉处理中,CCD与外界条件有一定关系,物体通过透镜成像到CCD面阵上一点的RGB值可以通过三刺激值计算公式来求解^[5]

$$\left. \begin{aligned} S_R &= \int_{\lambda} S(\lambda) \beta(\lambda) \tau(\lambda) r(\lambda) d\lambda \\ S_G &= \int_{\lambda} S(\lambda) \beta(\lambda) \tau(\lambda) g(\lambda) d\lambda \\ S_B &= \int_{\lambda} S(\lambda) \beta(\lambda) \tau(\lambda) b(\lambda) d\lambda \end{aligned} \right\} \quad (1)$$

式中: S_R 、 S_G 、 S_B 为RGB的积分值; $S(\lambda)$ 为光源功率分布系数; $\beta(\lambda)$ 为物体光谱辐射亮度; $\tau(\lambda)$ 为透镜的光谱透射比; $r(\lambda)$ 、 $g(\lambda)$ 、 $b(\lambda)$ 为三原色的标准光谱刺激值。

由式(1)可以看出,不同的光源具有相异的 $S(\lambda)$ 。即使是太阳光线,在清晨与中午的功率分布系数也是有很大差异的。可见,物体同一块区域的表面在CCD上一点成像的RGB值是随光源的改变而变化的。虽然在常用的照明设备下这种改变量不是太大,但其幅值足够影响到阈值滤波的精度。因此,仅靠颜色识别方法很难选择一个确定的颜色阈值来适应不同的光照条件。若将阈值滤波范围设置得较宽松,那么会无法滤去背景噪声,从而导致位置、姿态求解精度的下降或失败,而如果滤波阈值一定,当物体在不均匀的光场中运动时,标记区域的大小、形状将发生变化,此时获得的视觉特征的可靠性较差。

为适应不同光线下的颜色识别,本文采用的方法是首先对视觉图像进行增强处理,滤除背景的跳动和椒盐噪声,同时使图像中的各颜色块更加集中,这样不但可以加快后续的图像处理速度,而且易于进行颜色的识别分割。此时颜色分割阈值设立得较宽松,这样可能会出现误分割的情况,但由于有后续的其他特征子集的处理,并不会对最终的识别造成影响。

令 S_{xy} 表示中心在 (x, y) 点、尺寸为 $m \times n$ 的矩形子图像窗口的坐标组,均值滤波的过程就是计算由 S_{xy} 定义的区域中被干扰图像 $g(x, y)$ 的平均值,即在任意点 (x, y) 处,复原图像 $f(x, y)$ 的值就是用 S_{xy} 定义的区域像素计算出来的算术均值

$$f(x, y) = \frac{1}{nm} \sum_{(s, t) \in S_{xy}} g(s, t) \quad (2)$$

均值滤波的目的是简单去除图像中的随机噪

声. 为尽可能使图像中的颜色色块集中并减弱噪声的影响,在均值滤波后加入中值滤波. 中值滤波方法类似于卷积,但计算不是求加权和,而是把邻域中的像素按灰度级进行排序,然后选择该组的中间值作为输出像素值. 如图 2c 所示,滤波后的图像颜色色块更加集中,并尽量保持了图像细节的边缘.

在得到了增强处理的图像后,进行 RGB 向量分割. 假如目标是在 RGB 图像中分割特殊彩色区域的物体,给定一个感兴趣彩色的有代表性的彩色样品集,可得到一个期望分割的平均色 a ,令 z 代表 RGB 空间中的任意一点,如果它们之间的距离小于特定的阈值 D_0 ,表示 z 与 a 是相似的,即分割出了期望的颜色(见图 2d), D_0 将作为颜色子特征集的阈值 H_c .

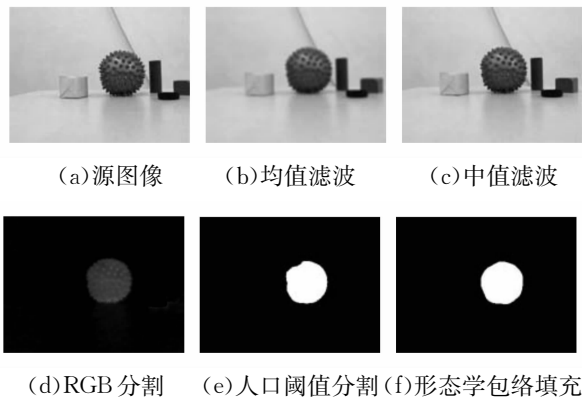


图 2 图像的增强与分割

2.2 分割目标

当物体与背景有较强对比时,采用传统的固定阈值或利用双峰直方图对物体进行分割,一般可以获得较满意的结果. 然而,在进行机器人视觉跟踪时,物体和背景的灰度是不断变化的,这时,一个在图像中某一区域效果良好的阈值在其他区域却可能效果很差,而且在摄像机随机机械臂一起运动的过程中,直方图也可能受到噪声的干扰而难以确定波谷阈值的数值. 为实现运动物体特征的稳定提取,本文提出一种结合人口统计与区域生长的区域分割方法.

首先确定一个人口密度基数 ρ_0 作为一个地区的拥挤程度,然后选择每个像素的 $m \times n$ 邻域作为人口扫描区域,通过比较该区域中存在的前景统计值与 ρ_0 ,就获得了一个地区的人口密度指数 ρ . 在完成人口统计后,根据事前定义的准则,将具有高密度指数的区域进行区域生长,从而聚合成更大的区域,而低密度指数区域将被设为背景. 这样,便实现了以

人口密度指数作为依据的人口增长,确保了运动过程中最核心的区域被提取出来. 其中, ρ_0 将作为运动物体分割子特征集的阈值 H_m .

由图 2d 可以看出,光线的投影造成了蓝色球体的阴影,这些阴影在颜色分割过程中一起被识别出来,通过基于人口统计的阈值分割(如图 2e 所示),将这些阴影确定为低人口区域,从而得以去除.

2.3 形态学及形状识别

区域分割完成后,虽然实现了运动物体的聚类,但聚类的形状复杂且不规则. 为实现形状识别,需要对得到的聚类图形进行形态学运算.

首先沿垂直方向对图像采样,获得最外侧的采样点,然后连接这些采样点,连接的原则是轮廓上任意 3 个点的连线夹角要小于 180° ,且连接物体内部任意 2 个点的直线均落在物体内部. 连接完成后,就得到了一个具有连续边界的凸多边形,在对多边形中可能存在的孔洞进行填充后,即可进行物体形状的分析.

传统的形状识别方法需要精确标定的模板,所以必须完全知晓物体的先验模型,且模板相关性匹配十分耗时,因此不适用于物体模型未知且环境不确定的运动物体跟踪. 本文根据图像的统计信息对物体形状进行近似估计,然后将估计的结果与其他子特征集一起融合.

在进行识别判断前,首先要统计物体边界和内部的像素面积. 由于在形态学中已经将物体描述成凸多边形,因此物体的一个边界的像素一半落在物体内部,而另外一半落在物体外. 多边形的面积

$$A = N_a - [(N_b/2) + 1] \quad (3)$$

式中: N_a 为所有像素的数量; N_b 为边界像素的数量. 计算完面积后,通过比较物体面积 A_0 与最小外接矩形面积 A_r ,即可获得物体的矩形度

$$R = A_0/A_r \quad (4)$$

R 反映了一个物体对其最小外接矩形的充满程度. 对于矩形物体, R 取最大值 1;对于圆形物体, R 取 $\pi/4$;对于弯曲、纤细的物体, R 的取值更小. 本文以 R 作为形状识别子集的阈值 H_s .

3 粒子分析

获取了物体的颜色、形状及运动分割子特征集后,即可有效地分割出一般的运动物体,然而当场景中出现极其近似的干扰物体时,仍然可能无法做到精确的识别. 如图 3a 所示,2 个矩形的大小、颜色、形状均极其相似,要想利用视觉分辨出这 2 个物体,

是具有相当难度的. 为此, 本文提出一种模糊化粒子分析的方法来对多种特征进行融合.

对于具有高级智能的人类而言, 识别一个物体并不需要完备的描述, 仅利用一些模糊化的局部信息就能做到有效的识别, 如拥有“是恒星”、“具有一定体积”、“距离水星最近”3 个特征, 即可判别该物体是“太阳”, 但是只拥有任何一个单一的特征, 则无法做到准确判别. 与此类似, 我们将特征子空间的每一项也定义为模糊的, 例如, 颜色的阈值可以是“黄色”、“近似橙色”、“近似橙红”, 形状的阈值可以是“矩形”、“近似圆形”、“具有圆角的矩形”. 如果选择的每一项子特征集的模糊化程度越高, 这种识别的智能化程度也就越高, 即越接近高级智能识别, 但这需要更多的子特征集来提供支持. 这样, 虽然特征子空间的某些项可能偏离特征空间, 但利用多特征间冗余信息的相互补充就可以获得期望的输出结果.

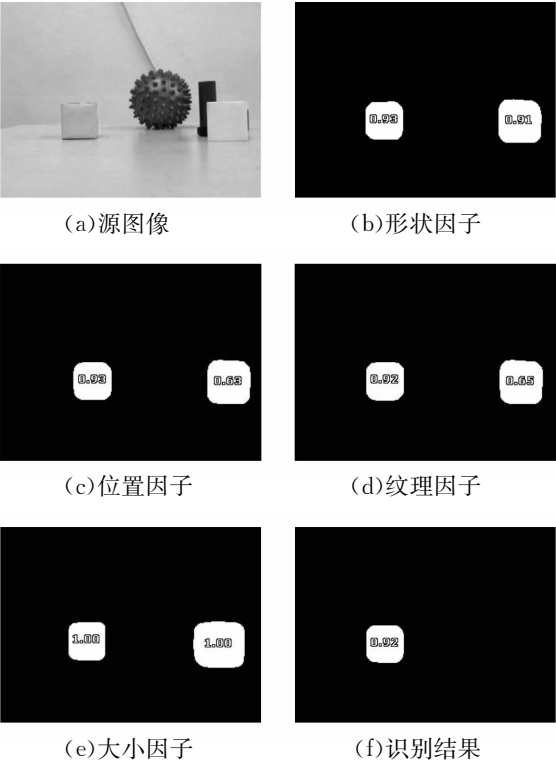


图 3 图像多特征因子

在获得了足够的特征子集之后, 利用已有的先验知识对各特征进行加权处理

$$\Sigma = m_1 h_1 \cap m_2 h_2 \cap \cdots \cap m_n h_n \quad (5)$$

式中: Σ 为最终的输出结果; $h_1, \dots, h_n$ 表示特征子空间中的特征子集, 分别对应颜色、形状、纹理等; $m_1, \dots, m_n$ 为对应的子特征阈值, $m_1 \cup m_2 \cup \cdots \cup m_n \in \eta$, η 为先验知识空间. 由于特征子空间中特征子集相

交, 而特征阈值隶属于先验知识空间, 所以 η 中的每一项并不需要完备的信息, 即 $m_1, \dots, m_n$ 是一种模糊化的信息. 以图 3 的近似方形物体识别为例, 利用图 3b~3e 的多信息融合识别, 就可以在复杂且具有干扰的情况下做到准确的特征识别. 图中的数字表示每一项因子对应的阈值.

4 试验结果分析比较

本文的研究是基于哈尔滨工业大学中德空间机器人技术联合实验室自主研发的 4 自由度空间机械臂完成的. 图 4 为双机器人视觉跟踪系统, 右侧机器人为 RX60 工业机器人, 左侧为空间机械臂.

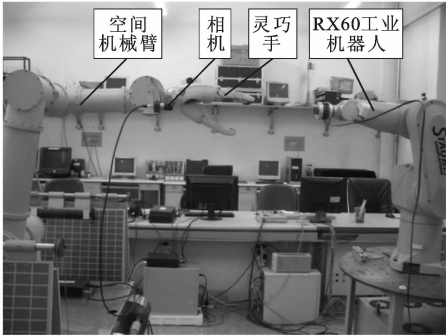


图 4 双机器人视觉跟踪系统

为分析不同光照与部分遮挡对目标识别的影响, 首先将本文方法与张正友^[6]提出的利用平面靶标提取角点特征方法进行比较. 图 5a 中的图像序列依次表示清晨、均匀灯光以及点光源照射下的原始图像, 图 6 为相应的彩色直方图. 对比图 5、图 6 可

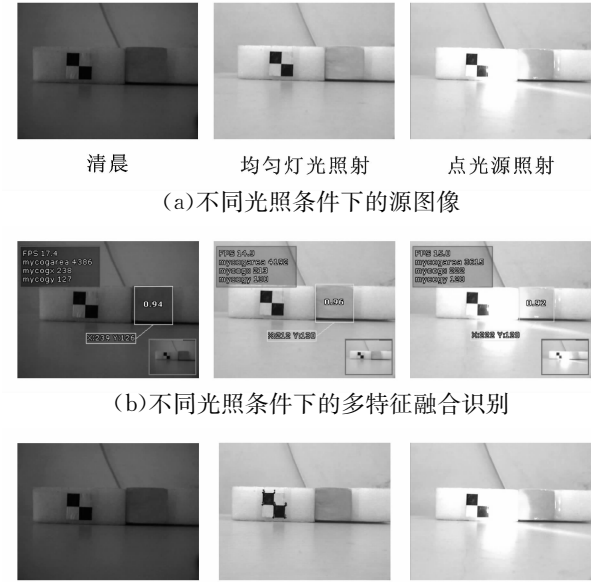


图 5 光照条件对识别的影响

以看出,清晨时刻图像整体偏暗,R、G、B 的像素在 80~160 灰度范围内均匀分布.当场景处于点光源强光照时,整个图像的亮度非常高,几乎所有的 R、G、B 的像素都集中在高亮的灰度区域.角点识别方法是首先利用一个固定阈值对图像进行二值化,再利用二值化后的图像进行形状遍历识别,由于选择的固定阈值明显与外界光线密切相关,因此只有在均匀光场条件下才能够获得满意的结果,而当外界光线发生变化时,则无法做到特征的有效识别.在图 5c 中,中间一幅图像表示在均匀光场下的有效识别,角点上的数字标示识别出的特征点,而在其他 2 种光照条件下,角点均未识别出.与之相比,本文的方法对于光线变化具有良好的适应能力,如图 5b 所示,黄色方框标示识别区域的包容盒.

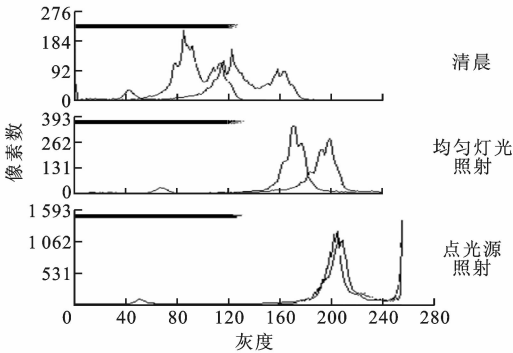


图 6 不同光照条件下的彩色直方图

图 7 表示在不同遮挡条件下采用本文方法进行的双机器人视觉跟踪试验,被跟踪物体为 RX60 机器人末端的蓝色部件(见图 7a).当 RX60 机器人以 20 mm/s 的速度运动时,在目标无遮挡以及部分遮挡的情况下,利用空间机械臂对目标进行了平面视觉跟踪.由图 7b 可以看出,采用本文提出的方法能够有效地克服部分遮挡对视觉跟踪的干扰.

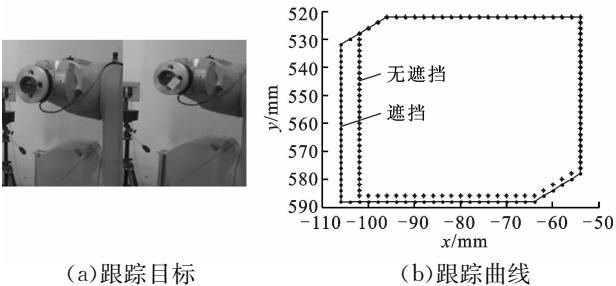


图 7 不同遮挡条件下的视觉跟踪

5 结论及展望

(1)本文提出了一种模糊化多视觉信息融合的

视觉跟踪策略:综合多个反映目标特征的视觉信息,通过对每个特征子集的模糊化加权融合处理,实现了特征在复杂的动态环境中的鲁棒提取和跟踪.相对于传统的局部特征识别方法,本文方法对周围环境噪声具有抑制性,在不均匀光场及部分遮挡的情况下仍可以得到较为满意的跟踪效果.

(2)针对运动物体分割阈值受噪声干扰的问题,提出了一种结合人口统计与区域生长的区域分割方法,实现了运动区域特征的稳定聚类.

(3)虽然机器人视觉跟踪的理论和实验研究已经取得了很大的进展,但其在工业领域、尤其是航空航天领域的应用仍十分有限.本文提出的多特征融合方法在提高特征识别的鲁棒性和智能化方面进行了有益的尝试,但是多特征的引入必然会带来特征选择及搜索方面的问题,因此如何选择性能最优的特征进行融合,如何缩小特征的搜索范围,都是需要进一步研究的课题.

参考文献:

[1] ODA M. ETS-VII: achievements, troubles and future [C]//Proc of the 6th International Symposium on Artificial Intelligence and Robotics & Automation in Space: i-SAIRAS. Quebec, Canada: Canadian Space Agency, 2001: 18-22.

[2] PAI C, TYAN H, LIANG Y, et al. Pedestrian detection and tracking at crossroads [J]. Pattern Recognition, 2004, 37(5): 1025-1034.

[3] 王麟琨,徐德,谭民. 机器人视觉伺服研究进展[J]. 机器人, 2004, 26 (3): 277-282.

WANG Linkun, XU De, TAN Min. Survey of research on robotic visual servoing [J]. Journal of Robot, 2004, 26(3): 277-282.

[4] JEPSON A, FLEET D, EI-MARAGHI T. Robust online appearance models for visual tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296-1311.

[5] 庄健,王孙安. 视觉伺服移动机器人中的图像处理研究[J]. 西安交通大学学报, 2002, 36(3): 287-290.

ZHUANG Jian, WANG Sun'an. Study of image processing of visual servoing mobile robot [J]. Journal of Xi'an Jiaotong University, 2002, 36(3): 287-290.

[6] ZHANG Zhengyou. A flexible new technique for camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (11): 1330-1334.

(编辑 葛赵青)