

# 双目主动视觉监测平台下的目标识别

孔令富<sup>1</sup>, 连秀梅<sup>1</sup>, 赵立强<sup>2</sup>

(1. 燕山大学信息科学与工程学院, 河北 秦皇岛 066004; 2. 河北科技师范学院, 河北 秦皇岛 066004)

**摘要:** 利用对目标旋转、尺度变化、视角变化等具有稳定性的尺度不变特征变换(SIFT)算法, 提出一种适用于基于圆轨的基线可调双目主动视觉监测平台的目标识别方法, 通过离线建立物体的多侧面 SIFT 特征点数据库, 将三维空间的目标转换为二维特征描述, 利用二维特征描述实现三维空间目标的识别, 以提高匹配识别效率。实验结果表明, 该方法能实时准确地识别目标。

**关键词:** 尺度不变特征变换; 目标识别; 二维特征描述; 双目主动视觉; 特征点数据库

## Object Identification in Binocular Active Visual Monitoring Platform

KONG Ling-fu<sup>1</sup>, LIAN Xiu-mei<sup>1</sup>, ZHAO Li-qiang<sup>2</sup>

(1. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China;

2. Hebei Normal University of Science & Technology, Qinhuangdao 066004, China)

**【Abstract】** Scale Invariant Features Transform(SIFT) algorithm is stable to rotating, scale changes and visual angle changes of the target, so this paper proposes an object identification method for a new visual monitoring platform—the baseline based on circle track adjustable binocular active vision platform. By establishing the off-line database of multifaceted SIFT features, it converts three-dimensional object into two-dimensional characterization, and uses it to realize object recognition three-dimensional space, so that the matching and recognition efficiency is improved. Experimental results show that the method can identify objects real time and accurately.

**【Key words】** Scale Invariant Feature Transform(SIFT); object recognition; two-dimensional characterization; binocular active vision; feature point database

DOI: 10.3969/j.issn.1000-3428.2011.14.047

### 1 概述

基于圆形导轨的双目主动视觉监测平台是由燕山大学智能并联机器人实验室的研究团队在孔令富<sup>[1]</sup>教授带领下设计的, 目的是提高并联机器人工作的主动认知能力及其智能性, 使其可以更好地实用化和商业化, 从而利用机器视觉技术有效地解决并联机器人的“盲”工作问题。在该平台中, 承载摄像机的装置有 3 个自由度, 其中, 2 个自由度用于调整摄像机的姿态; 另一个自由度控制摄像机在圆轨上滑动, 调整 2 台摄像机的位置。该视觉平台的基线不仅长度、位置、方向均可调, 而且在一环域内具有 360°全视角, 由于具有灵活主动的双目结构, 因此可以实现全空间、多视角监测, 具有视觉跟踪、避让能力, 对提高并联机器人运动准确性及智能性、推动并联机器人实用化和扩展其应用具有重要的理论意义和技术价值。

根据该主动视觉平台下摄像机的全视角特点, 在对目标进行识别和定位时, 2 台摄像机会从不同视角捕捉目标图像, 即需要对物体的三维侧面进行识别, 因此, 该平台对识别能力在尺度变化和视角变化等方面具有较高的要求。

针对该视觉平台, 本文给出了一种基于特征点匹配的识别方法<sup>[2]</sup>。它考虑的是像素点邻域的灰度变化。常用的特征点提取方法有 Harris 角点提取<sup>[3]</sup>、SUSAN 角点提取<sup>[4]</sup>等, 但这些角点都不具有尺度不变性, 而尺度不变特征变换(Scale Invariant Feature Transform, SIFT)<sup>[5]</sup>描述子对尺度变化、视角变化、光照变化和遮挡问题等具有稳定性, 符合该平台的灵活性。因此, 本文通过采用 SIFT 算法提取特征点, 建立特征点离线特征库, 实现了双目主动视觉监测平台下的多目标识别。

### 2 SIFT 特征点提取与匹配原理

#### 2.1 图像的多尺度表示

为了模拟人类在不同距离观察事物的过程, 提出了多尺度空间方法<sup>[6-7]</sup>。本文用高斯函数作为尺度空间的内核函数。一幅二维图像的尺度空间定义为:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

其中,  $I(x, y)$  是图像  $I$  的像素位置;  $\sigma$  是尺度因子;  $G(x, y, \sigma)$  是尺度可变高斯函数:

$$G(x, y, \sigma) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

为了在尺度空间有效检测到稳定的关键点, 提出了高斯差分尺度空间, 利用不同尺度的高斯差分核与图像卷积生成:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3)$$

#### 2.2 SIFT 特征向量的生成

SIFT 特征向量的生成步骤如下:

(1) 检测尺度空间极值点, 以初步确定关键点位置和所在尺度。通过与相邻的 26 个点比较得到, 此时的特征点称为候选特征点。

(2) 精确定位极值点。通过拟合三维二次函数精确确定关键点的位置和尺度(达到亚像素精度), 同时去除低对比度的

**基金项目:** 国家自然科学基金资助项目(60975062); 河北省教育厅自然科学研究计划基金资助项目(Z2009115)

**作者简介:** 孔令富(1957—), 男, 教授、博士、博士生导师, 主研方向: 智能控制, 机器人技术, 并行分布式系统; 连秀梅, 硕士研究生; 赵立强, 教授、博士

**收稿日期:** 2010-12-15 **E-mail:** lianxioumei@yahoo.com.cn

关键点和不稳定的边缘响应点(因为高斯差分算子会产生较强的边缘响应),以增强匹配稳定性,提高抗噪声能力。

(3)利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数,使算子具备旋转不变性,如式(4)、式(5)所示,其中,  $L$  所用的尺度为每个关键点各自所在的尺度:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{4}$$

$$\theta(x,y) = a \tan 2\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{5}$$

(4)生成关键点描述子。为了增强匹配的稳健性,文献[5]提出对每个关键点使用  $4 \times 4$  共 16 个种子点进行描述,这样,对于一个关键点就可以产生 128 个数据,即最终形成 128 维的 SIFT 特征向量。此时 SIFT 特征向量已经去除了尺度变化、旋转等几何变形因素的影响,再将特征向量进行归一化,则可以进一步去除光照变化的影响。

### 2.3 SIFT 特征点匹配

2 幅图像的 SIFT 特征向量生成后,采用关键点特征向量的欧氏距离作为 2 幅图像中关键点的相似性判定度量。取图像 1 中的某个关键点,并找出其与图像 2 中欧氏距离最近的前 2 个关键点。由于每个关键点生成 128 维的特征向量,因此  $n=128$ 。在这 2 个关键点中,如果最近的距离除以次近的距离小于某个比例阈值,则接受这一对匹配点。降低这个比例阈值, SIFT 匹配点数目会减少,但更加稳定。

## 3 匹配识别

本文以课本、冰红茶、茶叶盒、字典这 4 种物体作为待识别目标来举例说明识别过程。具体识别步骤如下:

### (1)特征库的建立

为了满足该平台在一环域内  $360^\circ$  全视角下的目标识别的需求,需要得到全视角下目标的 SIFT 特征点。由于 SIFT 算法本身的复杂性,提取 SIFT 特征点需要一定的时间,因此在识别目标之前对待识别目标单独地离线建立特征点数据库,即先将各个待识别目标的 SIFT 特征点提取并存储,由此建立 SIFT 特征点库。较直接与图像库进行匹配识别相比,本文方法在实时性方面具有很大的优越性。

特征库的建立过程如下:

1)以不具有 SIFT 特征点的简单背景为背景,单独地对每个待识别目标从不同视角采集图像,然后分别进行离线的 SIFT 特征点提取,最后生成特征向量,并将这些特征向量以数据文件的形式存储。

2)在某个目标不同视角的图像集中,以其中一幅图像中目标的 SIFT 特征点特征向量存储的数据文件为基准(即基文件),依次将其他视角下得到的 SIFT 特征点特征向量的数据文件整合到这个基文件中。

整合的具体过程为:将除生成基文件之外的任一视角的目标数据文件的特征向量按照上述匹配原理依次与基文件中存储的特征向量进行匹配,此时,匹配的欧氏距离的比例阈值设为 1,这样就能将所有描述相同的点的 SIFT 特征点匹配起来,匹配成功的点说明该视角拍摄的目标部分已经有该 SIFT 特征点存在,属于 2 个视角的公共 SIFT 特征点,即公共部分,因此,将匹配成功的特征向量从待整合的数据文件中去掉,只将其余的特征向量追加到基文件中即可。以此类推,直到把其余不同角度得到的目标的不同 SIFT 特征向量都整合到基文件中。这样就完成了第 1 个待识别目标的 SIFT 特征点数据文件的建立。

接下来,用同样的方法得到其他 2 个目标的 SIFT 特征点特征向量的数据文件,保存待识别目标的 SIFT 特征点信息,该信息表示该目标。这样就建立了每个待识别目标的 SIFT 特征点数据库。

由于从各个角度拍摄的图像中可能相邻角度有一部分重合,因此整合特征点的数据文件一方面可以去掉冗余的 SIFT 特征点,从而大大提高识别的实时性,另一方面可以减少误匹配,从一定程度上提高识别的准确性。下面以茶叶盒特征库的建立过程为例,分别从 6 个不同角度对其拍摄图像,其特征库建立过程如图 1 所示,其中的数字代表特征点个数。

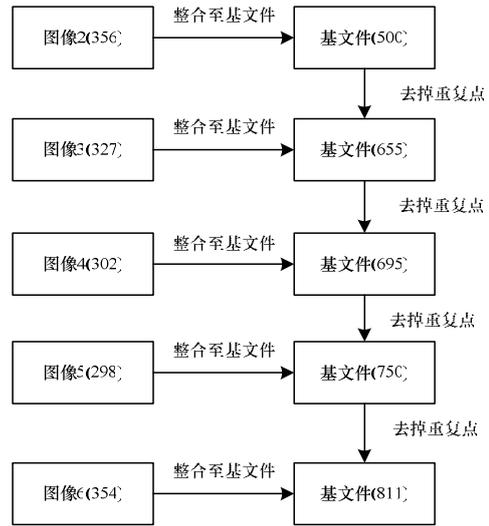


图 1 茶叶盒的特征库建立过程

### (2)待识别图像的特征提取

输入一幅待识别图像,其中可能含有数据库中已训练好的一个或者几个待识别的目标。该图像中每个目标的编号依次为:  $X_1, X_2, X_3, X_4$ ,由于 SIFT 特征点对光照、噪声和畸变等具有稳定性,因此该过程无需对图像进行预处理,可直接提取图像的 SIFT 特征点,并生成 SIFT 特征点特征向量的数据文件,以便下一步通过匹配实现目标识别。因为图像预处理也是一个较耗时的操作过程,所以该算法的应用在这一步起到了省时的作用,从而提高了识别效率,达到系统所要求的实时性。

### (3)目标的匹配识别

本文的多目标识别流程如图 2 所示。

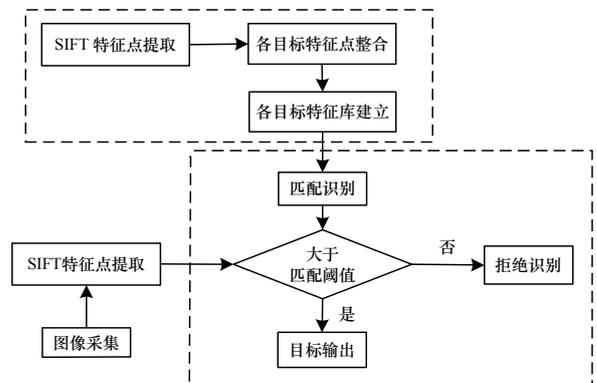


图 2 多目标识别流程

具体步骤如下:

1)对于步骤(2)中生成的待识别图像特征点数据文件中的特征向量与数据库中不同目标的特征点数据文件中的特征向

量, 采用关键点特征向量的欧氏距离作为 2 幅图像中关键点的相似性判定度量进行匹配, 每匹配一个数据库中目标的数据文件, 都记录下成功匹配的点对数目  $S(i)(i=1,2,3,4)$  以及此时匹配的数据库中目标特征的数据文件编号  $P(i)$ 。

2)特征库扫描完毕之后, 查看  $S(i)$  的大小, 若  $S(i)$  大于指定阈值, 说明输入图像中存在数据库中已训练的目标, 此时,  $S(i)$  对应的  $P(i)$  即为要识别的目标。识别效果可通过以下方法验证: 检查  $P(i)$  是否与  $X1、X2、X3、X4$  中的一个对应, 若对应, 表示识别成功, 若不对应且存在  $P(i)$ , 表示识别失败, 若不存在  $P(i)$ , 则拒绝识别。不同目标对匹配识别所取的阈值不同, 本文根据建立特征库和待识别目标的识别情况, 阈值取为 10。

#### 4 实验与结果分析

为了验证 SIFT 算法在该平台上识别的高效性和准确性, 分别选取了不同视角共 58 幅不同的图像进行识别, 并且这些图像中待识别的目标个数、种类以及观察视角都不一样。该目标识别实验用 Matlab 仿真实现。由于篇幅有限, 选取了其中 3 幅输入图像的识别结果, 如图 3~图 5 所示。从图 3 和图 5 可知, 由于摄像机高度的调整, 目标图像在尺度上发生了变化, 而摄像机绕圆轨滑动时拍摄视角的变化可从茶叶盒观测到, 匹配结果如表 1 所示。从实验结果可以看出, SIFT 算法能很好地适应目标的多视角和多尺度等变化, 并能通过匹配准确地识别图像中的目标。从实时性上来看, 本文采用了离线的 SIFT 特征点提取, 通过存储数据文件, 并整合每个目标在各视角下的 SIFT 特征点的特征向量, 建立了所有待识别目标特征点的特征向量数据库, 在很大程度上提高了实时性, 否则, 在匹配识别时不仅需要对待识别图像进行在线的特征提取, 还需要对图像数据库中的图像进行在线的特征提取, 这就很难满足该平台的实时性要求。



图3 课本、字典、红茶、茶叶盒的识别实验



图4 茶叶盒、字典、红茶的识别实验



图5 课本、红茶、字典的识别实验

表1 匹配识别结果

数据库 中目标	图 3		图 4		图 5	
	匹配 结果	识别 结果	匹配 结果	识别 结果	匹配 结果	识别 结果
课本	52	课本	1		51	课本
茶叶盒	60	茶叶盒	49	茶叶盒	0	
字典	98	字典	21	字典	89	字典
红茶	91	红茶	81	红茶	87	红茶

#### 5 结束语

本文对基于图像特征点提取的 SIFT 算法进行了研究并给出了一种基于该算法在新型双目立体视觉平台上进行识别的方法。通过全视角的特征提取建立特征向量数据库, 并将每个待识别目标的所有特征点(即从不同视角拍摄的所有特征点)进行整合, 这不仅满足了该平台全视角获取图像的需要, 且不必重复匹配特征点, 将三维空间的目标转换为二维特征描述, 通过二维特征描述实现三维空间目标的识别, 同时大幅提高了识别的效率。该方法还能在一幅图像中准确识别多个目标。

#### 参考文献

- [1] 张世辉, 孔令富. 一种新型 6-PUS 并联机构雕刻机[J]. 机器人, 2005, 27(4): 313-318.
- [2] 曹 健, 王武军, 韩 飞, 等. 基于局部特征的目标识别技术研究[J]. 计算机工程, 2010, 36(10): 203-205.
- [3] Harris C. Geometry from Visual Motion[M]//Active Vision. Blake A, Yuille A. [S. l.]: MIT Press, 1992: 263-284.
- [4] Smith S M, Brad Y M. SUSAN: A New Approach to Low Level Image Processing[J]. International Journal of Computer Vision, 1997, 23(1): 45-78.
- [5] Lowe D G. Distinctive Image Features from Scale Invariant Key Point[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [6] Mikolajczyk K. Detection of Local Features Invariant to Affine Transformations[D]. Avenue Félix Viallet, France: Institute National Polytechnique de Grenoble, 2002.
- [7] Witkin A P, Baudin B J. Uniqueness of the Gaussiankernel for Scale-space Filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(1): 26-33.

编辑 张 帆