

足球机器人视觉目标识别的 PCA-SIFT 算法

李彤斐, 杨马英*

(浙江工业大学 信息工程学院 浙江 杭州 310023)

摘要: 针对 Robocup 类人组比赛的特殊环境, 为了识别比赛双方机器人, 并克服光照变化、图像扭曲等问题给识别目标带来的干扰, 选用尺度不变特征变换(SIFT) 算法进行特征提取。由于 SIFT 算法的实时性差, 提出改进的 PCA-SIFT 目标识别算法, 将主成分分析法(PCA) 用于对经典 SIFT 算法中的 128 维描述子进行降维, 并用最近邻法进行特征点的匹配。通过实验发现, 经过 PCA 改进的描述子不仅提高了运算速度, 匹配度也保持在较好的状态。

关键词: Robocup 类人组比赛; 尺度不变特征变换; 主成分分析法; 最近邻方法

中图分类号: TN 911.73 **文献标志码:** A **文章编号:** 1671 - 7147(2013)05 - 0601 - 06

Improved PCA-SIFT Algorithm for Soccer Robots Visual Target Identification

LI Tong-fei, YANG Ma-ying*

(College of Information Engineering Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: This article aims to identify targets in the Robocup humanoid league competition environment considering the problems of illumination changes image distortions etc. In order to distinguish teammates from other players, we choose the scale invariant feature transform (SIFT) algorithm to extract the feature. Due to the heavy computational load of the SIFT algorithm, we proposed a PCA-SIFT target identification algorithm in which the principal component analysis method (PCA) is applied to the 128-dimensional descriptors of the classical SIFT algorithm to reduce the dimensionality, and the nearest neighbor method is used to match the feature points. Experimental results illustrated that the improved PCA descriptors can enhance computational efficiency with a high matching degree.

Key words: Robocup humanoid league, scale invariant feature transform, principal component analysis, the nearest neighbor method

Robocup 即机器人世界杯足球锦标赛, 是目前国际上最具影响的世界杯机器人足球赛^[1]。为了让一个机器人球队像人类一样进行足球比赛, 包含了很多领域技术的融合, 比如机械电子学、机器人学、传感器信息融合、智能控制、通信、计算机视觉、计算机图形学、人工智能等。

计算机视觉相当于足球机器人的眼睛, 为了让

机器人定位自身和其他目标并实现合作与对抗, 图像的特征提取以及目标识别是极为关键的一步。特征提取有基于边缘、角、区域和脊的方法。目前已经成熟的特征提取算法有: (1) Hough 变换^[2]; (2) Harris 角点检测^[3]; (3) SIFT 算法^[4]; (4) SURF 算法^[5]; (5) Mean Shift 算法^[6]等。Hough 变换于 1962 年由 Paul Hough 提出, 对圆和直线比较敏感;

收稿日期: 2013 - 06 - 29; 修订日期: 2013 - 08 - 13。

基金项目: 浙江省重点科技创新团队项目(2009R50014)。

作者简介: 李彤斐(1989—), 女, 浙江杭州人, 系统工程专业硕士研究生。

* 通信作者: 杨马英(1966—), 女, 浙江海宁人, 教授, 硕士生导师。主要从事机器人控制与过程控制等研究。

Email: myyang@zjut.edu.cn

Harris 角点检测在 1988 年由 Harris 等人提出,是基于模板的角点检测算法,具有旋转不变性。SIFT 算法由 D. G. Lowe 1999 年提出,其特征能对旋转、尺度缩放、仿射变换、视角变化、光照变化等图像变化因素保持一定的不变性,而对物体运动、遮挡、噪声等因素也保持较好的可匹配性,从而可以实现差异较大的两幅图像之间的特征匹配,可以说应用范围非常广泛,也比较全面。SURF 算法由 Herbert Bay 和 Luc Van Gool 等人在 2006 年提出,是 SIFT 算法的快速算法^[7]。Mean Shift 算法中的 Mean Shift 概念最早是由 Fukunaga 等人于 1975 年在一篇关于概率密度梯度函数的估计中提出来的,后由 Comaniciu 等人发展,在图像平滑和图像分割中 Mean Shift 都得到了很好的应用^[8]。

考虑到球场上待识别的物体不只是球、门柱,为了更好地帮助决策系统作出攻守策略判断,识别比赛双方机器人也是非常关键的环节,同时比赛场地存在光照条件变化、场外物体干扰等,移动目标的视频图像也会出现扭曲、模糊等情形,所以认为 SIFT 算法的全面性能满足类人机器人足球比赛图像处理和目标识别的要求。但是 SIFT 算法的实时性比较差,这会直接影响到机器人的决策。因此文中借鉴了 Yan Ke 和 Rahul Sukthankar 提出的 PCA-SIFT 方法^[9]。将主成分分析(PCA)算法用于对 SIFT 算法描述子进行降维,生成低维的 PCA-SIFT 描述子,并用最近邻法进行特征点匹配,实现机器人场地目标识别。

1 背景知识

1.1 尺度不变特征变换算法原理

尺度不变特征变换(SIFT)算法的计算过程如下:首先构造一个尺度空间,即图像金字塔,并对其差分计算;然后粗略地检测差分金字塔空间中的极值点;再去除一些边缘响应等干扰,以进行精确定位;最后计算每个极值点的梯度值和梯度方向,形成描述子。具体各步骤的描述如下:

1) 构造尺度空间 图像的尺度空间 $L(x, y, \sigma)$ 定义为一个变化尺度的高斯函数 $G(x, y, \sigma)$ 与原图像 $I(x, y)$ 的卷积

$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y)$$

其中

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m/2)^2 + (y-n/2)^2}{2\sigma^2}}$$

m, n 表示高斯模板的维度。

为了高效地检测极值点, Lowe 用高斯差分算子

代替拉普拉斯算子进行极值检测:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times$$

$$I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma)$$

在实际计算中,使用高斯金字塔每组中相邻上下两层图像相减,得到高斯差分图像(DoG),如图 1 所示。

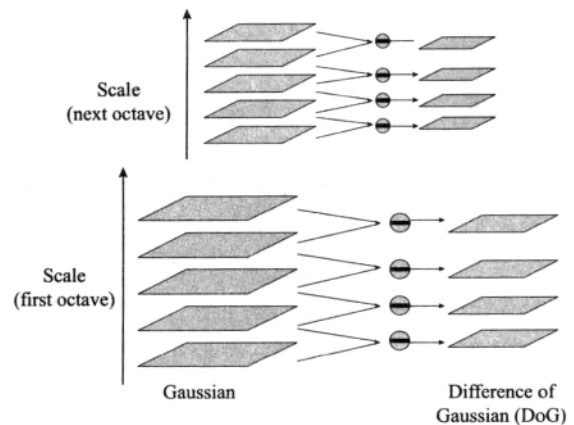


图1 高斯图像金字塔到差分高斯图像金字塔的变换
Fig.1 Transformation from Gaussian image pyramid to differential Gaussian image pyramid

2) 检测尺度空间极值点 这里指的是初步确定的关键点,即差分高斯金字塔内的局部极值点。如图 2 所示,共要比较 26 个相邻点。

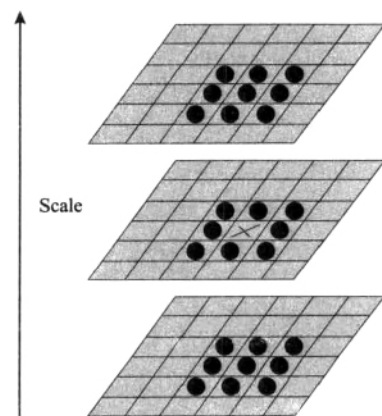


图2 DoG 空间极值检测
Fig.2 Space extremum detection of DoG

3) 精确定位极值点 由于 DoG 值对噪声和边缘较敏感,因此,在上面 DoG 尺度空间中检测到的局部极值点还要进行进一步的检验才能精确定位。

为了提高关键点的稳定性,需要对 DoG 空间函数进行拟合:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X$$

4) 为每个关键点指定方向参数 通过计算每个极值点的梯度值来确定方向。像素点的梯度表示:

$$\text{grad}I(x, y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$$

梯度幅值:

$$m(x, y) =$$

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

梯度方向:

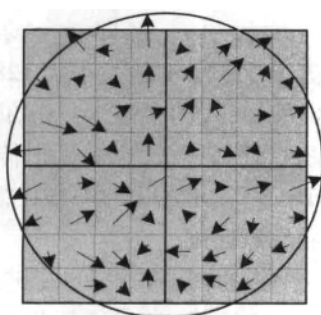


图3 关键点方向

Fig.3 Direction histogram of key points

5) 关键点描述子生成 首先,确定计算描述子所需的图像区域。其次,将坐标轴旋转为关键点的方向,以确保旋转不变性。再将领域内的采样点分配到8个方向上,计算其权值。插值计算每个种子点8个方向的梯度。如上统计的 $4 \times 4 \times 8 = 128$ 个梯度信息即为该关键点的特征向量。

1.2 PCA 方法概述

PCA 是对数据空间降维进行多元统计分析的方法。它计算主成分的目的是将高维数据投影到较低维空间,一般步骤如图4所示。



图4 PCA 方法降维的一般步骤

Fig.4 PCA-based dimensionality reduction steps

2 改进的 SIFT 算法

因为传统的 SIFT 算法产生的关键点的描述子包含 128 维的梯度信息,特征匹配时计算量相当庞大,直接影响到算法实时性。文中提出改进的 PCA-SIFT 目标识别算法,使用 PCA 方法将 128 维描述子降维,以提高匹配速度。

改进的 SIFT 算法主要工作分为 3 部分,分别如

$$\theta(x, y) = \tan^{-1} \left[\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right]$$

完成关键点的梯度计算后,使用直方图统计领域内像素的梯度和方向。梯度直方图将 $0 \sim 360^\circ$ 的方向范围分为 36 个柱,其中每柱 10° 。如图3所示,直方图的峰值方向代表了关键点的主方向。

以下各小节描述。

2.1 生成 SIFT 特征点描述子

改进的 SIFT 算法中,首先要生成 128 维的 SIFT 特征点描述子,这里用的是传统 SIFT 算法:分别预处理原图像与待匹配的图像,若不为灰度图像就转为灰度图像;然后构造尺度空间,并生成差分图像金字塔;检测尺度空间极值点,为每个极值点确定方向参数;最后生成 128 维的 SIFT 特征点描述子。特征向量形成后,为了去除光照变化的影响,需要对它们进行归一化处理。

其中,检测以及精确定位极值点时,为了找到这些点,每一个像素点都要与它的相邻点比较,判断该点是否为极值点。中间的检测点和它同尺度的 8 个相邻点与上下相邻尺度对应的 9×2 个点,共 26 个点比较,以确保在尺度空间和二维图像空间都检测到极值点。为了提高关键点的稳定性,需要对 DOG 空间函数进行拟合:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X$$

其极值点

$$\bar{X} = (x, y, \sigma)^T$$

在计算过程中,分别对图像的行、列及尺度进行修正,去除那些对比度较低的不稳定极值点。

对于 DoG 产生的边缘响应的点,用 Hessian 矩阵计算主曲率:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

DoG 的主曲率和 H 的特征值成正比,为了避免直接

计算这些特征值,令特征值为 α, β ($\alpha > \beta$) 则

$$\frac{\alpha}{\beta} = r > 1$$

$$\text{Tr}(H) = D_{xx} + D_{yy}$$

$$\text{Det}(H) = D_{xx} \times D_{yy} - D_{xy} \times D_{xy}$$

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(r+1)^2}{r}$$

时,将关键点保留,反之剔除。

2.2 PCA 对 SIFT 特征描述子降维

改进的 SIFT 算法第 2 步是用 PCA 方法对 SIFT 特征描述子进行降维。具体步骤如下:

1) 输入两幅待匹配图像中所有关键点(设为 n 个)的 128 维 SIFT 特征描述符,将输入的这 n 个特征描述符作为样本,写出样本矩阵为 $[x_1, x_2, \dots, x_n]^T$,其中 x_i 表示第 i 个特征点的 128 维特征向量。

2) 计算 n 个样本的平均特征向量:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

3) 计算所有样本点的特征向量与平均向量的差,得到差值向量 $d_i = x_i - \bar{x}$, $i = 1, 2, \dots, n$ 。

4) 构建协方差矩阵

$$C = \frac{1}{n} \sum_{i=1}^n d_i d_i^T = \frac{1}{n} Q Q^T$$

其中 $Q = [d_1, d_2, \dots, d_n]$ 。

5) 求协方差矩阵的 128 个特征值 λ_i 和 128 个特征向量 e_i 。

6) 将求出的 128 个特征值按从小到大的顺序进行排列 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{128}$ 和对应的特征向量 $[e_1, e_2, \dots, e_{128}]$ 。

7) 选取对应 t 个最大特征值的特征向量作为主成分方向。

8) 构造一个 $128 \times t$ 的矩阵 A , 它的列由 t 个特征向量组成。

9) 把原始的 128 维描述符根据下式投影到所计算出的 n 维子空间 M 中,就可以得到 PCA-SIFT 的描述符 y_1, y_2, \dots, y_n 了,即 $y_i = x_i \times A$ 。

实验中,矩阵 A 的大小为 $128 \times t$, x_i 的大小为 1×128 ,即 $x_i \times A$ 的大小为 $1 \times t$,即每一个 y_i 就是一个 t 维的特征描述子,也就是把传统的 128 维 SIFT 描述子降成了 t 维的 PCA-SIFT 描述子。

2.3 特征点匹配

在提取完特征点之后,改进的 SIFT 算法第 3 步,即特征点匹配。这里选用最近邻的方法进行匹配。对于待匹配图像中的一个特征点,把它逐一与已知图像中的特征点计算欧氏距离,找出距离最近与次近的点,当最近点与次近点的比值小于一个阈值时,认为这两点是匹配的。计算时运用以下公式:

$$\frac{D_1}{D_2} < r$$

其中 D_1 为最近点距离, D_2 为次近点距离。

3 实验及结果

实验需要在机器人足球比赛场地抓取一些图像(像素为 500×836)。如图 5 所示,左侧全部为原图,右侧分别为扭曲后、光照变暗以及用于识别比赛双方机器人的图像。

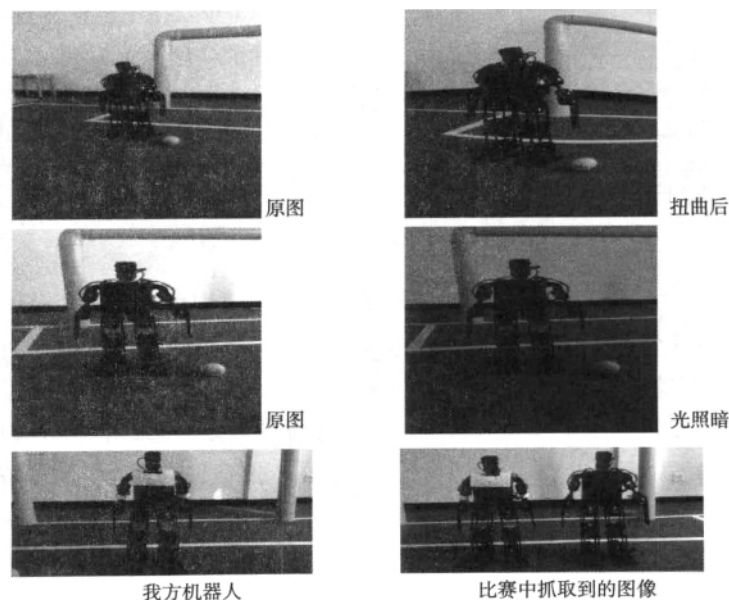


图 5 待处理的 3 对图像

Fig. 5 Three pairs of images to be processed

实验配备的计算机为 Mobile AMD Sempron(tm) Processor 3200 + 处理器, 内存为 1.60 GHz, 960 MB。仿真平台为 Matlab 7.0.1。

针对上述存在图像扭曲、光照变化以及双方机器人的图像, 分别运用 SIFT 算法与 PCA-SIFT 算法提取特征点, 并用最近邻法进行匹配, 观察算法的计算效率与匹配结果。计算中 PCA-SIFT 的参数 t 取 10, 匹配算法的参数 r 取 0.6。

图 6 为图像扭曲后与原图的匹配效果。可以发现 (b) 图明显比 (a) 图的匹配对数少, 即经过改进

的 PCA-SIFT 算法有明显的降维作用, 但是 (b) 图正确匹配率不如 (a) 图高, 说明改进的 PCA-SIFT 算法在图像扭曲的情况下效果不是很好。

图 7 为双方机器人的识别效果。图中机器人分别与原图的我方机器人进行匹配(用线条连接两幅图像中认为匹配的点)。匹配度高的一方为我方机器人, 反之为对方机器人。可以发现, 图 7(a), (c) 显然比图 7(b), (d) 的匹配对数多, 此时认为图像中左侧机器人为我方机器人, 右侧机器人为对方机器人。

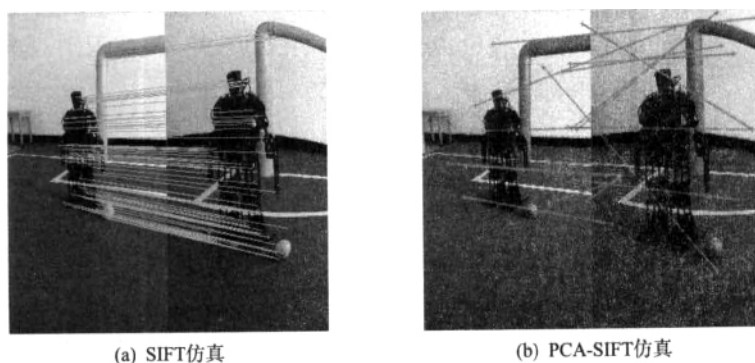
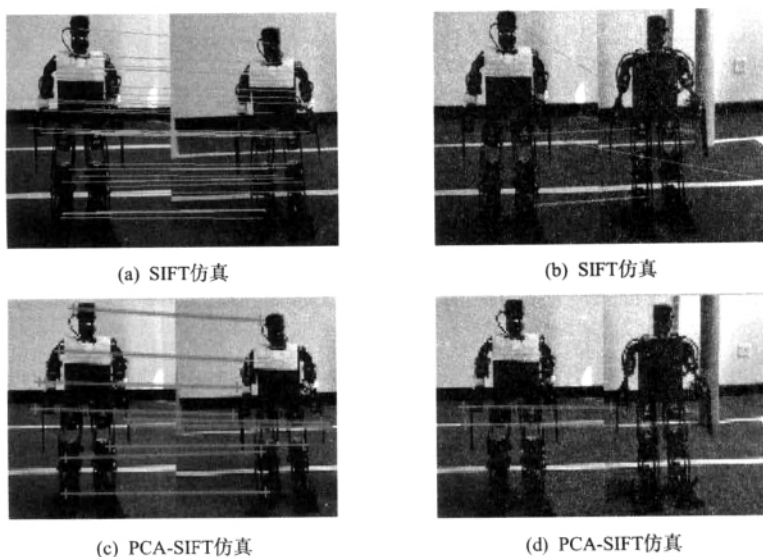


图 6 在图像扭曲的情况下的匹配结果

Fig. 6 Matching results under image distortion



(a), (c) 为我方机器人与抓图左侧机器人的匹配; (b), (d) 为我方机器人与抓图右侧机器人的匹配。

图 7 双方机器人的匹配识别结果

Fig. 7 Matching result of robots recognition

从表 1 统计的数据可以发现, PCA-SIFT 算法在图像扭曲与双方机器人识别方面于 SIFT 算法相比较, 匹配对数明显下降, 有很显著的降维作用; 而在光照变化的条件下降维不太明显。另外, 在双方机器人识别方面, 图 7b, d 比图 7a, c 的匹配对数少很多, 可以判断图 7b, d 为对方机器人。

从表 2 统计的数据中可以发现, 在 3 种条件下 PCA-SIFT 都能保持较好的匹配率, 特别是在光照变化和双方机器人识别方面占有优势。另外, 用匹配率的数据同样可以判断是否为我方机器人, 本实验数据显示对方机器人的匹配率为零, 则匹配率不为零的为我方机器人。

受限于计数器的精度,文中采集的图像均为 500 × 836 像素,而在 Robocup 比赛中机器人抓图得到的图像像素较低,所以相对而言统计出来的匹配时间较长。从表 3 可以发现,PCA-SIFT 算法普遍比 SIFT 算法快,特别是在光照变化的条件下省下的时间比较多。

表 1 不同条件下 2 种算法的匹配对数

Tab.1 Number of matching pairs

单位:对

不 同 算 法	图 像 扭 曲	光 照 变 化	双方机器 人识别 (图 6a c)	双方机器 人识别 (图 6b d)
SIFT	97	478	39	6
PCA-SIFT	18	207	9	2

表 2 正确匹配率

Tab.2 Correct matching rate

单位:对 / 对

不 同 算 法	图 像 扭 曲	光 照 变 化	双方机器 人识别 (图 6a c)	双方机器 人识别 (图 6b d)
SIFT	80/97	470/478	33/39	0/6
PCA-SIFT	10/18	199/207	8/9	0/2

表 3 不同条件下 2 种算法的匹配时间

Tab.3 Matching time under different conditions

单位: s

不 同 算 法	图 像 扭 曲	光 照 变 化	双方机器 人识别 (图 6a c)	双方机器 人识别 (图 6b d)
SIFT	18.5	17.7	3.9	4.2
PCA-SIFT	15.8	9.0	2.6	3.1

综合以上结果,可以得出结论: SIFT 算法用于机器人足球的场景来识别比较复杂的目标时有很好的识别率;经过改进的 SIFT 算法,即 PCA-SIFT 算法较 SIFT 算法有降维加速计算的效果,同时能保持较高的匹配率,且在识别双方机器人方面有较强的优势。

4 结 语

文中提出改进的 PCA-SIFT 目标识别算法,将经典 SIFT 算法与 PCA 方法相结合,获取低维的 PCA-SIFT 描述子,之后用最近邻方法完成特征点匹配。该方法与传统 SIFT 算法在多种场景条件下作了对比,实验证明在机器人足球比赛场景下经过改进的 PCA-SIFT 目标识别算法既保持了匹配率,又缩短了匹配时间,是一种机器人足球比赛适用的提取方法。

参考文献(References):

[1] Vadakkepat P, Sin N B, Goswami D. Soccer playing humanoid robots: processing architecture, gait generation and vision system [J]. Robotics and Autonomous Systems, 2009, 57: 776-785.

[2] 梁霄颖, 李永新, 张杰, 等. Hough 变换在足球机器人视觉处理中的应用 [J]. 机械设计与制造, 2008(1): 167-169.
LIANG Xiao-ying, LI Yong-xin, ZHANG Jie, et al. Applications of Hough transform in soccer robot vision processing [J]. Mechanical Design and Manufacture, 2008(1): 167-169. (in Chinese)

[3] Harris C, Stephens M. A combined corner and edge detector [C]//Manchester: Proceedings of the 4th Alvey Vision Conference. Manchester, UK: The University of Sheffield Printing Unit, 1988: 147-151.

[4] Lowe D G. Distinctive image features from scale-invariant key-points [J]. International Journal of Computer Vision, 2004, 17(1): 43-76.

[5] Bay H, Gool V. SURF: speeded up robust features [C]//European Conference on Computer Vision. Graz, Austria: Springer-Verlag, 2006, 1: 404-417.

[6] Comaniciu D, Meer P. Mean shift: a robot approach toward feature space analysis [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002(5): 603-619.

[7] Khan N Y, McCane B, Wyvill G. SIFT and SURF performance evaluation against various image deformations on benchmark dataset [C]//International Conference on Digital Image Computing: Techniques and Applications. Queensland, Australia: IEEE Computer Society, 2011, 9: 501-506.

[8] ZHOU H Y, YUAN Y, SHI C. Object tracking using SIFT features and mean shift [J]. Computer Vision and Image Understanding, 2009, 113: 345-352.

[9] Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors [C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2004.

(责任编辑: 杨 勇)