文章编号 2095-1531(2017)06-0719-07

基于改进的加速鲁棒特征的目标识别

龙思源[1,2],张 葆[1*],宋 策[1],孙保基[1,2]
(1. 中国科学院 长春光学精密机械与物理研究所 航测部重点实验室,吉林 长春 130033;
2. 中国科学院大学, 北京 100049)

摘要:为了提高加速鲁棒特征(SURF)算法的实时性和准确性,本文提出了一种结合 AGAST 角点检测和改进的 SURF 特征描绘算法。首先利用 AGAST 角点检测模板检测特征点,再使用增加对角信息的哈尔小波响应来生成特征点的描述子,之后利用特征袋对产生的描述子进行编码并生成新的特征向量,最后利用支持向量机(SVM)对特征向量进行分类,完成识别。本文以 SIFT 和 SURF 算法为对照,分别进行不同视角、光照和尺度的识别实验。实验结果表明,本文算法的平均识别率为 98.0%、96.9%、97.1%,平均时间分别为 66.1 ms、79.3 ms、41.0 ms,在识别率上较优于 SURF 算法,所耗时间约是 SURF 算法的 1/3。

关 键 词:图像处理;目标识别;加速鲁棒特征;AGAST 角点检测
中图分类号:TP391.4 **文献标识码**:A doi:10.3788/CO.20171006.0719

Object detection based on improved speeded-up robust features

LONG Si-yuan[1,2], ZHANG Bao[1*], SONG Ce[1], SUN Bao-ji[1,2]
(1. *Key Laboratory of Airborne Optical Imaging and Measurement, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun* 130033, *China*;
2. *University of Chinese Academy of Sciences, Beijing* 100049, *China*)
* *Corresponding author, E-mail: alfred_lsy@*163. *com*

Abstract: To improve the real-time performance and the accuracy of the SURF algorithm, an algorithm combined with AGAST corner detector and improved SURF feature descriptor is proposed. Firstly, feature points are detected by using AGAST corner detection template. Secondly, the Haar wavelet response with increased diagonal information is used to generate descriptor of feature points. Then, the generated descriptor is encoded by the feature bag and a new feature vector is generated. At last, the classification is fulfilled by Support Vector Machine(SVM). Finally, SVM is used to classify the feature vectors to complete the detection. Detection experiments for different view-points, illumination and scales are conducted respectively using SIFT and SURF algorithm as a control. The results show that the average detection rate of this algorithm is 98.0%, 96.9% and 97.1%, and the average time is 66.1 ms, 79.3 ms and 41.0 ms, respectively, which is better than that of SURF algorithm, and the time consumption is about 1/3 of the SURF algorithm.

Key words: image processing; object detection; Speeded-Up Robust Features; AGAST corner detection

收稿日期:2017-05-11;**修订日期**:2017-07-13
基金项目:中国科学院长春光学精密机械与物理研究所重大创新资助项目(No. Y3CX1SS14C)
Supported by Major Innovation Project of CIOMP, CAS, China(No. Y3CX1SS14C)

1 引 言

近年来,目标自动识别已经成为图像处理领域的热门方向,其核心是利用计算机视觉技术,对图像或视频中的内容进行识别和描述。它在人机交互、武器自动跟踪、智能视频监控等领域有着很好的应用前景和潜在的经济价值。然而,在图像或视频的采集过程中,经常会受到环境影响,例如光照变化、视角变化、目标尺度变化等,使得待检测目标发生光照、尺度和形状上的变化,从而导致识别失败[1]。

普遍来说,目标的自动识别主要分为两个步骤:(1)目标图像的特征向量提取;(2)训练分类器。由于局部特征往往都具有亮度、旋转和尺度不变性,因此,在目标识别和图像匹配中有着极为广泛的应用[2]。所选取的目标特征的稳定与否,决定了该自动识别方法的优劣。

2004年,加拿大学者David Lowe首次提出尺度不变特征变换(Scale Invariant Feature Transform,SIFT)算法[3]。该算法先利用高斯差分(Difference of Gaussian,DoG)算子找到图像的特征关键点,再在关键点的4×4邻域的每个小块中计算其8个方向的灰度梯度方向直方图。这样就在每个关键点上构建了一个128维的特征向量。SIFT特征算法具有较好的视角、尺度、亮度不变性,但是由于DoG算子的特性和特征向量的维数较高[4],该算法的实时性能较差。在SIFT的基础上,2006年Herbert Bay提出了加速鲁棒特征(Speeded-up Robust Feature,SURF)算法[5],其用快速海森矩阵为模板提取关键点,并利用Haar小波特征生成64维的特征向量,从而大大提高了算法的实时性[6]。2011年,Stefan Leutenegger等人提出的BRISK(Binary Robust In variant Scalable Keypoints,BRISK)算法和2012年Alexandre Alahi等人提出的FREAK(Fast Retina Keypoint,FREAK)算法都是基于二值化特征提取的算法。这两种算法大大提高了特征提取的时间效率。然而,由于使用的是二值特征,丢失了很多图片原有信息,它的识别效果远不如SIFT和SURF[7-8]。所以,目前世界上广泛研究和使用的算法仍是SURF算法。

为了进一步提高SURF算法的实时性和稳定性[9],本文以AGAST(Adaptive and Generic Accelerated Segment Test,AGAST)角点提取方法代替快速海森矩阵模板法,并对SURF特征描述子加以改进,此外还加入了特征袋编码,最后使用支持向量机(Support Vector Machine,SVM)进行特征分类,最终实现目标识别。实验证明,该算法在实时性和不同尺度下的识别率都要高于原来的SURF算法。

2 SURF算法简介

2.1 特征点检测

在SURF算法中,图像的显著度由该像素点的Gaussian-Hessian矩阵行列式的值决定。图像的不同尺度需要利用高斯核$G(t)$卷积得到,再与Hessian矩阵结合,这样就可以得到$\boldsymbol{x}(x,y)$,外尺度为σ的高斯海森矩阵,如式(1):

$$H(\boldsymbol{x},\sigma)=\begin{bmatrix}\dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\partial y}\\ \dfrac{\partial^2 f}{\partial x\partial y} & \dfrac{\partial^2 f}{\partial y^2}\end{bmatrix}[I(\boldsymbol{x},\sigma)*G(\boldsymbol{x},\sigma)]=I(\boldsymbol{x},\sigma)*\begin{bmatrix}L_{xx}(\boldsymbol{x},\sigma) & L_{xy}(\boldsymbol{x},\sigma)\\ L_{xy}(\boldsymbol{x},\sigma) & L_{yy}(\boldsymbol{x},\sigma)\end{bmatrix},\tag{1}$$

式中,与图像卷积的矩阵即为高斯海森矩阵。其中$L_{xx}(\boldsymbol{x},\sigma)$、$L_{xy}(\boldsymbol{x},\sigma)$、$L_{yy}(\boldsymbol{x},\sigma)$表示高斯函数的二阶微商。

为了进一步加快运算速度,SURF算法将$L_{xx}(\boldsymbol{x},\sigma)$、$L_{xy}(\boldsymbol{x},\sigma)$、$L_{yy}(\boldsymbol{x},\sigma)$3种模板进行近似处理,将原先的模板变换为$D_{xx}(\boldsymbol{x},\sigma)$、$D_{xy}(\boldsymbol{x},\sigma)$、$D_{yy}(\boldsymbol{x},\sigma)$,如图1所示。

加入补偿系数之后,其行列式如公式(2)所示:

$$\det(H)=D_{xx}D_{yy}-(0.9D_{xy})^2,\tag{2}$$

其中0.9是Bay给出的一个经验值。利用该行列式的值,同时结合非极大值抑制方法,就可以得到

一系列特征点的位置和尺度。

2.2　特征描述

检测出关键点之后,在生成描述向量之前需要确定主方向。SURF 算法中,主方向的选取方法是统计特征点邻域内的 Haar 小波响应。具体方法是:在尺度 s 下,在特征点附近半径为6s的圆内,统计圆心角为 60°的扇形内所有点的水平、垂直 Haar 小波响应并进行加权求和,然后以一定的角度旋转扇形重复上述操作,最终得到一系列的模值,模值最大的方向即为该关键点的主方向。

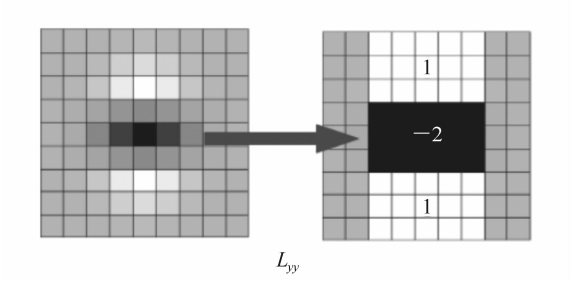


图 1　高斯海森矩阵中的 L_{yy} 积分模板的近似 D_{YY}

Fig. 1　Approximate D_{YY} of L_{yy} integral template in Gaussian-Hessian matrix

确定关键点的主方向之后,在特征点周围取一个 4 ×4 的正方形邻域,每一个正方形小邻域包含有 5 ×5 个像素,在每个小邻域中统计 25 个像素的水平方向和垂直方向的 Haar 小波响应,将它们作为特征描述(这里,水平和垂直均是相对于主方向)。这些 Haar 小波特征描述值为:水平方向值的和 Σd_x,水平方向绝对值的和 $\Sigma|d_x|$,垂直方向的和 Σd_y,垂直方向绝对值的和 $\Sigma|d_y|$。

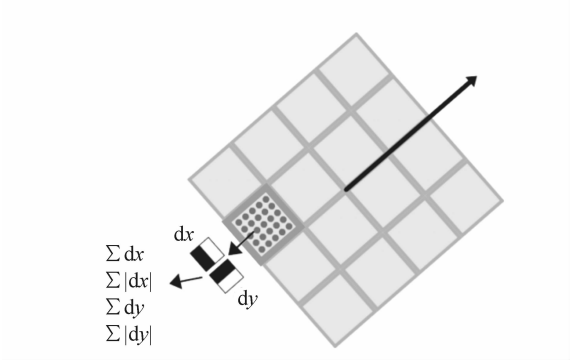


图 2　SURF 描述子的生成示意图

Fig. 2　Schematic diagram of SURF descriptors generating

这样,每个小区域就有 4 个值,而 4 ×4 个小邻域内共有 64 个值,即生成了该特征点的 64 维的描述向量。如图 2 所示。

SURF 算法在保证特征鲁棒性和灰度变化信息的同时,将 SIFT 的 128 维描述向量减少到 64 维,大大缩短了特征向量的生成时间,相对于 SIFT 算法,极大地提高了实时性。

3　结合 AGAST 的改进 SURF 算法

3.1　特征点检测

SURF 算法在特征点检测中使用近似的快速海森矩阵模板进行检测,虽然速度比 SIFT 的 Harris 模板有了显著的提高,但是实时性并不是很好。在尺度不变要求下,AGAST 算法有着更高的计算效率和实时性[11]。

对于 AGAST 算法,为了较好实现尺度不变要求,需构建尺度空间金字塔。它由 n 个 octave 层 c_i 和 n 个中间层 d_i 组成($i=0,1,\cdots,n-1$)。c_i 是通过对原始图像 c_0 的依次半采样得到的,而中间层 d_i 位于 c_i 和 c_{i+1} 之间,第一个内层 d_0 是对原始图像 c_0 的 1.5 倍下采样得到的,其他内层是依次对 d_0 半采样得到的。

完成尺度空间金字塔构建后,可以进行关键点的检测。用 FAST9-16 模板分别作用于每个层和每个内层,用于找到潜在的感兴趣区域。使用该模板的前提是,在周围 16 个像素的圆周上至少

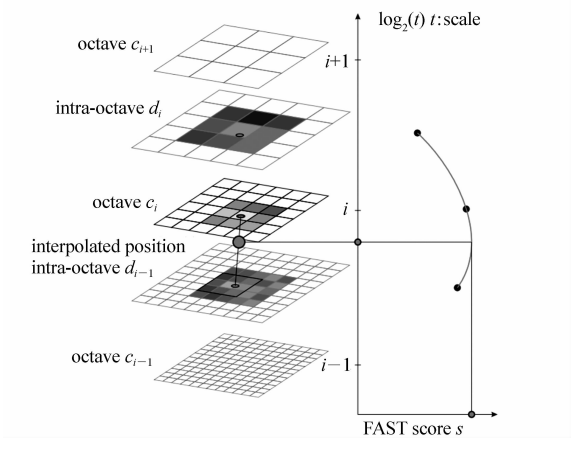


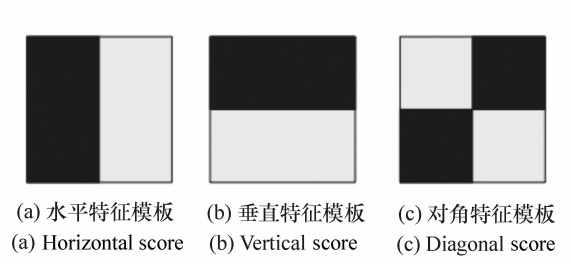
图 3　AGAST 在不同尺度下的层和内层

Fig. 3　Octaves and intra-octaves of AGAST algorithm in different scales

需要有9个连续的像素点比圆心像素充分亮或者充分暗。之后进行非极大值抑制,非极大值抑制:(1)该点的FAST模板积分 s 须大于同一层周围邻域中8个点的积分 s;(2)该点积分 s 必须大于相邻层同一点的积分。满足以上两个条件的点才能被检测为关键点。经实验证明利用该方法检测出的关键点,有着较好的尺度不变性,同时,较SURF具有更好的实时性,大大缩短了关键点的检测时间[12]。

3.2 特征描述

在主方向确定后,SURF算法利用Haar小波特征的水平和垂直模板对正方形小邻域进行卷积和加权求和,构造了64维向量对关键点进行描述。但是,仅仅采用水平方向和垂直方向这两个方向的Haar小波响应统计量过于简单。为了进一步提高识别准确率,本文加入了对角特征的采样模板[13],如图4所示。



(a) 水平特征模板 (a) Horizontal score (b) 垂直特征模板 (b) Vertical score (c) 对角特征模板 (c) Diagonal score

图4 Haar特征模板

Fig. 4 Scores of Haar feature

图4(c)即为对角Haar小波响应的模板,黑色块的值为-1,白色块的值为1,采用与SURF算法类似的方法进行加权求和,在SURF向量中加入对角向量的和值以及对角向量的绝对值的和值。这样改进后的SURF描述子就变为(Σd_x, $\Sigma|d_x|$, Σd_y, $\Sigma|d_y|$, Σd_d, $\Sigma|d_d|$,描述关键点的特征向量相应地变为96维。虽然特征向量的维数有一定提高,但是得到的描述子包含更多的图像信息,使得图像的区分度更高。

3.3 特征袋编码

为了进一步提高算法的鲁棒性,减少分类器的学习和分类时间,需对大量的特征向量进行特征袋(Bag of Features,BoF)编码。BoF的原理是利用已提取的全部局部特征的统计信息,对特征向量重新编码,然后生成新的描述向量[14]。对特征进行BoF编码不仅不会丢失原有的图像信息,而且会进一步提高算法的实时性。

具体过程是,首先,对所有训练样本中的局部特征向量进行K-均值分类,把 k 个聚类中心 $\{\alpha_1, \cdots, \alpha_k\}$ 作为视觉词汇构成编码本(k 是人为选取);然后,找到与待检测图片样本中每一个局部特征向量最接近的视觉词汇 α_i,并以 α_i 代替该特征向量;再统计 $\{\alpha_1, \cdots, \alpha_k\}$ 在该图片样本中出现的频率直方图,以该频率直方图为基础,即可生成该图片新的描述向量。

虽然选取的编码本视觉词汇数量远大于特征向量维数,例如选取1 000个词汇数量远大于SURF的向量维数64,但是一个图像的所有特征向量经编码后就是一个1 000维的带有图像局部特征统计信息的向量。而如果直接使用SURF,每个特征点均有一个64维向量,对于一幅分辨率为384×288的图像,则会检测出150~180个特征点,那么待处理的数值就有上万个,不仅时间性能较差,而且不利于后面学习分类的数据使用。

3.4 特征分类

特征分类的方法有很多,如最近邻分类器、神经网络分类器、Adaboosting分类器、支持向量机等。

SVM是一种二类分类模型,基本模型定义为特征空间上间隔最大的线性分类器,其是基于机器学习的分类算法,学习的策略即算法的核心是间隔最大化,即先将样本通过一个变换核映射到更高维度的空间,然后在该空间中找到一个超平面将其分为两类。

由于SVM有着较好的精确性和较低的时间复杂度,而且根据文献,SVM对SIFT/SURF特征的分类效果较好,故本文选用SVM分类器对样本进行训练和分类[15]。

综上所述,整个识别算法分为两部分,首先进行样本的训练和BoF字典的生成,再利用SVM对特征向量进行识别分类,整个识别过程如图5所示。

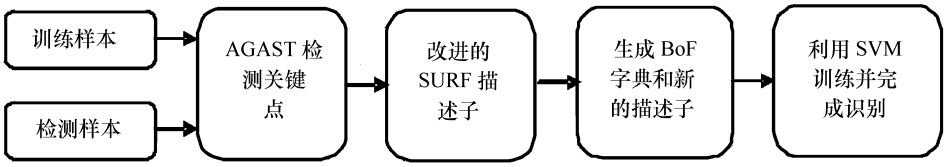


图 5　识别算法流程

Fig. 5　Flowchart of proposed detection algorithm

4　实验与结果

为了检验本文提出的目标自动识别算法的性能,在 Visual Studio 2012 的软件平台上,分别使用 SIFT 特征、SURF 特征和本文提出的特征算法对仿真图像进行目标识别实验,使用识别率、识别时间两种定量指标对算法性能进行评价。本文选取 ALOI 数据库中的图像进行目标识别,分别使用 3 种算法,对不同光照、视角和尺度的目标图像进行识别。

4.1　不同视角实验

首先,每隔 5°拍摄一幅物体的照片,每隔 45°视角选取一个正样本,共 7 个正样本,进行训练,其他的作为检测样本,如图 6 所示。

图 6　不同角度的样本目标

Fig. 6　Objects of different vew-points

分别使用 SIFT、SURF 和本文提出的算法对 102 个检测样本进行识别,正确数、正确率和平均识别时间见表 1。

由表 1 可以看出,本文算法在正确率上高于 SURF 算法,但是略低于 SIFT 算法。在时间性能上,本文算法远高于 SIFT 和 SURF 算法,平均所需时间接近 SURF 算法的 1/3;SIFT 算子由于特征描述复杂度更高,其平均耗时远大于 SURF 算子,是本文算法的近 9 倍。

表 1　不同视角下 3 种算法的性能比较

Tab. 1　Performances comparison of three algorithms under different view-points

	SIFT	SURF	本文算法
正确数	102	99	100
正确率/%	100	97.0	98.0
平均时间/ms	568.7	287.2	66.1

4.2　不同光照实验

本实验中选取光照均匀亮度适中的图像作为训练样本,然后改变光照的角度和强度,作为待检测样本,如图 7 所示。

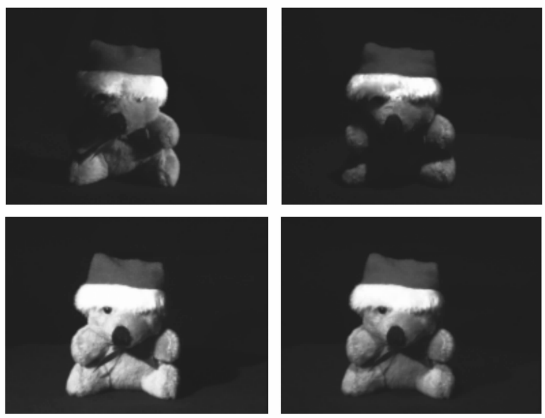


图 7　不同光照的样本目标

Fig. 7　Objects of different illumination

同样使用 3 种算法对 49 幅图像进行识别检测,得到结果见表 2。

表2 不同光照下3种算法的性能比较

Tab. 2 Performances comparison of three algorithms under different illumination

	SIFT	SURF	本文算法
正确数	38	47	47
正确率/%	77.6	95.9	95.9
平均时间/ms	590.6	272.9	79.3

从表2可以看出,在光照变化的条件下,SIFT算子的识别率大幅降低,而本文算法和SURF算法仍然保持较高的识别率,同时,相对于SURF算法,时间性能仍可大幅提高。

4.3 不同尺度实验

下面,将视角和尺度相结合进行检测。训练样本均选取不同视角实验中的训练样本,对不同视角实验中的待检测样本进行下采样,将分辨率由384×288降为192×144,将这些下采样图片作为待检测样本,然后再使用3种算法对102幅图像进行识别检测,结果如表3所示。

表3 不同尺度下3种算法的性能比较

Tab. 3 Performances comparison of three algorithms under different scales

	SIFT	SURF	本文算法
正确数	98	96	98
正确率/%	96.1	94.1	96.1
平均时间/ms	294.1	146.2	41.0

在不同视角条件下,由于图像被降采样,可能会丢失一些特征,导致同种算法识别率有一定程度降低。但是横向比较可发现,本文算法在保持了SIFT算子的高识别率下,大幅缩短了所需处理时间。

由上述3个对不同视角、不同光照、不同尺度的目标进行的识别实验结果可以看出,本文的改进SURF算法,由于在描述子中加入了新的特征,识别率相对于原SURF算法略有提高,但在尺度和视角的鲁棒性上比SIFT算法略低。本文提出的算法的最大优势是所需的平均时间相对于SURF算法缩短了近60%,是SIFT算法所需时间的近1/9,算法的实时性有了极大地提高。

5 结 论

本文在SURF算法的基础上,对其特征点的检测和特征点的描述进行了改进;使用速度更快的AGAST特征点检测算法,大大提高了算法的实时性;而在特征点的描述中,在Haar小波特征的垂直和水平特征基础上增加了对角特征以提高描述子的信息量、保证识别算法的鲁棒性和准确率。然后再配合特征袋编码算法和支持向量机分类器完成识别。对ALOI数据库图像的实验结果显示,在不同视角、光照、尺度的情况下,本文算法的识别率均能保持在95%以上,总体上优于SIFT和SURF算法,而其平均时间约为SURF算法的1/3,SIFT算法处理时间的1/9,实时性有了极大地提高。本文算法相对于SURF算法有很好的实时性,同时对于视角和尺度的变化也保持了较高的鲁棒性,在实际工程应用中具有重要意义。

参考文献:

[1] 贾平,徐宁,张叶. 基于局部特征提取的目标自动识别[J]. 光学 精密工程,2013,21(7):1898-1905.
JIA P,XU N,ZHANG Y. Automatic target recognition based on local feature extraction[J]. *Opt. Precision Eng.*,2013,21(7):1898-1905. (in Chinese)

[2] 赵春阳,赵怀慈. 多模态鲁棒的局部特征描述符[J]. 光学 精密工程,2015,23(5):1474-1483.
ZHAO CH Y,ZHAO H C. Multimodality robust local feature descriptors[J]. *Opt. Precision Eng.*,2015,23(5):1474-1483. (in Chinse)

[3] LOWE D. Distinctive image features from scale-invariant keypoints[J]. *International J. Computer Vision*,2004,60(2):91-110.

[4] HARRIS C,STEPHENS M. A combined corner and edge detector[J]. *Alvey Vision Conference*,1988(3):147-151.

[5] BAY H,TUYTELAARS T,VAN GOOL L. SURF:speeded-up robust features[J]. *Computer Vision & Image Understand-*

ing,2006,110(3):404-417.

[6] MIKOLAJCZYK K,SCHMID C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2005,27(10):1615-1630.

[7] 王飞宇,邸男,贾平.结合尺度空间 FAST 角点检测器和 SURF 描绘器的图像特征[J].液晶与显示,2014,29(4):598-604.
WANG F Y,DI N,JIA P. Image features using scale-space FAST corner detector and SURF descriptor[J]. *Chinese J. Liquid Crystals and Displays*,2014,29(4):598-604. (in Chinese)

[8] 成培瑞,王建立,王斌,等.基于多尺度区域对比的显著目标识别[J].中国光学,2016,9(1):97-105.
CHENG P R,WANG J L,WANG B,*et al*.. Salient object detection based on multi-scale region contrast[J]. *Chinese Optics*,2016,9(1):97-105. (in Chinese)

[9] 武治国,李桂菊.动态目标识别中的实时复杂巡航场景运动检测[J].液晶与显示,2014,29(5):844-849.
WU ZH G,LI G J. Real-time complex cruise scene detection technology in target recognition[J]. *Chinese J. Liquid Crystals and Displays*,2014,29(5):844-849. (in Chinese)

[10] MAIR E,HAGER G,BURSCHKA D,*et al*.. Adaptive and generic corner detection based on the accelerated segment test [J]. *Computer Vision-ECCV*,2010,2010:183-196.

[11] 王丽.融合底层和中层字典特征的行人重识别[J].中国光学,2016,9(5):540-546.
WANG L. Pedestrian re-identification based on fusing low-level and mid-level features[J]. *Chinese Optics*,2016,9(5):540-546. (in Chinese)

[12] ZHANG H,HU Q. Fast image matching based-on improved SURF algorithm[C]. International Conference on Electronics,Communications and Automation Control,IEEE,2011:1460-1463.

[13] ROSTEN E,PORTER R B,DRUMMOND T,*et al*.. Faster and better:a machine learning approach to corner detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2010,32(1):105-119.

[14] CSURKA G,DANCE C,FAN L,*et al*.. Visual categorization with bags of keypoints[C]. 8th European Conference on Computer Vision,Progue,Czech Republic,2004:1-22.

[15] DONGHOON K,ROZENN D. Face components detection using SURF descriptors and SVMs[C]. IEEE International Machine Vision and Image Processing Conference,IEEE,2008:51-56.

作者简介:

龙思源(1992—),男,湖北襄阳人,硕士研究生,主要从事图像处理和目标识别方面的研究。E-mail:Alfred_lsy@163.com

张　葆(1966—),男,吉林磐石人,博士,研究员,博士生导师,主要从事航空光电成像技术方面的研究。E-mail:cleresky@vip.sina.com