一种改进的具有仿射不变性的 SIFT 描述符

章明珠[1]，廖开阳[2]
（1.西安翻译学院，西安 710105；2.西安理工大学，西安 710048）

摘要：目的 针对 SIFT 描述符在具有仿射变换的图像中会产生大量误匹配问题，提出一种改进的 SIFT 描述符，以提高其鲁棒性和可区分性。**方法** 使用 Hessian Affine 检测子提取特征点，在特征描述阶段通过椭圆规一化成圆形的邻域来增强仿射不变性，运用仿射尺度空间增加尺度不变性，采用对数极坐标变换在圆形区域计算梯度直方图提高描述子的旋转不变性。**结果** 选择具有缩放和旋转、摄像机角度的变化、图像模糊、JPEG 图像压缩、光照变化的图像，重叠误差 ε_0 小于 0.5，其 recall & 1−precision 曲线图显示性能提高了 2%~10%。**结论** 与 SIFT, GLOH, PCA-SIFT 等描述符相比，文中改进的描述符更具有鲁棒性和可区分性。

关键词：局部特征；描述符；仿射不变；SIFT
中图分类号：TP391.41 **文献标识码：**A **文章编号：**1001-3563(2018)07-0194-05
DOI：10.19554/j.cnki.1001-3563.2018.07.035

An Improved Affine Invariant SIFT Descriptor

ZHANG Ming-zhu[1], *LIAO Kai-yang*[2]
(1.Xi'an Fanyi University, Xi'an 710105, China; 2.Xi'an University of Technology, Xi'an 710048, China)

ABSTRACT: The work aims to propose an improved SIFT descriptor to improve the robustness and distinguishability of SIFT descriptor which produces a large number of mismatches in an image with affine transformation. Hessian Affine detector was adopted to extract feature point. In the characterization phase, the elliptical was normalized into neighboring region to enhance the affine invariance. Scale invariance was improved by affine scale space. The gradient histogram in the circular region was calculated by the log-polar transform to improve the rotation invariance of descriptor. The image with scaling and rotation, change of camera angle, image blurring, JPEG image compression and illumination change was selected, with overlapping error εo less than 0.5. Its recall & 1-precision curve graph display performance increased by 2%~10%. Compared with SIFT, GLOH and PCA-SIFT descriptor, etc., the proposed improved descriptor is more robust and more distinguishable.

KEY WORDS: local feature; descriptor; affine invariant; SIFT

特征提取是图像处理技术中的重要研究领域，目前已广泛应用于计算机视觉、目标识别与跟踪等方面[1—3]。特征提取分为全局特征和局部特征，全局特征在有遮挡或者尺度变换等情况下检测性能不佳；局部特征对平移、旋转和尺度变换具有良好的鲁棒性，近年来其逐渐成为图像匹配领域的研究热点。基于局部特征的提取由特征检测子和特征描述子等部分组成，其中特征检测子用来确定提取特征的位置及局部区域，特征描述子用来描述该局部区域[4]。

早期的特征检测子主要有 Harris 角点检测子[5]、Harris-Laplace 算子[6]、DOG 算子[7]等，但是这些算子不能有效地对抗仿射变换。为此，近年来许多学者对局部特征的仿射不变性进行了深入研究，提出了具备仿射不变性的特征检测子，如 Harris-Affine 检测子、Hessian-Affine 检测子、基于边缘区域（EBR）的检测子、基于图像区域（IBR）的检测子、最大稳定极值区域（MSER）检测子等。Mikolajczyk 等[8]对现有的各种局部特征检测子的性能进行了比较和评测，指出

收稿日期：2017-07-02
基金项目：陕西省教育厅科研计划（17JK0990）；陕西省自然科学基金（2016JM6022）
作者简介：章明珠（1981—），女，硕士，西安翻译学院讲师，主要研究方向为图像处理、特征提取。

hessian-affine 和 MSER 是目前最好的检测子。

特征描述子有形状上下文描述子[9]、尺度不变特征 SIFT 描述子[10]、不变矩、复杂的滤波器描述符等。而在众多的描述子中尺度不变特征变换（SIFT）凭借其尺度、旋转、光照不变性，成为目前应用最广泛的特征描述子。SIFT 描述子维度较高计算复杂，不具有仿射不变性，为此很多学者对此进行了改进。Ke 和 Sukthankar[11]运用主成分分析法（PCA）对 SIFT 算法进行降维，提出了 PCA-SIFT 描述子。颜雪军[12]等对 PCA-SIFT 中对本征向量空间的求解进行了降维操作，提出了 2DPCA-SIFT 有效提高了计算效率。Mikolajczyk[13]等采用环形扇区代替矩形区域构建梯度位置直方图提出了（GLOH）描述子。Bay 等[14]利用小波变换和 Hession 矩阵增加特征描述子的鲁棒性，提高时间效率，提出一种类似于 SIFT 算法的 SUR 算法。Fan 等[15]采用灰度划分的方法，提出了多支持区域梯度直方图（MROGH）。黎剑兵[16]等改进了 MROGH 算子提出一种结合环状划分的（CLCGH），提高了描述子的仿射不变性。虽然近来研究者提出了一些新的方法，但是，对具有仿射变换的图像还是会产生大量的误匹配，为提高其鲁棒性和可区分性，文中提出一种改进的 SIFT 描述符。

1　新描述子的构造

1.1　椭圆区域提取

SIFT 描述符在围绕特征点的圆形区域内计算主方向上的梯度方向直方图。圆形区域的大小是由特征点的尺度来决定，因此具有尺度不变性，但对于不同的图像结构这样的邻域基本上没有仿射不变性的。局部邻域的形状可以用像素点附近像素的灰度梯度的二阶矩矩阵来描述它们的梯度分布，因此，可以利用二阶矩矩阵来估计图像中特征点的椭圆邻域结构，以获得描述子的仿射不变性。

对任意给定灰度图像 $I(x)$，其仿射高斯尺度空间表示见式（1）。

$$L(x,\Sigma)=g(x,\Sigma)\times I(x) \tag{1}$$

式中：$x\in R^2$；Σ为一个对称半正定协方差矩阵，并且与特征点的尺度对应。非均匀高斯核函数 $g(x,\Sigma)$ 定义见式（2）。

$$g(x,\Sigma)=\frac{1}{2\pi\sqrt{\det\Sigma}}\exp\left(-\frac{x^{\mathrm{T}}\Sigma^{-1}x}{2}\right) \tag{2}$$

在仿射尺度空间中二阶矩矩阵在图像中任一点的定义为：

$$\mu(x,\Sigma_{\mathrm{I}},\Sigma_{\mathrm{D}})=\begin{pmatrix}\mu_{11} & \mu_{12}\\ \mu_{21} & \mu_{22}\end{pmatrix}$$

$$=\det(\Sigma_D)g(x,\Sigma_{\mathrm{I}})\times((\nabla L)(x,\Sigma_{\mathrm{D}})(\nabla L)(x,\Sigma_{\mathrm{D}})^{\mathrm{T}}) \tag{3}$$

式中：T 为转置操作符；$\sum_{\mathrm{I}}$ 和$\sum_{\mathrm{D}}$ 为高斯核的积分高斯核和微分高斯核；∇L 为梯度算子，定义见式（4）。

$$\nabla L(x,\Sigma_{\mathrm{D}})=\begin{pmatrix}L_x(x,\Sigma_{\mathrm{D}})\\ L_y(x,\Sigma_{\mathrm{D}})\end{pmatrix} \tag{4}$$

借鉴 Mikolajczyk[8]提出通过迭代的方法求二阶矩矩阵：

$$M=\mu(x,\Sigma_{\mathrm{I}},\Sigma_{\mathrm{D}}) \tag{5}$$

$$\Sigma_{\mathrm{I}}=\sigma_{\mathrm{I}}M^{-1} \tag{6}$$

$$\Sigma_{\mathrm{D}}=\sigma_{\mathrm{D}}M^{-1} \tag{7}$$

为了确保特征点的椭圆邻域内的每个样本点能够正确映射到相应的块中，文中利用特征点的二阶矩矩阵求取椭圆区域的参数，然后，使用二阶矩矩阵的平方根将椭圆图像区域内数据都转换到圆形区域：

$$x'=M^{1/2}x \tag{8}$$

1.2　基于椭圆区域的仿射尺度空间

对于已经归一化的圆形的邻近区域 $I'(x)$ 上的特征点 x，仿射尺度空间图像的尺度接近于这个特征点的局部尺度 σ，因此可以通过与标准高斯核函数的卷积来生成：

$$L(x,\sigma)=g(x,\sigma)\times I'(x) \tag{9}$$

式中标准高斯核函数 $g(x,\sigma)$ 定义见式（10）。

$$g(x,\sigma)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^{\mathrm{T}}x}{2\sigma^2}\right) \tag{10}$$

这个高斯核是圆对称的而且它的参数由一个尺度因子 σ 确定。尺度图像可以对图像进行高斯平滑来得到。这个过程可以重复使用在上一级平滑过的图像上，得到一组不同分辨率下连续变化的图像。

1.3　对数极坐标变换

归一化的图像上每个像素点可以用笛卡尔坐标 $z(x, y)$ 表示，也可以用极坐标 $p(r,\theta)$ 表示，设坐标原点为（0, 0），则两坐标满足如下关系：

$$z=x+\mathrm{i}y=r(\cos\theta+\mathrm{i}\sin\theta)=r\,\mathrm{e}^{\mathrm{i}\theta} \tag{11}$$

$$r=\sqrt{x^2+y^2},\quad \theta=\arctan(y/x) \tag{12}$$

设点 $z(x,y)$ 的对数极坐标为 $\omega(\rho,\varphi)$，令 $\omega=\ln z$，则可推导出笛卡尔坐标与对数极坐标的关系为：

$$\rho=\ln r+\mathrm{i}\theta=0.5\ln(x^2+y^2),\quad \varphi=\theta \tag{13}$$

假定点 $z(x, y)$ 缩放 r_0 倍，并旋转 θ_0 度，由对数极坐标的性质可知，变换后的点 $z'(x',y')$，ρ,φ 满足如下关系：

$$z'(x',y')=r_0 r\mathrm{e}^{\mathrm{i}(\theta+\theta_0)} \tag{14}$$

$$\rho=\ln r+\ln r_0,\quad \varphi=\theta+\theta_0 \tag{15}$$

由式（14—15）可以看出，笛卡尔坐标下图像的缩放或旋转，相当于对数极坐标中图像的垂直或水平

平移，见图 1，有效克服了图像的旋转和尺度不变性。

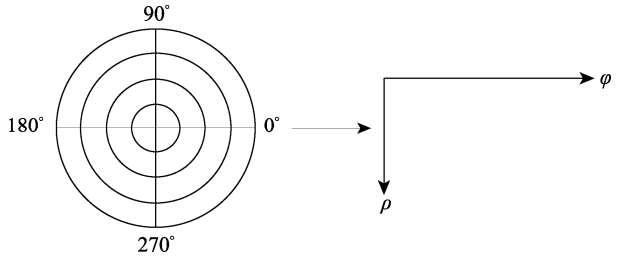

图 1　对数极坐标变换原理

Fig.1 The principle of log-polar transformation

1.4　在圆形区域上计算梯度直方图

SIFT 描述符在方形区域上求梯度直方图，当图像发生旋转变化后，如若直接选取同样方形区域构建描述符，势必会产生较大误差，见图 2。考虑圆形区域在不同的主方向中所包含的像素是完全一致的，文中利用对数极坐标方法在圆形区域上计算梯度直方图，主要步骤如下所述。

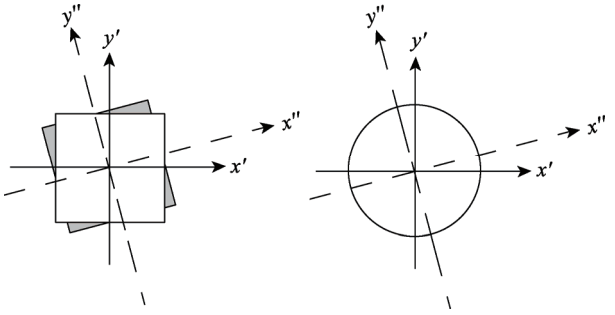

图 2　旋转后方形区域和圆形区域误差

Fig.2 Error of the square and circular regions after rotation

1）在归一化的图像区域上，计算每一个关键点 $L(x,y)$ 的梯度值 $M(x,y)$ 和方向 $\theta(x,y)$。

$$M(x,y)=\ln\sqrt{\left(L(x+1,y)-L(x-1,y)\right)^2+\left(L(x,y+1)-L(x,y-1)\right)^2} \tag{16}$$

$$\theta(x,y)=\arctan\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right) \tag{17}$$

2）以主方向为起点，将圆形邻近区域划分成 4 个同心圆，每个同心圆在均匀划分成 4 个扇形网格，每个扇格内计算 8 个方向（0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°）的梯度累加值，见图 3，形成 4×4×8 个特征向量。

3）为避免外围点对特征点造成较大的影响，将每个像素的梯度通过均匀高斯函数进行加权运算。

4）将 0 号扇区内 r=1 的子环内 8 个梯度累加值排序后，编码为 1—8 号向量；将 0 号扇区内 r=2 的子环内 8 个梯度累加值排序后，编码为 9—16 号向量；将 0 号扇区内 r=3 的子环内 8 个梯度累加值排序后，

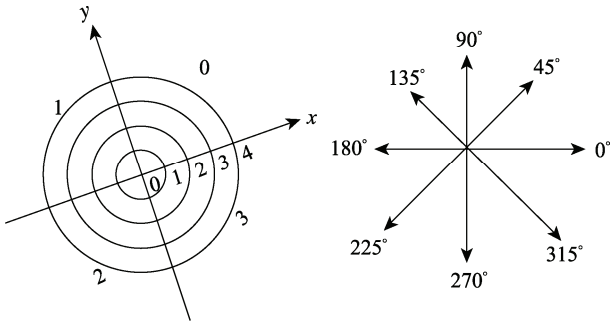

图 3　改进的 SIFT 特征描述符

Fig.3 Improved SIFT feature descriptor

编码为 17—24 号向量；以此类推，将 3 号扇区内 r=4 的子环内 8 个梯度累加值排序后，编码为 121—128 号向量，形成 128 维向量描述符。

2　实验结果及分析

评价描述子性能的指标，文中采用 recall 和 1−precision 曲线。recall, 1−precision 的定义见式（21—22）。两幅图中一致区域的个数

$$\text{recall}=\frac{\text{正确的匹配对个数}}{\text{两幅图中一致区域的个数}} \tag{21}$$

$$1-\text{precision}=\frac{\text{错误匹配对个数}}{\text{总的匹配对个数}} \tag{22}$$

通过计算重叠误差 ε_o 评价一对匹配区域的相似程度，定义见式（23）。

$$\varepsilon_o=1-\frac{\boldsymbol{A}\cap\boldsymbol{H}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{H}}{\boldsymbol{A}\cup\boldsymbol{H}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{H}} \tag{23}$$

式中：$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 2 个图像区域的像素矩阵；$\boldsymbol{H}$ 为这 2 个图像区域之间的平面映射变换矩阵。如果 2 个图像区域的重叠误差小于选择阈值认为匹配是正确的，否则匹配是错误的，文中选用的阈值小于 0.5。

为评价改进描述符的性能，文中从标准 INRIA 数据库中选取测试图像，见图 4。测试图像中涵盖了基本的图像变换：缩放和旋转、摄像机角度的变化、图像模糊、JPEG 图像压缩、光照变化。同时，与原始的 SIFT 描述、梯度位置方向直方图（GLOH）、PCA-SIFT 描述符、文献[16]提出的 CLCGH 描述符进行了对比，图 5—9 分别是对图 4a—e 中每种图像变换的 recall & 1−precision 曲线。实验中 5 种描述符分别用 SIFT, GLOH, PCA-SIFT, CLCGH 和文中提出的方法 MSELF 来表示。

旋转角度大约为 45°，尺度变化因子大约是 2.0 的图像（图 4a）的 recall & 1−precision 曲线见图 5。从图 5 中可以看出，文中改进的描述符的性能优于其他描述符，此结果验证了使用对数极坐标变换，在圆形区域的扇形网格内求方向直方图的方法，有效减少了旋转对方形区域的误差的影响，增强了描述符的抗旋转能力。

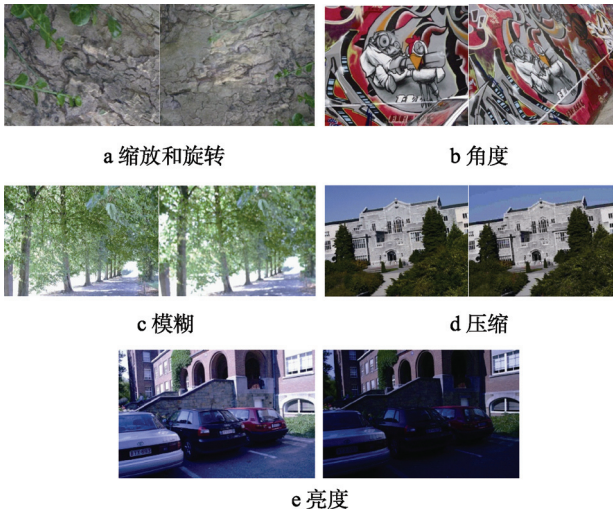

图 4　测试图像
Fig.4 Test image

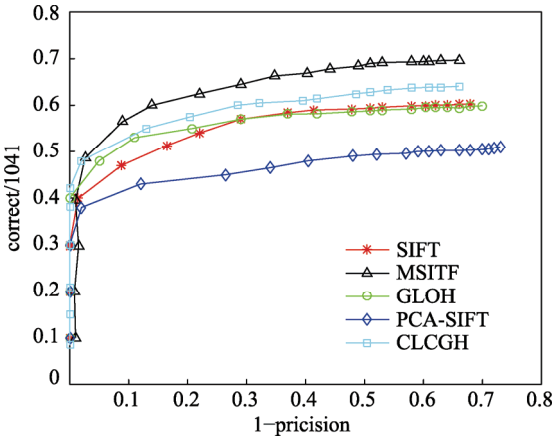

图 5　缩放和旋转变换下实验结果
Fig.5 Experimental results of scaling and rotation transformation

摄像机的视角发生 45°变换的图像（图 4b）进行性能评比见图 6。从图 6 中可以看出，文中改进的方法性能优于原始的 SIFT 描述、梯度位置方向直方图（GLOH）、PCA-SIFT 和 CLCGH 描述符。这个可能与文中采用了椭圆区域规一化有关，有效增强了描述符的仿射不变性。

在具有大量模糊的图像上（图 4c）上进行对比实验见图 7，图像中的模糊是由于改变照相机的焦点而产生的。从实验结果可以看出，即使没有在图像中加入任何的几何变换，文中改进的方法、原始的 SIFT 描述、梯度位置方向直方图（GLOH）、PCA-SIFT、CLCGH 描述符都严重受这类图像变换的影响。像素灰度和局部形状的结构变化是一个不可预知的变换方式，这种情况下描述符不鲁棒。

在具有 JPEG 图像压缩变换图像（图 4d）上进行对比实验见图 8，其中图像的质量是原参考图像的 5% 左右。在具有光照变化的图（图 4e）进行对比的实验结果见图 9，这种光照变化是通过改变亮度以及光源的位置来模拟的。从实验结果可以看出，文中改进的方法性能优于原始的 SIFT 描述、梯度位置方向直方图（GLOH）、PCA-SIFT、CLCGH 描述符，区分能力更强，尤其是纹理突出的图 4a, b 等图像中表现更明显。其原因可能是文中改进算法采用了圆形区域上对数极坐标方向直方图提供更多的鲁棒信息。

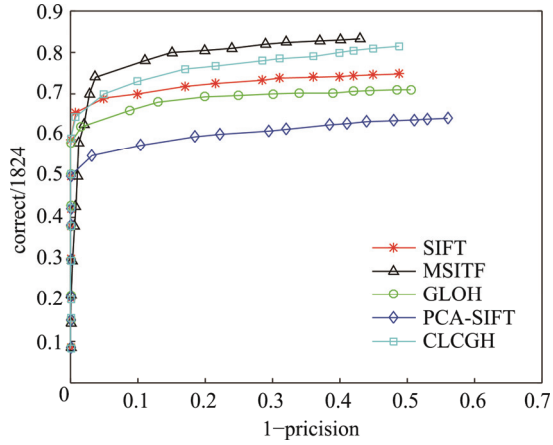

图 6　视角变换下的实验结果
Fig.6 Experimental results of angle transform

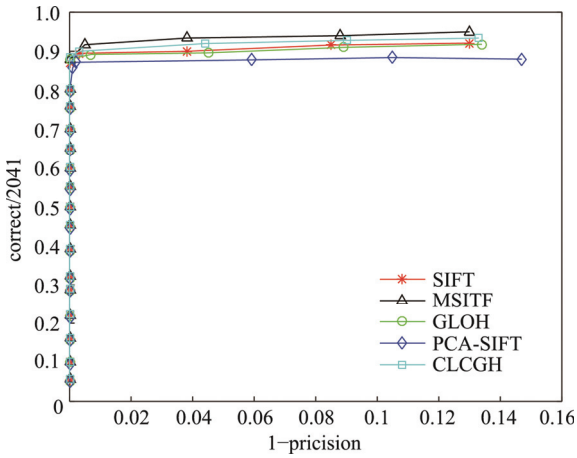

图 7　模糊变换下的实验结果
Fig.7 Experimental results of fuzzy transform

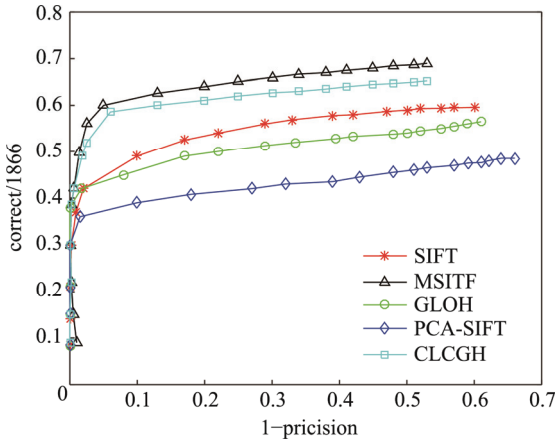

图 8　压缩变换下的实验结果
Fig.8 Experimental results of compression transform

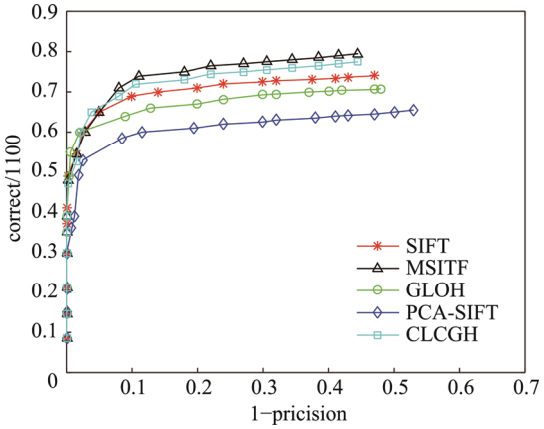

图 9　亮度变换下的实验结果

Fig.9 Experimental results of brightness transform

3　结语

提出了一种新的描述符，从规一化椭圆邻近区域转换到仿射尺度空间，然后采用极坐标方向直方图提高描述符的精确度来改进 SIFT 描述符，并且在图像匹配中将这个描述符与经典的 SIFT 算子、GLOH、PCA-SIFT、CLCGH 作比较。实验结果表明，文中的改进算法在缩放和旋转、摄像机角度的变化、图像模糊、JPEG 图像压缩、光照变化等图像变换中都能表现出较好的匹配性能，具有更强的区分性和鲁棒性。

参考文献：

[1] 张焕龙, 胡士强, 杨国胜. 基于外观模型学习的视频目标跟踪方法综述[J]. 计算机研究与发展, 2015, 52(1): 177—190.
ZHANG Huan-long, HU Shi-qiang, YANG Guo-sheng. Video Object Tracking Based on Appearance Models Learning[J]. Journal of Computer Research and Development, 2015, 52(1): 177—190.

[2] 曹明伟, 李书杰, 贾伟, 等. 面向大规模 SFM 的快速鲁棒特征跟踪算法[J]. 仪器仪表学报, 2017, 38(1): 252—260.
CAO Ming-wei, LI Shu-jie, JIA Wei. Fast and Robust Feature Tracking Algorithm for Large-Scale SFM[J]. Chinese Journal of Scientific Instrument, 2017, 38(1): 252—260.

[3] 陈青, 王飞. 基于 SIFT 的 NSCT 域抗几何攻击水印算法[J]. 包装工程, 2017, 38(5): 178—182.
CHEN Qing, WANG Fei. NSCT-domain Watermarking Algorithm Robust to Geometric Attacks Based on SIFT[J]. Packaging Engineering, 2017, 38(5): 178—182.

[4] 王鹏, 王平, 沈振康, 等. 一种基于 SIFT 的仿射不变特征提取新方法[J]. 信号处理, 2011, 27(1): 87—93.
WANG Peng, WANG Ping, SHEN Zhen-kang, et al. A Novel Algorithm for Affine Invariant Feature Extraction Based on SIFT[J]. Signal Processing, 2011, 27(1): 87—93.

[5] BROWN M, HUA G, WINDER S. Discriminative Learning of Local Image Descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(1): 43—57.

[6] MIKOLAJCZYK K, SCHMID C. Scale & Affine Invariant Interest Point Detectors[J]. International Journal of Computer Vision, 2004, 60(1): 63—86.

[7] 颜雪军, 赵春霞, 袁夏, 等. 一种鲁棒的基于图像对比度的局部特征描述方法[J]. 电子与信息学报, 2014, 36(4): 882—887.
YAN Xue-jun, ZHAO Chun-xia, YUAN Xia, et al. A Robust Local Feature Descriptor Based on Image Contrast[J]. Journal of Electronics & Information Technology, 2014, 36(4): 882—887.

[8] MIKOLAJCZYK K, TUYTELAARS T, SCHMID C, et al. A Comparison of Affine Region Detectors[J]. International Journal of Computer Vision, 2005, 65(1/2): 43—72.

[9] 冯祥卫, 冯大政, 侯瑞利. 高斯形状上下文描述子[J]. 西安电子科技大学学报, 2016, 43(4): 45—50.
FENG Xiang-wei, FENG Da-zheng, HOU Rui-li. Gaussian Shape Context Descriptors[J]. Journal of Xidian University, 2016, 43(4): 45—50.

[10] LOWE D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91—110.

[11] KE Y, SUKTHANKAR R. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004 (502): 506—513.

[12] 颜雪军, 赵春霞, 袁夏. 2DPCA-SIFT: 一种有效的局部特征描述法[J]. 自动化学报, 2014, 40(4): 675—682.
YAN Xue-Jun, ZHAO Chun-Xia, YUAN Xia. 2DPCA-SIFT: An Efficient Local Feature Descriptor[J]. Acta Automatica Sinica, 2014, 40(4): 675—682.

[13] MIKOLAJCZYK K, SCHMID C. A Performance Evaluation of Local Descriptors[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2005, 27(10): 1615—1630.

[14] BAY H, ESS A, TUYTELAARS T, et al. Speeded-Up Robust Features (SURF)[J]. International Journal on Computer Vision and Image Understanding(S1077-3142), 2008, 110(3): 346—359.

[15] FAN B, WU F, HU Z. Rotationally Invariant Descriptors Using Intensity Order Pooling[J]. IEEE Transactions on Pattern Analysis and Machine intelligence, 2012, 10(34): 2031—2045.

[16] 黎剑兵, 肖潇, 李小平, 等. 一种新的结合环状区域划分的特征描述子[J]. 西安电子科技大学学报, 2016, 43(2): 64—69.
LI Jian-bing, XIAO Xiao, LI Xiao-ping, et al. Local Coordinate Gradient Descriptor Based on Cyclic Division[J]. Journal of Xidian University, 2016, 43(2): 64—69.