doi:10.11959/j.issn.1000-436x.2015089

改进的图像局部特征区域描述方法

朱仁欢，高清维，卢一相，孙冬
（安徽大学 电气工程与自动化学院，安徽 合肥 230601）

摘 要：针对描述子的性能与维数相矛盾的问题，提出了一种顽健的图像局部特征区域的描述方法。首先按照像素排序将局部特征区域分割为若干个子区域，然后利用基于阈值的分段局部描述子设计方法计算描述子，并采用纹理谱加权方法累加局部描述子得到子区域描述子，最后连接各部分子区域描述子得到最终的特征描述子。该方法综合了全局信息和局部信息，在保证描述子维数较小时对噪声具有一定的顽健性。实验结果表明该方法不仅对单调强度变化和旋转变化具有不变性，而且对其他几何和光学变换具有较好的顽健性。
关键词：局部特征；特征检测；特征描述；旋转不变性
中图分类号：TP39 **文献标识码**：A

Improved method for local image feature region description

ZHU Ren-huan, GAO Qing-wei, LU Yi-xiang, SUN Dong
(College of Electrical Engineering and Automation, Anhui University, Hefei 230601, China)

Abstract: A robust method for local image feature region description was proposed based on the problem of contradiction between performance and dimension of descriptor. First, the local feature region was divided into several subregions according to their intensity order. Then, the method of segmented local feature description based on threshold was used to compute the descriptor, and the method of the weighted texture spectrum was used to cumulate the local descriptor. At last, concatenate the descriptors of every subregion together to get the final descriptor of the interest region. This method combined overall information and local information together, and robust to noise when ensure the dimension of descriptor was small. The experimental results show that the proposed descriptor is not only invariant to monotonic intensity changes and image rotation but also robust to many other geometric and photometric transformations.
Key words: local feature; feature detection; feature description; rotation invariant

1 引言

图像局部特征描述是近年来计算机视觉和模式识别领域的研究热点之一，在目标识别[1]，宽基线匹配[2]，图像检索[3]，全景图像拼接[4]等具体应用中发挥着重要作用。图像局部特征描述在应用中的基本思路是：首先从待匹配图像中检测出特征点或者特征区域，然后对特征点或特征区域的邻域设计具有一定不变性的局部特征描述子，最后通过匹配这些描述子来确定图像对应点。利用局部特征的优点是对各种光照和几何变化具有较好的顽健性。

在局部特征描述方法中，最受关注的是由 Lowe 提出的 SIFT（scale invariant feature transform）描述子[1]。这种特征描述子的构造是通过对特征点的邻域建立一个三维梯度直方图来实现的。SIFT 不仅对图像的旋转变化和尺度变换具有不变性，而且对光照和几何形变也具有较强的适应性。在此基础上，许多学者对 SIFT 进行了改进，例如 PCA-SIFT[5]、GLOH[6]、SURF[7]等。近年来，也出现了一些较新的局部描述子，如 CS-LBP[8]利用中心对称的局部二值模型 LBP（local binary pattern）构造描述子，由于利用中心对称的采样点，从而避免了描述子维数

收稿日期：2014-01-08；**修回日期**：2014-04-19
基金项目：国家自然科学基金资助项目（61370110）；国家自然科学青年基金资助项目（61402004, 61402003）
Foundation Items: The National Natural Science Foundation of China (61370110); The Youth Project of the National Natural Science Foundation of China (61402004, 61402003)

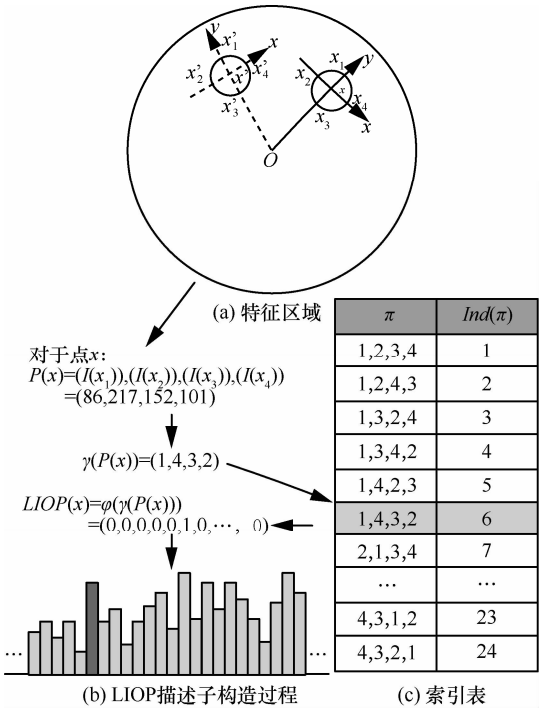
过大。HRI-CSLTP[9]将 CS-LBP 扩展为三值编码，使描述子具有更强的辨别能力。DAISY[10]是一种类似于 SIFT 和 GLOH 的描述子，但是计算速度比 SIFT 要快，并且有效地利用于宽基线立体匹配中。从文献[6,11]中可以看出类似 SIFT 设计思想的描述子性能是最好的，即对特征点邻域或特征区域分块，然后分别构造局部描述子，最后连接各子区域的描述子。

但是在 SIFT 等算法中构造描述子时需要估计主方向，而主方向的估计是造成性能下降的一个主要原因[12]。在文献[12]中，作者详细讨论了由于主方向估计而使描述子性能下降的原因，在主方向估计误差太大时导致最终图像匹配效果很差，因此作者提出了一种新的描述子构造框架，并且提出了基于局部梯度和像素强度的 2 种描述子：MROGH（multisupport region order-based gradient histogram）和 MRRID（multisupport region rotation and intensity monotonic invariant descriptor）。这 2 种描述子采用多层支撑区域重复计算描述子，构造过程比较复杂，计算量较大。文献[13]中作者提出了 LIOP（local intensity order pattern）描述子，其构造框架与文献[12]类似，但是描述子计算方法比文献[12]要简单。LIOP 描述子在构造时，仅利用邻域像素排序即可计算出描述子，该描述子不仅对均匀光照和旋转变化具有固有不变性，而且对尺度变化、视角变化和复杂光照变化具有很好的顽健性。但是由于利用了像素排序，因此该描述子对噪声较敏感，而且计算描述子时利用邻域像素较少，但是如果利用较多邻域像素时，描述子的维数将会较大。鉴于此，本文提出了 IWLIOP（improved weighted local intensity order pattern）描述子，降低了描述子对噪声的敏感性，并且保证了在更多地利用局部像素的前提下维数较低。

2　LIOP 描述子

2.1　描述子计算框架

在构造描述子之前，首先检测特征区域，一般检测到的特征区域是椭圆形区域，因为圆形区域具有固有的旋转不变性，所以一般将椭圆形区域规范化为一个固定半径的圆形区域。然后将圆形区域内的像素进行排序，按照像素顺序将整个区域分为若干个子区域，每个区域所包含的像素个数相同，每个子区域内的像素不必是彼此相邻的。因为经过像素排序分块，因此这种方法使最终的描述子对均匀强度变化具有固有的不变性。对于每一个子区域内的每一个像素点建立局部坐标系，如图 1(a)所示。圆形区域圆心为 O，对于某像素点 x，以 $\vec{Ox}$ 为 y 轴正方向，x 轴垂直于 y 轴，且正方向为 y 轴正方向顺时针旋转 90°的方向。在 x 的半径为 R 的邻域上均匀采样 N 个点，其中第一个在 y 轴正半轴上，其余的点按照逆时针方向均匀采样，由这 N 个点的像素值计算 LIOP 描述子。如图 1(a)中，x'为 x 旋转后的点，按照同样的方式采样 N 个点，然后计算描述子。由于 x 和 x'邻域采样点完全相同，所以得到的描述子完全相同，因此该局部坐标系的建立保证了旋转不变性。累计每个子区域内所有点的描述子，得到每个子区域的描述子，最后将各个子区域的描述子顺次连接起来构成最终的描述子。



π	$Ind(\pi)$
1,2,3,4	1
1,2,4,3	2
1,3,2,4	3
1,3,4,2	4
1,4,2,3	5
1,4,3,2	6
2,1,3,4	7
…	…
4,3,1,2	23
4,3,2,1	24

图 1　LIOP 描述子构造过程

2.2　LIOP 描述子的设计

对于某个像素点 x，在其半径为 R 的邻域上按照 2.1 节所述的采样规则采样 N 个点：$\{x_1,x_2,\cdots,x_N\}$，得到 N 个像素值：$P(x)=\{I(x_1),I(x_2),\cdots,I(x_N)\}$，$I(x_i)$ 表示 x_i 点的像素值。将这 N 个像素值按照从小到大的顺序排序，其序号构成一个如式(1)所示的 N 位序列 π。

$$\pi=\gamma(P)=(i_1,i_2,\cdots,i_N)\quad I(x_{i_1})<I(x_{i_2})<\cdots<I(x_{i_N}) \tag{1}$$

然后建立索引表，索引表的第一列是排列序号的可能情况，即 1 到 N 这 N 个数排列组合的所有情况，第二列是对应的索引值，如图 1(c)所示的索引表是 $N=4$ 时的情况。对于 N 个像素值，其序号排列组合有 $N!$ 种情况。由序列 π 找到对应的索引值 $Ind(\pi)$，将 $Ind(\pi)$ 按照式(2)转化为一个 $N!$ 维的向量。

$$\boldsymbol{\phi}(\pi)=\boldsymbol{V}_{N!}^{Ind(\pi)},\pi\in\Pi^N \tag{2}$$

其中，$\boldsymbol{V}_{N!}^{Ind(\pi)}$ 是只有第 $Ind(\pi)$ 个数为 1，其余均为 0 的 $N!$ 维向量，即关于点 x 的 LIOP 描述子如式(3)所示。

$$\begin{aligned}LIOP(x)&=\phi(\gamma(P(x)))\\&=V_{N!}^{Ind(\gamma(P(x)))}\\&=(0,\cdots,0,\ 1\ ,0,\cdots,0)\\&\qquad(Ind(\gamma(P(x))))\end{aligned} \tag{3}$$

再将每个子区域内各点计算的描述子进行累加得到该子区域的描述子，最后顺次连接各子区域的描述子得到最终的描述子，如式(4)所示。

$$\begin{aligned}&LIOP\ \ descriptor=(des_1,des_2,\cdots,des_B)\\&des_i=\sum\nolimits_{x\in bin_i}LIOP(x)\end{aligned} \tag{4}$$

其中，bin_i 表示第 i 个子区域，des_i 表示第 i 个子区域的描述子，B 为特征区域被分割的子区域个数，则最终描述子维数为 $B\times N!$。

3 IWLIOP 描述子

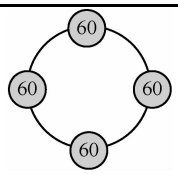
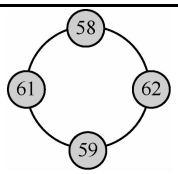
在第 2.2 节中，对 N 个像素值排序得到一个 N 位数的排列组合，当存在噪声时，邻域像素的顺序可能发生很大的变化，由式(1)计算得到的 π 不同，则最终的描述子也不同，特别是对于平坦区域，更易受噪声影响，如表 1 所示，当存在噪声时得到的 LIOP 描述子与没有噪声时得到的 LIOP 描述子完全不同。因此，由于经过像素排序，LIOP 描述子受噪声的影响。针对这种情况，本文提出了基于阈值的 LIOP 描述子，即 T-LIOP (threshold-local intensity order pattern)。

在对邻域像素值排序时，不是简单地从小到大的顺序排序，首先设定一个阈值 T，只有当 $I(x_i)-I(x_j)>T$ 时才认为 $I(x_i)>I(x_j)$，即按照式(5)得到 π。

$$\begin{aligned}&\pi=\gamma(P)=(i_1,i_2,\cdots,i_N)\\&I(x_{i_{j+1}})-I(x_{i_j})>T\ \ (j=1,2,\cdots,N-1)\end{aligned} \tag{5}$$

表 1 是 $T=5$ 时平坦区域中 LIOP 描述子和 T-LIOP 描述子的对比结果。从中可以看出，通过设置一定的阈值，使最终的描述子在有噪声和无噪声时是完全相同的。因此 T-LIOP 描述子对噪声具有较好的抗干扰性。

表 1 LIOP 与 T-LIOP 示例（T=5）

描述子	 纹理平坦区域	 带噪声的纹理平坦区域
LIOP	π=(1,2,3,4) $Ind(\pi)$=1 $LIOP(x)$=(1,0,⋯,0)	π=(1,3,2,4) $Ind(\pi)$=3 $LIOP(x)$=(0,0,1,0,⋯,0)
T-LIOP	π=(1,2,3,4) $Ind(\pi)$=1 $T\text{-}LIOP(x)$=(1,0,⋯,0)	π=(1,2,3,4) $Ind(\pi)$=1 $T\text{-}LIOP(x)$=(1,0,⋯,0)

由文献[13]的实验结果可以知道，当邻域采样点数越多时，描述子的性能越好，这是因为采样点数越多，利用的局部信息越丰富，描述子所表达的信息更加精确。但是邻域采样点数越多，描述子维数越大，为了能更多地利用局部信息，同时又不使描述子维数过大，本文提出了 IWLIOP 描述子。首先将某点邻域的采样像素值分为 2 部分来进行描述子计算，再将其串联起来得到该点的描述子。

对于某像素点 x，其邻域采样点数为 N（偶数），将其对应的 N 个像素值按照奇偶序号分为 2 部分，如式(6)所示。

$$\begin{aligned}P^{(1)}(x)&=\{I(x_1),I(x_3),\cdots,I(x_{N-1})\}\\P^{(2)}(x)&=\{I(x_2),I(x_4),\cdots,I(x_N)\}\end{aligned} \tag{6}$$

然后由式(5)得到对应的 $\pi^{(1)}$ 和 $\pi^{(2)}$，再由式(7)得到 $T\text{-}LIOP^{(i)}$ 描述子。

$$\begin{aligned}T\text{-}LIOP^{(i)}(x)&=\phi(\gamma(P^{(i)}(x)))\\&=V_{(N/2)!}^{Ind(\gamma(P^{(i)}(x)))}\\&=(0,\cdots,0,\ 1\ ,0,\cdots,0)\\&\qquad(Ind(\gamma(P^{(i)}(x))))\end{aligned} \tag{7}$$

其中，$i=1,2$，得到的 $T\text{-}LIOP^{(i)}$ 描述子是 $(N/2)!$ 维向量。最后将 2 部分描述子串联得到该像素点的描述子如式(8)所示。

$$T\text{-}LIOP(x)=(T\text{-}LIOP^{(1)}(x),T\text{-}LIOP^{(2)}(x)) \quad (8)$$

如果邻域采样点数较多，但是因为分成了 2 部分来计算描述子，维数不会因为采样点数的增多而增大，最终的描述子又能包含更多的局部信息，因此能够更加准确地描述图像特征。

因为纹理变化区域比纹理平坦区域更加稳定[13]，因此并不是将各像素点的描述子直接累加，而是采用了一种纹理谱加权计算方法，用某点的邻域所有采样像素值之差的平方和作为该点描述子的权值，权值的计算如式(9)所示。

$$w(x)=\sum_{i,j}(I(x_i)-I(x_j))^2 \quad (9)$$

然后加权累加得到子区域的描述子，最后将各个子区域的描述子顺次串联起来得到最终的特征区域描述子，如式(10)所示。

$$IWLIOP=(des_1,des_2,\cdots,des_B)$$
$$des_i=\sum_{x\in bin_i}w(x)T-LIOP(x) \quad (10)$$

其中，bin_i 表示第 i 个子区域，des_i 表示第 i 个子区域的描述子。

若局部特征区域被分割的子区域个数为 B，则最终描述子维数为 $2B(N/2)!$。当 $B=6,N=4$ 时，LIOP 的维数为 144（$6\times4!$），IWLIOP 的维数为 24（$2\times6\times(4/2)!$），可以看出，在相同参数时，IWLIOP 描述子的维数较小。

4　实验与结果分析

为了验证本文算法的性能，做了仿真实验。所有实验均在 CPU 为 Intel 奔腾双核 2.6 GB，内存 4 GB（2.99 GB 可用）的 32 位 Windows 7 系统的计算机上采用 Matlab2012a 编程实现。首先说明了实验数据和评价标准，然后做了仿真实验。

4.1　数据与评价标准

本文利用标准数据库[14]中的图像对算法进行性能评价，选取了数据库中包含以下 6 种几何和光照变换的图像：视角变化、尺度变化、图像旋转、图像模糊、光照变化和 JPEG 压缩，其中每种包括 6 张图像。为了研究本文算法对复杂光照变化的性能，本文利用了文献[13]采用的图像：desktop、corridor 和由 desktop 求平方根和平方得到的图像 desktop-syn，它们具有更剧烈的光照变化，desktop 和 corridor 包含的 5 张图像分别是不同曝光程度的图像，而 desktop-syn 包含 3 张图像，分别是 desktop 的第 5 张图像和对其平方根和平方运算得到的 2 张图像，所有测试图像如图 2 所示。

利用文献[6]中的评价标准来评价描述子，即基于 2 幅图像中正确匹配数与错误匹配数的 Recall-Precision 准则进行评价。匹配策略是利用欧式距离作为相似性度量，使用最近邻与次近邻距离比(NNDR, nearest neighbor distance ratio)[1]。计算召回率(recall) 和精度(precision)按照式(11)计算。

$$recall=\frac{\#correct\ marches}{\#correspondences}$$
$$1-precision=\frac{\#false\ matches}{\#all\ matches} \quad (11)$$

其中，*#correspondences* 是真实的正确匹配数目，按照文献[15]中的方法进行计算。

4.2　仿真实验

首先利用 Harris-Affine 仿射不变特征区域检测方法[16]检测特征区域，得到椭圆形的特征区域后将其规范化为直径为 41 像素的圆形区域。将圆形区域内的像素排序，然后按照像素顺序均匀分为 B 个子区域，邻域采样的点数为 N，采样半径为 R，按

(a) bikes　(b) boat　(c) graf　(d) leuven　(e) ubc　(f) wall

(g) desktop　(h) corridor　(i) desktop-syn(平方根)　(j) desktop-syn(平方)

图 2　实验图像

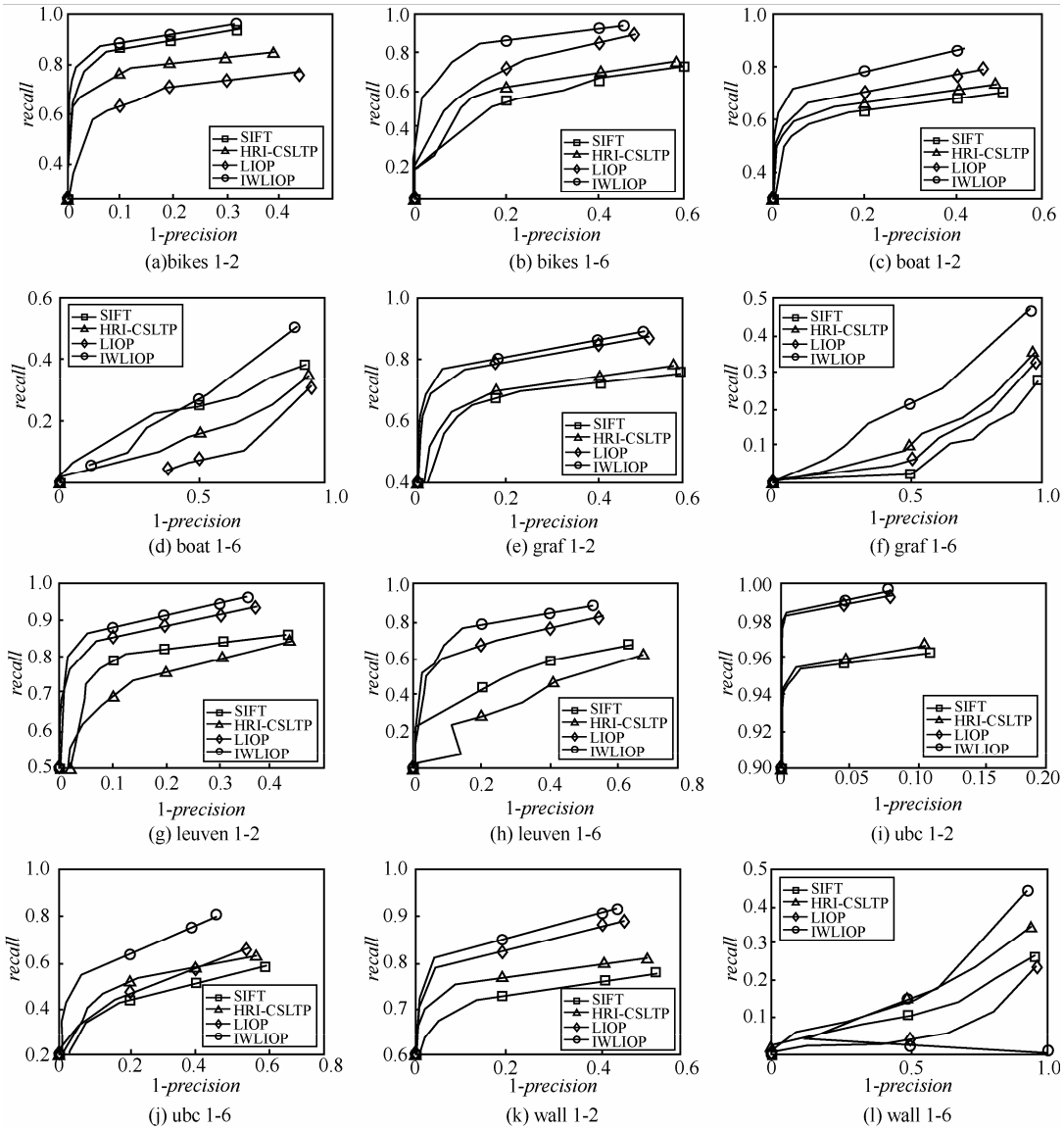


图 3 实验结果

照本文改进的算法计算 IWLIOP 描述子。对测试图像经过多次实验，当阈值 $T=5$ 时实验效果最好。因此，算法所涉及到的主要参数取值为：$B=6$，$N=6$，$R=4$，$T=5$。

将 IWLIOP 与 SIFT[1]、HRI-CSLTP[9]、LIOP[13]对比，实验结果如图 3 所示。图中仅显示了每种图像中变化最不剧烈和最剧烈的 2 组图像的实验结果，如图 3（a）bikes 1-2 表示 bikes 图中第一张图和第二张图的匹配 PR 曲线。从实验结果可看出，几乎在所有的结果中，IWLIOP 描述子的性能都比其他 3 种描述子的性能要好，特别是对于变化剧烈的图像，IWLIOP 描述子的性能更加突出，这是因为本文首先通过设定阈值减弱了噪声对描述子的影响，然后由于利用更多的邻域像素，即利用更多的局部信息，因此描述局部特征更加准确，最后又采用了更能反映图像纹理特性的描述子加权累加方法。

对光照变化剧烈的图像，实验结果如图 4 所示。从图中也可以看出，IWLIOP 描述子的性能比其他 3 种描述子的性能更加突出，特别是对于平方根和平方这 2 种非线性运算后的图像，该描述子仍具有更好的顽健性。

然而尽管利用较多的局部像素，但是由于采用了本文所提出的描述子分段计算再串联的描述子构造方法，因此 IWLIOP 描述子的维数相对其他 3 种描述子的维数较小。4 种描述子的维数如表 2 所

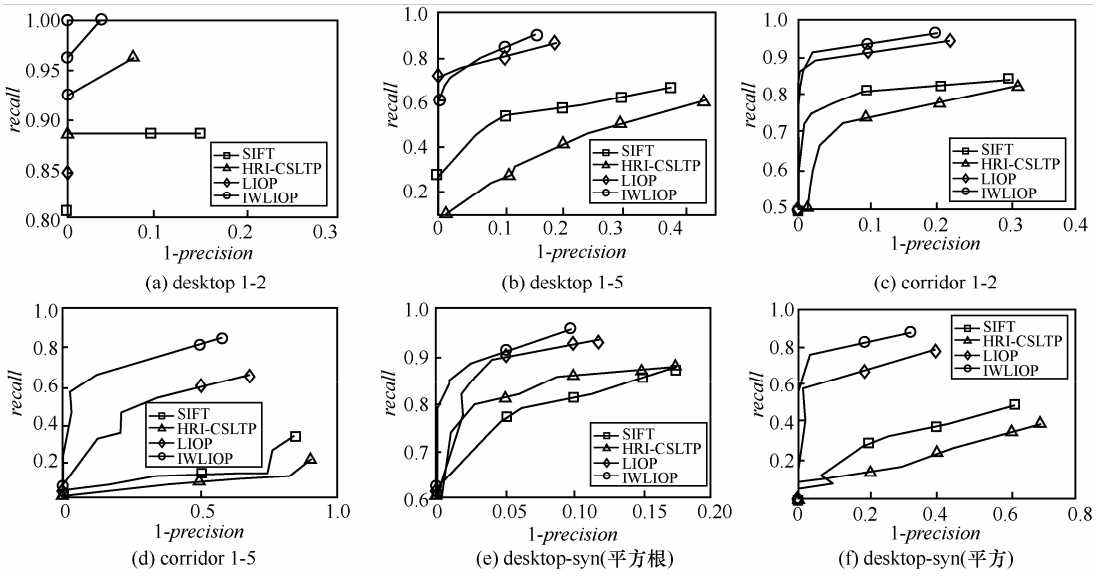


图 4　光照变化剧烈图像的实验结果

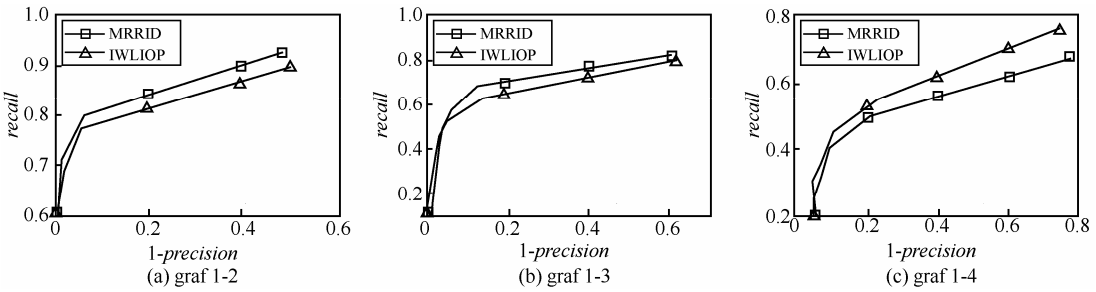


图 5　IWLIOP 描述子与 MRRID 描述子的对比实验结果

示，从表中可以看出 IWLIOP 描述子比 LIOP 描述子利用更多的邻域像素值，但是维数却是一半。表中还统计了构造一个描述子所需要的时间。从表中可以看出 IWLIOP 描述子计算时间比 LIOP 要稍长，但是在维数和性能方面更具有优越性。

表 2　各描述子维数与单个描述子计算时间

描述子	维数	时间/ms
SIFT	128	1.4
HRI- CSLTP	384	3.2
LIOP	144	3.4
IWLIOP	72	3.9

文献[12]采用多层支撑区域重复计算描述子，IWLIOP 与 LIOP 描述子相同，都只采用单层支撑区域计算描述子，因此 IWLIOP 描述子的性能比文献[12]中的 MROGH 和 MRRID 描述子的性能稍差，如图 5 是以 graf 图为例，显示了 IWLIOP 描述子与 MRRID 描述子的对比实验结果，从结果可以看出，MRRID 描述子性能较好。但是由于 MRRID 描述子计算复杂，所需要的计算时间较长，构造一个 MRRID 描述子需要 16.8 ms，而构造一个 IWLIOP 描述子只需要 3.9 ms。因此，虽然性能方面 IWLIOP 描述子比文献[12]中的描述子性能略差，但是在计算速度方面 IWLIOP 描述子更具有优越性。

在性能方面，IWLIOP 描述子比 SIFT、HRI-CSLTP 和 LIOP 描述子要好，而且描述子维数更小，更利于后续利用；虽然 IWLIOP 描述子比 MROGH 和 MRRID 描述子性能稍差，但是在计算速度方面，IWLIOP 描述子比 MROGH 和 MRRID 描述子要快，将更有利于在如目标跟踪等对计算速度要求较高应用中的利用。因此，综合描述子性能、维数和计算速度，本文提出的 IWLIOP 描述子对图像局部区域具有更好的描述能力。

5　结束语

本文提出了一种顽健的图像局部特征区域描述子的构造方法，并将其应用到图像匹配中。该方法综合了全局信息和局部信息，利用局部像素值之

间的相互关系计算描述子，采用分段描述子计算方法和加权累加描述子构造方法，在充分利用局部信息的同时维数较小，而且由于阈值的设定使该算法对噪声的敏感性较小。IWLIOP 描述子维数较小，因此后续利用中计算量和存储空间更小。通过实验对标准测试图像进行评价，实验结果显示本文的描述子对各种几何和光照变换具有更好的不变性，对光照复杂的图像同样具有较好的顽健性。

参考文献：

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

[2] TUYTELAARS T, VAN GOOL L. Matching widely separated views based on affine invariant regions[J]. International Journal of Computer Vision, 2004, 59(1): 61-85.

[3] NISTER D, STEWENIUS H. Scalable recognition with a vocabulary tree[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. New York, NY, USA, 2006. 2161-2168.

[4] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features[J]. International Journal of Computer Vision, 2007, 74(1): 59-73.

[5] KE Y, SUKTHANKAR R. PCA-SIFT: A more distinctive representation for local image descriptors[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Washington, DC, USA, 2004. 506-513.

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

[7] BAY H, TUYTELAARS T, VAN GOOL L. Surf: Speeded up Robust Features[M]. Berlin Heidelberg. Springer-Verlag, 2006.

[8] HEIKKILA M, PIETIKAINEN M, SCHMID C. Description of interest regions with local binary patterns[J]. Pattern recognition, 2009, 42(3): 425-436.

[9] GUPTA R, PATIL H, MITTAL A. Robust order-based methods for feature description[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco, CA, USA, 2010. 334-341.

[10] TOLA E, LEPETIT V, FUA P. Daisy: An efficient dense descriptor applied to wide-baseline stereo[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2010, 32(5): 815-830.

[11] DAHL A L, AANAES H, PEDERSEN K S. Finding the best feature detector-descriptor combination[A]. International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission[C]. Hangzhou, China, 2011. 318-325.

[12] FAN B, WU F, HU Z. Rotationally invariant descriptors using intensity order pooling[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2012, 34(10): 2031-2045 .

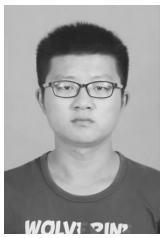
[13] WANG Z, FAN B, WU F. Local intensity order pattern for feature description[A]. IEEE International Conference on Computer Vision[C]. Barcelona, Spain, 2011. 603-610.

[14] Available online[EB/OL]. http://www.robots.ox.ac.uk/~vgg/research/affine/.

[15] 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.

[16] MIKOLAJCZYK K, SCHMID C. Scale & affine invariant interest point detectors[J]. International Journal of Computer Vision, 2004, 60(1): 63-86.

作者简介：



朱仁欢（1989-），男，安徽安庆人，安徽大学硕士生，主要研究方向为模式识别、图像特征提取等。

高清维[通信作者]（1965-），男，安徽合肥人，安徽大学教授、博士生导师，主要研究方向为数字信号处理、小波变换、多尺度几何分析、混沌与分形理论等。E-mail: qingweigao@ahu.edu.cn。

卢一相（1980-），男，安徽阜阳人，安徽大学博士生，主要研究方向为图像处理、模式识别等。

孙冬（1983-），男，安徽阜阳人，安徽大学博士生，主要研究方向为图像处理、模式识别等。