

基于 K 近邻一致性的特征匹配内点选择算法

肖春宝¹ 冯大政²

(西安电子科技大学计算机学院 西安 710071)¹

(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)²

摘要 宽基线图像特征匹配是计算机视觉应用中一项极具挑战性的工作。由于图像之间存在较大的差异,宽基线图像初始特征匹配的结果中不可避免地包含大量的外点。提出了 K 近邻一致性算法来实现从宽基线图像初始匹配结果中快速选出高可靠性的点对。该算法采用仿射不变的结构相似度来衡量两组 K 近邻特征点的结构相似性。K 近邻一致性算法采取由粗到精的策略,通过 K 近邻对应一致性检测和 K 近邻结构一致性检测两个步骤来选择内点。实验结果表明,提出的算法在查准率、查全率和运行速度等方面接近或优于当前几种最新的内点选择算法,可适用于存在大范围的视点、尺度和旋转变化的宽基线图像。

关键词 宽基线图像,特征匹配,内点选择,K 近邻,结构相似度

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.12.062

Inlier Selection Algorithm for Feature Matching Based on K Nearest Neighbor Consistency

XIAO Chun-bao¹ FENG Da-zheng²

(School of Computer Science and Technology, Xidian University, Xi'an 710071, China)¹

(National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)²

Abstract Feature matching for wide baseline images is an extremely challenging task in computer vision applications. A large number of outliers are inevitably included in the initial matching results due to significant changes between views of wide baseline images. An inlier selection algorithm called K nearest neighbor consistency (KNNC) was proposed to efficiently select matches with high reliability from initial feature matching results of wide baseline images. An affine-invariant structure similarity is utilized to measure the degree of structure similarity between two groups of K nearest neighboring features. Adopting the coarse-to-fine strategy, KNNC algorithm selects inliers by the processes of K nearest neighbor correspondence consistency checking and K nearest neighbor structure consistency checking. Experimental results show that the proposed algorithm approximates or surpasses several state-of-the-art inlier selection algorithms in performance on precision, recall and computational time, and is applicable to wide baseline images with large differences in viewpoint, scale and rotation.

Keywords Wide baseline image, Feature matching, Inlier selection, K nearest neighbor, Structure similarity

宽基线图像^[1,2]具有获取方便、存储量小、应用范围广等优点,在目标识别、运动分析、三维重构等领域中起重要作用。然而宽基线图像之间的较大差异也增加了特征匹配的难度。宽基线图像特征匹配是计算机视觉应用中一项极具挑战性的工作。特征匹配亟待解决的准确性和效率问题对于宽基线图像尤为突出。常规的特征匹配方法首先分别从每幅图像中提取特征,然后根据特征描述子^[3,4]对不同图像中的特征进行匹配。

当前流行的特征提取算法,如 SIFT(Scale Invariant Feature Transform)^[5]、SURF(Speed-up Robust Features)^[6]和 BRISK(Binary Robust Invariant Scalable Keypoints)^[7]等,可以从图像中提取大量局部不变性特征^[8,9]。由于图像之间存在因视角、缩放、旋转和光照等方面的变化而引起的差异,初

始匹配的结果中不可避免地包含大量的外点(错误的点对)。当前最常用的内点(正确的点对)选择方法是 RANSAC(Random Sample Consensus)^[10]。RANSAC 方法的缺点是当外点比例很高时效率变低,因而不适用于宽基线图像特征匹配。

目前已有一些针对宽基线图像(或高外点比例)特征匹配的内点选择算法。GLTCC(Global and Local Transformation Consistency Constraints)^[11]利用全局特征点的尺度与旋转一致性和局部邻近特征点的空间分布一致性来剔除宽基线图像初始匹配中的外点,具有较高的效率和准确性。Ishii 等^[12]提出一种联合基于特征和基于区域的匹配方法,利用 ASIFT(Affine-SIFT)和 POC(Phase-Only Correlation)实现宽基线图像的立体匹配。Wang 等^[13]提出一种基于轮廓的匹配方法,

到稿日期:2014-12-21 返修日期:2015-03-22 本文受国家自然科学基金项目(61271293)资助。

肖春宝(1976—),男,博士生,主要研究方向为计算机视觉,E-mail: xiaochunbao@163.com;冯大政(1959—),男,博士,教授,主要研究方向为雷达信号处理、智能信息处理。

可以获得高可靠性的点对,适用于无纹理的目标,但效率有待提高。TC(Topological Clustering)^[14]利用 SIFT 匹配点对的拓扑关系一致性,从宽基线图像初始特征匹配结果中选出具有较高正确率的匹配,但选出的内点数量偏少。KNN-TAR (Triangle Area Representation of K Nearest Neighbors)^[15]利用 KNN-TAR 描述子找出候选的外点,再通过迭代的方法根据局部结构和全局信息逐一检测并移除外点。KNN-TAR 可用于存在大的仿射变换、相似模式和低重叠区域的遥感图像,缺点是移除外点的过程计算量大、效率低,因而不适用于外点数量很大的情况。

针对上述算法中存在的选出的内点数量偏少、准确性和效率低下等问题,本文提出了 K 近邻一致性(K Nearest Neighbor Consistency, KNNC) 检测算法。KNNC 算法与 KNN-TAR 算法较为接近,基于区域面积比构造具有仿射不变性的相似度。KNNC 算法采用了一种简单有效的 K 近邻结构相似度来表示两组邻近特征点拓扑结构的相似性。KNNC 算法采取由粗到精的策略,通过 K 近邻对应一致性检测和 K 近邻结构一致性检测来选择内点,在查全率、查准率和运行速度方面均得到显著提高,可以较好地适应于宽基线图像。

1 K 近邻相似度

本文定义了两种相似度:K 近邻对应相似度和 K 近邻结构相似度,分别作为 K 近邻对应一致性检测和 K 近邻结构一致性检测的依据。

1.1 K 近邻对应相似度

设待匹配图像 I 和 I' 经过特征提取和初始特征匹配后得到的候选对应点对集为

$$H = \{(x_i, x'_i) | i=1, 2, \dots, m\} \quad (1)$$

其中, x_i 和 x'_i 表示特征点在图像中的位置, m 为候选对应点对的数量。候选对应点对 (x_i, x'_i) 的 K 近邻结构如图 1 所示, x_i 和 x'_i 用圆点表示, x_i 和 x'_i 的 K 近邻点用菱形表示。

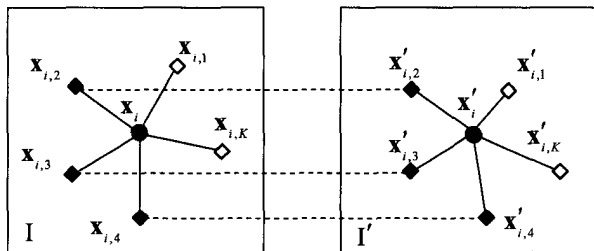


图 1 K 近邻结构示意图

设 x_i 和 x'_i 的 K 近邻点集分别为 P_i 和 P'_i , 其中

$$P_i = \{x_{i,j} | j=1, 2, \dots, K\} \quad (2)$$

$$P'_i = \{x'_{i,j} | j=1, 2, \dots, K\} \quad (3)$$

点对 (x_i, x'_i) 的 K 近邻对应一致性用存在于 P_i 和 P'_i 中的候选对应点对所占的比例来表示, 即

$$C_i = \frac{\text{Cardinality}(P_i \cap \psi(P'_i))}{K} \quad (4)$$

其中, $\psi(\cdot)$ 表示特征点集的对应点集, $\text{Cardinality}(\cdot)$ 表示集合的基数。

C_i 的值越高, 表明 P_i 和 P'_i 中存在对应关系的点所占的比例越大, (x_i, x'_i) 为内点的可能性也越大。在图 1 中, 存在对应关系的点用填充的菱形表示, 并且用虚线连接。不存在

对应关系的点称为单点, 用空心菱形表示。

1.2 K 近邻结构相似度

如果两幅图像中对应点对的两组 K 近邻点是相同场景的像, 其位置关系满足投影变换。由于 K 近邻点彼此之间的距离都很小, 它们来自场景中的一个较小的局部区域。假设该局部区域的范围远小于场景深度, 则两组对应的 K 近邻点的位置关系可以用仿射变换来近似。Alajlan 提出了 TAR (Triangle-Area Representation)^[16] 来表示形状边界的结构特征, 并用于形状匹配。TAR 具有仿射不变性, 对遮挡和非刚性变形有较高的鲁棒性。本文受 TAR 思想的启发, 提出了一种用于表示两组 K 近邻特征点结构相似性的相似度。

采用齐次坐标表示特征点, 设

$$x_i = [x_i \ y_i \ 1]^T \quad (5)$$

$$x'_i = [x'_i \ y'_i \ 1]^T \quad (6)$$

x'_i 和 x_i 之间存在仿射变换关系, 即

$$x'_i = \begin{bmatrix} A_{2 \times 2} & t_{2 \times 1} \\ 0^T & 1 \end{bmatrix} x_i \quad (7)$$

其中, A 是一个包含仿射变换参数的非奇异矩阵, t 是与平移参数有关的向量。

设特征点 x_i 和集合 P_i 中的两个点 $x_{i,j}$ 和 $x_{i,j+1}$ 不共线, 则三角形 $\Delta(x_i, x_{i,j}, x_{i,j+1})$ 的面积为

$$S_{i,j} = \frac{1}{2} |\det[x_i \ x_{i,j} \ x_{i,j+1}]|, j=1, 2, \dots, K \quad (8)$$

其中, \det 算子表示行列式运算, $|\cdot|$ 表示绝对值运算, 规定 $x_{i,K+1} = x_{i,1}$ 。

上述 3 个点经过仿射变换后得到的 3 个点分别为 x'_i , $x'_{i,j}$ 和 $x'_{i,j+1}$, 由式(7)和式(8)可计算出 $\Delta(x'_i, x'_{i,j}, x'_{i,j+1})$ 的面积为

$$S'_{i,j} = \frac{1}{2} |\det[A] \det[x_i \ x_{i,j} \ x_{i,j+1}]|, j=1, 2, \dots, K \quad (9)$$

由式(8)和式(9)可得对应的三角形面积比为

$$r_{i,j} = \frac{S'_{i,j}}{S_{i,j}} = 1/|\det[A]|, j=1, 2, \dots, K \quad (10)$$

由式(10)可知, 若点对 (x_i, x'_i) 及其两组 K 近邻点中不存在外点, 并且满足同一仿射变换, 则对应三角形的面积比相同, 其值与仿射变换的参数有关。

选定一个作为参照的三角形面积比, 如 $\Delta(x_i, x_{i,K}, x_{i,1})$ 和 $\Delta(x'_i, x'_{i,K}, x'_{i,1})$ 的面积比 $r_{i,K}$, 则其余三角形的相对面积比为

$$R_{i,j} = \begin{cases} r_{i,j}/r_{i,K}, & r_{i,j} \leq r_{i,K} \\ r_{i,K}/r_{i,j}, & r_{i,j} > r_{i,K} \end{cases}, j=1, 2, \dots, K-1 \quad (11)$$

$R_{i,j}$ 通过采用相对面积比, 消除了仿射变换的影响。

点对 (x_i, x'_i) 的 K 近邻结构相似度用相对面积比的均值来表示, 其计算公式为

$$\bar{R}_i = \frac{1}{K-1} \sum_{j=1}^{K-1} R_{i,j}, i=1, 2, \dots, m \quad (12)$$

\bar{R}_i 具有仿射不变性, 其值介于区间 $[0, 1]$ 。对应的两组 K 近邻的结构相似度越高, \bar{R}_i 越接近于 1。如果两个三角形的面积都非常小, 其面积比是不稳定的。因此规定面积小于 0.4 的三角形是无效的, 不参与面积比的计算。

2 KNNC 算法

本文基于特征点对的 K 近邻一致性, 提出了 KNNC 算法

来快速准确地检测出候选特征点对中的内点。该算法首先计算 $\{x_i | i=1, 2, \dots, m\}$ 的距离矩阵 $D_{m \times m}$ 和 $\{x'_i | i=1, 2, \dots, m\}$ 的距离矩阵 $D'_{m \times m}$, 其中元素 $D_{i,j}$ 表示 x_i 到 x_j 的距离, $D'_{i,j}$ 表示 x'_i 到 x'_j 的距离。分别找出 D 和 D' 的第 i 行中 $K+1$ 个最小的元素并按升序排序, 根据排序的列索引可以确定 x_i 和 x'_i 的 K 近邻集 P_i 和 P'_i 。

为了提高效率, KNNC 算法采用由粗到精的两个步骤实现内点选择。第一步进行 K 近邻对应一致性检测, 保留 K 近邻对应相似度大于阈值 T_C 的对应点对。 K 近邻对应一致性检测可以快速消除大部分外点, 减少后面步骤的计算量。第二步对通过 K 近邻对应一致性检测的候选点对进行 K 近邻结构一致性检测, 由每个点对的两组 K 近邻中存在对应关系的点构成的各三角形的面积计算出 K 近邻结构相似度, 保留结构相似度大于阈值 T_R 的点对。 KNNC 算法的流程如下。

Input: 候选对应点对集 H , 近邻数 K , K 近邻对应相似度的阈值 T_C , K 近邻结构相似度的阈值 T_R

Output: 内点集 H_1

initialize $H_1 = \emptyset$;

计算两幅图像的特征点的距离矩阵 D 和 D' ;

for $i=1$ to m do

 分别找出 D 和 D' 的第 i 行中的 $K+1$ 个最小元素并按升序排列, 将排序的列索引分别记为 B 和 B' ;

 根据 B 和 B' 找到 x_i 和 x'_i 的 K 近邻集 P_i 和 P'_i ;

 按式(4)计算 C_i ;

 if $C_i > T_C$ then

 移除 P_i 和 P'_i 中的单点;

 计算两组近邻点构成的各三角形的面积;

 按式(12)计算 R_i ;

 if $R_i > T_R$ then

$H_1 \leftarrow H_1 \cup (x_i, x'_i)$;

 end if

 end if

end for

3 实验结果与分析

实验所用的 8 组测试图片对如图 2 所示, 其中图 2(a) — 图 2(e) 所示的图片对取自 Oxford 数据集 (<http://www.robots.ox.ac.uk/~vgg/data/data-mview.html>), 图 2(d) — 图 2(f) 所示的图片对取自 Strecha 数据集 (<http://cvlabwww.epfl.ch/~strecha/multiview/denseMVS.html>), 图 2(g) — 图 2(h) 所示的图片对取自 INRIA 数据集 (<http://lear.inrialpes.fr/people/mikolajczyk/Database/>)。由于测试图片对存在较大的视角、尺度和旋转变换, 并且包含重复纹理, 对其进行准确的特征匹配具有一定的难度。实验程序采用 MATLAB R2011b 编写, 在 Pentium dual-core 2.7GHz CPU、4GB 内存的计算机上运行。

实验采用 SURF 算法来提取图像的特征, 利用余弦距离进行特征匹配, 最近邻距离比 (Nearest Neighbor Distance Ratio, NNDR)^[17] 的值设为 0.7。为了验证本文提出的算法的性能, 选择 KNN-TAR、GLTCC 和 TC 作为对比算法。几种算法的参数设置如下: KNN-TAR 算法设置 $K=7, \sigma=1, \alpha=5$; GLTCC 算法设置 $\tau_\sigma=1, \tau_\theta=0.5, \tau_{score}=0.7, K=7, \lambda=0.65$; TC 算法设置 $c=3, n_t=3$; KNNC 算法设置 $K=15, T_C=0.3, T_R=0.2$ 。

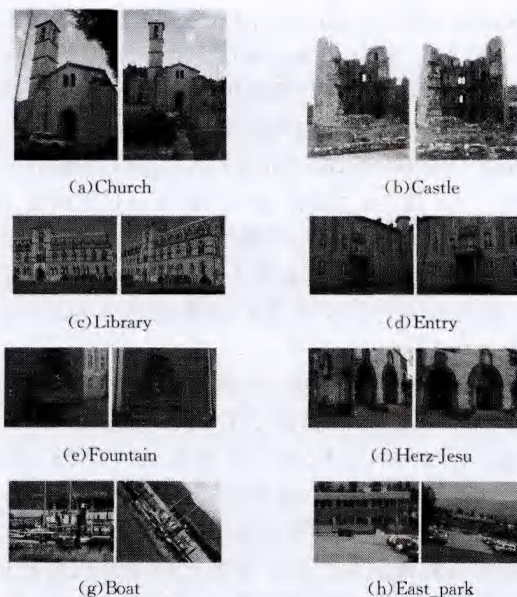


图 2 测试图片

本文采用查准率、查全率和计算时间 3 个指标对算法的性能进行评价。查准率和查全率的计算公式为

$$\text{查准率} = \frac{TP}{TP + FP} \quad (13)$$

$$\text{查全率} = \frac{TP}{TP + FN} \quad (14)$$

其中, TP 表示检测出的正确匹配的数量, FP 表示将错误匹配误检为正确匹配的数量, FN 表示将正确匹配误检为错误匹配的数量。

图 3 给出了图片对 Church 的初始匹配分别经过 4 种内点选择算法得到的匹配结果, 正确的匹配用实线连接, 错误的匹配用虚线连接。每幅图的小标题的括号内的两个数字分别表示正确匹配数/检测出的匹配数。从图 3 可以看出, KNNC 算法相对于其他 3 种算法可以获得更高的内点数量和查准率。

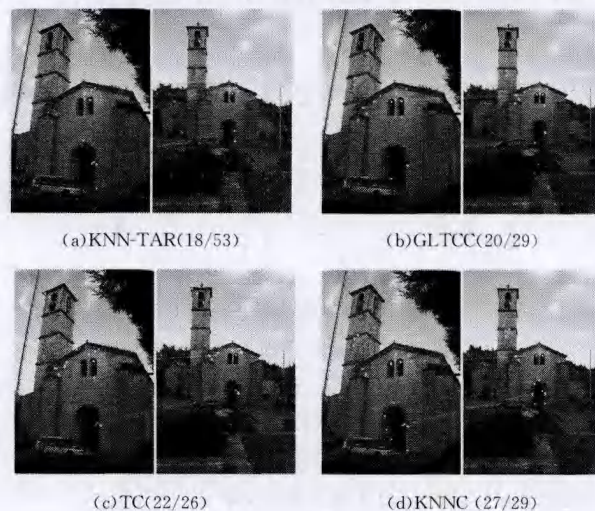


图 3 Church 匹配结果

实验的结果由对每组图片样本进行 200 次测试取平均得到。表 1 列出了 4 种内点选择算法对于每组测试图片对的查准率、查全率和计算时间, 最后一列是 8 组图片样本实验结果的平均值。初始匹配的内点数/候选点对数在表 1 的第一行中图片对名称下方的括号内标出。从表 1 可以看出, KNN-

TAR 算法耗时很多,其平均计算时间超过 28s。GLTCC、TC 和 KNNC 算法的平均计算时间较 KNN-TAR 算法有大幅的减少。GLTCC 算法的平均计算时间最少,只有 0.1506s。

KNNC 算法的平均计算时间与 GLTCC 算法接近,为 0.1786s。KNNC 算法的平均查准率为 95.36%,平均查全率为 63.50%,这两项指标明显优于其他 3 种算法。

表 1 内点选择算法性能比较

图片对 (内点数/候选点对数)	Church (44/164)	Castle (79/285)	Library (149/444)	Entry (494/1395)	Fountain (88/502)	Herz-Jesu (128/613)	Boat (432/692)	East_park (225/462)	平均值
KNN-TAR	查准率(%)	33.96	37.00	45.68	50.42	44.19	36.42	97.94	53.69
	查全率(%)	40.91	46.84	49.66	48.99	43.18	42.97	98.84	50.76
	时间(s)	2.5481	4.9965	10.4796	109.6819	15.3641	21.4145	44.0665	28.7964
GLTCC	查准率(%)	68.97	97.96	94.83	80.89	61.70	86.11	99.73	85.63
	查全率(%)	45.45	60.76	36.91	51.42	32.95	48.44	85.88	54.39
	时间(s)	0.0192	0.0350	0.0801	0.5871	0.0788	0.1222	0.1817	0.1506
TC	查准率(%)	84.62	90.00	52.48	78.29	55.26	56.76	94.24	75.01
	查全率(%)	50.00	34.18	35.57	40.89	23.86	16.41	53.01	40.24
	时间(s)	0.1313	0.5663	0.9464	12.3820	1.8265	3.2837	1.7010	2.7177
KNNC	查准率(%)	93.10	98.31	94.94	95.21	88.24	97.06	99.75	95.36
	查全率(%)	61.36	73.42	50.34	64.37	34.09	51.56	93.29	63.50
	时间(s)	0.0317	0.0497	0.1235	0.5784	0.1103	0.2032	0.1738	0.1786

结束语 本文基于特征点对的 K 近邻一致性,提出了 KNNC 算法来对宽基线图像初始匹配结果进行内点选择。该算法利用仿射不变的结构相似度,采取由粗到精的策略,通过 K 近邻对应一致性检测和 K 近邻结构一致性检测两个步骤来选择内点。KNNC 算法能从内点比例很低的初始特征匹配结果中快速选取高可靠性的对应点对,其效率不受内点比例变化的影响。实验结果表明,KNNC 算法具有较高的查准率、查全率和运行速度,可以适应宽基线图像之间大的视角、尺度和旋转变化的影响。

参 考 文 献

[1] Mayer H, Bartelsen J, Hirschmüller H, et al. Dense 3d reconstruction from wide baseline image sets [M] // Outdoor and Large-Scale Real-World Scene Analysis. Springer Berlin Heidelberg, 2012: 285-304

[2] Wu B, Zhang Y, Zhu Q. A triangulation-based hierarchical image matching method for wide-baseline images [J]. Photogrammetric Engineering & Remote Sensing, 2011, 77(7): 695-708

[3] Wu Jian, Ma Yue. Improved SIFT Algorithm [J]. Computer Science, 2013, 40(7): 270-272 (in Chinese)
吴建, 马跃. 一种改进的 SIFT 算法 [J]. 计算机科学, 2013, 40(7): 270-272

[4] Miksik O, Mikolajczyk K. Evaluation of local detectors and descriptors for fast feature matching [C] // 2012 21st International Conference on Pattern Recognition (ICPR). IEEE, 2012: 2681-2684

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

[6] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features [M] // Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417

[7] Leutenegger S, Chli M, Siegwart R Y. BRISK: Binary robust invariant scalable keypoints [C] // 2011 International Conference

on Computer Vision (ICCV). IEEE, 2011: 2548-2555

[8] Yang Heng, Wang Qing. A Novel Local Invariant Feature Detection and Description Algorithm [J]. Chinese Journal of Computers, 2010, 33(5): 935-944 (in Chinese)
杨恒, 王庆. 一种新的局部不变特征检测和描述算法 [J]. 计算机学报, 2010, 33(5): 935-944

[9] Sun Hao, Wang Cheng, Wang Run-sheng. A review of local invariant features [J]. Journal of Image and Graphics, 2011, 16(2): 141-151 (in Chinese)
孙浩, 王程, 王润生. 局部不变特征综述 [J]. 中国图象图形学报, 2011, 16(2): 141-151

[10] Harley R, Zisserman A. Multiple View Geometry in Computer Vision (2nd Edition) [M]. Cambridge: Cambridge University Press, 2003: 116-123

[11] Zhou H B, Zhang D Z, Chen C, et al. Discarding wide baseline mismatches with global and local transformation consistency [J]. Electronics letters, 2011, 47(1): 25-26

[12] Ishii J, Sakai S, Ito K, et al. Wide-baseline stereo matching using ASIFT and POC [C] // 2012 19th IEEE International Conference on Image Processing (ICIP). IEEE, 2012: 2977-2980

[13] Wang Z, Wang L. Wide-baseline correspondence from locally affine invariant contour matching [M] // Image Analysis and Recognition. Springer Berlin Heidelberg, 2011: 242-252

[14] Wang Y T, Zhang D Z, Tian J W. Discarding wide baseline mismatches via topological clustering [J]. Electronics Letters, 2008, 44(11): 670-671

[15] Zhang K, Li X, Zhang J. A Robust Point-Matching Algorithm for Remote Sensing Image Registration [J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(2): 469-473

[16] Alajlan N, El Rube I, Kamel M S, et al. Shape retrieval using triangle-area representation and dynamic space warping [J]. Pattern Recognition, 2007, 40(7): 1911-1920

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