

# 一种抗视角变换的 SURF 匹配算法

李 婕, 邓德祥, 石文轩, 颜 佳  
(武汉大学 电子信息学院, 武汉 430079)

**摘要:** 针对传统的 SURF 匹配算法在图像搜索、物体识别中对视角变化大的图像特征点无法匹配的缺陷, 提出了一种改进的 SURF 图像匹配算法。该算法利用模拟图序列的概念, 即由原始图像模拟各种可能的视角变换生成的一组图像序列, 对图像序列中的每幅图像进行 SURF 特征点提取; 剔除冗余特征点后, 将模拟图序列中的特征点映射到原始图像中, 这一过程增加了原始图像的特征点数量, 特征点描述向量改用原始图像和多个模拟图的 SURF 特征描述向量共同表示; 最后采用欧氏距离准则, 最终得到两幅图像的正确匹配对。实验结果表明, 本算法增加了特征点匹配对数, 改善了视角变换过大 (大于  $60^\circ$ ) 使得图像无法匹配的情况。

**关键词:** SURF; 视角变换; 模拟图序列; 图像匹配

中图分类号: TP391 文献标志码: A 文章编号: 1001-3695(2012)09-3536-04  
doi:10.3969/j.issn.1001-3695.2012.08.089

## Anti-viewpoint-changing image matching algorithm based on SURF

LI Jie, DENG De-xiang, SHI Wen-xuan, YAN Jia

(School of Electronic Information, Wuhan University, Wuhan 430079, China)

**Abstract:** This paper proposed an improved SURF algorithm for image matching to bridge the gap of conventional SURF matching algorithm in matching feature points when the viewpoint change was great in the applications of image searching and object identification. The improved algorithm simulated possible perspective transformations from the original image, generated a set of simulated image sequence, and extracted SURF feature points in the sequence image. Then, the feature points of the simulated images sequence were mapped to the original image. Composed the descriptors of the original image by the descriptors in original image and in the simulated images. Finally, used Euclidean distance criterion to get the matching point pairs. Experiment results show that this new algorithm increases the matched pairs of the image feature points, and overcomes the influence of viewpoint change (more than  $60^\circ$ ) in image matching.

**Key words:** SURF; viewpoint change; simulating image sequence; image matching

## 0 引言

随着图像匹配在三维重建、医学图像处理、图像搜索、物体识别等计算机视觉领域的广泛应用, 基于局部特征的匹配逐渐成为研究的热点。这些局部特征算子, 按属性可以分为基于点特征的检测方法, 如 Harris-Laplace、Hessian-Laplace<sup>[1,2]</sup>、SIFT<sup>[3]</sup>; 基于区域特征的检测方法, 如 Harris-affine、Hessian-affine<sup>[1,2]</sup>、MSER<sup>[4]</sup>等。文献[5,6]用仿射不变性、明暗变化、噪声等衡量各种局部算子, 得出 SIFT 是最出色的局部特征算子。2006 年, Bay 等人<sup>[7]</sup>提出的 SURF (speeded-up robust features) 特征算子在各个方面均接近或超越了 SIFT 的性能, 计算时间仅为 SIFT 的 1/3。虽然传统的 SURF 匹配算法有非常广泛的工程应用, 但对于图像搜索和物体识别中经常遇到的视角变换过大导致的图像不能匹配的问题还是显得无能为力。文献[8]提到的 3D 视角补偿的方法虽然在一定程度上解决了景象匹配中由于视角变换引起的难以匹配问题, 但此方案针对大视角变化图像的匹配并未作详细的阐述。本文的目的在于改进传统的 SURF 匹配算法, 使得算法在抗视角变换方面具有一定

的鲁棒性。借助尺度空间的思想, 阐述了视角变换与相机获得的图像之间的变换关系, 生成一组模拟图序列, 继而在模拟图序列中用传统的 SURF 特征点提取方法进行特征点提取和描述, 最后将模拟图的特征点映射到原始的图像中。用欧氏距离作为衡量两个特征点之间的相似性, 全局遍历所有的特征点。结果表明, 改进的 SURF 算法不仅继承了原始 SURF 在尺度、光照、旋转、噪声、缩放方面的鲁棒性, 而且能适应大于  $60^\circ$  的视角变换。

## 1 SURF 匹配算法及其缺陷

传统的 SURF<sup>[7]</sup> 匹配算法主要分为四个步骤:

a) 特征点检测。选取不同的箱式滤波器建立图像的尺度空间, 使用快速的 Hessian 矩阵检测每一层图像上的极值点, 对该点与邻近的  $3 \times 3 \times 3$  立体邻域内进行非极大值抑制, 大于邻近 26 个响应值的点称为 SURF 特征点。

b) 特征点主方向确定。以特征点为中心, 计算半径为  $6\sigma$  ( $\sigma$  为特征点所在的尺度值) 圆形邻域内的点在  $x, y$  方向的 Harr 小波响应, 并给这些响应值按距离赋予不同高斯权重系

收稿日期: 2011-12-15; 修回日期: 2012-02-14

作者简介: 李婕 (1984-), 女, 湖北宜昌人, 博士研究生, 主要研究方向为机器视觉 (jielonline@gmail.com); 邓德祥 (1961-), 男, 教授, 博导, 主要研究方向为空间图像处理、片上系统; 石文轩 (1983-), 男, 博士 (后), 主要研究方向为目标识别、机器视觉; 颜佳 (1983-), 男, 博士 (后), 主要研究方向为目标识别、机器视觉。

数,对加权后的 Harr 小波响应用直方图进行统计;再将圆形区域以每 5°为间隔,划分为 72 组,分别将每组 60°范围内的响应加起来形成一个新的矢量;遍历整个圆形区域,一共生成 72 个矢量,选择最长的矢量的方向为特征点的主方向。

c)特征点描述。以特征点为中心,将坐标轴旋转到主方向,按主方向选取  $20\sigma \times 20\sigma$  的正方形区域,将该窗口区域划分为  $4 \times 4$  个子区域,在每个子区域内计算 Harr 小波在  $y$  方向的响应  $d_x, d_y$ ,每个子区域在  $x, y$  方向的响应分别求和并生成了一个  $\sum d_x, \sum |d_x|, \sum d_y, \sum |d_y|$  的四维向量,这样每个特征点就有 64 维的描述向量,再进行归一化,形成最终的特征点描述子。

d)特征点匹配。采用基于最小欧氏距离来度量特征点间的相似性为

$$D = \sqrt{\sum_{k=1}^n (A_{ik} - B_{ik})^2} \quad (1)$$

其中:  $A_{ik}$  表示待匹配的第一幅图像的第  $i$  个特征点描述向量的第  $k$  维的值,  $B_{ik}$  表示待匹配的第二幅图像的第  $i$  个特征点描述向量的第  $k$  维的值,  $n$  表示特征点描述子是一个  $n$  维的描述向量。假设  $T_U, T_V$  分别为待匹配两幅图像的特征点集合,对于  $T_U$  中的任意一个特征点  $T_{U_i}$ ,若  $T_V$  中存在欧氏距离最小的两个特征点  $T_{V_j}, T_{V_j'}$ ,且  $T_{V_j} < T_{V_j'}$  ( $\lambda$  的范围一般为  $0.5 \sim 0.7$ ),则  $T_{V_j}$  认为是  $T_U$  的匹配对。最后采用穷举搜索算法找出所有的匹配对。

在步骤 a)中,尺度空间的建立使得算法对尺度缩放具有一定的鲁棒性,在立体区域里查找最大值去除了噪声对图像的影响;在步骤 b)中,特征点主方向的选取使得在匹配过程中特征点保持旋转不变性;步骤 c)中的描述向量归一化使得图像对光照强度不敏感,同时描述子按照不同的尺度用邻近的像素表示,使得图像对于小视角变换具有鲁棒性。但随着待匹配的两幅图像视角变换逐渐变大,传统的 SURF 匹配算法无法正确匹配。本文给出了两组视角变换较大的图像(图 1(a)的视角变换为  $60^\circ$ ,图 1(b)的视角变换为  $80^\circ$ ),匹配结果并不理想,两组图像的正确匹配对分别仅为 0 和 1。

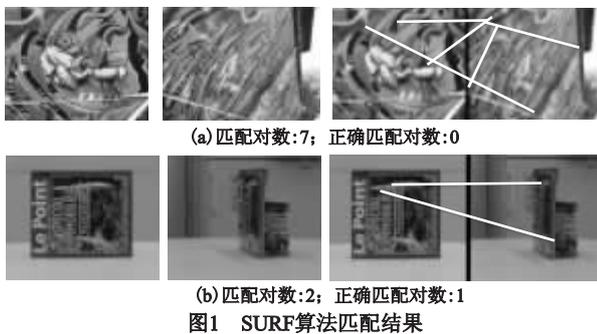


图1 SURF算法匹配结果

## 2 改进 SURF 算法介绍

改进的 SURF 算法有效地利用了尺度空间理论中模拟图序列的概念。模拟图序列是指由原始图像模拟各种可能的视角变换生成的一组图像序列,推导了原始图像与视角变化后的图像之间的数学关系,再利用实验得到生成模拟图序列所需要的抽样参数。

图像成像的数学过程如图 2 所示。假设  $u$  为相机所获取的数字图像,一般情况下可表示为

$$u = SGAu_0 \quad (2)$$

其中: $S$  表示相机 CCD 抽样函数, $G$  表示相机透镜的反走样高斯滤波函数, $A$  是平面仿射变换矩阵, $u_0$  表示相机正对物体时的无限平面。其中  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  作为一个线性仿射矩阵,也可以分解成

$$A = \lambda \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \quad (3)$$

式(3)可简写成

$$A = \lambda R(\varphi) T_t R(\phi) \quad (4)$$

其中: $\lambda, t, \varphi, \phi > 0$ ;图像平面旋转角度为  $\phi; \varphi \in [0, 180^\circ)$  为相机相对于光轴的旋转角; $\lambda$  为缩放参数; $t \in [1, 4\sqrt{2}]$  表示图像的倾斜程度。图 3 所示的相机视角变换可以更好地解释式(3),本文主要讨论与视角变换有关的参量  $t$ 。如图 3 所示,假设光轴与垂直于图像平面  $u_0$  的面所形成的夹角为  $\theta$ (以下简称为相机视角),当相机正对物体平面时  $t = 1, \theta = 0$ ;当相机视点移动时,  $\theta = \cos^{-1}(1/t), \theta \in [0^\circ, 80^\circ)$ 。因此,通过式(2)的数学变换模型可以模拟出任意相机从正对物体面所看到的图像平面  $u(x, y)$  到相机移动后所获取的图像  $u(A(x, y))$  的过程。

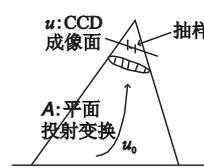


图2 相机成像模型

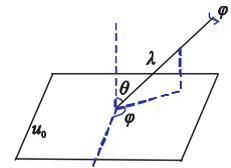


图3 相机视角变换

假设  $u(x, y) = SGAu_0$  和  $v(x, y) = SGBu_0$  是同一相机在不同视角拍摄  $u_0$  平面所获取的数字图像,  $A$  和  $B$  为两个线性映射矩阵。参照文献[10]得到仿射矩阵  $A, B$  存在以下关系:

$$BA^{-1} = H_\lambda R_1(\varphi_{12}) T_t R_2(\varphi_{12}) \quad (5)$$

其中: $\varphi_{12}$  为拍摄两幅图像时相机相对于光轴的旋转角之差,  $\phi_{12}$  则为图像平面旋转角度之差。令  $\tilde{A} = T_{\varphi}^x R_2 A, \tilde{B} = T_{\varphi}^x R_2(\frac{\pi}{2}) R_1^{-1} B, \tilde{u}, \tilde{v}$  为图像  $u, v$  经过平面变换和旋转后生成的模拟图,那么  $\tilde{u} = SG \tilde{A} u_0, \tilde{v} = SG \tilde{B} u_0$ , 因为  $T_t^x R_2(\frac{\pi}{2}) = R_2(\frac{\pi}{2}) T_t^x, T_t = T_t^y = H_t T_{t-1}^y$ , 那么

$$\tilde{B}^{-1} R_2(\frac{\pi}{2}) H_{\lambda} \tilde{A} = B^{-1} R_1 T_{\varphi}^y T_{\varphi}^x H_{\lambda} R_2 \tilde{A} = B^{-1} R_1 T_t^x R_2 A \quad (6)$$

将式(5)代入式(6)得到

$$\tilde{B} = R_2(\frac{\pi}{2}) H_{\lambda} \tilde{A} \quad (7)$$

因此,  $\tilde{u} = SG \tilde{A} u_0, \tilde{v} = SG H_{\lambda} \tilde{A} u_0$ , 模拟图  $\tilde{v}$  可以由模拟图  $\tilde{u}$  通过参数为  $\lambda \sqrt{t}$  的图像缩放得出。只需要选取的特征描述子能适应  $\lambda \sqrt{t}$  的缩放变换,  $\tilde{u}, \tilde{v}$  便具有相同的图像描述子。

以上说明了将视角变化的两幅图像演化为对原始图像进行抽样旋转和平面变换生成模拟图序列的过程。大量实验证明,  $\Delta t = \sqrt{2}, \Delta \phi = \frac{72^\circ}{t}$  为最合适的生成模拟图像的抽样参数。

## 3 改进的 SURF 算法流程

改进后的 SURF 算法引入了模拟图序列的概念,在一定程度上消除了视角变换对图像匹配的影响。其具体步骤如下:

a)生成视角变换模拟图序列。设要匹配的图像为  $u, v$ , 原

始图像的  $t = 1$ , 以  $\Delta t = \sqrt{2}$  进行横坐标  $x$  方向的图像抽样, 则生成的模拟图序列的  $t = \sqrt{2}, 2, \dots, 4\sqrt{2}$ 。数字图像沿坐标轴进行抽样前应作一个高斯卷积滤波, 高斯卷积的标准偏差  $c\sqrt{t^2 - 1}$ , 参照 SIFT<sup>[3]</sup>, 此处  $c = 0.8$ , 生成图像  $u_1$ 。在图像  $u_1$  基础上再进行  $\Delta\phi = \frac{72^\circ}{t}$  的平面旋转仿射, 在生成的模拟图序列  $u_{11}, u_{12}, \dots$ , 中,  $\phi = k\frac{72^\circ}{t}$ , 其中  $k = 1, 2, \dots$ , 且  $k\frac{72^\circ}{t} < 180^\circ$ 。最后生成模拟图序列  $U$  表示为

$$U = \bigcup_{i=1}^n (u_i \bigcup_{j=1}^m u_{ij}), V = \bigcup_{i=1}^1 (v_i \bigcup_{j=1}^k v_{ij}) \quad (8)$$

其中:  $u_i, v_i$  表示第  $i$  次纵坐标的抽样图像;  $u_{ij}, v_{ij}$  表示第  $i$  次纵坐标的抽样的基础上第  $j$  次旋转。本文在实验中取  $n, l = 3, m, k = \text{round}(5t)$ 。对于视角变换过大的图像, 应该获得更密集的模拟图序列, 这时  $n$  可以取 5。

b) 在生成的图像序列  $U, V$  中, 分别提取 SURF 特征点, 原始图像旋转形成的新图像序列在 SURF 特征点提取时会导致一部分特征点落在原始图像的边缘或原始图像之外, 应将这一部分特征点剔除。

c) 假设原始图像  $u$  用 SURF 算法得到特征点集合  $T_u$  为

$$T_u = \bigcup_{i=1}^n (L_i, M_i) \quad (9)$$

其中:  $L_i$  表示第  $i$  个特征点的位置向量,  $M_i$  表示第  $i$  个特征点的 64 维描述向量。依次在图像序列中计算 SURF 特征点, 然后再将序列图像 SURF 特征点映射到原始图像  $u$  中, 将式 (9) 改写为

$$T_u = \bigcup_{i=1}^n [L_i, (\bigcup_{j=1}^m M_{ij})] \quad (10)$$

其中:  $L_i$  依然表示第  $i$  个特征点的位置向量,  $M_{ij}$  表示在第  $i$  个特征点在第  $j$  张图像的 64 维特征描述向量。同理, 得到第二幅图像的特征点集合  $T_v$ 。

d) 采用传统 SURF 匹配算法的步骤 d) 进行特征点匹配, 得到  $T_u, T_v$  中的匹配对集合  $M_w$ 。

### 4 实验结果及分析

评价基于特征点图像匹配算法的匹配性一般采用 recall 和 1-precision 曲线<sup>[3,4]</sup>。定义为

$$\text{recall} = \frac{CM}{CP}, 1\text{-precision} = \frac{FM}{FM + CM}$$

其中:  $CM$  表示找到的匹配点对中正确的匹配对数目,  $CP$  表示两幅图像中存在正确匹配的特征点对数目,  $FM$  表示匹配点对中错误匹配的数目。在图像匹配的实际应用中, 往往设定参数使 1-precision 尽量小。本文实验中调试特征匹配参数, 控制参数在 0.1 左右。

下面比较传统 SURF 算法和本文算法得到的图像匹配对数。实验 1 采用图 1 所示的传统 SURF 匹配算法无法正确匹配的两组图像, 用本文算法匹配的结果如图 4 所示, 可以看出正确的匹配点在数量上分别从 0 对增加到了 25 对和 49 对。实验 2 采用了数据库<sup>[11]</sup>中的一组图像验证本文算法的有效性。取拍摄视角分别为  $0^\circ \sim 80^\circ$  的图片 (每  $10^\circ$  取一张图片), 采用视角为  $0^\circ$  图像分别与之匹配。生成的匹配效果如图 5 所示。图 6 展示了在 1-precision 尽量小的情况下, 分别用传统 SURF 算法和本文算法得到的图像特征点匹配对数

比较。由于图像本身纹理的差异, 不能定量地分析出本文算法得到的匹配对数是 SURF 算法的多少倍。但从图 5、6 可以明显看出, 改进的 SURF 算法除了得到的匹配对数远高于传统 SURF 算法外, 对于视角变换大 (大于  $60^\circ$ )、SURF 匹配算法无法正常匹配的图像能生成有效的匹配对。因此, 针对基于特征的物体识别、图像检索的应用中, 视角变化大的图像无法被检测出来的问题, 本文算法给出了比较好的解决方案。

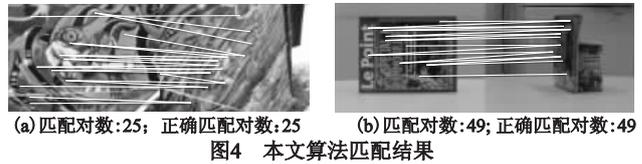


图 4 本文算法匹配结果

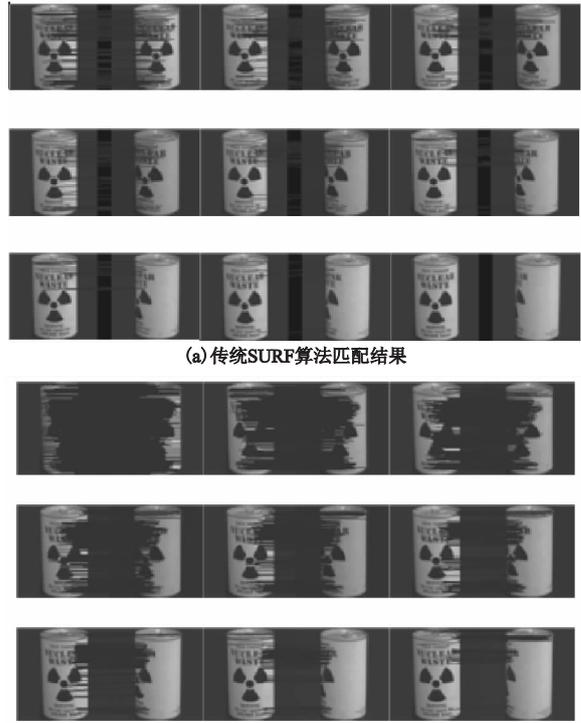


图 5 算法匹配结果

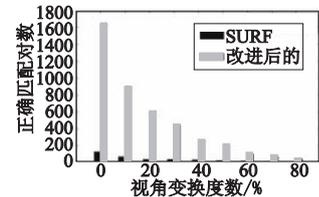


图 6 SURF 算法和改进后 SURF 算法的匹配对数比较

为了证明改进后的算法保留了传统 SURF 算法在光照、缩放、旋转、模糊方面的鲁棒性, 本文实验 3 采用数据库<sup>[12]</sup>的图片来进行测试。文献 [12] 中有三组经过不同程度光照、缩放、旋转、模糊后的图像。每组有一幅原始图像和五幅变换图像, 第一幅变换图像为原始图像经过光照、缩放、旋转、模糊变换中任意一种变换后生成的图像, 后四幅变换图像是在第一幅变换图像的基础上继续进行类型相同而幅度不同的图像变换所生成的变换图像。然后用传统 SURF 匹配算法和本文匹配算法分别对原始图像和经过变化的图像进行特征匹配, 如图 7 所示, 横坐标的值代表数据库中第几幅变换图像, 实验结果证明改进的算法保留了传统 SURF 算法在光照、缩放、旋转、模糊方面的鲁棒性。

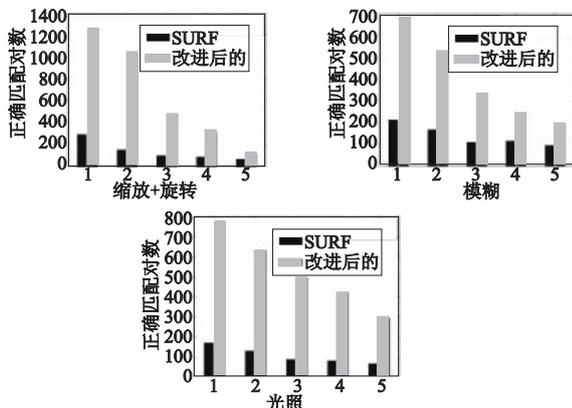


图7 SURF算法和改进后SURF算法在旋转、缩放、模糊、光照变化下匹配对数

## 5 结束语

本文针对原始 SURF 算法对于视角变换较大的图像不能进行有效匹配的缺陷,提出了一种改进的 SURF 算法:首先生成一组模拟图序列,对序列中的每一幅模拟图分别作 SURF 算法特征点提取;然后映射到原始图像,此时原始图像中的特征点数目明显增多;之后利用欧氏距离作为衡量特征点相似性的准则;最后遍历所有的特征点,得到最终的匹配对。实验证明改进的 SURF 算法在保证 SURF 算法原有的尺度、光照、旋转不变的基础上还能有效地抵抗视角变换。

本文算法的不足在于,每幅模拟图都进行特征提取,导致映射到原始图像的特征点存在冗余。怎样有效地消除冗余以及怎样提高算法的运算速度是进一步需要研究的问题。

### 参考文献:

[1] MIKOLAJCZYK K, SCHMID C. An affine invariant interest point detector[C]//Proc of the 7th European Conference on Computer Vision. London: Springer-Verlag, 2002:128-142.

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

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

[4] MATAS J, CHUM O, URBAN M, *et al.* Robust wide-baseline stereo from maximally stable extremal regions [J]. Image and Vision Computing, 2004, 22(10):761-767.

[5] MIKOLAJCZYK K, TUYTELAARS T, SCHMID C, *et al.* A comparison of affine region detectors[J]. International Journal of Computer Vision, 2005, 65(1):43-72.

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

[7] BAY H, TUYTELLARS T, Van GOOL L. SURF: speeded up robust feature[C]//Proc of the 9th European Conference on Computer Vision. Berlin: Springer-Verlag, 2006:404-417.

[8] 陈冰,赵亦工,李欣.基于快速鲁棒性特征的景象匹配[J].系统工程与电子技术, 2009, 31(11):2714-2718.

[9] MOREL J M, YU Guo-shen. ASIFT: a new framework for fully affine invariant image comparison[J]. SIAM Journal Imaging Sciences, 2009, 2(2):438-469.

[10] KE Yan, SUKTHANKAR R. PCA-SIFT: a more distinctive representation for local image descriptors[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2004:506-513.

[11] Columbia University. COIL-100[EB/OL]. [2011-05-08]. <http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php>.

[12] MIKOLAJCZYK K. Affine covariant features[EB/OL]. (2007-06-15)[2011-05-08]. <http://www.robots.ox.ac.uk/~vgg/research/affine/>.