

基于仿射迭代模型的特征点匹配算法

邓宝松 宋汉辰 杨冰 吴玲达

(国防科学技术大学多媒体研究开发中心, 长沙 410073)

摘要 图像序列中的特征点匹配是计算机视觉中的一个基本问题,也是目标识别、图像检索以及3维重建等问题的基础。为了提高图像匹配的精度,提出了一种针对两幅图像的高精度特征点自动匹配算法。该算法首先分析并提出两幅图像中相应特征点的邻域窗口之间的单应映射可以用仿射变换模型来近似;然后通过快速的基于仿射变换模型的迭代优化方法,不仅估计并矫正了相应邻域窗口之间的透视畸变,同时还补偿了在特征点检测阶段对相应特征点的定位误差,从而使匹配结果达到子像素级精度;最后通过真实图像的实验以及与现有算法的比较结果表明,该算法不仅得到了更多的匹配关系,还提高了特征点匹配的精度。

关键词 特征点匹配 单应映射 仿射变换 定位误差补偿 对极几何

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)04-0678-06

Feature Point Matching Based on Affine Iterative Model

DENG Bao-song, SONG Han-chen, YANG Bing, WU Ling-da

(Multimedia Research & Development Center, National University of Defense Technology, Changsha 410073)

Abstract Feature point matching is a key problem of computer vision and is frequently used in object recognition, image retrieval and 3D reconstruction and so on. In this paper, an accurate feature points matching method for two-frame images was proposed. Since it was proved that the homography between two windows of corresponding feature can be geometrically approximated by an affine transformation model. The projective distortion of windows of corresponding feature was estimated and rectified by a fast iterative scheme based on the affine transformation model. At the same time, the location error of corresponding feature points produced at the feature detection stage was compensated using the estimated affine parameters. The matching results of corresponding feature points can achieve sub-pixel precision, which effectively improve the precisions of the final epipolar geometry. Experimental results of real images and the comparisons with other methods strongly demonstrate the validity and accuracy of the algorithm.

Keywords feature point matching, homography, affine transformation, location error compensation, epipolar geometry

1 引言

众所周知,特征点匹配是计算机视觉中的经典问题,其在自动导航、目标识别、3维重建以及图像数据库检索等许多领域都有着重要的研究和应用价值。然而目前这一问题并没有被很好地解决,其主要原因包括成像过程中发生的透视畸变、亮度变化以及场景的相互遮挡等^[1,2],这些因素都给特征点

的匹配带了一定的困难。

现有匹配算法大部分是基于图像亮度相关性的^[2-4],其前提是假设同一场景点在不同图像上所成的像点是相似的,且与其邻域内的场景点所成像点的相对位置保持不变,然而这些性质在透视成像过程中并不能得到很好地保持。由于图像间存在较大的透视畸变,这类直接匹配方法很难自动矫正相应特征点邻域窗口(以下称邻域窗口)的局部形状,再加上两组待匹配特征点集之间真正可用的不变量

基金项目:国家自然科学基金项目(60473117)

收稿日期:2005-09-07;改回日期:2006-01-18

第一作者简介:邓宝松(1978~),男,2001年获国防科技大学工学学士学位,现为国防科技大学多媒体研究开发中心博士研究生。主要研究方向为计算机视觉、虚拟现实及可视化。E-mail: dbs310@163.com

很少,因此匹配结果并不尽人意。后来的研究方法其主要思想是先对相应邻域窗口之间的几何变换进行建模,然后通过求解模型参数来进行两个邻域窗口的畸变矫正。最初矫正使用的是相似变换模型,包括旋转^[5]和比例放缩^[6-8];后来扩展为仿射变换模型^[2,9-11],显然,仿射变换包含相似变换,它是相应邻域窗口之间透视畸变的一个很好近似,这一点在本文第 2 节中还会详细描述。然而,这些基于仿射变换模型的方法大多侧重于仿射不变性特征的提取和表示,而没有针对提高匹配结果的精度。此外,由于像素的量化误差以及特征点检测算法本身的弊端,致使相应特征点的定位误差较大,不能满足高精度应用的需要。

本文首先利用 Harris 算法^[12]分别在两幅图像中检测得到两组特征点集,然后通过区域搜索和迭代优化来自动估计并矫正相应邻域窗口间的透视畸变,同时补偿检测阶段相应特征点间的定位误差,使得匹配精度真正达到子像素级,从而大大提高了对极几何关系的估计精度。真实图像的实验结果以及与已有算法的比较验证了本文算法的正确性和高精度。

2 图像几何变换模型

如果将光滑场景表面上以点 P_0 为中心的局部邻域面片 G 近似认为是平面的话,其在两幅图像上所成的像之间满足 2 维平面上的单应映射(用 3×3 单应矩阵 H 表示),该变换将第 1 幅图像 I 上的一个圆形窗口变换到第 2 幅图像 \hat{I} 上的椭圆形窗口(如图 1 所示)。 H 由摄像机的内外参数和平面方程唯一确定,由 H 就可以实现相应邻域窗口之间的透视矫正。

设第 1 幅图像上特征点 p_0 的邻域窗口内某一图像点的坐标为 $p = (x, y)^T$ (以 p_0 为原点),其对应点在第 2 幅图像上的坐标为 $\hat{p} = (\hat{x}, \hat{y})^T$ (以 \hat{p}_0 为原点),则二者之间的映射关系可以表示为

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ 1 \end{bmatrix} \sim H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{v}^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

其中, \sim 表示齐次向量之间的相等; A 为 2×2 矩阵,表征邻域窗口之间的仿射变换;而 $\mathbf{v} = (v_1, v_2)^T$ 是 2 维向量,表征透视形变的大小。在实际应用中,由于 v_1 和 v_2 非常小,可以被忽略,而直接用仿射变换模

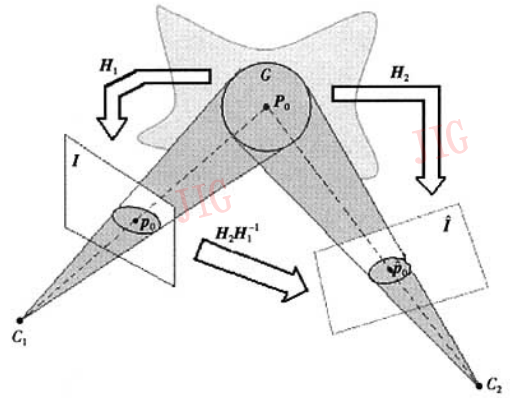


图 1 光滑场景表面面片在两幅图像上所成像之间的几何变换模型

Fig. 1 The homography between corresponding feature windows of the smooth patch on scene surface

型来近似表示这种映射关系^[9,13],即 $\hat{p} = Ap$ 。

如引言中所介绍,在第 2 幅图像 \hat{I} 上检测出的特征点 \hat{p}_0 并非与点 p_0 精确对应,二者之间存在着一定的偏移。现有的很多直接匹配算法中都忽略了这一偏移向量 t ,而直接将检测得到的特征点作为匹配结果的坐标,这在实际中是不合适的,而且有时甚至会导致较大的误差。考虑这种因素的影响,相应邻域窗口之间的映射关系可以进一步修正为

$$\hat{p} = Ap + t \quad (2)$$

3 特征点匹配

特征点匹配的的目的就是在两组待匹配特征点集中分别找到一个子集,使得两个子集之间存在一一映射,即映射两端都对应着同一场景空间点,而其依据就是二者及其邻域在亮度上的相似性。有很多的准则用来衡量两个邻域窗口的相关程度,常用的有规范化交叉相关系数(normalized cross correlation, NCC)^[4]和亮度差值的平方和(sum of squared difference, SSD)^[14],而这两个度量准则在一定程度上也是一致的。本文采用首先最小化两个邻域窗口中对应像素亮度差值的平方和,然后再用规范化交叉相关系数判断两个窗口相似程度的方法,即对两幅图像上的每一对相应特征点及其邻域窗口,找到合适的仿射变换参数,使得下式最小。

$$\varepsilon_{\min} = \min_{A, t} \iint_{p \in W} \|I(p) - \hat{I}(Ap + t)\|^2 dp \quad (3)$$

其中, W 表示第 1 幅图像上的邻域窗口, $I(p)$ 表示图像点 p 处的亮度值。如果能找到最优的仿射变换参数 A^{best} 和 t^{best} , 使得式 (3) 达到最小, 且此时两邻域窗口的规范化交叉相关系数大于给定阈值 δ_{max} , 则认为两个特征点是匹配的, 否则认为其不是由同一场景点投影所形成。

3.1 仿射迭代算法

本文采用迭代方法来最小化式 (3), 即首先将其在初始参数 $A = E_{2 \times 2}$ 和 $t = 0$ 处进行一阶泰勒展开, 即

$$\hat{I}(\hat{p}) = \hat{I}(Ap + t) \approx \hat{I}(p) + (\nabla I)^T \frac{\partial \hat{p}}{\partial a} (a - a_0) \quad (4)$$

其中, $a = [a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}, t_x, t_y]^T$ 为仿射变换的参数向量, $a_0 = [1, 0, 0, 1, 0, 0]^T$; $\nabla I = [I_x, I_y]^T$ 表示图像 I 在点 p 处的亮度沿水平和竖直方向上的导数, 在实际应用中, 为了加快算法的收敛速度, 并提高其稳定性, 该值可用两个邻域窗口内相应位置的图像亮度导数进行加权得到。令式 (3) 中的亮度差等于零, 并将式 (4) 代入即可以得到

$$[I(p) - \hat{I}(p)] + (\nabla I)^T \frac{\partial \hat{p}}{\partial a} a_0 \approx (\nabla I)^T \frac{\partial \hat{p}}{\partial a} a \quad (5)$$

在式 (5) 中未知参数为 a , 进一步整理可得到

$$[I(p) - \hat{I}(p)] + xI_x + yI_y \approx F(p)^T a \quad (6)$$

其中, $F(p)^T = [xI_x, yI_x, xI_y, yI_y, I_x, I_y]$, 将式 (6) 两边同乘以 $F(p)$, 并将两个窗口内的像素值进行加权即可以得到

$$Ha = h \quad (7)$$

$$\text{上式中 } h = \iint \omega(p) F(p) (I(p) + xI_x + yI_y) dp,$$

$H = \iint \omega(p) F(p) F(p)^T dp$, 其中 $\omega(p)$ 表示窗口内像素的权重函数, 由于可以通过该函数增大窗口中心像素的重要性, 因此合适的 $\omega(p)$ 能够大大加快算法的收敛速度。从式 (7) 可以直接求得 a , 以便更新所有的变换参数, 再以此为初值进行下一次迭代。在迭代过程中, 可以用单位矩阵 $E_{2 \times 2}$ 和零平移向量 0 作为初始值, 而由于相邻特征点的变换矩阵具有很大的相似性, 则可以将上一对正确匹配的变换矩阵作为下一对待匹配特征点的初始值, 从而可加快算法的收敛速度。

由于邻域窗口内的图像亮度变化相对于仿射变换参数而言, 具有极强的非线性特性, 且迭代过程中必然存在较多的局部极值点, 因此一味地增加迭代次数也并不能确保迭代过程能够跳过局部极值点,

也就是说, 迭代对相应邻域窗口的矫正作用会随着迭代次数的增加而迅速减小。因此, 本文在算法的实现过程中是采用统一的迭代次数 4, 如果在 4 次迭代之后两个特征点的规范化交叉相关系数 C_{normal} 仍没有大于给定的阈值 δ_{max} , 则认为二者不匹配。

3.2 定位误差补偿

由于通过特征点检测算法获得的两幅图像上的相应特征点不能实现空间上的精确对应, 因此这在很多高精度的应用中常常不能满足实际需要。而通过 3.1 节中的方法获得了最佳仿射变换参数 A^{best} 和 t^{best} 后, 就可以对这种定位误差进行精确的补偿。

图 2(a) 表示第 1 幅图像上的邻域窗口, 其特征点 p_0 位于窗口中心 o , 图 2(b) 为第 2 幅图像上的相应邻域窗口, 其检测的特征点位于窗口中心 \hat{o} 处, 而与 p_0 精确对应的像点 \hat{p}_0 并不位于窗口中心, 此时需要的就是偏移向量 d 。由于经过仿射变换后, \hat{p}_0 变换到窗口中心 (即原点 \hat{o} 处), 从而与图 2(a) 窗口实现了精确匹配 (如图 2(c) 所示), 由于 $A^{\text{best}} d + t^{\text{best}} = 0$, 从而有

$$d = - (A^{\text{best}})^{-1} t^{\text{best}} \quad (8)$$

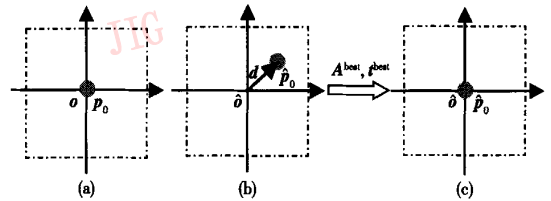


图 2 相应特征点的定位误差补偿示意图

Fig. 2 Compensation of location error for corresponding feature points

然后用 d 对第 2 幅图像上的特征点的检测结果进行误差补偿, 就可以大大提高相应特征点的定位精度, 即使两点之间的差别很大, 只要相距不是太远 (一般要求小于邻域窗口宽度的 $1/6$, 如果差距过大, 则很难保证两个特征点均能够收敛到窗口中心), 那么都能收敛到同一场景点, 这还可以在在一定程度上弥补那些未能检测到的特征点。

图 3 列举出了 3 组对应特征点的迭代过程和匹配结果, 其中第 1、2 列分别为两幅图像上相应特征点的邻域窗口, 第 3 至 6 列为第 2 列邻域窗口的 4 次迭代过程, 图像下面的数字为其与第 1 列邻域窗口之间的规范化交叉相关系数。由此可以看出, 经过仿射迭代和矫正, 相关性大大提高了。由迭代匹

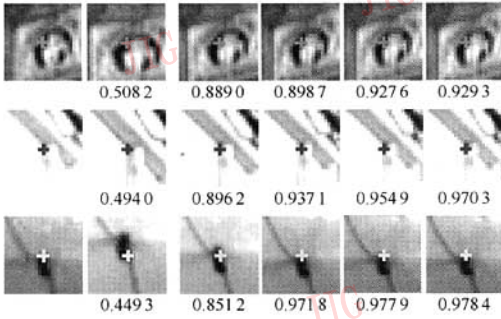


图 3 3 组相应邻域窗口的迭代和匹配过程
Fig. 3 The iteration and matching process of corresponding feature windows

配结果可见,虽然第 1 组的两个邻域窗口之间存在较大的透视畸变,第 2 组的两个邻域窗口之间存在较大的比例放缩,特别是第 3 组的两个邻域窗口之间存在着较大的定位误差,但这些都得到了很好的矫正和补偿,也充分显示了本文算法的鲁棒性。

3.3 算法实现流程

由于两幅图像的视点相近,即相应图像点之间的视差不会很大,因此给定第 1 幅图像上一个特征点,只需搜索第 2 幅图像上以该点坐标为圆心半径为 r 的区域内的特征点,如果二者的相关性系数大于给定的较小阈值 δ_{\min} ,并且按照 3.1 节和 3.2 节中的迭代矫正方法能够收敛到相关系数大于较大阈值 δ_{\max} ,则将其并入到匹配结果的集合中。

设第 1 幅和第 2 幅图像中检测出来的特征点集合分别为 $\{p_{1,i} | i = 1, \dots, n_1\}$ 和 $\{p_{2,j} | j = 1, \dots, n_2\}$,两个特征点 $p_{1,i}$ 和 $p_{2,j}$ 坐标之间的平面距离用 $\|p_{1,i} - p_{2,j}\|$ 表示,则两邻域窗口间的相关系数用 $C_{\text{normal}}(p_{1,i}, p_{2,j})$ 表示,图 4 给出了匹配算法的流程图,图中 M 和 S 均表示集合,其中 M 表示生成最终匹配点对的集合。

4 实验结果及分析

为了验证本文算法的性能,用 Matlab 语言实现了本文算法,并对多种类型的测试图像对进行了实验,结果均表现出很好的鲁棒性和很高的匹配精度。在实验过程中,邻域窗口直径大小 $w = 25$ pixels,迭代次数 $N = 4$,相关性阈值 $\delta_{\min} = 0.35, \delta_{\max} = 0.88$,特征点搜索区域窗口的半径 r 可视图像和视差大小而

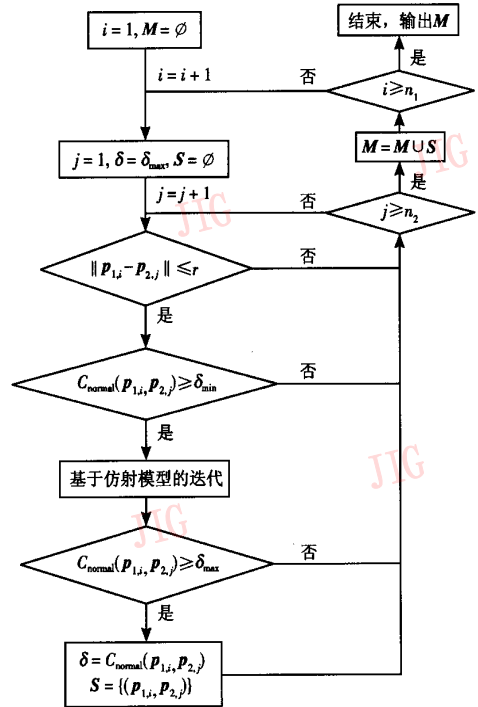


图 4 特征点匹配算法流程图
Fig. 4 The flow chart for feature point matching

定,一般取图像宽和高的最大值的 $1/6 \sim 1/4$ 。

由于图像特征点之间的自相似性,致使初始的匹配结果中总会存在一些误匹配,本文采用 RANSAC 方法^[15]和规范化 8 点算法^[16]来估计基本矩阵 F ,首先得到了初始对极几何关系,然后根据对极距离约束去除可能存在的误匹配;最后用所有正确的匹配重新估计基本矩阵作为最终的对极几何关系,并用平均对极距离 (average epipolar distance, AED) 作为比较算法性能优劣的误差度量,其定义如式(9)所示

$$\bar{d} = \sqrt{\frac{1}{2n} \sum_{k=1}^n (d^2(p_{1,k}, F^T p_{2,k}) + d^2(p_{2,k}, F p_{1,k}))} \quad (9)$$

其中 $(p_{1,k}, p_{2,k}), k = 1, \dots, n$ 表示所有正确匹配的点的序号, $d(p_{1,k}, F p_{2,k})$ 表示点 $p_{1,k}$ 到相应极线 $F p_{2,k}$ 的对极距离^[16]。

图 5 给出了牛津大学 Visual Geometry Group 提供的 3 组测试图像的实验结果,图 5(a) 是 House 序列中的第 0 幅和第 2 幅,两幅图像之间存在由于场景的旋转而产生的较大透视畸变,初始匹配结果共得到 142 对匹配点,其中正确匹配 117 对;图 5(b) 图像对中存在着由于镜头拉伸而导致的比例放缩,初始匹配结果共得到 173 对匹配点,其中正确匹配

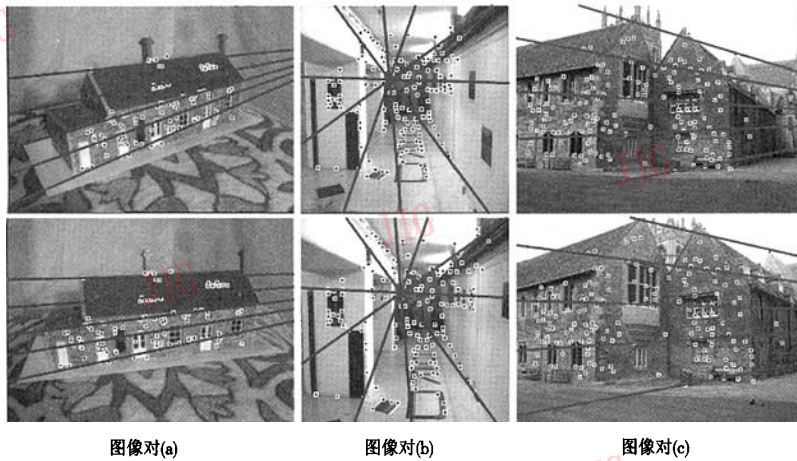


图 5 用本文算法对 3 组图像对的匹配结果

Fig. 5 Matching results of three pairs of images using the proposed method

150 对;图 5(c)图像对中同时存在透视畸变和比例放缩,初始匹配结果共得到 196 对匹配点,其中正确匹配 159 对。

表 1 给出了传统的基于窗口的直接匹配算法^[4]的平均对极距离以及本文算法进行定位误差补偿之前与之后的平均对极距离的比较,由表 1 可以看出,通过仿射迭代,在得到更多匹配特征点的同时,精度也提高了,再经过定位误差补偿,可使平均对极距离进一步减小,这些都有力地证实了本文算法的有效性和高精度性。

表 1 几种算法平均对极距离的比较

Tab.1 Comparisons of average epipolar distance with traditional method

	图像对(a)		图像对(b)		图像对(c)	
	正确匹配点数	平均对极距离	正确匹配点数	平均对极距离	正确匹配点数	平均对极距离
直接匹配算法	67	2.462 2	83	2.873 8	93	2.635 1
定位误差补偿前	117	1.593 1	150	1.757 5	159	1.002 8
定位误差补偿后	117	0.258 5	150	0.230 9	159	0.165 4

5 结 论

本文提出了一种基于仿射迭代模型的高精度特征点匹配算法,即在分析相应邻域窗口之间的单应映射可以用仿射变换模型表示之后,提出了基于仿

射迭代模型的特征点匹配方法,用于估计相应邻域窗口之间的仿射变换参数,以便对窗口的形状进行透视矫正,同时补偿了检测阶段相应特征点的定位误差,实验结果表明,本文算法的匹配结果具有很强的鲁棒性和很高的精度。

同时,笔者注意到在图 5(a)中,由于存在较大的旋转,使得地面上没有找到任何一对匹配,类似地,由于随着两幅图像之间基线以及视角的增大,不仅相应邻域窗口之间的透视畸变会更加明显,而且图像亮度也会发生较明显变化,此时如果仍采用直接优化的方法就很难保证迭代结果的全局性和收敛的快速性,因此,在保证速度的前提下选择合适的初始参数将成为大基线图像特征点匹配算法成功的关键,这也是今后进一步研究的目标。

致 谢 感谢牛津大学 Visual Geometry Group 提供本文所用的实验图像数据。

参 考 文 献 (References)

- 1 Pritchett P, Zisserman A. Wide baseline stereo matching[A]. In: Proceedings of International Conference on Computer Vision [C], India, 1998: 754 ~ 760.
- 2 Georgescu B, Meer P. Point matching under large image deformations and illumination changes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(6): 674 ~ 688.
- 3 Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J]. Artificial Intelligence, 1995, 78(1): 87 ~

- 119.
- 4 Zhou J, Shi J Y. A robust algorithm for feature point matching[J]. *Computer & Graphics*, 2002, **26**(3): 429 ~ 436.
- 5 Schmid C, Mohr R. Local greyvalue invariants for image retrieval [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligent*, 1997, **19**(5): 530 ~ 534.
- 6 Lowe D. Object recognition from local scale-invariant features[A]. In: *Proceedings of International Conference on Computer Vision[C]*, Kerkyra, Greece, 1999: 1150 ~ 1157.
- 7 Lowe D. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91 ~ 110.
- 8 Mikolajczyk K, Schmid C. Indexing based on scale invariant interest points[A]. In: *Proceedings of International Conference on Computer Vision[C]*, Vancouver, Canada, 2001: 525 ~ 531.
- 9 Mikolajczyk K, Schmid C. An affine invariant interest point detector [A]. In: *Proceedings of European Conference on Computer Vision [C]*, Copenhagen, Denmark, 2002: 128 ~ 142.
- 10 Tell D, Carlsson S. Wide baseline point matching using affine invariants[A]. In: *Proceedings of European Conference on Computer Vision[C]*, Dublin, Ireland, 2000: 814 ~ 828.
- 11 Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets [A]. In: *Proceedings of European Conference on Computer Vision[C]*, Copenhagen, Denmark, 2002: 414 ~ 431.
- 12 Harris C, Stephens M. A combined corner and edge detector[A]. In: *Proceedings of Alvey Vision Conference[C]*, Manchester, UK, 1988: 147 ~ 151.
- 13 Rothganger F, Lazebnik S, Schmid C, *et al.* 3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints[A]. In: *Proceedings of International Conference on Computer Vision and Pattern Recognition[C]*, Madison, Wisconsin, USA, 2003: 272 ~ 277.
- 14 Shi J, Tomasi C. Good features to track[A]. In: *Proceedings of International Conference on Computer Vision and Pattern Recognition [C]*, Seattle, WA, USA, 1994: 593 ~ 600.
- 15 Fishler M A, Boles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of the Association for Computing Machinery*, 1981, **24**(6): 381 ~ 395.
- 16 Hartley R. In defence of the eight-point algorithm [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(6): 580 ~ 593.