

基于初始尺度变换的 SIFT 匹配算法

张 静, 桑红石*

(华中科技大学 图像识别与人工智能研究所 多谱信息处理国家重点实验室, 湖北 武汉 430074)

摘要:直接使用检测到的 SIFT(Scale-Invariant Feature Transformation)特征点进行特征点匹配,匹配性能仍然有待提升.提出了改进的 SIFT 匹配算法,利用匹配特征点的尺度比直方图,估计出近似的图像尺度比 k ,然后将空间分辨率较高的图像初始尺度增大到 k 倍,再次提取特征点进行匹配.实验结果表明,相比于其它用尺度约束条件提升性能的匹配算法,基于初始尺度变化的 SIFT 匹配算法在处理结构型图像时性能得到了很大的提升.

关键词:尺度不变特征;特征点匹配;初始尺度

中图分类号:TP391.4 **文献标识码:**A

SIFT matching method based on base scale transformation

ZHANG Jing, SANG Hong-Shi*

(National Key Laboratory of Science & Technology on Multi-spectral Information Processing, Institute for Pattern Recognition & Artificial Intelligence, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: Performance of the matching method which directly using SIFT feature points needs to be improved. This paper proposed an improved SIFT matching method which firstly uses the scale ratio histogram of the matched feature points to estimate the approximate scale ratio k of the image pair, then increases the base scale up to k times and extracts feature points again for the higher resolution image, finally performs feature matching again to obtain the result. The experiment results show that, compared with the existing improved SIFT matching methods with scale difference, performance is improved significantly for structured image pairs.

Key words: scale-invariant feature transformation(SIFT), feature matching, base scale transformation

PACS: 42.30.Tz

引言

SIFT 算法^[1]是最好的特征点提取和描述算子^[2],它提取的特征点对尺度和旋转具有不变性,同时对于亮度和三维视角变化具有一定的不变性,因此被广泛地用于图像配准,目标检测等视频处理系统的前端.传统的 SIFT 匹配算法^[1],不加约束地在整个特征向量集合中搜索匹配的特征向量,运算量大而且搜索出明显错误匹配点的概率比较高,影响匹配性能.

利用正确匹配点间的空间约束条件^[3,4]和尺度约束条件^[5-7],在匹配完成后剔除一些明显错误的匹

配点,可以显著降低误匹配率. Alhwarin^[5]将所有的特征点按 octave 进行分组,然后在各组间进行匹配,匹配数量最多的两组 octave 的尺度比被认定为正确的尺度因子,其他各组的匹配直接剔除.该算法只是在 octave 间粗略地估计特征点间尺度关系,因此尺度因子只能局限于 2^n (n 为整数). Yi^[6]首先估计出匹配点尺度差的均值,然后以该均值为中心点设定一个窗口,尺度差落于该窗口的认定为正确匹配,窗口外的为错误匹配.上述两种算法先用特征点匹配的结果提取尺度约束条件,然后利用该约束条件剔除一些明显错误的匹配,因此称为尺度限制匹配(Scale restriction SIFT, SR-SIFT)算法. SR-SIFT 算

收稿日期:2012-10-09,修回日期:2013-09-15

基金项目:教育部支撑计划项目(625010107)

Foundation items: Supported by Key Program of Ministry of Education of China(625010107)

作者简介(Biography):张静(1981-),女,湖北枣阳人,博士研究生,主要研究领域为计算机视觉前端算法研究及 ASIC 并行实现等. E-mail: jingzh8189@gmail.com

* 通讯作者(Corresponding author): E-mail: sanghs@gmail.com

Received date: 2012-10-09, revised date: 2013-09-15

法在降低错误匹配率的同时也减少了正确匹配的对数.为解决该问题, Bastanlar^[7]提出了图像预处理匹配(Preprocessing SIFT, PP-SIFT)算法,在获得尺度约束条件后,对图像进行高斯平滑和降采样等预处理后再次进行匹配. PP-SIFT 算法降低了错误匹配率,却不可避免的增加了算法复杂度,同时,经过预处理的图像其特征点必须进行逆变换之后才能用于后续的图像空间变换模型的计算.

本文分析了传统 SIFT 匹配算法引起误匹配的原因,利用尺度约束条件,改变特征点提取的初始尺度,来降低传统 SIFT 匹配算法的误匹配率.相比于已知的用尺度约束条件提升性能的匹配算法来说,基于尺度变换的 SIFT 匹配算法更为简单有效,在处理结构型图像时,误匹配率显著下降.

1 传统的 SIFT 匹配算法

传统的 SIFT 匹配算法包含两个部分:(1)提取两幅图像的 SIFT 特征向量,(2)在图像间搜索匹配的特征向量.提取 SIFT 特征向量包含四个主要的运算步骤:首先在高斯差分(Difference of Gaussian, DoG)尺度空间中检测三维极值点作为候选的特征点;然后对三维极值点进行精确定位,同时剔除低对比度以及位于边缘处的特征点;对于检测到的每一个特征点,先在其邻域内统计尺度方向直方图,求得直方图的峰值作为特征点的主方向;最后以主方向为新的坐标系建立梯度方向的直方图作为特征向量.对于两幅图像 A 和 B,特征向量 a (属于图像 A)的具体匹配方法是:在 B 中搜索与之距离最近和次近的特征向量 b 和 c ,若距离之比 $d(a,c)/d(a,b)$ 大于预先设定的阈值,则 b 为 a 的匹配点,否则没有与 a 匹配的特征点.

传统 SIFT 匹配算法,为了找到匹配点,要在图像的整个特征向量集中寻找匹配点,由于 SIFT 特征点具有多量性的特点,导致找到的匹配点中存在大量明显错误的匹配点.对于图 1 所示的一组图像来说,设置匹配阈值等于 1.5 时,可获得 461 个匹配点,其中错误匹配点为 111 个.明显错误的匹配点(尺度比大于 3 的 1 个,小于 1 的 35 个)占错误匹配总数的 32% 之多.

为了分析这些错误匹配点的来源,首先列出图 2 所示的特征点分布直方图.考虑到两幅图像间尺度近似为 2 的关系,若图 1(a)中某个特征点的尺度为 s ,那么图 1(b)中与其匹配的特征点尺度必然在 $s/2$ 附近.显然,尺度小于 3 的特征点(见图 2(a)箭

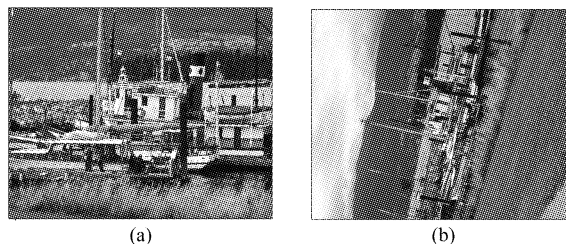


图 1 示例图像对(近似尺度为 2)
Fig. 1 Example image pair (approximate scale ratio is 2)

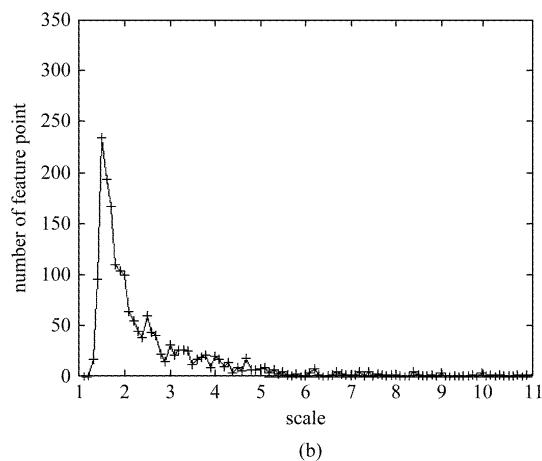
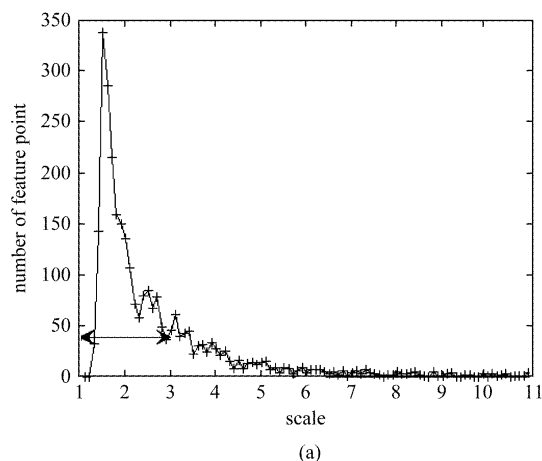


图 2 特征点分布直方图(a)图 1(a)的直方图,(b)图 1(b)的直方图

Fig. 2 Number of features distribution histogram (a) histogram of Fig. 1(a), and (b) histogram of Fig. 1(b)

头所示)在图 1(b)中找到匹配点的可能性几乎为 0,因为图 1(b)中不存在尺度小于 1.5 的特征点(见图 2(b)).为了从数量上说明这个区间的重要性,表 1 列出了两图像中检测出的特征点数量.图 1(b)中 octave 0 的特征点与图 1(a)中 octave 1 的特征点具有尺度对应关系,图 1(b)中 octave 1 的特征点与图 1(a)中 octave 2 的特征点具有尺度对应关系,依次类推.显然,图 1(a)中 octave 0 的特征点在图 1(b)

中没有与之对应的尺度,在图1(b)中找不到匹配点.然而,这部分特征点的数量高达2221个(特征点总数的78%),这些特征点在匹配阶段不但增加了运算量,而且增加了匹配出错的概率.

表1 图1中图像的SIFT特征点数量
Table 1 Number of SIFT features of images shown in Fig. 1

octave	scale	Fig. 1(a)	Fig. 1(b)
0	[1.6 3.2]	2221	1474
1	[3.2 6.4]	511	308
2	[6.4 12.8]	121	77
3	[12.8 25.6]	0	0

2 基于初始尺度变换的SIFT匹配算法

从上一节的分析可知,高分辨率图像的低尺度空间特征点数量巨大,却不可能找到匹配的特征点,在匹配阶段不但增大运算量,而且增加匹配出错的概率.如果将高分辨率图像的特征点集中在图2(a)所示箭头区域外的尺度范围,就可以降低较低尺度空间特征点数量,进而降低误匹配率.本文首先利用匹配特征点的尺度比直方图估计出图像间的尺度因子 k ,然后利用该尺度因子调节特征提取的初始尺度,再次提取特征点进行匹配,从而降低算法的误匹配率,因此称为基于初始尺度变换的SIFT匹配算法,其流程图如图3所示.

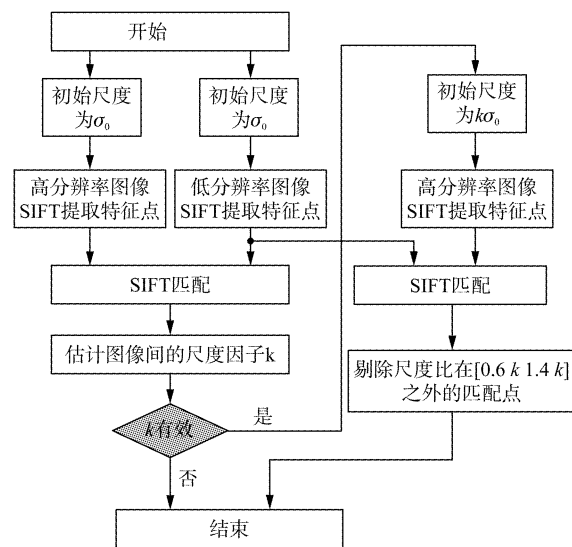


图3 基于初始尺度变换的SIFT匹配算法流程图
Fig. 3 Flow chart of SIFT matching method based on base scale transformation

基于初始尺度变换的SIFT匹配算法包括两次特征点匹配,共四个主要步骤:(1)第一次特征点匹

配设置相同的初始尺度 σ_0 ,分别提取两帧图像的SIFT特征向量,对特征点进行匹配;(2)根据特征点的匹配结果,估计图像间的尺度因子 k ,若 k 为有效的尺度因子,按照下面的步骤进行第二次匹配,否则第一次匹配结果即为最终匹配结果;(3)第二次特征点匹配设置初始尺度为 $k\sigma_0$,重新提取高分辨率图像的特征点,与低分辨率图像的特征点进行匹配;(4)对第二次匹配的结果,剔除尺度比落在区间 $[0.6k, 1.4k]$ 范围之外的点.

估计图像间的尺度因子 k 是新算法的一个关键环节.基于对线性尺度空间的研究,1998年,Lindeberg^[8-9]提出使用拉普拉斯算子尺度归一化后的算子 $\sigma^2 \nabla^2 G$ 来检测blob结构的特征点, $\sigma^2 \nabla^2 G$ 算子极值点所在的尺度可以反映blob结构的特征长度.Lowe^[1]以Lindeberg的研究为基础,使用DoG算子近似于 $\sigma^2 \nabla^2 G$ 算子,从而将blob结构的检测转换为求高斯差分尺度空间的极值点.因此,Lindeberg提出的线性尺度空间所具有的尺度不变性也适用于SIFT特征点.尺度不变性在特征点匹配上的具体表现是:对于尺度发生变换的图像来说,正确匹配的特征点尺度比应该近似于图像的尺度比 k .然而,图像间不止存在尺度变换,还存在旋转、视角和亮度等各种变换.对于尺度变换为主的图像来说,特征点尺度比直方图应该在 k 处具有最高的概率密度.某些视角发生较大变化的图像对,直方图不存在明显的峰值,采用直方图的极大值点作为图像尺度因子又不太合理.考虑到这些因素,提出了下面的尺度因子提取算法.使用该算法之前,要用传统的SIFT匹配算法,提取两幅图像的特征向量进行匹配,然后按照下面的步骤进行.

设图像A和B匹配的结果为 (a_i, b_i) , $1 \leq i \leq N$,其中 a_i 和 b_i 分别为图像A和B中的特征点, N 为匹配的特征点数量.特征点的尺度比定义如式(1)所示,其中 $scale(a_i)$ 和 $scale(b_i)$ 分别表示特征点 a_i 和 b_i 的特征尺度.

$$sr_i = \frac{scale(a_i)}{scale(b_i)} \quad (1)$$

提取图像尺度因子 k 的算法如下:

1)首先计算特征点尺度比的最大值 $srMax$ 和最小值 $srMin$ 方法是:对向量 (sr_i) 进行排序得到向量 (sro_i) ,分别取最大值5%及最小值5%的值为 $srMax$ 和 $srMin$,即 $srMin = sro_p$, $srMax = sro_q$,其中 $p = \lceil N \times 0.05 \rceil$, $q = N - p$, N 是向量内元素个数.

2)设置特征点尺度比区间间隔为 d ,统计落入

各个区间的匹配点数量,得到特征点尺度比直方图 $h(j)$,见式(2):

$$h(j) = \sum_{i=1}^N sr_i \geq (srMin + j \times d - d) \& sr_i < (srMin + j \times d) \quad , \quad (2)$$

其中, $1 \leq j \leq \frac{srMax - srMin}{d}$.

3) 获取直方图 $h(j)$ 最大值处的尺度比作为图像尺度因子 k ,见式(3);

$$k = srMin + (m - 1) \times d + \frac{d}{2} \quad , \quad (3)$$

其中 $\forall j, h(m) \geq h(j)$.

4) 最后判断上述所得尺度因子 k 是否为有效值,具体方法是:计算尺度比在区间 $[0.6k, 1.4k]$ 内的匹配点数量,若占总匹配点数的 75% 以上,说明 k 为有效的尺度因子;否则, k 为无效的尺度因子.

从获取直方图峰值的角度来看,选用的区间间隔 d 对于求取直方图的峰值有很大的影响, d 太小可能得不到极大值, d 太大又不够精确. 但是,本文工作的出发点是由粗到精提高匹配质量,即初始匹配时粗略获得两帧图像间的尺度关系,然后使用该尺度关系去限制匹配条件,从而提升匹配质量. 从这个角度理解,算法对尺度比 k 的精度要求不高,只要离预期的尺度比相差不大即可. 经过试验,本文提出的算法中选取参数 $d = 0.05$ 可以获得较好的匹配性能.

3 实验结果与分析

为了验证本文算法的有效性,从两个方面来考虑. 一是针对图 1 所示的一组图像,分析采用新算法如何降低误匹配率;二是比较多组场景的误匹配率,验证新算法相比于其它用尺度约束条件提升匹配性能的优势.

对图 1 所示的一组图像,按照上节所述方法进行第一次特征点匹配,获得特征点尺度比直方图如图 4 所示. 根据式(2)和式(3)计算峰值所处位置的尺度比为 1.91,其中落于区间 $[1.14, 2.66]$ ($0.6k = 1.14, 1.4k = 2.67$) 范围内的匹配点数量占总数量的 89.9%,说明提取的尺度因子 1.91 是有效的,可以用于下一步的初始尺度变换匹配. 显然,采用本文算法提取的尺度因子 k 与第一节提到的近似图像尺度 2 接近,是合理并且有效的. 另一方面,图 4 中直方图分布也进一步验证了第一节所述

传统 SIFT 匹配算法存在的问题. 直方图左侧部分对应于图 1(a)中低尺度特征点的匹配结果,这部分特征点在图 1(b)并不能找到匹配的点,即使找到了匹配点,尺度比也是不合理的,是明显错误的匹配,然而这部分明显错误的匹配结果却占到了总数量的 9% 之多.

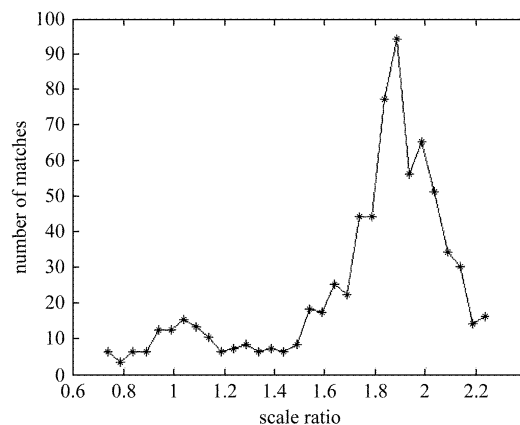


图 4 匹配的特征点尺度比直方图

Fig. 4 Scale ratio histogram of matched features

提取图像的尺度因子后,调整初始尺度为 $k\sigma_0$,再次对图像 1(a) 进行特征点提取,得到各尺度区间的特征点数量如表 2 所示. 在低尺度空间,图像 1(a) 检测到的特征点数量从表 1 中的 2221 降低到了 275,其他尺度空间的特征点分布变化不大.

表 2 初始尺度变换后图 1 的 SIFT 特征点数量
Table 2 Number of SIFT features after base scale transformation images shown in Fig. 1

octave	scale	Fig. 1a	Fig. 1b
0	[1.6 3.2]	275	1474
1	[3.2 6.4]	505	308
2	[6.4 12.8]	135	77
3	[12.8 25.6]	39	0

表 3 初始尺度变换后匹配数量的变化

Table 3 Number of matches after base scale transformation

	匹配数量	正确匹配数量	错误匹配数量	误匹配率
传统算法	461	350	111	0.241
新算法(无剔除操作)	396	349	47	0.115
新算法($[0.8k, 1.2k]$)	384	348	36	0.094
新算法($[0.7k, 1.3k]$)	400	361	39	0.097
新算法($[0.6k, 1.4k]$)	402	362	40	0.099
新算法($[0.5k, 1.5k]$)	403	362	41	0.102

第二次匹配后匹配点数量的变化如表 3 所示. 低尺度空间特征点数量的降低直接导致再次匹配

后,错误匹配点数由 111 降为 47 个.47 个错误匹配点中,明显错误的匹配只有 6 个,相比于传统算法的 36 个有了显著下降,这说明本文的基本思想——通过调整初始尺度来降低明显错误匹配个数是可行并且有效的.

最后一步剔除尺度比落于 $[0.6k, 1.4k]$ 外的匹配点.为了评估剔除区间对最终性能的影响,表 3 列出了选择不同剔除区间时误匹配率的情况.当选择较为苛刻的剔除区间时(如 $[0.8k, 1.2k]$),得到的错误匹配数量最低,当区间变得宽松时,错误匹配数量有所增加.苛刻的剔除区间可以剔除较多错误匹配的同时也导致一些正确匹配被剔除,降低了正确匹配点的数量.对于图 1 所示尺度变化类图像来说,特征点间主要是尺度的变化,因此尺度比直方图具有较好的密度分布,设置较为苛刻的剔除区间对误匹配率的影响不是特别大.但是为了提高算法的鲁棒性,本文设置一个折中的区间 $[0.6k, 1.4k]$ 作为最终剔除区间.

为了更进一步验证算法性能,对已知的用尺度约束条件提升匹配性能的算法进行实验,参与比较的算法有 SR-SIFT 算法和 PP-SIFT 算法,实验结果如图 6 所示.实验所用的 SR-SIFT 算法为文献[5]和[6]的改进版,采用尺度比值而不是尺度差作为约束条件,适用的尺度范围更大,也没有文献[6]要求的尺度比 2^n 限制.测试图像采用图 1 所示的“boat”场景序列和图 5 所示的“garden”和“graf”场景序列,其中 boat 和 graf 场景来自于 Mikolajczyk^[10], garden 场景为公园一角.每个场景序列中共有 6 帧图像,第一帧为高空间分辨率图像,其他各帧为尺度或者视角不断增大的图像.为了能对三个算法进行公平的测试,本文通过循环逼近的方法来寻找合适的阈值使得三个算法具有相同的匹配点数,然后比较三个算法的误匹配率.考虑到应用环境的需要,设置匹配点数分别为 50,60,70,80,90 和 100.

表 4 三种算法在各场景下的平均误匹配率

Table 4 Average false-positive rate of three algorithms in different scene

算法	boat 场景	garden 场景	graf 场景
PP-SIFT	0.139 4	0.212 9	0.181 2
SR-SIFT	0.090 4	0.173 4	0.192 9
proposed	0.046 8	0.156 5	0.172 2

新算法相比于其他算法的优势如图 6(a) 和图 6(b) 所示.图 6(a) “boat” 场景的曲线可以看出,新算法在不同的匹配点数下,误匹配率变化较为平稳,

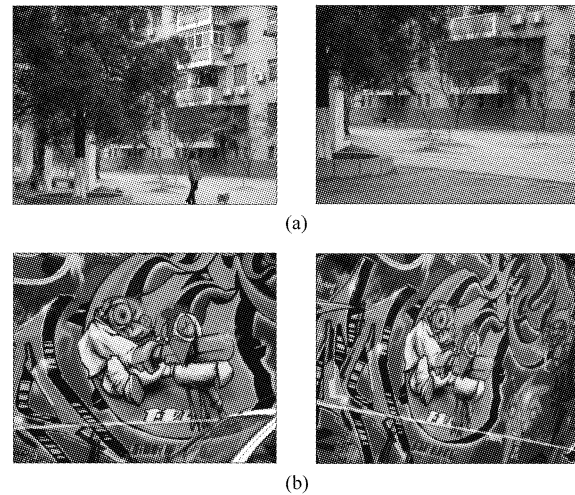


图 5 参与测试的场景图像(a)“garden”,(b)“graf”
Fig. 5 Image pairs for testing (a) garden, and (b) graf

而且相比于 SR-SIFT 和 PP-SIFT 分别有 0.1 和 0.05 的降低.图 6(b) “garden” 场景下,新算法的误匹配率曲线比较接近于 SR-SIFT,但是比 PP-SIFT 仍然有平均 0.05 左右的降低.

为了更清晰地展现新算法的优势,将图 6 中不同匹配对数得到的误匹配率进行平均得到表 4.对于“boat”和“garden”场景来说,SR-SIFT 算法相比于 PP-SIFT 算法,误匹配率有 0.05 左右的降低;而新算法相比于 PP-SIFT 算法则有 0.1 左右的降低.“graf”场景主要为视角的变化(见图 5(b)),因此各类基于尺度约束条件的匹配算法对匹配性能的提升比较有限.尽管如此,本文提出的新算法相比 PP-SIFT 算法仍有 0.02 左右的降低.

4 结论

对传统的 SIFT 匹配算法施加尺度上的约束,可以降低特征点的错误匹配率.基于该思想,本文提出了基于初始尺度变化的 SIFT 匹配算法:根据匹配特征点的尺度比直方图,估计两图像的尺度比 k ,然后将高分辨率图像的初始尺度设置为原值的 k 倍,再次提取特征点与另一图像进行匹配,可得到错误匹配率降低的匹配特征点.

从两个方面对本文新算法进行了实验验证,首先以一对图像为例分析新算法如何降低误匹配率,然后对多组场景试验进一步比较算法的误匹配率.实验结果证明,在相同匹配点数的情况下,对于结构型场景来说,本文算法相比于 SR-SIFT 和 PP-SIFT 算法具有明显的优势.

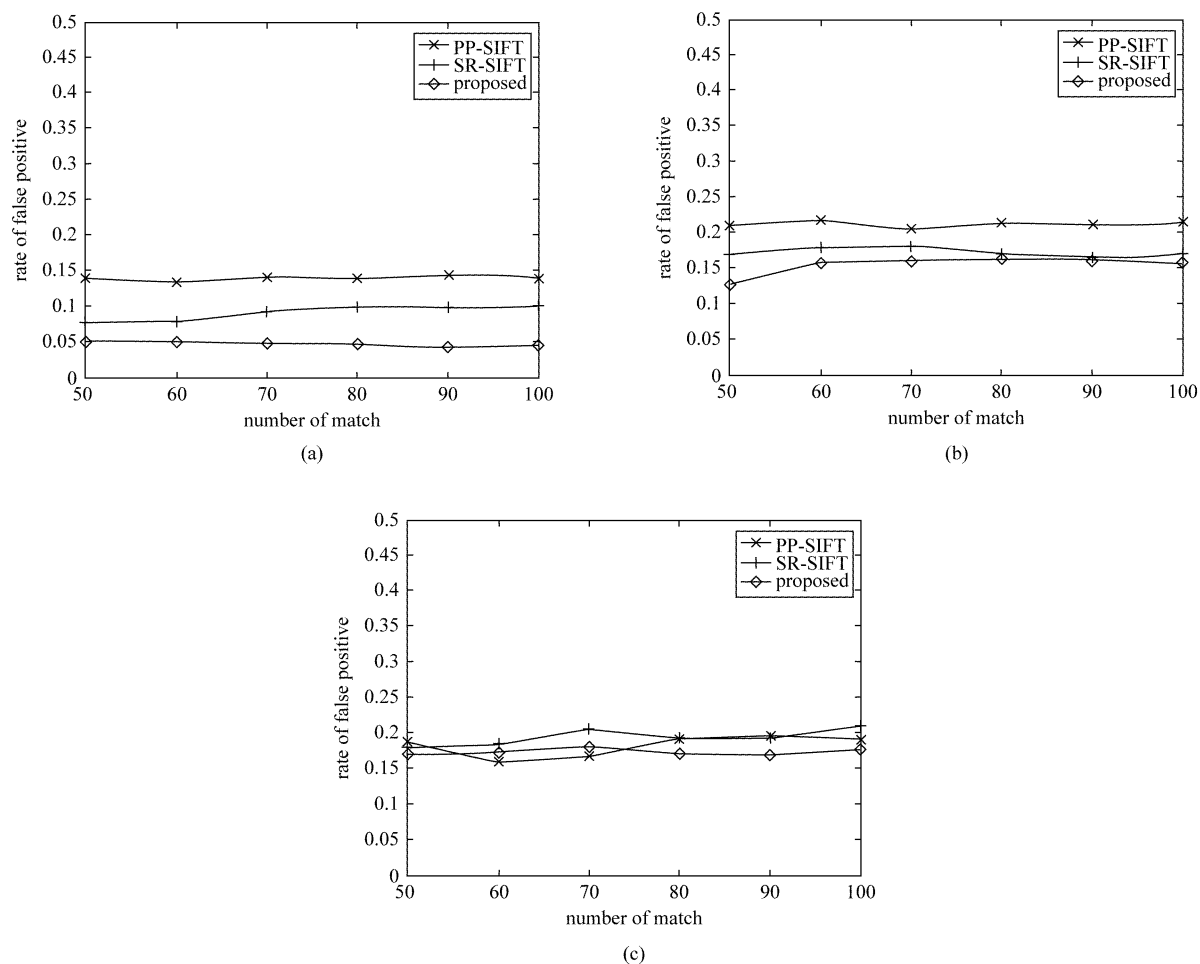


图6 三个测试场景的误匹配率(a) boat, (b) garden, (c) graf

Fig. 6 False-positive rate of three testing scene (a) boat, (b) garden and (c) graf

REFERENCES

- [1] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- [2] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1615-1630.
- [3] Torr P H S, Zisserman A. MLESAC: a new robust estimator with application to estimating image geometry[J]. *Computer Vision and Image Understanding*, 2000, **78**(1): 138-156.
- [4] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of the ACM*, 1981, **24**(6): 381-395.
- [5] Alharwin F, Wang C, Ristic-Durrant D, et al. A. Improved SIFT-features matching for object recognition[C]. *In Visions of Computer Science-BCS International Academic Conf.*, Imperial College, London, UK, September, 2008.
- [6] Yi Z, Zhiguo C, Yang X. Multi-spectral remote image registration based on SIFT[J]. *Electronics Letters*, 2008, **44**(2): 107-108.
- [7] Bastanlar Y, Temizel A, Yardimci Y. Improved SIFT matching for image pairs with scale difference[J]. *Electronics Letters*, 2010, **46**(5): 346-U4867.
- [8] Lindeberg T. Scale-space theory: A basic tool for analysing structures at different scales[J]. *Journal of Applied Statistics*, 1994, **2**(21): 224-270.
- [9] Lindeberg T. Feature detection with automatic scale selection[J]. *International Journal of Computer Vision*, 1998, **30**(2): 79-116.
- [10] Mikolajczyk K. Affine Covariant Features [EB-OL], <http://www.robots.ox.ac.uk/~vgg/research/affine/>