

基于 SIFT 算子的图像匹配算法研究

白廷柱, 侯喜报

(北京理工大学 光电成像技术与系统教育部重点实验室, 北京 100081)

摘 要: 针对目前基于 SIFT(scale invariant feature transform)的图像匹配算法在匹配相似区域较多的可见光图像时,匹配约束条件单一,没有有效剔除误匹配点,误匹配率高的问题,提出一种匹配改进算法,针对 128 维 SIFT 特征向量,采用距离匹配和余弦相似度匹配相结合的测度方法,利用特征点方向一致性进一步降低误匹配率。实验结果表明:改进算法对图像的缩放、旋转、光照、噪声和小尺度的视角变换均有较好的匹配效果。与原算法相比,在保证匹配点数和匹配时间的基础上,改进算法对旋转、缩放、噪声模糊和光照变换的误匹配率平均降低 10%~20%,对于小尺度的视角变换,误匹配率平均降低 5%。

关键词: SIFT; 图像匹配; 余弦相似度; 方向一致性; 校正误匹配

中图分类号: TP 391.9

文献标志码: A

文章编号: 1001-0645(2013)06-0622-06

An Improved Image Matching Algorithm Based on SIFT

BAI Ting-zhu, HOU Xi-bao

(Key Laboratory of Photoelectric Imaging Technology and System, Ministry of Education of China,
Beijing Institute of Technology, Beijing 100081, China)

Abstract: For matching visible image with many similar regions, the original image matching algorithm based on SIFT (scale invariant feature transform) has the disadvantages of limited matching constraints, high false matching rate and difficulty to effectively remove mismatching points. To overcome the shortcomings above, an improved algorithm was proposed in which a combined measure of distance similarity matching with cosine similarity matching was adopted to dealing with 128-dimensional feature vectors. Further, the orientation consistency of the image feature points was employed to reduce the false matching rate. Experimental results show that the proposed algorithm has a good matching result on the conditions of image zooming, rotating, lighting, noising and small-scale perspective transformation. Compared with the original algorithm, based on the premise of ensuring enough matching points and definite matching time, the improved algorithm achieves a 10% to 20% average reduction of the false matching rate for images zooming, rotating, lighting, noising transformation and 5% for small-scale perspective transformation.

Key words: scale invariant feature transform (SIFT); image matching; cosine similarity; consistency of orientation; mismatching calibration

数字图像匹配技术广泛应用于图像拼接与融合^[1]、目标识别与跟踪^[2]、摄影测量遥感^[3]、图像检索^[4]等图像处理和机器视觉领域。目前用于图像匹配的方法主要分为基于像素的匹配和基于特征的匹

配^[5]。近年来,提取图像局部不变特征的方法得到了较快地发展^[6-8]。哥伦比亚大学的 Lowe 在 1999 年提出了一种新的点特征提取算法——SIFT (scale invariant feature transform)算法^[2,6]。

收稿日期: 2012-09-15

作者简介: 白廷柱(1955—),男,博士,教授, E-mail: tzhbai@bit.edu.cn.

但在源图像和目标图像相似区域较多的情况下,由于 SIFT 算法仅采用欧式距离比这种单一的相似测度方法,且没有考虑空间的几何约束信息,因而导致误匹配率高,出现了明显的错配和乱配情况. 针对上述问题,在实时性要求不高的背景下,作者在原 SIFT 算子的基础上,辅以图像有效的几何限制和主运动方向信息,即采用向量空间余弦相似度测度和欧式距离比测度相结合的匹配方法建立约束条件,并充分利用特征点方向一致性作为图像的几何约束条件,在保证必须匹配点对个数和匹配时间的前提下,有效降低了误匹配率.

1 SIFT 算法基本原理

SIFT 图像匹配算法^[9]首先利用尺度空间的概念建立图像的尺度空间,在尺度空间中搜索出图像的局部极值点作为候选关键点,去除对比度低的关键点和不稳定的边缘响应点,确定关键点的主方向,生成关键点特征描述子,使每个关键点具有位置、尺度和方向信息,最后采用特征描述符向量之间的欧式距离度量两个特征点之间的匹配程度.

1.1 特征点检测

特征点检测可分为 3 步,尺度空间极值点的检测、关键点的精确定位、关键点方向的分配.

① 尺度空间极值点的检测. 高斯核是唯一可以产生多尺度空间的核^[10],一幅二维图像 $I(x,y)$ 的尺度空间 $L(x,y,\sigma)$ 可表示为

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y). \tag{1}$$

式中: σ 为尺度空间的空间尺度因子; $G(x,y,\sigma)$ 为高斯核函数,其定义为

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}\exp(-(x^2+y^2)/2\sigma^2). \tag{2}$$

为了在尺度空间检测稳定的关键点,使用高斯差分(DOG)算子近似尺度归一化的拉普拉斯-高斯(LOG)算子

$$D(x,y,\sigma)=L(x,y,k\sigma)-L(x,y,\sigma), \tag{3}$$

式中 k 为阈值. 这样,通过高斯平滑和降采样得到高斯金字塔,再由相邻尺度图像相减生成 DOG 金字塔后形成尺度空间. 在尺度空间内,每一个采样点和其同尺度相邻点及上下相邻尺度共 $8+9\times 2=26$ 个点进行比较,以确保在图像空间和尺度空间都检测到局部极值点.

② 关键点的精确定位. 为提高匹配的抗噪能力和稳定性,通过拟合三维二次函数更为精确的确

定关键点的位置和尺度,设定阈值去除对比度低的关键点. 由于 DOG 算子会产生较强的边缘响应,可通过利用 Hessian 矩阵计算主曲率设定阈值的方法,去除不稳定的边缘响应点.

③ 关键点方向的分配. 将关键点邻域像素的梯度方向分布作为每个关键点的指定方向信息,梯度值 $m(x,y)$ 和方向 $\theta(x,y)$ 计算表达式为

$$m(x,y)=\{[L(x+1,y)-L(x-1,y)]^2+[L(x,y+1)-L(x,y-1)]^2\}^{1/2}, \tag{4}$$

$$\theta(x,y)=\tan^{-1}\left[\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right]. \tag{5}$$

通过直方图统计关键点邻域像素的梯度方向,直方图的峰值代表该关键点的邻域梯度的主方向,作为该关键点的方向. 至此,特征点检测完毕,每一个特征点都包含位置、尺度和方向三方面信息.

1.2 特征点描述

特征点描述通过对关键点周围区域进行图像分块,计算各块内的梯度直方图,生成独特性的向量描述符,该向量描述符是该区域图像信息的一种抽象表示. 具体方法是:以关键点为中心取 (16×16) 像素的窗口,在窗口中取 (4×4) 像素的区域形成一个种子点,共 4×4 个种子点,在每个种子点区域内分别计算 8 个方向的梯度累加值,绘制梯度方向的方向直方图. 最终获得一个 $4\times 4\times 8=128$ 维的特征描述向量.

1.3 特征点匹配

SIFT 特征向量生成后,采用最近邻距离算法(NN)^[2]进行匹配,即采用和样本特征点最近邻的特征点的欧式距离与次近邻特征点的欧氏距离的比值与所设定阈值比较,若比值小于阈值,则认为特征点对匹配. Lowe 经过实验,设定阈值为 0.8. 在计算特征点之间的欧氏距离时,采用了 BBF 算法来处理 128 维的特征向量^[2].

2 匹配算法的改进

由于 SIFT 描述的是以计算图像的梯度为基础的局部特征,在匹配的过程中,若有些点的邻域信息特别相近,比如地图图像地名中相同的汉字,因为视角的不同,两个相近点的特征描述符的匹配程度甚至有可能会超过同一点在两幅图中的相似程度. 故当图像中存在大量形状相似的区域时,基于 SIFT 特征的图像匹配将会出现大量误匹配点对,导致图

像出现匹配误差. 不同分辨率、不同时间、不同传感器上获取的图像间的匹配更是如此, 如图 1 所示. 特征点匹配是图像匹配的关键步骤, 为了保证图像匹配精度, 必须采用有效的方法尽可能多地去除误匹配点对. 为此, 本文中考虑增加匹配约束条件减少误匹配点对数.

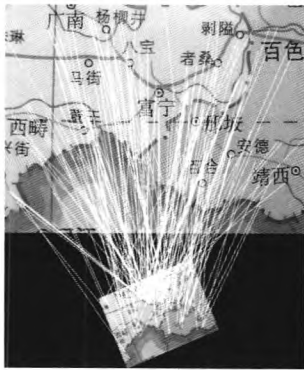


图 1 原算法匹配地图

Fig. 1 Matching map image using original algorithm

2.1 向量空间余弦相似度匹配

原算法中采用欧式距离度量两个特征向量之间的相似性, 计算待测图像中每一个特征向量与所有目标图像中全部特征向量的欧式距离, 如果最小欧式距离与次最小欧式距离的比值小于 0.8, 就认为源特征点和最小欧式距离特征点相匹配. 但是欧式距离没有考虑各向量分量之间的相关性, 体现单一特征的多个分量会干扰匹配结果^[11].

衡量两个向量之间的相似程度, 一般有距离测度法和相似性函数法^[11]. 距离测度用于衡量向量在空间上存在的距离, 距离越远说明向量间的差异越大. 相似性函数是用函数的方法表征两向量相似的程度, 其值越小, 说明向量间相似度越小, 差异越大. 欧氏距离是最常见的距离测度, 而余弦相似度则是最常见的相似性函数, 很多距离测度和相似性函数都是基于这两者的变形和衍生, 两者分别从数值特征和方向特征上体现了向量之间的差异. 所以, 作者在欧式距离测度的基础上加入了余弦相似度约束, 通过设定阈值减少误匹配点对的个数.

具体方法: 在某个向量的最小欧式距离与次最小欧式距离比值小于 0.8 的基础上, 进一步判断这两个向量的余弦相似度是否大于某一经验阈值, 大于则相匹配. 经验阈值 k 可通过实验得到. 余弦相似度是 2 个向量之间夹角的余弦值, 其不受参考坐标系旋转, 放大缩小的影响, 且规范化了向量的长度, 对于 2 个向量 x 和 y , 其余弦相似度表达式为

$$S(x, y) = \frac{(x, y)}{\|x\| \|y\|} = \frac{\sum_{i=1}^n x_i y_i}{(\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2)^{1/2}}. \quad (6)$$

2.2 方向一致性几何约束

由于原算法中两幅图像之间的比较退化成了特征向量集的比较, 忽略了各个特征之间的空间几何信息, 造成了大量的错误匹配. Lowe 指出^[2], 采用 RANSAC 算法可以有效剔除误配点, 但是在外点较多时其性能将受到很大的影响. 图像的主运动方向和特征点运动方向具有一致性, 如果不是正确的匹配点, 则满足这一事实的概率非常小, 由于 SIFT 算法的鲁棒性, 大部分匹配点是正确的, 其运动方向也大致相同, 所以可以认为偏离这个运动方向的匹配点为误匹配点.

作者在执行 RANSAC 算法之前, 将匹配点对方向一致性作为特征点对之间有效的空间几何约束限制, 进一步筛选正确匹配点. 匹配点对的方向一致性依据两幅图像中正确匹配点对的运动方向反映了图像主运动信息的事实, 通过建立匹配点对运动方向直方图, 认为偏离峰值的匹配点即与图像总运动趋势不一致的误匹配点. 具体方法如下.

① 计算已得到的匹配点对的方向差 ΔO_i , 其表达式为

$$\Delta O_i = O_{\text{img1}_i} - O_{\text{img2}_i}. \quad (7)$$

② 将方向差值移动到同一区间 $[0^\circ, 360^\circ]$, 其表达式为

$$\Delta O'_i = 360^\circ - \left(\frac{360^\circ \times \Delta O_i}{2\pi} + 360^\circ \right) \text{MOD } 360^\circ, \quad (8)$$

式中 MOD 表示求余运算.

③ 在区间 $[0^\circ, 360^\circ]$ 内建立得到的 $\Delta O'_i$ 直方图, 直方图以每 10° 方向为一个区, 共 36 个区, 使 $\Delta O'_i$ 落入对应的区中, 对应的区幅值累加 1.

④ 选取落入最大幅值区以及次大幅值区上所对应的匹配点对作为最终匹配点对, 若次大幅值区与最大幅值区不相邻, 则舍弃次大幅值区上的匹配点对.

图 2 为图 1 中两幅图像匹配点对的方向直方图, 依据上述方法, 最终只选取落入 2、3 区域的匹配点对, 排除可能的误匹配点对共 6 对.

3 算法实验与结果分析

实验平台为 Windows 7 操作系统、CPU 3.2

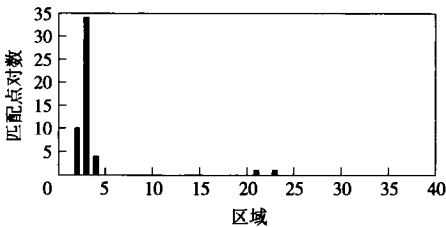


图 2 地图图像匹配点对的方向直方图
Fig. 2 Orientation histogram of matching points from map image

GHz、内存 2.00 GB 的 PC 机,编程环境为 Visual Studio 2008,并加入了开源库 Open CV. 为了能较为准确地计算出误匹配点的个数以及误匹配率,测试图像选用牛津大学机器人实验室的一组图像,该组图像包含了包括噪声模糊、视角变换、缩放和旋转变换、光照变换等各种图像变换。

3.1 寻找余弦相似度匹配的经验阈值

首先通过实验寻找余弦相似度匹配的经验阈值,分别设置噪声、视角、旋转和缩放、光照等不同种类变换的图像对余弦相似度匹配阈值为可变,认为当两向量夹角余弦值大于这一阈值时,即为正确匹配。通过不断变换阈值,拟合出阈值 k 和正确匹配点数 N 、阈值和误匹配率 P 之间的关系曲线。图 3 为一组光照变换的关系曲线。

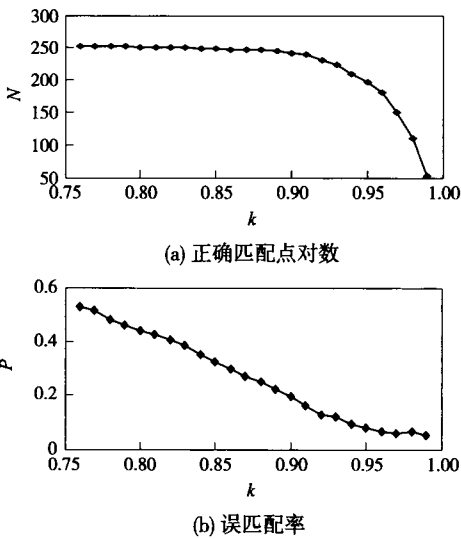


图 3 阈值和正确匹配点对数及误匹配率关系图
Fig. 3 Curves of threshold versus correct matching point and false matching rate

最优阈值应该满足正确匹配点数尽量多而误匹配率尽量小,由图 3 可以看出,随着阈值的增大,正确匹配点数逐渐减少,误匹配率逐渐减小。正确匹配点数减少的幅度开始很平缓,超过 0.95 后则急剧增大,误匹配率则相反。因而可以取该组变换的最

优阈值 k_{best} 为 0.95。

利用上面的方法寻找各种变换对应的最优阈值,结果如表 1 所示。为了使阈值具有普遍适用性,将阈值分别组内作平均,再总体加权平均,将最终结果 0.93 作为经验阈值。

表 1 各种变换的余弦相似度最优阈值 k_{best}

Tab. 1 Optimal thresholds of cosine similarity for different transformations

图像序号	k_{best}			
	旋转和缩放	视角	光照	噪声
1	0.96	0.94	0.95	0.96
2	0.96	0.86	0.94	0.95
3	0.94	0.93	0.96	0.94
4	0.93	0.84	0.96	0.95
5	0.90	0.87	0.95	0.94

3.2 改进算法性能测试

测试通过利用经典算法、加入余弦相似度约束、加入方向一致性约束以及最终改进算法进行匹配,统计总匹配点数、误匹配点数和匹配时间,计算误匹配率。表 2 和表 3 为图像库中 4 组缩放和旋转变换的实验结果,图像尺寸为(800×640)像素。

表 2 改进后算法与原算法误匹配率计算结果比较

Tab. 2 Comparison of false matching rates obtained from using improved algorithm and original algorithm

图像序号	P			
	原算法	加入方向约束	余弦相似度匹配	综合改进匹配
1	0.061 6	0.032 1	0.056 5	0.019 3
2	0.051 3	0.027 3	0.054 5	0.013 4
3	0.186 7	0.118 5	0.198 1	0.088 9
4	0.229 8	0.104 8	0.248 2	0.046 0

表 3 改进后算法与原算法匹配时间比较

Tab. 3 Comparison of matching time using improved algorithm and original algorithm

图像序号	t/ms			
	原算法	加入方向约束	余弦相似度匹配	综合改进匹配
1	1 061	1 186	2 464	1 202
2	795	733	2 106	734
3	811	796	1 748	686
4	812	718	1 326	717

部分图像原算法和改进算法匹配结果比较见图 4。图 5 为各种变换改进后算法和原算法效果比较。统计所有测试图像原算法和改进算法匹配时间,比较如图 6 所示。图 7 为初始地图图像的匹配结果。

由匹配结果结合实验数据可以得出:

① 相对于原算法,改进算法对于旋转和缩放、噪声模糊、光照等变换均有较好的效果,误匹配率平均降低 10%~20%;

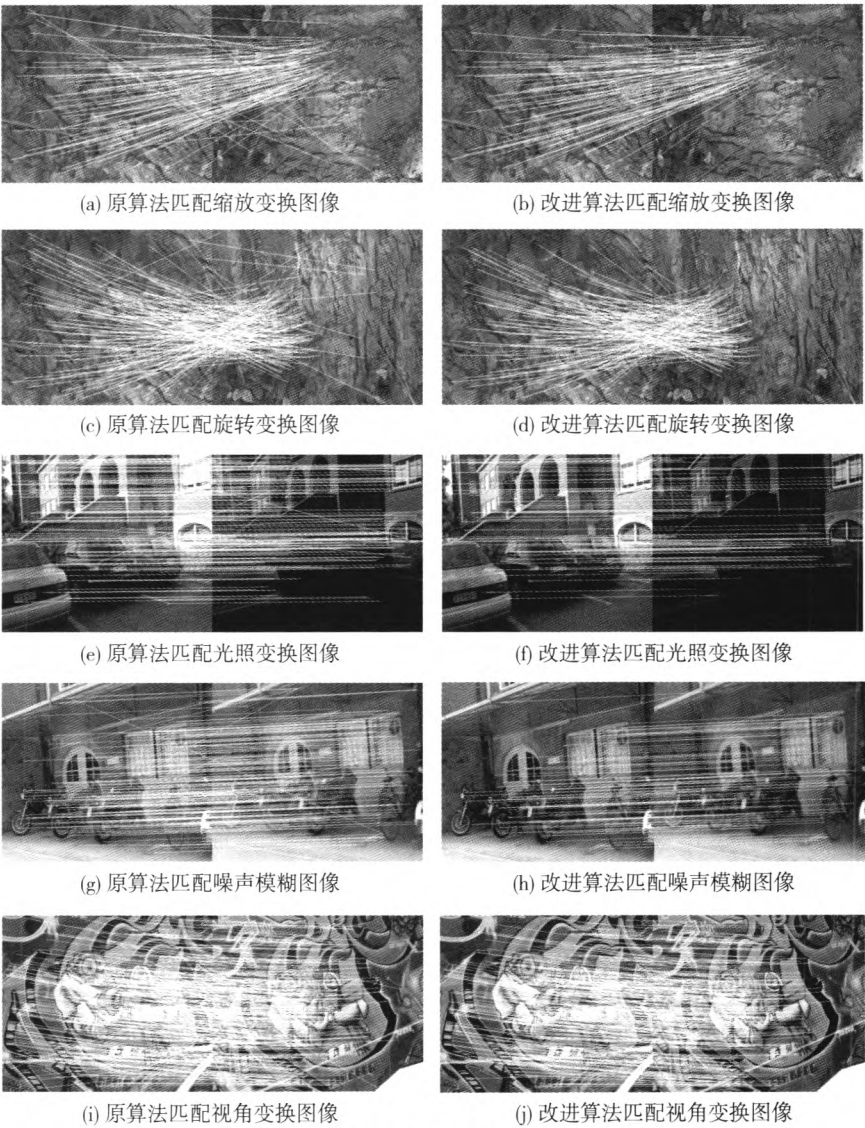


图 4 原算法和改进算法对各种变换匹配图像比较

Fig. 4 Comparison of matching images with different transformations obtained from using improved algorithm and original algorithm

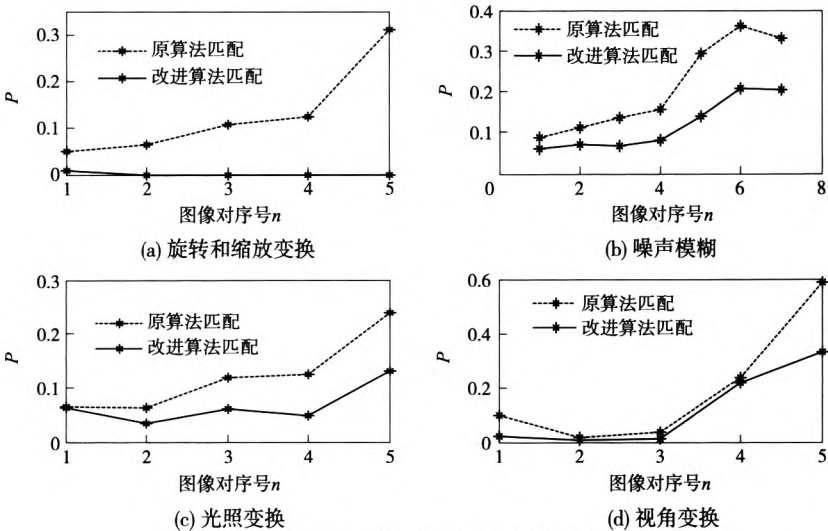


图 5 原算法和改进算法对各种变换误匹配率比较

Fig. 5 Comparison of false matching rate with different transformations using improved algorithm and original algorithm

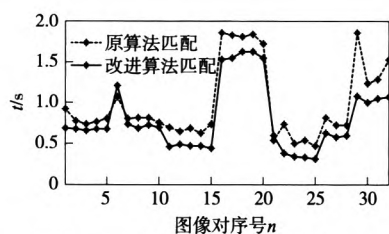


图 6 原算法和改进算法匹配时间比较

Fig. 6 Comparison of matching time using improved algorithm and original algorithm

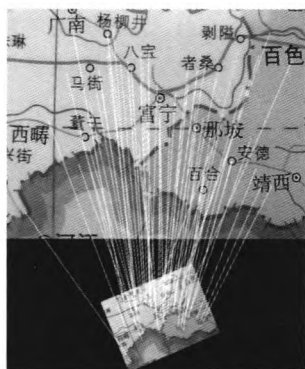


图 7 改进算法匹配地图图像

Fig. 7 Matching map image using the improved algorithm

② 改进算法和原算法的匹配时间相近;

③ 对于视角变换,改进算法效果不是很明显,对于小尺度的视角变换(40° 以内),误匹配率平均降低 5% 以内. 实际上,当视角超过 40° 后,原算法及改进算法误匹配率均超过 70%.

4 结束语

提出的采用余弦相似度匹配和距离匹配相结合的匹配方法建立约束条件,并加入方向一致性几何约束的改进方法,很好地解决了 SIFT 算法匹配阶段中距离比阈值约束小,没有考虑空间几何约束信息,导致误匹配率高,尤其对于相似区域较多的图像,出现明显的错匹配和乱匹配问题. 实验结果表明,本文中方法可以在保证匹配时间和正确匹配点数的基础上,对图像的旋转和缩放、噪声模糊、光照变化和小尺度视角变换具有较强的稳健型,与原算法相比,有效降低了误匹配率. 但对于视角变换较大时如何实现高精度的匹配,需要进一步的研究.

参考文献:

[1] Brown M, Lowe D G. Recognizing panoramas[C]// Proceedings of the 9th International Conference on

Computer Vision (ICCV03). Nice, France: IEEE, 2003:1218-1225.

- [2] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004,60(2):91-110.
- [3] 张春美,龚志辉,孙雷. 改进 SIFT 特征在图像匹配中的应用[J]. 计算机工程与应用,2008,44(2):95-97.
Zhang Chunmei, Gong Zhihui, Sun Lei. Improved SIFT feature applied in image matching [J]. Computer Engineering and Applications, 2008,44(2):95-97. (in Chinese)
- [4] Jegou H, Harzallah H, Schmid C. A contextual dissimilarity measure for accurate and efficient image search[C]// Proceedings of the Conference on Computer Vision & Pattern Recognition. Minneapolis, USA: [s. n.], 2007:1-8.
- [5] Zitova B, Flusser J. Image registration methods: a survey[J]. Image and Vision Computing, 2003, 21: 977-100.
- [6] Lowe D G. Object recognition from local scale-invariant features [C] // Proceedings of the International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999:1150-1157.
- [7] Mikolajczyk K, Schmid C. Scale & Affine invariant interest point detectors [J]. International Journal of Computer Vision, 2004, 1(60):63-86.
- [8] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(10):1615-1630.
- [9] 张炜,刘伟,普杰信. 一种基于 SIFT 和区域选择的图像拼接方法[J]. 微电子与计算机,2010,27(6):205-207.
Zhang Wei, Liu Wei, Pu Jiexin. A image stitching method based on SIFT and region selection[J]. Microelectronics and Computer, 2010, 27(6):205-207. (in Chinese)
- [10] Lindeberg T. Scale-space theory: a basic tool for analyzing structures at different scales[J]. Journal of Applied Statistics, 1994,21(2):225-270.
- [11] 张宇,刘雨东,计钊. 向量相似度测度方法[J]. 声学技术,2009,28(4):532-536.
Zhang Yu, Liu Yudong, Ji Zhao. Measurement of vector similarity [J]. Acoustic Technology, 2009, 28(4):532-536. (in Chinese)

(责任编辑:赵业玲)