

一种基于最佳相似点对的稳健模板匹配算法

王 刚^{1,2}, 孙晓亮^{1,2}, 尚 洋^{1,2*}, 于起峰^{1,2}

¹国防科学技术大学航天科学与工程学院, 湖南 长沙 410073;

²国防科学技术大学图像测量与视觉导航湖南省重点实验室, 湖南 长沙 410073

摘要 为了解决原始最好兄弟相似性(BBS)算法在剧烈非刚体变形、部分遮挡以及非均匀光照等复杂环境下匹配失败的问题,提出了一种更加稳健的模板匹配算法。将曼哈顿距离替代欧氏距离作为两个图像块之间的相似性度量,在此基础上,滑动窗口逐像素匹配得到新的由BBS响应值构成的置信度图,对该置信度图进行阈值筛选,并对剔除较小值后的置信度图滤波处理后,将最亮连通区域的中心位置定位为匹配结果。实验与分析结果表明,该算法可以有效地解决在弹性变形、相似区域干扰、部分遮挡与剧烈光照变化等变换与干扰存在情况下的图像匹配定位问题。

关键词 机器视觉; 模板匹配; 最好兄弟相似性; 置信度图

中图分类号 TP391.4 **文献标识码** A

doi: 10.3788/AOS201737.0315003

A Robust Template Matching Algorithm Based on Best-Buddies Similarity

Wang Gang^{1,2}, Sun Xiaoliang^{1,2}, Shang Yang^{1,2}, Yu Qifeng^{1,2}

¹College of Aerospace Science and Engineering, National University of Defense Technology,
Changsha, Hunan 410073, China;

²Hunan Provincial Key Laboratory of Image Measurement and Vision Navigation, National University of
Defense Technology, Changsha, Hunan 410073, China

Abstract To solve the problem of image matching failure under complex conditions including extreme non-rigid transformation, partial occlusion, and imbalanced illumination, a more robust template matching algorithm than the original best-buddies similarity (BBS) algorithm is presented. The similarity measurement between image patches is represented by Manhattan distance instead of Euclidean distance. On this basis, the new confidence map is constructed by sliding window to compute the BBS response values. The center of the brightest connected region is determined to be the last matching location after threshold filtering the confidence map by filtering process to eliminate the minor response values. Experimental and analysis results show that the proposed algorithm may be used to match the images with elastic deformation, similar region interference, partial occlusion, and extreme illumination change, etc.

Key words machine vision; template matching; best-buddies similarity; confidence map

OCIS codes 150.1135; 100.2960; 100.4994; 100.3008

1 引 言

模板匹配算法的任务是在目标图像上逐像素地滑动窗口找到与所选模板图像内容最接近的区域^[1]。模

收稿日期: 2016-10-18; 收到修改稿日期: 2016-11-08

基金项目: 国家自然科学基金(11332012)、国家重大科学仪器设备开发专项(2013YQ140517)、国家 973 计划(2013CB733100)

作者简介: 王 刚(1990—),男,博士研究生,主要从事大型结构变形测量与地形匹配制导方面的研究。

E-mail: wanggang13@nudt.edu.cn

导师简介: 于起峰(1958—),男,教授,中国科学院院士,主要从事空天图像测量与视觉导航方面的研究。

E-mail: yuqifeng@vip.sina.com

* 通信联系人。E-mail: shangyang1977@nudt.edu.cn

板匹配是计算机视觉、数字图像处理以及摄影测量与遥感等领域的一项重要处理步骤,并且有着极其广泛的应用范围,常被用于图像拼接、纹理合成、目标检测、目标识别分类、目标跟踪以及图像去噪等项目实践中^[2-10]。

目前的模板匹配算法主要包括基于灰度相关和基于几何特征两种。基于灰度的模板匹配算法通过计算像素之间的某种相似性来判断目标图像中是否存在与模板最相近的子图像。这一类方法原理简单、实现容易,但需要解决的问题是受光照影响大、抗干扰性不强、对噪声非常敏感。基于几何特征的模板匹配算法不直接使用图像灰度值计算,一般需要利用某种特征提取方法得到不变性特征,然后对这些特征进行比较。该类方法对于光照的影响不是很敏感,且能够有效地解决目标存在部分遮挡问题,通常对于比例、缩放、旋转、灰度变换具有不变性的特质。然而,基于几何特征的模板匹配算法复杂、计算量大、难以达到实时性要求,且应对不同场景需要选择各自适合的特征。

2015年,Dekel等^[11]提出了一种新颖的模板匹配算法,该算法的基本思想为统计模板与目标图像搜索窗口之间的相似点对的个数,将相似点对个数最多的窗口作为最终的匹配结果。统计相似点对的方法既可以只基于像素值,也可以基于几何特征,能够根据实际需求合理选择。与当前常用的模板匹配算法相比,如归一化互相关^[12]、统计直方图^[13]等,该方法稳健性更好、匹配精度更高,能够克服一定程度的几何变形、背景遮挡以及光照变化。然而,当模板中背景所占的比例较大或者目标存在较大遮挡时,该匹配方法将会失效。针对这些缺陷,本文提出对阈值筛选后的置信度图进行均值滤波,从而将搜索窗口周围的区域信息也考虑进来。实验结果表明,与上述方法相比较,本文方法能够有效地解决目标图像存在部分遮挡、相似性干扰或者目标背景区域所占面积较大等恶劣条件下的模板匹配问题。

2 基本原理

在模板匹配过程中,当感兴趣目标区域存在非刚体变形以及大量异常值(如背景杂乱、遮挡)时,如果按照传统的模板匹配思想,将候选窗口中所有的点都考虑进来进行相似性计算,将导致匹配到错误区域。正确的解决思路应该是只考虑模板与待选窗口中相似的点,基于这种思路,Dekel等^[11]提出了一种新的相似性度量准则,称为最好兄弟相似性(BBS)。

2.1 BBS 度量

定义两个点集 $P = \{p_i\}_{i=1}^N$ 和 $Q = \{q_j\}_{j=1}^M$, 其中 $p_i, q_j \in \mathbb{R}^d$, M, N 为点集中点的个数。存在一个点对 $\{p_i \in P, q_j \in Q\}$, 当 p_i 到点集 Q 的最邻近点为 q_j 并且 q_j 到点集 P 的最邻近点也为 p_i 时, 则称该点对为最好兄弟点对(BBP)。BBP 判别数学表达式为

$$p_{bb}(p_i, q_j, P, Q) = \begin{cases} 1, & N_N(p_i, Q) = q_j \wedge N_N(q_j, P) = p_i \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

式中 $N_N(p_i, Q) = \arg \min_{q \in Q} d(p_i, q)$, 表示 p_i 到点集 Q 的最邻近点, $d(p_i, q)$ 表示任意一种距离度量, \wedge 表示与运算。统计两个点集 P 和 Q 之间 BBP 的个数, 将 BBP 总数进行归一化处理便得到 BBS 相似性度量值,

$$S_{BB}(P, Q) = \frac{1}{\min\{M, N\}} \sum_{i=1}^N \sum_{j=1}^M p_{bb}(p_i, q_j, P, Q) \quad (2)$$

BBS 的关键特性: 1) 只依赖于待匹配点集中的 BBP, 其余的点都被当作离群点剔除; 2) 能在没有先验知识的情况下找到数据中的双向最邻近点对, 而不需要求解模板与搜索子图像之间的变形参数。

为了提高计算效率和降低噪声的干扰, 首先将目标图像和模板图像分割成多个 $k \times k$ 大小的子块, 然后将这些子块当作一个点, 由这些子块组成待匹配的点集。为了保证点集中能有足够多的点, 同时计算量又不会太大, 子块的尺寸 k 应根据模板大小调整, 当模板较大时, k 可取为 3 甚至更大; 当模板较小时, k 可取为 2 或直接将单个像素作为一个点。

计算两个点集之间的 BBS, 需要先得到每个点对之间的距离值。两个点之间的距离由两部分组成, 一部分是两点之间的颜色灰度值差异, 另一部分为两点之间的空间位置距离。目前为止, 在模板匹配中最常用的距离测度是欧氏距离。本文使用曼哈顿距离取代欧氏距离作为距离测度, 因为曼哈顿距离计算效率更高、差

异性更小,能够有效地描述两个像素块间的差异。每个点对之间的距离测度表示为

$$d(p_i, q_j) = |p_i^{(A)}, q_j^{(A)}| + \lambda |p_i^{(L)}, q_j^{(L)}|, \quad (3)$$

式中上标 A 表示颜色像素值,上标 L 表示像素点块在模板中的空间位置,归一化到 $[0, 1]$ 之间, λ 为权重系数,在实验中根据实验经验取为 1.5。

2.2 置信度图阈值筛选及均值滤波

一般的匹配搜索过程可以采用简单的逐点滑动窗口匹配,首先在搜索区域内部每一个像素点上依次计算目标图像与模板图像的相似性度量值(置信度),所有的计算结果就能构成一个置信度映射图,然后在这个映射图集合中排序置信度值,从而找到最佳匹配位置点坐标。

原始的 BBS 算法就是采用这种匹配策略,在逐个窗口计算得到 BBS 相似性度量值之后,选取所有窗口中相似性度量值最大的位置作为最终的匹配结果。然而实验结果证明,这种单纯考虑将全局最大值作为最佳匹配点的策略只利用了单个搜索窗口的信息,易受到目标姿态变化、模板背景所占比例较大以及目标中存在相似性区域干扰的影响。接下来可以通过对计算生成的置信度图进行处理,将搜索窗口周围的有效信息也考虑进来,从而提高算法的抗干扰能力。图 1 所示为模板匹配中不同阶段的置信度图,其中图 1(a)为待匹配图像,图 1(b)为原始 BBS 算法置信度图,图 1(c)为阈值筛选后置置信度图,图 1(d)为均值滤波后置置信度图。

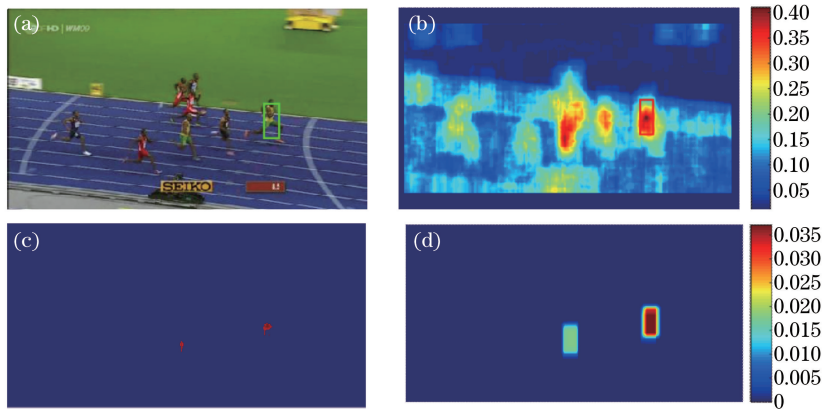


图 1 模板匹配中不同阶段的置信度图

Fig. 1 Confidence maps of different stages in template matching

如图 1(b)所示,通过计算每个搜索窗口与模板图像之间的 BBS 值,可以得到目标图像对应的 BBS 响应值置信图。首先同样需要对该置信度图的所有值进行排序,然后只保留置信度值排在前一部分的点,如图 1(c)所示。这样就缩小了待匹配的范围,剔除掉了不太可能的匹配位置。该比率的阈值按照经验通常设为 0.1%,可以表示为

$$S_{BB}(i, j) = \begin{cases} S_{BB}(i, j), & S_{BB}(i, j) > T_{0.1\%N} \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

式中 (i, j) 表示 BBS 响应值置信图的平面坐标, N 表示置信图中所有点的个数, $T_{0.1\%N}$ 表示阈值,等于排序后的置信图上第 $0.1\%N$ 个点的 BBS 响应值。

对剩下匹配可能性较大的候选点构成的置信度图进行均值滤波处理,滤波核的大小就为模板图像大小,表达式为

$$C_M = \frac{1}{MN} \text{ones}(M, N) \otimes C_{M, S_{BB} > T}, \quad (5)$$

式中 $\text{ones}(M, N)$ 表示 $M \times N$ 的单位矩阵, \otimes 表示卷积运算, $C_{M, S_{BB} > T}$ 表示 BBS 响应值大于阈值的置信图。

将滤波完成的置信度图进行阈值分割,得到与模板大小相近的多个连通区域,如图 1(d)所示,然后剩下的两个连通区域中较亮区域的中心位置作为最终的匹配结果。将连通区域中心位置代替响应最大值来定位模板匹配结果,能够有效地克服噪声或异常点的干扰,使匹配更加准确。

具体的算法流程如图 2 所示,本文模板匹配算法主要分为图像预处理、相似性图像构造以及寻找最优结

果三个部分。1)图像预处理部分:按照输入的模板大小设置图像子块的尺寸,根据图像子块的尺寸重新调整初始目标图像以及模板的大小,使其长宽均为子块尺寸的整数倍,然后将模板与目标图像均分割成多个图像子块。2)置信度图像构造部分:由(3)式计算模板与候选窗口图像子块之间的空间距离以及像素值差异,得到所有点对之间的距离测度。根据得到的距离测度由(1)式寻找模板与候选窗口间所有的最好兄弟相似点对。滑动窗口,统计所有候选窗口与模板之间的相似点对个数,并除以图像子块个数进行归一化,从而得到置信度响应值图像。3)寻找最优结果部分:设定阈值剔除相似性响应值较小的点,并对筛选后的相似性响应值图像进行均值滤波,从而得到最终的置信度图。将滤波完成的置信度图进行阈值分割,得到与模板大小相近的多个连通区域,然后剩下连通区域中较亮区域的中心位置作为最终的匹配结果。

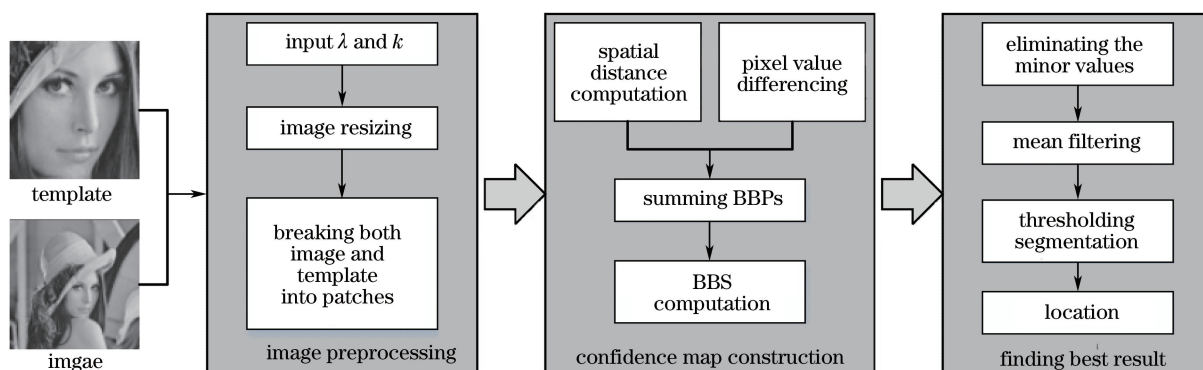


图2 模板匹配算法流程图

Fig. 2 Flow chart of template matching algorithm

3 实验结果及分析

对所提出的改进 BBS 算法在实际图像数据上做了详细的比较测试,分别进行了定性与定量分析。采用的数据集与文献[11]相同,都是从 Wu 等^[14]制作的标准视频序列图像库衍生而来的。为了方便地进行模板匹配,每段视频数据随机选出 3 组图像对,标准视频库中共有 35 段彩色视频,因而最终生成 105 组测试图像数据。模板图像与目标图像间隔 20 帧。该数据集场景复杂,具有很强的挑战性。

本文方法与包括原始 BBS 在内的 8 种类似的常用模板匹配算法进行比较,包括:1) 差值的平方和(SSD);2) 差值的绝对值和(SAD);3) 归一化互相关系数(NCC);4) 颜色直方图匹配(HM);5) 土堆移动距离^[7](EMD);6) 双向相似性^[15](BDS);7) 方向梯度直方图^[16](HOG);8) BBS。

3.1 定性评价

原始 BBS 模板匹配方法在测试算法性能时,在数据集上与上述几种经典模板匹配算法进行了比较,具体的实验结果已经发布在原文作者的个人主页上。本文主要是针对 BBS 的一些缺点进行了改进,因此这里主要列出几组改进算法与原始 BBS 算法的匹配结果对比,与其他方法的比较就不一一列出,部分结果如图 3 所示,图 3(a)为模板图像,图 3(b)则为匹配结果,其中绿色方框表示预先人工标记的真值(GT),红色方框表示 BBS 算法匹配结果,蓝色方框则表示改进算法匹配结果。

通过实验发现,在由 105 组待匹配图像对组成的测试集中,本文改进算法在原始 BBS 算法成功匹配的基础上多正确匹配了 9 组图像对,这里只列出了具有代表性的 5 组匹配结果。在这 5 组数据中,原始 BBS 算法都匹配失败,分别是由于待匹配图像中存在近似目标干扰、剧烈人体姿态变形、模板图像背景所占比例过大、部分遮挡以及非均匀光照等影响。尽管存在这些复杂的环境条件,但本文改进算法都能够成功地匹配到正确位置,充分说明了本文算法的稳健性和精确性。

3.2 定量评价

模板图像与待搜索目标图像之间的匹配的精度一般由匹配重叠率来衡量,可以表示为

$$A_{cc} = \frac{\text{area}(B_e \cap B_g)}{\text{area}(B_e \cup B_g)}, \text{ 其中 } B_g \text{ 为匹配的真值,通常由人工手动标记; } B_e \text{ 为估计解算出的矩形匹配区域; } \cap$$

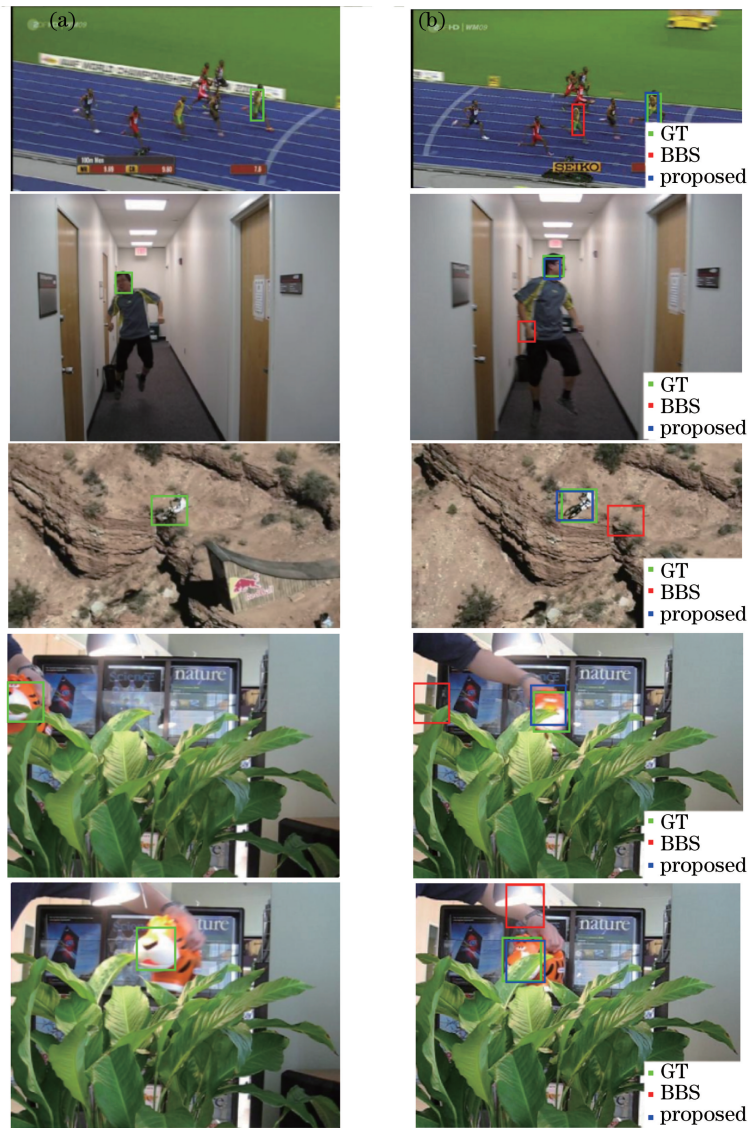


图 3 (a)本文算法与(b)BBS 算法匹配结果对比

Fig. 3 Matching results contrast between (a) proposed and (b) BBS algorithm

表示交集, \cup 表示并集。对于整个匹配图像库而言, 每组图像对都可以得到一个匹配重叠率。计算得到所有图像对的匹配重叠率之后, 统计这些匹配重叠率大于某个阈值所占总的图像对的比例, 阈值在 $[0, 1]$ 之间等间隔离散选取, 从而得到受试者工作特征 (ROC) 曲线。通过 ROC 曲线可以简单直观地分析各模板匹配算法性能, 曲线越靠近左上方, 匹配精度越高。ROC 曲线与平面坐标轴相交构成的闭合区域的面积叫做 ROC 曲线下面积 (AUC), AUC 值可以用来衡量整个数据集的平均匹配精度。

所有测试方法在上述数据集下的 ROC 曲线如图 4 所示。可以明显看出, 在所有测试方法中, 本文方法获得了最高的 AUC 得分 0.6043, 相较于目前性能最好的 BBS 方法来说, 精度提升了 5.62%。本文方法的匹配精度较高, 70% 的图像对匹配重叠率都达到了 55%。本文方法的 ROC 曲线随着重叠率阈值的增加, 下降速度较其他方法缓慢, 说明其算法性能更加稳健、稳定。

4 结 论

为了解决原始 BBS 算法在剧烈非刚体变形、部分遮挡以及非均匀光照等复杂环境下匹配失败的问题, 利用曼哈顿距离替代原来的欧氏距离作为像素点组成的图像块间的相似性度量, 并对生成的置信度图进行阈值筛选以及滤波处理, 充分考虑了搜索窗口周围的有效信息, 得到了一种更加稳健的模板匹配算法。真实

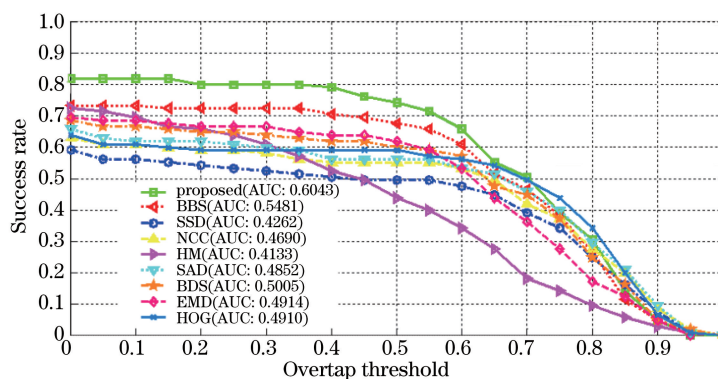


图 4 不同模板匹配方法的成功率曲线

Fig. 4 Success rate curves by different template matching methods

复杂场景下的测试集实验结果表明,与几种经典的模板匹配方法相比,该算法的匹配精度最高,并且能够有效地适应在光照不均匀、模板中异常点较多与旋转变形等多种恶劣条件下的图像匹配定位问题。该算法有望推广应用于目标跟踪、目标检测、图像检索等实际需求中。

然而,该算法的计算复杂度较高、匹配速度较慢,不能达到实时性需求,需要在接下来的工作中改进。

参 考 文 献

- [1] Ouyang W, Tombari F, Mattoccia S, *et al.* Performance evaluation of full search equivalent pattern matching algorithms[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 34(1): 127-143.
- [2] Li Qi, Dong Guofeng, Wang Qi. Object classification simulation for lidar-passive-infrared imaging combination[J]. Chinese J Lasers, 2007, 34(10): 1347-1352.
李 琦, 董国峰, 王 骥. 用于激光成像雷达和被动红外成像复合的目标分类仿真[J]. 中国激光, 2007, 34(10): 1347-1352.
- [3] Ge Liang, Zhu Qingsheng, Fu Sisi, *et al.* Application of Laws' masks to stereo matching[J]. Acta Optica Sinica, 2009, 29(9): 2506-2510.
葛 亮, 朱庆生, 傅思思, 等. Laws 纹理模板在立体匹配中的应用[J]. 光学学报, 2009, 29(9): 2506-2510.
- [4] Zhang Jianwei, Zhang Qiheng. Extended object tracking method based on straight-line edge feature matching[J]. Acta Optica Sinica, 2009, 29(10): 2826-2831.
张建伟, 张启衡. 基于直线边缘特征匹配的扩展目标跟踪方法[J]. 光学学报, 2009, 29(10): 2826-2831.
- [5] Zhao Shan, Wang Biao, Tang Chaoying, *et al.* Arm vein feature extraction and matching based on chain code[J]. Acta Optica Sinica, 2016, 36(5): 0515003.
赵 珊, 王 彪, 唐超颖, 等. 基于链码表示的手臂静脉特征提取与匹配[J]. 光学学报, 2016, 36(5): 0515003.
- [6] Dubuisson M P, Jain A K. A modified Hausdorff distance for object matching[C]. Conference A: Computer Vision & Image Processing, 1994, 1: 566-568.
- [7] Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval[J]. International Journal of Computer Vision, 2000, 40(2): 99-121.
- [8] Xue Feng, Zhang Yousheng, Jiang Julang, *et al.* Fast texture synthesis on arbitrary surfaces using texture extension and triangular texture matching[J]. Journal of Computer aided Design and Graphics, 2007, 19(2): 221-226.
薛 峰, 张佑生, 江巨浪, 等. 基于纹理延伸和三角块拼接的快速曲面纹理合成[J]. 计算机辅助设计与图形学学报, 2007, 19(2): 221-226.
- [9] Liu Wenjin, Zhang Lei. Stitching technique and its implementation of panoramic image based on template and mutual information[J]. Laser Journal, 2014, 35(9): 62-65.
刘文进, 张 蕾. 结合模板与互信息的全景图拼接技术研究 with 实现[J]. 激光杂志, 2014, 35(9): 62-65.
- [10] Zhao Haiying, Zhang Xiaoli, Li Xiongfei, *et al.* Image denoising algorithm based on multi-scale Meanshift[J]. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(5): 1417-1422.
赵海英, 张小利, 李雄飞, 等. 基于多尺度 Meanshift 图像去噪算法[J]. 吉林大学学报(工学版), 2014, 44(5): 1417-1422.
- [11] Dekel T, Oron S, Rubinstein M, *et al.* Best-buddies similarity for robust template matching[C]. IEEE Conference on

-
- Computer Vision and Pattern Recognition, 2015: 2021-2029.
- [12] Yoo J C, Han T H. Fast normalized cross-correlation[J]. Circuits Systems & Signal Processing, 2009, 28(6): 819-843.
- [13] Pérez P, Hue C, Vermaak J, *et al.* Color-based probabilistic tracking[J]. 7th European Conference on Computer Vision, 2002: 661-675.
- [14] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418.
- [15] Simakov D, Caspi Y, Shechtman E, *et al.* Summarizing visual data using bidirectional similarity[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2008: 1-8.
- [16] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 1: 886-893.