

基于 Leap Motion 的三维动态手指姿势识别

严利民^{1,2} 杜斌^{1,2*} 潘浩^{1,2} 郭强^{1,2}

¹上海大学新型显示技术及应用集成教育部重点实验室, 上海 200072

²上海大学微电子研究与开发中心, 上海 200072

摘要 为实现基于运动轨迹信息的动态手指姿势的识别,促进人机交互和虚拟现实交叉领域的发展,介绍了一种基于 Leap Motion 的三维(3D)空间内的指势关键特征点轨迹识别的方法。将 Leap motion controller 获取的模型化的手部信息经过手部骨架算法提取出手部骨架目标,然后,利用细化算法得到手指直线示意图,并提取指势关键特征点及运动轨迹。利用支持向量机(SVM)对 3D 基本动态指势进行识别和评估。实验结果表明,该方法能够在较为复杂的 3D 空间内,即使手部出现自遮挡的情况下,实现动态指势的识别,准确率较高,稳健性强。

关键词 图像处理; 人工智能; 动态指势识别; 三维空间; 关键特征点; 运动轨迹; 支持向量机

中图分类号 TP751.1 **文献标识码** A

doi: 10.3788/LOP53.111001

Recognition of Three-Dimensional Dynamic Finger Gesture Based on Leap Motion

Yan Limin^{1,2} Du Bin^{1,2} Pan Hao^{1,2} Guo Qiang^{1,2}

¹Key Laboratory of Advanced Displays and System Application, Ministry of Education, Shanghai University, Shanghai 200072, China

²Microelectronics Research and Development Center, Shanghai University, Shanghai 200072, China

Abstract In order to achieve the recognition of dynamic finger gestures based on motion trajectory and promote the development of the cross-cutting areas of human-computer interaction and virtual reality, a recognition method of the key feature points' trajectory is described, which is about finger gesture in three-dimensional (3D) space on basic of Leap Motion. The hand skeleton is extracted by skeleton algorithms with the modeling hand information, which is obtained by the Leap motion controller. Then, the thinning algorithm is used to get straight finger schematic, and extracts feature points and motion trajectory. The basic dynamic 3D finger gestures are identified and evaluated by means of support vector machine (SVM). Experimental results show that the identification of dynamic 3D finger gestures can be achieved in a more complex 3D space or hand block by themselves, and has a high accuracy and robustness.

Key words image processing; artificial intelligence; dynamic gesture recognition; three-dimensional space; key feature points; motion trajectory; support vector machine

OCIS codes 100.3008; 100.2000; 100.4996

1 引言

随着人机交互与虚拟现实技术的飞速发展,虚拟操作人机交互应运而生。传统人机交互技术在很大程度上依赖于鼠标、键盘、触摸屏等输入设备,这种接触式交互已经无法满足更多的需求。而虚拟操作人机交互过程中采用计算机视觉获取人手信息,在交互的自然性与成本上有很大的优势,是现如今人机交互领域发展的主要趋势^[1]。

收稿日期: 2016-05-25; **收到修改稿日期:** 2016-07-25; **网络出版日期:** 2016-10-21

基金项目: 国家自然科学基金(61376028)

作者简介: 严利民(1971—),男,博士,副教授,主要从事集成电路应用、智能检测和控制等方面的研究。

E-mail: yanlm@t.shu.edu.cn

* **通信联系人.** E-mail: 18717865628@163.com

在基于机器视觉的人机交互中,人手的基本信息无法直接被计算机获取,手指姿势参数的估计成为了至关重要的工作,只有较为准确的手指姿势参数才能驱动三维(3D)空间内的虚拟化的手部模型,从而保证人机交互虚拟和现实的一致性。

3D 信息获取及计算一直是视觉手指姿势估计的瓶颈问题。武汇岳等^[2]提出了 3D 动态手势状态转移模型,利用 Kinect 自带的 SDK 获取由 20 个关节所构成的人体骨骼模型,并将人体左右手的关节点作为特征点进行跟踪识别,但其准确性和精精度无法得到保证。蔡文等^[3]将红外 LED 和定向反射材料结合,形成定向反射空间,从而检测到手指在空间中的位置和轨迹,但其对设备和环境依赖性太强,移植性较差。

国外研究人员对该领域的研究也日益深入,推进了 3D 手指虚拟操作人机交互的快速发展。Raheja 等^[4]利用 Kinect 摄像机采集手部图像虽然可以获得稳定的 RGB-D 图像序列,但必须经过复杂的图像处理才能最终识别出手部信息,并且其硬件价格高昂,SDK 相对封闭,性价比不高是其巨大的软肋;Ren 等^[5]通过建立 3D 手部图像对手指进行分析,但计算量很大,难以达到实时性的要求。

针对上述存在的问题,本文介绍了一种基于开源设备 Leap Motion 的虚拟操作人机交互姿势识别方法,将获取的手部 3D 信息精简处理,提取指势关键特征点的连续 3D 坐标并计算出其运动轨迹,利用支持向量机(SVM)对 Leap motion controller 中基本姿势进行识别和评估。除此之外,Leap motion controller 具有运行速度高达 100 frame/s,精度高达 1/100 mm 的优势。经实验验证,该方法有较高的准确率、精精度以及较好的实时性。

2 运动目标的检测

指势分析可以分为三大模块:手部目标检测、关键特征点提取和识别评估,其中目标检测是识别方案的基础。目前已有的方案大多是建立在对手部图像的 3D 重建后手指特征点的全部提取,这样虽然有利于提高识别率,但在数据量本来就很庞大的 3D 空间内更进一步地增大了计算量,复杂程度随着特征点的增多而剧烈增大。

Leap motion controller 内置手部识别模块对手部模型进行识别,判别出手指方向矢量,并通过两个超广角红外传感器实时收集的数据不断验证这个模型^[6]。更强大的是 controller 可以在手指没有完全暴露的情况下跟踪手指的位置,即当手指被其他手部遮挡时,controller 根据已经收集的数据匹配出手部模型,如图 1 所示。

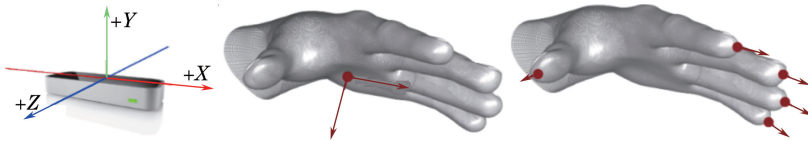


图 1 Leap motion controller 及手部模型

Fig. 1 Leap motion controller and hand model

在 Leap motion controller 25~600 mm(呈倒立四棱锥体状)的有效探测范围,结合人体手部骨骼结构对人手的 9 块骨头、29 个关节点进行探测^[7]。每个手指由 4 个骨节组成,分别为远端指骨、中间指骨、近端指骨以及掌骨,可以将人手的 3D 信息精简为手部骨架信息,实现代码如下:

```
Leap::FingerList fingers = frame.hands()[0].fingers();
for(Leap::FingerList::const_iterator fl = fingers.begin(); fl != fingers.end(); fl++)
{
    Leap::Bone bone;
    Leap::Bone::Type boneType;
    for(int b = 0; b < 4; b++)
    {
        boneType = static_cast<Leap::Bone::Type>(b);
        bone = (*fl).bone(boneType);
        std::cout << "Finger index: " << (*fl).type() << " " << bone << std::endl;
    }
}
```

图 2(a)为手掌照片图。鉴于拇指没有掌骨的特殊性,在不影响手指姿势的前提下,将掌骨综合成类似矩形的手掌区域,以便更好地分离出手指,并进行特征点提取和分割,效果如图 2(b)所示。可以看出,手指、手掌目标已经明朗,并且简化后的目标可以有效去除不同人手掌不规则形状对识别的影响。

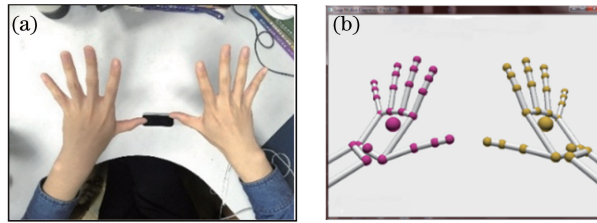


图 2 运动目标检测示意图

Fig. 2 Schematic of moving target detection

3 特征点提取

3.1 特征点提取算法

得到简化后的手部骨架图后,为更好地获取手指信息并剥离对指势识别影响较小的手掌信息,减少计算量及简化跟踪过程,本文将简化后的骨架图经过细化算法,以手指关节处球状示意图的球心为端点,将手指细化为单像素手指直线图,并且标出手指方向,如图 3(a)所示。图 3(b)所示为将单像素的手指直线在特征点提取后标记为不同的颜色。

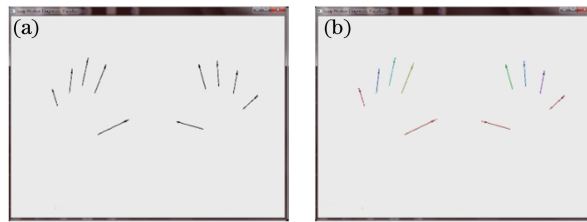


图 3 手指直线示意图

Fig. 3 Schematic of the straight finger

单像素手部直线图虽然去除了大量特征点,但是关于手指运动轨迹的特征点并没有去除,手指上每一个像素的运动轨迹都与指势运动相关联,其中处于手指末端的端点(指尖点)就是手势关键特征点。

如图 4 所示,设像素点 $I(I_x, I_y, I_z)$ 为某一帧手指直线图中的手指指向端的一个像素,其中 I_x, I_y, I_z 分别代表像素点 I 的空间坐标, $I_x \leq X, I_y \leq Y, I_z \leq Z$, X, Y 和 Z 分别表示图像的长、宽和高。 $I_n(I_x, I_y, I_z)$ 为后续第 n 帧单像素骨架图, $n=0, 1, 2, \dots$ 。像素点示意图如图 4 所示。

1) 判断 $I(I_x, I_y, I_z)$ 是否为 0,若是,去除点像素点 I ;若不是,保存像素点 I ,得到骨架完整像素,进入 2)。

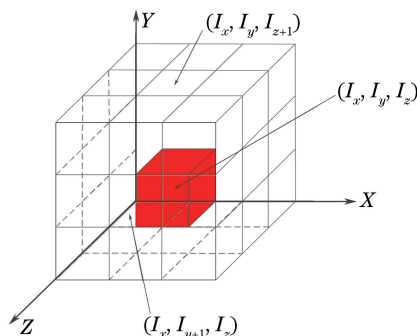


图 4 像素点示意图

Fig. 4 Schematic diagram of pixel points

2) 令 $I(I_x, I_y, I_z)$ 空间周围的其他 26 个像素点像素之和为 Q , 表达式为 $Q = I(I_{x\pm 1}, I_y, I_z) + I(I_x, I_{y\pm 1}, I_z) + I(I_x, I_y, I_{z\pm 1}) + I(I_{x\pm 1}, I_{y\pm 1}, I_{z\pm 1}) + I(I_{x\pm 1}, I_y, I_{z\pm 1}) + I(I_{x\pm 1}, I_{y\pm 1}, I_z) + I(I_{x\pm 1}, I_{y\pm 1}, I_{z\pm 1})$, 判断 Q 是否大于 3, 若是, 去除像素点 I ; 若不是, 保存像素点 I , 进入 3)。

3) 当 $Q=1$ 时, 手指直线图的 I 点就是端点, 即特征点(指尖点), 记为 $I'(I'_x, I'_y, I'_z)$ 。

4) 当 $Q=2$ 时, $I(I_x, I_y, I_z)$ 周围的两个像素点为 $I(I_{x1}, I_{y1}, I_{z1})$ 和 $I(I_{x2}, I_{y2}, I_{z2})$, 只要判断此两个像素点在 $I(I_x, I_y, I_z)$ 的同侧, 则可知 I 点就是端点, 即 $\prod_{n=1}^2 (I_{xn} - I_x)$, $\prod_{n=1}^2 (I_{yn} - I_y)$, $\prod_{n=1}^2 (I_{zn} - I_z)$ 同时满足大于等于 0, 则记为 $I'(I'_x, I'_y, I'_z)$ 为提取的特征点。

5) 当 $Q=3$ 时, 由 step4 可知, 周围三个像素点中任意两个点都在 $I(I_x, I_y, I_z)$ 的同侧, 则这三个点就都在 $I(I_x, I_y, I_z)$ 同侧, 同样可得手指的端点, 记为 $I'(I'_x, I'_y, I'_z)$ 。

得到手势动作关键特征点后, 就可以提取关键特征点的坐标, 然后获取每一帧关键特征点的数据。以“Swip”3D 动态指势为例, 为区分不同的手指, 将单像素的手指直线在特征点提取后标记为不同的颜色[图 3(b)], 得到运动轨迹, 如图 5 所示。

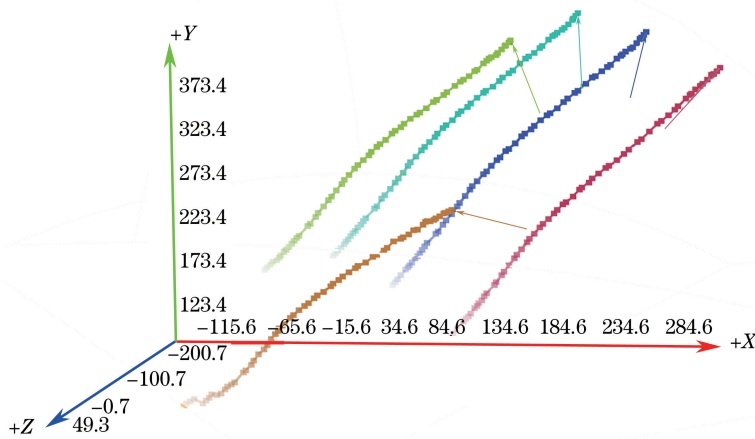


图 5 “Swip”动态指势关键特征点轨迹

Fig. 5 "Swip" dynamic finger point trajectory

3.2 算法对比

针对 3D 空间的指尖定位识别算法, 采样 20 个不同的个体, 以双手为例进行测试, 每个个体以任意姿势重复 10 个手指各 10 次, 共计 2000 次作为样本总量。获得最终结果与传统 Kinect 和多目摄像头的 3D 空间指尖识别方法进行对比, 如表 1 所示。

表 1 本文算法与文献[8]、[9]算法性能对比

Table 1 Performance comparison of proposed algorithm and Ref. [8], Ref. [9] algorithms

Detection method	Basic equipment	Number of fingers	Self-occlusion	Recognition range /cm	Recognition accuracy	Recognition rate /%
Proposed method	Leap motion	Multiple	Identifiable	2.5~60	1/100 mm	98.50
Ref. [8] method	Kinect	Single	Unidentifiable	50~150	1 cm	94.40
Ref. [9] method	Two cameras	Single	Unidentifiable	30~120	Low	92.00

从表 1 可以看出, 使用 Leap Motion 作为获取设备的识别率高达 98.50%, 远远高于其他两种方案。除此之外, 支持多手的跟踪识别以及抗自遮挡的功能也是传统算法无法比拟的。虽然基于 Leap Motion 的识别方案的识别范围有限, 约为 2.5~60 cm 的倒立四棱锥体状空间, 但是较高的识别精度使其可以捕捉到较为微小的手指移动信息。Kinect 和多目摄像头获取的信息需要经过复杂的算法进行预处理, 如肤色、滤波等, 在 3D 空间数据量较大的固有属性下大大增加了计算的复杂度, 而 Leap Motion 采样信息匹配手部后逐步进行简化, 计算量相对较少。Leap Motion 的性能优势为下一步动态指势识别的准确性打下良好的基础。使用同样的测试数据基于相同的 Leap Motion 设备, 而采用不同的关键特征点提取算法(指尖特征点、关节

特征点、手掌模型)进行对比,结果如表 2 所示。

表 2 本文算法与文献[10]、[11]算法对比

Table 2 Comparison of proposed algorithm and Ref. [10], Ref. [11] algorithms

Detection method	Algorithm scheme (number)	Frame rate / (frame/s)	Real-time	Algorithm complexity	recognition rate /%
Proposed method	Fingertip feature points (1)	90 ± 5	Good	Simple	98.50
Ref. [10] method	Joint feature points(21)	65 ± 5	General	Complex	99.20
Ref. [11] method	Palm model (100+)	40 ± 5	Poor	Very complex	94.33

从表 2 可以看出,采用简化的指尖特征点作为关键点的性能表现较为优越。关节特征点的关键点数量约为 21,即手指骨节关节数量,手掌模型的特征点的数量则远远大于其他两种,关键特征点的数量大小直接影响到算法处理的速度及算法的复杂性。本文算法的实时性较好,帧率约为(90 ± 5) frame/s 左右,虽然识别率略低于关节特征点的 99.20%,但是权衡其他几个方面性能,依旧是较为理想的算法选择。文献[11]的手掌模型将手掌所有的像素点作为关键点,包含大量无关指势的特征点和影响指势判别的特征点,导致其实时性较差、运行速率较慢且识别率明显低于其他算法。

4 识别和评估

为了验证本文特征点运动轨迹的动态指势识别方法的有效性,针对 7 种基本动态指势进行识别和评估。

4.1 测试数据

实验是在配置为英特尔(R)Core(TM) i3-4160@3.60 GHz 双核处理器,8G 1600 MHz 金士顿 DDR3 内存,Windows 7 旗舰版 64 位操作系统的台式机上完成。基本的动态指势包括画圈、点击、按、拉/推、滑动、拍手、双手滑动,如图 6 所示。总的测试数据为 1400,分别由 20 个人重复 10 次获得。

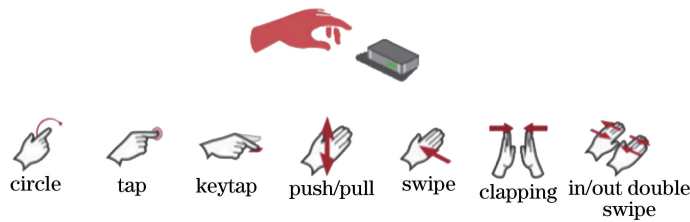


图 6 基本动态手指姿势

Fig. 6 Basic dynamic finger gestures

4.2 分类器

动作识别技术最后一步是分类识别,目前常用的有 SVM^[12]、神经网络^[13-14]、KNN^[15]等方法。神经网络存在一些缺陷,如参数采用经验方法确定,具有盲目性,难以获得全局最优参数,并且收敛的速度比较慢,这些不足导致神经网络在 3D 空间识别中,识别率时高时低,结果不稳定。而 KNN 算法对样本的不平衡较为敏感,样本量大小对其结果又有较大影响,针对样本量较小的类域误差较大,且计算量较大。相比之下,SVM 能够较好地解决小样本非线性的分类问题,不存在局部极优解等缺陷,计算量相对较小,因此采用 SVM 作为分类模型用于 3D 空间指势的识别。

4.3 评估方法

本文方法选最常用的留一法^[16]。在样本不大的情况下,将全部样本分为 N 份,用 $(N-1)$ 份构建预测模型,最后的 1 份作为基准测试集,平均 N 次实验性能作为评估的结果,更具有说服力。

5 实验结果及分析

运用上文方案提取出特征点,并对基础动态指势跟踪国际,按照留一法对实验进行评估。图 7 给出了基于 Leap Motion 的关键特征点运动轨迹的指势识别性能的混淆矩阵。

从识别结果的混淆矩阵中可以直观地看出,针对简化后的骨架图进行的手指直线细化算法较为完好地

保留了手部手指信息,平均识别率约为 94.29%,这也说明了本文方法对动态指势识别具有较高的稳健性。其中, key tap 和 tap 手指动作幅度较小,收集获取的信息量较少,导致了 key tap 和 tap 之间的误判较大。

	circle	swipe	key tap	screen tap	pull/push	clapping	double swipe
circle	0.98		0.02				
swipe		1					
key tap			0.91	0.09			
screen tap			0.1	0.90			
pull/push		0.02			0.95	0.03	
clapping						0.94	0.06
double swipe						0.08	0.92

图 7 动态指势识别的混淆矩阵

Fig. 7 Dynamics finger recognition of confusion matrix

6 结 论

Leap Motion 作为新兴的科技产品,其优越的计算性能克服了传统设备的实时性差、移植性差、性价比低等缺陷,渐渐成为了研究的热点。基于 Leap Motion 获取人体手部目标,结合手部生理构造简化出骨架图,细化出手指直线图,寻找出关键特征点作为动态手指姿势的描述子,结合 SVM 对 7 种基本动态指势进行了识别和评估。实验数据表明,选取 Leap Motion 为数据源,指尖提取算法以及 SVM 分类最终识别出特定指势具有较为理想的效果。

整体解决方案使用 Leap Motion 作为数据采集源,并通过数据比照选择出较为理想的算法,融合成一套 3D 动态手指姿势识别系统。所采用的方案在较为复杂的 3D 空间内具有实时性较好、识别率较高、移植性较好的独特优势。在现有的虚拟操作人机交互方法中,本文方法更具有现实意义,优势突出,且进一步促进了虚拟现实和人机交互交叉领域的发展。

参 考 文 献

- Cai Jiabin, Feng Guocan, Tang Xin, *et al.* Human action recognition by leaning pose dictionary[J]. Acta Optica Sinica, 2014, 34(12): 1215002.
蔡加欣, 冯国灿, 汤鑫, 等. 基于姿势字典学习的人体行为识别[J]. 光学学报, 2014, 34(12): 1215002.
- Wu Huiyue, Zhang Fengjun, Liu Yujin, *et al.* Vision-based gesture interfaces toolkit for interactive games[J]. Journal of Software, 2011, 22(5): 1067-1081.
武汇岳, 张凤军, 刘玉进, 等. 基于视觉的互动游戏手势界面工具箱[J]. 软件学报, 2011, 22(5): 1067-1081.
- Cai Wen, Tian Feng, Tian Jing, *et al.* True three-dimension touch control system based on optical[J]. Electronic Measurement Technology, 2013, 36(2): 47-50.
蔡文, 田丰, 田晶, 等. 基于光学的真三维触控系统设计[J]. 电子测量技术, 2013, 36(2): 47-50.
- Raheja J L, Chaudhary A, Singal K. Tracking of fingertips and centers of palm using Kinect[C]. 2011 Third International Conference on Computational Intelligence, Modelling and Simulation (CIMSIM), 2011: 248-252.
- Ren Z, Yuan J, Zhang Z. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera[C]. Proceedings of the 19th ACM International Conference on Multimedia, 2011: 1093-1096.
- Spiegelmock M. Leapmotion development essentials[M]. Birmingham: Packt Publishing Ltd, 2013.
- Huang Jun, Jing Hong. Gesture control research based on leap motion[J]. Computer Systems & Applications, 2015, 24(10): 259-263.
黄俊, 景红. 基于 Leap Motion 的手势控制技术初探[J]. 计算机系统应用, 2015, 24(10): 259-263.
- Yan Hao, Zhang Mingmin, Tong Jing, *et al.* Real time robust multi-fingertips tracking in 3d space using Kinect[J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(12): 1801-1809.
晏浩, 张明敏, 童晶, 等. 基于 Kinect 的实时稳定的三维多手指跟踪算法[J]. 计算机辅助设计与图形学学报, 2013, 25(12): 1801-1809.

- 9 Guo Kangde, Zhang Mingmin, Sun Chao, *et al.* 3D fingertip tracking algorithm based on computer vision[J]. Journal of Computer Research and Development, 2010, 47(6): 1013-1019.
郭康德, 张明敏, 孙超, 等. 基于视觉技术的三维指尖跟踪算法[J]. 计算机研究与发展, 2010, 47(6): 1013-1019.
- 10 Hu Hong, Chao Jiangang, Yang Jin, *et al.* Key point model for hand pose estimation based on Leap Motion[J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(7): 1211-1216.
胡弘, 晁建刚, 杨进, 等. Leap Motion 关键点模型手姿态估计方法[J]. 计算机辅助设计与图形学学报, 2015, 27(7): 1211-1216.
- 11 Moldovan C C, Staretu I. Motion Leap compared to data gloves in human hand tracking [M]. Berlin: Springer International Publishing, 2015: 195-202.
- 12 Gieser S N, Boisselle A, Makedon F. Real-time static gesture recognition for upper extremity rehabilitation using the leap motion[M]. Berlin: Springer International Publishing, 2015: 144-154.
- 13 Yang Jiang, Zhao Jian, Yang Chunni, *et al.* All-fiber perimeter security system based on michelson interferometer and pattern recognition[J]. Chinese J Lasers, 2014, 41(11): 1105005.
杨江, 赵健, 杨春妮, 等. 基于迈克耳孙干涉和模式识别的全光纤周界安防系统[J]. 中国激光, 2014, 41(11): 1505005.
- 14 Zhou Liangji, Li Qingwu, Huo Guanying, *et al.* Face image recognition method based on the NSCT and bionic pattern [J]. Laser & Optoelectronics Progress, 2015, 52(3): 031001.
周亮基, 李庆武, 霍冠英, 等. 基于 NSCT 和仿生模式的人脸图像识别方法[J]. 激光与光电子学进展, 2015, 52(3): 031001.
- 15 Hou Yuting, Peng Jinye, Hao Luwei, *et al.* Research of classification method for natural images based on adaptive feature-weighted K-nearest neighbors[J]. Application Research of Computers, 2014, 31(3): 957-960.
侯玉婷, 彭进业, 郝露微, 等. 基于 KNN 的特征自适应加权自然图像分类研究[J]. 计算机应用研究, 2014, 31(3): 957-960.
- 16 Gao Zan, Zhang Hua, Cai Anni. Discussion on the assessment strategy of action recognition algorithms[J]. Journal of Optoelectronics • Laser, 2012, 23(6): 1166-1172.
高赞, 张桦, 蔡安妮. 动作识别算法的评估策略探讨[J]. 光电子 • 激光, 2012, 23(6): 1166-1172.