

DOI: 10.11992/tis.201708003

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180402.1805.012.html>

对偶树复小波与空域信息的手势识别分类研究

贾鹤鸣¹, 朱传旭¹, 张森¹, 杨泽文², 何东旭²

(1. 东北林业大学 机电工程学院, 黑龙江 哈尔滨 150040; 2. 哈尔滨工程大学 自动化学院, 黑龙江 哈尔滨 150001)

摘 要: 为提高手势识别中特征获取的有效性, 本文提出空域特征与对偶树复小波变换特征相结合的融合特征, 主要包括水平位置、竖直位置、长宽比、矩形度、Hu 矩 7 个分量, 及 11 维空域特征与对偶树复小波变换的 16 维特征进行融合后得到的 27 维特征。针对分类器优化算法, 提出进行训练样本优选的最优距离-支持向量机 (BD-SVM) 分类方法。最后的实验结果表明, 对“1~9”手势进行测试, 当采用径向基核函数时, 平均识别精度最高, 为 90.33%, 平均识别时间为 0.026 s, 说明所提出的方法能够较好地进行静态手势识别, 具有较高的训练速度和辨识精度。

关键词: 手势识别; 空域特征; 对偶树复小波; 特征融合; 分类器优化; BD-SVM; 径向基核函数; 静态测试
中图分类号: TP273 **文献标志码:** A **文章编号:** 1673-4785(2018)04-0619-06

中文引用格式: 贾鹤鸣, 朱传旭, 张森, 等. 对偶树复小波与空域信息的手势识别分类研究[J]. 智能系统学报, 2018, 13(4): 619-624.

英文引用格式: JIA Heming, ZHU Chuanxu, ZHANG Sen, et al. Research on gesture recognition and classification of dual-tree complex wavelet and spatial information[J]. CAAI transactions on intelligent systems, 2018, 13(4): 619-624.

Research on gesture recognition and classification of dual-tree complex wavelet and spatial information

JIA Heming¹, ZHU Chuanxu¹, ZHANG Sen¹, YANG Zewen², HE Dongxu²

(1. College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China; 2. College of Automation, Harbin Engineering University, Harbin 150001, China)

Abstract: To improve the validity of features obtained in gesture recognition, in this paper, we propose a fusion feature that combines spatial and dual-tree complex wavelet transform features. These features mainly include seven components (horizontal position, vertical position, aspect ratio, rectangular degree, Hu moments, etc.) and 27 dimensional features, comprising 11 dimensional spatial features and 16 dimensional dual-tree complex wavelet transform features. We employ the optimal distance support vector machine (BD-SVM) classification method to optimize training samples for the classifier optimization algorithm. The experimental results show that, in a test of gestures “1~9” using the RBF kernel function, the highest average recognition accuracy is 90.33% and the average recognition time is 0.026 s. These results reveal that the proposed method demonstrates excellent static gesture recognition, a high training speed, and accuracy in identification.

Keywords: gesture recognition; spatial feature; dual-tree complex wavelet; feature fusion; classifier optimization; BD-SVM; radial basis kernel function; static test

收稿日期: 2017-08-03. 网络出版日期: 2018-04-03.

基金项目: 中央高校基本科研业务费专项资金项目 (2572014BB03);
国家自然科学基金项目 (31470714, 51609048); 黑龙江省研究生教育创新工程项目 (JGXM_HLJ_2016014).

通信作者: 贾鹤鸣. E-mail: jiaheminglucky99@126.com.

手势语言作为一种常用的交流语言, 通过不同手势的组合、不同手形的变化, 能够表达多种复杂的含义, 在非声音传递信息的方式中, 如特

警手语、聋哑人交流、远程指挥等方面,起到重要作用^[1]。特征提取是手势识别的关键环节,即通过对不同的手势进行数学描述,表示独特的特征关系。常用的特征有几何特征、形状特征、Hu矩特征等。

国外方面, Davis 等^[2]采用不变矩和神经网络,进行数字手势的识别; Miyashita 等^[3]针对静态手势具有平移和旋转性问题,采用基于边界信息特征描述的方法; Kapuściński 等^[4]提出改进的梯度直方图方法进行静态手势识别,通过加强图像坐标系的局部方向,使之能够找到最佳的关键向量,提高识别率。国内方面,谭台哲等^[5]将深度与肤色信息结合,根据手指和手掌质心判断方向,从而完成对手势和左右手的识别,鲁棒性强;刘淑萍等^[6]在传统的 HOG 方法的基础上加入肤色模型和手指检测算法,将识别率提高 20%;赵磊等^[7]对传统 DTW 算法进行改进,以欧氏距离为元素组成矩阵,结合事先录入的手势模板,实现手势识别。通过上述国内外研究现状的分析不难发现研究问题主要集中于空域特征的提取。实际上,频谱能在描述纹理特征时体现重要作用^[8]。常用的频域变换主要有傅里叶变换、小波变换等。前期研究中,大多仅采用空域特征或者仅采用小波变换系数特征,但由于传统小波变换在获取频谱特征时,不具有方向性,因此,拟采用对偶树复小波变换提取频域特征,实现 -75° 、 -45° 、 -15° 、 15° 、 45° 和 75° 等6个方向的特征度量,最后为提高特征的完备性,本文将空域特征与频域特征进行融合。

分类器通常采用人工神经网络,常用的有 BP 神经网络、支持向量机(SVM)等。经过许多学者研究证实,在图像识别方面, SVM 比 BP 精度更高,且不易陷入过拟合状态^[9-12]。但是,由于 SVM 在高维空间中计算量远远大于 BP,因此,当样本较多、特征维数较大时, SVM 速度较慢。本文引入 BD(best distance)优选 SVM 训练样本的算法,一方面减少样本数,提高训练速度,另一方面,优化样本质量,提高训练精度。

1 特征提取

1.1 空域特征

1) 位置。由前面部分手势跟踪,可获取手势的位置信息,用掌的面积中心表示。设手掌像素的坐标为 (x_i, y_i) ,其中 $(i=0,1,\dots,n-1; j=0,1,\dots,m-1)$,那么其质心 (\bar{x}, \bar{y}) 表示为

$$\bar{x} = \frac{1}{mn} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} x_i, \quad \bar{y} = \frac{1}{mn} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} y_j$$

2) 长宽比。手所在的最小矩形,长度定义为 L ,宽度定义为 W ,长宽比 r 定义为

$$r = W/L$$

3) 矩形度。手势所占面积 S_0 与最小外接矩形面积 S_m 之比。矩形度 a 表示为

$$a = S_0/S_m$$

4) Hu 矩特征。Hu 矩特征主要描述图像旋转的不变性,包含7个分量。对于一幅 $N \times M$ 的图像 $f(x, y)$,则其 $(p+q)$ 阶矩为

$$M_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y)$$

通过计算函数的重心,得到:

$$\bar{x} = M_{10}/M_{00}, \quad \bar{y} = M_{01}/M_{00}$$

对中心矩进行归一化后,使用二阶和三阶矩构造7个 h 分量矩:

$$\begin{aligned} h_1 &= \eta_{20} + \eta_{02} \\ h_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ h_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ h_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ h_5 &= (\eta_{03} - 3\eta_{12})(\eta_{30} + \eta_{12}) + [(\eta_{30} + 3\eta_{12})^2 - \\ &\quad 3(\eta_{21} + \eta_{03})^2](3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \cdot \\ &\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ h_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &\quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ h_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - \\ &\quad 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{30})(\eta_{21} + \eta_{03}) \cdot \\ &\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

1.2 单位的书写规则

用函数 $\psi_h(t)$ 和 $\psi_g(t)$ 将对偶树复小波定义为

$$\psi_c(t) = \psi_h(t) + j\psi_g(t)$$

且满足希尔伯特变换,即:

$$\psi_g(\omega) = \begin{cases} -j\psi_h(\omega), & \omega > 0 \\ \psi_h(\omega), & \omega < 0 \end{cases}$$

用树 A 和树 B 表示,结构图如图1所示。

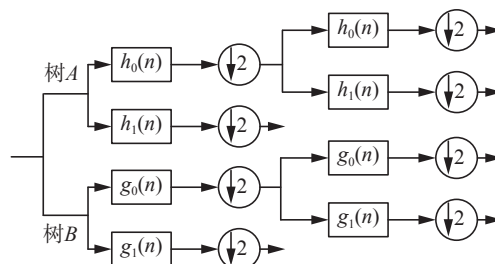


图1 对偶树复小波变换结构图

Fig. 1 Dual tree complex wavelet transform structure

通过 $\psi(t) = \psi_h(t) + j\psi_g(t)$, 计算二维对偶树复小波变换:

$$\begin{aligned}\psi(x, y) &= \psi(x)\psi(y) = \\ &[\psi_h(x) + j\psi_g(x)][\psi_h(y) + j\psi_g(y)] = \\ &\psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) + \\ &j[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)]\end{aligned}$$

输入图像后, 采用滤波器 $h_0(n)$ 与 $h_1(n)$ 、 $g_0(n)$ 与 $g_1(n)$ 进行两步分解操作, 获得每个子带对应的系数, 通过做差求取对偶树复小波变换的低频分量和高频分量。对偶树复小波变换后, 实部与虚部分量方向相同, 因此可进行6个方向的度量, 分别对应 -75° 、 -45° 、 -15° 、 75° 、 45° 、 15° , 如图2所示。

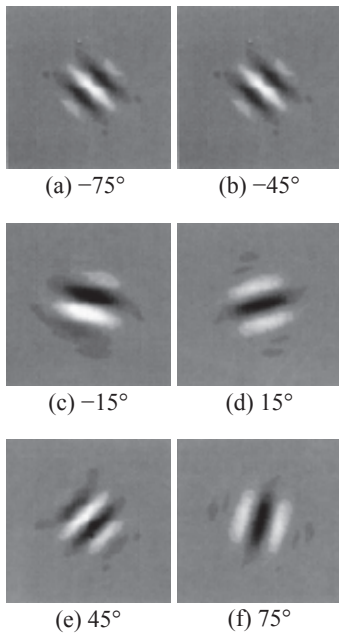


图2 对偶树复小波变换的6个方向性

Fig. 2 Six directions of dual tree complex wavelet transform

2 BD-SVM 分类器

根据传统 SVM 算法原理, 根据几何特征, 支持向量即同类距离最小, 异类距离最大的两类样本, 若训练样本未进行优化, 则寻找支持向量的过程就会耗时严重, 因此, 改进思路聚集在对训练样本的几何特征上。本文将训练样本的点间距离作为度量, 通过优化训练样本集, 提高 SVM 的训练速度, 也因样本集的品质提高, 从而间接提高分类器的识别精度。具体数学描述如下。

将样本点分布于空间中, 定义任意两点的距离为

$$\begin{aligned}d(\mathbf{x}_i, \mathbf{x}_j) &= \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)} = \\ &\sqrt{2 - 2K(\mathbf{x}_i, \mathbf{x}_j)}\end{aligned}\quad (1)$$

高维空间点定义为 z_k , 则 z_k 对应一对距离值

(d_k^+, d_k^-) , 具体表示如下:

$$\begin{cases} d_k^+ = \frac{1}{l^+} \sum d(z_k, \mathbf{x}_i) \\ d_k^- = \frac{1}{l^-} \sum d(z_k, \mathbf{x}_j) \end{cases}$$

则存在临界值 c_1 和 c_2 , 使得 $P(d_k^+ > c_1) = t$, $P(d_k^- < c_2) = t$, 支持向量就是满足 $\{z_k | d_k^+ > c_1, d_k^- < c_2\}$ 的 z_k 点。算法设计的具体步骤如下。

1) 设置比例系数 $t(0 < t < 1)$ 。

2) 在高维空间中, 按照式 (1) 计算距离矩阵 $D = \{(d_{ij})_{l \times l}\}$ 。

3) 令 $l^+ = \frac{1}{l^+} e_{l^+ \times 1}$, $l^- = \frac{1}{l^-} e_{l^- \times 1}$, 计算平均距离矩阵 $V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}_{(l^+ + l^-) \times 2}$ 。

4) 对矩阵 V 进行按序排列, 其中, V_{11} 和 V_{22} 降序, V_{12} 和 V_{21} 升序, 提取 $1 \cdot t$ 个样本, 得到新的训练样本集 $|T^{\text{new}}|$ 。

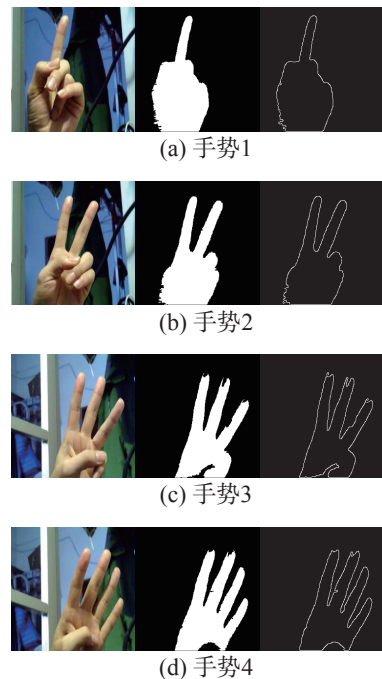
5) 在新样本集上利用 SVM 建立模型, 并进行测试。

3 结果与分析

3.1 特征提取

静态手势识别选取 60 幅“1~9”手势的图像作为训练样本, 30 幅作为测试样本。包括空域及频域特征提取, BD-SVM 分类器精度, 输出识别正确率和分类时间。

进行空域特征提取, 首先需对输入的图像进行预处理, 然后进行分割和检测, 最后, 计算相应的空域特征。某一样本“1~9”手势的预处理后的图像及分割检测后的图像如图3所示。



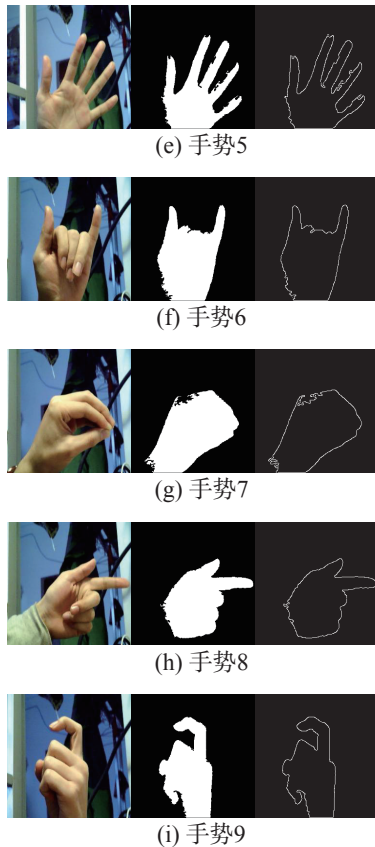


图3 “1~9”手势图像及分割检测图像

Fig. 3 "1~9" gesture image and split detection image

按照前文给出的公式,计算空域特征,包括位置、长宽比、矩形度、Hu矩的7个分量。取某一

表1 某一样本“1~9”手势的特征提取

Table 1 The feature extraction of a sample “1~9” gesture

特征	1	2	3	4	5	6	7	8	9
水平位置	91.17	111.32	145.21	144.29	126.43	114.07	136.97	129.64	119.11
竖直位置	180.70	199.94	209.10	216.43	216.43	185.29	157.81	169.26	170.17
长宽比	0.49	0.55	0.71	0.72	0.86	0.71	1.15	1.15	0.53
矩形度	0.55	0.55	0.43	0.50	0.47	0.63	0.56	0.48	0.62
Hu1	-0.64	-0.56	-0.56	-0.60	-0.61	-0.71	-0.72	-0.71	-0.64
Hu2	-1.75	-1.51	-1.62	-1.75	-2.29	-2.35	-2.03	-2.25	-1.72
Hu3	-2.23	-2.27	-2.90	-2.74	-3.15	-3.29	-3.25	-3.20	-2.80
Hu4	-2.57	-2.50	-3.02	-3.35	-3.03	-3.65	-4.29	-3.59	-3.46
Hu5	-4.47	-4.61	-5.85	-6.44	-5.81	-7.68	-7.36	-6.56	-6.24
Hu6	-3.45	-3.26	-3.85	-4.35	-4.17	-4.90	-5.39	-5.01	-4.46
Hu7	-5.48	-5.46	-6.20	-6.94	-8.61	-6.70	-7.50	-6.85	-7.53

选取高频分量输出的 16×16 矩阵,作为该手势图像的特征参数,方便计算,将该矩阵进行归一化处理,然后,取主对角线元素作为特征,共计16个分量。

对于每个样本,将空域特征的11个分量与频域的16个分量进行融合,组成新的矢量,则新矢量的维度为27,然后,以60个样本作为训练集,输入分类器中训练,30个样本预测集,用于验证

样本的“1~9”手势的特征,如表1所示。

对偶树复小波具有平移不变性和方向多样化的特性,按照对偶树复小波变换进行频域特征提取,dtwavexfm2函数中可以选择4个滤波器,本文选legall,进行变换分解。

以其中一幅手势图像为例,对偶树复小波特征提取的结果如下。某一示例图像“1”如图4所示,其他“2~9”手势特征提取过程相同。

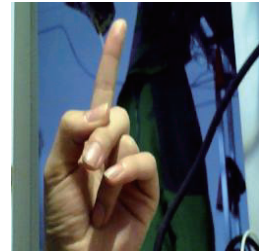


图4 示例图像

Fig. 4 Sample image

经过分解后的结果如图5所示。

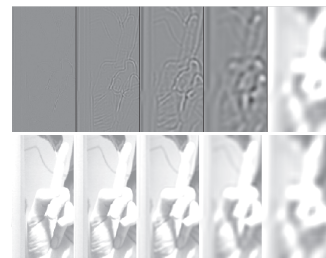


图5 分析结果

Fig. 5 Results of the analysis

分类精度。

3.2 BD-SVM 识别结果

本节主要测试手势“1~9”的识别效果。利用BD算法优化训练样本,设置BD-SVM的比例系数 t 为0.3,然后,将新的训练样本作为SVM训练集,最后,选取线性核函数、多项式核函数和径向基核函数分别进行实验。 γ 设置为0.1,惩罚参数设置为10,具体实验结果如下所示。

样本优选过程进行二维可视化后, 如图 6 所示。

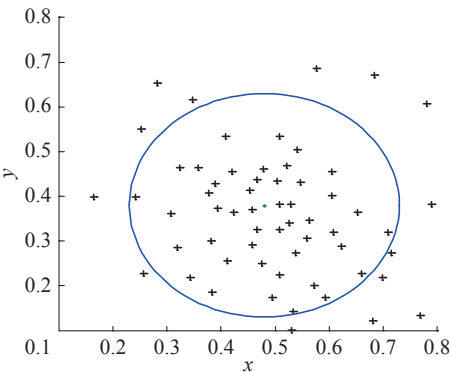


图 6 优选的训练样本

Fig. 6 Preferred training samples

圈内的样本为优选后的样本, 作为新的训练集, 挑选出 46 个新的训练样本。

对于每个核函数, 输出识别精度和识别时间, 最后计算平均正确率和平均识别时间。采用线性核函数的识别结果如表 2 所示。

表 2 线性核函数识别结果
Table 2 Linear kernel function recognition result

结果	正确率/%	时间/s	结果	正确率/%	时间/s
手势 1	82	0.021	手势 6	81	0.022
手势 2	84	0.023	手势 7	80	0.022
手势 3	86	0.021	手势 8	83	0.024
手势 4	87	0.021	手势 9	88	0.021
手势 5	85	0.023			

由表 2 得出, 线性核函数的平均正确率为 84%, 平均识别时间为 0.022 s。

采用多项式核函数的识别结果如表 3 所示。由表 3 得出, 线性核函数的平均正确率为 86.67%, 平均识别时间为 0.032 s。

表 3 多项式核函数识别结果
Table 3 Polynomial kernel function recognition result

结果	正确率/%	时间/s	结果	正确率/%	时间/s
手势 1	85	0.032	手势 6	88	0.033
手势 2	86	0.030	手势 7	89	0.032
手势 3	88	0.033	手势 8	84	0.033
手势 4	87	0.032	手势 9	87	0.033
手势 5	86	0.031			

采用径向基核函数的识别结果如表 4 所示。由表 4 得出, 线性核函数的平均正确率为 90.33%, 平均识别时间为 0.026 s。

通过实验得出, 选择径向基核函数的识别率

最高, 运算速度也较快, 因此, 进行静态手势识别, 采用径向基核函数。

表 4 径向基核函数识别结果
Table 4 Radial basis kernel function recognition result

结果	正确率/%	时间/s	结果	正确率/%	时间/s
手势 1	90	0.027	手势 6	89	0.027
手势 2	89	0.028	手势 7	88	0.025
手势 3	92	0.026	手势 8	90	0.028
手势 4	93	0.026	手势 9	91	0.025
手势 5	91	0.027			

4 结束语

本文主要采用空域特征和对偶树复小波相融合的特征, 空域特征包含水平位置、竖直位置、长宽比、矩形度、Hu 矩 7 个分量, 计 11 维特征, 频域特征采用对偶树复小波变换的 16 维特征, 特征融合后, 共计 27 维特征; 采用 BD 算法进行训练样本优选, 优选出 46 个新的训练样本, 再将新的训练样本进行 SVM 分类器的训练, 实验表明, 采用径向基核函数识别精度最高, 速度最快, “1~9”的静态手势识别效果最好。

参考文献:

- [1] BARROS P, PARISI G I, JIRAK D, et al. Real-time gesture recognition using a humanoid robot with a deep neural architecture[C]//2014 14th IEEE-RAS International Conference on Humanoid Robots. Madrid, Spain, 2014: 646–651.
- [2] DAVIS C C. Chronicler of the cavaliers: Some letters from and to William Alexander Caruthers, M. D. (1802-1846) [J]. The Virginia magazine of history and biography, 1947, 55(3): 213–232.
- [3] MIYASHITA T, SHINOZAWA K, HAGITA N. Gesture translation for heterogeneous robots[C]//2006 6th IEEE-RAS International Conference on Humanoid Robots. Genova, Italy, 2006: 462–467.
- [4] KAPUŚCIŃSKI T, OSZUST M, WYSOCKI M. Hand gesture recognition using time-of-flight camera and viewpoint feature histogram[M]//KORBICZ J, KOWAL M (eds). Intelligent Systems in Technical and Medical Diagnostics. Berlin, Germany: Springer, 2014: 403–414
- [5] 谭台哲, 韩亚伟, 邵阳. 基于 RGB-D 图像的手势识别方法[J]. 计算机工程与设计, 2018, 39(2): 511–515.
TAN Taizhe, HAN Yawei, SHAO Yang. Gesture recognition method based on RGB-D image[J]. Computer engineering and design, 2018, 39(2): 511–515.

- [6] 刘淑萍, 刘羽, 於俊, 等. 结合手指检测和 HOG 特征的分层静态手势识别[J]. 中国图象图形学报, 2015, 20(6): 781–788.
- LIU Shuping, LIU Yu, YU Jun, et al. Hierarchical static hand gesture recognition by combining finger detection and HOG features[J]. Journal of image and graphics, 2015, 20(6): 781–788.
- [7] 赵磊, 周亦敏. 一种基于改进 DTW-IMP 算法的手势识别[J]. 软件导刊, 2017, 16(11): 12–15, 21.
- ZHAO Lei, ZHOU Yimin. A gesture recognition based on improved DTW-IMP algorithm[J]. Software guide, 2017, 16(11): 12–15, 21.
- [8] 薛乐, 李立轻, 汪军. 应用频谱图的机织物纹理分析[J]. 纺织学报, 2015, 36(7): 55–60.
- XUE Le, LI Liqing, WANG Jun. Study on woven fabric texture using frequency spectrum[J]. Journal of textile research, 2015, 36(7): 55–60.
- [9] 汪丹, 张亚非. SVM 和 BP 算法在气体识别中的对比研究[J]. 传感技术学报, 2005, 18(1): 201–204.
- WAN Dan, ZHANG Yafei. Research of Gas Classification Based on SVM Compared with BP[J]. Journal of transclution technology, 2005, 18(1): 201–204.
- [10] 陈浩, 陈立辉, 毕笃彦, 等. BP 网络和支持向量机在非线形函数逼近中的应用[J]. 航空计算技术, 2004, 34(3): 27–30.
- CHEN Hao, CHEN Lihui, BI Duyan, et al. The appliance of BP-network and SVM in approach of non-linear function[J]. Aeronautical Computer Technique, 2004, 34(3): 27–30.
- [11] LIANG Haonan, ZHANG Hanqi. Identification of slope stability based on the contrast of BP neural network and SVM[C]//2010 3rd IEEE International Conference on

Computer Science and Information Technology. Chengdu, China, 2010: 347–350.

- [12] 周宇晴, 秦梦芝, 马志宏. 支持向量机和 BP 神经网络在可见-近红外光谱检测药品中的比较与应用[J]. 天津农学院学报, 2016, 23(2): 49–52.

ZHOU Yuqing, QIN Mengzhi, MA Zhihong. Comparison and application of support vector machine and BP neural network in visible-near infrared spectroscopy detection of drugs[J]. Journal of Tianjin agricultural university, 2016, 23(2): 49–52.

作者简介:



贾鹤鸣, 男, 1983 年, 副教授, 博士, 主要研究方向为非线性控制理论与信息检测技术。



朱传旭, 男, 1993 年, 硕士研究生, 主要研究方向为智能控制与信息处理技术。



张森, 男, 1994 年, 硕士研究生, 主要研究方向为智能控制与检测技术。