

面向人体动作识别的随机增量型混合学习机模型^{*}

常 征, 班晓娟, 马博渊, 邢一鸣



(北京科技大学 计算机与通信工程学院, 北京 100086)

通讯作者: 班晓娟, E-mail: banxj@ustb.edu.cn

摘 要: 针对自然人机交互应用中的人体动作识别问题,总结了传统机器学习模型在识别人体动作时的缺点,然后在此基础上针对自然人机交互应用的独特要求提出了面向人体动作识别的随机增量型混合学习机模型.该模型将误差反向传播模型、增量型极限学习机模型和双向增量型极限学习机模型相结合,克服了传统方法在识别人体动作时的不足.详细阐述了针对面向人体动作识别的随机增量型混合学习机模型的算法理论、模型合理性和实现方案.最后通过对比识别实验结果,验证了随机增量型混合学习机模型在识别人体动作问题上具有更好的鲁棒性、实时性和准确性.

关键词: 自然人机交互;机器学习;极限学习机;人体动作识别

中文引用格式: 常征,班晓娟,马博渊,邢一鸣.面向人体动作识别的随机增量型混合学习机模型.软件学报,2016,27(Suppl. (2)):137-147. <http://www.jos.org.cn/1000-9825/16029.htm>

英文引用格式: Chang Z, Ban XJ, Ma BY, Xing YM. Random increased hybrid learning machine oriented human body movement identification. Ruan Jian Xue Bao/Journal of Software, 2016,27(Suppl.(2)):137-147 (in Chinese). <http://www.jos.org.cn/1000-9825/16029.htm>

Random Increased Hybrid Learning Machine Oriented Human Body Movement Identification

CHANG Zheng, BAN Xiao-Juan, MA Bo-Yuan, XING Yi-Ming

(School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100086, China)

Abstract: Focusing on the problem of human movement identification in the application of natural human-computer interaction, this paper summarizes the shortcomings of the traditional machine learning model in the identification of body movement. Based on the unique requirements of natural human-computer interaction application, it proposes Random Increased Hybrid Learning Machine for human body movement identification. Combined with the Error Back Propagation Model, the Increased Extreme Learning Machine and Bidirectional Extreme Learning Machine, the model overcomes the shortcomings of traditional methods. This paper describes in detail the algorithm theory, model rationality and implementation scheme of the Random Increased Hybrid Learning Machine. Finally, by comparing the experimental results, the paper verifies the Random Increased Hybrid Learning Machine's a better robustness, accuracy and timeliness in identification of human body movement.

Key words: natural human-computer interaction; machine learning; extreme learning machine; human body movement identification

随着机器学习和传感器技术的发展,人机交互技术的研究也朝着越来越自然和易用的方向快速发展.触摸屏、声音控制、体感控制等技术也使得计算机系统变得越来越聪明,甚至能够学习并识别人体的自然协同交互动作,构建出一个更加“聪明”的虚拟现实,提升用户使用体验.而这种基于自然人机交互技术的现实应用场景非常广泛,例如伤残帮助、虚拟现实^[1]、区域监控^[2]、智能交通、体育竞技分析^[3]等.所以对于自然人机交互技术的研究极具理论研究意义和实际应用价值^[4].本文就针对自然人机交互过程中的人体动作识别问题提出了一

* 基金项目: 国家自然科学基金(61272357, 61300074, 61572075); 国家重点研发计划(2016YFB0700502, 2016YFB1001404)

Foundation item: National Natural Science Foundation of China (61272357, 61300074, 61572075); National Key Research and Development Plan (2016YFB0700502, 2016YFB1001404)

收稿时间: 2016-05-01; 采用时间: 2016-11-21

种机器学习模型,克服了传统方法的缺点和不足.

1 问题

因为人体动作呈现出随意性、灵活性和多义性,所以人体动作数据的特点也表现出单次数据量较小、多批次增加、数据维度较高和数据多变.基于传统机器学习方法的人体动作识别过程的识别效果往往不太理想,容易出现分类过杂、训练速度缓慢和观察位置敏感等问题^[5,6],无法满足自然人机交互的实时性和鲁棒性要求.

针对本文研究的人体动作的学习和识别过程,本文曾尝试应用传统的机器学习方法(例如 BP 网络模型、支持向量机模型、深度置信网络模型)去实现,发现一些问题,例如:模型训练时间过长、网络参数初始值需要靠经验选取、网络泛化能力差、容易产生过拟合现象等^[7,8].所以传统方法不能满足实际应用中的实时性要求.

于是本文曾尝试使用固定极限学习机模型去解决以上问题,但是通过实验发现一个问题:网络隐含层节点最佳数目不好确定,隐含层节点数目过少则识别效果不稳定,节点数目过多则浪费计算资源^[9].

针对隐含层节点最佳数目不易确定问题,本文在尝试使用增量型极限学习机模型学习和识别人体四肢协同交互动作.通过实验发现,这一模型学习和识别人体四肢协同交互动作过程中有两个问题.

(1) 在网络隐含层节点的随机生成过程中,网络模型中有些节点的输出权重 β 值很小.隐含层节点输出权重 β 过小,会使得该节点对于网络最终输出的影响很小.如果将这些对于网络最终输出影响不大的网络隐含层节点除去,就能够进一步加快极限学习机网络模型的学习效率.

(2) 增量型极限学习机模型每次固定递增一个隐含层节点,网络增长速度较慢,模型训练过程较长.如果能够在一次迭代过程中,网络递增多个隐含层节点,加快网络的增长,就能够进一步加快模型学习效率.

针对这些问题,本文提出一套新的面向人体动作识别问题的机器学习模型,克服传统模型在这一问题上的各种不足,为自然人机交互技术的发展做出贡献.

2 方法

为了解决以上两个问题,进一步加快学习效率,本文参考了 Huang 提出的最小化误差极限学习机模型 EM-ELM(error minimized extreme learning machine)^[10,11],以及 Yang 提出的双端增量型极限学习机模型 B-ELM(bidirectional extreme learning machine)^[12].

最小化误差极限学习机模型可以很好地解决增量型极限学习机模型每次递增一个隐层节点的问题,但是对于低效隐层节点无能为力.

双端增量型极限学习机模型可以很好地解决低效隐层节点问题,但是网络每次递增只增加一个隐层节点,网络增长效率太慢.

本文在这两个模型基础上,结合眼部与四肢的协同交互动作的特点,提出了随机增量型混合学习机模型 RIHLM(random increased hybrid learning machine).

2.1 随机增量型混合学习机模型的原理

随机增量型混合学习机模型总结了误差反馈模型(BP)、固定极限学习机模型(ELM)和增量型极限学习机模型(I-ELM)在眼部与四肢的协同交互学习和识别中的各种不足.随机增量型混合学习机模型具有以下优点.

- (1) 网络训练时间短,识别精度高.
- (2) 不需要靠经验选取网络初始参数.
- (3) 网络可以通过自生长方式确定最佳隐层节点数目.
- (4) 网络中低效隐层节点数目少.
- (5) 网络每次递增多个隐层节点,自增长速度快.

随机增量型混合学习机模型对于偶数个隐含层节点利用网络学习误差逆向传播的方式,通过每次网络增长的网络误差反馈矩阵来计算网络中的部分隐含层节点的连接参数;对于奇数个隐含层节点,则利用最小化误

差极限学习机模型的原理,通过误差最小化计算更新网络连接参数.随机增量型混合学习机模型的算法流程结构示意图如图 1 所示.

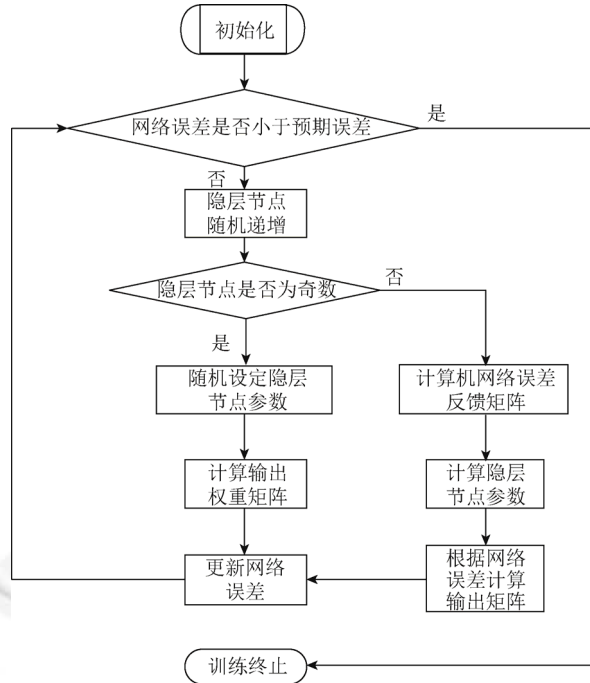


Fig.1 Random increased hybrid learning machine training process

图 1 随机增量型混合学习机训练流程图

2.2 随机增量型混合学习机模型的合理性证明

若一个单隐含层前馈神经网络,其网络激励函数 $g(R) \rightarrow R$ 满足在定义区间内分段可微的,大小为 N 的训练数据样本集合 $x_i = [x_{i1}, x_{i2}, \dots, x_{iN}]^T$,对于任意连续目标函数 $f()$,有目标类别集合为 $T = [f(x_{i1}), f(x_{i2}), \dots, f(x_{iN})]^T$,网络训练误差要求为 η ,网络初始输出误差 $e_0 = T$,隐含层节点数目为 L 时的网络输出误差为 e_L ,网络输出矩阵为 H_L ,增量网络迭代次数 k ,第 k 次迭代之后网络隐含层节点数目为 L_k .

当 $L=1$ 时,网络迭代次数为 $k=1$,根据最小化误差极限学习机模型原理,则网络的输出误差为

$$e_1 = \min \|H_1 \beta_1 - T\| \leq T = e_0 \tag{1}$$

假设当 $L=m$ 时,网络迭代次数为 k ,网络输出误差 e_m 仍然满足单调递减性质,即网络输出误差 $e_m \leq e_{L_{k-1}}$.

当第 $k+1$ 次迭代时,随机增长的隐含层节点数目为 r ,则网络隐含层节点数目为 $L_{k+1} = L_k + r = m + r$.

如果 L_{k+1} 为奇数,根据最小化误差极限学习机模型(EM-ELM)原理,因为 $\beta_{L_{k+1}}$ 是 $\|H_{L_{k+1}} \beta_{L_{k+1}} - T\|$ 的最小二乘解,则网络误差满足:

$$e_{L_{k+1}} = \min \|H_{L_{k+1}} \beta_{L_{k+1}} - T\| = \min \left\| \begin{bmatrix} H_{L_k} \\ H_r \end{bmatrix} \beta_{L_{k+1}} - T \right\| \leq \min \left\| \begin{bmatrix} H_{L_k} \\ H_r \end{bmatrix} \begin{bmatrix} \beta_{L_k}^T \\ 0 \end{bmatrix}^T - T \right\| = \min \|H_{L_k} \beta_{L_k} - T\| = e_{L_k} \tag{2}$$

即当随机增长后网络隐含层几点为奇数时,网络输出误差满足单调递减性质,即网络输出误差 $e_{L_{k+1}} \leq e_{L_k}$.

如果 L_{k+1} 为偶数,根据两端增量型极限学习机模型原理,则神经网络的误差反馈矩阵和网络输出权重矩阵为 H

$$H_{L_{k+1}}^e = e_{L_k} \cdot (\beta_{L_k})^{-1} \tag{3}$$

$$\beta_{L_{k+1}} = \frac{\langle e_{L_k}, H_{L_k} \rangle}{\|H_{L_k}\|^2} \quad (4)$$

根据误差反馈矩阵,可以得到网络误差为

$$\|e_{L_{k+1}}\| = \|e_{L_k} - H_{L_{k+1}}^e \cdot \beta_{L_{k+1}}\| \quad (5)$$

则网络误差的迭代差值为

$$\Delta = \|e_{L_k}\|^2 - \|e_{L_{k+1}}\|^2 = \|e_{L_k}\|^2 - \|e_{L_k} - H_{L_{k+1}}^e \cdot \beta_{L_{k+1}}\|^2 = 2\beta_{L_{k+1}} \langle e_{L_k}, H_{L_{k+1}}^e \rangle - \|H_{L_{k+1}}^e\|^2 \cdot (\beta_{L_{k+1}})^2 = \|H_{L_{k+1}}^e\|^2 \left(\frac{2\beta_{L_{k+1}} \langle e_{L_k}, H_{L_{k+1}}^e \rangle}{\|H_{L_{k+1}}^e\|^2} - (\beta_{L_{k+1}})^2 \right) \quad (6)$$

根据网络输出权重矩阵计算公式(4),上式可以推导出:

$$\Delta = \|H_{L_{k+1}}^e\|^2 \left(2\beta_{L_{k+1}} \beta_{L_{k+1}} - (\beta_{L_{k+1}})^2 \right) = \|H_{L_{k+1}}^e\|^2 (\beta_{L_{k+1}})^2 \geq 0 \quad (7)$$

即可以得到 $\|e_{L_k}\|^2 \geq \|e_{L_{k+1}}\|^2$, 网络输出误差同样满足单调递减性质.

综上随机增量型混合学习机模型的网络误差满足单调递减性质,也就是说随着多次迭代,网络的误差会逐渐变小.

进一步推导,根据公式(3)和公式(6),可以得到:

$$\|e_{L_k}\|^2 - \|e_{L_{k+1}}\|^2 = \|H_{L_{k+1}}^e\|^2 \cdot (\beta_{L_{k+1}})^2 = \left(\frac{\beta_{L_{k+1}}}{\beta_{L_k}} \right)^2 \|e_{L_k}\|^2 \quad (8)$$

对上式进行整理,可以得到:

$$\frac{\|e_{L_{k+1}}\|^2}{\|e_{L_k}\|^2} + \frac{(\beta_{L_{k+1}})^2}{(\beta_{L_k})^2} = 1 \quad (9)$$

随机增量型混合学习机模型的网络输出权重 $\beta_{L_{k+1}}$ 和 β_{L_k} 和网络的输出误差 $e_{L_{k+1}}$ 和 e_{L_k} 满足椭圆曲线公式,即输出权重-误差椭圆曲线公式.根据输出权重-误差椭圆曲线公式我们可以推论得到随机增量型混合学习机模型的学习速度只是跟 $\frac{\beta_{L_{k+1}}}{\beta_{L_k}}$ 有关,因为如果 $\lim \left(\frac{(\beta_{L_{k+1}})^2}{(\beta_{L_k})^2} \right) \rightarrow 1$, 则有 $\lim \left(\frac{\|e_{L_{k+1}}\|^2}{\|e_{L_k}\|^2} \right) \rightarrow 0$, 即每通过一次迭代的网络的输出误差会明显变小.

2.3 随机增量型混合学习机模型的实现

算法 1. 随机增量型混合学习机算法.

Input: 大小为 N 的训练样本集合 (x, t) , 隐含层神经元个数 $L=0$ 和激励函数 $h(x)$;

Output: 隐含层节点和输出层节点之间的权重矩阵 β .

1 start;

2 设定人工神经网络的网络初始化参数,包括:大小为 N 的网络训练数据样本集合 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$,

网络训练误差 η , 网络初始输出误差 $e_0=T$, 隐含层节点数目为 L 时的网络输出误差为 e_L , 网络输出矩阵为 H_L , 增量网络迭代次数 $k=0$, 第 k 次迭代之后网络隐含层节点数目为 L_k ;

3 增量网络迭代次数 $k=k+1$, 隐含层节点数随机递增 $L_k=L_{k-1}+r$;

4 while 网络误差 $e_{L_k} > \eta$

5 {

6 if L_k 为奇数

7 {

8 随机给定隐含层的输入权重 a_L, b_L ;

- 9 计算输出矩阵 $H_{L_k} = [H_{L_{k-1}}, H_r]$;
- 10 更新输出权重矩阵 $D_{L_k} = ((I - H_{L_k} H_{L_k}^{-1}) H_r)^{-1}$, $U_{L_k} = H_{L_k}^{-1} - H_{L_k}^{-1} H_r^T D_{L_k}$, $\beta_{L_k} = \begin{bmatrix} U_{L_k} \\ D_{L_k} \end{bmatrix} T$;
- 11 }
- 12 else if L_k 为偶数
- 13 {
- 14 计算误差反馈矩阵 $H_{L_k}^e = e_{L_{k-1}} \cdot (\beta_{L_k})^{-1}$;
- 15 计算隐含层神经节点参数 $a_{L_k} = h^{-1}(u(H_{L_k}^e)) \cdot x^{-1}$, $b_{L_k} = \sqrt{\text{mse}(h^{-1}(u(H_{L_k}^e)) - a_{L_k} \cdot x^{-1})}$, 其中 $\text{mse}(\cdot)$ 为方差运算, $u(\cdot)$ 为区间 $[0,1]$ 上的归一化函数;
- 16 更新输出权重矩阵 $H_{L_k} = u^{-1}(h(a_{L_k} \cdot x + b_{L_k}))$;
- 17 计算输出权重矩阵 $\beta_{L_k} = \frac{\langle e_{L_{k-1}}, H_{L_k} \rangle}{\|H_{L_k}\|^2}$;
- 18 }
- 19 计算网络输出误差 $e_{L_k} = e_{L_{k-1}} - \beta_{L_k} \cdot H_{L_k}$;
- 20 }
- 21 end;

本文提出的随机增量型混合学习机模型具有极限学习机所具有的万能逼近性质,而且因为结合了网络误差逆向传播和单次迭代递增多个节点的特点,所以网络中的低效隐含层网络节点数目减少,网络递增速度也得到提高,进一步缩短了网络的训练耗时。

3 实验

为了证实本文中提出的面向人体动作识别问题的随机增量型模型的性能,这里我们从两个不同的角度进行实验对比,以便证明该方法的良好性能。本文研究受到国家自然科学基金支持,所有的对比实验都是在微软的 Windows 7 环境下,通过 Matlab 2010b 环境下进行模拟仿真。运行的计算机系统配置为 Inter Core 4 Quad 2.9-GHZ CPU,内存大小为 6-GB RAM。

3.1 在通用公共数据集上的实验对比

为了验证本文提出的新模型在通用公共数据集上的良好性能,这里与其他的常用机器学习算法在常见的通用公共数据集进行了对比实验。在对比实验中,我们按照样本数量从 UCI 公共数据集中选择了 6 组分类问题数据集进行对比实验,见表 1。其中大样本量的数据集合有 Census Income(集合 1)和 Bank Marketing(集合 2),中等样本数量的数据集合有 Wine Quality(集合 3)和 Abalone(集合 4),小样本量的数据集合有 Student Performance(集合 5)和 Kidney Disease(集合 6),这些对比实验的结果可以参考后文。

3.1.1 不同学习算法学习效果对比

本次实验中,为了证明新方法在没有损失过多学习准确率的前提下,缩短了在线增量学习的过程,我们对比了误差反向传播算法模型、支持向量机算法模型(SVM)、深度置信网络模型 DBN(deep belief nets)、经典固定极限学习机模型、增量极限学习机模型和随机增量型混合学习机在 6 个机器学习公共集合上的分类效果。其中误差反向传播算法模型(BP)和支持向量机算法模型采用 Python Sklearn 1.4 机器学习库中的算法实现。深度置信网络作为深度学习的典型代表,这里本文采用的是 Matlab 的 DeepLearn Toolbox 实现。固定极限学习机模型和增量极限学习机模型采用的是新加坡南洋理工大学的教授 Huang 提供的算法模型实现。学习模型中的核函数均采用 Sigmoid 函数,算法中参数均保持默认。

Table 1 Public training data set in comparison test**表 1** 对比实验公共训练数据集

数据集	特征数目	总数目	训练数目	测试数目
集合 1	14	48 842	28 842	20 000
集合 2	17	45 211	25 211	20 000
集合 3	12	4 898	2 898	2 000
集合 4	8	4 177	2 177	2 000
集合 5	33	649	349	300
集合 6	25	400	200	200

对于经典误差反向传播算法模型和固定极限学习机算法模型,本文中对这些经典机器学习算法使用的是 1 000 个网络隐含层节点上限.对于增量极限学习机模型,我们使用了一次递增一个网络隐含层节点的网络递增方式,网络隐含层节点增长上限为 1 000.对比实验中,仿真程序设定的误差阈值为 0.001.一旦满足误差阈值要求或者达到隐含层节点数目上限,增量型极限学习机和双端增量极限学习机就停止网络增长.为了排除对比实验中的偶然奇异值对实验结果的影响,我们反复执行了 50 次仿真实验之后,将对比实验的结果求取统计平均值之后的结果.

Table 2 RMSE of network learning**表 2** 网络学习标准误差结果

公共数据集	模型学习标准误差(RMSE)					
	BP	SVM	DBN	ELM	I-ELM	RIHLM
集合 1	0.099 2	0.043 7	0.037 6	0.059 5	0.048 2	0.042 3
集合 2	0.099 4	0.049 0	0.037 5	0.047 5	0.044 8	0.041 0
集合 3	0.108 6	0.058 5	0.045 4	0.054 1	0.047 1	0.044 5
集合 4	0.110 3	0.061 3	0.053 3	0.059 5	0.058 2	0.054 3
集合 5	0.169 2	0.086 4	0.073 4	0.077 5	0.077 0	0.071 0
集合 6	0.159 4	0.078 4	0.069 4	0.073 1	0.071 1	0.062 5

从表 2 中列出的对比实验结果可以看出,在 6 组选定的测试公共数据集上,6 种算法模型的学习标准误差(RMSE)相近.误差反向传播模型在学习误差上面比较不理想,不过值得一提的是误差反向传播模型的稳定度相当好,50 次仿真实验得到的结果比较一致,没有出现过大的起伏.深度置信网络模型在大样本量数据集上学习误差最小,可见深度置信网络模型的多层网络结构能够很好地发掘并学习大样本量数据集中的隐藏知识.固定极限学习机和增量极限学习机学习误差在 6 组数据集上结果处于中等,但是他们的实验效果不稳定,在 50 次实验中结果出现多次较大幅度起伏.随机增量型混合学习机模型则在小样本量较高维度的测试数据集上的学习误差处于领先地位,比较符合本文对于人体眼部与四肢动作识别应用场景中的数据特点.

从上面表 3 中列出的对比实验结果可以看出,在 6 组选定的测试数据集上随机增量型混合学习机训练时间比传统 BP 算法快了大约 30 倍,比经典的支持向量机算法快了大约 22 倍,比深度置信网络快了大约 25 倍.

通过上面列举的学习误差和训练时间两项对比实验结果,我们可以得知,随机增量型混合学习机在没有损失过多的学习准确率的前提下,缩短了模型学习的过程.特别是在本文的实际应用场景中的中小样本量较高维度数据集上,随机增量型混合学习机模型做到了学习误差和训练时间之间很好的平衡.

Table 3 Network Learning training time**表 3** 网络训练时间结果

公共数据集	模型训练时间(s)					
	BP	SVM	DBN	ELM	I-ELM	RIHLM
集合 1	122.152	103.761	109.143	3.222	3.386	4.232
集合 2	121.669	104.251	108.712	3.260	3.336	4.193
集合 3	56.869	41.349	45.813	1.779	1.850	2.109
集合 4	55.453	40.938	44.859	1.752	1.846	2.098
集合 5	28.947	15.847	19.835	0.594	0.621	0.854
集合 6	27.548	14.013	18.441	0.589	0.618	0.855

3.1.2 不同隐含层节点数目对比

本次实验中,为了证明新方法可以有效降低隐含层节点数目而不影响网络分类效果,我们对经典误差反向传播算法模型、深度置信网络模型、经典固定极限学习机模型、增量极限学习机模型、和随机增量型混合学习机模型分别在大样本量数据集 Census Income(集合 1)、中样本量数据集 Wine Quality(集合 3)、小样本量数据集 Student Performance(集合 5)上的不同隐含层节点数目的分类效果。

对于经典误差反向传播算法模型、深度置信网络和经典固定极限学习机算法模型,本实验中进行了多次测试,分别选取了不同的隐含层节点数目进行试验。对于增量型极限学习机模型我们使用了一次递增一个网络隐含层节点的网络递增方式,网络隐含层节点增长上限为 1000。对比实验中,仿真程序设定的误差阈值为 0.001。一旦满足误差阈值要求或者达到隐含层节点数目上限,就停止网络增长。

根据图 2~图 4 中的不同隐含层节点数目的学习误差对比实验结果,我们可以得知随机增量型混合学习机模型可以有效降低隐含层节点数目而不影响网络分类效果。在上面大中小 3 类样本数据集中,我们都可以看出,随机增量型混合学习机模型相比较其他的经典机器学习算法模型,即使在少量隐含层节点时,也能够达到比较满意的学习误差范围要求。该模型在没有损失过多的学习准确率的前提下,缩短了在线增量学习的过程。随机增量型混合学习机模型因为应用极限学习机模型的思想,所以不需要经过反复调整网络参数,相比较误差反向传播模型和深度置信网络模型模型要更快,虽然学习的准确率没有深度置信网络模型高,但是快速的学习过程正好符合本文中关于眼部与四肢动作识别的应用场景要求。相比较同一算法体系中的固定极限学习机模型与增量型极限学习机模型,随机增量型混合学习机模型能够在较少隐含层节点数目的情况下就达到比较不错的分类准确率,固定极限学习机模型与增量型极限学习机模型则需要更多的隐含层节点才能达到学习误差的要求,验证了随机增量型混合学习机模型对于无效隐含层节点的排除效果。

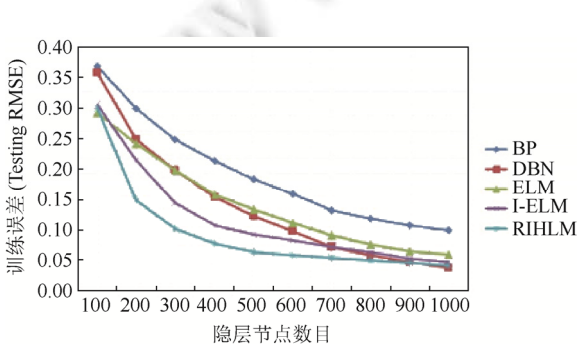


Fig.2 Census Income data set testing RMSE

图 2 数据集 Census Income 的训练误差

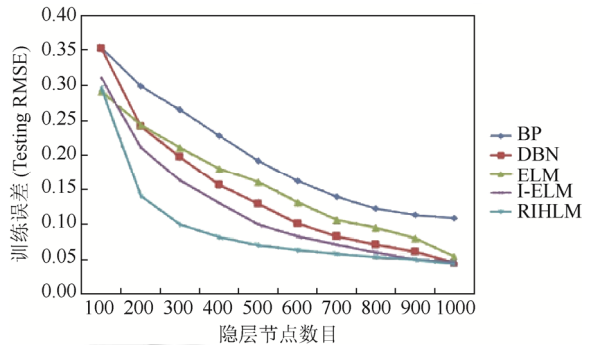


Fig.3 Wine Quality data set testing RMSE

图 3 数据集 Wine Quality 的训练误差

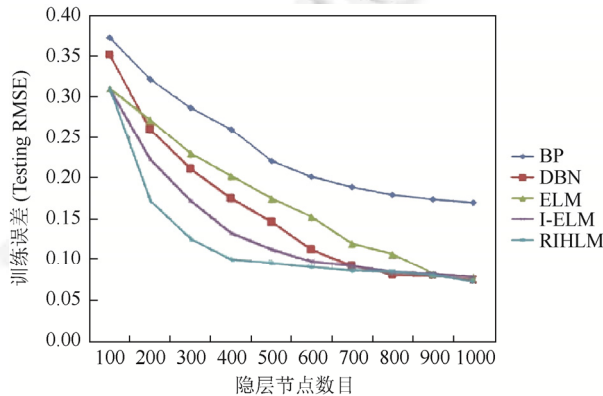


Fig.4 Student Performance data set testing RMSE

图 4 数据集 Student Performance 的训练误差

3.2 在人体动作数据集上的实验对比

为了验证本文提出的新模型在人体动作识别问题上的良好性能,这里与其他机器学习算法在特定的人体动作识别动作数据集进行了对比实验.

本实验使用微软的结构光传感器 Kinect 捕捉人体动作的三维运动轨迹数据.为了确保动作数据的可辨识性和完整性,本文在采集人体动作的三维运动轨迹数据时,需要目标人体正面朝前,身体直立于水平面,并且保证目标人体到结构光传感器之间有 1.2m~2m 的距离.

这里本文选取了 6 个具有代表性的交通指挥手势进行了对比实验,实验中一共列出了 6 种交通指挥手势.

- (1) 停止信号,目标车辆停止,不准通行.
- (2) 直行信号,右方直行的车辆通行.
- (3) 左转弯信号,目标车辆可以左转弯,或者是在安全的情况下允许车辆调头.
- (4) 变道信号,目标车辆减速,慢行并入其他车道.
- (5) 减速慢行信号,目标车辆减速,缓慢行驶.
- (6) 靠边停车信号,目标车辆减速并停车至道边.

对于以上展示的 6 种不同的代表性交通指挥动作,本文应用三维动作历史图像和三维 Hu 不变矩模型^[13],分别进行人体动作的三维建模和特征提取.每个动作对应的三维动作历史图像如下.

- (1) 停止信号(如图 5 所示).
- (2) 直行信号(如图 6 所示).

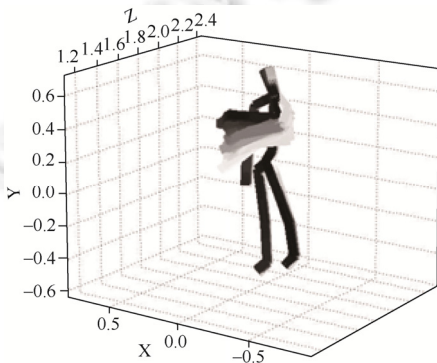


Fig.5 Stop traffic signal identification image
图 5 停止信号交通指挥手势系统识别图

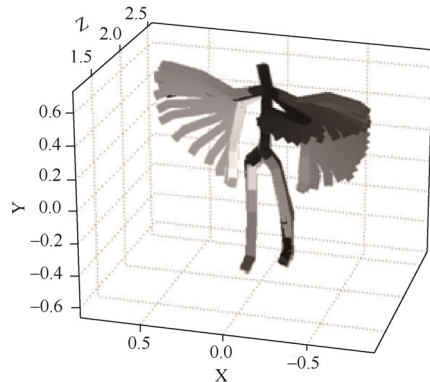


Fig.6 Straight traffic signal identification image
图 6 直行信号交通指挥手势系统识别图

- (3) 左转弯信号(如图 7 所示).
- (4) 变道信号(如图 8 所示).

这里将人体动作的三维 Hu 不变矩阵作为人体动作的三维特征矩阵输入到随机增量型混合学习机模型中,对应的人体动作的类别作为对应特征矩阵的目标输出类别.

实验中,本文采集了课题团队中 10 位同学的四肢动作运动轨迹数据.这 10 位同学的身高 155cm~180cm,其中包括 2 位女性,8 位男性,年龄 23 岁~29 岁,身体健康均无明显外伤或畸形.每一位参与实验的同学做每一类的四肢动作都需要重复 20 次,这样每类四肢动作就能产生 200 组样本,所以 6 种动作共计 1 200 组数据样本.然后我们从 1 200 组样本中随机选择 600 个样本作为测试样本数据集,剩下的其余四肢动作样本就作为训练样本数据集.

为了保持实验条件一致,我们在不同的人体四肢动作识别算法中使用相同的动作样本数据和实验环境,具有相同的最大的隐含层节点数目(即隐含层节点数目上限相同,1 000 个隐含层节点)和相同的人工神经网络激活函数(标准 Sigmoid 函数).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{10}$$

图 11 和图 12 展示了在不同人体四肢动作识别算法对于每一类动作的识别准确率和消耗的时间.

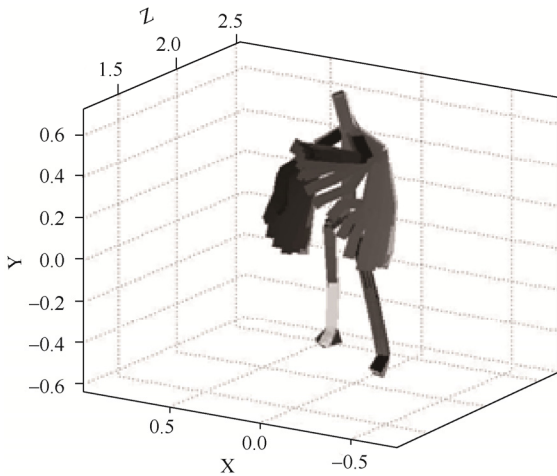


Fig.7 Turn left traffic signal identification image
图 7 左转弯信号交通指挥手势系统识别图

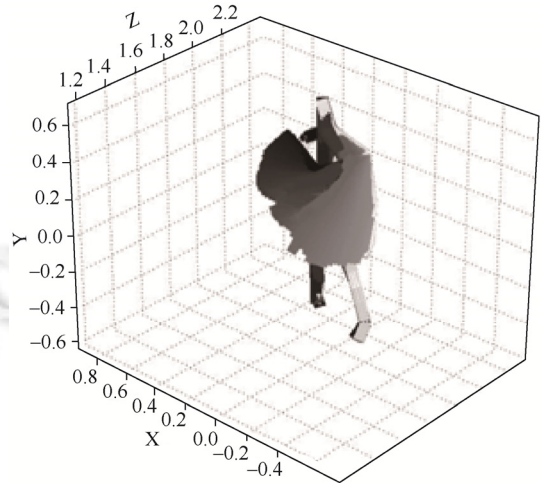


Fig.8 Change line traffic signal identification image
图 8 变道信号交通指挥手势系统识别图

- (5) 减速慢行信号(如图 9 所示).
- (6) 示意车辆靠边停车信号(如图 10 所示).

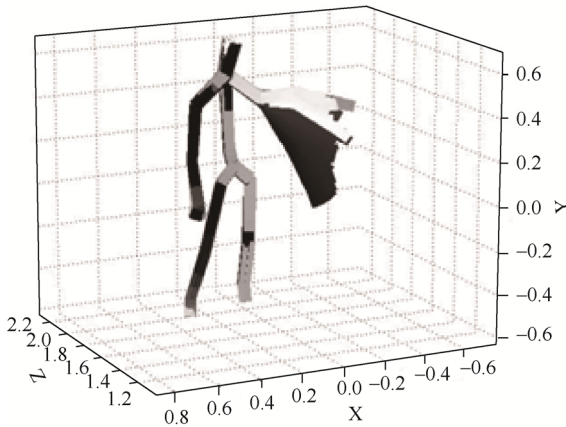


Fig.9 Slow traffic signal identification image
图 9 减速慢行信号交通指挥手势系统识别图

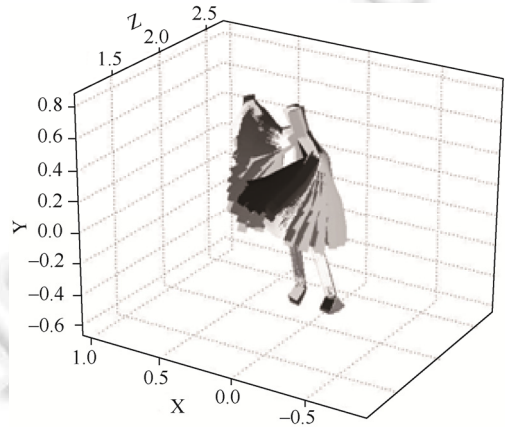


Fig.10 Side stop traffic signal identification image
图 10 靠边停车信号交通指挥手势系统识别图

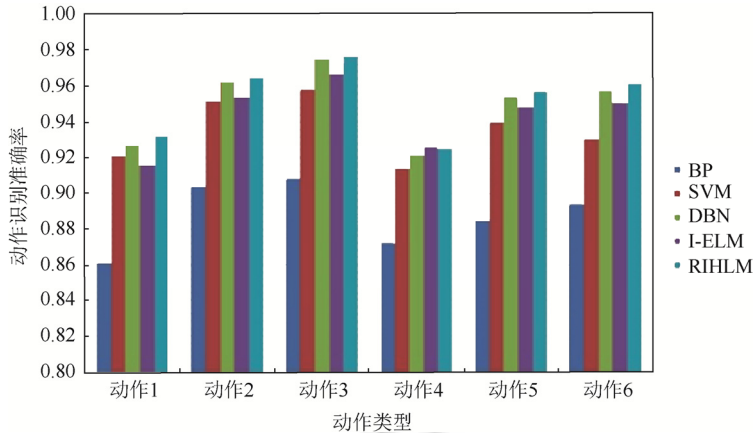


Fig.11 Identification rate of different learning algorithms

图 11 不同学习算法的识别准确率

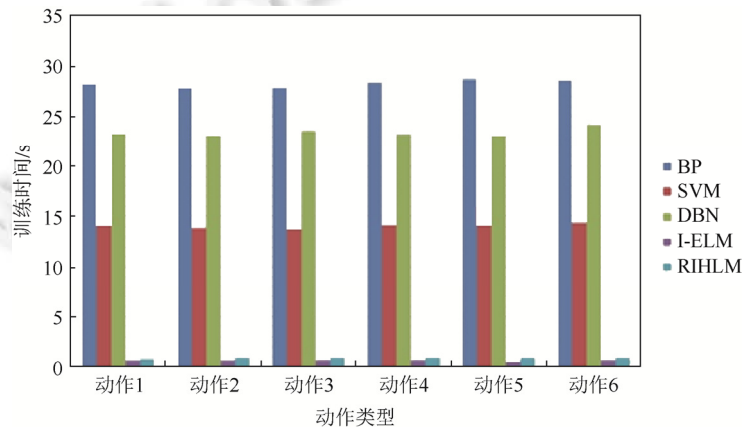


Fig.12 Training time of different learning algorithms

图 12 不同学习算法的训练时间

正如图 11 和图 12 所示,经典的人体四肢动作识别算法的准确率无明显差异,其中随机增量型混合学习机模型在对于动作 1、动作 2、动作 3、动作 5 和动作 6 的识别上都位居第一,而且随机增量型混合学习机模型在对于 6 个动作的训练时间的消耗上面明显少于误差反向传播模型、支持向量机和深度置信网络模型等机器学习算法.从以上两个方面来分析,我们可以得出结论本文提出的人体四肢动作识别模型对于人体四肢动作的学习和识别上面具有更好的学习效果.

4 结 论

为了满足自然人机交互应用场景下的人体四肢协同交互动作的学习、识别和理解过程对鲁棒性、实时性和准确性的要求,本文提出了基于随机增量型混合学习机模型的人体四肢协同交互动作的学习和识别模型.本文结合理论推导和实验分析,证明了新模型很好地克服了传统方法对于四肢协同交互动作建模和识别上的问题,并使得协同交互动作的学习和识别过程高效快捷,保证了系统整体实时性,同时降低了由机器学习产生的过程开销.

References:

- [1] Nijholt A. Meetings, gatherings, and events in smart environments. In: Proc. of the VRCAI. 2004. 229–232.
- [2] Du YT, Chen F, Xu WL, Li YB. A survey on the vision-based human motion recognition. Acta Electronica Sinica, 2007,35(1): 84–90 (in Chinese with English abstract).
- [3] Ruan TT, Yao MH, Qu XY, Lou ZW. A survey of vision-based human motion analysis. Computer Systems and Applications, 2010,20(2):245–253 (in Chinese with English abstract).
- [4] Bashir FI, Khokhar AA, Schonfeld D. Object trajectory based activity classification and recognition using hidden Markov models. IEEE Trans. on Image Processing, 2007,16(7):1912–1919.
- [5] Leung MK, Yang YH. First sight: A human body outline labeling system. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1995,17(4):359–377.
- [6] Gavril DM. Vision-Based 3-D Tracking of Human in Action. Maryland: University of Maryland, 1996.
- [7] Bashir F, Khokhar A, Schonfeld D. A hybrid system for affine-invariant trajectory retrieval. In: Proc. of the ACM SIGMM Multimedia Information Retrieval Workshop. 2004.
- [8] Ohya J, Kishino F. Human posture estimation from multiple images using genetic algorithm. In: Proc. of the ICPR. 1994.
- [9] Chang Z, Ban XJ, Shen Q, Guo J. Research on three-dimensional motion history image model and extreme learning machine for human body movement trajectory recognition. Mathematical Problems in Engineering, 2015. 1–15.
- [10] Feng G, Huang GB, Lin Q, Gay R. Error minimized extreme learning machine with growth of hidden nodes and incremental learning. IEEE Trans. on Neural Network, 2009,20(8):1352–1357.
- [11] Lan L, Soh YC, Huang GB. Random search enhancement of error minimized extreme learning machine. In: Proc. of the European Symp. on Artificial Neural Networks. 2010. 327–332.
- [12] Yimin Yang. Research on Extreme Learning Theory for System Identification and application. Changsha: Hunan University, 2013
- [13] Chang Z, Ban XJ. Recognition of human body movements trajectory based on the three-dimensional depth data. In: Proc. of the 19th IFAC world Congress, Vol.19. 2014. 12331–12336.

附中中文参考文献:

- [2] 杜友田,陈峰,徐文立,李永彬.基于视觉的人的运动识别综述.电子学报,2007,35(1):8490.
- [3] 阮涛涛,姚明海,翟心昱,楼中望.基于视觉的人体运动分析综述.计算机系统应用,2010,20(2):245–253.
- [12] 杨易旻.基于极限学习的系统辨识方法及其应用研究[博士学位论文].长沙:湖南大学,2013.



常征(1986—),男,河南商丘人,博士,工程师,主要研究领域为人工智能与人工生命理论研究,可视化技术,计算机动画与仿真技术研究,智能软件理论及应用。



班晓娟(1970—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为人工智能与人工生命理论研究,可视化技术,计算机动画与仿真技术研究,智能软件理论及应用。



马博渊(1992—),男,硕士,主要研究领域为人工智能与人工生命理论研究,可视化技术,计算机动画与仿真技术研究,智能软件理论及应用。



邢一鸣(1986—),女,助教,主要研究领域为人工智能与人工生命理论研究,可视化技术,计算机动画与仿真技术研究,智能软件理论及应用。