

基于核函数的在线序列 ELM 算法的姿态识别

韩莹¹, 张浩¹, 刘健², 陈立平¹

(1 中国科学院 微电子研究所, 北京 100029; 2 北京科技大学 计算机与通信工程学院, 北京 100029)

摘 要: 姿态识别是许多应用的基础(医学、运动、游戏、安全), 传统的识别算法采用批学习的方式去训练网络, 但是数据量庞大且数据不会一次性获取, 这会导致这类算法花费大量的学习时间且网络权重也不能在线更新. 对此利用一种基于核函数的在线序列极限学习机 OS-KELM(Online Sequential Kernel Extreme Learning Machine)算法实现人体姿态的分类识别. 为降低学习难度和提高学习效率, 使用了基于 Fisher 准则和特征聚类的方法进行特征选择. 用手机的三轴加速度计和陀螺仪数据识别人走路、下楼、上楼、站立、坐和躺下的姿态, 平均识别精度达到 91.89%.

关键词: 在线序列 ELM; 核函数; 人类姿态识别; 模式识别; Fisher 准则; 特征聚类

中图分类号: TP31

文献标识码: A

文章编号: 1000-7180(2018)01-0091-05

Activity Recognition Based on Online Sequential Kernel Extreme Learning Machine

HAN Ying¹, ZHANG Hao¹, LIU Jian², CHEN Li-ping¹

(1 Institute of Microelectronics of the Chinese Academy of Sciences, Beijing 100029, China;

2 School of Computer & Communication Engineering, University of Science & Technology, Beijing 100029, China)

Abstract: Activity recognition (AR) is the basis for many applications concerning health care, sports, security and gaming industry. Traditionally, batching learning recognition algorithms is adopted to train network. However, the amount of data is considerable and not all training data arrives instantly, the learning procedure is time-consuming and the network weights cannot be updated online. In this paper, Classification of the human activities is performed with Online Sequential Kernel Extreme Learning Machine (OS-KELM). The method of feature selection based on Fisher criterion and feature clustering has been adopted to reduce difficulty and improve efficiency of learning. A tri-axial accelerometer and gyros data from a user's smart phone are used to recognize walking, waling downstairs, walking upstairs, standing, sitting and laying. Experimental results with an average accuracy of 91.89% are achieved.

Key words: online sequential ELM; kernel function; human activity recognition; pattern recognition; Fisher criterion; feature clustering

1 引言

过去十年,随着移动设备和新型传感器的体积缩小、价格降低和功耗降低,其在应用领域体现出巨大的优势.而使用手机传感器识别人的日常动作也成了医学、运动、游戏、安全等领域的热门研究方向^[1].在医学方面,通过监控观察痴呆或者其他脑疾

病患者的日常活动,可以达到预防意外的目的.另外,人的姿态识别也可以运用于人与虚拟环境的交互,识别特定的动作时从而触发不同的事件,例如识别到上楼下楼的动作灯就自动打开,而识别到躺下的时候灯就自动关闭等.姿态识别也可用在最近热起来的体感游戏、VR 等应用中.

许多学习算法已经被用到人的姿态识别当中,

文献[2]使用支持向量机(SVM)识别出走路、下楼,识别精度达85%,文献[3]使用BP神经网络识别出站立、坐和下楼,平均识别精度达到98.5%。然而他们采用的都是批学习和训练的网络,时间消耗大,网络的权重也不能在线更新。文献[4]指出实际应用中,不是所有数据都会同时得到,而是随时间一块一块产生的,所以序列学习机制就显得尤为重要,还提出了使用在线序列极限学习机 OS-ELM(Online Sequential Extreme Learning Machine)算法对走路、下楼、上楼、站立、坐和躺下进行识别,平均识别精度达到82%。文献[5]在 OS-ELM 中加入核函数,对算法进行了优化,提高了分类精度。本文提出了使用 OS-KELM 分类算法实现人走路、下楼、上楼、站立、坐和躺下的姿态分类识别,通过实验数据验证,OS-KELM 的分类效果比 OS-ELM 更好。

特征选择是提高算法学习效率、识别精度的有效方法之一。特征选择是从原始特征集合中去除不相关特征和冗余特征,得到一个低维度的次优特征子集。首先使用单个特征的 Fisher 比作为去除不相关特征的准则,筛去噪声和鉴别力较差的特征;然后采用一种基于特征之间的相关性系数的分层聚类方法去除冗余特征;最后从每个聚类结果中选择一个特征组成特征子集。

2 系统模型

图1显示了整个姿态识别过程。首先采集手机里传感器数据,包括三轴加速度计的加速度数据和三轴陀螺仪的角速度数据,经过滤波、分窗等预处理,然后分别在时域和频域上面提取特征。经过特征选择后,把特征子集输入到网络进行训练,得到网络输出权重矢量。最后用一组测试数据测试算法的识别精度。

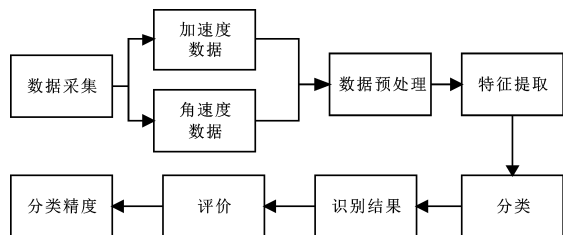


图1 姿态识别流程框图

3 算法原理

3.1 基于 Fisher 准则和特征聚类的特征选择

采集到的加速度和角速度数据经过预处理后,

对每段数据进行特征提取,文献[6]中给出了包括平均值、协方差、信号幅度区(SMA)、能量、熵、矢量间角度等561种特征。然而原始特征集中含有许多冗余特征甚至是噪声特征,这些特征不仅会增大学习算法训练时间和计算复杂度,还可能降低分类的准确度。合理的特征选择可以提高分类算法的效率,减少计算复杂度^[7]。

Fisher 准则主要思路是鉴别能力较强的特征表现为样本类内距离尽可能小,类间距离尽可能大。给定训练数据集 $N = \{x_i, t_i \mid x_i \in R^m, t_i \in R^l, i = 1, 2, \dots\}$, 将其分成6类,每一类包含 n_i 样本,单特征的 Fisher 准则表示为

$$J_{\text{fisher}}(k) = S_B^{(k)} / S_W^{(k)}$$

式中, $S_B^{(k)}$ 和 $S_W^{(k)}$ 分别表示该维特征在训练样本集上的类间方差和类内方差。某特征的 Fisher 比越大,说明该特征包含的鉴别信息越多。

设初始特征集 $S = \{F_i, i = 1, 2, \dots, D\}$, $S_1 = S_2 = \text{NULL}$

(1) 在训练数据集上计算每一维特征的 $J_{\text{fisher}}(F_i)$;

(2) 计算所有特征 Fisher 比的总和 $\text{FisherSum} = \sum_{F_i \in S} J_{\text{fisher}}(F_i)$;

(3) 对 S 中特征按照 Fisher 比从大到小排序,按照顺序将 S 中的特征依次添加到 S_1 中,直到 $\sum_{F_i \in S_1} J_{\text{fisher}}(F_i) > \rho * \text{FisherSum}$;

原始特征集经过 Fisher 准则初步筛选,去除 Fisher 比特小,即鉴别能力特别差的特征。

预选所得的特征集再通过分层聚类去除冗余特征。聚类算法是非监督模式识别的一种方法,使用分层聚类中的合并(自底而上)方法,达到减少类别数目的目的。

本文用相关系数来度量特征之间的冗余度:

$$\rho_{xy} = \frac{\sum_i (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_i (x - \bar{x})^2} \sqrt{\sum_i (y - \bar{y})^2}}$$

ρ 的绝对值表示 x, y 之间的相关度,其值越大 x, y 相关度越大,也就表示冗余度越大。当两个聚类 φ_i, φ_j 分别包含 m, n 个特征时,这两个聚类的相似度量定义为 $\text{SIM}_{ij} = \frac{1}{mm} \sum_{x \in \varphi_i} \sum_{x \in \varphi_j} |\rho_{xy}|$ 。每一次聚类都将 SIM_{ij} 值最大的两个聚为一类,直到类别之间的相似度的最大值达到设定的阈值,聚类过程结束。

(1) 对 S_1 中的特征进行分层聚类,结束条件为相似度 SIM_{ij} 的最大值达到阈值 C_i ;

(2) 从每一个聚类结果里面选择 Fisher 比最大的那一维特征添加到 S_1, S_2 即为选好的特征子集。

3.2 OS-ELM 模型

在线序列极限学习机(OS-ELM)是在极限学习机(ELM)^[8]上的改进,OS-ELM 可以将数据一个一个或者一块一块地添加到网络中,在线学习的模式更适用于处理大量数据和实时数据处理的应用中。

给定训练数据集

$N = \{x_i, t_i \mid x_i \in R^m, t_i \in R^l, i = 1, 2, \dots\}$, 激活函数 $g: R \rightarrow R$, 隐层节点个数 L , 则 ELM 分类器的输出函数可写为

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = \mathbf{h}(x) \boldsymbol{\beta}$$

式中, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]$ 是输出权重矢量, 更直观的表达如下:

$$\mathbf{H} \boldsymbol{\beta} = \mathbf{T}$$

$$\text{式中, } \mathbf{H} = \begin{bmatrix} g(a_1 x_1 + b_1) & \cdots & g(a_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(a_1 x_N + b_1) & \cdots & g(a_L x_N + b_L) \end{bmatrix}, \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}.$$

根据 Bartlett 理论, 可以求解出:

$$\boldsymbol{\beta} = \mathbf{H}^* \mathbf{T}$$

$\mathbf{H}^*$ 是 $\mathbf{H}$ 的 Moore-Penrose 逆, 然后通过递推方式更新 $\boldsymbol{\beta}$, 从而达到在线序列学习的目的, 算法详细过程见文献[8]。

3.3 OS-KELM 模型

从优化效率和准确度的角度, 文献[9]将核函数应用于 ELM 算法, 提出了基于核函数的极限学习机 KELM(Kernel Extreme Learning Machine)算法, 使得原始 ELM 算法不用确定隐层输出矩阵, 而只需通过计算核函数来计算最终分类输出, 简化了分类算法输出形式, 也提高了分类精度。文献[5]在此基础上提出了 OS-KELM 算法, 使得 KELM 具有在线序列学习的特点。文献[5]中指出 KELM 分类器输出可总结为

$$f_L(\mathbf{x}) = \sum_{i=1}^L \beta_k k(\mathbf{x}, \mathbf{x}_i) = \mathbf{K}_L \boldsymbol{\beta}$$

式中, $k(\mathbf{x}, \mathbf{x}_i)$ 是隐层神经元输出, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]$ 是输出权重矢量, 映射采样 $\mathbf{X}_L = [\mathbf{x}_1, \dots, \mathbf{x}_L]$ 从训练集 $\mathbf{X}$ 随机选取。

对比可看出, KELM 输出中, 核矩阵 $\mathbf{K}_L$ 代替了原有的隐层输出矩阵和 $h(x)$, 消除了原网络中的隐层输入矩阵参数随机性, 这也是 KELM 分类精度比 ELM 高的原因之一。

为了使 KELM 具有在线学习的特性, 文献[5]推导了 OS-KELM 算法, 具体过程如下:

初始化: 用给定训练数据集 $N = \{x_i, t_i \mid x_i \in R^m, t_i \in R^l, i = 1, 2, \dots\}$ 中的一小块数据 $N_0 = \{x_i, t_i\}_{i=1}^{N_0}$ 初始化模型。

1. 从初始训练数据集 $N_0 = \{x_i, t_i \mid x_i \in R^m, t_i \in R^l\}_{i=1}^{N_0}$ 随机选取子集 X_L ;

2. 计算初始核矩阵 $\mathbf{K}_0 = \kappa(\mathbf{X}_0, \mathbf{X}_L)$;

3. 估计初始输出权重 $\boldsymbol{\beta}^{(0)} = \mathbf{G}_0 \mathbf{K}_0^T \mathbf{T}_0$

式中, $\mathbf{G}_0 = (\mathbf{I}/C + \mathbf{K}_0^T \mathbf{K}_0)^{-1}$, $\mathbf{T}_0 = [t_1, \dots, t_{N_0}]$;

设 $k = 0$;

序列学习:

4. 第 $(k+1)$ 块数据加入时:

$$N_{k+1} = \{(x_i, t_i) \mid i = \sum_{j=0}^k N_j + 1, \dots, \sum_{j=0}^{k+1} N_j\}$$

式中, N_{k+1} 代表第 $(k+1)$ 次加入的数据;

5. 计算第 $(k+1)$ 块数据的核函数:

$$\mathbf{K}_{k+1} = \kappa(\mathbf{X}_{k+1}, \mathbf{X}_L);$$

6. 设 $\mathbf{T}_{k+1} = [t_{\sum_{j=0}^k N_j + 1}, \dots, t_{\sum_{j=0}^{k+1} N_j}]^T$;

7. 计算输出权重 $\boldsymbol{\beta}_{k+1}$

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} + \mathbf{G}_{k+1} \mathbf{K}_{k+1}^T (\mathbf{T}_{k+1} - \mathbf{K}_{k+1} \boldsymbol{\beta}^{(k)})$$

$$\mathbf{G}_{k+1} = \mathbf{G}_k - \mathbf{G}_k \mathbf{K}_{k+1}^T (\mathbf{I} + \mathbf{K}_{k+1} \mathbf{G}_k \mathbf{K}_{k+1}^T)^{-1} \mathbf{K}_{k+1} \mathbf{G}_k$$

8. 设 $k = k + 1$, 跳到步骤 5。

4 实验及结果分析

4.1 实验数据

本实验中, 数据选自 UCI 数据集^[6]。该数据集来自 30 位年龄在 19~48 岁之间的志愿者, 每个人腰上佩戴手机按照规定实现六种预定动作(例如:平地走路、上楼、下楼、坐、站立和躺下)。一共采集 10 299 组数据, 得到的数据集被随机分成两组, 70% 的数据被用来训练, 30% 的数据被用来测试。试验中使用带有三轴加速度计和陀螺仪的 Samsung Galaxy S2 智能手机, 以 50 Hz 的采样率采集加速度和角速度信号。原始信号经过一个低通滤波器滤除噪声, 然后被分为 50% 覆盖率 2.56 s 固定长度的窗, 即每行数据包括 128 个值(50 Hz × 2.56 s)。对每个窗口长度的数据, 总共提取了 561 类特征, 包括均值、标准误差、最小值、最大值、协方差、能量、熵等。

已有数据集给出的 561 个特征, 远多于原始数据个数(128 个), 所以对特征进行一个选择。经过基于 Fisher 准则和特征聚类的特征选择过程后, 得到优化的特征子集。Fisher 准则参数 ρ 设为 0.9, 初步选择后得到 210 个特征; 然后用特征之间相似度对特征进行聚类时 SIM_{ij} 的阈值设为 0.85, 最后得到

的特征子集为 64 个,其中包括加速度的均值、标准差、最大值、最小值、能量、加速度与重力的夹角等,角速度的标准差、最大值、最小值、能量、四分差等.

同时本文在相同特征/神经元条件下比较 OS-ELM 和 OS-KELM 的精度和训练时间,从而评价这两个分类算法. OS-KELM 核函数选择高斯核函数 $\kappa(x,y)=\exp(-\gamma||x-y||^2)$, OS-ELM 的激活函数选择 Sigmoidal 函数 $G(a,b,x)=\frac{1}{1+\exp(-(ax+b))}$.

4.2 结果分析

实验中以逐渐递增特征/神经元的形式对模型进行训练,确定最佳隐层节点个数. 图 2 显示识别精度随着隐层神经元数量的增加而提高,在神经元个数大于 110 以后,识别精度不再随神经元个数增加而提高,最后在 0.85~0.95 之间波动,在 200 个神经元/64 个特征的时候达到最高精度 91.89%,高于文献[4]的最高精度 82%.

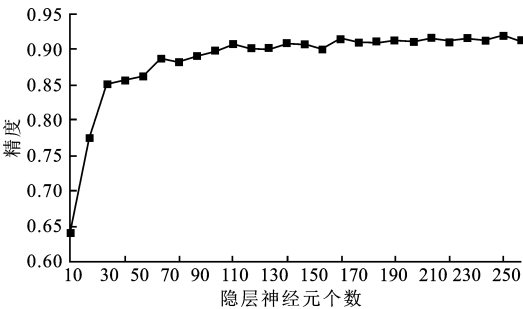


图 2 随神经元数量增加 OS-KELM 精度

在 200 个神经元/64 个特征的条件对 OS-ELM 和 OS-KELM 进行比较.

基于 OS-ELM 识别精度混淆矩阵如表 1 所示,第一列总共有 496 组走路数据被测试,总共有 438 组走路数据被正确识别,识别精度为 88.30%,其余的姿态以此类推,识别精度在 84.02%~97.02%之间,平均识别精度为 88.00%.

表 1 基于 OS-ELM 神经网络识别精度混淆矩阵

姿态	走路	上楼	下楼	坐	站立	躺下
走路	438	62	16	2	2	0
上楼	18	401	19	3	1	1
下楼	40	8	385	0	0	0
坐	0	0	0	402	82	0
站立	0	0	0	83	447	15
躺下	0	0	0	1	0	521
精度%	88.30	85.14	91.67	81.87	84.02	97.02

基于 OS-KELM 识别精度混淆矩阵如表 2 所示,第一列总共有 496 组走路数据被测试,总共有 452 组走路数据被识别正确,识别精度为 91.13%,其余的姿态以此类推,识别精度在 86.56%~97.02%之间,平均识别精度为 91.89%.

相比于 OS-ELM 算法,OS-KELM 算法在每个姿态的识别精度上都有所提高,总的平均识别精度也提高了 3.89%.

表 2 基于 OS-KELM 神经网络识别精度混淆矩阵

姿态	走路	上楼	下楼	坐	站立	躺下
走路	452	24	12	1	3	0
上楼	22	433	27	2	0	1
下楼	22	14	381	0	0	0
坐	0	0	0	425	31	0
站立	0	0	0	61	498	13
躺下	0	0	0	2	0	523
精度%	91.13	91.93	90.71	86.56	93.61	97.39

在姿态识别中,识别的精度和学习的时间是评价两个分类算法的重要指标. 比较完基于 OS-ELM 和 OS-KELM 的分类算法的分类精度后,再来比较他们的学习时间和测试时间.

如表 3 所示,在训练和测试时间上 OS-KELM 都比 OS-ELM 的时间要长. OS-ELM 的训练时间都在 1 s 以内,而 OS-KELM 的训练时间在 2 s 左右,这是引入核函数后造成的计算开销. 系统运行环境为 MATLAB 2015 b/Intel i5 2.3 GHz 处理器.

表 3 OS-ELM 和 OS-KELM 神经网络训练和测试时间比较 s

算法	训练时间	测试时间
OS-ELM	0.765 6	0.031 3
	0.828 1	0.031 3
	0.906 3	0.031 3
	0.796 9	0.031 3
	0.843 8	0.031 3
	0.859 4	0.015 6
OS-KELM	0.687 5	0.031 3
	1.937 5	0.203 1
	2.031 3	0.250 0
	2.062 5	0.218 8
	2.031 3	0.265 6
	1.921 9	0.3125
	2.000 0	0.281 3
	2.062 5	0.218 8

5 结束语

本文使用使用基于 Fisher 准则和特征聚类的方法进行特征选择,使用 OS-KELM 实现对人的六种姿态分类识别. 特征选择大大降低了特征集的维度,降低了学习难度. OS-KELM 在 OS-ELM 的基础上引入核函数,补偿 OS-ELM 算法因隐层输入随机矩阵带来的误差. 随后对这两个算法从识别精度和训练、测试时间上进行了比较. 实验数据来自智能手机中的加速计和陀螺仪传感器. 实验结果显示 OS-KELM 的识别精度在 64 个特征/200 个神经元的条件下可达到 91.89%,相较已有成果的 82%和选取 Sigmoidal 函数为激活函数的 OS-ELM 算法的 88.00%,在精度上都有所提高. 同时,由于核函数的引入在训练时间和测试时间上也略有增加.

本文使用 OS-KELM 算法对人的姿态分类识别进行了探索,考虑进一步的研究方向如下:(1)连贯姿态的识别,很多动作是连贯在一起的,例如走路后上楼,然后下楼再站立等,这些类似情况会导致两个或多个姿态的交叉融合;(2)可以更进一步探索不同的核函数对识别精度和训练时间的影响.

参考文献:

- [1] Perez A J, Labrador M A, Barbeau S J. G—sense: a scalable architecture for global sensing and monitoring [J]. IEEE Network, 2010, 24 (4): 57-64.
- [2] Shi X, Xiong Q Y, Lei L N. Fall detection system based on pressure sensor[J]. Chinese Journal of Scientific Instrument, 2010, 31(3): 715-720.
- [3] Liu Z G, Song Y N, Shang Y. Posture recognition algorithm for theelderly based on BP neural networks[C]

// Proceedings of 27th Control and Decision Conference (CCDC). IEEE, 2015: 1447-1449.

- [4] Al Jeroudi Y, Ali M A, Latief M, et al. Online sequential extreme learning machine algorithm based human activity recognition using inertial data[C]// Control Conference (ASCC), 2015 10th Asian. IEEE, 2015: 1-6.
- [5] Deng W Y, Ong Y S, Tan P S, et al. Online sequential reduced kernel extreme learning machine[J]. Neurocomputing, 2016, 174: 72-84.
- [6] Anguita D, Ghio A, Oneto L, et al. A public domain dataset for human activity recognition using smartphones[C]//ESANN. Bruges, 2013.
- [7] 王颢, 郑链. 基于 Fisher 准则和特征聚类的特征选择 [J]. 计算机应用, 2007, 27(11): 2812-2813.
- [8] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 70(1): 489-501.
- [9] Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012, 42(2): 513-529.

作者简介:

韩 莹 女, (1991-), 硕士研究生. 研究方向为数字信号处理、无线通信.

张 浩(通讯作者) 男, (1975-), 副研究员. 研究方向为复杂环境下的线性和非线性信号处理技术.

E-mail: haozhang@ime. ac. cn.

刘 建 男, (1977-), 副教授. 研究方向为认知无线网络、移动网网络、物联网技术.

陈立平 男, (1983-), 助理研究员. 研究方向为信号处理、无线通信.

(上接第 90 页)

- [8] 孟祥志, 杨萌, 来金梅. 基于累积频数统计的 FPGA 互连资源时序库[J]. 计算机工程, 2013, 39(2): 254-259.

作者简介:

段 芬 女, (1993-), 硕士. 研究方向为静态时序分析.

E-mail: 18216405639@163. com.

金湘亮 男, (1974-), 博士, 教授. 研究方向为 SOC 设计、系统设计.