

文章编号: 1006-2475(2014)04-0064-05

基于多分类器投票机的人体姿态识别算法

陈慧杰, 谢毅雄

(厦门大学信息科学与技术学院, 福建 厦门 361005)

摘要: 为了获得准确的人体姿态识别结果, 满足智能视频监控的需求, 提出一种融合 Hu 不变矩特征和傅里叶描述子特征的人体姿态识别算法, 并将 ReliefF 算法引入特征选择过程中, 区分特征的重要性, 然后使用大数投票法构建多分类器投票机制进行姿态识别, 该机制很好地发挥了各个分类器的优势, 提高了识别的准确率。实验结果表明, 提出的算法对各种姿态取得了很好的分类效果。

关键词: 姿态识别; Hu 不变矩; 傅里叶描述子; 多分类器

中图分类号: TP391 文献标识码: A doi: 10.3969/j.issn.1006-2475.2014.04.015

Human Posture Recognition Method Based on Multiple Classifiers

CHEN Hui-jie, XIE Yi-xiong

(School of Information Science and Engineering, Xiamen University, Xiamen 361005, China)

Abstract: To get an accurate body posture recognition result and meet the needs of intelligent video surveillance, we present a fusion of Hu invariant moments and Fourier descriptors characteristics of human posture recognition algorithm. The ReliefF feature selection algorithm is introduced into the process, for distinguishing the importance of characteristics. Voting mechanism of using multiple classifier is constructed with large numbers to vote, plays good advantages of the various classifiers to improve the recognition accuracy. Finally, experimental results show that the proposed algorithm achieves good classification results for various postures.

Key words: posture recognition; Hu invariant moments; Fourier descriptors; multiple classifier

0 引言

人体姿态识别是智能视频监控系统中的一种关键技术, 为稳健而有效的行为识别和事件发现提供保障。当前人体姿态识别方法主要有 2 种: 基于模型匹配^[1]的方法和基于状态空间^[2]的方法。模型匹配法首先需要将视频图像系列转换成一系列的运动模式, 然后根据先验知识训练并构建模板。在识别的时候对输入的视频系列与构建好的模板进行特征相似度的计算。该方法计算量小、操作较为简单, 但是没有考虑相邻帧之间的动态特性, 鲁棒性较差。Davis 和 Bobick^[3]提出了时空模板方法, 将目标信息转换成能量图像和历史图像。Efros^[4]等人提出用校正和模糊的方法提高光流法对噪声的鲁棒性。Gritai^[5]利用 13 个点的人体模型描述人体姿势。基于状态空间的方

法是将视频每一帧中的行为姿态作为一个个的状态节点, 然后将这些节点根据发生的概率统计联系起来。在识别的时候将动作看成是对这些行为姿态状态节点的遍历, 以此来计算出这些节点的联合概率, 以概率的大小作为分类的标准。Buccolieri^[6]等人用神经网络通过分析人体轮廓来识别姿态, Liu^[7]等人利用隐马尔可夫模型来描述人体的运动状态。为了准确地描述人体的姿态以及克服单分类器的局限性, 本文提出多特征融合的姿态描述与多分类器的投票机制相结合的算法, 并将 ReliefF 特征选取算法引入到识别过程中, 有效地提高了姿态识别的识别率。

1 人体姿态识别算法

本文提出的人体姿态识别过程分为 4 步。(1) 前景分割, 采用 VIBE^[8]背景建模方法完成前景目标

收稿日期: 2014-03-03

作者简介: 陈慧杰(1982-) 男, 河北石家庄人, 厦门大学信息科学与技术学院硕士研究生, 研究方向: 计算机视觉, 行人检测与跟踪, 人脸识别; 谢毅雄(1987-) 男, 福建安溪人, 硕士研究生, 研究方向: 计算机视觉, 图像处理, 行人检测。

的分割;(2)提取人体姿态特征,共提取了2类15个特征,并用Relieff算法进行特征筛选^[9];(3)多分类器投票机的构建,选用了朴素贝叶斯^[10]、决策树^[11]、随机森林3种分类方法进行训练构建分类模型;(4)输出姿态识别的结果。

1.1 前景分割

该步骤使用VIBE算法来进行背景建模。背景模型为每个背景点存储了一个样本集,然后将每一个新的像素值与样本集进行比较来判断是否属于背景点。如果一个新的观察值属于背景点,那么它应该和样本集中的采样值比较接近。

记 $P_t(x)$ 为像素点 x 在时刻 t 的值。为了对 $P_t(x)$ 进行分类,将它与它所属的样本像素集做比较,如果在欧式距离空间中与它足够接近的样本点数超过一个阈值,即以 $P_t(x)$ 为中心, $SR(P_t(x))$ 为半径的区域内的样本点数大于一个预先设定的阈值 \min ,则把 $P_t(x)$ 归类为背景像素。背景模型的初始化通过第一帧图像完成,对于一个像素点,随机地选择它的邻居点的像素值作为它的模型样本值。提取的前景效果如图1所示。



(a) 近景图 (b) 远景图

图1 前景提取图

1.2 特征提取

当人体做出不同动作的时候,人体姿态的形状和轮廓会产生不一样的变化,因此本文选用表征形状和轮廓的2种特征Hu不变矩^[12]和傅里叶描述子^[13]作为姿态识别的特征,然后用Relieff算法对提取的特征进行筛选。

1.2.1 Hu不变矩

Hu不变矩是一种表征形状的特征。二维随机变量的 $p+q$ 阶矩的定义如下:

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy, p, q = 0, 1, 2, \dots \quad (1)$$

其中,图像函数 $f(x, y)$ 是分段的连续有界函数, $f(x, y)$ 随着平移、旋转或尺度的变化而变化时, k 阶矩也可能随之变化。故为了获得不变的特征,定义中心矩如下:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy, p, q = 0, 1, 2, \dots \quad (2)$$

其中 $\bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}}$ 是图像 $f(x, y)$ 的中心点,可知中心矩对于图像的平移具有不变性。

尺度不变性可以通过归一化获得,归一化中心矩公式如下:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}}}, \gamma = (p+q+2)/2, p+q=2, 3, \dots \quad (3)$$

基于归一化的中心矩,引入7个Hu不变矩如下:

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &\quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

这7个Hu不变矩对于尺度、平移和旋转具有不变性,表1为不同姿态的Hu不变矩。

表1 对不同姿态取Hu不变矩

姿态	H ₁	H ₂	H ₃	H ₄	H ₅	H ₆	H ₇
	0.245239	0.006240	0.000095	0.000086	-0.000000	-0.000002	-0.000000
	0.222360	0.011196	0.000459	0.000003	-0.000000	-0.000000	0.000000
	0.477028	0.139629	0.019957	0.000785	0.000003	0.000281	0.000001
	0.403808	0.097080	0.007311	0.000001	-0.000000	-0.000000	-0.0000001
	0.269680	0.028974	0.002485	0.000275	-0.000000	-0.000022	0.000000
	0.283093	0.018934	0.003287	0.000265	0.000000	-0.000031	-0.000000

1.2.2 傅里叶描述子

傅立叶描述子(FD)是物体边界曲线的傅里叶变换系数,它是物体边界曲线频域信息的分析结果。少量的傅里叶描述子就可以描述物体边界的整体形状,并且它对图像旋转、平移、缩放和初始点的选取均不

敏感。

傅里叶描述子定义如下: 令 $x[m]$ 和 $y[m]$ 是一个给定二维图像边界上的第 m 个像素点的坐标, 定义一个复数:

$$z[m] = x[m] + iy[m] \quad (4)$$

则该图像的形状的傅里叶描述子(FD)定义如下:

$$Z[k] = \text{DFT}[z[m]] = \frac{1}{N} \sum_{m=0}^{N-1} z[m] e^{-j2\pi mk/N}, k=0, \dots, N-1 \quad (5)$$

其中 N 是图像边界上像素点的总数, 傅里叶描述子可以反映曲线的形状特征, 而且傅里叶变换具有能量集中性, 较少的傅里叶描述子就可以用来表征完整的边界曲线。低频系数集中了傅里叶描述子的大部分能量, 反映了轮廓曲线的整体形状。因此, 可以选用低频系数的 8 个傅里叶描述子作为特征。

本文首先使用 Canny 算子^[14]对运动目标提取边缘, 然后再进行边缘傅里叶描述子的计算, 表 2 为对不同姿态取 8 个低频傅里叶描述子。

表 2 对不同姿态轮廓取 8 个低频傅里叶描述子

姿态	FD ₁	FD ₂	FD ₃	FD ₄	FD ₅	FD ₆	FD ₇	FD ₈
	0.213467	0.249688	0.153127	0.202424	0.142016	0.171571	0.188637	0.083986
	0.221345	0.216283	0.201103	0.301324	0.190826	0.025634	0.026168	0.129071
	0.157087	0.197067	0.139511	0.076835	0.056398	0.035651	0.050639	0.077684
	0.154758	0.231615	0.154001	0.080205	0.065141	0.073142	0.129333	0.116277
	0.084562	0.157482	0.039263	0.097382	0.016461	0.160881	0.038625	0.069876
	0.074522	0.150363	0.023008	0.163409	0.053563	0.120720	0.037785	0.035857

1.2.3 ReliefF 特征筛选

ReliefF 算法是一种典型的特征筛选模型算法, 它可以有效且正确地评估特征的质量。ReliefF 算法主要是根据特征值在区分接近样本的能力上对特征的质量进行评估。假设有样本空间 R_1, R_2, \dots, R_N , 这个样本空间由属性集 $F = \{f_1, f_2, \dots, f_m\}$ 来描述, m 是特征属性的数目, 并且标注上目标的类标签 T 。在这

种假设下, 给定一个随机选择的样本 R_i , ReliefF 算法搜索与之最相近的 2β 个样本, 其中 β 个来自同一类, 把它称作 $HITH_j(R_i)$, 另 β 个来自不同的类, 把它称作 $MISSM_j(R_i)$, 然后通过计算 $R_i, M_j(R_i)$ 和 $H_j(R_i)$ 各个特征值上的间隔并累加起来更新所有特征 F 的质量权重 $\lambda(F)$ 。如果样本 R_i 和 $H_j(R_i)$ 在 F 上距离较远, 那么我们减小质量权重 $\lambda(F)$, 另一方面, 如果样本 R_i 和 $M_j(R_i)$ 在 F 上距离较远那么表明特征 F 将 2 个不同类的样本区别开来了, 因此增加质量权重 $\lambda(F)$ 。用 $F_i = \{f_{i1}, f_{i2}, \dots, f_{im}\}$ 表示第 i 个样本的 m 个属性值, λ_f^l 表示第 l 次更新时属性 f 的权值。其权值更新函数定义如下:

$$\lambda_f^{l+1} = \lambda_f^l - \sum_{j=1}^{\beta} \text{diff}(f, R_i, H_j(R_i)) / m\beta + \frac{\sum_{T \neq \text{class}(R_i)} \frac{p(T)}{1 - p(\text{class}(R_i))} \text{diff}(f, R_i, M(T))}{m\beta} \quad (6)$$

如果属性是离散值, 那么:

$$\text{diff}(f, R_i, R'_i) = \begin{cases} 0, & R_{if} = R'_{if} \\ 1, & R_{if} \neq R'_{if} \end{cases} \quad (7)$$

如果属性值是连续的, 那么:

$$\text{diff}(f, R_i, R'_i) = \frac{|R_{if} - R'_{if}|}{\max(f) - \min(f)} \quad (8)$$

其中 $\max(f)$ 与 $\min(f)$ 为属性 f 的上下界。

本文在选用的 15 个特征上使用 ReliefF 特征选取算法, 每个特征的权重如表 3 所示。

表 3 特征权值

属性序号	特征名称	属性权重	权重排序
1	H ₁	0.053519	1
2	H ₂	0.0240419	2
3	H ₃	0.0031145	10
4	H ₄	0.0009652	11
5	H ₅	0.0000608	13
6	H ₆	0.0006873	12
7	H ₇	-0.000055	15
8	FD ₁	0.0000513	14
9	FD ₂	0.0148706	6
10	FD ₃	0.0116442	7
11	FD ₄	0.0150962	5
12	FD ₅	0.0229138	3
13	FD ₆	0.0097971	9
14	FD ₇	0.0115406	8
15	FD ₈	0.0188737	4

从表 3 可以看出, 特征 7、8、5、6 的权重最小, 即

对分类效果的贡献最小。因此,将特征7、8、5、6剔除,选取特征1、2、12、15、11、9、10、14、13、3、4作为特征集。

1.3 构建多分类器投票机

本文选取朴素贝叶斯、决策树、随机森林3种运算量较小的分类器,然后再通过投票机制来构建多分类器投票机。

朴素贝叶斯分类器^[15]的最大优点是,对于分类它只需要少量的训练数据来做参数(变量的均值和方差)的估计。决策树分类器选择C4.5算法^[16-17]进行属性选择度的度量。随机森林分类器^[18]在建立每一棵树的时候都构建了一个随机的特征选择子集。

一般来说,比较3种分类器,随机森林会获得最好的性能。然而对于不同的数据集,不同的分类器的优劣性有所不同。因此,本文采用投票的机制来进行最后的分类决策。当前主要的投票方法有:

(1) 平均法:它将所有分类器算出来的分布累加后平均,求得的平均值作为决策的结果。

(2) 大数投票法:记录下投票的结果,记录投票得票最多的票数是多少,选取得票最多的类别。为了解决在多个类别值所得票数一样这种情况,还必须记录得票最多的几个类别值索引号,再随机选择其中之一,设置为最终结果。

(3) 中位数投票法:顾名思义是取分类器结果的中位数结果作为决策。

本文使用大数投票法进行最终的决策。这样可以发挥各个分类器的优势,提高识别的准确率。

2 实验结果与分析

本文采用的视频数据为自主拍摄的多人行为视频。针对视频进行了走、跑、摔倒、蹲、弯腰走姿态的训练以及测试。

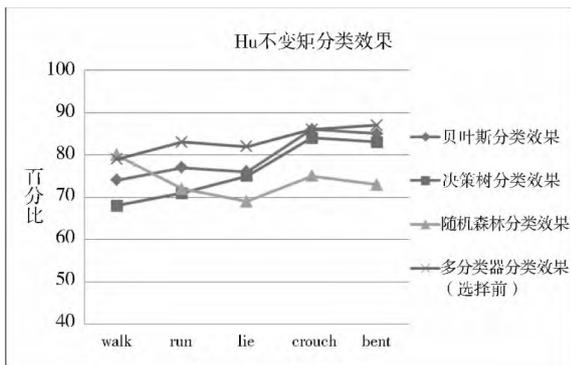


图2 Hu不变矩分类结果

图2和图3分别为单独使用7个Hu不变矩和单独使用8个低频傅里叶描述子进行识别的结果,从图中可以看出,对于不同的姿态,单一分类器均有不同的优势,在使用了多分类器投票机之后,识别率有了较大的提高。

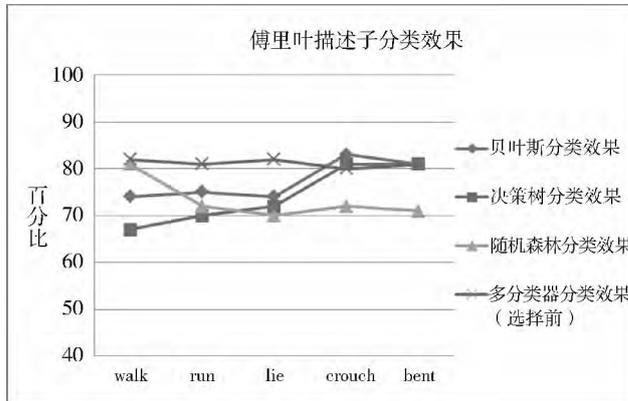


图3 傅里叶描述子分类结果

图4为进行了特征融合后分别使用3种不同的分类器以及多分类器分类的效果,从图中可以看出经过特征融合后明显提高了整体的识别率。

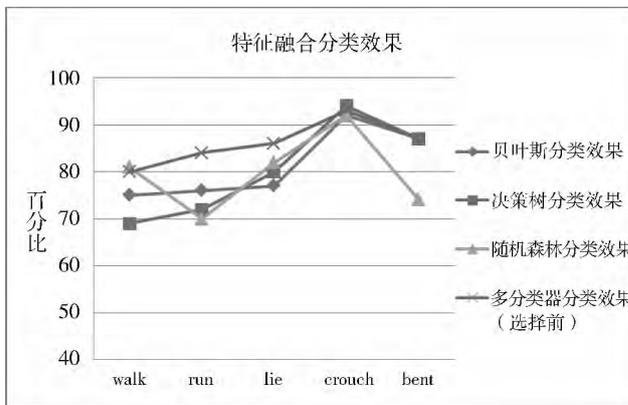


图4 特征融合分类结果

使用Relieff算法对特征进行权重评估后,选用1.2.3节的结果作为特征集进行实验,图5为特征筛选前后的分类效果,从图中可以看出在特征筛选之后分类效果有了进一步的提高。

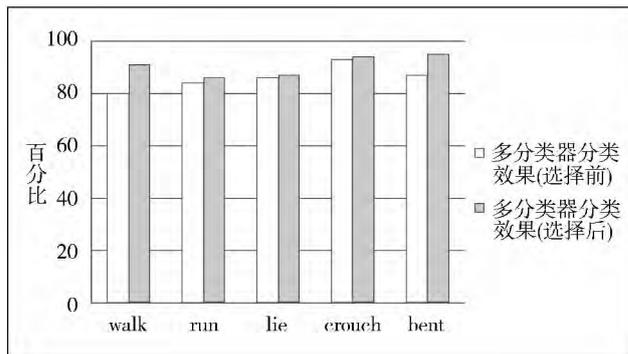


图5 特征筛选前后分类结果

3 结束语

本文提出一种多特征融合的人体姿态识别算法,并将特征选择算法和多分类器投票机引入姿态识别。选择的 Hu 不变矩特征对于图像的尺度、平移和旋转具有很好的鲁棒性,低频傅里叶描述子完整地反映了轮廓曲线的整体形状。使用大数投票法构建的多分类器投票机制,很好地发挥了各个分类器的优势,提高了识别的准确率。最后的实验结果表明算法对提出的各种姿态取得了很好的分类效果。对于实际场景的智能监控中的应用,所提出的人体姿态的种类还不够丰富,有待于扩大姿态特征库,从而更加完善其在实际中的应用。

参考文献:

- [1] Silva G B S, Mello M P, Shimabukuro Y E, et al. Multi-temporal classification of natural vegetation cover in Brazilian Cerrado [C]// Proceedings of the 6th International Workshop on the Analysis of Multi-temporal Remote Sensing Images. 2011: 117-120.
- [2] Cao Long-bing, Ou Y, Yu P. Coupled behavior analysis with applications [J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(8): 1378-1392.
- [3] Davis James W, Bobick A F. The representation and recognition of human movement using temporal templates [C]// Proceedings of IEEE 1997 Conference on Computer Vision and Pattern Recognition. 1997: 928-934.
- [4] Alexei A Efros, Alexander C Berg, Greg Mori, et al. Recognizing action at a distance [C]// Proceedings of the 9th IEEE International Conference on Computer Vision. 2003, 2: 726-733.
- [5] Gritai Alexei, Sheikh Yaser, Shah Mubarak. On the use of anthropometry in the invariant analysis of human actions [C]// Proceedings of the 17th International Conference on Pattern Recognition. 2004, 2: 923-926.
- [6] Buccolieri F, Distanto C, Leone A. Human posture recognition using active contours and radial basis function neural network [C]// Proceedings of the IEEE 2005 Conference on Advanced Video and Signal Based Surveillance. 2005: 213-218.
- [7] Liu Chin-de, Chung Y N, Chung P C. An interaction-embedded HMM framework for human behavior understanding: With nursing environments as examples [J]. IEEE Transactions on Information Technology in Biomedicine, 2010, 14(5): 1236-1246.
- [8] Olivier Barnich, Marc Van Droogenbroeck. ViBe: A universal background subtraction algorithm for video sequences [J]. IEEE Transactions on Image Processing, 2011, 20(6): 1709-1724.
- [9] 张勇. 基于 ReliefF 算法的模糊聚类新算法 [J]. 华南金融电脑, 2009, 17(1): 43-46.
- [10] Statsoft. Naive Bayes Classifier [EB/OL]. <http://www.statsoft.com/textbook/naive-bayes-classifier/>, 2014-03-03.
- [11] Julien Boeuf, Pascal Belin, Henri Maitre. Decision Tree Classifier [EB/OL]. http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/lguo/decision-Tree.html, 2013-03-15.
- [12] Huang Zhihu, Leng Jinsong. Analysis of Hu's moment invariants on image scaling and rotation [C]// Proceedings of the 2nd International Conference on Computer Engineering and Technology (ICCTET). 2010, 7: 476-480.
- [13] Julien Boeuf, Pascal Belin. Overview of the Fourier Descriptors Methods [EB/OL]. http://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/descript_fourier/Part1.html, 2014-03-03.
- [14] Canny J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 679-698.
- [15] Ling C X. Learnability of augmented naive Bayes in nominal domains [C]// Proceedings of the 18th International Conference on Machine Learning. 2001: 617-623.
- [16] 吴楠, 宋方敏. 用 C4.5 算法对局域网数据报进行行为分类 [J]. 计算机技术与发展, 2006, 16(7): 1-3.
- [17] 李楠, 段隆振, 陈萌. 决策树 C4.5 算法在数据挖掘中的分析及其应用 [J]. 计算机与现代化, 2008(12): 160-163.
- [18] Ho T K. The random subspace method for constructing decision forests [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(8): 832-844.