

DOI: 10.15923/j.cnki.cn22-1382/t.2017.6.06

FD+HOG 融合人体行为识别

史东承, 衡瑶瑶*

(长春工业大学 计算机科学与工程学院, 吉林 长春 130012)

摘要: 利用全局特征提取与局部特征提取相融合的方法, 将傅里叶变换特征与方向梯度直方图特征相融合的算法结合支持向量机(SVM)来识别人体行为。首先将 KTH 和 Weizmann 数据集中提取出来的图像做预处理, 分别提取出每张图像的傅里叶描述子和 HOG 特征, 然后利用主成分分析法对 HOG 提取特征降维, 最后用 FD+HOG 的融合特征放入 SVM 分类器中分类识别。实验结果表明, 该算法识别率可达 86% 以上。

关键词: 方向梯度直方图; 傅里叶变换; 支持向量机

中图分类号: TP 391 **文献标志码:** A **文章编号:** 1674-1374(2017)06-0550-05

Human behavior recognition based on FD+HOG fusion

SHI Dongcheng, HENG Yaoyao*

(School of Computer Science & Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: Combing global with local feature extraction, we apply both Fourier transform features and direction gradient histogram into Support Vector Machine (SVM) to recognize human behavior. Images from the KTH and Weizmann data sets are processed to extract Fourier Descriptor (FD) and HOG feature of each image, and then principal component analysis is used to extract feature dimensionality from HOG. Finally, fusion features of FD+HOG are put into the SVM classifier. Experimental results indicate that the recognition rate of the algorithm is greater than 86%.

Key words: Histogram of Oriented Gradient (HOG); Fourier Transform (FT); Support Vector Machine (SVM).

0 引言

随着人类社会的稳步发展, 计算机视觉领域的蒸蒸日上, 人体行为识别技术也慢慢深入到我

们生活的方方面面, 人类的衣食住行已离不开科技产品的辅助。人体行为识别技术的发展对社会公共设施提供了便利, 例如养老院、医院、警察局等处处体现了该技术的重要性。人体行为识别主

收稿日期: 2017-05-25

基金项目: 吉林省教育厅“十三五”规划项目(吉教合字[2016]第 349 号)

作者简介: 史东承(1959-), 男, 汉族, 吉林长春人, 长春工业大学教授, 硕士, 主要从事图像处理与机器视觉方向研究, E-mail: shidongchen@ccut.edu.cn. * 通讯作者: 衡瑶瑶(1992-), 女, 汉族, 山西晋城人, 长春工业大学硕士研究生, 主要从事图像处理与机器视觉方向研究, E-mail: 514403733@qq.com.

要研究特征提取和分类识别两个方面。近年来,研究者们已经提取出了解决特征提取识别率低的方法。在特征提取研究中主要是根据图像的纹理、梯度、形状等特征提出了许多特征描述子,主要包括局部二值模式(Local Binary Patterns, LBP)、方向梯度直方图(Histogram of Oriented Gradient, HOG)、尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)以及傅里叶描述子(Fourier Descriptor, FD)等。

图像形状的特征方法:

1)区域的形状特征提取,如Hu不变矩^[1]分别提取训练样本和测试样本的HOG和加权Hu矩自适应融合参数将两者融合后,进行模板匹配达到识别的效果对噪声干扰有很好的鲁棒性。

2)边界的形状特征提取,如HOG^[2]是对织物疵点的检测,通过疵点的灰度和纹理特征用HOG提取的纹理特征,超像素分割提取灰度特征,用低秩分解融合特征来检测疵点。和傅里叶变换^[3]将人体姿势的二维轮廓进行傅里叶变换生成傅里叶描述子结合SVM进行姿势识别,其他方法,如小波^[4]使用离散小波纹理特征获得均匀完整局部二值模式变换特征提取技术产生了最好的分类精度。

在分类识别研究方面主要有K最近邻(K Nearest Neighbor, KNN)、隐马尔可夫模型(Hidden Markov Model, HMM)、支持向量机(Support Vector Machine, SVM)等方法。文献

[5]证明了KNN分类器的性能显著地提高了在给定数据集中成对比较特征时基于训练类组统计的两个标准的使用。文献[6]提出了一种交互的嵌入式隐马尔可夫模型框架的自动检测和分类人类个体行为和群体互动。文献[7]提出了一种结合测地线图和支持向量机(SVM)的人工关节估计算法。文献[8]提出一种基于隐马尔可夫模型和支持向量机混合模型的音频分类方法,用于语音、音乐、语音+音乐、静音四类音频分类。考虑到提取特征复杂度和识别率,文中提出了基于改进后的HOG特征+FD特征融合的方法表征图像的特征+SVM完成分类识别。该方法首先分别提取出人体行为数据库中每帧图像的HOG特征和FD特征,然后用PCA对HOG特征降维,将降维后的HOG特征改进与FD特征融合放入SVM中训练识别。实验表明,文中算法在人体行为识别中识别率可达到86%以上。

1 研究方法

文中算法大概流程有三部分:目标检测、特征融合、分类识别。人体行为数据库为小段视频,首先用背景差分法将视频帧中的前景目标提取出来保存成图片格式,然后对图片进行预处理获取人体的轮廓形状,将人体形状作为输入提取图像的特征,最后进入SVM分类识别。

人体行为识别算法流程如图1所示。

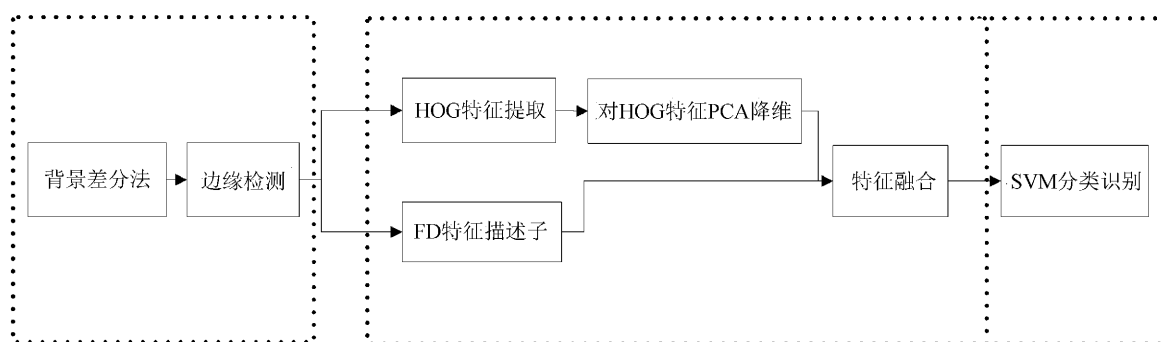


图1 人体行为识别算法流程

1.1 背景差分法提取目标前景

背景差分法实际上就是根据视频序列的图像构建一个背景模板,后面的图像与其做减法的方法。文中采用中值背景建模的方法,选视频帧的前15帧来作为背景建模,然后从16帧开始与背

景图形做差分运算,选择出合适的阈值对差分图像作二值化。得到的目标图像做边缘检测得到目标轮廓。

KTH数据库中提取的轮廓图如图2所示。



图 2 KTH 数据库中提取的轮廓图

1.2 傅里叶描述子

FD 描述全局特征的矩阵,表示图像轮廓形状的特征,将处理后的图像轮廓曲线用坐标点构建成一维序列后,对其进行一维的傅里叶变换,从而获得可以描述轮廓信息的一系列傅里叶系数。其能量主要集中于少数几个低频傅里叶系数,因此少数系数就可以描绘该序列特征。人体行为轮廓通过上面方法已经得到,取轮廓任一点 (x_0, y_0) 开始,沿着图像轮廓逆时针方向前进,轮廓上的点表示为 $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$,构成一条封闭轮廓曲线,在 X-Y 坐标平面看成一个复平面,则可以看成一系列一维的复数序列 c_n ,即 $c_n = x_n + jy_n (n=1, 2, \dots, n-1)$ 。接下来对该离散序列做傅里叶变换。轮廓的傅里叶描述子见下式:

$$c(k) = \sum_{n=0}^{N-1} c(n) e^{-j\pi kn/\pi} \quad (1)$$

$$n = 0, 1, 2, \dots, N-1$$

1.3 HOG 特征提取

HOG 是描述局部特征的算法。它提取图像特征的方式是计算图像局部区域的梯度和方向,然后进行统计得到。在一幅图像中,图形局部的边缘或梯度的方向密度分布能够很好地描述图像的形状。对图像进行 HOG 特征提取的步骤如下:

- 1) 标准化。输入提取轮廓后的图像,将像素点灰度值通过伽马校正归一化 $[0, 1]$ 范围。
- 2) 计算每个像素的梯度。采用 $[-1, 0, 1]$ 的梯度算子来计算图像的水平 and 垂直的梯度。
- 3) 图像分割为单元格。把整个图像分割为一个一个的 Cell 单元格(16×16 像素)。
- 4) 单元格构建梯度直方图。将图像分成若干的 Cell 单元格(16×16 像素),把 $0^\circ \sim 360^\circ$ 分成 9 个 bin 作为梯度的方向,对 cell 内每个像素用梯度方向在直方图中进行加权投影,就可以得到这

个 cell 的梯度方向直方图了,即对应 9 维特征向量。

5) Block 块内梯度直方图。把各个细胞单元组合成大的、空间上连通的区间(blocks)。这样,一个 block 内所有 cell 的特征向量串联起来便得到该 block 的 HOG 特征。

6) 收集特征。将图像中的所有 Block 特征描述子收集起来就是整张图像的特征描述子。

KTH 人体行为数据库中跑步动作的梯度方向直方图如图 3 所示。

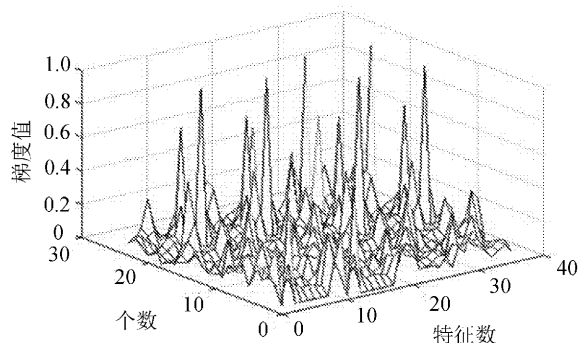


图 3 KTH 人体行为数据库中跑步动作的梯度方向直方图

1.4 对 HOG 特征降维

主成分分析(Principal Component Analysis, PCA),通常是用来为高维特征降维。PCA 的步骤为:

- 1) 文中将图像的尺寸归一化为 120×90 大小的图片,用 HOG 提取图像特征后得到 30×24 维的特征描述子;
- 2) 通过协方差公式^[9]可以得到 24×24 维的 HOG 的协方差矩阵,得到特征向量和特征根;
- 3) 根据贡献量从大到小排列得到新的矩阵,将每张图像降维到 576 维。

1.5 FD+HOG 特征融合

FD 是一个基于全局提取特征描述子的算法, HOG 是一个基于局部提取特征描述子的算法, 将两个算法中的改进部分相融合, 可以有效地提高人体行为识别的识别率。傅里叶变换可以将能量集中在极少的几个低频系数上, 即用少数的几个特征描述子就可以很好地保证序列帧的特征。经过多次试验结果证明, 在 KTH 数据库中采用 6 维的 FD 可以集合 SVM, 达到 80.33% 的识别率。HOG 特征经过 PCA 降维后得到 576 维的特征描述子, 根据 PCA 的原理特征, 矩阵的特征值是由大到小排列的, 根据其贡献量大小排列。文中算法选取其贡献量的 90% 来识别效果最好, 可以达到 83.33% 的识别率。文中的算法是将两个特征级联融合生成一个新的特征矩阵。

1.6 SVM 分类器分类识别

SVM 是可以用在很多领域的判别方法。它是一种监督的学习模型, 在模式识别、分类和回归的研究上扮演着重要的角色。SVM 方法是通过一个非线性映射, 把样本空间映射到一个高维乃至无穷维的特征空间中, 非线性可分问题转换到线性可分问题。其实就是将输入变量全部映射的一个高维空间, 在构造出最优分类超平面用作分类识别。分类函数为:

$$f(x) = w^T x + b \quad (2)$$

其中, w 是与空间维数有关的, 对于样本点 (x_i, y_i) , 样本点到分类超平面的距离为 $y_i(w^T x_i + b)$, 对 w 和 b 进行归一化, 用 $w/\|w\|$ 和 $b/\|w\|$ 分别代替 w 和 b , 即几何间隔为 $f(x)/\|w\|$, 求最大几何间隔为等价于:

$$\min \frac{1}{2} \|w\|^2 \quad (3)$$

$$\text{s.t. } (w^T x_i + b) \geq 1, i = 1, 2, \dots, n$$

式(3)为线性分类情况, 非线性分类按照式(3)进行分类可能会降低准确率, 因此, 加入松弛变量和惩罚因子, 则非线性的分类如下:

$$\begin{aligned} \min \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i \\ \text{s.t. } y_i(w^T + b) \geq 1 - \xi_i \\ i = 1, \dots, n; \xi_i \geq 0, i = 1, 2, \dots, n \end{aligned} \quad (4)$$

另

$$w = a_1 x_1 y_1 + a_2 x_2 y_2 + \dots + a_n x_n y_n$$

其中, n 为总样本数, 则最优分类超平面的求取为:

$$\begin{aligned} f(x) = \left\langle \sum_{i=1}^n (a_i x_i y_i), x \right\rangle + b = \\ \sum_{i=1}^n a_i y_i \langle x_i, x \rangle + b \end{aligned} \quad (5)$$

即转换成求内积^[10]。

2 实验与分析

2.1 实验数据

文中实验是在 Matlab 的仿真环境下完成的, 采用 KTH 和 Weizmann 人体行为库来验证文中算法的有效性。从 KTH 数据库中选出 6 个动作类, 包括拳击、鼓掌、画圈、慢跑、散步、跑步。在 Weizmann 库中也选出 6 个动作类, 包括跳起鼓掌、跳跃、单腿跳、跑步、散步、快速双腿跳。在数据库的每类动作中随机抽取 60 张图像, 其中 50 张作为训练集, 剩下的 10 张作为测试集。

2.2 实验过程及分析

首先, 将数据库中选取的图像目标检测后得到特征提取算法的输入图像, 处理过后的示例如图 4 所示。

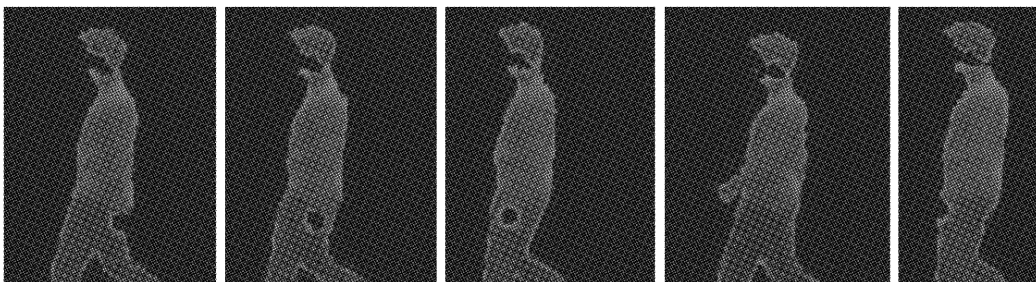


图 4 KTH 目标检测后的示例图

完成第一步后,把结果作为输入分别提取出每张图像的 FD 和 HOG, HOG 特征描述子用 PCA 降维后按照其贡献率大小来排列,故选取其 90% 能量的描述子与 FD 级联,然后将生成的新的描述子放入 SVM 中进行训练。训练时依次把一个动作类别的全部样本归为一个动作类,把其

他动作类的样本归为另外一类,文中两个库中选用的 6 个动作类就可以构造出 6 个不同 SVM。有了各自的分类器后,就可以进行最后的分类。将文中算法与提取单个特征的算法相比,识别率占明显优势。不同算法基于 KTH、Weizmann 数据集的识别率的比较见表 1。

表 1 不同算法基于 KTH、Weizmann 数据集的识别率的比较

算法	KTH	Weizmann
HOG+SVM	35%(21/60)	40%(24/60)
HOG+PCA+SVM	83.33%(50/60)	76.66%(46/60)
FD+SVM	80.33%(49/60)	66.66%(40/60)
文中算法	91.66%(55/60)	86.33%(53/60)

3 结 语

综上所述,文中将全局特征与局部特征相融合算法用在人体行为识别中。通过对提取的 HOG 特征描述子降维与 FD 想融合构造一个全新的特征矩阵,弥补了单独采用全局描述子或局部描述子的缺陷,有效地提高了分类器的识别率,在 KTH 和 Weizmann 数据集的试验中,文中算法得到了证实,识别率最高,可达到 91.66%。

参考文献:

- [1] 梁琛华,常青,胡谋法.基于加权 Hu 矩和 HOG 特征的自适应融合人体行为识别新方法[J].现代电子技术,2014(1):14-18.
- [2] 刘洲峰,闫磊,李春雷,等.基于特征融合与低秩分解的织物疵点检测[J].棉纺织技术,2017(10):1-4.
- [3] 崔广才,窦凤平,王春才,等.基于傅里叶与局部特征结合的人体姿态识别方法研究[J].长春理工大学学报:自然科学版,2016(1):82-87.
- [4] Yadav A R, Anand R S, Dewal M L, et al. Multi-resolution local binary pattern variants based texture feature extraction techniques for efficient classification of microscopic images of hardwood species[J]. Applied Soft Computing,2015,32(C):101-112.
- [5] Bhattacharya G, Ghosh K, Chowdhury A S. Granger causality driven AHP for feature weighted kNN [J]. Pattern Recognition,2017,66:425-436.
- [6] Liu C D, Chung Y N, Chung P C. An interaction-embedded HMM framework for human behavior understanding: with nursing environments as examples[M]. [s.n.]: IEEE Press,2010.
- [7] Kim H, Lee S, Kim Y, et al. Weighted joint-based human behavior recognition algorithm using only depth information for low-cost intelligent video-surveillance system[J]. Expert Systems With Applications an International Journal,2016,45(C):131-141.
- [8] 史东承,韩玲艳,于明会.基于 HMM/SVM 的音频自动分类[J].长春工业大学学报:自然科学版,2008,29(2):178-182.
- [9] Kim J M, Chung K, Kang M. Continuous gesture recognition using HLAC and low-dimensional space [J]. Wireless Personal Communications,2015,86(1):1-16.
- [10] Cervantes J, García Lamont F, López-Chau A, et al. Data selection based on decision tree for SVM classification on large data sets[J]. Applied Soft Computing,2015,37(C):787-798.