

基于区域形状与运动特征的实时行为识别

裴利沈¹ 董乐¹ 赵雪专² 任鹏³

(电子科技大学计算机科学与工程学院 成都 611731)¹

(中国科学院成都计算机应用研究所 成都 610041)²

(电子科技大学空天科学技术研究院 成都 611731)³

摘要 提出了一种基于推广的Hu不变矩特征的实时行为识别方法。首先,对Hu不变矩进行改进,使其在离散情况下同时具有平移、旋转和比例不变性。然后,结合运动目标的速度将目标行为刻画成结合Hu矩新特征和速度特征的13维特性向量。其中,Hu矩新特征表征了行为的区域形状特性,速度特征反映了行为的运动特性。随后采用预先定义的一些行为作为先验知识样本训练支持向量机,并最后使用其对待检测行为进行分类以达到行为识别的效果。所提方法计算效率高,能够实时检测人体行为。在处理实拍视频数据的实验中,该方法表现出了理想的处理效率以及识别精度。

关键词 行为识别,区域形状,Hu不变矩,运动特征

中图法分类号 TP391 **文献标识码** A

Real-time Action Recognition Based on Zone Shapes and Motion Features

PEI Li-shen¹ DONG Le¹ ZHAO Xue-zhuan² REN Peng³

(College of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)¹

(Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610041, China)²

(Aerospace Science and Technology Research Institute, University of Electronic Science and Technology of China, Chengdu 611731, China)³

Abstract This paper presented an efficient action recognition method based on Hu moment invariant features. Firstly, the Hu moment invariants were refined to be new features that are translation, rotation and scale invariant. Then an action was characterized by a 13-dimensional feature vector consisting of both Hu moment features and action speed features. The Hu moment features represent the Zone shape of the action, and the action speed features exhibit certain motion characteristics. Finally, a support vector machine(SVM), which is trained using labeled action frames, was applied to classify test sample actions into different categories. The proposed method is performed on real-world videos and achieves acceptable recognition rates with desirable computational efficiencies.

Keywords Behavior recognition, Zone shape, Hu moment invariant, Motion feature

1 引言

目前,智能监控系统广泛应用于各种场合,如火车站、居民小区、商场等。相对于传统监控系统仅用于事后采证的性能来说,智能监控系统能够早发现、早预警,因此可以减少大量犯罪行为的发生,在很大程度上保障了社会公共安全。

行为识别是智能监控的一种关键技术,运用不同的行为特征已有多种实现方法。国内外众多学者在此方向做了很多工作。目前研究行为识别的主要方法可以分为两大类,第一类是基于模板匹配的行为识别方法^[1-3],该方法的两种典型的模板是运动历史图像与运动能量图像^[1]。运动历史图像是一个标量值图像,它的每一个像素的亮度都是这一点处最近运动历史的函数。运动能量图像是一个二值图像,它记录了在

图像序列中运动发生的位置。第二类是基于跟踪的行为识别方法^[4],该方法依赖于跟踪人的运动轨迹进行行为的判别。然而,无论是基于模板的行为识别方法还是基于跟踪的行为识别方法都有其自身的局限性。基于模板的方法一般计算量比较大,速度相对较慢,很难做到实时检测。基于跟踪的方法在很大程度上依赖于跟踪的效果,目前,在复杂场景中的跟踪仍是一个难题,因此基于跟踪的行为识别方法很难在实际中得到应用。本文提出了一种基于模板匹配的快速方法,即结合区域形状特征与运动速度特征对行为进行描述,从而进行快速识别的方法。

为了对行为进行实时的检测,本文采用推广的Hu矩特征向量对区域形状特征进行描述。1962年,M. K. Hu给出了连续函数矩的定义和关于矩的基本性质^[5],证明了在连续函

到稿日期:2012-07-19 返修日期:2012-11-07 本文受国家自然科学基金重大研究计划(91024026),国家自然科学基金(61003123,61105005),中央高校基本科研业务费(ZYGX2011X014,23401039)专项资金资助。

裴利沈(1988-),女,博士生,主要研究方向为模式识别、计算机视觉、行为识别;董乐(1980-),女,博士,副教授,主要研究方向为数字媒体、人机交互、机器人与模式识别,E-mail:ledongisi@gmail.com;赵雪专(1986-),男,博士生,主要研究方向模式识别、计算机视觉。

数中有关矩的平移不变性、旋转不变性以及比例不变性等性质,并给出了具有平移、旋转和比例不变性的7个不变矩的定义,而且在计算机上采用两个简单字母的二值图像进行了识别实验。1978年R. Y. Wong进一步给出了在离散状态下各阶矩的计算方法^[6],并利用图像进行了实验,结果表明在比例因子小于2、旋转角度小于45°的情况下,基本保证了M. K. Hu所定义的7个矩的不变性。从此,Hu矩特征得到了广泛的应用,后人也对该特征进行了各种各样的改进。随后,丁明跃等人对于R. Y. Wong提出的用于图像匹配的不变矩法进行分析,给出了在离散情况下平移和旋转不变性的理论证明,同时指出比例因子的不变性在离散情况下是不成立的^[7]。吕洪涛等人在丁明跃理论证明的基础上,分析了比例因子对不变矩的影响方式,给出了两种使矩特征具有比例不变性的方法^[8]。刘进等人提出了一种快速有效的推导不变矩的方法,导出了5个新的不变矩表达式^[9]。近年来,鉴于Hu不变矩的简单有效性,其仍被广泛应用于图形图像领域的各个方面。同时,很多研究者将Hu矩特征进行改进,用于各种检测系统中,例如,印勇等对提取的运动人体轮廓求取不变矩特征,通过模板匹配的方法区分正常与异常行为^[10];G. Y. Zhang等人把Hu不变矩特征应用于机动车安全系统中,实时进行人眼检测,当人眼对路况信息不能给予足够的关注时便予以警报^[11];C-L. Hwang等人将Hu不变矩特征应用于手势识别以实现基于手势的功能控制^[12];H. Mu将不变矩应用于木材的缺陷检测^[13]。

本文对刘进在文献^[9]中提出的推广的Hu不变矩特征进行改进,使其在离散情况下同时具有尺度不变性、旋转不变性与平移不变性。采用改进后的推广的Hu不变矩方法提取运动目标的Hu矩特征,然后结合其速度特征,获得描述运动目标所处行为模式的特征向量,最后采用SVM分类器^[14,15]对特征向量进行分类,以识别其行为模式。该方法结合了运动的空间特征与时间信息,有效地描述了人体的运动模式。推广的Hu矩特征相比于Hu矩特征更好地描述了人体在空间上的一些细节信息,采用二维的速度特征相比一维的速度特征更好地分割了运动人体在垂直方向与水平方向的运动,二者的结合将人体运动描述为一个时空向量,有效地表征了人体的运动模式。实验结果表明,该方法计算效率高,能够实时检测人体行为。此外,该方法能够有效地应用于多类别多实例的行为场景中,实现多类别的多种行为同时检测。

2 识别算法

本文采用的算法框架如图1所示,该算法框架主要由两部分组成:训练与识别。在训练阶段首先处理训练集视频,获得运动人体的二值图像,根据获得的人体二值图像序列,计算Hu矩特征与运动特征向量,然后使用这些特征向量构造SVM分类器。在识别阶段,首先采用与训练阶段相同的方法获得运动人体的Hu矩特征与运动特征向量,然后使用训练阶段构造的SVM分类器对求得的特征向量进行分类,以识别当前人体的运动模式。

本文采用的行为识别方法的具体步骤如下:(1)运动人体提取,采用背景差分法获取运动人体的二值图像;(2)特征提取,将7个Hu不变矩特征推广到十二维,并加以改进,使其成为同时具有比例不变性、旋转不变性与尺度不变性的十一

维特征向量(该向量与运动人体的二维速度特征组合成为十三维的特征向量,以表征人体的运动模式);(3)分类识别,利用行为模板库中生成的特征向量对6种行为构造15个SVM分类器,采用SVM分类器对待检测视频中提取的特征向量进行分类,并根据分类结果判断当前视频中的行为模式。

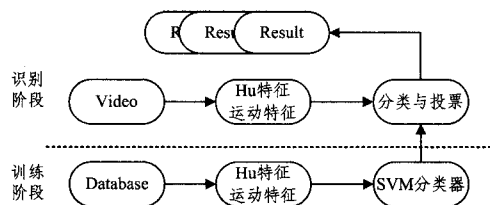


图1 算法框架

2.1 运动人体提取

提取运动前景有多种方法,如背景差分法、帧差分法、光流法等,其各自的计算复杂度与前景提取效果在一定程度上成反比。在摄像头静止的场景中,背景差分法是一种最简单快速的提取运动目标的方法,考虑到智能监控大都应用到摄像头静止的场景中,这里采用背景差分法来获取运动目标,以获得其二值图像。综合考虑计算复杂度与前景提取效果,本文提出一种基于像素灰度区间归类的背景重构方法来构建背景,运用背景差分法来提取运动人体目标。

基于像素灰度区间归类的背景重构方法是在“背景在图像序列中总是最经常被观测到”的假设前提下提出的,该算法共分为4步:首先,划分像素点的灰度平稳区间;其次,分别计算各平稳区间的平均灰度值;然后,将灰度平均值相近的灰度平稳区间归为一类,并统计该类区间出现的次数;最后,将出现次数最多的灰度区间的平均灰度值作为该像素点的背景灰度值。实验中,采用该方法构建背景,然后运用背景差分法来提取运动目标。

现取lena弯腰视频的第32帧与lena挥手1视频的第21帧进行处理,其结果见图2、图3,图2与图3的左图分别是lena弯腰视频的第32帧与lena挥手1视频的第21帧的原图像,从图2的右图中可以清楚地看到白色矩形框内的区域是人弯腰的姿态,而图3右图中的白色矩形标出的区域是人挥手1的姿态,由此看出,可以根据人体的姿态来判断其所处的行为模式。

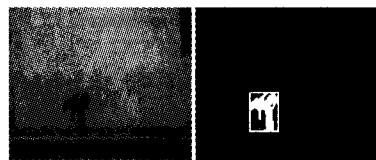


图2 弯腰视频第32帧及其二值图像

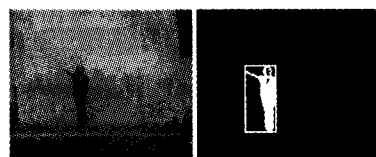


图3 挥手1视频第21帧及其二值图像

2.2 特征向量提取

首先对十二个Hu矩特征 φ_k 进行一些变换,使其成为在离散情况下同时具有旋转、平移、比例不变性的十一个Hu矩特征 φ_k^* ,将此特征与运动人体的速度特征 (V_k, V_v) 结合,构

造一个十三维的特征向量 \$(V_h, V_v, \varphi_1^*, \dots, \varphi_{11}^*)\$, 对提取的运动人体的二值图像求这个十三维特征向量, 并用该向量来描述人体的行为模式。

2.2.1 Hu 矩特征

在离散状态下, 二维函数 \$f(m, n)\$ 的 \$p+q\$ 阶矩定义如下:

$$m_{pq} = \sum_{m=1}^M \sum_{n=1}^N m^p \cdot n^q \cdot f(m, n) \quad (1)$$

式中, \$p, q=0, 1, 2, \dots\$。

函数 \$f(m, n)\$ 的 \$p+q\$ 阶中心矩形式如下:

$$\mu_{pq} = \sum_{m=1}^M \sum_{n=1}^N (m-\bar{m})^p \cdot (n-\bar{n})^q \cdot f(m, n) \quad (2)$$

式中, \$\bar{m}=m_{10}/m_{00}, \bar{n}=m_{01}/m_{00}\$。

函数 \$f(m, n)\$ 的 \$p+q\$ 阶归一化中心矩定义为:

$$\eta_{pq} = \mu_{pq} / \mu_{00} \quad (3)$$

这里, \$r=(p+q)/2+1\$。

根据上面的定义, M. K. Hu 给出了 7 个不变矩公式如下:

$$\varphi_1 = \eta_{20} + \eta_{02} \quad (4)$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (5)$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{33})^2 \quad (6)$$

$$\varphi_4 = (\eta_{30} + 3\eta_{12})^2 + (\eta_{21} + \eta_{33})^2 \quad (7)$$

$$\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})[(3\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] \quad (8)$$

$$\varphi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] \quad (9)$$

$$\varphi_7 = (\eta_{20} - \eta_{02})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] \quad (10)$$

此外, 还需要一些高阶矩特征来表示人体的区域形状的一些细节特征, 根据刘进对 Hu 矩特征的推广, 及二阶矩、三阶矩、四阶矩的定义又导出了 5 个不变矩公式如下:

$$\varphi_8 = 2\{\eta_{11}[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] - (\eta_{20} - \eta_{02})(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{33})\} \quad (11)$$

$$\varphi_9 = [(\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})](\eta_{20} - \eta_{02}) + 2\eta_{11}[(3\eta_{21} - \eta_{33})(\eta_{30} + \eta_{12}) - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{33})] \quad (12)$$

$$\varphi_{10} = [(3\eta_{21} - \eta_{33})(\eta_{30} + \eta_{12}) - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{33})](\eta_{20} - \eta_{02}) - 2\eta_{11}[(\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})] \quad (13)$$

$$\varphi_{11} = (\eta_{40} + \eta_{40} - 6\eta_{22})^2 + 16(\eta_{31} - \eta_{33})^2 \quad (14)$$

$$\varphi_{12} = (\eta_{40} + \eta_{40} - 6\eta_{22})[(\eta_{20} - \eta_{02})^2 - 4\eta_{11}^2] + 16\eta_{11}(\eta_{31} - \eta_{33})(\eta_{20} - \eta_{02}) \quad (15)$$

根据丁明跃对 7 个不变矩的不变性的理论证明, 作者同样证明了 \$\varphi_8 - \varphi_{12}\$ 在离散情况下具有平移、旋转不变性, 同时也求出了这 12 个不变矩在发生比例因子失真情况下的变化, \$\varphi_i'\$ 为失真后求得的不变矩, 其结果如下:

$$\varphi_i' = \begin{cases} \rho^2 \varphi_i, (i=1) \\ \rho^4 \varphi_i, (i=2) \\ \rho^6 \varphi_i, (i=3, 4) \\ \rho^8 \varphi_i, (i=6, 8, 9, 10, 11, 12) \\ \rho^{12} \varphi_i, (i=5, 7) \end{cases} \quad (16)$$

为使这些不变矩特征在离散情况下具有比例不变性, 这

里重新定义一组不变矩特征 \$\varphi_i^*\$ (\$s=1, 2, \dots, 11\$)。

$$\varphi_i^* = \begin{cases} \varphi_{s+1}/\varphi_1^2, (s=1) \\ \varphi_{s+1}/\varphi_1^3, (s=2, 3) \\ \varphi_{s+1}/\varphi_1^4, (s=5, 7, 8, 9, 10, 11) \\ \varphi_{s+1}/\varphi_1^6, (s=4, 6) \end{cases} \quad (17)$$

2.2.2 速度特征

从图 2 与图 3 可以看出, 在发生弯腰这个动作时, 人体轮廓的质心在垂直方向上有移动, 在水平方向上变动较小, 但这个动作恰好相反, 而跳这个动作在水平方向与垂直方向均有移动。考虑到不同的行为动作的这些特征, 采用动作水平方向速度、垂直方向速度来表示一种动作的速度特征。

这里用图像矩来求解运动目标的质心, 并将它作为目标的坐标中心, 计算各质心的位移, 从而求得人体运动速度的两个特征分量。

图像矩的定义与式(1)相同, 为了更清楚地说明速度特征的求法, 这里给出其新的定义:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y) \quad (18)$$

式中, \$f(x, y)\$ 表示像素点 \$(x, y)\$ 的像素值, \$x^p\$ 表示 \$x\$ 的 \$p\$ 次矩, \$y^q\$ 表示 \$y\$ 的 \$q\$ 次矩。\$M\$ 为图像区域的列数, \$N\$ 为图像区域的行数。

用零阶矩表示二值图像区域的面积, 即:

$$m_{00} = \sum_{x=1}^M \sum_{y=1}^N f(x, y) \quad (19)$$

则运动目标的质心位置为:

$$\begin{cases} \bar{x} = \frac{m_{10}}{m_{00}} \\ \bar{y} = \frac{m_{01}}{m_{00}} \end{cases} \quad (20)$$

第 \$K\$ 帧的质心坐标位置记为 \$(x_k, y_k)\$, 根据下面的式子可以求出每隔 \$N\$ 帧在水平与垂直方向的位移:

$$S_h = |X_{K+N} - X_K| \quad (21)$$

$$S_v = |Y_{K+N} - Y_K| \quad (22)$$

由于采用的视频帧率为 25f/s, 那么 \$N\$ 帧的时间间隔为 \$N/25s\$, 此时的速度为:

$$V_h = 25S_h/N \quad (23)$$

$$V_v = 25S_v/N \quad (24)$$

2.3 SVM 分类识别

将区域形状特征与运动特征相结合, 得到了一个十三维的特征向量。下面利用训练集中 6 种行为的特征向量构造 SVM 分类器, 这里采用了“一对一”的径向基核函数支持向量机, 即在 \$M\$ 个类中任意两个类之间均构造一个分类器, 共 \$M(M-1)/2\$ 个分类器。然后采用投票的方法来判断所测试行为所处的行为模式, 即每个分类器根据其分类结果给相应的类别投票, 在所有分类器的投票结果中, 若类 \$c_i\$ 的得票最多, 则类 \$c_i\$ 为最终分类。

训练每个分类器所用的样本由相应的两类训练样本构成, 对第 \$i\$ 类与第 \$j\$ 类构成的分类器, 假设 \$n_i, n_j\$ 分别表示第 \$i\$ 类与第 \$j\$ 类的样本数, 第 \$i\$ 类为正类, 第 \$j\$ 类为负类, 则训练样本构成的训练集可记为:

$$S_{ij} = \{(x_{ij,k}, c_{ij,k} | k=1, 2, \dots, n_i + n_j, c_{ij,k} \in \{+1, -1\})\}$$

式中, \$x_{ij,k}\$ 表示该样本来自第 \$i\$ 类或第 \$j\$ 类, 在训练集 \$S_{ij}\$ 中编号为 \$k, c_{ij,k}\$ 表示样本 \$x_{ij,k}\$ 在构造第 \$i\$ 类与第 \$j\$ 类样本之间的

分类器时属于正类或负类。

构造每两个类别之间的分类器时都采用径向基核函数的支持向量机,得到决策函数:

$$f_{ij} = \text{sgn} \left(\sum_{k=1}^{n_i+n_j} \alpha_{ij,k}^* c_{ij,k} K(x, x_{ij,k} + b_{ij}^*) \right) \quad (25)$$

式中, $\alpha_{ij,k}^*$ 与 b_{ij}^* 是最优化问题的解,若 $f_{ij} = 1$,则该样本属于第 i 类,若 $f_{ij} = -1$,则该样本属于第 j 类,当 $f_{ij} = 0$ 时,该样本的类别无法判断,这里采用投票的方法,当一个样本属于某一个类别时,将该类别投票数加 1,在每一个分类器对当前样本都进行分类之后,将投票数最高的那个类别作为该样本的最终类别。

3 实验

3.1 数据集与实验步骤

将以色列 Weizmann 科学院的人体运动数据库作为本文实验的第一个数据集,该数据库是在摄像头静止的场景下拍摄的,包含 10 组行为,每一组行为都由 9 个人完成。该数据库包含的 10 种行为为走、跑、跳(双脚并拢向前跳)、侧向跑,原地双脚蹦跳、原地双手挥舞蹦跳、单腿跳、挥手 1(原地单手挥舞)、挥手 2(原地双手挥舞)、弯腰。在本文的实验 1 中选取了 6 种行为进行训练与识别,即选择了走、跑、跳、弯腰、挥手 1、挥手 2 的部分视频数据作为训练集,选择这 6 种行为的其他部分视频数据作为测试集,进行算法准确率验证实验以及与其他方法(HU 矩、Zernike 矩、R 变换)的识别率对比实验。

实验 2 选用 Kect Gesture Data Set 作为数据集来验证文中所提算法的有效性。Kect Gesture 数据集是军事信号中部分手势的数据集,共 14 种手势,该数据集由 3 个人分别来完成这 14 种动作,并且每种动作连续执行 3 次,共 126 个视频序列。本文选取其中具有代表性的 6 种手势进行实验,包括 G1(turn left)、G2(turn right)、G5(fl原因)、G7(stop right)、G8(stop both)、G10(start)。

本文的第 3 个数据集是作者自行拍摄的一组多类别多实例的运动视频,用以测试该算法在多类别多实例的行为场景中的应用。该视频同样在摄像头静止的场景下拍摄,同一视频中有多个类别的行为,同时有一个类别的行为的多个实例。实验 3 将拍摄的一组单目标单行为的视频作为训练集,在多类别多实例行为的视频上进行测试,即将基于区域形状与运动特征的行为识别应用于多类别多实例的场景中。

本文的实验步骤如下:首先,对训练集中的视频提取其二值图像,利用文中的方法提取这些行为的十三维特征向量,并构造各种行为两两之间的 SVM 分类器;然后,以同样的方法提取测试集中各种行为的十三维特征向量,用训练好的每个 SVM 分类器对这些特征向量进行分类投票,再根据分类投票的结果来判断当前的测试模式属于哪种运动模式。

3.2 实验结果与分析

根据文中的实验步骤,按照实验 1 的实验设置,得到了弯腰、跳、跑、走、挥手 1、挥手 2 的识别结果。其效果见图 4,从图中可以看出,用文中的方法能够有效地识别训练好的 6 种行为,识别出当前人体所处的行为模式之后,在运动目标的中心位置用红色字体标出运动目标当前所处的行为模式。当然,该方法也有检测错误的情况,如图 5 左图错误地将挥手 2

识别为挥手 1,右图错误地将走识别为跳。

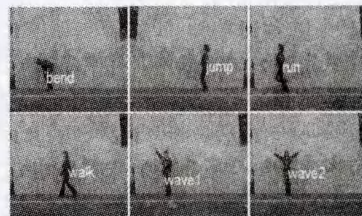


图 4 6 种行为识别结果

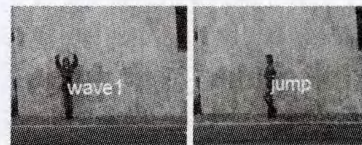


图 5 行为识别错误

用文中的方法对实验 1 中的测试集进行测试,识别率见表 1,从表中可以看出对跑、跳、弯腰、挥手 1 这 4 种行为的识别率较高,而对走、挥手 2 的识别率较低。这是由于行为的分类原则一方面是基于空间信息,即人体行为的姿态,另一方面又考虑到了时间信息,即人体的运动特征,因此较好地地区分了跑与走这两种运动,但是该方法也有其局限性,当两种运动姿态较为相近、速度也相差不多时,容易误分,从而出现了将走分类到跳,将挥手 2 分类到挥手 1 这种现象。

表 1 实验 1 中 6 种行为在测试集中的识别率

识别率	走	跑	跳	弯腰	挥手 1	挥手 2
走	80%	0	20%	0	0	0
跑	3.33%	93.33%	3.33%	0	0	0
跳	3.08%	0	96.92%	0	0	0
弯腰	0	0	0	95%	0	5%
挥手 1	1.25%	0	0	0	97.50%	0
挥手 2	0	0	0	0	10%	90%

在 Weizmann 的数据集下,我们也做了不同方法的比较实验。仍然选取上述的 6 种行为,9 个人中选取 1 个人(实验中选 lena)的行为作为训练样本,其他 8 个人的行为是测试样本。经过实验,得到的实验结果见表 2。

表 2 不同方法在 6 种行为中的识别结果

方法 \ 行为	走	跑	跳	弯腰	挥手 1	挥手 2
Hu 矩	正确	0	8	1	7	8
	错误	8	0	7	1	0
Zernike 矩	正确	5	7	4	5	4
	错误	3	1	4	3	4
R 变换	正确	7	7	7	3	6
	错误	1	1	1	5	2
本文	正确	8	8	5	7	8
	错误	0	0	3	1	0

根据表 2 中的实验结果,可以得到 4 种方法在相同的训练集和同样的测试集下的识别率。即:Hu 矩的识别率为 64.6%;Zernike 矩的识别率为 66.7%;R 变换的识别率为 77.1%;本文方法的识别率为 91.7%。通过对比表 1 与表 2 中本文所提算法的识别效果可以看出,本算法对模板的选择有一定的依赖性,选择不同的行为姿态模板会得出不同的识别效果。

实验 2 中 Kect Gesture 数据集是军事信号中部分手势的数据集,共 14 种手势。本实验采用 6 种手势来验证算法的有效性,包括 G1(turn left)、G2(turn right)、G5(fl原因)、G7(stop

right)、G8(stop both)、G10(start)。每种手势共有 9 个视频序列,这里选取其中一个视频序列作为训练样本,其他的 8 个序列作为测试样本集。6 种军事手势的部分识别结果见图 6,表 3 展示了对 Kect Gesture 数据集的识别结果。



图 6 Kect Gesture 部分行为识别结果

表 3 实验 2 中 14 种行为识别结果

	G1	G2	G5	G7	G8	G10
G1	6	0	0	0	0	0
G2	0	6	0	0	0	0
G5	0	0	8	0	0	0
G7	0	2	0	8	0	0
G8	0	0	0	0	8	0
G10	2	0	0	0	0	8

相比于其他的基于模板匹配的行为识别方法,本文的方法简单、计算量少、速度快。实验 1 训练了走、跑、跳、弯腰、挥手 1、挥手 2 这 6 种行为,并构造了 15 个分类器,在处理器为 Pentium(R) Dual-Core CPU, 2.93GHz、内存为 4GB 的机器上,用文中的方法对测试集中的行为进行了识别,速度达到 24.12~25.26 帧/秒,因此该方法可以用于智能监控中的实时监测。

本文提到的方法可以很容易地应用到多类别多实例的行为场景中,只需将文中的方法应用于运动人体提取阶段来提取的人体运动区域,即可判断该人所处的行为模式。在场景中的多个运动目标不存在遮挡与重叠的情况下,该方法能够轻松地检测出运动目标所处的行为模式。实验 3 的结果如图 7 所示,从检测结果看来,文中的方法能够有效检测同时发生的多类型的多个行为模式。



图 7 多行为多实例行为检测

结束语 本文实现了一种新的基于 Hu 矩特征与运动特征的行为识别方法,该方法运用了改进后的推广的 Hu 矩在离散情况下所具有的尺度、旋转、平移不变性特征,有效地描述了行为的空间信息,另外,将速度描述为二维特征,分开考虑垂直方向与水平方向的运动,充分地描述了运动的速度信息,二者的组合有效地识别了走、跑、跳、弯腰、挥手 1、挥手 2 这 6 种行为。此外,该方法能够有效地应用于多类别多实例

的行为识别中。

但该方法也有其局限性,当人的两种行为的姿态相近、速度也相差不大时容易出现误判的情况。此外,该方法仅适用于摄像头静止的场景中,考虑到智能监控中的摄像头大都是静止的,该方法仍具有广泛的应用场景,当然,也可以在运动目标提取部分更换其他的方法,以使该算法应用于运动场景中。

参考文献

- [1] Bobick A F, Davis J W. The recognition of human movement using temporal templates[C]//IEEE International Conference on Computer Vision, 2009
- [2] Shechtman E, Irani M. Space-time behavior based correlation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2009
- [3] Rodriguez M D, Ahmed J, Shah M. Action mach a spatio-temporal maximum average correlation height filter for action recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2008
- [4] Hu Yu-xiao, Cao Liang-liang, Lv Feng-jun, et al. Action detection in complex scenes with spatial and temporal ambiguities[C]//IEEE International Conference on Computer Vision, 2009
- [5] Hu M K. Visual pattern recognition by moment invariants[C]//IRE Transaction Information theory, 1962;179-182
- [6] Wong Y R. Scene matching with invariant moments[J]. Computer Graphics and Image Processing, 1978, 8:16-24
- [7] 丁明跃,常金玲,彭嘉雄. 不变矩算法研究[J]. 数据采集与处理, 1992, 7(1):1-9
- [8] 吕洪涛,周继成. 离散状态下的不变矩算法研究[J]. 数据采集与处理, 1993, 8(2):151-155
- [9] 刘进,张天序. 图像不变矩的推广[J]. 计算机学报, 2004, 27(5): 668-674
- [10] 印勇,张毅,刘丹平. 基于改进 Hu 矩的异常行为识别[J]. 计算机技术与发展, 2009, 19(9):90-92
- [11] Zhang Guang-yuan, Cheng Bo, Feng Rui-jia, et al. Real-time driver eye detection method using Support Vector Machine with Hu invariant moments[J]. IEEE International Conference on Machine Learning and Cybernetics, 2008(5):2999-3004
- [12] Hwang C-L, Lee H-W. The command control by hand gesture with Hu and contour sequence moments and probability neural network[C]//IEEE International Conference on Systems, Man, and Cybernetics, 2011;2056-2061
- [13] Mu Hong-bo, Qi Da-wei. Pattern Recognition of Wood Defects Types Based on Hu Invariant Moments[C]//2nd International Congress on Image and Signal Processing, 2009;1-5
- [14] Safari M, Harandi M T. A SVM-based Method for Face Recognition Using a Wavelet-PCA Representation of Faces[C]//IEEE International Conference on Image Processing, 2004;853-856
- [15] Schuld C, Laptev I. Recognizing Human Actions: A local SVM Approach[C]//International Conference on Pattern Recognition, 2004