

基于运动矢量分析的动作识别技术

The Action Recognition Technology Based on Motion Vector Analysis

梁家海, 黄雪燕

LIANG Jia-hai, HUANG Xue-yan

(钦州学院数学与计算机科学系, 广西钦州 535000)

(Department of Mathematics and Computer Science, Qinzhou University, Qinzhou, Guangxi, 535000, China)

摘要:改进人体行为识别传统模板匹配方法的算法,提出一种基于运动矢量分析动作识别技术,并对其识别效果进行实验验证。该技术把标准人体动作的百分比运动矢量作为模板,将待识别动作的百分比运动矢量与已知的模板进行对比,从而得到动作识别结果。该技术可以正确识别摆手、点头和摇头动作,动作重复3次的识别率可以达到95%以上。该技术进行实时动作识别具有效果好、算法简单、识别速度快、抗干扰性强等优点。

关键词:动作 识别 矢量 分析

中图分类号:TP391.4 **文献标识码:**A **文章编号:**1002-7378(2008)03-0256-04

Abstract: A technology of the motion-recognition based on motion vector analysis was proposed by improving the algorithms of human behavior recognition of traditional methods of template matching. The recognition effect of this technology was verified by experiments. The technology uses the percentage of vector of standards body motion as a template. Recognition action will be to identify by comparing the percentage of vector and the known template. The technology can correctly recognize shaking head, shaking head and nod. If recognition action can be repeated three times, the recognition rate of this action will be more than 95 percent. During conducting real time action, the advantages of this technology is effective identification, simple algorithm, fast recognition and strong anti-jamming.

Key words: action recognition, vector, analysis

随着计算机应用的迅速发展,基于视频序列的人体行为识别逐渐成为计算机视觉领域研究的热点,在人机接口系统、智能家居应用、智能监控、运动分析等方面得到了广泛的应用^[1]。人体视频行为识别存在计算复杂度和抗干扰性的矛盾。常见的视频人体行为识别的方法有两种:一种是模板匹配方法^[2,3],这种方法将图像序列转换为一组静态形状模式,然后在识别过程中和预先存储的行为标本进行比较得到识别结果,如文献^[2~4]利用二维网格的

特征进行人的动作识别。模板匹配技术虽然计算复杂度低、实现简单,但是存在对噪声比较敏感、识别依赖于模板的选择、应用于非特定人的复杂场合比较困难等缺点。另一种是状态空间方法^[5~13],这种方法将每个静态的姿势作为一种状态,这些递增状态之间通过某种概率联系起来,任何运动序列可以看作这些姿势不同状态之间的一次遍历过程,在这些遍历期间计算联合概率,其最大值被选择作为分类行为的标准。状态空间法受噪声影响较少,能更好刻画动作本质特征,应用范围较广,但是涉及迭代运算,计算相对较为复杂。

本文改进传统模板匹配方法的算法,提出一种基于运动矢量分析的动作识别技术,以求在保持传统模板匹配方法算法简单的基础上,能够有效地提

收稿日期:2008-03-19

修回日期:2008-06-12

作者简介:梁家海(1968-),男,讲师,硕士,主要从事计算机测控技术和模式识别研究工作。

高动作识别的抗干扰能力。

1 运动检测及运动轨迹的提取

运动检测采用目前常用的相邻帧相减法进行,通过将相邻两帧图像进行相减,滤去相同部份(背景),差值部份即为运动部份^[1]。为了便于分析,通过设置阈值将差值图进行二值化处理^[5]得到二值图。

若一个动作有 N 帧图像,由于相邻两帧可以相差得到一幅二值图,那么共得到 $(N - 1)$ 幅二值图;为了提取动作的运动轨迹,我们将二值图的质心作为运动轨迹的关键点,将所有的质心按顺序连接起来构成动作的运动轨迹。由于二值图是不规则的,所以将质心定义为二值图中所在白色像素的中心,如图 1 所示。质心通过如下数学模型^[6]求得。

设 $m \times n$ 维二值图像 $I(x, y)$, 其中目标部分为 A , 背景部分为 B , 即

$$I(x, y) = \begin{cases} 1, (x, y) \in A, \\ 0, (x, y) \in B, \end{cases} \quad (1)$$

目标的质心 (x_0, y_0) 定义为:

$$x_0 = \frac{\sum_{(x,y) \in A} xI(x, y)}{\sum_{(x,y) \in A} I(x, y)}, y_0 = \frac{\sum_{(x,y) \in A} yI(x, y)}{\sum_{(x,y) \in A} I(x, y)}. \quad (2)$$

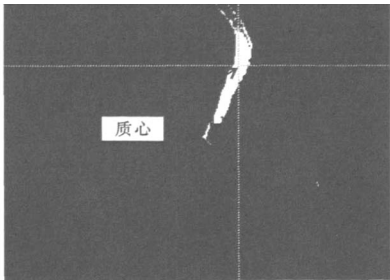


图1 摆头运动的二值图及其质心

质心的连线即为动作的运行轨迹,图 2(a)、(b) 分别为左右摆头、摇头动作的质心运行轨迹,相同的动作其质心运行轨迹相似。

2 基于运动矢量分析的动作识别

2.1 运动矢量统计

由于动作是周期性的,所以动作的运行轨迹是闭合的,因此运动轨迹描述为

$$L' = x_0 \vec{0}^\circ + x_1 \vec{1}^\circ + x_2 \vec{2}^\circ + \dots + x_n \vec{n}^\circ, \quad (3)$$

其中 x_0, x_1, x_2, x_n 为方向为 0 度、 1 度、 2 度、 n 度上的位移量, $+$ 表示线段连接。

表达式(3)同时也表达了动作幅度,但是在动作的识别中,往往不需考虑动作的幅度,即不管是

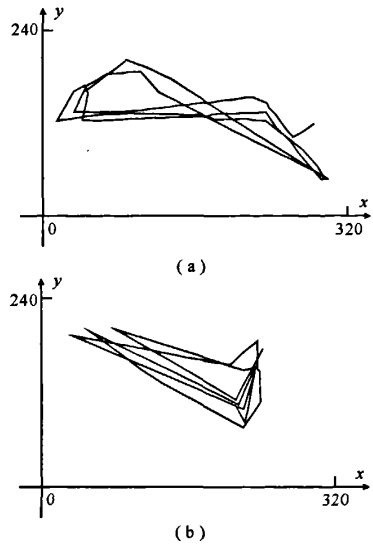


图 2 两种动作的质心运动轨迹
(a) 左右摆头; (b) 摇头

“轻轻地点头”还是“重重地点头”都认为是点头动作。为了提高动作的识别效果,提高抗干扰性,将表达式(3)进行转换成没有动作幅度成分,即统计 360 度方向范围内每个角度区间内的位移量占总位移量的百分比。当某个动作重复的次数越多,该动作的质心运动轨迹的矢量百分比越接近实际,识别的效果越好。取得矢量百分比的过程如下,设 M_n, M_{n+1} 为任意相邻二值图的质心, L_n 为两点间的位移(见图 3),则这两个质点的运动方向 α_n 和位移 L_n 为:

$$\alpha_n = \arctg\left(\frac{dy}{dx}\right), (0 \leq \alpha_n \leq 360);$$

$$L_n = \sqrt{(x_{n+1} - x_n)^2 + (y_{n+1} - y_n)^2}.$$

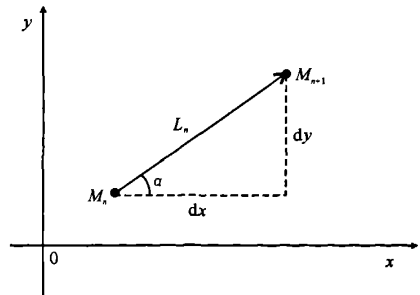


图 3 质心位移示意

在实时视频图像采集时,定时进行图像的采集,每采集到一帧图像都与前一帧进行差分得到二值图,从而求出质心,并以前一帧二值图的质心进行上述运算,求出两个质点的运动方向 α_n 和位移 L_n 。然后将实时产生的每组 (α_n, L_n) 进行统计,求出每个方向上位移的总和并记录在 $360/N$ 个元素的一维

数组 M 中, N 表示为方向分组系数, 如当 $N = 4$ 时, 则将 4 作为一组进行统计, 则 M 数组共有 $360/4 = 90$ 个元素, 如图 4 所示。

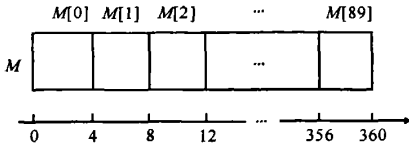


图 4 运动方向与数组对应关系

将数组 M 进行标准化, 生成滤去动作幅度信息的矢量百分比记录在数组 p 中, 其数学模型为

$$p[k] = \frac{M[k]}{\sum_{i=0}^{n-1} M[i]}, (k \in 0 \sim n - 1),$$

式中 $p[k]$ 、 $M[k]$ 为 p 和 M 中的第 k 个数组的元素。显然, 数组 p 可拟合为以方向角度为自变量的曲线 $y = F(\alpha)$ 。

2.2 动作识别

动作识别的过程就是将当作运作的矢量百分比与模板库中的矢量百分比进行差分, 差分的值最小且小于某一阈值时, 可以认为该动作是此时模板对应的动作, 差分的数学模型为

$$D \int_0^{360} F(\alpha) - T(\alpha) d\alpha, \tag{4}$$

(4) 式中, $F(\alpha)$ 表示当前动作百分比函数, $T(\alpha)$ 是模板库中某动作的百分比函数。

3 实例分析

动作识别程序采用 C# 设计; 摄像头为普通上网视频聊天用的 Camera 301x 摄像头, 成像效果一般, 视频采样周期为 0.1s, 即 10 帧/秒; 实验用计算机配置为 Ceron 1.3GHz/256M。

为简单起见, 实验以左右摇头、点头、摇头进行识别为例。首先通过训练建立左右摇头、点头、摇头这 3 个动作的矢量百分比模板; 然后进行动作的实时采集视频序列图像, 统计其矢量百分比, 并与模板进行比较进行动作识别。实验中, 阈值取 0.2, 差分数值最小且与次小差分数值大于 0.2 者即为此时模板所对应的动作。

图 5、图 6、图 7 所示为某次识别摇头、点头、摇头动作时的效果, 图中数据为当前动作与模板的差分结果。

实验表明, 当前动作与模板中相同的动作差分的值较小, 而与其不同的动作差值较大, 说明动作识别的稳定性较好。同时实验还表明, 识别动作重复的

次数越多识别越准确, 如果重复 3 次, 识别率可以达到 95% 以上。

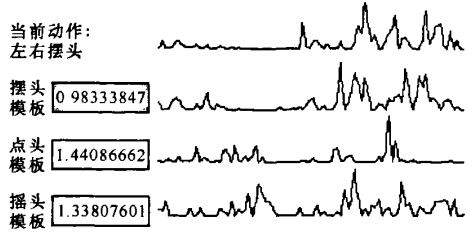


图 5 摇头动作识别

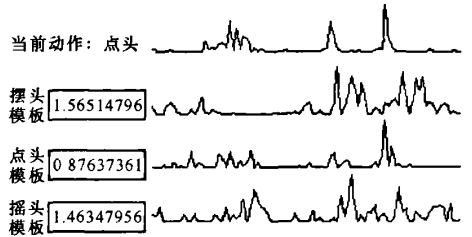


图 6 点头动作识别

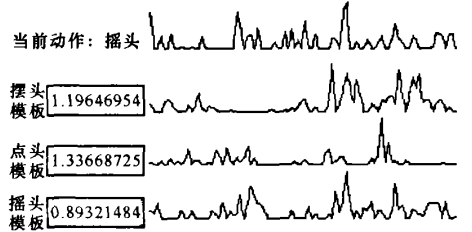


图 7 摇头动作识别

4 结束语

本文提出一种基于运动矢量分析的动作识别技术, 它通过将标准人体动作的运动轨迹抽象成运动的矢量, 并对运动矢量标准化为百分比运动矢量, 将每个标准动作的百分比运动矢量作为模板。识别过程中将待识别的动作点抽象出其百分比运动矢量, 然后与已知的模板进行对比, 加权差值最小的即为此时模板所对应的动作。实例研究表明, 使用基于运动矢量分析的动作识别技术来描述人体的动作是可行的, 该技术进行实时动作识别具有效果好、算法简单、识别速度快、抗干扰性强等优点。但是由于该方法依赖于已知的动作模板, 不具备识别未知动作的能力, 所以较适合应用在人机智能接口方面。

参考文献:

[1] 王亮, 胡卫明, 谭铁牛. 人运动视觉分析综述[J]. 计算机学报, 2002, 25(3): 225-237.
 [2] Bobick A, Davis J. Real-time recognition of activity using temporal templates: proc IEEE Workshop on Application of computer Vision, Sarasota[C]. Florida:

- [s. n.], 1996:39-42.
- [3] Davis J, Bobick A. The representation and recognition of action using temporal templates[R]. MIT Media lab, Perceptual Computing Group: Technical Report 402, 1997.
- [4] Polans R, Nelson R. Low level recognition of human motion; proc IEEE Workshop on Motion of Non-Rigid and Articulated Objects[C]. Austin: TX, 1994:77-82.
- [5] 赵国英, 李振波, 邓宇, 等. 基于检索的人体运动识别的模拟[J]. 计算机研究与发展, 2006, 43(2):368-374.
- [6] 雷蕾, 李言俊, 张科. 图像目标质心快速搜索算法[J]. 红外技术, 2007, 29(9):548-551.
- [7] Cui Y, Weng J. Hand segmentation using learning based prediction and verification for hand sign recognition; proc IEEE Conference on computer Vision and pattern Recognition[C]. Puerto Rico; IEEE Computer Society Press, 1997:88-93.
- [8] Rabinier L. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceeding of the IEEE, 1989, 77(2):257-285.
- [9] Starner T, Pentland A. Real-time American sign language recognition from video using hidden Markov models; proc International Symposium on Computer Vision, Coral Gables [C]. Florida: IEEE Computer Society Press, 1995:265-270.
- [10] Yamato J, Ohya J, Ishii K. Recognizing human action in time-sequential image using hidden Markov model; proc IEEE Conference on Computer Vision and Pattern Recognition[C]. Champaign: IEEE Computer Society Press, 1992:379-385.
- [11] Brand M, Oliver N, Pentland A. Coupled hidden Markov models for complex action recognition; proc IEEE Conference Computer Vision and Pattern Recognition[C]. Puerto Rico; IEEE Computer Society Press, 1997:994-999.
- [12] Bregler C. Learning and recognizing human dynamics in video sequences; proc IEEE Conference on Computer Vision and Pattern Recognition[C]. Puerto Rico; IEEE Computer Society Press, 1997:568-574.
- [13] Campbell L, Bobick A. Recognition of human body motion using phase space constraints; proc International Conference on Computer Vision[C]. Cambridge; IEEE Computer Society Press, 1995:624-630.

(责任编辑: 韦廷宗)

(上接第255页)

扣计算, 其它情况须税务人员核准审批后才可以进行抵扣。收购发票进项税的计算方法为: 收购发票进项税 = 收购发票金额 \times 13%。

运输发票审核比对子系统。当运输发票所对应的收购发票能够进行进项税抵扣, 那么, 运输发票也可以进行进项税抵扣。其它情况则不允许抵扣。运输发票进项税计算方法为: 运输发票进项税 = 运输发票金额 \times 7%。

增值税发票监开子系统。对增值税专用发票的票号、发售及回收情况进行管理。跟踪每一张发票的使用及回收情况, 对超期未回收的发票进行预警。

销售合同备案子系统。对企业的产品销售合同、销售发票进行录入、查询与监控, 当销售价格出现异常时, 自动报警; 另一方面, 计算企业的销项税, 以曲线图的形式监控销项税的情况。销项税额 = 产品销售收入 \times 17%。

库存商品监控子系统。实时监控企业库存、产出及销售, 预测库存税收情况, 并计算每天的出糖率。如果出糖率过低, 则表明有可能是出糖口数据异常或进厂糖料异常。通过出糖率的监控可以及时发现糖厂生产的异常情况, 避免税收流失。

综合数据分析子系统。对相关的监控指标, 通过

图表、报表的方式, 及时给国税人员提供相应的信息支持。

3 结束语

本文提出的制糖企业税控信息系统, 通过从“购、产、存、销”4个环节上进行监控, 从源头上杜绝了信息的失真, 并且可以将采集的信息进行加工利用, 可以有效的监控制糖企业的税收情况。该系统已开发成功, 并在广西某市国税局运行, 成功解决了该市对制糖企业税收征管的三大难题, 效益显著, 具有良好的推广价值。

参考文献:

- [1] 吕欣, 贺亮军. 全区蔗糖业创历史新高[N]. 广西日报, 2008-01-26(2).
- [2] 邱伟年, 王政. 税控装置的国际经验及在我国推广应用的借鉴[J]. 企业经济, 2006(8):173-176.
- [3] 赵磊, 余建桥, 金良锋, 等. 福利企业远程指纹税控系统的设计与实现[J]. 计算机系统应用, 2007(11):1-5.
- [4] 王文清, 田柏林, 于明. 矿山企业产量税控系统的研发和应用[J]. 煤炭科学技术, 2007(10):61-63.

(责任编辑: 韦廷宗)