

文章编号: 1007-130X(2009)05-0056-03

基于感知评价的三维信号识别的一种在线算法^{*}

An Online Algorithm for 3D Signal Recognition Based on Ethesis Evaluation

龚元浩, 杨晨晖

GONG Yuan-hao, YANG Chen-hui

(厦门大学计算机科学系, 福建 厦门 361005)

(Department of Computer Science, Xiamen University, Xiamen 361005, China)

摘要:目前,各种传感器广泛应用于各种智能系统中,但对这些传感器信号的识别大多以传统的模式识别方法来完成,而这些方法一般都需要训练样本集进行学习或者需要一些先验知识。本文提出了一种新的基于感知评价的方法。该方法通过对加速传感器的三维信号进行提取、分割和识别来完成动作的识别。在信号识别阶段,我们把曲线转为一种加权图结构,我们称之为简单线性图。通过这个结构,构造了线性简单图之间相似性的一个评价函数,并用来对两个图结构进行相似性评价,最终完成动作的识别。该方法使用了三维信号的空间拓扑结构信息和极小样本,而不依赖于先验知识和大量训练样本。而且该方法是一个在线算法,能够满足很多应用的实时性要求。最后,给出了该方法的实验结果。

Abstract: Nowadays, there are so many kinds of sensors which have been widely used in intelligent systems. The signals of these sensors are recognized by the traditional pattern recognition methods, which need huge training sample sets or experts' knowledge. In the article, we propose a new method which is mostly based on ethesis evaluation. It obtains original signals, does segmentation and recognition of the signals and then recognizes the action what the signal means. In the stage of signal recognition, it uses comparison and similarity of two curves. We turn the curves to weight graphs which we call simple linear graphs. Using this structure, we construct an evaluation function for comparison. Our method uses the space topological information and very little samples instead of huge training sets or experts' knowledge. Moreover, it is an online algorithm, which benefits many real-time applications. Finally, we present the results of our experiment.

关键词: 模式识别; 动作识别; 相似性; 在线算法; 简单线性图

Key words: pattern recognition; action recognition; similarity; online algorithm; simple linear graph

中图分类号: TP391.4

文献标识码: A

1 引言

目前,各种传感器广泛应用于多种智能系统中。例如,车载系统中的温度传感器和加速传感器、手持设备上的重力加速传感器等都在智能系统中起到了重要的作用^[1]。对这些传感器采集到的信号进行识别是一项非常重要的工作。传统的方法对这些信号的识别需要进行学习,发掘特征向量和待匹配模板,然后从待测信号中分离出特征向量,利用聚类的方法加以分类^[2]。这些方法的关键在于特征向量和训练集的选取,而这两者如何选取却难于把握。本文提出了一种新的方法,不依赖于特征向量和训练集,而是基于信号表

示动作本身的空间结构信息。该方法克服了传统方法中特征向量和训练集难于选择的问题,集中分析动作信号本身的结构和极少量样本,这对于动作识别有重要意义。例如,在游戏控制器上集成的加速传感器可以通过无线网络将信号传输给主机,但由于不同年龄的玩家可能对同样动作的表达方式不一致,如果采用基于训练集的方法,不但需要大量的训练数据,还需要大量的计算来寻找特征向量。而本文提出的方法,基于图结构的动作信号表达能够很好地解决这类问题。而且本文提出的方法是一种在线算法,运行效率高,能够满足很多实时性要求高的应用需求。

形状之间的距离如何度量一直是一个难点。文献[3]从图结构的角度介绍了一种旋转距离度量。在结构化模式

* 收稿日期: 2008-02-26; 修订日期: 2008-06-02

基金项目: 国家自然科学基金资助项目(40627001)

作者简介: 龚元浩(1984-),男,硕士生,工程师,IEEE会员,研究方向为计算机视觉;杨晨晖,博士,教授,研究方向为计算机视觉。

通讯地址: 361005 福建省厦门市厦门大学计算机科学系; E-mail: gongyuanhao@ieee.org

Address: Department of Computer Science, Xiamen University, Xiamen, Fujian 361005, P. R. China

识别中,编辑距离也被认为是有效的图距离度量方法,文献[4]给出了图的编辑距离的二次规划算法。从集合的角度看,也有一种 Lattice 相似度量^[5~7]。从形状的角度看,还有 Hausdorff 距离、Hamming 距离等度量。文献[8]给出了相似性度量的比较。但是,这些方法一般都需要全局信息,计算量较大。本文提出的评价函数也是对加权简单线性图结构的度量,计算简单。我们没有选择内积结构的相对相似性度量,而是选择了一个评价函数来计算绝对几何度量。这样,能够减少计算量,以满足实时性需求。

本文提出的算法具有线性累加性,是一个在线算法,能够在分割信号的同时完成信号的识别。

2 基于感知的识别

我们采集到的信号是三维加速信号,分别对应于传感器自身局部坐标系的 x 、 y 、 z 方向的三个分量值。首先,这个数据也包含了重力加速度。由于重力加速度的存在,使得三个轴采样得到的数据并非位置无关。这对于车载系统加速传感器等位置相对固定的传感器几乎没有影响,但对于手持设备上的传感器来说,却会因为位置的不同而有重大的影响。我们使用向量的位置无关性来消除这种位置有关性。我们整体的处理框架分为两个模块 Module1 和 Module2。Module1 判断信号表示中动作是否开始和结束,着重考虑信号处理的实时性。如果动作开始,则获取信号的同时将信号同步传给 Module2 进行处理。若判断动作信号结束,则停止向 Module2 输入信号。为了保证实时性,我们使用了一种在线算法。

加速度信号可以完全表示空间中的曲线。

定理 1 给定加速度函数 $g(t)$ 、初始时间 t_0 、初始速度 v_0 以及初始位置 C_0 ,该物体在空间中的运动轨迹 $C(x, y, z, t)$ 是唯一的。

证明 由于三维空间的正交性,只考虑一维情形。

由 $\frac{\partial v}{\partial t} = g(t)$, $v(t_0) = v_0$, 知 $v(t) = \int_{t_0}^t g(t) dt + v_0$ - $g(t) dt / t=t_0$, 从而速度是确定的。

$x(t) = x_0 + \int_{t_0}^t v(t) dt$, 从而位置是确定的。证毕。

待识别信号的初始速度和初始位置都是确定的,从而对三维加速信号的识别也就是对动作在空间划过的曲线路径的识别。而且由于积分的关系,该识别有更好的冗余性。使用加速信号进行动作识别,不仅基于上述定理,而且基于动作的运动特征,因为加速度直接反映力学特性。所以,对加速度信号的识别更能反映动作的运动和力学特征。

2.1 信号的分割

Module1 的工作比较简单,我们使用了在线算法以保证实时性。首先,我们有如下定义:

定义 1 对于一维离散数字信号 $(t_0, t_1, t_2, \dots, t_n)$, 如果 t_i 满足如下条件之一,则称 t_i 为关键点:

- (1) $i = 0$ 或者 $i = n$;
- (2) $t_{i-1} < t_i$ 且 $t_i < t_{i+1}$, $i \in \{1, 2, \dots, n-1\}$;

$$(3) t_{i-1} = t_i \text{ 且 } t_i > t_{i+1}, i \in \{1, 2, \dots, n-1\};$$

$$(4) t_{i-1} = t_i \text{ 且 } t_i < t_{i+1}, i \in \{1, 2, \dots, n-1\};$$

$$(5) t_{i-1} > t_i \text{ 且 } t_i = t_{i+1}, i \in \{1, 2, \dots, n-1\}.$$

基于差分的方法能够获取动作加速度信号的关键点,也可以获取无动作情况下的加速度。这个加速度可以作为背景信息在一段时间以后被更新,我们记其为 g_0 。一般情况下可以认为 g_0 是重力加速度,从而可以将局部坐标系转为世界坐标系,以消除位置相关性。但是,对于游戏手柄上集成的加速传感器等背景信息不是十分充足的情况,这种分割信号的方法就不能满足我们的要求。事实上,我们利用矢量的位置无关性可以直接消除位置相关性,而不需要额外的计算。本文使用的分割算法是一种在线算法,在获取信号的同时探测关键点并完成计算,当分割完成时评价函数也同时完成。这满足了很多实时信号处理的需求。

2.2 信号的识别

Module2 同步地从 Module1 接收信号并完成实时性计算。当信号结束时,返回计算结果。在接收信号的同时进行关键点的探测,之后累加三角面所含的信息量。由于这种累加性,该算法是一个在线算法。

2.2.1 关键点探测以消除位置相关性

该探测算法非常简单,按照上述定义在线扫描一维信号即可。对于三维信号 (x, y, z) ,我们对其各个分量分别进行关键点检测,依次得到集合 S_1, S_2, S_3 。 $S_i (i = 1, 2, 3)$ 中每一个元素 a_j 都唯一对应于一个三维信号 b_j 。 $T = \{b_j | (b_j) = (a_j, a_j, a_j) \in S_i\}$ 即为我们检测到的关键点,其中 π 为投影函数。

通过关键点,我们可以构造矢量来表示动作的拓扑结构。检测出关键点以后,相邻两个关键点之差表示的矢量消除了重力加速度的影响,从而达到了与局部坐标系无关的目的。基于空间中三维矢量的结构是相对稳定的,该结构反映了动作的所有力学特征和基本空间特征(定理 1)。

2.2.2 简单线性图及其评价函数

由于该三维空间的特殊结构,我们定义简单线性图 $G = (V, E)$ 如下:

定义 2 V 中节点为一个有序序列, $V = (v_1, v_2, \dots, v_n)$, $d^+(v_i) = 1, i = 1, \dots, n-1$, $d^-(v_i) = 1, i = 2, \dots, n$, 其中 d^+ 为出度函数, d^- 为入度函数。

定理 2 对于两个简单线性图 $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, 若 $|V_1| = |V_2|$, 则 G_1, G_2 同构。

证明 只需考虑 V_1, V_2 的有序性即可。证毕。

显然,动作的三维信号是一个简单线性图。为了便于评价简单线性图 $G_0 = (V_0, E_0)$, 我们构造一个映射 f 。 $f(G_0) = G$, 其中 G 为一个加权简单线性图。 f 将 E_0 中的边映射为 G 中的顶点,将 E_0 中相邻两条边之间的夹角映射为 G 中的边。其中, G 的顶点权值为 E_0 中对应边的长度, G 的边权值为 E_0 中相应夹角的响应函数。

如何评价两个简单线性图之间的相似程度是整个系统的关键。当然,我们可以使用基于曲线相似性的方法,利用文献[8]中的关键点删除技术来进行相似性评价。但是,我们认为这些方法都不能反映力学特征,我们要构造新的评价准则以适合加速传感器的力学特征和人眼的视觉感知特

征。基于曲线论^[9]和信息论^[10],定义如下评价函数:

$$E = \frac{1}{n-1} + \frac{\sum_{i=1}^{n-2} (i+1) \sin(i) y(i)}{\sum_{i=1}^{n-1} (i+1)^2} \quad (1)$$

其中, i 为 G 中的顶点权值, θ_i 为 i 与 $i+1$ 之间的夹角, $y(i)$ 为转角信息度量, $(i+1) \sin(i)$ 为该度量的权重(实际上为该曲线的局部面积), (L, θ) 为长度 L 与角度 θ 之间信息量的比函数。反映的是人眼对于长度变化感知和角度变化感知之间的比例关系。

由于这种累加性,我们可以在接收信号的同时进行计算,而不需要得到整体信号后才进行计算,从而就提高了实时性。

为了使该评价函数不依赖于加权简单线性图的“大小”,我们将式(1)改写如下:

$$E = \frac{1}{n-1} + \frac{\sum_{i=1}^{n-2} (i+1) \sin(i) y(i)}{\sum_{i=1}^{n-1} (i+1)^2} \quad (2)$$

其中, $y(i)$ 为转角信息度量,满足 $y(0) = 0$ 。从信息论的角度,可以将 $y(i)$ 定义为 $y(i) = -\log(\dots)$ 。另外,为了简化计算,将 λ 人为设定为常数。这样,就得到了一个近似的评价函数:

$$E^* = \frac{1}{n-1} - \frac{\sum_{i=1}^{n-2} (i+1) \sin(i) \log(\lambda)}{\sum_{i=1}^{n-1} (i+1)^2} \quad (3)$$

3 实验结果

对于动作信号,我们随机选取三个动作,每个动作随机选取两组信号来进行实验。三个动作分别是“画叉”、“斜劈”和“画圆”。由于有两组信号,我们分别用 $c1$ 、 $c2$ 、 $p1$ 、 $p2$ 、 $q1$ 、 $q2$ 表示该信号。得到的结果如图 1 所示。

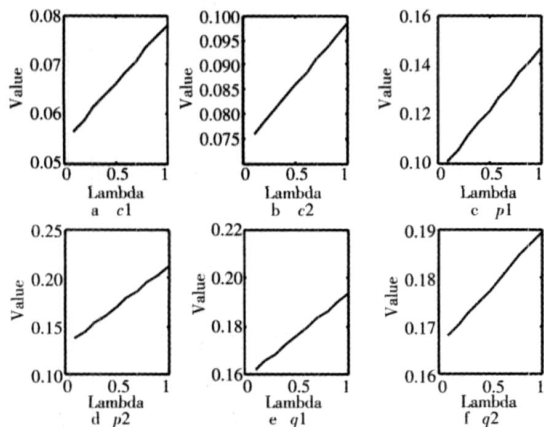


图 1 Lambda 与评价函数值在不同动作中的关系

经过实验,我们取 $\lambda = 0.385$,分别计算各个信号在评价函数下的函数值,如表 1 所示。

另随机抽取一组数据,重复上述实验,得到的结果如表 2 所示。

表 1 各组动作的信息评价 1

信号代码	关键点数量	评价函数值	类内均值	类内距离	类间距离
c1	38	0.068 8	0.078 8	0.010 00	0.084 55(p) 0.111 35(q)
c2	35	0.088 8		0.010 00	0.064 55(p) 0.091 35(q)
p1	18	0.126 9	0.153 4	0.026 45	0.048 10(c) 0.053 25(q)
p2	12	0.179 8		0.026 45	0.101 00(c) 0.000 35(q)
q1	35	0.180 1	0.180 2	0.000 05	0.101 30(c) 0.026 75(p)
q2	36	0.180 2		0.000 05	0.101 40(c) 0.026 85(p)

表 2 各组动作的信息评价 2

信号代码	关键点数量	评价函数值	类内均值	类内距离	类间距离
c1	43	0.058 7	0.072 95	0.014 25	0.057 90(p) 0.110 25(q)
c2	26	0.087 2		0.014 25	0.029 40(p) 0.081 75(q)
p1	33	0.104 6	0.116 60	0.012 00	0.031 65(c) 0.064 35(q)
p2	28	0.128 6		0.012 00	0.055 65(c) 0.040 35(q)
q1	32	0.175 1	0.168 95	0.006 15	0.102 50(c) 0.058 50(p)
q2	38	0.162 8		0.006 15	0.089 85(c) 0.046 20(p)

由以上数据可以看出,该方法有较好的区分度。类内均值有较好的稳定性,类间距离普遍大于类内距离。对于极个别情况,区分不够明显。对于这种情况,可能的改进方法是:引入挠度信息和引入 (L, θ) 的非常值逼近函数。

4 结束语

由于本文算法是三维到一维再到常数的压缩算法,也是一个在线算法,所以保证了实时性。利用了拓扑结构信息和简单线性图图结构,使得该方法能够很快完成计算和评价,而且有较好的区分度。但是,由于没有引入挠度信息^[9],可能使得在三维中不一致的信号得到同样的压缩映像。接下来的工作主要在于引入挠度信息和加快计算方面,同时对 (L, θ) 的逼近问题进行深入的研究。

参考文献:

- [1] Junker H, Amft O, Lukowicz P, et al. Gesture Spotting with Body-Worn Inertial Sensors to Detect User Activities [J]. Pattern Recognition, 2008, 41(6): 2010-2024.
- [2] 王珏,周志华,周傲英. 机器学习及其应用[M]. 北京:清华大学出版社,2006.
- [3] Chartrand G, Zhang P. Introduction to Graph Theory[M]. 北京:人民邮电出版社,2007.
- [4] Neuhaus M, Bunke H. A Quadratic Programming Approach to the Graph Edit Distance Problem[C]. LNCS '07, 2007: 92-102.
- [5] Guo Zhang-lin, Tian Jing. Discussion About Similarity Measures in Pattern Recognition of Fuzzy Information[C]. Proc of Int'l Conf on Computational Intelligence and Security Workshops, 2007: 299-303.
- [6] Wang Peizhuang. Synopsis About Fuzzy Mathematics () [J]. Mathematics in Practice and Theory, 1980(2): 45-59.
- [7] Liu Laifu. Identifying Wheat Parents by Using Fuzzy Mathematics[J]. Journal of Beijing Normal University(Natural Science), 1979(3): 78-85.
- [8] Masood A. Optimized Polygonal Approximation by Dominant

(下转第 61 页)

步骤8 若 k 满足 $E(k) < E(i^*)$, 则令 $i^* = k, k = 1$, 以 i^* 为新的初始点, 利用BP神经网络算法(经加入动量因子改进)得到 $E(\cdot)$ 的另一个局部极小点 i_2^* , 令 $i^* = i_2^*$, 返回步骤3; 否则进入下一步。

步骤9 减小 μ (如 $\mu = 0.1 \times \mu$)。

步骤10 若 $\mu < \mu_L$, 则 $k = 1$, 令返回步骤5; 否则算法停止, i^* 为全局极小点。

算法包括两个阶段: 阶段1为局部极小化阶段, 利用BP神经网络算法(经加入动量因子改进)得到 $E(\cdot)$ 的一个局部极小点 i^* ; 阶段2为填充阶段, 构造填充函数 $p(x, x_i^*, \mu)$ 且极小化 $p(x, x_i^*, \mu)$, 当找到满足 $E(k) < E(i^*)$ 的且在定义域内的 k 时, 算法跳出阶段2并返回阶段1, 并把 k 作为新的初始点寻找 $E(\cdot)$ 的另一个局部极小点 i_2^* (如果存在的话)。

算法在步骤3中令 $\mu = 1$, 以上两个阶段如此循环, μ 逐步减小直到 $\mu < \mu_L$ 时, 算法再不能找到更好的 $E(\cdot) < E(i^*)$ 时, 则认为最后找到的局部极小点即为全局极小点, 算法终止。

5 在金华市智能交通监控集成系统中的应用

在智能交通监控集成系统中将分割的字符大小归一化后, 通过算法模块自动识别其车牌号码。金华市智能交通监控集成系统从实际抓拍到的车辆中选取了具有代表性的图片, 通过车牌自动识别, 结果如图1所示。

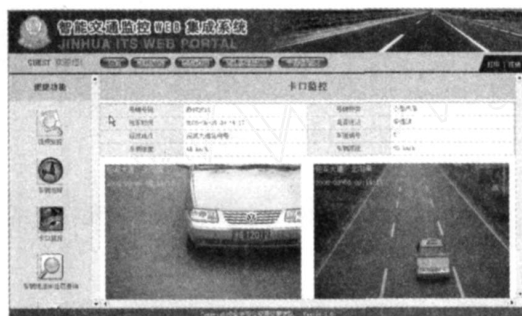


图1 智能交通监控集成系统中的车牌识别实例

字符识别的正确率是衡量算法性能最重要的指标。系统对车牌识别算法在智能交通监控系统的应用实例进行了一系列的测试, 选取字符分割正确的500张图像作为实验样本, 采用具有一个隐含层的三层BP神经网络。测试结果如表1所示, 并与文献[5]中的基于小波变换和神经网络的车牌识别系统进行了比较。

表1 字符识别测试

网络	图片数	识别数	识别率(%)		平均训练时间(ms)	
			本系统	文献[5]	本系统	文献[5]
汉字网络	500	438	87.6	85.7	290	305
字母/数字	500	460	92.0	90.7	189	210
字母网络	500	469	93.8	93.0	250	297
数字网络	500	478	95.6	94.3	143	185

从表1中可见, 基于填充函数法训练BP神经网络的车牌字符识别算法对字符的识别率可以达到95.6%, 平均训练时间最长仅为290ms; 基于加入动量因子改进的BP神经网络的字符识别算法^[5]对字符的识别率最高为94.3%, 平均训练时间最长为305ms。测试数据表明, 本车牌

字符识别算法在准确率和识别时间上都有了较大的提高。其原因主要为本车牌字符识别算法利用填充函数法能够迅速地跳出局部最优, 并在理论上能保证找到全局最优, 而基于加入动量因子改进的BP神经网络并不能保证一定找到全局最优。

6 结束语

字符识别是车牌识别系统的关键问题, 字符识别算法中大都运用BP算法, 但BP算法的收敛速度较慢、容易陷入局部最优。本文提出一种新的基于填充函数法训练BP神经网络的车牌字符识别算法, 该算法中的填充函数法是确定型全局优化, 能够在极小化问题搜索过程中检测出局部极小点后跳出该局部极小点, 得到一个更低的极小点, 重复此过程求得全局最优解, 从而提高算法的收敛速度, 增强其全局寻优能力。测试结果表明, 该车牌字符识别算法的准确率较高, 在实际系统中达到了较好的运行效果。

参考文献:

- [1] 胡占义, 吴福朝. 基于主动视觉摄像机标定方法[J]. 计算机学报, 2002, 25(11): 1149-1156.
- [2] 安勇, 张高伟. 基于灰度图像的车牌识别系统[J]. 计算机工程与科学, 2004, 28(2): 61-62.
- [3] 周亮, 刘云. 基于数学形态学的汽车牌照提取算法[J]. 青岛科技大学学报, 2007, 28(1): 69-72.
- [4] 任洪海, 张飞侠, 肖刚强. 一种基于垂直纹理特征的车牌分割方法[J]. 计算机与数字工程, 2007, 35(5): 173-174.
- [5] 苏厚胜. 车牌识别系统的设计与实现[J]. 可编程控制器与工厂自动化, 2006(3): 103-107.
- [6] Fan J Y. An Approach to Fault Diagnosis of Chemical Processes via Neural Networks[J]. Journal of AIChE, 1993, 39(1): 34-38.
- [7] 李俊民, 李靖, 薄立军. 一类非线性动态系统的自适应模糊小波神经网络控制[J]. 工程数学学报, 2006, 23(1): 37-44.
- [8] 陈凤娟. 时滞细胞神经网络的指数稳定性[J]. 浙江师范大学学报(自然科学版), 2007, 30(1): 34-38.
- [9] Ge R P. The Theory of Filled Function Methods for Finding Global Minimizers of Nonlinearly Constrained Minimization Problems[J]. Journal of Computer Mathematics, 1987, 5(1): 1-9.

(上接第58页)

- [10] Point Deletion[J]. Pattern Recognition, 2008, 41(1): 227-239.
- [9] 陈省身, 陈维桓. 微分几何讲义[M]. 北京: 北京大学出版社, 2001.
- [10] Attneave F. Some Informational Aspects of Visual Perception[J]. Psychological Review, 1954, 61(3): 183-193.
- [11] Latechi L J, Lakamper R. Shape Similarity Measure Based on Correspondence of Visual Parts[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1185-1190.
- [12] Armstrong M A. 基础拓扑学[M]. 北京: 北京大学出版社, 1983.