

基于 AdaBoost 算法和光流匹配的实时手势识别

王 凯¹, 于鸿洋¹, 张 萍²

(1 电子科技大学 电子科学技术研究院, 四川 成都 611731; 2 电子科技大学 电子工程学院, 四川 成都 611731)

摘 要: 着眼于更宽泛和更便捷的应用需要, 提出了基于 AdaBoost 算法和光流匹配的实时手势识别方案. 只需连接到计算机的摄像头读取二维手势视频片段就能对手势作为较为准确的识别. 其中, 采用 AdaBoost 算法遍历图像, 完成静态手势的识别工作; 在动态手势的识别过程中, 运用了光流法结合模板匹配的方法. 整个系统对静态和动态手势的识别均具有较强的鲁棒性.

关键词: 手势识别; AdaBoost; 光流; 模板匹配

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2012)04-0138-04

Real-time Gesture Recognition Based on AdaBoost Algorithm and Optical Flow Matching

WANG Kai¹, YU Hong-yang¹, ZHANG Ping²

(1 Research Institute of Electronic Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, China; 2 School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Focusing on more general and more convenient application, a novel real-time gesture recognition method based on AdaBoost algorithm and optical flow matching was put forward. In detail, the AdaBoost algorithm was used to traverse the whole image for the recognition of static gestures. As to the dynamic gestures, the method combining optical flow with template matching was utilized. The whole system has strong robustness in the recognition of both static and dynamic gestures.

Key words: gesture recognition; AdaBoost; optical flow; template matching

1 引言

在日常生活中, 手势是一种常用的自然直观的交流方式. 因此, 手势识别技术也越来越受到关注. 手势识别包括静态的姿态识别和动态的手势识别. 其中动态的手势识别更具有挑战性. 手势识别按输入设备的不同可分为基于数据手套的手势识别和基于视觉的手势识别. 由于基于数据手套的识别需要使用用户带上笨重的识别设备, 限制了交互的自然性和方便性, 并不实用, 因此当前主要研究基于计算机视觉的手势识别技术, 而基于计算机视觉的手势识别主要由收拾分割、手势特征提取、手势识别等阶段

组成^[1].

1911 年富士通实验室完成了 46 个手势符号的识别^[2]; Starnier 等对美国手势中带有词性的 40 个词汇随机组成的短句子识别率达到 99.2%^[3-4]; 利用数据手套等典型传感设备的方法中, 卡耐基-梅隆大学的 Christopher Lee 和 Xu 在 1995 年完成了一个操纵机器人的手势控制系统^[4]. 在我国, 文献^[5]中给出了 ANN 与 HMM 的混合方法作为手语训练识别方法, 以增加识别方法的分类特性和减少模型的估计参数的个数. 将 ANN-HNN 混合方法应用于 18 个传感器的 CyberGlove 型号数据手套的中国手语识别系统中, 孤立词识别率为 90%, 简单语句识

收稿日期: 2011-07-12; 修回日期: 2011-09-18

基金项目: 中央高校基本科研业务费专项资金(ZYGX2009J024)

别率为 92%^[5].

本文提出了一种基于 AdaBoost 算法和光流模板匹配的手势识别新方法. 该方法使用训练好的分类器遍历图像来定位静态手势, 并作为识别; 采用光流矢量统计和模板匹配相结合的方法来识别动态手势.

2 静态手势识别

相对动态手势, 静态手势的识别要简单一些. 要达到实时识别, 对算法的计算速度和准确性有较高要求. 因此, 本文采用了速度较快并且识别准确度较高的 AdaBoost 算法.

2.1 AdaBoost 算法

AdaBoost 算法的核心思想是将容易得到的正确率刚刚超过 50% 的弱分类器整合成为正确率远远超过 50% 的强分类器. AdaBoost 算法能够自适应的调节训练样本权重的分布, 不断挑选出当前样本权重分布下表现最佳的弱分类器, 并整合所有得到的弱分类器, 让它们按照一定的权重投票, 组成一个强分类器^[6].

AdaBoost 算法训练了 T 个弱分类器 $h_t, t \in \{1, \dots, T\}$. 这些弱分类器大多情况只包含一次分裂或仅有几次分裂的决策树. 最后做决定的时候将赋值权重 α_t 给每个分类器. AdaBoost 训练时输入的特征向量是 x_i , 向量的类别标签是 y_i (这里 $i=1, \dots, M, M$ 是样本总数), 且 $y_i \in \{1, -1\}$. 首先, 初始化数据样本的权值 $D_t(i)$ 来决定分类器将一个数据点分类错误的代价是多少. AdaBoost 算法的主要特征是, 在训练过程中, 这个代价将会更新, 使得后来的弱分类器更加关注与前面的分类器没有分对的数据点. 算法如下:

(1) $D_1(i) = 1/M, i=1, \dots, M$

(2) 针对 $t=1, \dots, T$:

① 寻找使得权重为 $D_t(i)$ 的总错误最小的分类器 h_t .

② 求 $h_t = \operatorname{argmin}_{h_j \in H} \epsilon_j$, 这里 $\epsilon_j = \sum_{i=1}^m D_t(i) (其中 $y_i \neq h_j(x_i)$), 如果最小错误满足 $\epsilon_j < 0.5$, 则继续; 否则退出.$

③ 设置 h_t 的权重 $\alpha_t = \log[(1 - \epsilon_t)/\epsilon_t]$, 这里 ϵ_t 是②中的最小错误.

④ 更新数据点权重: $D_{t+1}(i) = [D_t(i) \exp(-\alpha_t y_i h_t(x_i))]/Z_t$, 这里 Z_t 将所有数据点的权重归一化.

在②中, 若找不到小于 50% 错误率的分类器, 则停止, 表明可能需要更好的特征.

当前面的训练算法结束后, 最后的强分类器接受输入向量 x , 使用所有的弱分类器的加权和来进行分类.

$$H(x) = \operatorname{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right)$$

这里, 符号函数把所有正数变为 1, 把所有的负数变为 -1 (0 仍然是 0)^[7].

2.2 Haar 分类器

Haar 分类器的前身是 Viola-Jones 分类器. Viola-Jones 分类器使用 AdaBoost 算法, 但将其组织为筛选式级联分类器, 每个节点是多个树构成的分类器, 且每个节点的正确识别率很高, 正确拒绝率很低. 在任一级计算中, 如果一旦获得目标“不在类别中”的结论, 则计算终止, 宣布该位置没有目标. 因此只有通过分类器中的所有级别, 才认为目标被检测到. 目标出现频率较低的时候, 筛选式的级联分类器可以显著地降低计算量, 因为大部分被检测的区域都很早被筛选掉, 迅速判断出此处无目标. 接着, 该分类器被对角特征扩展, 成为 Haar 分类器^[8]. Haar 分类器是一个监督分类器, 使用类 Haar 的小波特征, 该特征由矩形图像区域的加减组成. 先对图像进行直方图均衡化并归一化到同样大小, 然后标记里面是否包含要检测的物体. Haar 分类器中的筛选式级联如图 1 所示. 静态手势分类器制作过程如图 2 所示.

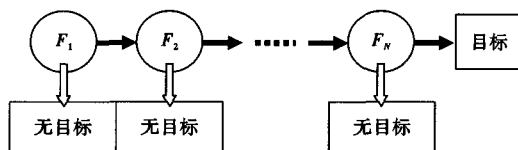


图1 Haar 分类器的筛选式级联

3 动态手势识别

光流是基于像素点定义的, 所有光流的集合称为光流场. 光流场的计算最初由 Horn 和 Schunk 提出来, 光流场计算的基本等式是在相邻图像之间的时间间隔很小并且图像中灰度变化很小的前提条件下成立的^[9]. 目前主要有两种经典的光流算法: Horn-Schunck 算法和 Lucas-Kanade 算法. 通过在 OpenCV2.0 下针对多组图像的光流计算结果分析, 本文采用的金字塔 Lucas-Kanade 算法可以得到更精确的光流图像. 具体的跟踪方法是: 在图像金字塔

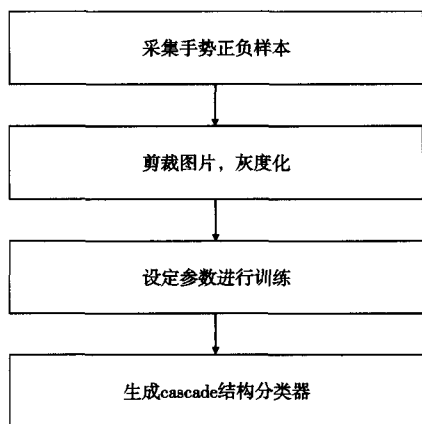


图2 静态手势分类器制作流程

的最高层计算光流. 用得到的运动估计结果作为下一层金字塔的起始点, 重复这个过程直到到达金字塔的最底层. 这样就将不满足运动假设的可能性降到最小, 从而实现对更快和更长的运动的跟踪^[9]. 其原理如图3所示.

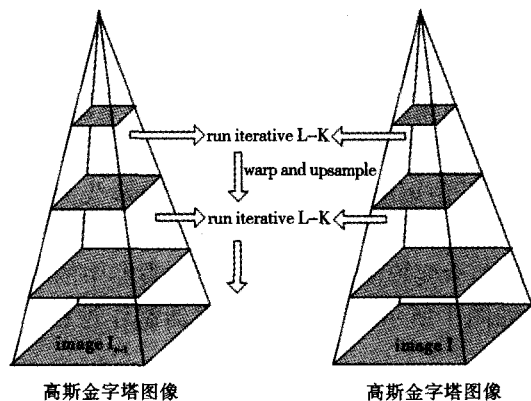


图3 金字塔 Lucas-Kanade 光流

模板匹配是图像识别中最具代表性的方法之一. 具有算法实现容易、匹配速度快的特点. 将从待识别的图像提取的若干特征量与模板对应的特征量进行比较, 计算图像和模板特征量之间的距离, 用最小距离法判定所属类^[10]. 将实时得到的光流数据与事先准备好的光流数据模板作比较即可对动态手势作出识别. 动态手势识别流程如图4所示.

4 实验过程与结果

本文对3个静态手势的分类器训练过程均采用了5000个正样本和10000个负样本. 样本由多种光照条件和背景下的图片组成, 其中, 提出了“添加A以外手势的部分正样本作为A手势负样本”的改进训练方法, 进一步降低了识别错误率. 经测试, 准

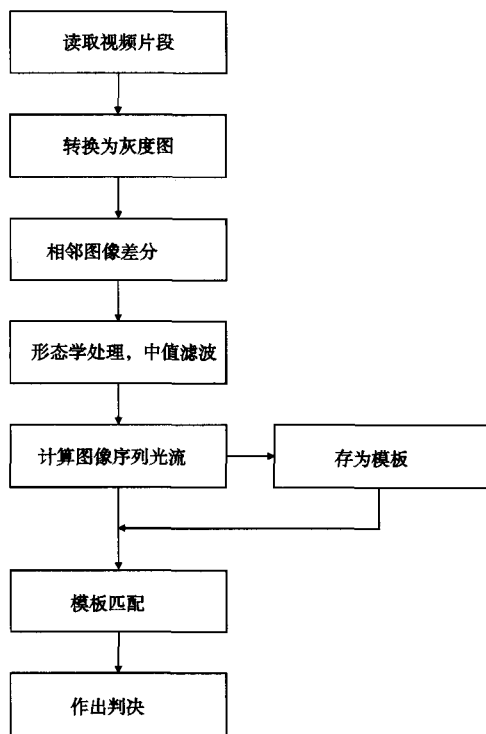


图4 动态手势识别流程图

确识别率普遍能达到95%左右.

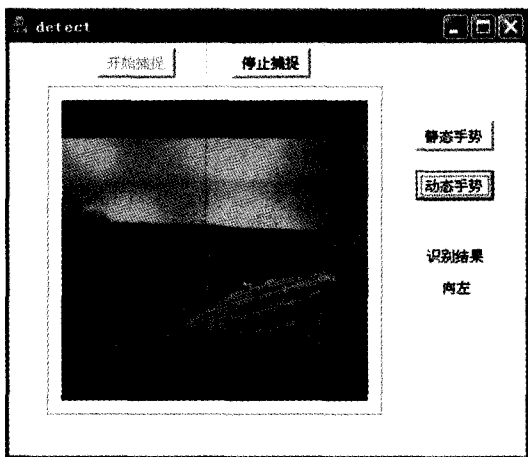
本文将每个动态手势持续时间定为1s, 摄像头通过USB连接到电脑, 采样率为30帧/s, 分辨率为640×480. 将读取到的某个动态手势的30帧图像转为灰度图像(转化为灰度图像可大大降低计算量, 但对识别效果影响较小)后进行前后差分, 再通过腐蚀、膨胀、中值滤波等一系列降噪处理后即为运动前景图像. 然后用上述金字塔 Lucas-Kanade 算法计算前后图像间的光流, 把得到的较明显的光流(过小及过大的光流矢量均可判定为噪声)矢量按方向阈值均匀划分区间并记录保存. 识别过程中, 将实时得到的一系列光流与事先保存好的光流模板进行比较(文中采用对应元素差的平方和作为衡量标准)并作出判定.

实验测试平台采用自行在 Visual Studio 2008 编译环境中运用 OpenCV2.0 函数库编写的“detect”软件. 在编写过程中, 运用到了 OpenCV 中文论坛上分享的 CCameraDS 类来处理摄像头打开、关闭、读取图像和调整图像大小等操作, 其中包含的 directshow 技术对图像的采集性能有较大提升.

软件界面主要由控制视频捕捉、切换静态手势识别、显示识别结果和视频显示窗口等几个实用部分构成. 图5分别为软件测试过程中对某个静态手势和某一动态手势的准确识别.



(a)识别到静态手势“拳头”



(b)识别到动态手势“向左”

图5 两个典型界面

5 结束语

本文提出的方法对少量的静态手势和动态手势的识别效果较佳,但在要求识别大量手势动作的环境中,随着各手势间的差异变小,判定将容易出现错误.如果想取得比较好的效果,必须加大正负样本数量来投入训练过程,动态识别中的模板数也要相应增加,计算量和识别过程耗时也将增大.而且,在处理动态手势识别过程中,背景的一些较大变化(若光照明暗变化)将影响到识别效果.增加模板容量可使上述情况有一定改观,但同样将增加计算量和识别耗时.下一步将研究解决上述两个问题的更好途径.

参考文献:

- [1] 雷海龙,杨胜,刘红新. 融合图像轮廓矩和傅立叶描叙子的手势识别[J]. 微电子学与计算机,2011(4):103-105.
- [2] Takahashi T, Shino F K. Hand gesture coding based on experiments using a hand gesture interface device [J]. SIGCHI Bulletin, 1991, 23(2): 67-73.
- [3] Starner T, Pentland. A visual recognition of american sign language using hidden markov models[R]. Technical Report TR306, Media Lab, MIT, 1995.
- [4] Starner T, Pentland. A real-time american sign language recognition from video using hidden markov models[R]. Technical Report TR375, Media Lab, MIT, 1996.
- [5] 吴江琴,高文,陈熙霖. 基于数据手套的汉语手指字母的识别[J]. 模式识别与人工智能, 1999, 12(1):74-78.
- [6] 严超,王元庆,张兆扬. 基于 AdaBoost 和 Kalman 算法的人眼检测与跟踪[J]. 南京大学学报:自然科学, 2010, 46(6): 682-684.
- [7] Freund Y, Schapire R E. Experiments with a new boosting algorithm[C]//Proceedings of the Thirteenth International Conference. San Francisco: IEEE, 1996: 148-156.
- [8] Viola P, Jones M J. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition[s. l.]: IEEE, 2011:511-518.
- [9] 夏侯玉娇,龚声蓉,刘纯平,等. 结合 Gaussian 分布和 LK 光流法的视频对象分割算法[J]. 微电子学与计算机,2009(6):240-243.
- [10] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C]// Proceedings of the 1981 DARPA Imaging Understanding Workshop, Vancouver, British Columbia: ACM, 1981: 121-130.

作者简介:

王凯 男,(1986—),硕士研究生.研究方向为数字视频处理.

于鸿洋 男,(1964—),副教授.研究方向为数字视频、图像传输和算法.

张萍 女,(1976—),博士.研究方向为嵌入式系统、数字视频处理.