

# 考场异常行为检测算法

戴金波<sup>1,2</sup>,龙曼丽<sup>3</sup>,赵宏伟<sup>1</sup>,陈奋君<sup>1</sup>

(1. 吉林大学 计算机科学与技术学院,长春 130012; 2. 长春师范学院 计算机科学与技术学院,长春 130032;  
3. 吉林大学 公共外语教育学院,长春 130012)

**摘要:**以智能视频监控理论为依据,结合考试现场特点,提出了一种能够进行考场异常行为检测的高效算法。该算法从考生信息结构和内容方面作了科学设计,提出了行为覆盖区、3维考场关注度等概念。仿真实验针对分析器的准确率和效率两方面进行,并特别比较了本算法设计的分析器和普通方式分析器的效率。实验结果表明,本算法能很好地挖掘视频帧间的历史关系,与未采用本算法的普通方式相比检测效率有较大提高。

**关键词:**计算机应用; 行为覆盖区; Latent SVM; 行为模型; 三维考场关注度

**中图分类号:**TP391.41    **文献标志码:**A    **文章编号:**1671-5497(2012)Sup. 1-0236-05

## Algorithm of the exam abnormal behavior detection

DAI Jin-bo<sup>1,2</sup>, LONG Man-li<sup>3</sup>, ZHAO Hong-wei<sup>1</sup>, CHEN Fen-jun<sup>1</sup>

(1. College of Computer Science and Technology, Jilin University, Changchun 130012, China; 2. Department of Computer Science and Technology, Changchun Normal University, Changchun 130032, China; 3. School of Foreign Language Education, Jilin University, Changchun 130012, China)

**Abstract:** A efficient algorithm was presented for abnormal behavior detection in examroom based on intelligent video surveillance theory and examroom characters. The concepts of Behavior Coverage Region and Three - Dimensional Attention(TDA) were given in this algorithm from the structure and content of examinee information. And the efficiency between analyzer designed by this algorithm and ordinal methods had been compared after the simulation experiments to analyzer's accuracy and efficiency. The result suggested that this algorithm can reveal the history relationship of video inter-frame and can also accelerate greatly comparing with ordinal methods.

**Key words:** computer application; action coverage area (ACA); latent SVM; viewpoint action model (MVAM); three-dimensional attention (TDA)

随着视频监控技术的不断发展,智能视频监控技术在多学科融合的基础上,作为其发展的第三个阶段之中的“机器脑”部分出现,并因所具备的自动识别能力在不同的领域中得到广泛应

用。目前,基于各类考试的公平性、透明性角度考虑,视频监控技术在考试现场也得以发挥作用,但智能化程度还受到一定的约束,视频信息的过滤、检索和存储效率也成为其发展面临的瓶颈。如何

收稿日期:2012-03-26.

基金项目:国家自然科学基金项目(61101155);吉林省科技发展计划资助项目(20101504).

作者简介:戴金波(1971-),女,副教授,博士研究生. 研究方向:智能信息系统与嵌入式技术.

E-mail:daijinbo@sina.com

通信作者:赵宏伟(1962-),男,教授,博士生导师. 研究方向:智能信息系统与嵌入式技术.

E-mail:zhaohw@jlu.edu.cn

高效而有针对性地处理海量数据是其亟待解决的问题。对此,本文在观察和统计了考场异常行为的基础上设计了一种用于分析视频帧序列的算法。该算法包含了考生信息结构和内容的设计,将空间、时间、诚信度三者结合在一起建立考核标准,确立了 3 维考场关注度 TDA (Three-Dimensional Attention)。从考生数据结构出发发掘视频帧间的历史关系,进行视频帧提取,从而提高检测效率,为考场情况和考生个人表现的评价提供数据基础。

## 1 考场行为模型的建立

通过对 Part-based 模型、可变形模板理论和小姿态(Poselet)<sup>[1]</sup>表示法的分析,本文设计了多视角行为模型 MVAM(Multiple viewpoint action model)。该模型由正面、左/右 45°、左/右侧和背面 4 个视角部件构成。MVAM 能够应对同类行为因视角变化导致的外观差异。采用行为检测效果良好的 HOG 特征提取方法<sup>[2]</sup>来提高模型在复杂背景下检测行为目标的准确度。为让模板能简洁地表达行为的内容,提出了行为覆盖区 ACA (Action coverage area)的概念,即能够完整描述与行为相关的人体部分和物体的视觉区域。同时,为了提高模型在目标不完整的情况下具有较强的鲁棒性及提高检测效率,在行为覆盖区内又增加了行为核心 AC(Action core),它是一个比 ACA 更小的区域,具有更快的运算速度,可以作为视频分析的辅助。考虑到包括人体和行为本身的变化等行为的非刚性变化,本文使用 Part-based 模型和可变形模板来详细设计了行为覆盖区和行为核心。并使用 Latent SVM<sup>[3-4]</sup>进行分类获得了高性能的行为模型。Latent SVM 在线性 SVM 的基础上引入了潜在变量。在训练的过程中本文将行为的视角和各个 Part-base 部分的位置作为分类器的潜在变量。Latent SVM 方法提供了高效的训练集挖掘方法,加快了过程收敛速度,同时它高效的目标检测算法为行为模型应用于视频分析提供了可靠的技术保障。

## 2 考场异常行为智能检测的设计

考场异常行为或者说可疑行为定义为非正常答卷行为。因考生正常答卷时的身体状态都是大同小异的,故行为模型训练只针对正常的答卷行为。与此相反,考生考试时的作弊行为形式多样

化,难以一一罗列出来进行训练。即使能够罗列出来整理归类,训练得到的模型数量较多,匹配的时间复杂度会很高,将直接影响系统的实时性。按照高考考场守则常见的考场作弊行为有:交头接耳、传递答案、自行传递个人用品等。本文智能检测设计的基本功能需求是能够把这些作弊行为排除在正常的考试行为之外。设可罗列的考场行为集合为

$$A = \{b_1, \dots, b_m, c_1, \dots, c_n, d_1, \dots, d_k\} \quad (1)$$

式中:  $b$  为正常的考试行为, 种类数目为  $m$ , 一般较小;  $c$  为异常的考试行为, 种类数目为  $n$ , 一般很大;  $d$  为不含  $b$  和  $c$  的行为,  $k$  是无穷的, 但实际收集到的样本有限。考试行为模型的训练以行为  $b$  为正样本最为合适, 负样本则用  $c$  和  $d$ , 其中  $c$  非常重要, 它与考场相关度最高, 即出现的概率较大, 模型随时都需要对其进行判别。 $d$  样本作为训练的补充。收集以上图片数据, 用 AnnotationV2.0 处理生成训练样本, 采用本文的行为模型表示法并经过 Latent SVM 分类器训练得到考试行为模型。

行为模型是在静态图像上训练得到的, 它的检测也需在静态图像上进行。给定一个视频  $V$ , 实现对视频的检测首先要把它转换成帧的集合, 表示为式(2)。

$$F = \{f_1, \dots, f_n\} \quad (2)$$

式中:  $f$  是有序的, 携带其原来在视频  $V$  中的时间戳。但是, 视频帧的数据量巨大, 处理时需要最大限度地减小数据量。实际观察中, 考生答卷时的身体姿态变化频率不高, 一个作弊动作很少能够在少于 0.5 s 内完成, 在 1 s 内提取 2 帧完全能够满足对考生行为的全程监测。假设视频的帧率是 24 帧/s, 则每秒提取 1 或 2 帧, 一分钟处理 60 或 120 帧, 这样, 1 min 的视频帧里只需分析其中的 8.3%。对  $F$  的分辨率重新调整为合适的尺寸后即可作为分析的输入变量。

把帧按次序输入分析器中, 检测正常考试行为为目标, 统计数目, 设帧中考生的数目为  $m$ , 检测到的数目是  $k$ , 如果  $k$  小于  $m$  则该帧被标记为可疑帧, 发送警示。

显然这样的处理过于简单, 连续的帧包含重要的历史信息, 帧与帧之间必然存在有价值的联系, 若将它们完全割裂开来处理不可取。为使检测器具备更多的智能, 充分地应用视频帧中的历史信息和帧间的联系, 本文在对考场现场进行观

察和统计的基础上,提出了一种高效的智能检测算法。即在帧中把在场考生用矩形框标记出来,当框中的人物存在作弊嫌疑时用红框表示,否则用绿框表示或者不用标示出来。检测器的检测过程也进行了更多细节上的设计,流程如图 1 所示。检测器的初始化任务主要是初始化检测目标信息。目标信息的结构如图 2 所示。

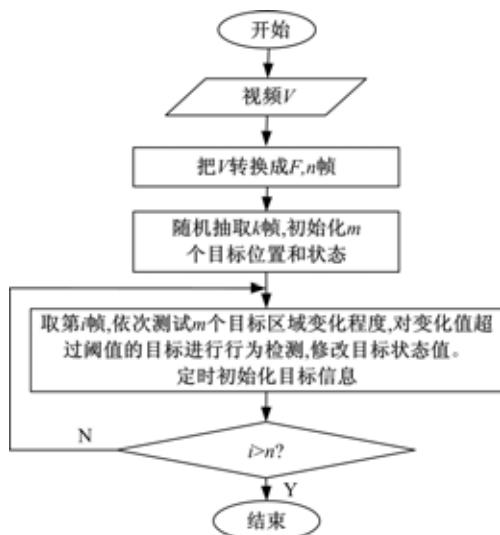


图 1 视频帧分析流程

Fig. 1 Process of frame analysis

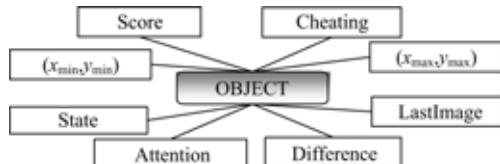


图 2 目标信息结构

Fig. 2 Structure of object information

一个目标对应一个考生。从图 2 可以看到一个目标由 8 个信息块组成。

(1)  $(x_{\min}, y_{\min})$  和  $(x_{\max}, y_{\max})$ : 这两个坐标限定了目标的基本活动区域。在检测器进行初始化时,先提取  $k$  个帧,把行为模型的阈值降低,对这  $k$  个帧进行目标检测,由于阈值降低了,与考试行为接近的行为都会被识别出来,这就达到定位考生的作用。随机取  $k$  帧进行分析是为了防止仅通过 1 帧定位所获得的信息不具代表性甚至出现考生漏检情况。为了避免这种非模型性能本身引发的低级失误扩大至整个智能检测过程,在分析器运行一定帧数后还要重新定位所有目标。目标活动区域大小的调节由参数 Range 控制,它的值表示基本活动区域的缩放比例。

(2) Score: 目标每帧的行为得分,它由考试行

为模型给出。它的高低反应考生的诚信度,一个考生总是保持高分,对他的关注可以小些,从而节省资源。通过统计全场考生的得分情况能够反映整个考场的考风或者宽严程度。

(3) Cheating: 记录考生当前是否异常,异常的次数,每次异常对应的帧编号。每个考生从考试开始到考试结束全程考评,将每个考生的历史 Score 规范化以后可以得到曲线图。该统计数据能够为智能监控系统的其他操作提供参考,同时也能帮助改进分析器的设计。

(4) LastImage: 目标在上一帧活动区域的灰度图像。它的作用是和当前帧的活动区域做比较。

(5) Difference: LastImage 和当前活动区域的差异值。计算方式可以直接采用平均像素差或其他方法。这是一个对分析器运行速度影响很大的变量。分析器内设一个阈值 DetectionThresh,指示是否调用行为模型对当前目标进行打分。调用一次行为模型就要进行一系列的目标检测运算,这是分析器内最耗时的操作,所以当 Difference 大于 DetectionThresh 的情况下才使用。若 Difference 小于 DetectionThresh 则目标保持当前状态。通过设置 DetectionThresh 的大小来控制调用行为模型的检测的频率。

(6) State: 目标的活跃状态,它是一个历史累加值,当超过一定阈值时就强制对目标进行检测。观察发现,考试过程中考生小幅度动作很频繁时往往存在作弊的嫌疑,由于幅度小,Difference 值不一定能够反映出来,需要定量地对这些目标进行检测。另外信用度比较高的目标被检测的次数相对较少,但他们的活跃状态累加到一定值以后就必须接受检验,这保证智能检测的公平性。

(7) Attention: 在智能检测中提出关注度的定义很有意义。从考场空间上看,有些座位的作弊行为相对而言不易被监考员察觉,而有些座位旁边有人作弊时也需要重点关注,对这些位置的关注度必需高于其他的,理想的情况是检测系统能够知道现场监考员当前所在位置进而评估空间关注度;从时间上看,通过观察发现考试刚开始 10 min 内作弊现象较少,考试结束前 5 min 作弊现象也不多,如图 3 所示。在考试中间过程一个考生也不可能频繁地作弊而不被抓到,一般作弊行为都会有一定的时间间隔,所以时间关注度就是要在不同的时段评估考生被关注的程度;从考

生个人考场历史来看,不同考生作弊倾向程度不一样,对不良记录考生当然需要提高关注度。综上,从空间、时间和个体诚信 3 维 ( $S, T, Q$ ) 上定义了考场关注度(根据实际需要可增删维度),本文称之为 3 维考场关注度 TDA (Three-Dimensional Attention)。分析器根据实际情况在 3 维空间定义了一个球体,球的半径  $R$  给出了分析器对目标信任临界球面。若考生的 TDA 半径  $r > R$ ,则分析器对其提高检测频率。



图 3 作弊概率  $P$  与时间  $t$  的关系曲线

Fig. 3 Curve for relation between possibility of cheating and time

通过以上对目标信息结构的设计,为分析器提供了更多的参考信息。设监控教室里有 9 个考生,分析器会根据考生的 TDA 决定间隔多少帧检测一次他的活动区域,同时根据考生活动区域的变化程度决定是否使用行为模型对其进行检测。如图 4 所示,红色目标区域表示有作弊嫌疑,带颜色填充的区域表示在该帧中需要重新检测的目标活动区域。可见在该帧中只需要检测 3 个考生,这节省了系统资源,提高了帧的检测速度。

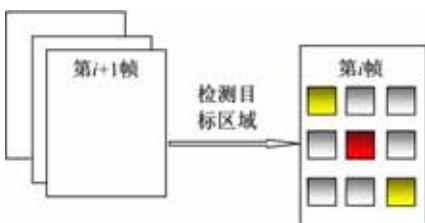


图 4 检测帧中目标的方法

Fig. 4 Method for detecting the objects in the frames

### 3 检测算法的测试

根据检测器的设计,本文在 Ubuntu10 操作系统下用 UNIX MATLAB R2010a 编程实现。计算机 CPU 2.4 GHz, 内存 2 G。实验视频由吉林大学认知计算研究室成员在教室内录制,视频含 4 位考生,以其中两位考生的行为作主要的分析目标,视频内容包含了典型的作弊行为,如偷看他人试卷、传递答案、交头接耳、打小抄、用手机

等。视频提取成最大尺寸为 500 像素的帧序列。

考试行为模型的训练数据集由本文在网络收集并使用本文开发的 AnnotationV2.0 行为标记工具对其进行信息标记。同时使用了 PASCAL VOC<sup>[5]</sup> 提供了开发工具包,内含标记好的图像及其文件信息,同时包含部分基础的数据读取代码。模型的分类和行为的检测参考了 Latent SVM 版本 4<sup>[6]</sup>。

实验设计测试内容为分析器的准确率和效率两方面。考察分析器对所设置的作弊行为的识别情况和处理指定数目帧所需要的时间。特别比较了本文设计的分析器和普通方式的效率。

为了验证本文设计的算法对提高系统分析速度是否有效,特与普通方式做了对比实验,用 5 组帧序列,共 1776 幅图像,完成分析所用时间如表 1 所示。

表 1 检测算法的加速效果测试

Table 1 Evaluation of detection algorithm efficiency

帧组号	帧数	本文算法/s	普通方式/s	加速效果/%
1	318	19.00	243.81	92.21
2	486	15.78	372.60	95.76
3	486	37.80	370.26	89.79
4	486	30.86	371.87	91.70
5	486	26.60	371.65	92.84

由表 1 知,本文所设计的检测算法是非常有效的,与未采用本算法的普通方式相比较平均加速 92.46%。观察普通方式用时变化不大,在帧数相同的情况下第 2、3、4 和 5 组用时皆为 371 s 左右,平均 0.97 s/帧(试验中每隔 6 帧分析一帧,实际上每隔 12~20 帧处理 1 帧也不会丢失动作),这是因为普通方式只是机械地检测帧内容是否有可疑行为,没有保留历史信息,也不做任何判断。与此不同本文算法分析第 2 组帧序列用时与帧数相同的第 3、4 和 5 组用时相差高达 22 s,是第 2 组的 2 倍多,产生这种现象的原因是第 2 组帧序列图像内容变化频率相对较小,即考生的状态相对稳定,本文算法能根据各个目标信息做出相应判断,省去了重复的模型匹配运算。图 5 直观地反映了上述情况。用时曲线呈振荡形式,波峰越多说明帧的内容越活跃。据此可知,若帧数增多,本文算法与普通方式相比优势更大,同时因实验所需特在短时间内做出大量作弊行为,这与实际考场是不相符的,即实际画面会更平静一些,算法速度会更快。

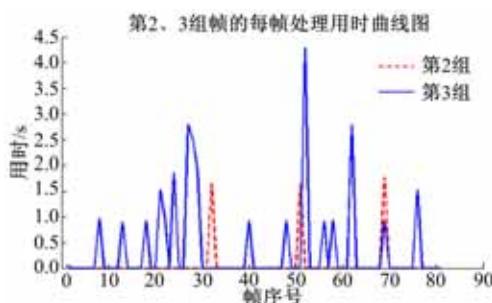


图 5 算法处理每帧的用时

Fig. 5 Time consumed by detection algorithm to analyse each frame

在进行的 5 组实验中,普通方式在准确度上并无优势,其处理速度与本文算法差距太大,故只对本文算法的准确度进行了如下考查。

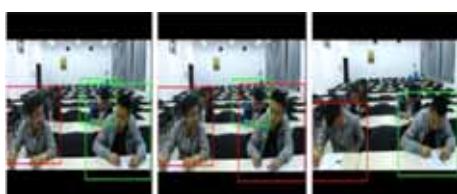
(1) 对偷看与传答案的检测结果如图 6(a)所示。

分析器会自动标出帧中考生的活动区域,红框表示有作弊嫌疑的目标,绿框表示行为正常。对转身、侧身和伸脑袋偷看他人答案的明显动作无漏检情况。

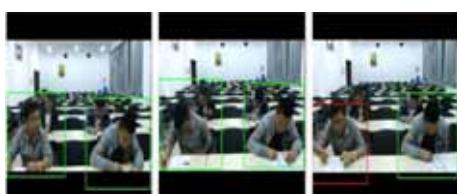
(2) 对交头接耳、传递答案和打小抄的检测情况如图 6(b)所示。



(a) 偷看与对答案



(b) 交头接耳、传答案、抄袭



(c) 未检测出的作弊行为

图 6 作弊检测结构  
Fig. 6 Detection results of cheating

(3) 未能及时检测出来的情况如图 6(c)所示。此处未能及时检测出来的行为都比较细微,其中第 1 个和第 3 个未被检测到,原因是这两个作弊行为主要表现在眼球的活动上,行为模型本身难以区分。而第 2 个行为尽管幅度较大,但未被第一时间检测出来,经过了若干帧后才被发现,造成这种滞后性的原因是分析器只在考生活动区域差异度累积到一定值时才调用行为模型对其进行检测。

#### 4 结束语

实验结果证明基于静态图像的行为识别可应用于视频监控系统中,本文所设计的考生信息数据结构和对应的检测算法提高了系统的效率,实现了对常规作弊行为的识别。为减少系统产生的无效警示需要针对考场的常见非作弊行为进行归类和建模,同时需要调整摄像头的高度和角度以扩大监控的范围。考虑到实际考场还存在考生和监考员的走动,需要为算法增加新的变量来应对复杂的动态场景。

#### 参考文献:

- [1] Bourdev L, Poselets J Malik. Body part detectors trained using 3D human pose annotations[C]// In ICCV, 2009: 1365-1372.
- [2] Dalal N, Triggs B. Histogram of oriented gradients for human detection[C]// In CVPR, 2005: 886-893.
- [3] Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model[C]// In CVPR, 2008: 1-8.
- [4] Felzenszwalb P, Girshick R, McAllester D, et al. Object detection with discriminatively trained part based models[J]. In IEEE PAMI, 2009: 1627-1645.
- [5] PASCAL VOC 数据工具包[EB/OL]. [2010-10-20]. <http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2008/index.html#devkit>.
- [6] Felzenszwalb P F, Girshick R B. Discriminatively trained deformable part models, release4[EB/OL]. [2010-10-15]. <http://people.cs.uchicago.edu/~pff/latent-release4>.