

文章编号: 1002-0268 (2008) 07-0128-04

基于类 Haar 特征的驾驶者人眼 疲劳状态的检测方法研究

林土胜, 陈少珠

(华南理工大学 电信学院, 广东 广州 510640)

摘要: 使用摄像头采集视频图像, 对输入图像做预处理 (图像灰度化、中值滤波等); 通过学习训练的方法构造基于类 Haar 特征的层叠式分类器, 利用基于类 Haar 特征的层叠式分类器从输入图像中直接定位人眼; 把人眼部分的图像截取出来, 二值化人眼图像; 然后计算二值化图像中垂直方向上瞳孔的宽度大小, 从而判断眼睛的状态; 最后通过多次的捕捉, 计算眼睛闭合的频率来得出其疲劳状态。试验验证了上述算法的有效性。

关键词: 交通工程; 疲劳检测; 类 Haar 层叠式分类器; 眼睛定位; 图像处理

中图分类号: U491.2⁺54

文献标识码: A

Study on the Method for Driver's Eyes Fatigue State Detection Based on Haar-like Feature

LIN Tu-sheng, CHEN Shao-zhu

(School of Electronic & Engineering, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: With help of Haar-like cascaded classifier designed by Adaboost algorithm and constructed by training method, driver's eyes were located by the pretreated input video images. Fatigue states of driver's eyes were classified by the frequency of blinking. In order to calculate the frequency, the open degrees of the driver's eyes were calculated first, and the eyes' images were transferred into binary images, then the pupil widths on vertical direction of the binary images were calculated to judge the fatigue states of eyes. Experiments proved its validity.

Key words: traffic engineering; fatigue state detection; Haar-like cascaded classifier; eyes location; image processing

0 引言

司机疲劳驾驶已经成为潜在的交通隐患, 驾驶者的疲劳状况可以从眼睛的闭合频率和眼皮的张开程度反映出来。一般地, 驾驶者精神饱满时, 其眨眼频率和眼皮张开程度都有统计的正常值, 而当驾驶者处于疲劳状态下, 眼皮的张开程度较小, 眨眼频率也明显降低^[1]。因此, 利用眼睛的状态就可以判断出驾驶者的疲劳程度。在车辆中装入监控摄像头, 从监控录像

中截取图像, 对图像进行实时处理。为了达到实时的效果, 必须寻找一种快速收敛的算法进行眼睛定位以获取眼睛状态的信息。

本文采用普通 USB 摄像头作为输入设备, 利用 DirectShow 视频捕捉技术^[2], 使用基于 Adaboost 算法^[3]的层叠式分类器搜索图像中眼睛的位置, 得到包含眼睛的一个矩形区域。后面的处理就只对眼睛矩形区进行处理判断, 大大降低了运算量。

收稿日期: 2007-05-21

基金项目: 国家自然科学基金项目 (60472006)

作者简介: 林土胜 (1945-), 男, 广东化州人, 教授, 博士生导师, 研究方向为信号与图像信息处理、电路与系统的理论与技术。

(eetshlin@scut.edu.cn)

1 预处理

自然条件下获得的图像总是存在一些噪声, 在进行眼睛定位之前先对图像进行去噪处理。可采用中值滤波, 它对于干扰脉冲和点状噪声有良好的抑制作用; 本文采用 3×3 的中值滤波器, 经过中值滤波后的图像噪声得到了明显的抑制。

2 眼睛定位

为了得到眼睛图像的信息, 必须进行眼睛定位。现有的眼睛定位方法有很多, 基于肤色模型的方法^[4], 灰度投影法^[5], 基于模板的方法^[6], 基于神经网络的方法^[7], 基于特征提取^[8]等。本文利用 AdaBoost 算法构造一个基于类 Haar 特征的层叠式分类器来进行人眼的定位。类 Haar 特征^[9]通常由 2~4 个矩形组成, 如图 1 所示。

2-矩形特征表示两个形状大小相同的矩形区域内所有像素和的差值, 如 A, B; 3-矩形特征表示外面的两个矩形的像素和减去中间矩形的像素和; 4-矩形特征表示对角矩形内像素和之差。为了方便计算, 在这里引进积分图像概念:

$$ii(x, y) = \sum_{u < x, v < y} i(u, v) \quad (1)$$

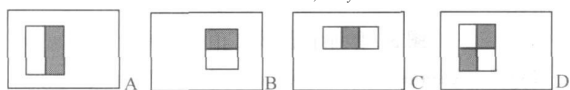


图 1 类 Haar 特征矩形

Fig 1 Haar-like feature rectangle

这样, 这些矩形特征就能用积分图像中的几个点表示^[10]。然而, 一幅图像的矩形特征远远比它的像素的个数要多, 必须有一个高效的分类器才能有效地筛选出有用的特征。Freund 等人在 Boosting 算法的基础上提出了 AdaBoost 算法^[3], 其主要思想是给定一弱学习算法和一训练集, 通过训练后将弱学习算法提升为强学习算法。这样只要找到一个比随机猜测好一点的弱学习算法提升为强学习算法, 就不必直接去找通常情况下很难获得的强学习算法。利用 AdaBoost 算法, 从所有的矩形特征中选择一个, 训练出一组比随机猜测略好一点的分类器, 加权组合成一个层叠式分类器。层叠式分类器是一个退化的决策树结构^[10], 如图 2 所示。

假设每个分类器可识别 50% 的非目标模式, 且目标模式的错分概率为 0.2%, 则 $h = 0.998$, $f = 0.5$ 。如果 $N = 12$ 那么最终识别正确率约为 97.6%,

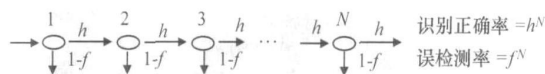


图 2 层叠式分类器结构

Fig 2 Schematic diagram of the cascaded classifier

误检测率约为 0.024%。在算法的训练过程中, 最好的分类器不断地被加进来, 直到达到所需的误检测率为止。

Viola 等的试验结果表明^[10], 在人脸检测过程中, AdaBoost 算法选出的第 1 个特征是“眼睛所在的矩形区域比眼睛下面的脸颊区域的灰度值小”, 第 2 个特征是“两只眼睛的灰度值比中间鼻梁的灰度值要小”。

在训练分类器时, 需要很大数量的目标图像集合和非目标图像集合来构造正、负样本库。具体的步骤如下:

(1) 构造正负样本图片库。

(2) 每一层训练 Boosting 训练: 从所有特征中选取加权后整体错误率最小的弱分类器; 通过样本的训练进行权重更新, 直到达到需要的检测率和误判率。

(3) 将每一层训练好的强分类器串联起来组成层叠分类器, 用于目标检测。

此算法的大部分时间用来训练检测具体目标的分类器, 分类器训练好后用来检测目标能够达到实时性要求。因此利用此方法来检测驾驶者的眼睛状态能够达到很好的实时性。

在眼睛的检测过程中, 检测子窗口是固定的 (采用 22×5), 图像中人眼睛的尺寸是变化的。为了尽可能多的检测到各种尺寸的眼睛, 采用图像金字塔的扫描方式。将图像在一定范围内依次进行尺寸的缩放, 将缩放后的图像作为眼睛检测器的输入。

本文在 324×146 的原始图像上, 以 1.25 的因子依次进行图像的缩小, 进行 12 次之后原始图像就缩小为 22×10 。这个过程就相当于检测子窗口进行了从 22×5 到 320×73 上 12 个梯度的放大, 在 CPU 为 2 GHz, 512 MB 内存的计算机上, 检测一幅 324×146 的图像可达到 10 帧/s, 检测的效果如图 3 所示。可以根据输入图像的大小以及具体的运用, 在检测之前使用 Gaussian 金字塔分解对图像向下采样, 以降低运算量。

3 判断眼睛状态

截取出眼睛图像之后, 对眼睛图像进行分析。由于判断人眼疲劳状态的重要指标是眼皮的张开程度和眨眼的频率, 其实眨眼频率也是计算眼皮张开程度的

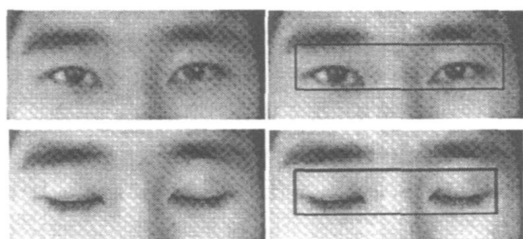


图3 眼睛检测

Fig 3 Results of eyes detection

变化频率,因此,计算眼皮的张开程度就成为判断眼睛状态的主要任务。

从先验知识可以得出眼皮张开的程度可以等价于眼皮之间瞳孔的大小,即垂直方向上瞳孔夹在眼皮之间可看见的直径大小。由于眼睛的特殊特点:眼睛瞳孔的灰度值比周围的小很多。介于此特点,可以通过二值化眼睛图像来分割出瞳孔。二值化的操作如下,将灰度值小于 T 的归到目标类,即为 0;灰度值大于 T 的归为背景,其灰度值都为 255。

$$g(x, y) = \begin{cases} 255, & g(x, y) > T, \\ 0, & g(x, y) < T. \end{cases} \quad (2)$$

二值化的关键是确定阈值 T , 阈值如果选择不好,就会影响目标物体和背景的错分率。在图像采集过程中由于光照不同的影响,最佳阈值可能不是同一个值。因此,本文采用自适应阈值法来确定 T , 避免了不同光照条件下使用同一阈值而增加错判率,具体的方法是使类间方差最大来分割物体。方差是灰度分布均匀性的衡量标准,方差越大,说明目标物体和背景的差别越大,越容易分割出来。当部分目标错分为背景或者部分背景错分为目标都会导致两部分的差别变小,因此,使类间方差最大就意味着错分的几率最小。

假设目标样本像素和与背景样本像素和分别占总体样本像素和的 w_0 、 w_1 , 且各自的平均灰度值为 g_0 、 g_1 , 图像整体平均灰度值为 g , 则两类方差表示为

$$w_0(g_0 - g)^2 + w_1(g_1 - g)^2. \quad (3)$$

为了减少计算量,可将该式化简,即将 $w_0 + w_1 = 1$ 和 $w_0g_0 + w_1g_1 = g$ 代入上式,可得到

$$w_0w_1(u_0 - u_1)^2. \quad (4)$$

求解使上式达到最大值的灰度阈值就是所求的二值化阈值,此方法在试验中达到了很好的效果。

图4是对图3中截取的眼睛图像进行自适应阈值法的二值化图像,可以看到,此方法能够有效地分割出目标物体。



图4 利用自适应阈值法的二值化图像

Fig 4 Binary images from adaptive threshold value method

状态中的两个极端。从这两幅图像可以看出,眼睛的张开程度可以以眼睛瞳孔的垂直宽度作为衡量。图5是眼睛不同状态的一系列图像,其中包括半开合的中间状态。

分割出目标物体之后,计算垂直方向两眼皮之间瞳孔的宽度就变得很容易了。从左到右扫描图像,计算垂直方向瞳孔的宽度,保存其最大值。由于眼睛的对称性,只需要扫描左眼或者右眼进行计算就行了,具体的算法描述如下:

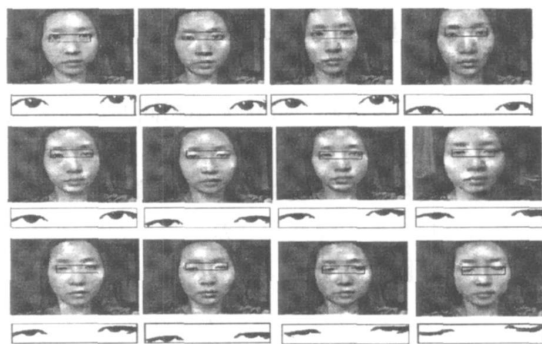


图5 眼睛检测及二值化结果

Fig 5 A series of eyes image detection and binary images

```
flag ← 0, max ← 0, eyeheight ← 0
```

```
for i ← image.left to image.left + image.width/2 // 扫描左眼
```

```
for j ← image.bottom to image.top // 计算垂直方向黑色线段的长度
```

```
if g(i, j) = 0 && flag = 0
```

```
point.x ← j // 记下黑色线段的起始位置
```

```
flag ← 1
```

```
if g(i, j) = 255 && flag = 1
```

```
point.y ← j // 记下黑色线段的结束位置
```

```
list.AddTail(point) // 将标记黑色线段的起始合结束位置的点加到列表中
```

```
flag ← 0
```

```
switch(list.GetCount())
```

```
case 0: // 列表空, 退出
```

```
break
```

```
case 1: // 列表中只有一个点, 即只有一段黑色线段
```

```

temp ← list.GetAt(0)
eyeheight ← temp.y - temp.x // 计算该线段的长度
if eyeheight > max // 记下整个左眼图像的最大值
max ← eyeheight
break
default: // 如果列表中有多个线段
mergeLine( ) // 合并间隔很小的线段
for k ← 0 to list.GetCount() - 1 // 计算各线段的长度, 保存最大值
temp ← list.GetAt(k)
eyeheight ← temp.y - temp.x
if eyeheight > max
max ← eyeheight
list.RemoveAll()
return max

```

从左到右扫描图像的过程中, 如果在垂直方向有多个黑色线段, 那么相邻两个线段间隔很小的话, 将它们合并成一个线段。如图 4 中的瞳孔由于光照问题出现了反光现象, 即二值化的图像中瞳孔中间出现白色的斑点。在计算垂直方向上黑色线段的长度时, 这些白色的斑点将一个黑色线段分隔成两个黑色线段, 这样有可能造成检测出来的黑色线段不是最长的瞳孔宽度。如果将相邻很紧的黑色线段合并就可以避免这样的问题, 合并相邻很紧的黑色线段可以改善反光的问题, 图 6 是上述算法检测的结果。

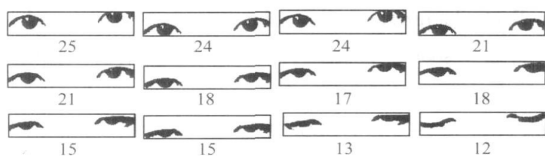


图 6 检测眼睛的张开程度

Fig 6 Detection of eyes' open states

4 结论

基于类 Haar 矩形特征检测目标的方法可以快速地检测到驾驶者的眼睛。由于类 Haar 矩形特征是利用灰度信息检测目标物体, 输入图像是灰度图像就可以了。虽然没有利用到颜色信息, 但是由于灰度图像是单通道, 减少了计算量, 提高了效率。本文提出了一种判断眼睛状态且易于实现的算法, 即利用可见瞳孔的大小计算眼睛张开的程度。另外本文回避了驾驶者戴深色眼镜的情况, 有待在以后的学习中改进。

参考文献:

References:

- [1] 王荣本, 郭克友, 储江伟, 等. 适用驾驶员疲劳状态监测的人眼定位方法研究 [J]. 公路交通科技, 2003, 20 (5): 111-114
WANG Rong-ben, GUO Ke-you, CHU Jiang-wei, et al. Study on the Eye Location Method in Driver Fatigue State Surveillance [J]. Journal of Highway and Transportation Research and Development, 2003, 20 (5): 111-114.
- [2] 陆其明. DirectShow 开发指南 [M]. 北京: 清华大学出版社, 2003.
LU Qi-ming. DirectShow Developing Guide [M]. Beijing: Tsinghua University Press, 2003.
- [3] TERRILLON J C, SHIRAZI M N, FUKAMACHI H, et al. Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images [C] // Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000: 54-61.
- [4] FREUND Y, SCHAPIRE R E. A Decision-theoretic Generalization of On-line Learning and An Application to Boosting [J]. Journal of Computer and System Sciences, 1997, 55 (1): 119-139.
- [5] KANADE T. Picture Processing System by Computer Complex and Recognition of Human Faces [D]. Kyoto University, 1973.
- [6] YUILIE A L, COHEN D S. Feature Extraction from Faces Using Deformable Templates [C] // Proceedings of IEEE Computer Society Conference on CRPR 1989: 104-109.
- [7] WAITE J, VINCENT J. A Probabilistic Framework for Neural Network Facial Feature Location [J]. British Telecom Technology Journal, 1992, 10 (3): 20-29.
- [8] LAM K M, YAN H. Locating and Extracting the Eye in Human Face Images [J]. Pattern Recognition, 1996, 29 (5): 771-779.
- [9] PAPAGEORGIOU C, OREN M, POGGIO T. A General Framework for Object Detection [C] // In International Conference on Computer Vision. 1998: 555-562.
- [10] VIOLA P, JONES M. Rapid Object Detection Using a Boosted Cascade of Simple Features [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Hawaii: 2001: 511-518.