

# 基于投影峰的眼睛快速定位方法<sup>\*</sup>

戴景文 刘丹 苏剑波

(上海交通大学 自动化系 上海 200240)

**摘 要** 提出一种基于投影峰的眼睛快速定位算法. 该算法首先根据人脸的几何分布特征设置眼睛候选窗口. 再通过直方图阈值分割出眼睛. 然后经过灰度积分投影得到一系列投影峰. 经过比较筛选得到眼睛的精确位置. 算法简单高效, 不需要训练和学习过程. 实验结果表明, 该算法能够快速准确地定位眼睛位置, 满足实时自动人脸识别的需要.

**关键词** 眼睛定位, 阈值分割, 投影函数, 投影峰

中图法分类号 TP 391.4

## Rapid Eye Localization Based on Projection Peak

DAI Jing-Wen, LU Dan, SU Jian-Bo

(Department of Automation, Shanghai Jiaotong University, Shanghai 200240)

### ABSTRACT

An eye localization method based on projection peak is presented. Firstly, according to the proportional relationships of face features, the eye region from an face image is segmented by setting appropriate candidate window. Then, by means of histogram analysis of the eye region image, a threshold is obtained and binary transform is performed to segment the eyes out of the eye region. Next, a series of projection peak is derived from vertical and horizontal gray projection curves of the binary image. Through comparison and analysis, the exact coordinates of the eyes is finally confirmed. The presented method has no need of any previous knowledge and training process. Experimental results on three face databases show that the proposed method is effective, accurate and rapid in eye localization and it meets the requirement of real-time face recognition system entirely.

**Key Words** Eye Localization, Threshold Segmentation, Projection Function, Projection Peak

<sup>\*</sup> 国家自然科学基金项目 (No. 60675041)、教育部新世纪优秀人才支持计划项目 (No. NCET-06-0398) 资助

收稿日期: 2008-04-18; 修回日期: 2008-12-22

作者简介: 戴景文, 男, 1983年生, 硕士研究生, 主要研究方向为人脸识别. E-mail: daijingwen@sjtu.edu.cn; 刘丹, 女, 1984年生, 硕士研究生, 主要研究方向为 лица 标定与识别. 苏剑波, 男, 1969年生, 教授, 博士生导师, 主要研究方向为机器视觉、机器学习与人机交互、多传感器信息融合和网络机器人等.

# 1 引言

人脸识别广阔的应用前景吸引很多研究人员开展大量研究并取得较大进展,已逐渐出现一些商业化的人脸识别系统.一个实用的人脸识别系统通常包括以下3个步骤:人脸检测与跟踪,面部关键特征点的定位与配准,人脸特征的提取与比对.

很多研究人员在做识别实验时都假定面部关键特征点是已给定的,或在条件允许的情况下要求用户进行一定程度的交互(如手动标定眼睛的位置).目前的主流人脸识别算法,都要依赖于面部特征(如眼睛位置)的严格配准来归一化人脸以便提取人脸描述特征,从而进行比对识别.因此人脸特征的误配准,对识别率有很大影响.文献[1]指出,目前较流行的人脸识别算法——Fisherface的识别率会随着误配准的增大而急剧下降,即使在只有1个像素平移偏差的情况下,识别性能也下降了10个百分点.同时,眼睛的定位也是其他特征点(如鼻子,嘴巴等)定位的基础,因此精确而快速的眼睛定位是实用人脸识别系统的关键环节之一.

眼睛等面部关键特征点的定位方法主要包括<sup>[2]</sup>,模板匹配、灰度积分投影法、Snake方法、可变形模板法、Hough变换、弹性图匹配、区域增长搜索方法、主动形状模型(Active Shape Model ASM)、主动外观模型(Active Appearance Model AAM)和AdaBoost<sup>[3]</sup>等.灰度积分投影是脸部特征领域的一类经典算法,它利用脸部特征灰度值较皮肤低的特点,先统计出 $X$ 或 $Y$ 方向上灰度值的和或灰度函数值的和,找出特定的变化点.然后根据先验知识将不同方向上的变化点位置相结合,得到脸部特征的位置.Kanade等人首先将这一方法应用于人脸识别<sup>[4]</sup>.文献[5]提出Generalized Projection Function文献[6]根据人脸图像的面部比例特征,提出一种基于候选窗口的灰度积分投影的方法来定位眼睛.灰度积分投影法计算量较低,但在大的姿态变化、光照变化或人脸有配饰(如眼镜)时容易失效.本文在灰度积分投影的基础上,提出一种基于投影峰的眼睛定位算法,在人脸姿态和光照变化及佩戴眼镜的情况下,可准确快速地实现眼睛定位.

## 2 基于投影峰的眼睛定位算法描述

本文算法是基于二值化后的人脸图像的,因此需要首先把窗口内的眼区图像截取出来,然后对眼

区图像通过阈值分割实现二值化.本节将对算法的各个步骤分别进行描述.

### 2.1 候选窗口的选取

人脸检测的方法有很多,如AdaBoost<sup>[3]</sup>等.在得到人脸的大致位置后,可将人脸图截取出来,然后在截取出来的图像上定位眼睛的位置.

人脸的五官位置具有一定的统计规律,如双眼一定位于人脸的上半部分,眉毛一定位于眼睛上面,双眼一定对称分布于正面人脸的中轴线两侧等.这些先验知识可帮助缩小眼睛的搜索区域范围,这样不但可去除一些干扰,而且还可提高搜索速度.

确定的眼睛区域需要对人脸的姿态发生偏转或旋转变化时有一定的鲁棒性.眼睛区域过小时,当人脸姿态发生变化,眼睛就不在候选窗口内.但如果窗口取得过大,将会引入过多的干扰,如眉毛,鬓角,刘海,眼镜框等,这些干扰的灰度值也相对较低,对通过使用灰度投影法来定位眼睛有很大影响<sup>[2]</sup>.

设人脸图像的宽度为 $width$ ,高度为 $height$ .实验中我们按照如下规则选择双眼候选窗口:左眼窗口原点为 $(width/12, height/12)$ ,右眼窗口原点为 $(11 \times width/21, height/12)$ ,双眼窗口大小均为 $(5 \times width/12, 5 \times height/12)$ .这种选择候选窗口的方法保证人脸姿态发生变化的情况下,眼睛依然在候选窗口内.但会引入少量的干扰(如眉毛,鬓角,刘海,眼睛框等),因此我们需要把这些干扰尽量去除.

候选窗口选定后,就可以把眼区截取出来.

### 2.2 阈值分割

图像阈值分割<sup>[7]</sup>是一种广泛使用的图像分割技术,其关键在于阈值的选取.常用的阈值选取方法有直方图阈值分割法、类间方差阈值分割法、二维最大熵值分割法等.为了把灰度值较低的眼睛分割出来,本文采用直方图分析的方法来选取阈值,令灰度值最低的 $p\%$ 个像素点的值为255,其他像素点的值为0.

### 2.3 灰度积分投影

理想情况下,眼睛的中心横坐标和纵坐标分别为垂直灰度投影曲线和水平灰度投影曲线峰值点对应坐标.但过大的候选窗口会引入干扰,如眉毛,刘海,鬓角,眼镜边框等,它们的灰度值可能甚至比眼睛瞳孔的灰度值还要低.另外由于光照的原因,使得眼区内会出现阴影,这同样会对灰度投影产生影响.所以不能简单地认为水平和垂直灰度投影的峰值点为眼睛的坐标值.

### 2.4 投影积分峰分析

为解决上述问题,首先需要分析可能存在的干

扰. 水平方向从上到下眼区内可能存在灰度值较低的点依次为刘海、眉毛、眼镜上框、眼睛瞳孔、眼镜下框. 垂直方向从左到右眼区内可能存在灰度值较低的点依次为鬓角、眼镜外侧边框、眼睛瞳孔、眼睛内侧边框. 当然, 在阈值分割之后, 这些干扰点有时会同时出现, 有时个别出现, 有时甚至全部不会出现. 其次, 眉毛和瞳孔在眼区内的灰度相对较低, 而且和上面所说的干扰比起来, 所占的面积也较大. 所以通过选取合适的阈值, 总可以把它们分割出来. 如果眉毛较稀疏或颜色较淡, 可能分割不出来, 但瞳孔总是可以分割出来的.

以由阈值分割后的二值化图像得到的水平灰度投影曲线  $P(y)$  为例. 通过分析可以发现, 投影曲线上会呈现若干个峰, 记作  $[y_{11}, y_{12}]$ ,  $[y_{21}, y_{22}]$ ,  $\dots$ ,  $[y_{n1}, y_{n2}]$ , 其中  $y_{11}, y_{21}, \dots, y_{n1}$  为投影峰的上升沿,  $y_{12}, y_{22}, \dots, y_{n2}$  为投影峰的下降沿. 眼睛、眉毛和其他的一些干扰, 分别与这些峰相对应, 而且眼睛对应的峰宽度较宽, 面积较大且偏离图像中心的位置较小. 如果做垂直灰度投影, 可以得到同样结果.

根据这一特性, 我们提出一种基于投影峰的快速眼睛定位算法. 对于每个投影峰, 我们定义评价值  $U_n$  如下:

$$U_n = \alpha W_n + \beta S_n + \gamma D_n \tag{1}$$

其中,  $W_n, S_n, D_n$  分别为投影峰的宽度、面积和与图像中心的距离,

$$W_n = y_{n2} - y_{n1}, S_n = \sum_{y=y_{n1}}^{y_{n2}} P(y), D_n = \left| \frac{(y_{n2} + y_{n1})}{2} - Y_c \right|$$

$Y_c$  是图像沿  $y$  轴方向的中心,  $\alpha, \beta, \gamma$  为权重, 一般取  $\alpha = \beta = \gamma = 1/3$

对所有的投影峰求  $U$  值, 并排序.  $U$  值最大的峰可认定是眼睛所在区域对应的投影峰. 在此峰内求最大值点, 即可得到眼睛瞳孔中心位置的横坐标或纵坐标.

本文算法步骤如下 (以求眼睛中心纵坐标为例).

- step 1 对阈值分割出来的图像做水平灰度积分投影.
- step 2 消除积分曲线上的毛刺.
- step 3 搜索曲线上的投影峰.
- step 4 求取所有投影峰的  $U$  值, 排序,  $U$  值最大的投影峰对应眼睛所在区域.
- step 5 在确定的眼睛所处区间内, 求灰度积分值最大的点, 则该点对应眼睛瞳孔中心的纵坐标.

瞳孔中心横坐标的求取与上述步骤类似. 算法流程如图 1 所示.

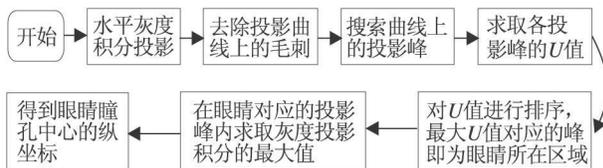


图 1 投影峰算法流程图

Fig 1 Flow chart for projection peak algorithm

### 3 实验结果与分析

为了测试本文算法的有效性, 用 C++ 编程实现上述算法, 进行实验. 实验中采用文献 [8] 给出的眼睛定位精度评价标准, 设  $C_1$  和  $C_r$  为手工标定的眼睛位置,  $C'$  和  $C'_r$  为自动定位的眼睛位置,  $d_1$  为  $C_1$  和  $C'$  间的欧式距离,  $d_r$  为  $C_r$  和  $C'_r$  间的欧式距离, 则眼睛定位的相对误差为

$$err = \frac{\max(d_1, d_r)}{\|C_1 - C_r\|}$$

如果  $err < 0.1$  则认为定位是准确的. 所以对于包含  $N$  张人脸图片的数据库, 定位的准确率为

$$rate = \sum_{\substack{i=1 \\ err_i < 0.1}}^N \frac{1}{N} \times 100\%$$

其中  $err_i$  为第  $i$  张人脸图片的相对误差.

#### 3.1 标准人脸库测试

FERET<sup>[9]</sup> 人脸库中共包含 14 051 张人头肩部的图像, 其中手工标定双眼坐标的图像有 3 816 张. BioID 人脸库<sup>[10]</sup> 中包含 1 521 张手工标定双眼坐标的人头肩部的图像. JAFFE 人脸库<sup>[11]</sup> 中包括 213 张头肩部人像, 这些人脸图像有较明显的表情变化, 该人脸库用于测试本文方法在不同表情下的定位效果. 在上述人头肩部的图像中通过 Adaboost 人脸检测把人脸区域截取出来构成眼睛定位方法的测试集.

测试结果如表 1 所示, 表中  $p$  为直方图阈值分割的阈值. 实验表明当  $p = 5$  时, 本文方法对 3 个标准人脸库的人眼定位准确率相对最高.

通过本文方法定位出双眼位置的部分人脸图像如图 2 所示. 图 2(a)、(d)、(f) 分别为 FERET、BioID 和 JAFFE 人脸库中部分正脸样本的定位的结果. 实验结果表明, 当光照、人脸表情发生变化时, 本文方法仍可准确定位出眼睛的位置. 图 2(c) 为 FERET 人脸库中部分侧脸样本的定位结果, 当人脸姿态发生变化时, 本文方法依然有效. 图 2(b)、(e) 分别为 FERET 和 BioID 人脸库中部分佩戴眼镜的人脸样本的定位结果. 结果表明, 本文方法可有效去除眼镜的

表1 标准人脸库上算法定位的准确率

Table 1 Localization accuracy on standard face database

	%		
$p$	FERET	BioID	JAFFE
1	77.34	70.12	84.21
3	91.56	84.10	89.63
5	97.03	93.92	98.78
8	94.72	92.23	96.54
10	93.61	84.82	96.12



(a) FERET 库内正脸样本

(a) Samples of frontal face in FERET



(b) FERET 库内戴眼镜的样本

(b) Samples with glasses in FERET



(c) FERET 库内侧脸样本

(c) Samples of side face in FERET



(d) BioID 库内正脸样本

(d) Samples of frontal face in BioID



(e) BioID 库内戴眼镜的样本

(e) Samples with glasses in BioID



(f) JAFFE 库内的样本

(f) Samples in JAFFE



(g) 部分错误定位的样本

(g) Samples with wrong location of eyes

图2 本文方法对不同样本进行眼睛定位的结果

Fig. 2 Some eye localization results of different samples by the

干扰,实现眼睛位置的精确定位.图2(g)为部分错误定位的样本,通过对这些样本的分析,发现当人脸受到强光照射,若样本图像中的人脸佩戴了眼镜,由于镜片的反光,使镜片区域的图像变白,遮挡了眼睛瞳孔,使其不完全可见,从而出现错误定位.

### 3.2 与其它方法的比较

文献[3]中提出的 Adaboost 是一种通用且有效的物体检测方法.我们为了得到用于检测眼睛的 Adaboost 分类器,选取 4 532 张  $20 \times 20$  的眼睛图片为正样本,每层取 2 236 张不含眼睛的图片为负样本,训练出 16 层的分类器.在上文设置的候选窗口内使用这个分类器作眼睛定位,并在上述的 3 个标准人脸库上测试,得到的实验结果见表 2.文献[5]方法和本文方法的测试结果也列于表 2 中.比较 3 种方法的实验结果可以发现,本文方法在保证较高定位准确率的同时,平均耗时也较少.此外本文方法不需要训练和学习过程,便于实现.

表2 3种方法的定位准确率及平均耗时的比较

Table 2 Comparison of localization accuracy and average time cost among 3 methods

	定位准确率/%			平均耗时 /ms
	FERET	BioID	JAFFE	
本文方法	97.03	93.92	98.78	0.56
文献[3] 方法	97.16	94.17	98.96	10.36
文献[5] 方法	94.52	90.12	96.27	0.49

### 3.3 实时系统测试

本文提出的眼睛定位算法,在 Intel P4 1.8G CPU,512MB 内存配置的 PC 机上,对于  $128 \times 128$  大小的人脸图像,定位双眼中心坐标耗时约 0.56ms.本文算法已应用于 RCIR@SJTU 自主开发的人脸识别系统上(图3).准确的人眼定位是人脸识别系统具有高识别率的重要保障.



图3 实时环境下的眼睛定位

Fig. 3 Real-time eye localization

## 4 结束语

本文算法充分利用人脸五官特征的统计规律, 通过简单的运算实现眼睛中心坐标的准确快速定位. 针对额发、鬓角、眉毛、眼镜等干扰, 提出基于投影峰的定位方法, 有效去除上述干扰, 使得定位算法更具鲁棒性. 此外当人脸姿态发生俯仰 ( $\pm 15^\circ$ )、旋转 ( $\pm 15^\circ$ ) 等变化时, 同样可实现精确定位. 通过研究误定位的样本发现, 当人脸受到强光照射, 若样本图像中的人脸佩戴了眼镜, 由于镜片的反光, 使镜片区域的图像变白, 使眼睛瞳孔不完全可见, 导致算法失效. 对于这种情况, 需要进一步研究解决.

### 参 考 文 献

- [1] Shan Shiguang, Chang Yizheng, Gao Wen, *et al.* Curse of Misalignment in Face Recognition: Problem and a Novel Misalignment Learning Solution // Proc of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004: 314–320
- [2] Lin Weikun, Pan Gang, Wu Zhaohui, *et al.* A Survey on Facial Features Localization. Journal of Image and Graphics, 2003, 8(8): 849–859 (in Chinese)  
(林维训, 潘纲, 吴朝晖, 等. 脸部特征定位方法. 中国图象图形学报, 2003, 8(8): 849–859)
- [3] Viola P, Jones M. Rapid Object Detection Using a Boosted Cascade of Simple Features // Proc of the IEEE International Conference on Computer Vision and Pattern Recognition, Hawaii, USA, 2001: 511–518
- [4] Kanade T. Picture Processing System by Computer Complex and Recognition of Human Faces. Ph. D. Dissertation, Kyoto, Japan, Kyoto University, Department of Information Science, 1973
- [5] Zhou Zhuhua, Geng Xin. Projection Functions for Eye Detection. Pattern Recognition, 2004, 37(5): 1049–1056
- [6] Li Genhua, Cai Xuanping, Li Xianshuai. An Efficient Face Normalization Algorithm Based on Eyes Detection // Proc of the IEEE/RST International Conference on Intelligent Robots and Systems, Beijing, China, 2006: 3843–3848
- [7] Gonzalez R C, Wood R E. Digital Image Processing, 3rd Edition. Upper Saddle River, USA: Prentice Hall, 2008
- [8] Jesorsky O, Kirchberg K, Frischholz R W. Robust Face Detection Using the Hausdorff Distance // Proc of the International Conference on Audio- and Video-Based Biometric Person Authentication, Halmstad, Sweden, 2001: 90–95
- [9] Phillips P J, Wechsler H, Huang H, *et al.* The FERET Database and Evaluation Procedure for Face-Recognition Algorithms. Image and Vision Computing, 1998, 16(5): 295–306
- [10] B i o I D. The B i o I D Face Database [DB/OL]. [2008-04-08]. <http://www.bioid.com/downloads/facedb/facedatabase.html>
- [11] Lyons M J, Akamatsu S, Kamachi M, *et al.* Coding Facial Expressions with Gabor Wavelets // Proc of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998: 200–205