

文章编号 1004-924X(2006)02-0320-07

一种应用于人脸正面图像的眼睛自动定位算法

李粉兰, 徐可欣

(天津大学 精密测试技术及仪器国家重点实验室, 天津 300072)

摘要: 基于人眼轮廓所具有的椭圆特性, 提出了一种眼睛定位算法。首先对人脸图像边缘提取, 并对边缘所有点进行最小平方椭圆拟合, 根据眼睛宽度与高度比例, 给定一个阈值, 确定眼睛候选区域; 然后运用训练过的支持向量机(SVM)分类器验证眼睛区域; 最后在眼睛区域, 采用一种基于方向性圆形霍夫变换(Circle Hough Transform (CHT))的检测算子定位眼球中心, 并对 ORL 及自建人脸库进行了测试。结果表明: 分别在 1.25 s, 3.0 s 内, 验证率为 91.6%, 95% 的情况下, 该方法可以获得 7.26 和 3.45 个像素的平均定位误差, 同时也证实了该方法对人脸的平面内旋转、尺度、表情等变化具有较强的鲁棒性。

关键词: 边缘提取; 椭圆拟合; 支持向量机(SVM); 检测算子

中图分类号: TP391.4 **文献标识码:** A

An algorithm applied in frontal-view face images for automatically localizing eyes

LI Fen-lan, XU Ke-xin

(State Key Laboratory of Precision Measuring Technology and Instruments,
Tianjin University, Tianjin 300072, China)

Abstract: Based on the ellipse-like outline of human eyes, an algorithm for localizing eyes was presented. After edge extracting from the face image, least square ellipse fitting was performed on all points on the edges. According to the proportions of eyes width and height, a threshold was given to confirm eyes candidate regions. A trained Support Vector Machine (SVM) classifier was applied to identify eyes regions. Finally, a circle detection operator based on directional Circle Hough Transform (CHT) was used to localize the eyeball center in the regions resulting from SVM classifier. The experiments were carried out on ORL and self-built face database. The results indicate that the method can accomplish the localization in 1.25 s and 3.0 s respectively, while accurate identification can reach 91.6% and 95%, the average localization errors are 7.26 and 3.45 pixels respectively. At the same time, the results also demonstrate that the method is very robust to face's in-plane rotation, scale and facial expression variation.

Key words: edge extraction; ellipse fitting; Support Vector Machine (SVM); detection operator

1 引言

眼睛自动定位是人脸识别^[1]中一个非常重要的课题,因为与人脸其他部件如鼻子,嘴巴相比,眼睛区域不仅包含了丰富的,区别于其他人的重要信息,而且是对图像进行旋转校正,归一化和均衡化的前提,同时也是人脸其他部位检测和提取的基础。目前人眼定位的算法有:可变形模板法,灰度投影法,边缘提取法^[2]等。可变形模板法虽能同时得到眼睛的位置和形状信息,但定位速度慢,而且其中包含的寻优过程是否收敛往往依赖于参数初始值的选择。灰度投影法是根据波峰波谷的分布信息来定位眼睛,定位速度较快,但波峰波谷的分布对不同的人脸和姿态的变化非常敏感,定位精度较差,并易陷入局部最小而导致定位失败。边缘提取法是首先对人脸图像进行边缘提取,然后用霍夫变换检测眼球,构造一个包括眼睛、眼睑的眼部模板,用一系列函数从能量角度找出眼睑,该方法需要做大量预处理,参数过多的眼部模板不适用于个体差异太大的人脸。

本文根据眼睛轮廓所具有的椭圆特性提出一种简单、易行的定位算法。首先对人脸图像进行边缘提取,并对边缘上的点进行椭圆拟合,因眼睛宽度和高度的比例是在一定范围之内的,因而可以对椭圆的长轴和短轴设定一个阈值,以确定眼睛的候选位置,再运用支持向量机分类器粗略定位眼睛区域,最后在该区域内应用基于方向性圆形霍夫变换(Circle Hough Transform (CHT))的检测算子来定位眼球中心。该方法对图像平面内旋转、尺度、表情的变化具有较强鲁棒性。通过对不同的人脸库进行实验,得到了令人满意的结果。

2 眼睛检测

2.1 边缘提取

边缘提取的算子主要有 Robert、Sobel、Prewitt、LOG 和 Canny 算子,它们对人脸图像的边缘提取效果如图 1 所示。

从理论上讲,为了减少计算量及头发,背景对眼睛检测的影响,所选择的算子应具备:在所提取的边缘图像中,在包含眼睛轮廓的前提下,其他边

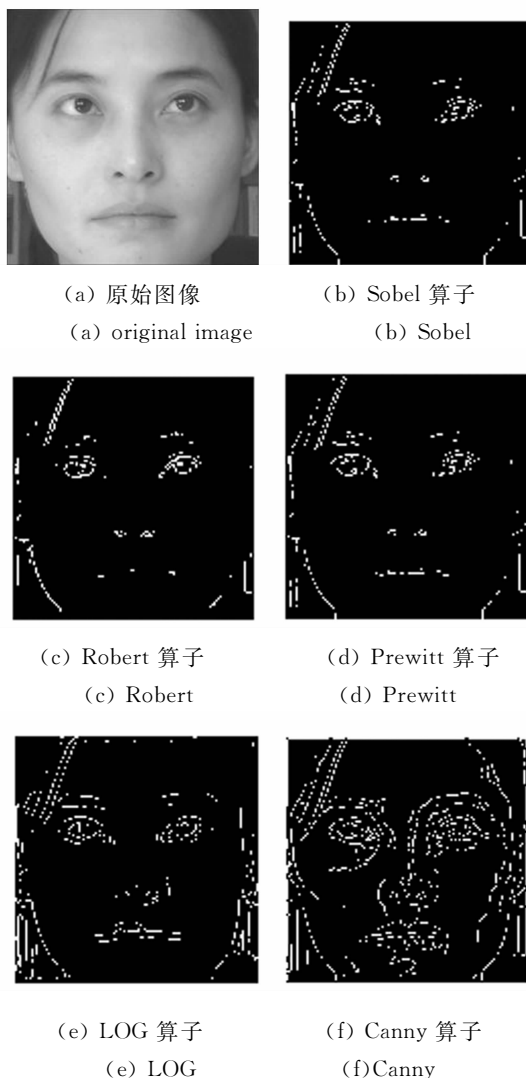


图 1 对原始图像使用不同边缘提取算子的结果比较
Fig. 1 Comparative results of each edge extraction operator on original image

缘点应尽可能的少。但确保真正的眼睛区域在眼睛的候选区域是实现眼睛检测的前提,所以为安全起见,本文选择眼睛区域边缘点较为丰富的 Canny 边缘检测算子进行边缘提取。

2.2 最小平方椭圆拟合

椭圆拟合的方法主要有两大类:聚类拟合和最小平方拟合,其中最小平方拟合最为常用。它主要是通过找到一组参数,使待拟合的数据点与椭圆之间的距离达到最小来实现的。最小平方拟合中,常用的是二次曲线拟合,一般来讲,二次曲线可以由一个二次多项式来表示:

$$F(G, X) = a \cdot x = ax^2 + bxy + cy^2 + dx + ey + f = 0, \quad (1)$$

其中: $G=[a b c d e f]^T$, $X=[x^2 x y y^2 x y 1]^T$, $F(G, X_i)$ 称为点 (x_i, y_i) 到曲线 $F(G, X)=0$ 的“代数距离”。椭圆拟合就是通过最小化 N 个数据点 x_i 到曲线的“代数距离”平方和来实现的:

$$D_A(G) = \sum_{i=1}^N F(X_i)^2, \quad (2)$$

为了避免平凡解 $G=0$, 或同一曲线存在多解问题, 需要对参数矢量 G 进行约束。Rosin^[3]和 Gander^[4]加入了 $a+c=1$ 的约束条件。Taubin^[5]将 G 约束成二次: $G^T C G=1$, 其中 C 是一个 6×6 的限制矩阵。Bookstein^[6]认为如果对参数进行二次约束, 式(2)的最小化问题可以简化为特征值的求解问题:

$$H^T H G = \lambda C G, \quad (3)$$

其中, $H=[X_1, X_2, \dots, X_N]^T$ 称为设计矩阵。尽管对此做了大量的研究, 但都并不是专门针对椭圆的, 而是在需要进行椭圆拟合时, 采用二次曲线拟合方法, 或者迭代方法, 使估计值接近于椭圆状。为了保证式(2)有解的同时, 实现真正意义的椭圆拟合, 对参数矢量 G 进行约束时, 使得这些参数所表述的二次曲线逼近于椭圆。众所周知, 恰当的约束条件就是: $b^2 - 4ac < 0$ 。通过在约束条件中加入缩放因子, 可以得到一个等价的约束条件, 即: $4ac - b^2 = 1$ 。这是一个二次约束条件, 用矩阵形式 $G^T C G=1$ 可以表示成:

$$G^T \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} G = 1, \quad (4)$$

根据 Bookstein 的理论, 椭圆拟合问题简化为: $G^T C G=1$ 条件下, 最小化 $E = \| H G \|^2$ 。这里的 H , 前面已定义。引入一个拉格朗日因子 λ 并微分, 最小化问题可以简化为如式(3)所示的特征值求解问题, 只不过此时加入一个约束条件: $G^T C G=1$ 。在此条件下, 要使 $\| H G \|^2$ 最小, 只能产生一个解, 该解对应于一个唯一的椭圆。相对于其他拟合方法, 该方法计算量少, 拟合效果好, 而且对噪声具有一定的鲁棒性, 更重要的是, 它是专门针对椭圆的拟合。

眼睛轮廓所表示的椭圆, 其长度和宽度都是在一定范围之内的, 通过对自建人脸库测试, 得到

眼睛的宽度约为 40 个像素, 高度在 11~20 个像素之间。所以在对边图像进行椭圆拟合后, 对长轴和短轴设定一个范围, 筛选出符合条件的椭圆区域作为眼睛的候选区域。这不仅可以提高眼睛检测速度, 同时可以减少其他因素对检测结果的影响。图 2 是长轴 $\in [15 \ 50]$, 短轴 $\in [5 \ 30]$ 的椭圆区域(即眼睛候选区域)在边图像中的位置及其在原图像中的对应位置。

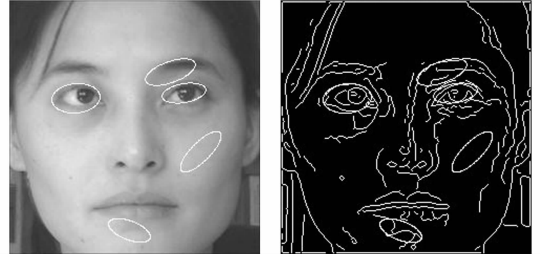


图 2 眼睛候选区域在边图像和原始图像中的位置
Fig. 2 Positions of eyes candidate regions in edge image and original image

2.3 支持向量机(SVM)

SVM^[7-8]是建立在经验风险最小的基础上, 通过找到一个最佳决策超平面来实现的两类分类的方法。假定存在一个可分的超平面, 它能将属于该类(正类)的样本与不属于该类(负类)的样本分开。对一组有标示的训练样本集 (x_i, y_i) , 其 x_i 是输入的特征矢量, y_i 代表其所属的类, 仅有两个值: -1 和 $+1$, 输入到 SVM 中, SVM 将会产生 N_s 个支持向量 s_i , 支持向量的权重因子 α_i , 支持向量的类标示 y_i 和一个常数 w_0 。对于线性可分情况, 超平面定义为:

$$w \cdot x + w_0 = 0, \quad (5)$$

其中 x 是超平面上一点, \cdot 表示点积运算。 w 是超平面的法向, w_0 是超平面到原点的距离。经过训练样本的训练之后, w 可表示成:

$$w = \sum_{i=1}^{N_s} \alpha_i y_i s_i, \quad (6)$$

给定 w 和 w_0 , 输入一个矢量 x , 就可以将其分到两类中的其中一类, 这取决于 $w \cdot x + w_0$ 是大于 0 还是小于 0。然而通常情况下, 类之间并不是线性可分的。在这种情况下, 可以通过核函数的非线性变换转换为某个高维空间中的线性问题, 再在变换空间中求最优分类面。常用的核函数主要有高斯核, 多项式核, 决策函数可以表示

成:

$$\sum_{i=1}^{N_i} \alpha_i y_i K(s_i, \mathbf{x}) + \omega_0 = 0, \quad (7)$$

2.3.1 SVM 的训练

为了能使用 SVM,需要一些训练数据来获得最优超平面。一幅眼睛图像可以用一个含有像素值的矢量表示。眼睛的训练集分成正图像集和负图像集,正图像集包含了不同姿态,不同人,不同表情下,戴眼镜、没戴眼镜的眼睛图像,非眼睛图像放在负图像集,图像大小是 45×70 。图 3 列出了训练集中一些眼睛和非眼睛图像。



图 3 训练集中的眼睛图像和非眼睛图像

Fig. 3 Eye images and non-eye images in the training set

本文采用的训练集中含有 500 幅正图像和 500 幅负图像。为了能达到最高的分类准确率,需要选择最佳的 SVM 参数,表 1 中列出了不同参数下的两种不同核函数的 SVM 分类器,对 100 幅眼睛候选图像验证了结果,从结果来看,使用高斯核函数, $\sigma=3$ 的情况下,可以获得最高为 95% 的正确率。

表 1 不同核函数和参数的 SVM 对 100 幅图像的验证结果

Tab. 1 Validation results of SVM with different kernel functions and parameters for 100 images

核的类型	幂	σ	准确度
Gaussian(高斯)		1	89%
Gaussian(高斯)		2	84%
Gaussian(高斯)		3	95%
Gaussian(高斯)		4	94%
Gaussian(高斯)		5	92%
Polynomial(多项式)	2		91%
Polynomial(多项式)	3		93%
Polynomial(多项式)	4		91%

2.3.2 眼睛验证

根据眼睛的候选区域,以其拟合椭圆的中心为中心,从原图像中分割出和训练图像同样大小的子图像并形成矢量,输入到训练过的 SVM 中, SVM 将会将输入矢量分到眼睛类或非眼睛类。

图 4 是 SVM 分类器正确辨别眼睛区域的情况,相对于图 2,已经去除了非眼睛区域。

SVM 对眼睛验证的结果非常好,尤其是对同一人种的人。对于与训练图像的人种大大不同的人来说, SVM 验证效果差,需要对 SVM 重新训练。

3 眼球中心定位

本文采用一种基于方向性 CHT 的检测算子来定位眼球中心。因为虹膜通常要比周围的巩膜黑,所以检测虹膜的边缘相对来说会比较容易。将虹膜的边缘视为一个圆。虹膜直径因人而异,而且随着人与摄像机距离的不同而不同,因此实验将圆半径限制在 $[R_{\min}, R_{\max}]$ 范围之内。一般眼球直径在 18~30 个像素之间,即 $R_{\max} = 15$, $R_{\min} = 9$ 。

该算子是通过掩模与梯度图像的卷积来实现圆形检测的。掩模^[7]如图 5 所示。它表示了每个点按其到圆环中心的距离进行折合的半径矢量的方向。掩模维数是 $(2R_{\max} + 1) \times (2R_{\max} + 1)$ 。将梯度图像与掩模卷积,可以估算出图像中有多少个点,它的梯度方向与圆环梯度方向一致。卷积结果的最大值给出了图像中半径在 $[R_{\min}, R_{\max}]$ 范围内的圆的中心位置。卷积过程可以表示成:

$$u(x, y) = \iint_{D(x, y)} e(\alpha, \beta) \mathbf{O}(\alpha - x, \beta - y) d\alpha d\beta / 2(R_{\max} - R_{\min})^2, \quad (8)$$

其中:

$$D(x, y) = \{(\alpha, \beta) \in R^2 \mid R_{\min}^2 \leq (\alpha - x)^2 + (\beta - y)^2 \leq R_{\max}^2\}, \quad (9)$$

e 是梯度向量: $e(x, y) = [E_x(x, y) / |E|, E_y(x, y) / |E|]^T$, \mathbf{O} 是掩模所表示的矢量:

$$\mathbf{O}(x, y) = [\cos(\tan^{-1}(y/x)) / \sqrt{x^2 + y^2}, \sin(\tan^{-1}(y/x)) / \sqrt{x^2 + y^2}]^T, \quad (10)$$

使用梯度向量,是为了使卷积结果与每个点的梯度强度无关。掩模的矢量中包含了一个归一化因子,是为了保证在轮廓最为完整的圆的中心,卷积结果达到最大。在式(8)中,加入 $2(R_{\max} - R_{\min})^2$,是为了确保不管圆的半径为多大,最后的卷积结果都在 $[-1, +1]$ 之间。

检测算子分别与图 4 中左、右眼子图像的梯度图像进行了卷积,卷积结果最大时的 x, y 坐标

值就是左、右眼球中心的位置。眼球定位结果见图 6。

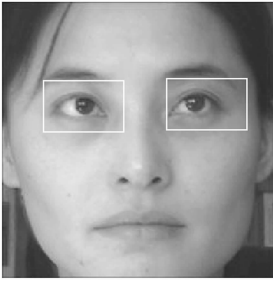


图 4 标有眼睛定位结果的原始图像
Fig. 4 Image marked with identified eyes

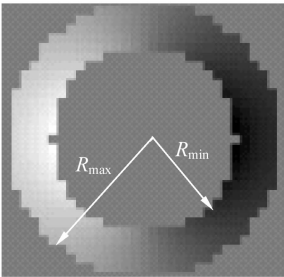


图 5 掩模
Fig. 5 Mask

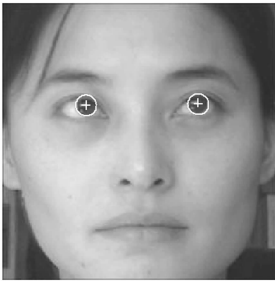


图 6 CHT 的眼球中心定位结果
Fig. 6 Localizing results of eyeball center by CHT

4 实验结果及分析

本文使用 ORL 和自建人脸库进行眼睛定位的实验。因为本文是根据眼球边缘是圆形这一特征来检测的,所以实验前,去除了库中眼睛眯拢厉害或者闭合的部分图像。

ORL 库中包含了 40 个人,每人 10 幅正面图像,图像大小为 112×92 ,图像中有表情,姿态和尺度变化,从中每人选取 6 幅图像进行测试。自建库包含了 10 个人,每人 10 幅正面图像,其中 3

人佩戴眼睛。图像大小是 410×420 ,图像是在同一视场背景,同一光照条件下,被摄对象不同姿态,不同表情变化的情况下拍摄完成的。对该库中所有图片进行了测试。

2,3 节中是针对自建库中的图像选择边缘检测算子及 SVM 训练的。因 ORL 库图像与自建库图像大小,人种相差悬殊,为了真实,充分地验证本文算法的定位效果,本文按照上述流程也为 ORL 库选择了相应的边缘检测算子(Canny),并对 SVM 进行了重训练(高斯, $\sigma=4$)。图 7 是两库中的图像及其定位结果实例。

从图 7 可以看出,本文的定位方法对姿态,表情,尺度,是否佩戴眼镜等都具有一定的鲁棒性。为了量化本文方法的定位准确度,将自动定位的结果与手工定位的结果进行了比较。首先引入两



(a) 自建库的定位结果
(a) Results of self-built face database



(b) ORL 库的定位结果
(b) Results of ORL face database

图 7 眼球中心定位结果实例(用“+”标志)
Fig. 7 Examples of eyeball center localization results (marked with “+”)

个概念:验证准确率和平均定位误差。验证准确率是指 SVM 分类器正确验证眼睛的图像数与总图像数之比。一幅图像验证正确与否是以验证区域内是否含有眼球为准。平均定位误差是指眼球自动定位坐标与手工定位坐标的欧式距离之和与总正确验证图像之比。设左右眼球中心位置为 E_l, E_r , 自动定位的位置是 E_l', E_r' 。 d_l 是 E_l 与 E_l' 之间的欧式距离, d_r 是 E_r 与 E_r' 之间的欧式距离。平均定位误差定义为:

$$\text{Errors} = \frac{\sum_{i=1}^N \sqrt{d_{li}^2 + d_{ri}^2}}{N} \quad (11)$$

其中, N 为正确验证的图像数。对两个库的眼球定位结果见表 2。

表 2 眼球中心定位结果
Tab. 2 Results of eyeball center localization

人脸数据库	验证结果	平均定位误差(像素)	定位时间(s)
ORL	91.6%	7.26	1.25
Self-built(自建)	95%	3.45	3.0

从两个库的实验结果来看,自建库的定位准确度明显优于 ORL 库,这可能与 ORL 库的图像质量,对比度差有关系。因为自建库的图像维数较大,所以定位时间较长。

实验中,SVM 在对一幅图像进行验证时,可能会出现以下两种情况:(1)只产生一个包含眼球区域;(2)产生两个以上包含眼球区域。针对第一种情况,可在这个区域确定眼球中心后,根据眼睛对称性,确定另一眼球的中心,具体见文献[9],在第二种情况中,即一个眼睛同时处于两个或两个以上的验证区域,此时应该对所有的验证区域都使用圆形检测算子定位眼球中心,找出它们之间距离最大的作为最终的两眼眼球中心位置。

5 结 论

为了实现眼睛的自动定位,本文根据眼睛轮

廓所具有的椭圆特性,对人脸图像的边缘图像进行椭圆拟合,以确定眼睛可能所在的区域。使用训练过的 SVM 对候选区域进行验证,最后使用基于 CHT 的检测算子定位眼球中心。实验结果证实该方法不仅能快速实现眼睛定位,而且可以获得高达 95% 的验证率,平均定位误差在 3 个像素左右。另外该方法对姿态、尺度、表情等变化都具有一定的鲁棒性。

该方法同其他眼睛定位方法相比具有以下优点:1) 相对于可变形模板法,该方法无需设定参数,整个定位过程简单易行;2) 相对于灰度投影法,该方法对姿态不敏感,且不受头发,眉毛等部位的影响;3) 相对于边缘提取法,它无需大量的预处理,无需建立模板;4) 相对于文献[7]提出的直接对人脸图像进行圆形检测,该方法不仅避免了对整幅图像的扫描,节省定位时间,而且可以减少背景,人脸其他部位对定位结果的影响。

参考文献:

- [1] 梁毅雄,龚卫国,潘英俊,等. 基于奇异值分解的人脸识别方法[J]. 光学精密工程, 2004, 12(5):543-549.
LIANG Y X, GONG W G, PAN Y J, et al. Singular value decomposition-based approach for face recognition[J]. *Optics and Precision Engineering*, 2004, 12(5):543-549.
- [2] 闫丽,袁家虎,刘江,等. 光学子波用于图像处理[J]. 光学精密工程, 2000, 8(3):225-230.
YAN L, YUAN J H, LIU J, et al. Optical-wavelet transform application in image processing[J]. *Optics and Precision Engineering*, 2000, 8(3):225-230.
- [3] ROSIN P L. A note on the least squares fitting of ellipses[J]. *Pattern Recognition Letters*, 1993, 14:799-808.
- [4] GANDER W, GOLUB G H, STREBEL R. Least-square fitting of circles and ellipses[J]. *BIT*, 1994, 34: 558-578.
- [5] TAUBIN G. Estimation of planar curves, surfaces and non-planar space curves defined by implicit equations, with applications to edge and range image segmentation[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1991, 13(11):1 115-1 138.
- [6] BOOKSTEIN F L. *Fitting conic sections to scattered data*[M]. Computer Graphics and Image Processing, 1979, 9

(1): 56—71.

- [7] TAKEUCHI K, COLLIER N. Bio-medical entity extraction using support vector machines[C]. *Proc ACL 2003 Workshop on NLP in Biomedicine, Sapporo*, 2003.
- [8] JOACHIMS T. Learning to classify text using support vector machines:[Dissertation][C]. *Norwell, MA, USA: Kluwer Academic Publishers*, 2002.
- [9] ORAZIO T D, LEO M, CICIPELLI G, *et al.* An algorithm for real time eye detection in face images[C]. *Proceedings of the 17th International Conference on Pattern Recognition, Cambridge*, 2004.

作者简介:李粉兰(1978—),女,江苏人,博士研究生,主要研究方向为计算机视觉,图像处理与模式识别,E-mail:orchidli@126.com;

徐可欣(1956—),男,安徽人,博士,教授,博士生导师,主要研究方向为图像处理,血糖检测,牛奶成分检测。