

一种高性能近红外光人脸检测和眼睛定位算法^①

张昌明, 童卫青, 王燕群

(华东师范大学 计算机科学技术系, 上海 200241)

摘要:提出了一种新型的近红外光人脸检测和眼睛定位算法,该算法先用 AdaBoost 分类器进行人脸区域定位,然后在人脸区域内进行眼镜检测:若检测为不戴眼镜,使用基于数理形态学的 N-Quoit 滤波器进行眼睛精确定位;若检测为戴眼镜,采用“人脸可信值最大”准则进行眼睛定位。实验结果表明,此算法检测率高而且眼睛定位准确,有效的克服了镜片反射近红外光干扰眼睛精确定位的问题,同时达到了实时性的要求。

关键词:人脸检测;眼睛定位;眼镜检测;AdaBoost;N-Quoit 滤波器

A High-Performance Algorithm of Face Detection and Eye Localization Based on Near-Infrared Image

ZHANG Chang-Ming, TONG Wei-Qing, WANG Yan-Qun

(Department of Computer Science and Technology, East China Normal University, Shanghai 200241, China)

Abstract: This paper proposes a novel face detection and eye localization algorithm based on near-infrared (NIR) image. Initially, the AdaBoost algorithm is introduced here to detect face, and then detection of the presence of glasses is performed in the candidate face area. If glasses are not present in face area, N-Quoit filter will be used to locate eyes. Otherwise, eye location can be determined by the “Faceness” criterion. Experimental results demonstrate that our method is capable of achieving high accuracy and real-time performance, and is an effective solution to the issue of specular reflections of NIR lights on glasses in eye localization.

Keywords: face detection; eye localization; glasses detection; AdaBoost; N-Quoit filter

1 引言

近红外光图像人脸识别是人脸识别领域的一个新的研究热点。近红外光图像在很大程度上解决了光照变化对识别性能的影响,弥补了可见光人脸识别的不足。人脸检测作为所有人脸识别系统的一个前提步骤,它的性能直接影响整个人脸识别系统的性能。现有的多数人脸识别系统都要依赖于眼睛中心位置的严格配准来归一化人脸,以便提取人脸描述特征^[1]。因此,眼睛的精确定位具有极其重要的意义。

本文主要研究近红外光图像的人脸检测和眼睛定位算法。国内外诸多学者在这方面进行深入研究,而且提出了许多算法。Rowley 提出了基于神经网络的人脸检测方法^[2],此方法是一个行之有效的方法,但是无法达到实时性的要求。Osuna 提出了基于支持向量机的人脸检测方法^[3],仍无法达到实时检测的要求。

Viola 提出的基于 Haar 特征和 AdaBoost 算法的人脸检测方法^[4]是第一个实时人脸检测方法。此后的人脸检测研究基本都是基于此方法之上的,主要的改进有基于 MB-LBP 特征的人脸检测方法^[5]、实时多视角人脸检测方法^[6]等方面。

近红外光图像的眼睛定位方法主要有: Dowdall 使用人脸图像在水平方向和垂直方向的投影来确定眼睛的位置^[7]。Zhao 使用数理形态学的开运算进行瞳孔定位^[8]。王使用基于数理形态学的 Quoit 滤波器进行眼睛定位^[9],该方法并不适用戴眼镜时的情况。Li 使用基于 Haar 特征和 AdaBoost 算法的人眼检测器^[10]进行眼睛定位,考虑到了戴眼镜与不戴眼镜时的问题,在满足实时性的要求下取得了较好的结果。

本文在总结前人研究方法的基础上,提出了一种新的近红外光图像眼睛定位方法,此方法不仅可以有

① 收稿时间:2010-07-25;收到修改稿时间:2010-08-26

效的定位眼睛，而且达到了实时性的要求，在一定程度上解决了戴眼镜时眼睛定位的难题。该方法先对人脸区域进行眼镜检测，若检测为不戴眼镜，则采用基于数理形态学的 N-Quoit 滤波方法进行眼睛定位；若检测为戴眼镜，则采用“人脸可信值最大”准则进行眼睛定位。

本文有以下 3 个贡献：(1)针对近红外光图像具有“瞳孔高亮”的特征，提出了基于数理形态学的 N-Quoit 滤波器的眼睛定位方法。此方法相对王的 Quoit 方法^[9]，在速度上提高了很多。(2)为了进行人脸图像戴眼镜与否的预判断，提出了简洁高效的眼镜检测方法。(3)当人脸图像戴眼镜情况下，镜片反射近红外光干扰 N-Quoit 滤波器眼睛定位时，采用“人脸可信值最大”准则进行眼睛定位。

2 算法介绍

本文的算法由以下两个部分组成：人脸区域检测和眼睛精确定位。人脸区域检测部分采用基于 Haar 特征和 AdaBoost 算法的方法，眼睛精确定位采用基于 N-Quoit 滤波器的方法和“人脸可信值最大”准则。

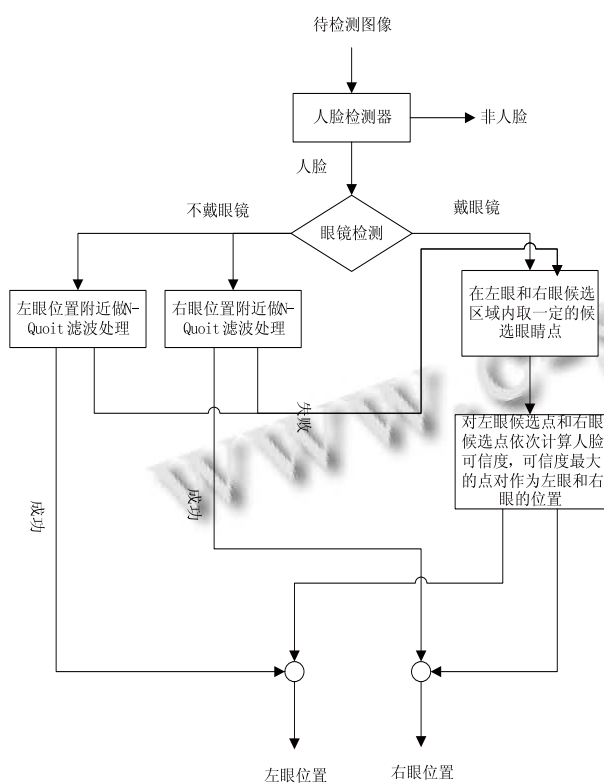


图 1 算法流程图

算法流程如图 1 所示：当检测到人脸区域后，采用眼镜检测器进行戴眼镜/不戴眼镜的预判断。如果检测为不戴眼镜，则采用 N-Quoit 滤波器分别在左眼和右眼区域进行眼睛精确定位。若眼睛定位失败，则转入“戴眼镜模式”的眼睛定位模块。如果检测为戴眼镜，则进入“戴眼镜模式”模块，采用“人脸可信值最大”准则进行眼睛定位。

3 基于 Haar 特征和 AdaBoost 的人脸检测

人脸检测部分，本文采用经典的基于 Haar 特征和 Adaboost 算法的方法^[4]。

3.1 Haar 特征

Haar 特征具有结构简单、计算快速的特点。Haar 特征的特征值定义为白色区域内像素总和减去黑色区域内像素总和。借助于积分图像^[4]，计算一个特征值只需要十几次加减运算即可完成。本文仅使用以下四类 Haar 特征(图 2)。A、C 类 Haar 特征可以提取垂直边缘，B 类特征可以提取到水平边缘，D 类可以提取到对角特征。这些特征用于表征图像的基本结构。

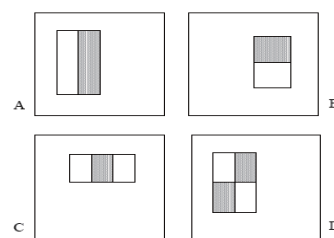


图 2 四类基本的 Haar 特征

3.2 级联 AdaBoost 算法

级联 AdaBoost 学习算法^[4]的主要过程如下：

输入：1) 每级的最大误检率 f ，每级的最小检测率 d ，级联分类器的目标误检率 F_{target}

初始化：1) 所有正例样本分成两部分，一部分构成训练用正例样本集 S^+ ，另一部分构成验证用正例样本集。2) 随机选取预定义数量的反例样本，构成训练用初始反例样本集 S_0^- 。3) $i=1$ ， $F_0=1.0$ ， $D_0=1.0$ 。（ F_i 表示前 i 级级联分类器在验证集上的误检率， D_i 表示前 i 级级联分类器在验证集上的检测率）

While $F_i > F_{target}$

Begin

(1) 初始化 $n_i = 1$

(2) 用当前训练集训练当前强分类器 H_i

a)根据 AdaBoost 算法步骤, 训练一个具有 n_i 个特征的强分类器。

b)调整本级强分类器阈值, 使当前生成的级联分类器在验证集上的检测率 D_i 不低于 $d \times D_{i-1}$ 。

c)用当前生成的级联分类器在验证集上测试误检率 F_i 。

d)若误检率 $F_i > f \times F_{i-1}$, 则增加一个弱分类器 $n_i = n_i + 1$, 重复步骤 a~c。

最终, 构成本级强分类器: H_i

$$H_i(x) = \text{sign}\left(\sum_{t=1}^{n_i} \alpha_t h_t(x) - \frac{1}{2} \eta_i \sum_{t=1}^{n_i} \alpha_t\right), \text{ 其中,}$$

n_i 是本级训练出的弱分类器个数, $\alpha_t = \log \beta_t^{-1}$ 是第 t 个弱分类器的权重, η_i 是阈值的松弛系数。

(3) 自举反例样本, 形成下一级反例样本训练集将训练用反例样本集 S^- 清空。从反例样本库中, 随机取出一个反例样本, 如果它被当前生成的级联分类器错误分类, 则将它添加到下一级的反例样本集中。重复此过程, 直到训练用反例样本集的数量达到预定义的反例样本数量。

End

3.3 AdaBoost 人脸检测器训练过程

根据上述级联 AdaBoost 算法, 人脸检测器训练过程如下。首先对训练用人脸样本进行归一化处理, 以消除图像背景和发型, 所有原始人脸图像都归一化到 24×24 像素。然后, 将归一化人脸样本分为两部分: 一部分作为训练用正例样本集, 数目为 8294; 另一部分作为验证用正例样本集, 数目为 8294。非人脸样本是从非人脸图像库中随机截取的, 大小同样为 24×24 像素。非人脸图像是指图像任何位置都不包含人脸的图像。归一化非人脸样本同样分为训练集和验证集, 每级训练用非人脸样本数目设定为 10000, 验证用非人脸样本数目为 24963192。

设定 $f=0.5$, $d=1.0$, $F_{\text{target}}=5.0 \times 10^{-4}$, 按照 3.2 节算法训练级联分类器。最终训练得到的级联分类器共 9 级, 每级的弱分类器个数分别为: 8, 16, 16, 30, 40, 40, 60, 100, 420。最终的级联分类器在验证集上的检测率为 1.0, 误检率为 4.04×10^{-4} 。

4 基于N-Quoit滤波器的眼睛定位算法

本章先介绍 Quoit 滤波器和 N-Quoit 滤波器的基本概念, 然后介绍基于 N-Quoit 滤波器的眼睛定位算法。

4.1 Quoit 滤波器的概念

Quoit 滤波器是一种数理形态学滤波器。从概念上说, Quoit 滤波器由两个滤波器(圆环滤波器和圆盘滤波器)叠加而成, Quoit 滤波器的输出值定义为圆盘和圆环做膨胀运算后的差值^[11]。对于一个孤立点, 圆环可以落下但是圆盘无法落下, 因此圆盘和圆环之间产生高度差(图 3), 所以 Quoit 滤波器可以检测到孤立点。

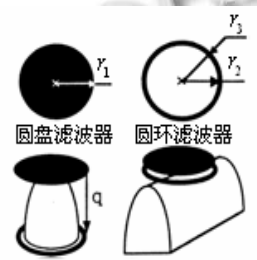


图3 Quoit 滤波处理

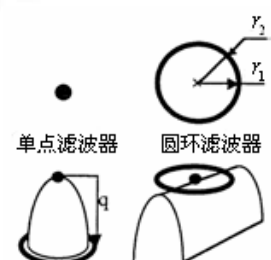


图4 N-Quoit 滤波处理

4.2 N-Quoit 滤波器的概念

当孤立点具有旋转对称性而且高度单调递减时, 要实现最大高度差, 圆盘内只有中心点起作用^[12]。因此, 单点滤波器可以取代圆盘滤波器(图 4)。其运算定义如下: $q = f - f \oplus R$ 。 f 为输入图像, R 是圆环结构元素, \oplus 为膨胀运算。因为单点滤波器就是图像本身, 不再需要膨胀运算, 所以 N-Quoit 滤波器的处理时间大幅缩短。

4.3 基于 N-Quoit 滤波器的眼睛定位

对于检测到的人脸区域, 使用 N-Quoit 滤波器在眼睛区域附近进行滤波处理, 得到滤波后的灰度图像, 然后对灰度图像进行二值化处理, 把白色区域的质心作为眼睛位置的备选点。若左眼和右眼区域各有一个备选点, 那么它们就作为左眼和右眼的位置。若左眼和右眼有多个备选点, 先利用一些假设进行筛选, 然后使用“人脸可信度最大”准则来确定最终的左眼和右眼位置。

图 5 为 N-Quoit 滤波器处理的结果, 从左到右为: 原始人脸图像、圆环滤波后的图像、两幅图像的差值图像、差值图像的二值图像。

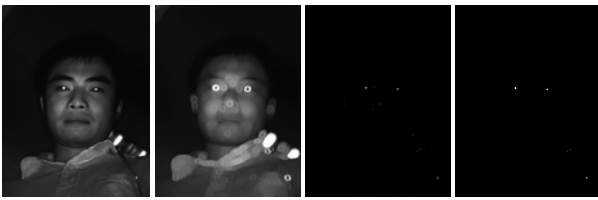


图 5 N-Quoit 滤波器处理结果

5 眼镜检测及戴眼镜时的眼睛定位算法

在近红外光图像中,理想的人脸图像具有明显的“瞳孔高亮”特征,使用 N-Quoit 滤波器可以容易的定位眼睛。但是,戴眼镜时的镜面反光给眼睛定位带来了很大的困难。因为戴眼镜时的反光区域有时也呈现孤立区域,它与“瞳孔高亮”特征极为相似,使得这些干扰点和真正的瞳孔点无法区分,这极大干扰了眼睛定位的精度。为了解决这个问题,本文提出了简洁高效的眼镜检测方法,它可以快速准确的识别出戴眼镜/不戴眼镜的状态,若检测到人脸是戴眼镜时,则采用“人脸可信值最大”准则进行眼睛定位,这极大的提高了眼睛定位算法的鲁棒性。

5.1 眼镜检测

眼镜检测的目标是判断人脸图像是否戴眼镜。眼镜的形状多种多样,但是眼镜在鼻梁上方的镜架横梁区域特征比较一致,因此本文选择该区域作为眼镜检测的特征区域。我们研究发现,近红外光图像中镜架横梁区域与皮肤呈现明显的灰度差异,而且镜架横梁区一般呈现水平方向。因此本文采用判断人脸图像鼻梁上方一定区域内是否存在眼镜横梁的方法进行眼镜检测。具体的算法流程如下:先检测人脸区域,然后对人脸区域鼻梁上方的一定区域内做边缘检测,再对边缘图像做二值化处理,然后对二值图像进行数理形态学处理来消除异物的干扰,最后判断出人脸是否戴眼镜的状态。流程图如下:

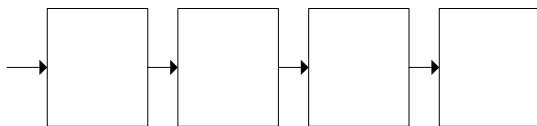


图 6 眼镜检测流程图

5.2 “人脸可信值最大”准则

戴眼镜时由于镜片和眼镜边框会反射近红外光,

使得瞳孔甚至眼睛部分不可见,这给眼睛定位带来了很大的困难。为了解决此问题,本文提出了“人脸可信值最大”准则进行眼睛定位。

人脸可信值 Faceness 定义如下:

$$Faceness(A, B) = \sum_{i=1}^N \sum_{t=1}^{n_i} \alpha_{it} h_{it}(x)$$

A, B 表示原始图像中的两个点, N 表示 AdaBoost 算法训练出来的分类器的级联数目, n_i 表示第 i 级强分类器中弱分类器的个数, α_{it} 表示第 i 级强分类器中第 t 个弱分类器的权重, $h_{it}(x)$ 表示第 i 级强分类器中第 t 个弱分类器。 x 表示原始图像中把 A, B 两点作为左右眼睛位置时,进行几何归一化后的图像。

从直观上来说,人脸可信值 Faceness 为每级强分类器的得分总和。当左眼候选区域(图 7 的 RA 区域)和右眼候选区域(图 7 的 RB 区域)内某两个点作为输入计算得到的 Faceness 最大时,表示此时的图像最像人脸,这两个点即可作为眼睛位置。其运算定义如下:如果

$$Faceness(Leye, Reye) = \max_{\substack{p_1 \in RA \\ p_2 \in RB}} \{Faceness(p_1, p_2)\}$$

则 $Leye$ 为左眼位置, $Reye$ 为右眼位置。其中 RA 为左眼坐标候选区域, RB 为右眼坐标候选区域。当检测到人脸区域后,即可确定 RA、RB 区域。 p_1 代表 RA 区域内的一个点, p_2 代表 RB 区域内的一个点。点 $Leye$ 和点 $Reye$ 代表使用“人脸可信值最大”准则确定的双眼坐标。此方法用全局的观点来确定眼睛的位置,后续的实验表明此方法有较好的效果。

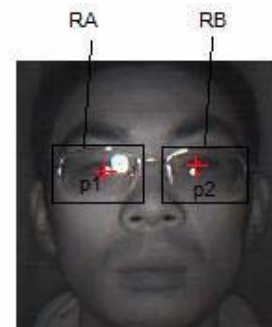


图 7 眼睛候选位置示意图

具体的眼睛定位方法如下:当检测到人脸区域后,即可确定左眼坐标候选区域 RA 和右眼坐标候选区域

RB。在 RA 区域内取 n 个点作为左眼坐标候选点，同样在 RB 区域内取 n 个点作为右眼坐标候选点，那么它们可以两两组成 n^2 个眼睛坐标点对。然后对这 n^2 个点分别计算 Faceness 值，并且计算 n^2 个 Faceness 值的最大值，此最大值对应的点对即为左右眼睛坐标。

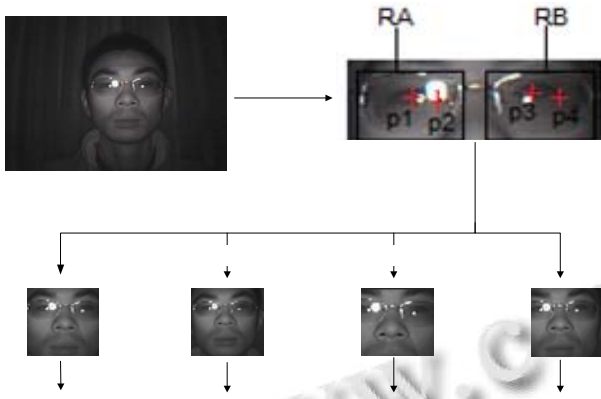


图 8 基于“人脸可信值最大”准则的眼睛定位示意图

图 8 表示基于“人脸可信值最大”准则的眼睛定位流程。左上角的图像表示原始人脸图像，右上角图像中的 p1、p2 点位于 RA 区域内，表示左眼候选坐标；p3、p4 点位于 RB 区域内，表示右眼候选坐标。那么候选眼睛坐标为 (p1, p3)、(p1, p4)、(p2, p3)、(p2, p4)，然后分别以它们为眼睛坐标进行几何归一化处理，并计算相应的 Faceness 值。图 8 下半部分的四幅图像为对应的归一化图像，它们下方的数值表示对应的 Faceness 值。可以观察到 (p1, p3) 对应的 Faceness 值最大，则 (p1, p3) 为最终的左右眼睛坐标。

6 实验结果 (p1, p3) (p1, p4)

本文算法使用 C++ 实现，开发平台为 Visual Studio 2005。

6.1 实验数据

为了说明并验证本文算法的有效性，我们在自建的近红外光人脸数据库上进行了实验。自建数据库是实验室采集的近红外光真实人脸图片（大小为 320×240 像素），共有 74 人，每个人分为 normal 和 glass 两个目录（normal 目录表示不戴眼镜，glass 目录表示戴眼镜），共 10171 幅图像。

6.2 实验方法及实验结果

使用通用的指标 $err^{[13]}$ 来定量的衡量算法的效果。假设 $d_l(d_r)$ 表示左眼（右眼）定位结果和真实左眼（右

眼）位置之间的欧式距离， D 定义为真实双眼位置之间的欧式距离，则 err 定义为： $err = \frac{\max(d_l, d_r)}{D}$ 。

以下实验为了说明并检验算法各个部分的作用而设计。

实验 1：为了单独检验眼镜检测算法的有效性，实验将全部 10171 个人脸样本作为输入进行人脸检测，实验仅对成功检测的人脸区域进行眼镜检测。实验结果如下：成功检测到的人脸总数为 10032（其中戴眼镜的个数为 5160，不戴眼镜的个数为 4872），戴眼镜的样本中被误判的个数为 149，不戴眼镜的样本中被误判的个数为 198，正确率达 96.5%。

实验 2：为了检验整个算法的效果，我们在以下三个数据集上进行实验。数据集 S1 中的图像为不戴眼镜的人脸图像，数目为 1429；数据集 S2 中的图像为戴眼镜的人脸图像，数目为 1373；数据集 S3 为 S1 和 S2 的并集，包括戴眼镜和不戴眼镜的样本，样本总数为 2802 个。在三个不同数据集上的测试结果如表 1 所示。

表 1 三个数据集上实验结果的比较

数据集	检测率 ($err \leq 0.25$)	平均误差 (像素)	运行时间 (ms)
S1	97.9%	1.3	24
S2	90.5%	5.8	44
S3	94.3%	3.5	34

数据集 S3 包含戴眼镜和不戴眼镜的人脸图像，比较符合真实的样本分布。在 CPU 为 2.1GHz 的 PC 机上，处理一幅图像的时间为 34ms，其中 AdaBoost 人脸检测耗时 21ms，眼睛定位耗时 13ms。眼睛定位结果和真实眼睛位置的平均误差为 3.5 个像素。在 $err < 0.25$ 的条件下，检测率为 94.3%；当 $err < 0.15$ 时，检测率为 88.4%。图 9 详细的给出了实验 2 的结果。图 10 为部分近红外光图像的实验结果。

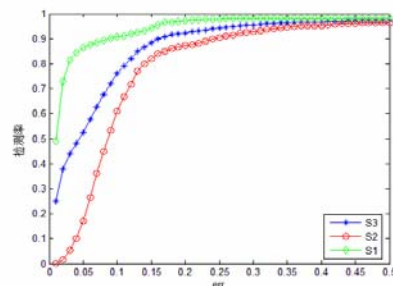


图 9 实验结果

6.3 实验分析

从上述的实验结果看, 本文的算法有较好的效果, 但是仍然存在不少定位不准或者判断错误的情况。AdaBoost 人脸检测时, 当眉毛区域被头发遮挡较多时容易发生漏检。眼镜检测的错误主要有: 头发等异物落入眼镜横梁区域上方时, 引起戴眼镜的误判断。有些眼镜的横梁区灰度与皮肤太接近时, 也会造成不戴眼镜的误判断。这些都是算法需要改进的地方。

从表 1 中可以看出, 本文算法在数据集 S1 上的效果优于在数据集 S2 上的效果, 前者的运行时间较短、且具有较低的平均误差。这主要是因为数据集 S1 中的全部图像都是不戴眼镜的人脸图像, 使用 N-Quoit 滤波器可以快速准确地定位眼睛; 而数据集 S2 的图像都为戴眼镜的人脸图像, Faceness 准则从全局信息出发, 眼睛定位时间相对较长。

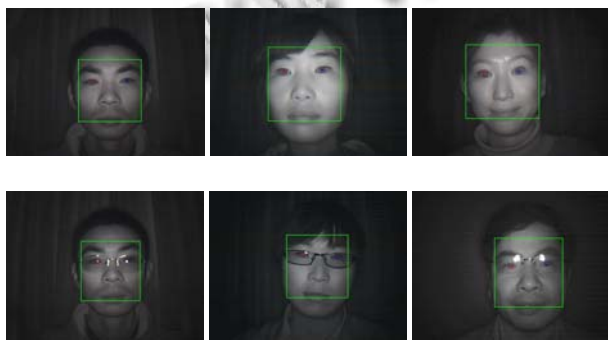


图 10 部分实验结果

7 结论

针对近红外光图像, 本文提出了实时高精度人脸检测和眼睛定位算法。该算法先对输入图像进行人脸检测, 然后对检测到的人脸区域进行眼镜检测, 若检测为不戴眼镜, 使用 N-Quoit 滤波器进行眼睛定位; 若检测为戴眼镜, 使用“人脸可信值最大”准则进行眼睛定位。实验结果表明, 此算法有较好的效果。

我们将在后续的研究中考虑加入其他面部特征点(例如鼻孔位置)来提高眼睛定位的精度, 以期能够得到更高性能的人脸检测系统。

参考文献

- 1 山世光, 高文, 唱轶钲, 曹波, 陈熙霖. 人脸识别中的“误配准灾难”问题研究. 计算机学报, 2005, 28(5): 782-791.
- 2 Rowley HA, Baluja S, Kanade T. Neural network-based face

- detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998, 20(1): 23-38.
- 3 Osuna E, Freund R, Girosi F. Training support vector machines: an application to face detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Puerto Rico: IEEE, 1997: 130-136.
- 4 Viola P, Jones M. Robust real time object detection. Second International Workshop on Statistical and Computational Theories of Vision. Vancouver: IEEE, 2001: 137-154.
- 5 Zhang L, Chu R, Xiang S, Li SZ. Face detection based on multi-block LBP representation. Proc. of IAPR/IEEE International Conference on Biometrics (ICB-2007). Seoul: IEEE, 2007: 11-18.
- 6 Huang C, Ai H, Li Y, Lao S. High-performance rotation invariant multiview face detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(4): 671-686.
- 7 Dowdall JB, Pavlidis I, Bebis GN. Face detection in the near-IR spectrum. Image and Vision Computing, 2003, 21(7): 565-578.
- 8 Zhao SY, Grigat RR. An automatic face recognition system in the near infrared spectrum. Proc. of Machine Learning and Data Mining in Pattern Recognition. Leipzig: Springer, 2005: 437-444.
- 9 王基帆, 童卫青. 基于数理形态学的近红外光图像实时人脸检测. 华东师范大学学报(自然科学版), 2010, 3: 39-47.
- 10 Li SZ, Chu R, Liao S, Zhang L. Illumination invariant face recognition using near-infrared images. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(4): 627-639.
- 11 Yamamoto S, Matsumoto M, Tateno Y, Iinuma T, Matsumoto T. Quoit filter: a new filter based on mathematical morphology to extract the isolated shadow, and its application to automatic detection of lung cancer in X-ray CT. Proc. of 13th International Conference on Pattern Recognition (ICPR'96). Vienna: IEEE, 1996: 2-7.
- 12 Miwa T, Kako J, Yamamoto S, Matsumoto M, Tateno Y, Iinuma T, Matsumoto T. Automatic detection of lung cancers in chest CT images by the variable N-Quoit filter. Systems and Computers in Japan, 2002, 33(1): 53-63.
- 13 Zhou ZH, Geng X. Projection functions for eye detection. Pattern Recognition, 2004, 37(5): 1049-1056.