

采用动态 HMM 概率分布模型的人眼精确定位

王洪群^{1),2)} 彭嘉雄¹⁾ 于秋则¹⁾

¹⁾(华中科技大学模式识别与人工智能研究所图像信息处理与智能控制教育部重点实验室, 武汉 430074)

²⁾(解放军信息工程大学电子技术学院, 郑州 450004)

摘要 从含有复杂背景的单幅灰度图像中精确定位人眼仍是一个尚未完全解决的复杂问题。为了对图像中的人眼进行精确定位,提出了一种基于动态 HMM 概率分布模型的人眼精确定位方法,该方法采用了含状态持续时间的动态观测符号概率分布 HMM 模型,首先用虹膜网格采样方式和特殊的特征提取办法来抽取观测序列;然后通过对观测序列进行评估来控制采样网格大小,并动态修正观测符号概率分布。这样无需对图像进行旋转、缩放和匹配运算,即可对图像中的人眼进行精确定位。实验结果表明,该方法检测效率较高、算法鲁棒,并具有较高的定位精度。

关键词 隐马尔可夫模型(HMM) 人眼定位 人脸检测

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2006)01-0026-07

A Precise Eye Localization via a Dynamic Probability Distribution HMM Model

WANG Hong-qun^{1),2)}, PENG Jia-xiong¹⁾, YU Qiu-ze¹⁾

¹⁾(Institute for Pattern Recognition & Artificial Intelligence, State Education Commission Key Laboratory for Image Processing & Intelligent Control, Huazhong University of Science & Technology, Wuhan 430074)

²⁾(Institute for electric technology, University of information technology of the PLA, Zhengzhou 450004)

Abstract Accurately locating the human eyes in a single and gray level image with cluttered background remain a difficult problem in recent year. In this paper, we use a discrete Hidden Markov Model(HMM) that included explicit state duration density to locate the human eye precisely. At first, a retinal sampling grid and a special method are adopted to extract the observation sequence. We do not need rotation and scaling transformation and matching operation, for we use a dynamic observation symbol probability distribution in state to adapt the various angles of human eyes in image, and a dynamic sampling model that controls the sizes of the retinal sampling grid through the evaluation of the observation sequence to adapt to the various sizes of human eyes in image. The experiment results show that our algorithm is effective and robust and have high positioning accuracy.

Keywords hidden markov model (HMM), localization of human eyes, face detection

1 引言

人脸局部特征的精确定位在实际应用中更具有意义^[1],因为人脸识别、表情分析、人机交互、双目视觉人脸 3 维重构中的配准要求等,都需要进行各种人脸局部特征的精确定位。人眼作为人脸最显著

的特征,与嘴、鼻相比,能够提供更可靠、更重要的信息,而静态图像的人眼的精确定位,一直以来都是一项具有挑战性的任务。目前有关单帧复杂背景灰度图像的人脸检测和识别已有很多报道^[2,3],其检测和识别方法大致可分为基于全局特征和局部特征的两大类^[4],而双眼的定位算法实际上也可分为以下两类:①通过人脸的全局特征在人脸模板中找眼睛

收稿日期:2004-06-03;改回日期:2005-03-26

第一作者简介:王洪群(1968~),男,解放军信息工程大学讲师。1997 年获西北农业大学工学硕士学位,现为华中科技大学博士研究生。主要研究方向为模式识别与人工智能等。E-mail:whq_lanzhou@163.com

的算法,如变形模板法^[5];②利用局部特征直接定位眼睛的算法,如文献[6]介绍的算法。另外,也可分为粗定位和细定位两类^[6]。本文算法可以直接定位人眼,当然也可以配合简单的人脸区域检测算法^[7,8],先找到人脸区域后,再用本算法来定位人眼。如今多数采用粗定位方法的文献对定位精度不太关心,只有少数文献,如文献[9,10]才使用眼睛的定位精度来衡量人脸的检测正确率。实验表明,本文提出的基于动态 HMM 概率分布模型的人眼定位技术,能在较高的定位精度和较小的运算量下,获得较高的检测率。

2 人脸眼部特征提取及模型描述

隐马尔可夫模型^[11] (hidden markov model, HMM)是一个统计分类模型。HMM 模型通常记为 $\lambda = (a, b, \pi)$,其中 a 为状态转移概率矩阵, b 为观测符号概率分布矩阵, π 为初始状态分布。HMM 模型适用于动态过程时间序列建模,并具有强大的时序模式分类能力,理论上讲,其可处理任意长度的时序。其最大的优点就是它的“平滑”效果,即所有的状态不需要精确的观测,即可实现模式分类。基于这些优点,本文将其应用于人眼的识别。下面详述如何将时序模式分类问题转化成人眼的 2 维平面图像的分类问题。

2.1 采用虹膜网格的采样

图 1 显示了针对左眼瞳孔位置的虹膜采样网格。从图 1 可以看出,如果采用虹膜网格采样模式,则能体现眼部及周围区域的变化特征。所谓虹膜网格,即离中心位置越近,采样间隔越小,采样密度越大;离中心位置越远,采样间隔越大,采样密度越小,而且从中心位置往外,采样间隔呈指数增长。这不仅符合人眼周围灰度由密到疏的变化情况,也正好适应了不同人脸图像眼睛特征变化不大,而周围区域变化随机性很大的实际情况。具体计算是在虹膜网格的节点位置上获得 $N \times 4N$ 个图像灰度值 $I(r_t, \psi_k)$, $t=1 \sim N, k=1 \sim 4N$,其中 r_t 为第 t 个虹膜网格节点的极坐标半径,它以指数形式增长,而在对数域上则等间隔增长,其计算公式为 $r_t = \exp\left(t \cdot \frac{\ln R}{N}\right)$, $t=1 \sim N$,共 N 个采样半径,其中 $R=r_N$ 为网格最大采样半径,它在后文中是可调节的; ψ_k 为第 k 个采样点方向角度,其计算公式为 $\psi_k = (k-1) \cdot 360^\circ / (4N)$,

$k=1,2,\dots,4N$,共 $4N$ 个采样方向,所以每个采样半径一次扫描 $4N$ 个采样点。

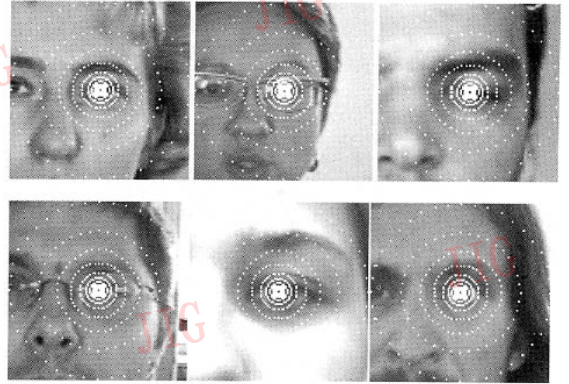


图 1 虹膜采样网格

Fig. 1 The retinal sampling grid

2.2 特征提取方式

对虹膜网格的每个半径进行扫描采样,并抽取该采样半径下采样点的一种特征,即观测符号。观测符号设计成 11 种符号类型,对应 11 种不同的采样特征。下面以图 2 为例说明符号的抽取方法。

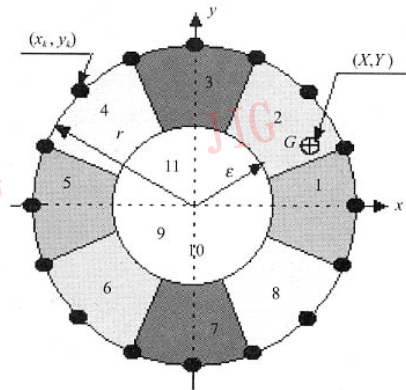


图 2 观测符号种类设计

Fig. 2 The design of the observation symbol

图 2 说明了对每个采样半径 $r=r_t$ 的观测符号种类的设计方法。图中黑点为采样点,共 $4N$ 个采样点。设同一采样半径下,第 k 个采样点的采样值为 $I(x_k, y_k)$,其中 x_k, y_k 表示第 k 个采样点的图像坐标。对采样值进行二值化,即如果 $I(x_k, y_k) > T$ (T 为阈值,取值方法见下文),则其二值化数据 $A_k=1$,否则 $A_k=0$ 。若已得到二值化采样数据 $A_k, k=1, \dots, 4N$,则观测符号的抽取按如下步骤进行:

(1) 在同一采样半径下, 计算采样值为 1 的采样点个数 $n = \sum_{k=1}^{4N} A_k$ 。如果第 t 个 $n \geq 2N$, 则第 t 个观测符号 $O_t = 9$; 如果 $n \leq 4N/10$, 则第 t 个观测符号 $O_t = 10$; 如果 $4N/10 \leq n \leq 2N$, 则进入下一步;

(2) 计算重心 G 点的坐标 X, Y

$$X = \frac{1}{n} \sum_{k=1}^n \{x_k | A_k = 1\}, Y = \frac{1}{n} \sum_{k=1}^n \{y_k | A_k = 1\}$$

和 G 点到原点的距离为 $\sqrt{X^2 + Y^2}$ 。设 ε 为一半采样点的 A_k 值连续为 1 时的重心 G 到原点的距离。如果 $\sqrt{X^2 + Y^2} < \varepsilon$, 即重心落入图 2 所示的白色部分, 则 $O_t = 11$; 否则, 转下一步;

(3) 计算 $\theta = \arctan(Y/X)$, 并判断重心 G 点落入图 2 所示哪个区域, 若落入第 $m (m = 1 \sim 8)$ 个区域, 则 $O_t = m$ 。

在 N 个半径上扫描采样, 并按上述各步抽取各自的观测符号类别就获得了长度为 N 的观测序列: $O = \{O_t, t = 1, \dots, N\}$, 其可作为提取的眼部区域的特征。

2.3 含有状态持续时间密度显函数 HMM 模型的设计

从图 1 可以观察到: 当以瞳孔位置为起始位置, 采样半径从小依次呈指数增大时, 灰度大小分布存在着某种变化规律, 即一般采样半径在瞳孔处的灰度值较低, 可能获得特征 $O_t = 9$; 在眼皮处常常对称分布, 则可能获得特征 $O_t = 11$; 再大一点, 可能的特征为 $O_t = 10$; 在眉毛处仅有一段灰度较低, 则 $O_t = m$; 在纯额部 (即眉骨上方不包含头发的部分), 采样的灰度值大部分都较高, 则可能 $O_t = 10$; 而到前额和头发位置时, 则又可能 $O_t = m$ 。这种变化规律致使可用 HMM。下面讨论如何建立这种模型:

首先, 若将虹膜采样半径经过的位置分为瞳孔、眼皮、眉、纯额部、额发部 5 种状态, 则 HMM 的状态转移模型如图 3 所示。根据观察可知, 采用含有状态持续时间概率密度显函数的隐马尔可夫模型更能反映状态转移的实际情况。

图 3 中, $S_1 \sim S_5$ 为设计的 5 个隐状态, 分别模拟 5 个不同半径下的采样位置, 所以要求 $N \geq 5$ 。图 1 及本文取 $N = 10$ 。 $a_{i,j}$ 表示状态转移概率, $p(i, d)$ 表示状态持续时间的概率密度显函数, $i = 1 \sim 5$ 表示状态标号, $d = 1 \sim d_{\max}$ 表示延长时间, 符号 d_{\max} 表示最大状态延迟时间, 而各状态的自转移系数都等于零。这种模型, 只有在一个状态已经产生出一定数

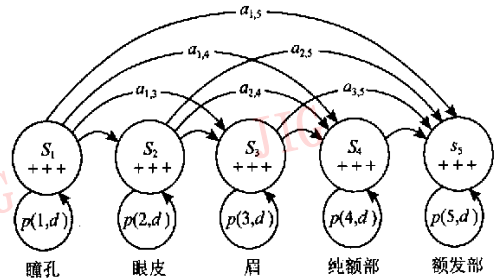


图 3 HMM 状态转移模型

Fig. 3 State transmit model of HMM

量的观测 (由状态持续时间密度确定) 之后才转移到另一个状态。

2.4 HMM 模型初始参数设计及参数重估问题

HMM 的根本问题是如何利用有限的训练数据来建造一个最佳的识别器, 即模型参数 $\lambda = (a, b, \pi)$ 的训练问题。经典方法是采用最大似然准则法 (ML) [11] 来进行训练模型参数的。其重估公式理论上可以给出与似然函数局部最大值相对应的 HMM 模型参数值。但是, 关键问题是如何选择 HMM 模型参数 π, a, b 合适的初始估计值, 以及状态持续时间概率密度显函数 $p(i, d)$, 以使局部最大值成为似然函数的全局最大值。其理论上如何选择, 目前基本上没有简单的答案 [11]。较为复杂的初始估计包括: ①人为地将观测序列分为一些状态, 并在状态内对观测进行平均和统计; ②将观测序列进行最大似然分段, 并进行平均; ③进行 K 均值分段和聚类等。可变持续时间 HMM 的所有参数的重估公式可参见文献 [11] 的第 269 ~ 270 页。参数训练时, 首先选择一些在大小和亮度上归一化的、无旋转的正面人脸, 然后确定合适的阈值 T 和 R , 最后在眼睛瞳孔处采样, 获得一些包含模型各状态的最佳的观测序列样本, 用于进行训练学习。对于旋转的人像问题, 本文后面将专门讨论。文献 [11] 中的重估公式是采用局部最大的最大似然估计法。应用统计决策理论设计的训练方法的前提是对先验概率和类概率密度函数有充分的先验知识, 或者有足够多的样本, 如果这些条件不满足, 则采用最优的方法训练出来的概率模型往往并不具有最优的性质, 有时甚至比采用简单方法设计的概率模型更差。对于本问题, 可通过对大多数人眼周围灰度值的几何分布特征进行观察来解决, 只要根据统计及经验设计的初始值能够接近全局最大就足够了。例如, 由于转移矩阵中, 从瞳孔位置跳转到眼皮的概率 $a_{1,2}$ 显然要

比从瞳孔位置直接跳转到额发的概率 $a_{1,5}$ 要大得多,所以,下面将直接给出参数的初始估计值:

(1) 初始转移矩阵设计为

$$a = \begin{bmatrix} 0 & 0.6 & 0.3 & 0.1 & 0 \\ & 0 & 0.6 & 0.39 & 0.01 \\ & & 0 & 0.5 & 0.5 \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

(2) 初始观测符号的概率分布设计为

$$b_{left} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0.1 & 0.8 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0.6 & 0.3 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$b_{right} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 & 0.2 \\ 0 & 0.1 & 0.8 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0.1 & 0.3 & 0.6 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

b_{left} 和 b_{right} 分别是左眼和右眼的正面无旋转的观测符号初始设计概率分布,其中行表示各状态,列表示观测符号类别。

(3) 初始状态分布设计为

$$\pi = [0.9 \quad 0.1 \quad 0 \quad 0 \quad 0]$$

(4) 状态持续时间不仅和人脸的大小有关,而且也最大虹膜网格采样半径 R 值的大小密切相关,即人脸越大,持续时间越大;同理,虹膜网格最大半径 R 越小,往往持续时间也越大。另外,在进行状态持续时间概率密度显函数的设计时,应该考虑以下 2 个近似条件:① $\sum_{i=1}^5 d_{max}(i) \approx N$;式中, $d_{max}(i)$ 表示使得第 i 个状态 S_i 的持续时间概率密度显函数 $p(i, d_{max})$ 最大的持续时间;② 从瞳孔状态 S_1 到额发状态 S_5 的 $d_{max}(i)$ 应该近似按指数速度递减。这样就可以人为地将观测序列样本分为一些状态后,再分段统计状态持续情况,并根据上述两个条件最后确定状态持续时间的概率密度显函数 $p(i, d)$ 。本文在 $N=10$,持续时间 $d=1 \sim 4$ 的情况下,将第 i 状态的持续时间概率密度显函数矩阵设计为

$$p = \begin{bmatrix} 0.3 & 0.1 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.8 & 0.2 & 0 & 0 \\ 0.9 & 0.1 & 0 & 0 \\ 0.8 & 0.2 & 0 & 0 \end{bmatrix}$$

其中,矩阵元素为 $p(i, d), i=1, \dots, 5, d=1 \sim 4$ 。

2.5 模型对图像亮度及人脸大小的适应性调整

2.5.1 对图像亮度和大小的适应

由于图像亮度的变化,经过上述采样计算过程所获得的观测序列受上述灰度阈值 T 的影响很大,即使采用人脸灰度统计模型中的阈值 $T=50$,也不一定能够获得很好的眼部周围特征的观测序列值。另外,人脸大小也是变化的,采用固定的最大虹膜采样半径 R 抽取的观测序列不一定能包含模型的几个状态特征。为了适应图像亮度变化,可采用 4 个灰度阈值: T 等于 50、60、70 和 80;为适应人脸大小变化,可采用 15、25、35 和 45,4 个 R 值来分别抽取观测序列,共获得 16 个观测序列。可通过设计以下评价准则来选择最佳观测序列:

准则 1 计算序列中观测值的变化次数,用 β 表示;求取观测序列的最大持续时间(即序列中相同观测符号连续重复出现的最大次数),用 γ 表示;则评价值 $E_1 = \beta/\gamma$;

准则 2 观测序列的前 1/4 长度的范围中至少有一个观测值是 9,则评价值 $E_2 = 1$,否则 $E_2 = 0$;

准则 3 观测序列中至少含有两个不相邻的观测值为 10 的观测,且第 2 个观测值为 10 的位置在观测序列的后 1/3 长度的范围内,则评价值 $E_3 = 1$,否则 $E_3 = 0$;

准则 4 在两个观测值为 10 的观测之间至少含有一个小于 9 的观测值(将靠近第 2 个观测值 10,且小于 9 的观测值记为 φ ,后文称 φ 为人脸旋转方向值),则评价值 $E_4 = 1$,否则 $E_4 = 0$ 。

而总评价值 E 可取几个评价值的权重和来获得:

$$E = \alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3 + \alpha_4 E_4$$

如权重可取 $\alpha_1 = 1000, \alpha_2 = 100, \alpha_3 = 10$ 和 $\alpha_4 = 1$ 。取上述 16 个观测序列中总评价值最大的观测序列来作为最佳观测序列 $O(T, R)$,其中 T 和 R 分别为获得的最佳观测序列的灰度阈值和最大虹膜网格半径。

为了获得包含各状态的更好的观测序列,先给最大虹膜采样半径 R 以很小的一个扰动微调(取 2 个像素),然后采样抽取观测序列 $O(T, R^{(-2)})$ 和 $O(T, R^{(+2)})$,分别计算它们的 HMM 输出值 P ,选取最大的作为该点的最终输出值:

$$H = \max \{ P(O(T, R^{(-2)}) | \lambda), P(O(T, R) | \lambda), P(O(T, R^{(+2)}) | \lambda) \}$$

λ 即 HMM 参数, 含状态持续时间的概率输出计算公式可参见文献[11]的第 269~270 页。

2.5.2 模型对头像旋转的适应

通过对人眼观测样本的分析可以看出, 状态 S_3 和 S_4 的观测并非完全独立, 而其他各状态之间是基本上接近独立的。这是由眼部灰度分布特征所决定的, 即眉毛和额发部位的采样值重心 G 在方向上存在着某种对应关系。例如, 图 2 的观测符号 1~8 对应 8 个方向, 对于无旋转的正面人脸, S_3 (即眉毛位置) 的观测值为 3, 在此观测方向及附近方向上的观测符号概率分布就很高 (见 b_{left} 的第 3 行), 而 S_5 (即额发位置) 的观测值为 1 或 2 的概率就很大 (见 b_{left} 的第 5 行), 而其他观测符号的概率就接近于零。如果人脸旋转了一定角度, 则 S_3 的观测的最大概率处就为其他值, 而 S_5 的最大概率处则也随其变化。对于左眼来说, 其方向则是将眉毛方向顺时针旋转 1 到 2 个方向。为了更好地反映这种特性, 本文设计了动态修正的观测符号概率分布, 即 S_3 和 S_5 的观测符号概率分布是根据实际的观测序列而定的, 并不采用重估值。由于观测的状态是隐态的, 必须附加人为的判断, 即在进行 HMM 的概率输出计算之前, 先扫描最佳观测序列 O 。当同时满足上述评价准则 3 和 4 时 (即 E_3 和 E_4 等于 1 时), 才使用修正的观测符号概率分布; 否则, 仍旧使用无旋转的观测符号概率分布。下面为左眼的观测符号动态修正办法: 表 1 中的符号 φ (人脸旋转方向值) 即为上文中在判断评价准则 4 时所记录的靠近第 2 个符号 10, 且小于 9 的观测值, 可将其作为眉毛的最大可能方向。根据此方向确定 S_3 和 S_5 的观测符号概率分布。在使用前向算法或后向算法迭代计算 HMM 输出时, 其计算公式并不需要大的改变。例如前向算法迭代公式中:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^5 \alpha_t(i) \alpha_{i,j} \right] b(j, O_{t+1})$$

$$1 \leq t \leq N-1, 1 \leq j \leq 5$$

其中, $b(j, O_{t+1})$ 为 j 状态的观测符号 O_{t+1} 的概率; 当 E_3 和 E_4 等于 1 且 $j=3$ 时, $b(3, O_{t+1})$ 的值根据 φ 值在表 1 中取得即可; 当 E_3 和 E_4 等于 1 且 $j=5$ 时, $b(5, O_{t+1})$ 根据 φ 值在表 2 中取得即可。

右眼的观测符号概率分布和左眼类似, 只是状态 S_3 和 S_5 观测符号概率分布的对应关系与左眼有所不同。如当 $\varphi=3$ 时, 状态 S_3 的观测符号概率分

布和左眼一样: $b_{\text{right}}(3, O_t) = [0, 0.1, 0.8, 0.1, 0, 0, 0, 0, 0, 0, 0, 0]$, 但状态 S_5 的概率分布变为 $b_{\text{right}}(5, O_t) = [0, 0, 0.1, 0.3, 0.6, 0, 0, 0, 0, 0, 0, 0]$ 。其他 φ 值按右眼的结构依次对应变化。

表 1 状态 S_3 观测符号概率分布 $b_{\text{left}}(3, O_t)$ (左眼)

Tab.1 The observation symbol probability distribution $b_{\text{left}}(3, O_t)$ in state S_3 (left eye)

φ	O_t										
	1	2	3	4	5	6	7	8	9	10	11
1	0.8	0.1						0.1			
2	0.1	0.8	0.1								
3		0.1	0.8	0.1							
4			0.1	0.8	0.1						
5				0.1	0.8	0.1					
6					0.1	0.8	0.1				
7						0.1	0.8	0.1			
8	0.1						0.1	0.8			

表 2 状态 S_5 观测符号概率分布 $b_{\text{left}}(5, O_t)$ (左眼)

Tab.2 The observation symbol probability distribution $b_{\text{left}}(5, O_t)$ in state S_5 (left eye)

φ	O_t										
	1	2	3	4	5	6	7	8	9	10	11
1	0.1						0.6	0.3			
2	0.3						0.1	0.6			
3	0.6	0.3	0.1								
4		0.6	0.3	0.1							
5			0.6	0.3	0.1						
6				0.6	0.3	0.1					
7					0.6	0.3	0.1				
8						0.6	0.3	0.1			

*注: 表 1 和 2 中空格中的数据为 0。

3 眼睛的定位

如果直接对原图像所有象素点进行计算, 则计算量较大, 不仅不利于实际应用, 而且亦无必要。为

提高运算速度,可分以下 3 步定位人眼:①使用文献[12]公式(1)形态学方法来找到潜在的人眼候选点;②分别使用上述左眼 b_{left} 和右眼 b_{right} 的观测符号概率分布参数来计算这些潜在的人眼候选点,并且仅计算原图像中灰度值小于最大阈值 $T=80$ 的点的上述模型输出: H_{left} 和 H_{right} ;③分别求取 H_{left} 和 H_{right} 中的最大值的位置,以作为左右眼的位置。

4 实验结果

为验证本文算法的有效性,本文采用 BIOD 人脸图像库^[10],对 1521 幅大小为 384×286 ,由 23 个不同人构成的灰度图像集进行了测试。该图像集含有复杂的背景、变化的光照条件以及变化的人脸姿势、大小等。为便于进行不同算法的定位效果比较,使用 Jesorsky 等人的定位精度度量参数 d_{eye} ^[10] 来衡量定位精度。 d_{eye} 被称为相对误差,当 $d_{eye} = 0.25$ 时,相当于眼睛的一半距离。图 4 为用本文算法在整幅图像中直接定位人眼的测试结果和用文献[10]的 Hausdorff Search + MLP 方法进行测试结果对比。横轴为 d_{eye} ,纵轴为检测正确率。当 $d_{eye} < 0.1$ 时,本文左右眼都定位准确的检测正确率为 83%,高于文献[10]的 80%。当 $d_{eye} < 0.25$ 时,本文检测正确率为 89%,文献[10]为 91.8%。

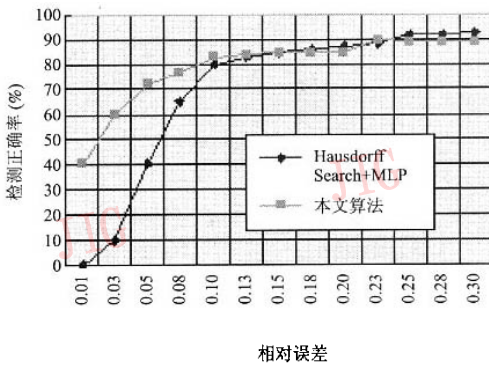


图 4 测试结果比较

Fig. 4 The comparison of detecting results

图 5 是 BIOD 图像库中的一幅精确定位的图例,可以看出图中人脸有较大的旋转。图 6 为定位失败的图像。

在主频为 850Hz 的奔腾 III 个人计算机上,本文算法对 BIOD 图像进行定位的运算时间为平均每幅 35ms。

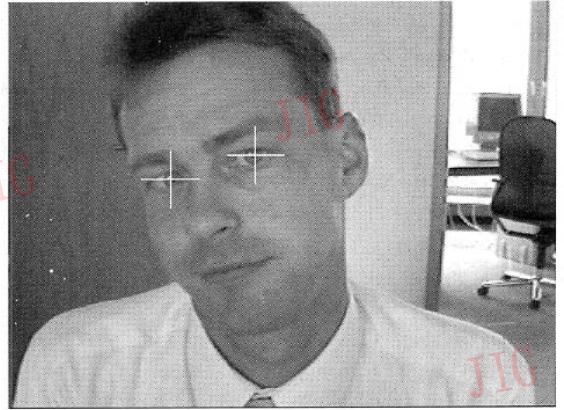


图 5 精确定位图例

Fig. 5 A result that located the eyes precisely



图 6 定位失败图例

Fig. 6 A result that failed to locate the eyes

5 结论

实验结果表明,本文算法在定位平均相对误差较小时,有较高的检测正确率,即具有较高的人眼定位精度,同时具有较快的运算速度。由于算法能够自适应图像亮度、人脸大小和旋转的变化,因此有较高的鲁棒性,适合于实际应用。另外,如果结合人脸区域分割算法,则可以进行多人脸的眼睛精确定位。

参考文献 (References)

1 LIN Wei-xun, PAN Gang, WU Zhao-hui, et al. A survey on facial features localization [J]. Journal of Image and Graphics, 2003, 8(8):849~859. [林维训,潘刚,吴朝晖等.脸部特征定位方法[J].中国图象图形学报,2003,8(8):849~859.]

- 2 YANG Ming-hsuan, Kriegman David J, Ahuja Narendra. Detecting faces in images: A survey [J]. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 2002, **24**(1):34 ~ 58.
- 3 Samai A, Iyengar P A. Automatic recognition and analysis for human faces and facial expressions; A survey [J]. *Pattern Recognition*, 1992, **25**(1):65 ~ 77.
- 4 WANG Lei, MO Yu-long, QI Fei-bu. Automatic localization of the center of IRIS[J]. *Journal of Infrared Millim Waves*, 1998, **17**(5): 349 ~ 356. [王磊, 莫玉龙, 戚飞虎. 眼球的自动定位[J]. *红外与毫米波学报*, 1998, **17**(5):349 ~ 356.]
- 5 AI Hai-zhou, LIANG Lu-hong. Face detection based on skin color and template[J]. *Journal of Software*, 2001, **12**(12):1784 ~ 1792. [艾海舟, 梁路宏. 基于肤色和模板的人脸检测[J]. *软件学报*, 2001, **12**(12):1784 ~ 1792.]
- 6 TAO Liang, ZHUANG Zhen-guan. Locating human eyes in complex background [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2003, **15**(1):38 ~ 42. [陶亮, 庄镇泉. 复杂背景下人眼自动定位[J]. *计算机辅助设计与图形学学报*, 2003, **15**(1): 38 ~ 42.]
- 7 WANG Hong-qun, PENG Jia-xiong, QIANG Zan-xia. A fast and precise method of human face localization based on the edge and texture feature[J]. *Computer Engineering and Applications*, 2004, **40**(7):27 ~ 32. [王洪群, 彭嘉雄, 强赞霞. 基于边缘和纹理特征相结合的快速人脸精确定位方法[J]. *计算机工程与应用*, 2004, **40**(7):27 ~ 32.]
- 8 YANG Guang-zheng, Huang Thomas S. Human face detection in complex background [J]. *Pattern Recognition*, 1994, **27**(1): 53 ~ 63.
- 9 Gloria Chow, LI Xiao-bo. Towards a system for automatic facial feature detection [J]. *Pattern Recognition*, 1993, **26**(12): 1739 ~ 1755.
- 10 Jesorsky Oliver, Kirchberg Klaus J, Robert W F. Robust face detection using the Hausdorff distance [A]. In: J. Bigun, F. Smeraldi, ed. *Proceedings Third International Conference on Audio- and Video-based Biometric Person Authentication, Lecture Notes in Computer Science, LNCS-2091* [C], Halmstad, Sweden: Springer-Verlag, 2001:90 ~ 95.
- 11 Rabiner I. R. A tutorial on hidden Markov models and selected applications in speech recognition [J]. *Proceedings of the IEEE*, 1989, **77**(2):257 ~ 286.
- 12 HAN Chin-chuan, LIAO Hong-yuan Mark, YU Cwo-jong, *et al.* Fast face detection via morphology-based pre-processing [J]. *Pattern Recognition*, 2000, **33**(10):1701 ~ 1712.