

眼球的自动定位*

王磊 莫玉龙** 戚飞虎

(上海交通大学计算机科学与工程系, 上海, 200030)

TP391.41

摘要 提出一种由粗到细自动定位人脸的一个重要特征即眼球圆心的方法. 该方法综合运用局部二值化、霍夫变换、像素聚类、边缘提取等技术, 对均匀或非均匀光照下拍摄的或戴眼镜的照片都能获得很高的定位精度, 适用于处理证件照.

关键词 霍夫变换、像素聚类、边缘提取、特征提取、人脸识别.

眼球定位

引言

由于人脸识别技术在保安和商贸系统等领域有着日益广泛的应用需要, 自九十年代开始再度成为研究热点^[1]. 当前的识别方法主要分为基于局部特征^[2,3]和基于全局特征^[4~6]的两大类. 无论哪类方法, 对基本特征如眼睛、鼻子和嘴巴等的准确定位都是必不可少的步骤, 尤其是眼球圆心的准确定位. 这是因为两眼间距受光照或表情变化的影响最小, 所以常被用于几何特征或图像尺寸的归一标准. 在精确定位眼球中心前一般要先找到眼睛的大致位置, 谓之粗定位, 常见的方法有: 神经网络法^[7]、边缘点积分投影曲线的极值位置判别法^[8]、对称性法^[9]、多分辨率的马赛克图法等^[10]等. 这些方法虽然各有特色, 但都不尽完善. 比如, 用神经网络方法总是需要大量充分考虑了各种变化情况的训练样本; 基于边缘的方法对受均匀光照且不戴眼镜的图像较有效^[8,11]. 然而实际使用的证件照片不可能都满足这些条件. 为此, 本文设计了一种基于局部二值化图像积分投影曲线和先验知识的粗定位方法. 至于眼球圆心的精确定位, 现有的若干种方法是: 基于霍夫变换^[12,13]、几何及对称性检测^[9]、弹性模板^[14,15]等. 由于霍夫变换是一种抗噪能力较强且算法简单的检测给定形状曲线的技术, 很适合用于精确定位眼球圆心. 在文献[12]中, 只用了6张实验室里拍摄的人脸像, 眼球附近没有其它明显的圆弧边缘干扰; 文献[13]也用了该方法, 为了排除眼球附近其它圆弧边缘的干扰, 加入了几个凭经验设置的约束条件. 以上实验都没有处理戴眼镜的照片. 因为镜片的反光点、镜脚与镜架的联结点及眼角阴影的边缘都可能是明显的圆弧线, 很难与眼球边缘线区分开, 若要凭经验找出区分它们的约束条件是很困难的. 本文提出的基于小灰度值聚类和用霍夫变换检测下半圆的方法无需设约束条件, 能准确定位实际使用的证件照的眼球圆心. 本实验用了51幅一寸免冠照, 其中8幅是戴眼镜的. 实验结果验证了本方法的正确性及有效性. 本方法的限制条件是要求人脸为正面, 基本布满整幅图像且基本位于正中.

* 国家自然科学基金(编号 69572026, 69475006)与上海市教委科技发展基金资助(编号 97AJ09)项目

** 上海大学(嘉定校区)通信与信息工程学院

稿件收到日期 1997-08-22, 修改稿收到日期 1998-02-17

1 眼球定位过程

眼睛在人脸上的位置有以下特点:(1) 在眉毛与鼻子的中间,离眉毛的距离小于离鼻孔的距离;(2) 两眼之间约有一只眼睛的距离;(3) 两眼对称于人脸对称轴.定位眼球时遵循以下由粗到细的过程:先局部二值化突出眼睛部位,然后根据二值图像的水平积分投影曲线的局部极小位置及上述先验条件自动粗定眼睛区域,接着用像素聚类方法找到眼球上的初始中心点,最后在此点附近用霍夫变换检测眼球的下半部分边缘以精确定出圆心和半径.

1.1 粗定眼睛区域

1.1.1 二值化

为了分散原始图像上的灰度级,用如下变换关系曲线将灰度值归到 0~255 范围:

$$Y = \left[\frac{X - \min(X)}{\max(X) - \min(X)} \right]^\gamma \quad 0 < \gamma < 1 \quad (1)$$

其中 X 是原始图像 ($m_x \times n_x$), $\min(X)$ 和 $\max(X)$ 分别为它的最小和最大灰度值, Y 为变换后的图像, γ 在 0~1 区间取值的目的是为了式(1)决定的曲线具有微微上凸的形状,以此分散低灰度值区域如眼睛等部位的灰度级,集中高灰度值区域如皮肤处的灰度级,这能拉开眼球与其周围的灰度差别并在一定程度上减少阴影的影响. γ 越小,曲线凸得越厉害,高灰度级处的灰度合并越明显,经实验测试取 $\gamma=1/2$ 即可.

为去除大部分头发及背景的干扰而简化定位过程并节省处理时间,从图像 Y 中取出第 $0.4m_x \sim m_x$ 行第 $0.25n_x \sim 0.75n_x$ 列(记为图像 Y_p)进行处理.取 $\gamma=1$,再用式(1)对图像 Y_p 进行灰度变换,得到图像 $Z(m_x \times n_x)$ 以进一步拉开灰度级.设其灰度分布直方图为 $p(r)$,其中 $r=0,1,\dots,255$ 对应各灰度级, p 为对应各个 r 的像素数.用于二值化的门限 t_c 定义为

$$t_c = \{r_i | \text{when}(\sum_{r=0}^{r_i} P(r_i)/(m_x \times n_x)) \geq p_c \%\}, \quad (2)$$

式(2)表示从最小灰度 0 开始累加像素点数 $p(r_i)$,当它占总数 ($m_x \times n_x$) 的百分比刚到或刚超过实验设定值 $p_c \%$ 时,对应的灰度级 r_i 即为门限 t_c .合理选择 p_c 的依据是:在图像 Z 中,眼睛、鼻孔、嘴巴,可能划进来的眉毛、少量头发和下巴处的阴影,深色衣领这些低灰度区域占整图的比例不会大于 $p_c \%$.由于图像 Z 上方 $1/3$ 行内必包含眼睛,而余下部分对定位眼睛没有作用,所以只对上方 $1/3$ 行二值化,将灰度值低于 t_c 的点置 0(黑),其它全部置 1(白),得到图像 C .

1.1.2 定位

积分投影是自动定位特征时常用的一种技术.设图像为 $I(i, j)$, i 和 j 分别表示行和列,则图像 I 在大小为 $(2 * b_x) \times (2 * b_y)$ 、中心位置为 (i_0, j_0) 窗口内的水平积分投影定义为

$$H(j_0) = \sum_{i=i_0-b_x+1}^{i=i_0+b_x} \sum_{j=i_0-b_y+1}^{j=i_0+b_y} I(i, j), \quad (3)$$

垂直积分投影定义为

$$V(i_0) = \sum_{i=i_0-b_1+1}^{i_0+b_1} \sum_{j=x_0-b_j+1}^{j_0+b_j} I(i, j). \quad (4)$$

当该窗口沿水平或垂直方向移动时就得到了水平和垂直积分投影曲线. 如图 1 所示, 在垂直积分投影曲线上对应眼睛、鼻孔、嘴巴和下巴处出现局部极小.

1. 1. 2. 1 定位眼睛区域的上下行

在图像 C 中, 黑色像素对应着眼睛、可能划进来的眉毛和刘海、不完整的眼镜框、眼镜脚和严重的阴影, 因此要自动正确地判定眼睛区域的中心行 $eyem$, 仅从图像 C 垂直积分投影分布曲线上局部极小对应的行来判断是不可能的, 必须结合水平投影的分布情况.

用一个宽为 n_x , 高为 $2 * b_{e1}$ 的窗口 WD_i 在图像 C 中自上往下逐行移动, 求出图像 C 的垂直积分投影曲线

$$V_e(i) = \left[\sum_{i_0=i-b_{e1}+1}^{i_0+i+b_{e1}} \sum_{j=1}^{j=n_x} C(i_0, j) \right] / (2 * b_{e1}),$$

$$i = b_{e1} : 1 : (m_x/3 - b_{e1}) \quad (5)$$

该曲线上 n_e 个局部极小处对应的行构成集合

$$\varphi_{e1} \{i_{e1} | V_e(i_{e1}) \text{ is a local minimum in } V_e(i), i = b_{e1} : 1 : (m_x/3 - b_{e1})\}. \quad (6)$$

即 $V_e(i_{e1})$ 是 $V_e(i)$ 中的一个局部极小. 将每个 i_{e1} 对应的窗口 WD_i 分成左、中、右 3 部分, 求出它们的值为

$$\left. \begin{aligned} He_{left}(i_{e1}) &= \left[\sum_{i_0=i_{e1}-b_{e1}+1}^{i_0=i_{e1}+b_{e1}} \sum_{j=1}^{j=0.4n_x} C(i_0, j) \right] / (2 * b_{e1}) / 0.4n_x, \\ He_{middle}(i_{e1}) &= \left[\sum_{i_0=i_{e1}-b_{e1}+1}^{i_0=i_{e1}+b_{e1}} \sum_{j=0.4n_x+1}^{j=0.6n_x} C(i_0, j) \right] / (2 * b_{e1}) / 0.2n_x, \\ He_{right}(i_{e1}) &= \left[\sum_{i_0=i_{e1}-b_{e1}+1}^{i_0=i_{e1}+b_{e1}} \sum_{j=0.6n_x+1}^{j=n_x} C(i_0, j) \right] / (2 * b_{e1}) / 0.4n_x. \end{aligned} \right\} \quad (7)$$

将上述 3 个值中不为 1 (即区域中有黑像素) 的最小和次小值对应的块置标记 1, 另一个置 0, 作出一个 $n_e \times 3$ 维标志矩阵 FLAG. 根据对称性可知一般眼睛和眉毛处的标记会是 [101], 但是如果额头两边有少量头发或镜框, 也会是 [101]. 在图像 Z 中, 如果这少量头发没有与眉毛粘在一起, 它对应的 $V_e(i_{e1})$ 不可能比眼睛、眉毛的更小; 如果粘在一起, 则只有一个合并的 [101]. 此外考虑眼睛, 在图像 C 中保留下来的一般是下镜框, 上镜框大都齐眉或和眉毛靠得很紧, 由于二值后的镜框比较细, 所以它对应的 $V_e(i_{e1})$ 也不可能比眼睛、眉毛的更小. 总之, 眼睛区域的 $V_e(i_{e1})$ 必然是所有 $V_e(i_{e1})$ 中最小或次小的一个. 基于此, 给出判断眼睛区域中心行 $eyem$ 的算法流程图如图 2.

在戴眼镜的图像上有可能出现 FLAG 中没有 [101] 的情况, 但这种可能性极小. 在实验

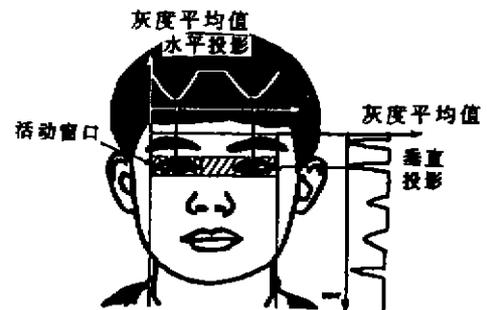


图 1 局部灰度平均值的垂直和水平投影分布示意图
Fig. 1 Diagram of vertical and horizontal of local average gray-scale projection curves

用的 51 幅照片中只碰到一张,原因在于中央镜架比较大,颜色又深,而此人的眼睛却比较小,所以在眼睛部位的标志是[110].从 $eyem$ 往上减或往下加 b_{e1} 行定出上下边界.记左眼的上下边界行为 leu 和 led ,右眼的上下边界行为 reu 和 red .

1.1.2.2 定位眼睛区域的左右列

要定位眼睛区域的左右列,首先要定出眼球的大致中央列.用一个宽为 $2 * b_{e2}$,高为 $(led - leu + 1)$ 的窗口 WD_2 在粗定的左眼区域内从左向右逐列滑动,算出左眼的水平积分投影曲线为

$$H_{le}(j) = \left[\sum_{i=leu}^{led} \sum_{j_0=j-b_{e2}}^{j+b_{e2}} C(i, j_0) \right] / (2 * b_{e2})$$

$$j = 1, 2, \dots, 0.4n_x \quad (8)$$

如果 $H_{le}(j)$ 只有一个局部极小,必对应了左眼球的中心列,但通常由于阴影和镜架的干扰使 $H_{le}(j)$ 有不只一个局部极小.阴影往往出现在左侧,因此如果式(8)中有两个或两个以上的局部极小,就分析全局极小 $H_{le}(j_1)$ 和全局次小 $H_{le}(j_2)$,若它俩的值相差比较小且 j_1 和 j_2 又分得比较开,说明干扰比较严重,根据左侧有阴影或镜脚的先验条件可知须取 j_1 和 j_2 中右边的列定为左眼眼球中心列.

用同样的方法定出右眼中心列,其判别情况与左眼的相反.从左右眼球中心列向左右分别扩出一定距离就定出了两眼的左右边界.至此,两眼的大致区域已被框出.对于无眼镜且光照均匀的照片,这样的定位方法必定正确,但对于阴影严重和戴眼镜的照片,眼球中心列的判断不一定全部正确.但是该区域必包含眼球部分,所以不会影响获得正确结果.

1.2 粗定眼球中心点

眼球是一个灰度值小的像素团,小灰度值聚类法无疑是一种简捷的定位眼球中心的方法.由于在许多照片的眼球上都有亮反光点,它们使小灰度值的聚类中心不能很好地落在眼球中心上,所以在聚类前要先对框出的左眼或右眼图像 $E(m_e \times n_e)$ 进行局部平均,以缩小眼球内像素灰度值的差别,而又不影响眼球与眼白的灰度值差别.用 9×9 窗口对图像 E 完成局部平均后,去掉宽度为 4 的边界得到图像 $E1(m_{e1} \times n_{e1})$.聚类的过程是将图像 $E1$ 中灰度值最小的 n 个像素按列从小到大的顺序排列,如果相邻两个像素的列数差值未超过预先设定的门限 T_1 ,说明只有一个聚类中心,求出这些像素行列的平均值就是要找的眼球中心点;若超过了门限 T_1 说明这 n 个像素可以聚成两类,对左眼,因为阴影、镜脚集中在左眼的左部,所以取两类中列数大的那类的平均值;对右眼,阴影干扰往往很少,偶有镜脚的干扰,所以取两类中列数小的那类的平均值即可. n 的选择必须根据图像 $E1$ 的总像素数目决定,如果太小,这些像素可能只是颜色较眼球更深的如头发上的点,如果太大,会引进过多的聚类团,影响正确判定,所以一般以最多聚成两类为宜.

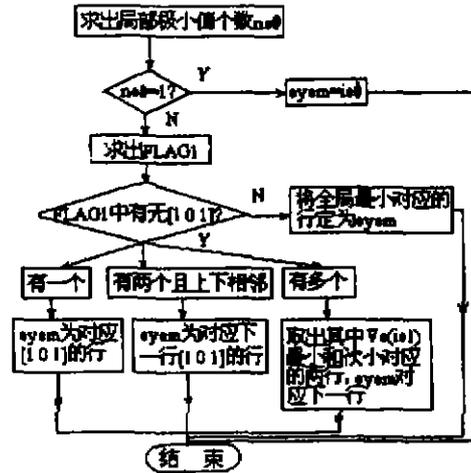


图 2 粗定眼睛区域中心行的算法流程图
Fig. 2 The flowchart of locating the coarse center row position of the eye region

1.3 用霍夫变换寻找精确的眼球中心

本文用 Canny 算子^[16,17]提取图像 $E1$ 的边缘. 由于眼球的上半部分较多地被眼皮覆盖, 所以用霍夫变换检测下半圆更可靠且省时. 设图像空间为 (x, y) 水平方向的坐标 x 对应图像矩阵空间的列, 垂直方向的坐标 y 对应图像矩阵空间的行. 三维变换空间为 (i, j, R) , 其中 i, j, R 分别代表眼球圆心的行和列, R 为半径. 下半圆表达式为

$$y = i_c + \sqrt{R^2 - (x - j_c)^2}. \quad (9)$$

变换空间的每一个坐标点 (i, j, R) 的值是图像空间中对应半圆上存在的边缘点总数. 变换空间中的峰值点坐标对应眼球半圆参数. 利用粗定的眼球上中心点, 将待检测的图像空间从图像 $E1$ 缩小为图像 $E2$, 眼球圆心 (i_c, j_c) 的检测范围也可从图像 $E2$ 缩小为图像 $E3$, 这不仅避免了周围圆弧的干扰, 而且节约运算时间. 它们的具体取值范围参见表 1. 实验证明, 霍夫变换确实具有抗干扰能力强的优点. 当眼球与眼白的对比度偏低或有严重的阴影干扰时, 提取出来的边缘是断裂或不很规则的, 即使如此, 仍能根据变换空间中的峰值点准确地定位眼球中心. 若有几个等值的峰值点, 求其平均即可.

表 1 实验所用参数(原图大小 512×384)
Table 1 Experimental parameters (image size : 512×384)

序号	参 数 内 容	参 数 值
1	粗定眼睛区域时式(2)中用于投定门限 t_c 的百分比 P_c	$P_c = 15$
2	粗定眼睛区域时式(5)中用于求垂直积分投影曲线的窗口宽度参数 b_{v1}	$b_{v1} = 5$
3	粗定眼睛区域时式(5)中用于求水平积分投影曲线的窗口宽度参数 b_{h2}	$b_{h2} = 10$
4	定位眼球初始中心点时用于聚类的点数 n	$n = 100$
5	定位眼球初始中心点时用于聚类的门限 T_1	$T_1 = n_c / 5$
6	精确定位眼球圆心时从图像 $E1(m_{e1} \times n_{e1})$ 中缩小搜索范围为图像 $E2$ 的顶行 u_1 、底行 d_1 、左列 l_1 和右列 r_1 (i_0, j_0)	$u_1 = i_0 - m_{e1} * 0.3; d_1 = i_0 + m_{e1} * 0.5; l_1 = j_0 - n_{e1} * 0.2; r_1 = j_0 + n_{e1} * 0.2$
7	精确定位眼球圆心时从图像 $E2(m_{e2} \times n_{e2})$ 中缩小搜索范围为图像 $E3$ 的顶行 u_2 、底行 d_2 、左列 l_2 和右列 r_2	$u_2 = i_0 - m_{e2} / 6; d_2 = i_0 + m_{e2} / 6; l_2 = j_0 - n_{e2} / 6; r_2 = j_0 + n_{e2} / 6$
8	精确定位眼球圆心时半径 R 的取值范围	$R = \{7, 8, 9, 10, 11, 12\}$

2 实验结果

本实验用 51 张异地拍摄的一寸免冠照进行实验, 其中 8 幅是戴眼镜拍摄的, 这些照片的光照条件不统一, 有的有淡胡须, 但没有人有很浓密的胡须, 扫图时使人脸基本布满整幅图像且基本位于正中, 全部图像均为 512×384 . 用表 1 列出的参数完成全部图像的实验处理后, 选择其中有代表性的图像置于图 3, 正确率和错误率的统计结果见表 2.

表 2 实验结果正确率及错误率的统计结果

Table 2 The statistical results of correctness rate and error rate of the experiment

内 容	实验总数	正确标准	正确率*	错误率*	备 注
定位眼睛的位置	51	正确定位两只眼睛	102(100%)	0	
粗定眼睛区域	102	眼球被全部或部分划入	100(98.0%)	2(2.0%)	只要有部分眼球被划入就不会影响后继工作
定位眼球初始中心	102	与目测中心的行或列相距±2个像素	98(96.0%)	4(4.0%)	只要该点在眼球上就能保证正确划出眼球区域
精确定眼球圆心	102	与目测圆心的行或列相距±1个像素	95(93.1%)	7(6.9%)	发生错误的7幅图像中有6幅的偏差是±2个像素,只有1幅是行列都错另1幅列偏3个像素

* 括号外数据表示结果正确或错误的图像的数目,括号内数据是对应的百分比

图 3 的每一组图像为同一人的双眼,上面是粗定区域、初定眼球中心(用一个白“+”字标志,有些与眼球内反光亮点的边缘有些重叠)、提取边缘的结果图;下面是用霍夫变换精确定位眼球圆心后重新划分眼球区域的结果.其中,(a1)、(a2)和(a3)代表那些眼球与眼白对

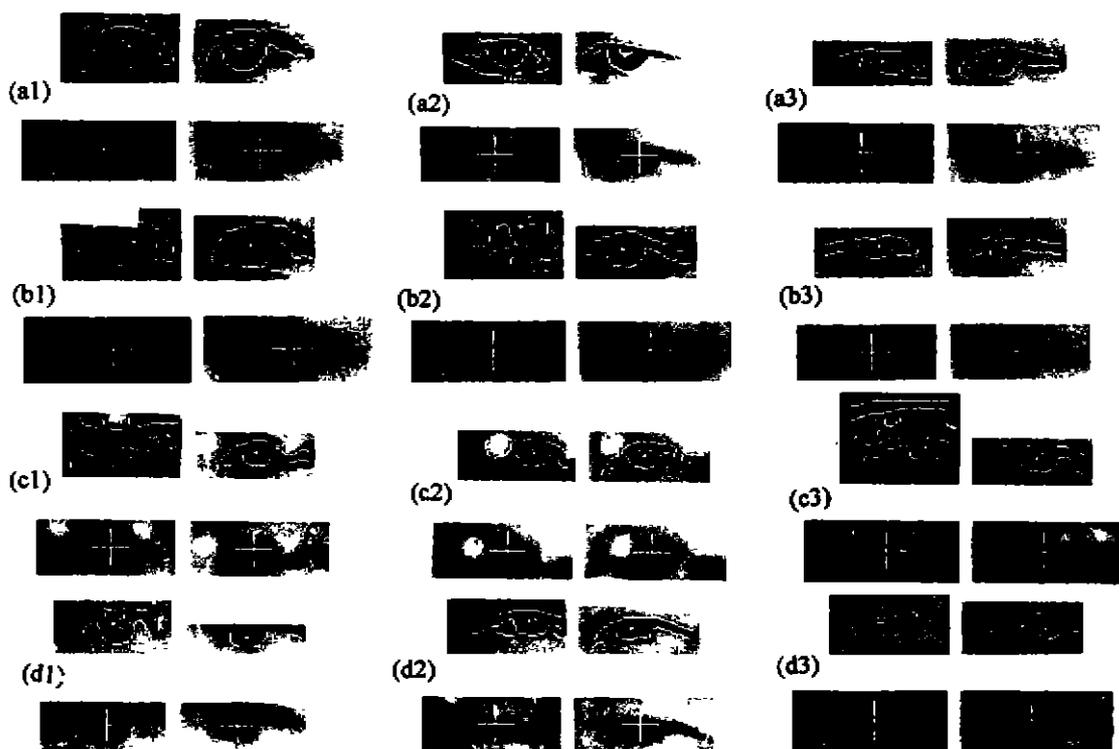


图 3 定位眼球的实验结果图

Fig. 3 Experimental results of locating iris centers

比度大从而使得边缘清晰、准确,定位精确高的图像。(b1)的眼角由于较为严重的阴影干扰(在灰度值没有归到0~255区间时,原图中的眼球和眼白难以分清)而使边缘断裂,但计算结果仍是准确的;(b2)的左眼角虽有阴影,但是边缘提取效果仍令人满意,当然结果也是准确的;(b3)的眼睛本身比较小,而这张照片本身对比度差,分辨率低,所以边缘变形比较厉害,但是即便如此,定位结果还是较准确的。(c1)~(c3)是戴有眼镜时的定位结果,在(c1)中,由于深色镜架及其投影的影响,导致粗定眼睛区域时发生了错误,好在眼球被框进去了-一部分,此外阴影的干扰使眼球上本已不完整的边缘的精确度又有所下降,所以左眼球圆心的列的定位略有偏差,这3张图上都有镜片上的反光亮点投影,特别是(c1)和(c2)上这些投影与眼球的大小相近,边缘也较清晰,如果不采用本文所用的先找眼球初始点,再在其附近定位圆心的方法,就需设定约束条件了。(d1)~(d3)是未能准确定位的错例中错误最明显的三张图,其中(d1)和(d2)的对比度较低,致使眼球边缘不准确,从而影响了精度,与目测圆心相比,(d1)的右眼圆心的行和列分别偏下和偏左2个像素距离;(d2)的右眼圆心列偏左2个像素距离。(d3)代表了那些眼睛细长而眼球大的例子,下眼敛的遮挡使得眼球下端的边缘线比较长而平坦,从而影响了精度,左眼和右眼的列分别偏右和偏左2个像素距离。另外要注意的是,(a2)左眼的粗定区域的左上端有头发,它的灰度与眼球相仿,但是用本文的像素聚类方法初定眼球中心时没有出错。

3 结论

图3中给出了误差最严重的图例,再结合表2的实验正确率和错误率的统计分析,可以看到用本方法定位眼球圆心的结果是令人满意的,此外本方法特别考虑了阴影严重或戴眼镜的情况,所以能处理实际使用的证件照,这说明本方法有实用价值。由前述理论和实验分析可以看到造成误差的一个重要原因是光照不均匀带来的影响,如果能保证光照均匀和对比度强这个条件,即在控制条件下摄入图像,显然用本方法检测到的眼球圆心的正确率更高。在后继工作中,我们将设法提取眼睛的其它特征,本文提出的这种由粗到细,逐步缩小搜索范围的定位方法最大限度地节省了计算时间。

REFERENCES

- 1 Chellappar, Wilson C L, Sirohey S. *Proc of IEEE*, 1995, **83**(5):704~740
- 2 Chen C W, Huang C L. *Inter J. of Pattern Recognition and Artificial Intelligence*, 1992, **6**:571~593
- 3 Kamd M S, Shen H C, Wong A K C, et al. *IBM. Syst J.*, 1993, **32**(2):307~320
- 4 Kirby M, Sirovich L. *IEEE Trans. on PAMI*, 1990, **12**(1):103~108
- 5 Turk M A, Pentland A P. *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, 586~591
- 6 Lin S H, Kung S Y, Lin L J. *IEEE Trans on Neural Networks*, 1997, **8**(1):114~131
- 7 Ham Y K, Lee S Y, Park R H. *IJCNN*, 1993, **3**:2125~2128
- 8 Brunelli R, Poggio T. *IEEE Trans. on PAMI*, 1993, **15**(10): 1042~1052
- 9 Intrator N, Reisfeld D, Yeshurun Y. *Pattern Recognition Letter*, 1996, **17**: 67~76
- 10 Seitz P, Bichsel M. *IEEE International Canadian Conference on Security Technology*, 1991, 77~83
- 11 Stringa L. *SPIE*, 1989, **575**:279~285
- 12 Nixon M. *SPIE*, 1989, **575**:279~285

- 13 Chow G, Li X B. *Pattern Recognition*. 1993,26(12):1739~1755
- 14 Lam K M, Yan H. *Pattern Recognition*. 1996,29(5):771~779
- 15 Yuille A L, Hallinman P W, Coken D S. *Inter J. Computer Vision*. 1992,8(2):99~111
- 16 Canny J. *IEEE Trans. PAMI*. 1986,8(6):679~698
- 17 WANG Lei, MO Yu-Long, QI Fei-Hu. *Chin. J. of Image and Graphics* (王磊,莫玉龙,戚飞虎. **中国图像图形学报**). 1996, 1(3):191~195

AUTOMATIC LOCALIZATION OF THE CENTRE OF IRIS*

WANG Lei MO Yu-Long** QI Fei-Hu

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract A coarse-to-fine approach was presented, which can locate automatically an important feature often used in face recognition, the centre of the iris on human face. It combines such techniques as local binarization, Hough transform, pixels cluster and edge extraction. By this method, one could attain high localization accuracy whether the photo is taken under uniform or nonuniform light condition, whether the person wears glasses or not. So it is suitable for processing identification photos.

Key words Hough transform, pixels cluster, edge extraction, feature extraction, face recognition.

* The project supported by the National Natural Science Foundation of China and the Science Foundation of Shanghai Educational Committee

** Institute of Communication and Information Engineering, Shanghai University (Jiading Campus)
Received 1997-08-22, revised 1998-02-17