

一种由粗到精的人眼定位方法

王 婷, 杨国胜, 申晓华

(河南大学先进控制与智能信息研究所, 开封 475001)

摘 要: 为提高人眼定位算法的实时性和抗噪性, 提出了一种基于眉眼区域内“凹陷”地形特征点检测和人眼方差滤波器的人眼定位方法(GPL)。该方法利用 Gabor 小波变换和直接提取“凹陷”地形特征点技术, 在眉眼区域内搜索眼睛候选点, 并且利用构造的人眼方差滤波器精确定位人眼。在有噪声和无噪声的人脸图像上进行了比较仿真试验。结果表明, 与纯粹基于地形特征匹配的人眼定位算法相比, GPL 在定位实时性、准确性和抗噪性方面都有显著提高。

关键词: Gabor 小波; 地形特征; 人眼方差滤波器; 人眼定位

Coarse-to-fine Eye Location Method

WANG Ting, YANG Guo-sheng, SHEN Xiao-hua

(Institute of Advanced Control and Intelligent Information Processing, Henan University, Kaifeng 475001)

【Abstract】 In order to improve the real-time and anti-noise performance of the eye location method, this paper presents a coarse-to-fine eye location method (Gabor-pit-location, GPL) based on directly detecting “pit” feature in brow-and-eye region and the eye variance filter. With the direct extraction of “pit” topographic features, the candidate’s eyes are searched in brow-and-eye region extracted from the face image by use of Gabor wavelet transform. The prior knowledge of face structure and the eye variance filter are employed to determine the real eye positions in the face image. Comparative experiments are performed on the images with and without noise respectively, and the results show that GPL is better in the real-time, accuracy and anti-noise performance than the eye location method purely based on the terrain feature matching.

【Key words】 Gabor wavelet; topographic feature; eye variance filter; eye location

1 概述

由于眼睛的精确位置信息在智能系统中的广泛应用, 其定位算法得到国内外研究者越来越多的关注, 出现了很多用于人脸图像中眼睛定位的方法, 主要有Hough变换^[1], 灰度投影^[2], 模板匹配^[3]等。尽管上述方法都可以得到较好的人眼定位效果, 但是每一种方法在定位的实时性、准确性和适用性方面各有其局限性, 特别是在复杂背景及人脸姿态变化情况下依然存在困难。

文献[4]提出了一种基于地形特征匹配的人眼定位算法(PTL), 该算法主要依据人眼瞳孔低灰度值的生物特性在人脸图像对应的3维地形图上显示出明显的“凹陷”地形特征。尽管该方法可以解决复杂背景及人脸姿态变化情况下的人眼定位问题, 定位准确率较高, 但也存在局限性: (1)将原图像转换为地形特征图, 在整幅地形特征图上搜索候选人眼点以及在得到的众多候选人眼点中筛选出真正的人眼位置都需要耗费大量的计算时间, 制约了算法的实时性; (2)赫森矩阵的特征值被用来判断像素点对应的地形特征, 但是计算赫森矩阵时需要求解对噪声非常敏感的二阶导数。所以, PTL 算法对要处理的人脸图像的去噪平滑要求较高, 否则会严重影响后续检测和定位的准确性。

为解决PTL算法的实时性和对噪声敏感的问题, 本文提出了一种基于眉眼区域内“凹陷”点检测和人眼方差滤波器的由粗到精的人眼定位算法(GPL)。借鉴文献[5]提出的利用Gabor小波变换突出显示眉眼区域的方法, 分割出人眼所在的“眉眼区域”; 文献[6]提出将主曲率方向和零交叉点^[7]相结

合, 直接从字符灰度图像中提取地形特征的方法, 在“眉眼区域”内先粗略提取“凹陷”地形特征点即定位出候选人眼点。在候选人眼点中, 基于人脸结构的先验知识, 利用方差滤波实现人眼的精确定位。经过这样处理后, 可以有效缩小候选人眼点的搜索范围, 回避计算对噪声比较敏感的赫森矩阵, 从而提高定位的实时性和抗噪性。

2 人眼粗定位

2.1 眉眼区域提取

Gabor 滤波器具有良好的带通特性。通过选择与人脸眉眼区域的频率及方向相一致的 Gabor 滤波器参数, 使得滤波后输出图像中突出显示所感兴趣的眉眼区域, 而其他区域信息受到抑制。因此, 对于大小为 $H \times V$ 的灰度图像 $I(x, y)$, 经 Gabor 小波变换后, 其对应的模值所形成的图像 K_m 为

$$K_m = \|I(z) * \psi_{\mu, \nu}(z)\| \quad (1)$$

其中, $\psi_{\mu, \nu}(z)$ 为 Gabor 小波核^[8]; μ 和 ν 分别表示 Gabor 变换核的方向和尺度。根据文献[6]并结合多次试验结果, 本文取 $\nu=1, u=6$ 。

应用“膨胀”算子 imdilate 扩展 K_m 中的高亮区域, 使其充分覆盖相应的眉眼区域; 然后, 结合人脸结构的几何知识,

基金项目: 河南省科委自然科学基金资助项目(0523020600); 河南省高校创新人才工程基金资助项目(2005KYCX012)

作者简介: 王 婷(1983 -), 女, 硕士研究生, 主研方向: 模式识别和图像处理; 杨国胜, 博士后、教授; 申晓华, 硕士研究生

收稿日期: 2006-12-31 **E-mail:** ygs@henu.edu.cn

眉眼区域位于人脸图像的上半部分，故在 K_m 的 $H/4$ 和 $2H/3$ 之间分别作水平和垂直方向上的灰度积分投影^[2]，根据水平和垂直积分投影曲线的特征确定出眉眼区域的上下和左右边界，并用集合 S 表示图像 I 中眉眼区域内的所有像素点。

2.2 人眼候选点检测

如图 1 所示，假设 $T \in S$ 的 8 邻域为 T_0, T_1, \dots, T_7 ，通过如下方法直接检测“凹陷”地形特征点即人眼候选点^[6]：(1)计算邻域像素 $T_i(i=0,1,\dots,7)$ 的梯度以及 T 在水平、竖直、右对角线、左对角线 4 个方向上的梯度的导数；(2)根据梯度的导数值确定 T 的主曲率方向和主正交元素^[6]；(3)判别人眼候选点：如果 T 在某一方向上的 2 邻域点的梯度值符号相异，称点 T 为该方向上的零交叉点。取 d_1 和 d_2 为主正交元素的方向，如果 T 在 d_1 和 d_2 方向上都是零交叉点，而且 T 在 d_1 和 d_2 方向上的 2 邻域点的梯度的符号变化都是从负到正，那么 T 就是一个凹陷地形特征点，即一个人眼候选点。用 $P = \{p_1, p_2, \dots, p_M\}$ 表示在眉眼区域中检测出的所有人眼候选点的集合， M 表示人眼候选点的个数。

T_0	T_1	T_2
T_3	T	T_4
T_5	T_6	T_7

图 1 T 的 8 邻域

3 人眼精确定位

遴选人眼候选点之后，下一步就是从这些候选点中确定真正的人眼位置。由于人脸的拓扑结构相对固定，而眼睛是由眼球和上下眼皮构成的活动对象，并且眼部区域的边缘特征非常丰富，灰度变化也比人脸其他部分更加明显。因此，本文结合人脸拓扑结构的先验知识，构建人眼方差滤波器^[9]，实现眼睛的精确定位。具体方法如下：

(1)选择 N 幅不同人、不同方向、大小为 15×27 的眼睛图像 $I_i(x, y)(i=1, 2, \dots, N)$ ，用于构建人眼方差滤波器，其中，眼球的中心被手工对齐到眼睛图像的中心。对于每一幅眼睛图像 I_i ，将其分成 3×3 的小块，计算每一个小块的方差，用这些方差值构成 I_i 对应的大小为 5×9 的方差图像 F_i^σ 。则人眼方差滤波器 F_e 由式(2)构成。

$$F_e = \frac{1}{N} \sum_{i=1}^N F_i^\sigma \quad (2)$$

(2)以 P 中任意一个点 $p_j(x, y)(j=1, 2, \dots, M)$ 为中心，从人脸图像中分割出大小为 $m \times n$ (如 15×27) 的矩形窗，记为 W_j ，按式(3)计算 W_j 通过 F_e 后的响应。

$$R(W_j^\sigma, F_e) = \frac{E[(W_j - E(W_j))(\xi_{F_e} - E(F_e))^T]}{\sqrt{D(W_j)D(F_e)}} \quad (3)$$

其中， W_j^σ 为 W_j 对应的方差图像； ξ_{W_j} 和 ξ_{F_e} 分别是将 W_j^σ 和 F_e 按行展开后得到的向量； $E(\cdot)$ 和 $D(\cdot)$ 分别表示数学期望和方差；记 e_1 为一个精确的人眼点，则取 e_1 为

$$e_1 \text{ s.t. } \max(R(W_j^\sigma, F_e), j=1, 2, \dots, M) \quad (4)$$

(3)将 e_1 从 P 中删除，在 P 的剩余元素中，重复步骤(2)，求出每一元素对应的方差图像与 F_e 的响应。在以下先验知识的约束下：1)双眼中心之间的距离应在一定范围内(如 $10 \sim 50$ 个像素)；2)眼睛中心与图像的边界有一定距离(如与图像左右边界的距离至少为 13 个像素，与图像上下边界的距离至少为

7 个像素)；3)双眼中心位置的倾角不超过 30° ，取响应最大的候选点为另一个精确的人眼点，记为 e_2 。

(4) e_1 和 e_2 构成了精确的人眼点对。

4 仿真实验结果与分析

为了比较 GPL 和 PTL 人眼定位算法的实时性和抗噪性(或定位的准确性)，本文进行了对比仿真实验。实验采用 AT&T 人脸库和 JAFFE 人脸表情库^[10]，从中选取两个眼睛瞳孔可见，眼部没有遮挡物影响的人脸图像，图像大小均为 112×92 。

4.1 无噪声时的仿真实验

从 AT&T 和 JAFFE 中随机选取不受噪声干扰的 5 幅原始人脸图像，分别采用 GPL 和 PTL 方法先对其进行人眼粗定位，结果如图 2 所示，其中，第 1 行是采用 GPL 方法获得的眉眼区域定位以及在眉眼区域中检测出的人眼候选点；第 2 行是采用 PTL 方法检测出的人眼候选点。图 2 中的每幅图像中检测出的人眼候选点个数如表 1 所示。

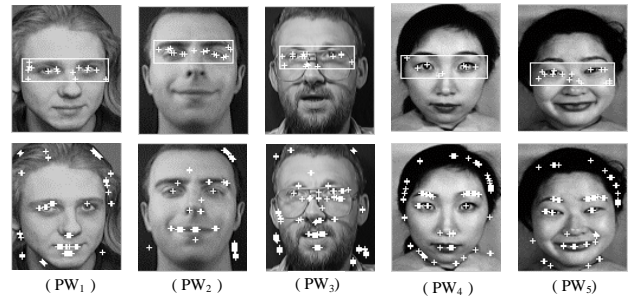


图 2 无噪声干扰的人眼粗定位结果

表 1 两种方法下检测出的人眼候选点个数

定位方法	人眼候选点个数				
	PW ₁	PW ₂	PW ₃	PW ₄	PW ₅
GPL	15	16	16	8	16
PTL	66	53	95	79	62

由表 1 可知，采用 GPL 方法可以准确定位出人脸图像上的眉眼区域，并且检测出的候选人眼点数远远少于用 PTL 方法检测出的候选人眼点数。这节省了对候选人眼点进行筛选以精确定位人眼位置所需的时间，提高了实时性。表 2 列出了采用 GPL 和 PTL 方法对图 2 中每幅人脸图像精确定位人眼所需的时间。

表 2 两种方法下精确定位人眼所需时间

定位方法	精确定位人眼所需时间/s				
	PW ₁	PW ₂	PW ₃	PW ₄	PW ₅
GPL	1.103	1.091	0.965	0.806	0.807
PTL	3.922	4.027	4.031	3.972	3.834

由表 2 可知，采用 GPL 方法进行人眼定位时，每幅图像的定位时间平均需要 0.954s，而采用 PTL 方法进行人眼定位时，每幅图像的定位时间平均需要 3.957s，由此可知，本文所提出的人眼定位方法 GPL 的实时性比 PTL 有了较大提高。

图 3 显示了分别采用 GPL 和 PTL 方法进行人眼精确定位的仿真实验结果，其中，白色“+”标记最终定位出的人眼位置；第一行为采用 GPL 方法的定位结果；第二行为采用 PTL 方法的定位结果。由图 3 可以定性看出，PTL 方法的准确性没有 GPL 方法的准确性高。为定量说明这一问题，本文引入定位准确性相对误差。设手工标定的左右眼睛瞳孔中心位置分别为 H_l 和 H_r ，算法最后定位出的左右眼睛位置分别为 H'_l 和 H'_r ，用 d_l 、 d_r 和 d_r' 分别表示 H_l 到 H'_l 、 H_r 到 H'_r 、 H_l 到 H_r' 的距离，则定位准确性的相对误差定义为^[9]

$$err = \frac{\max(d_l, d_r)}{d_{lr}} \quad (5)$$

通常, d_{lr} 约为眼睛宽度的 2 倍。根据式(5), 采用 GPL 和 PTL 方法时眼睛定位准确性的平均相对误差分别为 0.054 和 0.133。由比较可知, 本文所提出的 GPL 方法的定位准确性比 PTL 方法的定位准确性高。

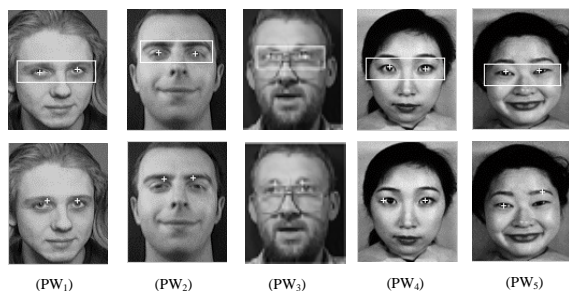


图3 无噪声干扰的人眼最终定位结果

4.2 有噪声时的仿真实验

为了检验 GPL 和 PTL 方法对噪声影响的敏感性, 本文还进行了下面的实验: 在从 AT&T 和 JAFFE 中随机选出的不受噪声干扰的 5 幅原始人脸图像上叠加特定类型的噪声(如: 均值为 0, 方差为 0.01 的高斯噪声), 得到有噪声干扰的人脸图像。分别采用 GPL 和 PTL 方法先对其进行人眼粗定位, 其结果如图 4 所示, 其中, 白色“+”标记人眼候选点; 白色边框标记眉眼区域; 第 1 行为采用 GPL 方法检测出的人眼候选点; 第 2 行为采用 PTL 方法检测出的人眼候选点。最终定位结果如图 5 所示, 其中, 白色“+”标记最终定位出的人眼位置; 第 1 行为采用 GPL 方法的定位结果; 第 2 行为采用 PTL 方法的定位结果。表 3 列出了图 4 的每幅图像中检测出的人眼候选点个数。

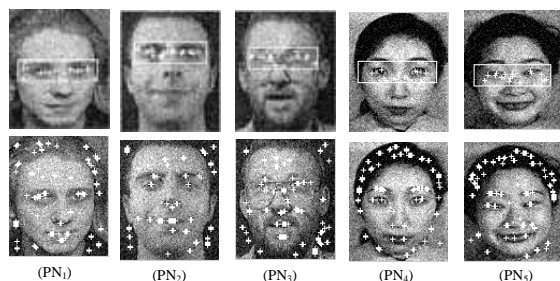


图4 有噪声干扰的人眼粗定位结果

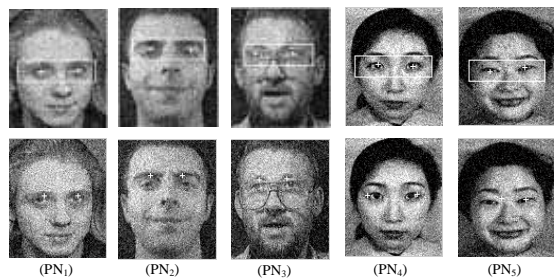


图5 有噪声干扰的人眼最终定位结果

表3 添加高斯噪声后两种方法下检测出的人眼候选点个数

定位方法	人眼候选点个数				
	PN ₁	PN ₂	PN ₃	PN ₄	PN ₅
GPL	15	18	17	10	16
PTL	95	88	106	99	110

将表 1 和表 3 中的数据对比可知, 对于添加高斯噪声前后的同一幅人脸图像进行人眼定位, 采用 GPL 方法得到的候

选人眼点个数变化不大; 而采用 PTL 方法时, 有噪声时检测到的候选人眼点个数远多于无噪声时检测到的候选人眼点个数。这是因为在 GPL 方法中直接提取地形特征时, 只用到了像素点的二阶导数值的符号信息, 而符号信息受噪声干扰不明显, 所以检测到的人眼候选点数受噪声干扰亦不明显; 而 PTL 方法在判断像素点对应的地形特征时用到了对噪声敏感的像素点的二阶导数, 导致检测到的人眼候选点数随着噪声的引入而增加。

比较图 3 和图 5 可知, 采用 GPL 方法对添加了噪声的图像进行人眼定位时的定位结果明显好于采用 PTL 方法对同一幅图像进行人眼定位的结果。根据式(5), 在该实验中, 采用 GPL 和 PTL 方法时眼睛定位准确性的平均相对误差分别为 0.074 和 0.193。

综上可知, PTL 方法受噪声干扰比较明显, 而 GPL 方法对噪声不敏感。

5 结束语

本文利用 Gabor 小波变换, “凹陷”地形特征点检测和人眼方差滤波器, 提出了一种改进的由粗到精的快速人眼定位方法。与文献[5]中的人眼定位方法相比, 该方法对噪声不敏感, 在定位实时性和准确性方面都有较大提高。仿真实验结果验证了本文所提算法的有效性。

参考文献

- 1 张晓飞, 赵瑞莲. 基于 Hough 变换圆检测的人眼精确定位方法[J]. 计算机工程与应用, 2005, 41(27): 43-45.
- 2 冯建强, 刘文波, 于盛林. 基于灰度积分投影的人眼定位[J]. 计算机仿真, 2005, 22(4): 75-104.
- 3 史慧荣, 张学帅, 梁彦, 等. 一种基于模糊分类的模板匹配眼睛定位方法[J]. 西北工业大学学报, 2005, 23(1): 55-59.
- 4 Wang J, Yin L J. Detecting and Tracking Eyes Through Dynamic Terrain Feature Matching[C]//Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005: 78-85.
- 5 Yang P, Du B, Shan S G, et al. A Novel Pupil Localization Method Based on Gaboreye Model and Radial Symmetry Operator[C]//Proc. of International Conference on Image Processing. 2004: 67-70.
- 6 Lee S W, Kim Y J. Direct Extraction of Topographic Features for Gray Scale Character Recognition[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1995, 17(7): 724-729.
- 7 Mallat S, Hwang W L. Singularity Detection and Processing and Processing with Wavelets[J]. IEEE Trans. on Information Theory, 1992, 38(2): 617-643.
- 8 Liu C J, Wechsler H. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant model for Face Recognition[J]. IEEE Transactions on Image Processing, 2002, 11(4): 467-476.
- 9 Fathi A, Manzuri M T. Eye Detection and Tracking in Video Streams[C]//Proc. of International Symposium on Communications and Information Technologies. 2004: 1258-1261.
- 10 Lyons M J, Akamatsu S, Kamachi M, et al. Coding Facial Expressions with Gabor Wavelets[C]//Proc. of the 3rd IEEE Int'l Conference on Automatic Face and Gesture Recognition. 1998: 200-205.
- 11 Jesorsky O, Kirchberg K J, Frishholz R W. Robust Face Detection Using the Hausdorff Distance[C]//Proc. of Lecture Notes in Computer Science on Audio and Video Based Person Authentication. 2001: 90-95.