

文章编号:1007-2780(2017)06-0499-07

基于改进的 SIFT 算子和 SVM 分类器的 瞳孔中心定位

田 耘*, 甄 雯, 赵海军
(北京电子科技职业学院, 北京 100176)

摘要:瞳孔定位的精确度很大程度取决于图片质量,但实际应用中通常要在低质量图片下进行瞳孔定位。我们的目标是在图片质量不佳的情况下进行精确的瞳孔中心定位。对于这个目标,本文提出一种基于改进 SIFT 特征和 SVM 分类器的瞳孔中心初始定位方法,并通过一个大小可变的修正矩形框得到最终瞳孔中心位置。实验结果表明,相比于其他国内外先进方法,本文的方法可以在低质量(光照不均、表情变化等)图片上拥有更高的瞳孔定位精度,定位结果在瞳孔区域内的精度为 87.32%。

关键词:瞳孔定位;分类器;SIFT 特征;修正矩形框

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20173206.0499

Accurate pupil center location with SIFT descriptor and SVM classifier

TIAN Yun*, ZHEN Wen, ZHAO Hai-jun

(School of Automotive Engineering, Beijing Polytechnic, Beijing 100176, China)

Abstract: The accuracy of gaze tracking largely depends on the quality of images, but additional constraints and large amount of calculation make gaze tracking impractical on high-resolution images. Our aim is to get the accurate localization of pupil center on low-resolution image. To this aim, we proposed a simple but effective method which can accurately locate pupil center in real time. The method first gets initial eye center based on improved SIFT descriptor and SVM classifier, and then gets final position of the pupil center through a size variable correction rectangular block. In this paper, comparing with the reported state-of-the-art methods, the experimental results demonstrate that our system can achieve a more accurate result on low-resolution images, the precision of positioning result in the pupil area was 87.32%.

Key words: pupil center localization; classification; SIFT feature; correction rectangular block

收稿日期:2016-10-28;修订日期:2017-02-13.

* 通信联系人, E-mail: zhitu 2004@163.com

1 引言

随着计算机技术的飞速发展,视线跟踪技术被应用于越来越多的领域。如今人们可通过对注视位置的实时检测,找到感兴趣区域,从而对产品进行评价。此外还可通过注视位置及眨眼频率评价驾驶员疲劳程度。瑞典 Tobii 公司与美国 SMI 公司研制出了基于人眼瞳孔中心定位的头戴式和非接触式两种眼动仪,其能实时监测被试者注视位置和眨眼频率。然而在光照变化等干扰因素下,要精确地检测出人眼位置仍然充满着挑战。

目前存在的人眼中心定位方法可分为 3 类^[1-2]:(1)眼电记录法:记录眼睛周围皮肤的电势差;(2)巩膜接触镜法:采用一些安装在接触透镜上的机械参考;(3)图像处理法:使用图像处理技术对人眼中心进行定位。考虑到大多数情况下不允许使用外界设备,所以第三种方法被广泛研究。其中基于训练的方法往往速度更快,准确率更高,因此近年来受到了研究者的广泛重视^[3-4]。

目前瞳孔中心定位的基本思想是首先找到人脸区域,然后在该区域中对瞳孔中心进行定位。文献[5-6]是目前比较先进的瞳孔定位方法。Valenti 等人^[5]提出了一种基于等照度线的定位方法,该方法利用尺度空间的变化来提高瞳孔定位精度和对图片大小变化的鲁棒性。Markus 等人^[6]使用随机回归树的集合来定位瞳孔中心,通过实验证实该方法可以得到一个精度很高的结果,并可以在移动设备上实时运行。然而以上两种方法都很难识别眨眼情况,这会在视线跟踪过程中引起很大的误差。

在使用高分辨率的正面人脸图片时,使用上文所提到的方法进行瞳孔定位可以得到较高的准确率。然而高分辨率图片会增加计算时间,为了保证每张图片的处理速度,一般采用低于 320×240 分辨率的图片进行处理。在对低分辨率图片使用上述提到的方法进行瞳孔定位时,当人脸出现较大表情变化或有较强光照影响时,定位精度往往会出现较大的偏差。

因此,本文提出了一种基于改进 PCA-SIFT 特征的瞳孔中心定位算法,该方法在低分辨率图片中可以快速、准确地对瞳孔中心进行定位,并且

能有效地识别出眨眼图片。本文方法对比之前的方法取得了更好的结果。此外,本文使用自己建立的闭眼数据库进行测试,同样取得了较好的鲁棒性。

2 方法

相比于人脸其他部位,瞳孔近似于圆形且灰度值较低。本文首先使用 opencv 中的人脸检测器找到人脸区域,然后将人脸高度的 $1/5 \sim 3/5$,宽度的 $1/10 \sim 9/10$ 作为瞳孔检测区域。

2.1 选取关键点

由于瞳孔近似为圆形,可以使用快速圆检测方法将瞳孔中心选取为关键点。Kothari 和 Mitchell 使用梯度检测圆心^[7-8],但这种方法不能解决由于眉毛、眼睑和眼镜框引起的误差。本文使用改进的 SIFT 算法可以快速地将瞳孔中心作为关键点检测出来。

SIFT 特征是一种局部特征,其在不同的尺度空间上查找关键点。尺度空间的获取需要通过高斯模糊来实现,并且使用高斯金字塔表示。在尺度空间建立完毕后,为了减少计算量,采用高斯差分(DoG)的方法来检测极值点。高斯差分尺度空间被定义为:

$$D(x, y, \sigma) = (G(x, y, \sigma(s+1)) - G(x, y, \sigma(s))) * I(x, y), \quad (1)$$

其中: $G(x, y, \sigma(s))$ 和 $I(x, y)$ 分别表示高斯函数和原始图片, s 为高斯金字塔的不同层面。因为高斯差分函数与归一化的高斯拉普拉斯函数近似(图 1),可以从高斯差分金字塔分层结构提

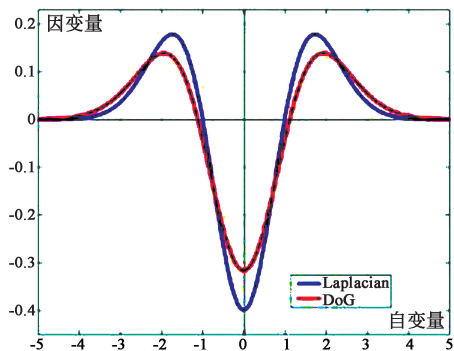


图 1 高斯拉普拉斯和高斯差分的比较

Fig.1 Comparison of gauss Laplace and gaussian difference

取出图像中的极值点作为候选的关键点。在已检测到的极值点中去除低对比度的关键点和不稳定的边缘响应点(因为高斯差分尺度空间会产生较强的边缘响应)记为所选取的关键点。

当高斯窗口大小能覆盖待检测圆时,由于SIFT算法中高斯函数的圆对称性,圆心点会被检测为DoG空间的一个极值点。调整好建立高斯金字塔的参数后,图片检测到的关键点为图形的边缘以及面积较小圆的圆心。所以瞳孔中心以及瞳孔/虹膜边缘会作为关键点被检测出来。

2.2 建立特征向量

本文瞳孔定位方法使用改进的SIFT特征和像素特征。使用这两个特征的原因:(1)SIFT特征对图像的缩放、平移和旋转具有不变性。并且对于光线、噪声、视角改变的鲁棒性很好,在图像匹配和识别中有着很好的效果^[9]。(2)灰度特征在图像识别中有着丰富的资源信息,且瞳孔部分像素的灰度值低于人脸其他部位像素的灰度值。

在以检测到的关键点为中心的邻域窗口中进行采样,并统计邻域像素的梯度方向,以每10度为一个单元,总共36个单元。统计每个单元内梯度方向总和,其中峰值处代表了关键点处邻域梯度的主方向,即为关键点的方向。将坐标轴旋转为关键点的方向,以确保旋转不变性。以特征点为中心取 8×8 的邻域作为采样窗口,每 2×2 邻域作为一个子单元,将采样点与特征点的相对方向通过高斯加权后归入8个方向直方图,最后获得 $4 \times 4 \times 8$ 的128维特征描述子。由于128维特征点描述向量维数较多,会造成计算量较大,检索效率低等问题。因此本文采用一种改进的SIFT和主成分分析(PCA)相结合的SIFT-PCA算法^[10],将128维的SIFT特征向量降低到36维。

像素灰度值特征向量由关键点自身灰度值和关键点周围9个点的灰度值组成,和36维SIFT特征向量组成一个46维的特征向量,记为该关键点的特征向量,并作为分类器的输入特征。使用SVM分类器对关键点进行分类,根据正例关键点位置计算瞳孔的初始位置。

2.3 瞳孔位置初定位

使用分类器对所有关键点进行分类后,可能会出现一只眼睛内得到的正例关键点数大于1个。这种情况大多是由于眼睫毛过重、黑色眼镜框和光线照射不均引起的,产生的错误分类关键

点大多处于虹膜边界,但有一小部分处于眼睛区域边缘。如果采用求取平均值的方法来确定瞳孔初始位置,处于眼睛区域边缘的错误分类关键点会使其远离真实瞳孔中心。因此以每个正例关键点到眼睛区域中心的欧式距离来分配权值 w_i :

$$d_i = \frac{1}{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}}, \quad (2)$$

$$w_i = \frac{d_i}{\sum_{i=0}^n d_i}, \quad (3)$$

其中: (x_0, y_0) 为眼睛区域的中心位置, (x_i, y_i) 为正例关键点位置, d_i 为每个关键点到眼睛区域中心的欧式距离的倒数。每个正例关键点分配好权值后,按公式(4)、公式(5)计算瞳孔的初始位置,得到的 (x, y) 即为瞳孔的初始位置。

$$x = \frac{\sum_{i=0}^n w_i x_i}{n}, \quad (4)$$

$$y = \frac{\sum_{i=0}^n w_i y_i}{n}. \quad (5)$$

该方法保证了眼睛中心区域的关键点可以获得较高的权值,瞳孔中心的初始位置不会过分偏离瞳孔中心的真实位置。若只有一只眼睛的区域内检测到了关键点,则以人脸中心为轴将该眼睛内瞳孔初始位置平移对称到另一只眼睛内,作为瞳孔中心的初始位置。

2.4 瞳孔位置修正

有时虹膜边界、瞳孔边界也会作为正例关键点被检测出来,这可能是因为SIFT特征较强边缘响应引起的,这会使瞳孔的初始定位偏离瞳孔中心。因此采用基于滑动窗口的方法对瞳孔中心位置进行修正。

瞳孔部分的灰度值低于瞳孔周围眼睑、巩膜部分的灰度值,所以以瞳孔初始位置为中心指定一个固定区域。连续地滑动该指定区域,并计算该区域内的平均灰度值。当固定区域内的平均灰度值最低时,认为该区域的中心为瞳孔中心的最终检测位置。

上文提到大多数情况是瞳孔边界引起的误差,因此瞳孔中心初始位置一般不会偏离出虹膜区域。由于人眼瞳孔直径一般为人脸宽度的1/30,最大移动距离设为人脸宽度的1/15。假设人

脸宽度为 N , 分别选取边长为 $N/30$ 、 $N/35$ 、 $N/40$ 的矩形框作为瞳孔中心定位的搜索框, 并比较这 3 种修正矩形框对瞳孔定位精度的影响。瞳孔最终定为结果如图 2 所示, 其中绿点为瞳孔初始定位结果, 红点为矫正后瞳孔定位结果。

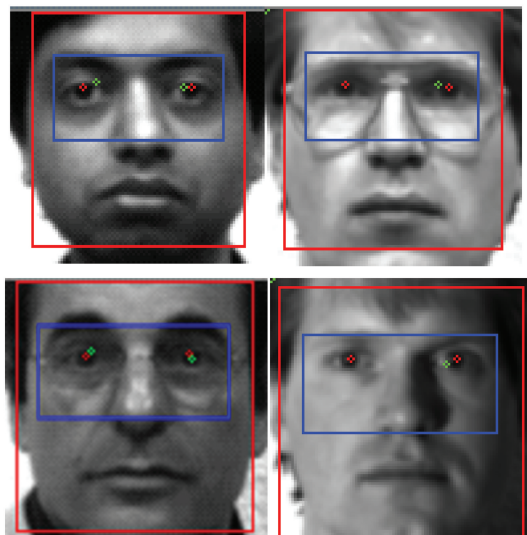


图 2 瞳孔位置修正

Fig.2 Pupil position correction

3 训练方法

3.1 训练数据

本文使用 BioID 和 Yale 公开数据库进行训练。此外, 我们建立了自己的数据库, 其中包括不同的学生和网上的图片。用于训练的图片总共有 3 200 张人脸图片, 瞳孔区域都已经被手工标定。

3.2 训练过程

在数据充足的情况下, Boosting 和随机森林通常是最好的分类器。但当数据有限的时候, 支持向量机 (SVM) 通常能取得较好的效果。Valenti 等人^[5]已经证明 KNN 分类器在 SIFT 特征中的分类效果可以达到最好。

考虑到本文使用的是基于 PCA-SIFT 特征, 因此分别采用 SVM 分类器和 KNN 分类器对上文提到的瞳孔检测区域内的关键点进行分类, 并比较这两种分类算法的正确率。我们对 BioID 和 Yale 数据库采用交叉验证的方法进行精度计算。将训练数据集分成 10 份, 轮流将其中 9 份做训练 1 份做测试, 10 次测试结果的均值作为对算法精

度的估计。分类结果如表 1 所示。

表 1 分类器交叉验证结果

Tab.1 Classifier cross validation results

方法	BioID 数据库	Yale 数据库	自己的数据库
SVM	91.65%	90.47%	89.02%
KNN	95.42%	96.16%	97.33%

从表 1 的分类结果中可以看出对本文使用的改进 PCA-SIFT 特征, KNN 分类器的分类效果确实优于 SVM 分类器。在下一节中, 我们将使用更加标准的方法对瞳孔最终定位精度的比较。

4 实 验

4.1 评价标准

归一化误差将会作为人眼定位精度的评价标准。该标准被 Jesorsky 等人^[16]所提出:

$$e = \frac{\max \{D_l, D_r\}}{D}, \quad (6)$$

其中 D_l 和 D_r 为检测到的左眼、右眼瞳孔中心到真实左眼、右眼瞳孔中心的欧式距离, D 为左眼、右眼真实瞳孔中心间的欧式距离。在这个标准中, $e \leq 0.25$ 对应于眼睛中心和眼角之间的距离; $e \leq 0.1$ 对应于虹膜的范围, 以及 $e \leq 0.05$ 对应于光瞳的范围。本文使用归一化误差比较我们的研究结果与其他已发表的研究结果。

4.2 实验结果

在本文中 BioID 和 ORL 数据库用于测试。BioID 数据库由 23 名不同被试, 总共 1521 张图片组成。每张图片的大小为 384×288 , 其中包含一部分被试者佩戴眼镜, 头部有较大的偏离, 闭眼或者眼睛被强烈光照所覆盖。由于这些条件, BioID 数据库被认为是一个较难、较真实的数据库。数据库中还提供了瞳孔的中心位置。

部分 ORL 数据库由 400 张正面人脸图片组成, 其中也包括了不同光照、姿势和位置的影响, 每张图片的大小为 100×100 。瞳孔真实位置由手工标定所得。图 3 和图 4 展示了不同被试者的测试结果。从图 3 和图 4 第二行可以看出, 如果眼睛在阴影中或者有较强的光照, 本文的方法将会定位失败。通过分析定位失败图片, 当距离真实瞳孔中心较远的正例关键点被提取, 本文的方法将会产生错

误定位。此外,眼睛区域中瞳孔位置的初始定位距离瞳孔真实位置过远,或者镜片反光造成瞳孔灰度信息缺失,2.3 节中介绍的瞳孔位置修正方法不能将瞳孔位置修正到真实瞳孔位置。

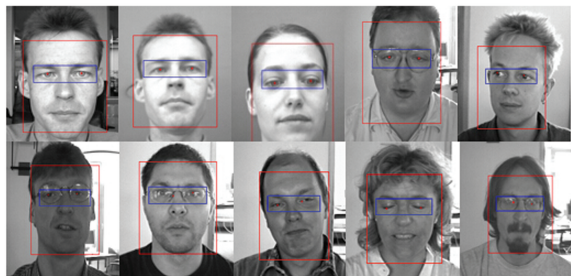


图3 BioID 数据库中瞳孔中心检测成功(第1行)和检测失败(第2行)例子

Fig.3 Successful and failed test examples in the BioID pupil center database

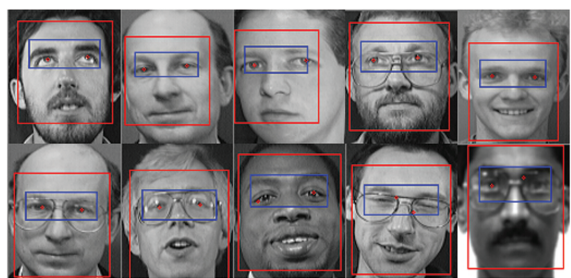
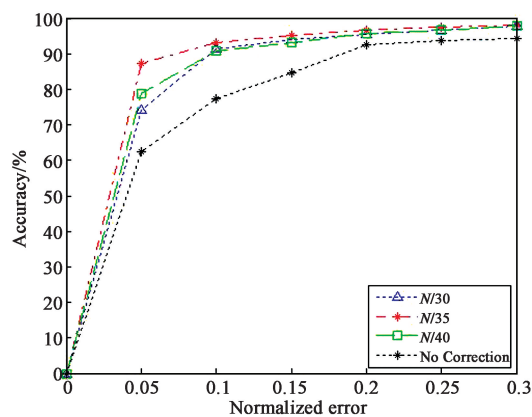


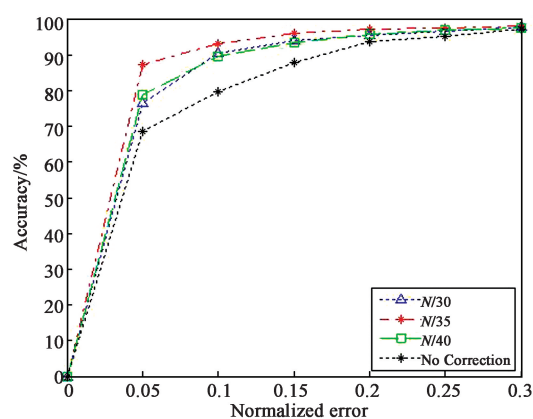
图4 ORL 数据库中瞳孔中心检测成功(第1行)和检测失败(第2行)例子

Fig.4 Successful and failed test examples in the ORL pupil center database

图5和图6对比了我们方法在加入不同尺寸修正矩形框时精度的变化。同时对比了加入修正矩形框对与不加修正矩形框的实验结果。从实验结果中可以看出,在归一化误差达到0.15时基本



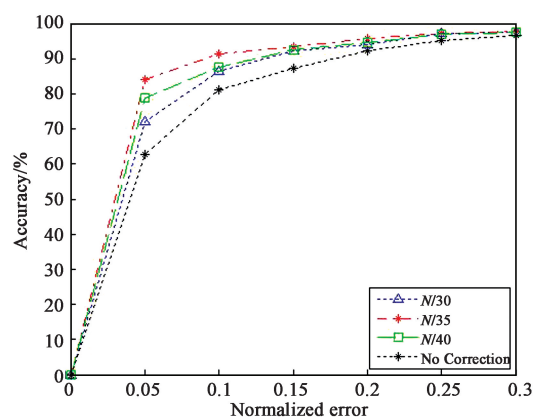
(a)SVM



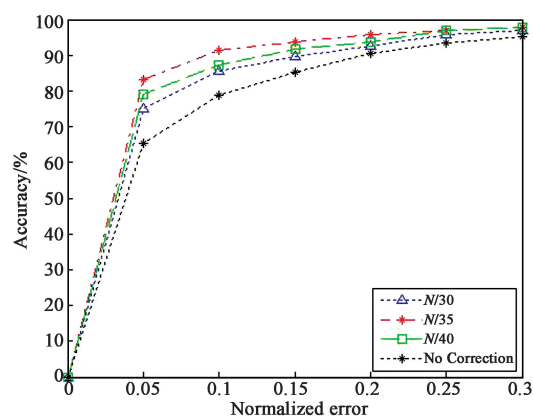
(b) KNN

图5 在 BioID 数据库中使用修正矩形框时的精确度变化

Fig.5 Accuracy change when using modified rectangular box in the BioID database



(a)SVM



(b)KNN

图6 在 ORL 数据库中使用修正矩形框时的精确度变化

Fig.6 Accuracy change when using modified rectangular box in the ORL database

达到稳定。这也证明了 2.3 节提到的关键点错误分类大多发生在瞳孔/虹膜边缘。结果还能证明在使用 $N/35$ 大小的修正矩形框时瞳孔定位精度最高。

通过比较图 5(a)和图 6(a)可以看出,BioID 数据库相比于 ORL 数据库可以得到更好的实验结果。这也证明了基于训练方法的实验结果很大程度上取决于训练数据。虽然 3.2 节已经证明在使用基于 sift 特征时 KNN 分类器会取得较好的效果,但在使用修正矩形框对瞳孔中心位置进行修正后,使用 KNN 分类器得到的实验结果相比于使用 SVM 分类器得到的实验结果没有明显的提升。这是由于修正矩形框能很好地修正瞳孔/虹膜边缘处错误关键点所引起的误差。考虑到 KNN 分类器的检测时间远远超过 SVM 分类器的检测时间,因此本文后续实验均采用 SVM 分类器和边长为 $35/N$ 的修正矩形框进行实验。

4.3 与其他方法比较

由于大多数先进方法使用 BioID 数据库进行的验证,本文对 BioID 数据库定位结果与其它先进方法进行了比较,如表 2 所示。

表 2 瞳孔中心定位精度比较

Tab.2 Comparison of pupil center positioning accuracy

Method	$e \leq 0.05$	$e \leq 0.1$	$e \leq 0.25$
Our results,SVM	87.32%	93.15%	97.38%
Our results,KNN	87.28%	93.26%	97.54%
Valenti ^[13] 基础方法	80.58%	85.18%	96.56%
Valenti ^[13] 延伸方法	86.09%	91.67%	97.87%
Timm ^[22]	82.50%	93.40%	98.00%
Markus ^[23] , $p=7$	85.7%	95.3%	99.7%
Markus ^[23] , $p=31$	89.9%	97.1%	99.7%

由于本文更关心的是瞳孔定位在瞳孔区域内的精度($e < 0.05$),修正矩形框选取的尺寸略小于瞳孔,对初始瞳孔初始位置定位在眼睛区域($e < 0.25$)的情况无法进行修正,因此在归一化误差 $e < 0.25$ 的情况和其他方法相比在精度上没有提高。但如果本文实验更关心 $e < 0.25$ 的精度,通过选取不同尺寸和形状的修正矩形框使用本文的方法进行检测,可以实现归一化误差比 $e < 0.25$ 更好的实验精度。

4.4 闭眼图片检测

在大多数先进方法中,瞳孔定位不能在闭眼图片下进行。由于本文的方法在训练过程中闭眼图片中的关键点全部按反例进行分类,因此在检测过程中如果两眼区域中均没有检测到关键点则说明该图片为闭眼图片。为了评价本文方法对闭眼图片的鲁棒性,从 BioID 数据库和通过摄像头采集总共 120 张闭眼图片。每张图片的大小为 384×288 。在对 120 张闭眼图片试验中,总共有 97 张图片没有检测到关键点,正确率为 80.83%。测试结果如图 7 所示。



图 7 闭眼图片识别成功(第 1 行)和识别失败(第 2 行)例子

Fig.7 Success (line 1) and failure examples (line 2) on closing eyes picture recognition

从检测失败的闭眼图片中可以看出,如果一只眼睛略微睁开,则会检测到正例关键点,另一只眼睛通过 2.4 节中提到的修正方法也会得到错误的瞳孔中心位置;此外如果在眼镜片上形成较强的反射光斑,也会产生错误的正例关键点,从而影响闭眼图片的判断。以上这两点所引起的错误判断,本文使用的方法不能进行识别。

5 结 论

本文提出了一种基于改进的瞳孔中心定位方法,其在对低分辨率图片进行检测时也能得到较高的精确性。通过调整生成 SIFT 关键点时尺度空间的参数,使瞳孔中心作为关键点被提取出来,并使用 SVM 分类器得到瞳孔中心的初始位置。然后利用一个大小可变的矩形框对瞳孔的初始定位进行修正,其提高了对低质量图片瞳孔中心定位的准确性。此外,本文方法对于实际情况中经常遇到的图片分辨率改变和闭眼问题均能表现出较好的鲁棒性。

参 考 文 献:

- [1] DUCHOWSKI A. *Eye Tracking Methodology: Theory and Practice* [M]. London: Springer, 2007.
- [2] SUGANO Y, MATSUSHITA Y, SATO Y. Appearance-based gaze estimation using visual saliency [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(2): 329-341.
- [3] CAMPADELLI P, LANZAROTTI R, LIPORI G. Precise eye localization through a general-to-specific model definition [C]//*Proceedings of the British Machine Vision Conference*. Edinburgh, Scotland: BMVC, 2006: 187-196.
- [4] TIMM F, BARTH E. Accurate eye centre localisation by means of gradients [C]//*Proceedings of the Sixth International Conference on Computer Vision Theory and Applications*. Vilamoura, Algarve, Portugal: VISAPP, 2011: 125-130.
- [5] VALENTI R, GEVERS T. Accurate eye center location through invariant isocentric patterns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(9): 1785-1798.
- [6] MARKUŠ N, FRLJAK M, PANDŽIĆ S, *et al.* Eye pupil localization with an ensemble of randomized trees [J]. *Pattern Recognition*, 2014, 47(2): 578-587.
- [7] KOTHARI R, MITCHELL J L. Detection of eye locations in unconstrained visual images [C]//*Proceedings of International Conference on Image Processing*. Lausanne: IEEE, 1996, 3: 519-522.
- [8] KROON B, HANJALIC A, MAAS S M P. Eye localization for face matching: is it always useful and under what conditions [C]//*Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval*. New York, NY, USA: ACM, 2008: 379-388.
- [9] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615-1630.
- [10] HE X F, CAI D, HAN J W. Learning a maximum margin subspace for image retrieval [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2008, 20(2): 189-201.

作者简介:田耘(1963—),女,北京人,学士,副教授,1987年毕业于北京联合大学获学士学位,主要从事机械设计制造, CAD/CAM,计算机视觉,图像处理。Email:zhitu2004@163.com

甄雯(1983—),女,北京人,硕士,讲师,2012年毕业于天津职业技术师范大学获得硕士学位,主要从事机械设计制造, CAD/CAM,计算机视觉,图像处理。Email:zhenwen_1101@163.com

赵海军(1961—),男,北京人,高级技师,2009年毕业于北京市委党校,主要从事机械设计制造,机械加工,计算机视觉,图像处理。Email:zhaohj21@126.com