

一种改进的基于随机森林的快速人眼定位方法

马 越, 贺光辉

(上海交通大学, 微电子学院, 上海 200240)

摘 要: 为了得到更强的鲁棒性和更好的实时性, 针对精确的人眼定位提出了基于随机森林的自适应梯度提升算法(Adaptive Gradient Boosting Random Forest), 同时引入了更加多样的训练样本来降低诸如光线强弱, 旋转角度等外部环境对定位的影响. 在 BioID 数据库中, 此方法的准确率为 92%, 定位时间少于 1 ms. 实验结果表明, 它已经满足在大多数不可控环境对鲁棒性和实时性的要求.

关键词: 人眼定位; 随机森林; 自适应梯度提升

中图分类号: TP39

文献标识码: A

文章编号: 1000-7180(2016)07-0001-04

A Revised Eye-Localization Algorithm Based on Random Forest

MA Yue, HE Guang-hui

(School of Microelectronics, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: In this paper, we present an enhanced Random Forest(RF) model for precise eye localization. To extend Random Forest, we 1) propose the randomized trees with adaptive gradient boosting for a more accurate eye localization 2) introduce a series of standard samples with random perturbation for the robustness to changes in illumination and head pose and eye rotation. Performance results of our methods showed that it can obtain an accuracy of 92 percent on BioID database and gives a frame processing time of less than 1ms because of the low computation cost. Experimental results on the challenging BioID database show that our model can locate eyes accurately and efficiently under a broad range of uncontrolled variations involving lightings, camera qualities, occlusions, etc.

Key words: eye localization; random forest; adaptive gradient boosting

1 引言

人眼作为面部最具特征的一个部分, 在人与机器的交流中占据非常重要的作用. 随着科技的发展, 它在身份鉴别、汽车防疲劳驾驶、视觉关注度检测等领域有着愈来愈重要的价值. 人眼定位是指从输入图像中提取出瞳孔的中心的位置. 尽管人眼有着十分明显的特征, 但是人眼本身并不是规则的几何图案, 并且变化的环境带来的极端情况对人眼定位产生新的挑战, 同时, 为了增强用户的体验, 当前对人眼定位的准确性和实时性提出了更高的要求.

人眼定位的方法主大致分为三类: 基于模板算法、基于几何特征算法以及基于机器学习的算法. K. Peng 使用了基于积分投影模板的算法^[1], 利用积分的局部极值中定位双眼的位置, 这种方法容易受到光照等外部条件的限制. R. Valenti 提出了利用面部的灰度等高线曲率的几何特征, 并以曲度为权重对等高线中心进行投票来确定瞳孔的位置^[2]. 近些年, 机器学习在人眼定位研究中得到了广泛的应用. 机器学习算法通过自身的训练过程得到人眼的特征. 其中基于 Haar 特征与 Adaboost 分类器通过级联多个简单分类器最终组合为一个强分类

器^[3]. 与 Adaboost 类似, 组合方法 (Ensemble methods) 如随机森林 (Random Forest) 也被应用于人眼定位. 针对其鲁棒性不强、准确度不高、收敛速度慢等问题, 提出了基于自适应梯度提升的随机森林模型以及更加多样化的样本处理方法. 实验验证了算法的性能.

2 随机森林算法与特征提取

随机森林算法是由 Leo Breiman 于 2001 年提出的, 随机森林是集成技术中的一种 (Ensemble Techniques), 其思想就是训练出的多个弱分类器, 并且将这些弱分类器分别得到的结果平均起来作为最终位置. 所以, 可以把每一颗决策树 (Decision Tree) 比作弱分类器, 而随机森林就是由这些树结合起来的强分类器. 本文的决策树特指回归树 (Regression Tree). 同时, 为了使不同分辨率和尺寸的图片可以得到一致的结果, 并且以相同的方式提取特征, 本方法将使用归一化的坐标体系, 将图片的左上角、左下角、右上角、右下角分别定义为 $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$.

如前文所述, 一颗决策树可以看做一个弱分类器, 或者是定位人眼的函数. 其输入为包含人眼的图片, 即输入图片的特征, 输出为人眼的坐标. 这个函数可以表示为 $y = f(x)$, 为输出的定位坐标, $y = (y_1, y_2)$ 为定位结果, $x = (x_1, x_2, \dots, x_R)$ 为输入图像的 R 维特征. 显然, 最直接的特征就是图像中每个像素点的灰度值. 本文采用的决策树是二叉树, 并且一般为满二叉树, 在根节点和每一个中间结点训练并保存基于不同点坐标之间灰度差分 $I(I_1) - I(I_2)$ 和一个阈值 T , 定义对图像 I 的二值测试 (binary test) 为

$$\text{bintest}(I; I(I_1), I(I_2), T) = \begin{cases} 0, & I(I_1) - I(I_2) < T \\ 1, & \text{otherwise} \end{cases}.$$

式中, $I(I_i)$ 是在坐标 I_i 上的灰度值; I_i 均被归一化为 $[0, +1]$.

假设树的深度为 d , 所以经过 d 次二值测试后, 回落到 2^d 个叶节点之一, 就得到了这次判定的人眼坐标. 通过训练样本得到每个节点保存的值, 而这些值都是对准确坐标的一个估计, 这些估计的聚集程度越高, 说明可信度也越高, 所以可以用误差的平方和来衡量分类的效果. 因此设定训练的目标为最小化误差:

$$Q_{\text{node}} = \sum_{p \in S_l} \left\| p - \frac{\sum_{q \in S_l} q}{|S_l|} \right\|^2 + \sum_{p \in S_r} \left\| p - \frac{\sum_{q \in S_r} q}{|S_r|} \right\|^2.$$

式中, S_l 和 S_r 分别为按照某种特征和阈值分类产生的左、右子节点中坐标组成的集合. 训练停止的条件为到达指定树的深度.

3 提出的基于梯度提升的决策树

定位精确与否是由训练得到的模型决定的. 训练的目的是最下滑损失函数, 其思想是不断地向损失函数的负梯度方向前进, 希望能够最快地得到最小的损失函数.

梯度提升决策树的算法如下所示.

Step1: 初始化拟合误差函数.

$$F_0(x) = \argmin \sum_{i=1}^N L(y_i, F(x_i)).$$

Step2: 计算向负梯度方向计算伪残差, 即

$$\tilde{y}_i = -[\partial L(y_i, F(x_i)) / \partial F(x_i)]_{F(x) = F_{m-1}(x)}, \quad i = 1 \dots N.$$

定义损失函数为

$$L(y_i, F(x_i)) = 1/2 (y_i - F(x_i))^2$$

代入得到

$$\tilde{y}_i = y_i - F(x_i).$$

Step3: 更新拟合伪残差.

$$\alpha_m = \argmin_{\alpha, \beta} \sum_{i=1}^N \|\tilde{y}_i - \rho h(x_i; \alpha)\|^2.$$

式中, $h(x_i; \alpha)$ 为拟合结果, 例如第一次的拟合结果为 $h(x_i; \alpha_1)$.

Step4: 更新拟合残差乘数.

$$\rho_m = \argmin \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \rho h(x_i; \alpha_m)).$$

Step5: 更新模型, $m = 1 \rightarrow M$, M 次后迭代结束.

$$F_m(x) = F_{m-1}(x) + \rho_m h(x; \alpha_m).$$

梯度提升算法的另外一个重要元素是引入学习率 (learning rate), 或者称为收缩因子 (shrinkage index). 式(9) 更新为

$$F_m(x) = F_{m-1}(x) + \gamma \rho_m h(x; \alpha_m).$$

式中, $1 < \gamma \leq 0$, γ 即是学习率, 当 $\gamma = 1$ 时为原算法, γ 的大小决定了迭代的收敛速度. 当 γ 接近 1 时, 第一次的拟合结果就已经比较接近目标值, 此时残差已基本被跳过. 因此通过缩小学习率可以逐渐放慢收敛速度, 残差也会起到建模所需的作用. 但学习

率取到很小时,虽然性能得到了提升,但使得收敛速度很慢,同时增加了训练和预测过程的计算消耗.因此需要选择一个适中的学习率.

显然单颗决策树一定是一个弱分类器,运用集成技术将这些树组成随机森林,形成强分类器.随机森林的思想就是用过差异化地训练决策树,保证训练结果的多样性,减小系统性误差.相比传统的随机森林通过随机的“正负抵消”的方式消除误差,改进的梯度决策树通过主动在提升最快的方向上对误差进行拟合,当决策树出现系统性误差,之后的决策树便会进行弥补.对于那些难以拟合的随机偏差,通过引入的学习率来控制收敛速度来解决.

图 1 模拟了由 4 颗决策树组成的定位器运用组合技术进行人眼定位的过程,眼球上的白点代表最终的最为结果.图 1(a)采用随机森林进行定位:因为是随机误差,所以会分布在准确位置的周围,只需将决策树的定位结果平均起来即可.图 1(b)和(c)分别展示了 $\gamma=1$ 和 $\gamma=0.4$ 时的梯度提升决策树,显然当 γ 较小时,步长减小,不断地修正前进方向,从而更精确地到达人样中心的位置,得到更好的效果.



图 1 运用集成技术进行人眼定位的过程

4 对集成方法以及训练样本的改进

类似 Adaboost 算法,在训练新的弱分类器时,会将上次分类错误的样本权重加强,同时上次分类正确的样本权重会降低.不同于 Adaboost 依靠分类正确与否来判断是否增加权重,本文通过精确定位结果的误差作为决定权重的依据,让那些定位结果偏差更远的样本具有更大的权重.采用直接将误差的欧式距离平方作为训练样本的权重 $\omega'_i = \|y_i - F_M(x_i)\|^2$,然后对 ω'_i 进行归一化,得到更新后的权重 $\omega_i = \omega'_i / (\sum_{i=1}^N \omega'_i)$. 所以只需要在每次迭代之前,即原算法 Step2 之前更新权重:

$$\omega_i = \|y_i - F_M(x_i)\|^2 / (\sum_{i=1}^N \|y_i - F_M(x_i)\|^2).$$

本文为了提高准确率,采用了多级定位结构,也称为金字塔结构,通过不断缩小定位的范围来提高精度:从一个固定的较大的 ROI (Region of Interest) 开始,依据当前的 ROI 定位结果作为下一级更小的 ROI 的区域中心,然后不断迭代直到返回最后一级 ROI 的定位结果,如图 2 所示.

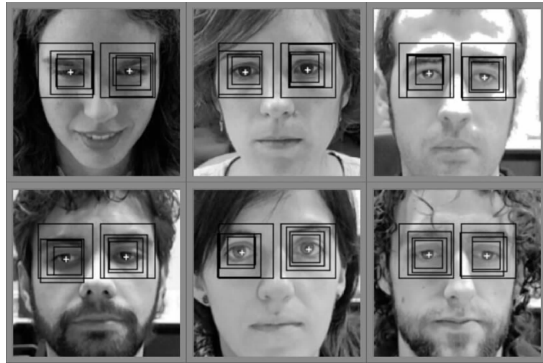


图 2 多级定位结构效果图

作为随机森林的“种子”,对训练样本的处理十分重要,为了训练出具有高定位精度和强鲁棒性的决策树,需要多样性的训练样本.本文采用业界比较通用的人脸库, BioID 和 LFW. 其中 BioID 库包括 1 项目 521 张黑白图片,分辨率 384×286 ,以正脸为主,该库中包含许多较难的环境,如光照、眼睛、闭眼等,是本领域最常用的数据库, LFW 库包含有 13 233 张分辨率为 250×250 的彩色图片,次库的图片姿态丰富.本文使用 LFW 作为训练集, BioID 是测试集.

为了使样本更加的丰富多彩,对 ROI 中心的横、纵坐标,以及尺寸这 3 个量加上 15% 的随机偏移,对 ROI 加以 15° 的随机转动.如图 3 所示.

5 实验结果与分析

5.1 人眼定位性能的评价标准

考虑到不同图像中的人眼大小不同,定位结果与标准位置之间的距离需要进行归一化,所以使用标定出的双眼距离作为归一化尺度,人眼定位的误差 (localization error) 定义为

$$\text{err} \triangleq \max(\|C_l - \tilde{C}_l\|, \|C_r - \tilde{C}_r\|) / \|C_l - C_r\| \triangleq \max(d_l, d_r) / \|C_l - C_r\|.$$

式中, C_l, C_r 分别为左眼、右眼在图像中的标准坐标位置; $\tilde{C}_l, \tilde{C}_r$ 分别是算法给出的坐标; d_l, d_r 分别为左右眼定为坐标与参考坐标之间的欧氏距离.当误差为 0.25 左右时,可以检测出人眼,当误差小

于 0.1 时,可以认为是在虹膜边缘的粗定位,当误差小于 0.05 甚至 0.025 时,误差大概是几个像素点的距离。



图 3 多姿态训练样本生成样例

5.2 实验结果

本文采用 BioID 作为测试集,实验表明采用集成技术之后性能得到提升。

本文提出的多级定位结果对模型的优化,如图 4 所示,采用了五级结构。可以看出随着搜索范围的不断缩小每次搜索的复杂度大大降低,性能也明显提高。

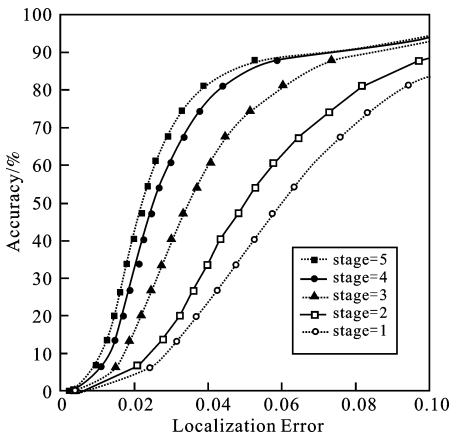


图 4 梯度提升决策树($\gamma=0.5$), 多级定位结构对定位性能的比较

图 5 给出了自适应算法对性能带来的提升. 同时将学习率调整为 $\gamma=0.2$,在梯度提升,多样性样本和多级定位的效果下,到达了比较好的效果. 本文最终采用了自适应梯度提升模型来进行人眼定位。

测试样本为整个 BioID 人脸库,运行在 CPU Core i7 下,鲁棒性和准确率达到了比较高的水平,实时性也达到了极好的表现. 表 1 给出了本文提出的人眼定位算法与其他最新方法的比较,其中测试集均是在 BioID. 在定位误差为 0.025 和 0.05 时,得到了最高的成功率。

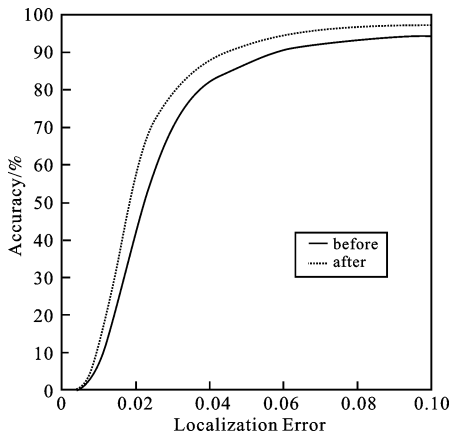


图 5 自适应算法对性能的提升($\gamma=0.2$)

表 1 本文方法与其他算法的比较

年份	定位方法	不同误差下的双眼定位成功率/%				
		err≤ 0.025	err≤ 0.05	err≤ 0.1	err≤ 0.15	err≤ 0.02
2015	本文	71	92	97.1	99.4	100
2014	文献[4]	61	89.9	97.1	98	98.9
2012	文献[5]	35	71.3	95.8	98.6	99.3
2011	文献[6]	54	81.5	99	99.5	99.8
2010	文献[7]	13	47	86	89	93
2009	文献[8]	30	81.5	98	99	99
2008	文献[9]	27	65	87	—	95
2008	文献[10]	45	77.2	82.1	86.2	93.8

6 结束语

本文在传统的随机森林算法的基础上提出了一种准确性较好、鲁棒性较强,同时实时性也达到要求的自适应梯度提升算法. 本文最终采用了 25 棵决策树组成的随机森林,学习率为 0.2,每张图片人眼定位时间小于 1 ms,在 BioID 库中,误差 0.05 的条件下,准确率为 92%,获得了较好的效果。

参考文献:

[1] Peng K. A robust algorithm for eye detection on gray intensity face without spectacles[J]. Journal of Computer Science & Technology, 2005,5(3):127-133.

[2] Valenti R, Gevers T. Accurate eye center location and tracking using isophote curvature[C]// Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference on. USA, Alaska, 2008: 1-8.

[3] Fasel I, Fortenberry B, Movellan J. A generative framework for real time object detection and classification[J]. Computer Vision and Image Understanding, 2005, 98(1):182-210.

22(2): 319-335.

- [4] Hirsch E A, Kojevnikov a. UnitWalk: A new SAT solver that uses local search guided by unit clause elimination[J]. *Annals of Mathematics and Artificial Intelligence*, 2005, 43(1): 91-111.
- [5] Dasgupta D, Nilo L F. Immunological computation: theory and applications [M]. Boca Raton: CRC Press, 2008.
- [6] Kelsey J, Timmis J. Immune inspired somatic contiguous hypermutation for function optimization[C]//Genetic and Evolutionary Computation Conference. Heidelberg, Springer, 2003: 207-218.
- [7] Jansen T, Zarges C. A theoretical analysis of immune inspired somatic contiguous hypermutations for function optimization[C]//Proc of the International Conference on Artificial Immune Systems, ICARIS. Berlin, Springer, 2009: 80-94.
- [8] Jansen T, Oliveto P S, Zarges C. On the analysis of the immune-Inspired B-Cell algorithm for the vertex cover problem[C]//ICARIS 2011. Springer, Heidelberg, 2011: 117-131.
- [9] Jansen T, Zarges C. Computing longest common subsequences with the B-Cell algorithm[C]//Proc of the

11th International Conference on Artificial Immune Systems (ICARIS '12). Italy, Taormina, 2012: 111-124.

- [10] Wegener I. Methods for the analysis of evolutionary algorithms on pseudo-Boolean functions[C]//In Evolutionary Optimization. Norwell, USA, Kluwer Academic Publishers, 2002: 349-369.
- [11] Neumann F, Witt C. Bioinspired computation in combinatorial optimization-algorithms and their computational complexity [M]. Series, Natural Computing Springer, 2010.
- [12] Droste S, Jansen T, Tinnefeld K, et al. A new framework for the valuation of algorithms for black-box optimization[C]//Foundations of Genetic Algorithms 7 (FOGA). Morgan Kaufmann, San Francisco, CA, 2003: 253-270.

作者简介:

夏小云 男,(1982-),博士,讲师.研究方向为计算智能与机器学习. E-mail:scutxxy@gmail.com.

周育人 男,(1965-),博士,教授,博士生导师.研究方向为计算智能、数据挖掘及社会网络.

(上接第4页)

- [4] Markuš N. Eye pupil localization with an ensemble of randomized trees[J]. *Pattern recognition*, 2014, 47(2): 578-587.
- [5] Valenti R, Sebe N, Gevers T. Combining head pose and eye location information for gaze estimation[J]. *Image Processing, IEEE Transactions on*, 2012, 21(2): 802-815.
- [6] Yi D, Lei Z, Li S Z. A robust eye localization method for low quality face images[C]// *Biometrics (IJCB)*, 2011 International Joint Conference on. USA, Washington DC, 2011: 1-6.
- [7] Asadifard M, Shanbezaheh J. Automatic adaptive center of pupil detection using face detection and cdf analysis[C]// *Proceedings of the International MultiConference of Engineers and Computer Scientists*. China, Hongkong, 2010: 3-5.
- [8] Tan X. Enhanced pictorial structures for precise eye localization under incontrolled conditions[C]// *Computer Vision and Pattern Recognition*, 2009. CVPR

2009. USA, Florida, Miami, IEEE Conference on, 2009: 1621-1628.

- [9] Kroon B, Hanjalic A, Maas S M. 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*. Canada, Niagara Falls, 2008: 379-388.
- [10] Valenti R, Gevers T. Accurate eye center location and tracking using isophote curvature[C]// *Computer Vision and Pattern Recognition, CVPR 2008*. IEEE Conference on. USA, Anchorage, Alaska, 2008.

作者简介:

马越 男,(1989-),硕士研究生.研究方向为数字图像处理、集成电路设计.

贺光辉(通讯作者) 男,(1980-),博士,副教授.研究方向为无线通信和多媒体系统的信号处理算法优化和 VLSI 架构设计. E-mail:guanghui.he@sjtu.edu.cn.