

基于级联 Adaboost 与示例投票的人脸检测

陈 骁,金 鑫

(南京航空航天大学 计算机科学与技术学院,江苏 南京 210000)

摘 要:传统的人脸检测算法在复杂背景、极端光照等非控条件下进行人脸检测的误检率较高。为有效降低误检率,文中提出一种级联 Adaboost 和示例投票的人脸检测方法。首先采用基于 LBP 特征的 Adaboost 算法初步定位人脸可能存在的区域,然后通过人脸示例集建立字典,使用稀疏编码的方法利用示例人脸对这些候选区域进行中心位置投票,根据得票数得到判别结果,排除非人脸区域,最终完成人脸检测。该方法的创新在于将基于字典学习的稀疏编码和基于部件模型的目标检测相结合,级联传统的 Adaboost 算法,实现非控环境下的人脸检测。在两个数据集上的实验结果表明,该方法在保持较高检测率的同时,有效降低了误检率,且鲁棒性较好。

关键词:人脸检测;LBP 特征;Adaboost 算法;稀疏编码;示例投票

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2015)12-0018-04

doi:10.3969/j.issn.1673-629X.2015.12.005

Face Detection Based on Cascade Adaboost and Exemplar Voting

CHEN Xiao, JIN Xin

(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics,
Nanjing 210000, China)

Abstract: In the conditions of complicated backgrounds and extreme illumination, face detection based on Adaboost algorithm usually has a higher false positive rate. Present a cascade of two algorithms in this paper, Adaboost and exemplar-based voting, to detect face in static images which is able to reduce the false-positives efficiently. This method utilizes LBP as features and a cascade Adaboost classifier is used to detect faces, and a voting method based on sparse coding is used as the final classifier to verify face or non-face. The innovation of the proposed method lies in combining sparse coding and part based model for face detection. The experimental result shows that this method can detect face with high detection rate, suppressing the error detection rate, with high robustness.

Key words: face detection; LBP feature; Adaboost algorithm; sparse coding; exemplar-based voting

0 引 言

人脸检测是计算机视觉领域的一个经典问题。近年来,研究人员在这一领域取得了显著进展^[1-2]。然而,由于角度、光照、表情、遮挡等复杂因素,非控环境下的人脸检测问题仍然没有得到很好的解决。现阶段比较成功的人脸检测是采用统计学习的方法,使用的分类器主要有 SVM、神经网络和 Adaboost。Osuna 等^[3]首先将 SVM 算法应用在人脸检测中,具有良好的分类效果,但 SVM 的非线性能力比较差。后来,Rowley 等^[4]采用局部连接的神经网络来检测人脸,而神经网络的结构复杂,检测效率较低。Viola 等^[5]提出一种有效的框架,将 Haar 特征和级联 Adaboost 算法应用于人脸检测,使得该问题得到了里程碑式的突破,让

人脸检测真正进入实用。但是,Harlike-Adaboost 方法仍存在某些缺陷:首先所选取的特征是最原始的矩形特征,无法获取图像内在的纹理及形状特性;其次 Adaboost 获得的强分类器不一定是最好的分类器,往往在达到一定的检测率的同时,也会产生很高的误检率。因此,如何在保证高检测率的同时最大程度地抑制误检率,以及如何有效提高检测的鲁棒性是当前对 Adaboost 算法改进的重点^[6-10]。

基于霍夫投票的方法,Leibe 等^[11]提出了 ISM 隐形状模型,采用无监督聚类的方式对局部信息生成码本模型,在检测时利用广义霍夫变换对目标中心进行概率投票^[12]。Shen^[13-14]将 ISM 模型结合图像检索和判别学习应用于人脸检测和人脸对齐,取得了良好

收稿日期:2015-03-11

修回日期:2015-06-17

网络出版时间:2015-11-19

基金项目:江苏省科研创新基金(KYLX_0289)

作者简介:陈 骁(1990-),男,硕士生,研究方向为计算机视觉与机器学习。

网络出版地址:<http://www.cnki.net/kcms/detail/61.1450.TP.20151119.1107.010.html>

的效果。但在实际中由于检测目标可能位于整幅图片的任意位置,对全局进行投票需消耗大量时间,使得其很难达到实时要求。

文中提出一种结合基于 LBP^[15] 特征的 Adaboost 算法和基于示例投票方法的人脸检测。具体而言,首先提取 LBP 纹理特征,以低阈值的 Adaboost 进行检测,降低漏检率;然后利用示例投票对 Adaboost 级联输出的人脸候选区进行最后的检测。实验结果表明,此方法在保证检测率和检测时间的同时,有效降低了误检率。

1 基于 LBP 的 Adaboost 人脸检测

1.1 特征提取

人脸与一般的背景图像相比,其纹理更具有有一定的规律。相比于 Haar 特征, LBP (Local Binary Patterns) 可以更好地提取人脸纹理特征,且对光照具有一定的鲁棒性^[16],因而文中采用 LBP 特征来进行人脸检测。

LBP 特征根据邻近区域的灰度像素分布差异来表示对目标轮廓的特征化描述,计算过程如图 1 所示。

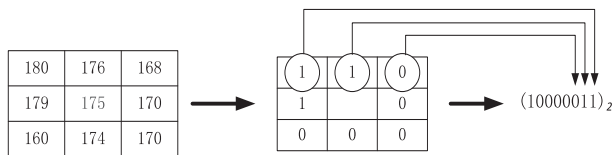


图 1 计算 LBP 示意图

在 3×3 的区域中,以中心像素为阈值,将其与邻近的 8 个像素进行比较。若中心像素小于邻近的像素则该位置的值置 0,否则置 1,然后按照次序计算出关于中心像素的二进制串表示,并转换成十进制值,即为其 LBP 值,并用这个值来反映该区域的纹理信息。文中采用 64×64 的基本检测窗口,将其分为 16 个单元,每个单元的大小是 16×16 像素,接着再把两个相连的单元重叠地组合成 1 个块并从中抽取 LBP 特征。这样得到的每个检测窗口有 3×3=9 个块,特征维数是 256×9=2 304。最终可以得到如图 2 所示的 LBP 图像。



图 2 输入图像及其 LBP 表示

1.2 Adaboost 算法

Adaboost^[17] 是一种具有自适应特性的 Boosting 学习算法,自动地从弱分类器空间中挑选出若干个弱分类器集成为一个强分类器。通过学习一系列的弱分类

器,每个弱分类器通过给予错误分类样本更大的权重,自适应地调整弱分类算法的错误率,经过若干次迭代后错误率能达到预定的期望值。

Adaboost 算法的具体描述如下:

设输入 n 个训练样本 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), x_i \in X, y_i \in \{0, 1\}$ 。已知训练样本中有 m 个负样本, l 个正样本,那么第 j 个特征生成的弱分类器为:

$$h_j(x) = \begin{cases} 1 & p f_i(x) < p_j \theta_j \\ 0 & \text{其他} \end{cases} \quad (1)$$

其中, h_j 为弱分类器的值; θ_j 为阈值; $f_i(x)$ 为第 i 个样本 x_i 的特征值。

(1) 对误差权重初始化。对 $y_i = 0$ 的样本,权重 $w_{1,i} = 1/2m$; 对应 $y_i = 1$ 的样本,权重 $w_{1,i} = 1/2l$;

(2) 对每轮 $t = 1, 2, \dots, T$ (T 为训练次数):

(1) 归一化。

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}} \quad (2)$$

(2) 对每个特征 j 训练相应的弱分类器 h_j , 即确定式(1)中的阈值 θ_j 以及偏置 p_j , 对应权重, 计算误差

$$\varepsilon_j = \sum_i w_i |h_j(x_i) - y_i| \quad (3)$$

(3) 选择误差最小的弱分类器 h_t 。

(4) 更新权重。

$$w_{t+1,i} = w_{t,i} \beta_i^{1-e_i} \quad (4)$$

其中, $\beta_i = \frac{\varepsilon_i}{1 - \varepsilon_i}$; 若正确分类则 $e_i = 1$, 否则 $e_i = 0$ 。

(5) 设 $\alpha_t = \log(1/\beta_t)$, 则最终 T 个弱分类器组成一个强分类器:

$$H(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{其他} \end{cases} \quad (5)$$

级联结构的 Adaboost 分类器如图 3 所示。

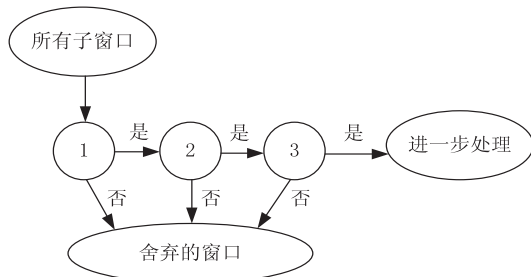


图 3 级联检测器示意图

每一级的分类器都采用 Adaboost 算法训练得到, 其中最初的几层使用较少的特征, 能快速排除大部分的负例样本, 而后面几层包含更多的弱分类器, 具有更强的分类能力。这样的层次结构使得人脸检测中大部分的背景区域在简单层次中被排除, 剩下少部分候选

区域通过复杂层次进行分类,保证了较高的检测速度,同时提高了检测精度。

由于最终得到的检测器采用级联分类器的结构,因此分类性能的好坏由整个级联分类器共同决定。默认的 Adaboost 阈值旨在训练过程中产生低错误率,一般而言,一个较低的阈值会产生更高的检测速率和更高的误检率。

文中在算法设计时采用低阈值的训练方法,放宽对误检率的要求,以降低漏检率,同时使得每一级弱分类器的个数减少,提高了检测速度。而对于在这里产生的误检样本,交给下一阶段通过投票方法将其排除,从而整体提高检测器性能。

2 基于示例投票的人脸检测

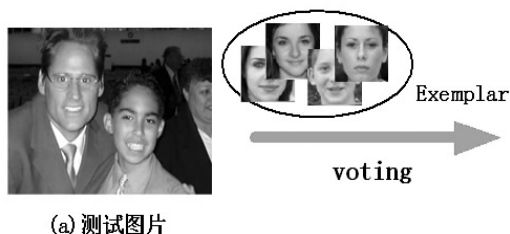
文中根据 ISM 模型的思想,利用训练样本来估计每个人脸局部区域与人脸中心的相对位置,通过对人脸各部分的检测和各部件到人脸中心相对位置关系对人脸中心进行投票。直觉上,虽然每张人脸外观上各有不同,但人脸各部分(如眼睛、嘴巴等器官)与人脸中心的相对位置没有太大偏差,所以采用人脸局部区域对中心位置的投票结果来检测人脸。

2.1 示例数据集

为了使用投票的方式来检测人脸,文中使用 500 张涵盖不同角度、表情、光照条件下的人脸建立了示例数据集。图 4 展示了一些示例人脸的例子。这些人脸数据大部分来自 LFW 数据集和 AFLW 数据集所标注出的人脸,这些图片较完整地包含了人脸的各种外观变化。



图4 包含各种不同的外观变化的人脸示例样例



(a) 测试图片



(b) 投票结果



(c) 投票响应图

图5 基于示例投票的检测结果

2.2 投票策略

在训练阶段,先将模板图片归一化为统一大小,采用基于 K -means 的方法对人脸示例局部区域特征进行聚类,将每个聚类中心作为字典的基,使特征由一个最接近的视觉单词来表示,得到具有较强判别能力的字典,不同的人脸局部区域对应生成字典中不同的基。与此同时,保持每个局部区域的位置信息,以在测试阶段进行投票。在投票阶段,通过训练阶段生成的字典得到图像中各待检测窗口的稀疏编码,并根据稀疏编码中对不同字典的基的响应值来判断待检测窗口是否属于人脸某一部区域。

对于给定的测试图片,用示例对测试图片进行投票时,采用文献[14]提出的相似度算法:

$$S(p, e_i) = \sum_k \sum_{\substack{f \in R(p), g \in e_i, \\ w(f) = w(g) = k}} \frac{\text{idf}^2(k)}{\text{tf}_{R(p)}(k) \text{tf}_{e_i}(k)} \quad (6)$$

其中, x 为测试图片的检测窗口; e_i 为示例数据集的第 i 个人脸图片; f 和 g 分别表示从 x 和 e_i 中提取出的局部特征; k 为字典中的第 k 个基; $\text{idf}(k)$ 表示逆向文件频率 (inverse document frequency); $\text{tf}(k)$ 表示词频 (term frequency)。

如果人脸示例中的某个特征 g 与检测窗口中的特征 f 相匹配,则这两个特征所属的相对位置也应该一致,这样通过计算出示例中特征 g 相对于示例中心点的距离,来估计检测窗口中特征点到检测窗口中心的距离。

在测试时,利用示例样本来估计人脸局部区域与人脸中心的相对位置。假设人脸示例某一块保留了该区域到人脸中心的相对位置,分块计算测试人脸与示例人脸的相似度,由示例给出人脸块的投票值,利用得到的相对位置关系对人脸中心进行投票。基于示例投票的结果及响应图如图 5 所示。

从图中可以看出,绝大部分投票结果是准确的,根据图中每个点的得票数可以得到投票响应图,投票响应图则较准确地反映出了测试图片中人脸中心的位置。

3 实验结果与分析

实验采用 CMU、FERET、MIT-CBCL 三个数据集中

的人脸作为正例样本,另外收集了 3 000 张不含人脸的自然图片,从中截取得到负例样本,这样共选取正例样本 4 820 张,负例样本 5 000 张。为便于训练,将所

有训练样本统一缩放到 64×64 像素的大小,然后提取 LBP 特征,再使用 Adaboost 训练分类器,调整每一级阈值使得漏检率尽可能减小。实验中,设置每一级最小检测率为 90%,最大的正例样本错误率为 50%,以尽可能降低这一阶段的漏检数。

在训练阶段先对所有人脸示例分块提取 SIFT 特征并用 K -means 聚类建立一个 1k 大小的字典。实验选取的分块大小为 20 像素,将人脸有重叠地划分为 $4\times4=16$ 个块,使得每个分块可以包含眼睛、鼻子、嘴巴等,训练和测试时均以此分块大小提取特征,有利于提升字典的判别能力。测试时首先将上一阶段检测出的人脸候选区域缩放成与示例人脸相同大小,计算候选区域块与示例人脸的相似度值,然后由相似度最大示例块给出测试图片块的投票策略,完成一次投票。所有示例投票结束后,得到最终的投票响应图,再根据中心位置的得票数判断该候选区域是否为人脸。最后,利用检测出的人脸窗口与真实人脸窗口的重叠面积来衡量检测精度,实验中规定当重叠区域面积大于真实人脸窗口面积的 50% 时为正确检测。

实验采用两个人脸数据集来对模型性能进行检测,分别为:Caltech 数据集,包括 450 张彩色图片,每张图片中包含一张正面人脸,图片环境比较复杂,包括室内室外、不同光照和某些遮挡;BioID 数据集,包括 1 521 张灰度图片,人脸的尺度、表情和视角都有较大变化。文中实验检测结果与基于 Haar 特征的 Adaboost 算法和基于 LBP 特征的 Adaboost 算法进行比较。表 1 和表 2 给出了在两种数据集上不同方法的检测率和误检数结果。

表 1 Caltech 数据集上不同方法的检测结果

方法	人脸数	正确检测数	检测率/%	误检数
Haar-like+Adaboost	450	442	98.2	90
LBP+Adaboost	450	440	97.7	88
文中方法	450	437	97.1	15

表 2 BioID 数据集上不同方法的检测结果

方法	人脸数	正确检测数	检测率/%	误检数
Haar-like+Adaboost	1 521	1 428	93.9	308
LBP+Adaboost	1 521	1 451	95.4	311
文中方法	1 521	1 402	92.2	86

从表中可知,文中方法在与经典方法检测率相当的情况下,有效抑制了误检数,取得了满意的效果。

图 6 展示了部分检测结果,其中(a)、(c)为传统 Adaboost 检测的结果,正确检测出人脸的同时,产生了一些误检;(b)、(d)为文中方法得到的结果,可以看出,基于示例投票的方法成功排除掉了误检窗口。

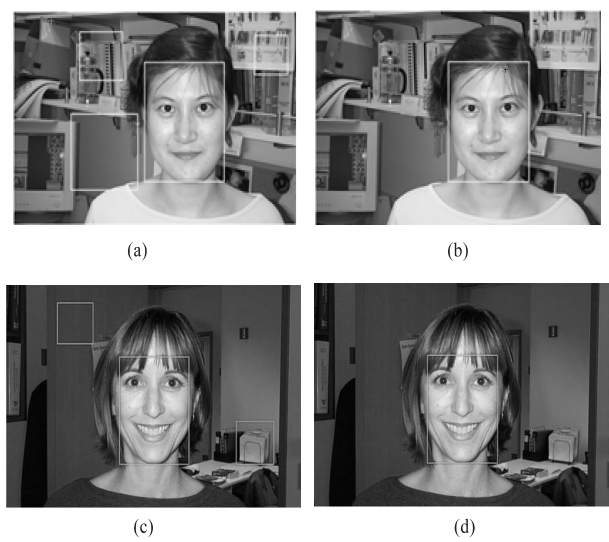


图 6 检测结果样例

4 结束语

文中对传统的人脸检测方法进行了改进,提出一种级联 Adaboost 与基于稀疏编码的示例投票的人脸检测方法。首先使用低阈值的 Adaboost 进行人脸粗检,缩小范围,产生候选区域,然后利用人脸示例训练出具有较强判别性的字典,通过稀疏编码的响应对这些候选区域投票进行验证,得到最终的人脸检测结果。实验结果表明,该方法在保持较高检测率的同时,有效抑制了误检率,对复杂背景环境具有一定的鲁棒性。然而,文中方法在人脸部分遮挡及角度偏转严重的情况下会出现较为明显的漏检,如何有效解决这个问题是后续工作的重点。

参考文献:

[1] 梁路宏,艾海舟,徐光祐,等.人脸检测研究综述[J].计算机学报,2002,25(5):449-458.

[2] 赵丽红,刘纪红,徐心和.人脸检测方法综述[J].计算机应用研究,2004,21(9):1-4.

[3] Osuna E, Freund R, Girosi F. Training support vector machines; an application to face detection[C]//Proc of IEEE conf on computer vision and pattern recognition. [s. l.]:IEEE,1997:130-136.

[4] Rowley H A, Baluja S, Kanade T. Neural network-based face detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1998,20(1):23-38.

[5] Viola P, Jones M J. Robust real-time face detection[J]. International Journal of Computer Vision,2004,57(2):137-154.

[6] 阮锦新,尹俊勋.基于人脸特征和 AdaBoost 算法的多姿态人脸检测[J].计算机应用,2010,30(4):967-970.

[7] 陈伟琦,梁一川,易强,等.基于肤色和 Adaboost 算法的人脸检测研究[J].计算机技术与发展,2008,18(12):44-46.

[8] 李明瑞,傅明,曹敦.基于肤色检测的 AdaBoost 人脸检测算法改进[J].计算机工程,2012,38(19):147-150.

度以及词语义距离相似度情况,有效规避了传统方法的不足,提高了关键词计算的精度,为文本特征的筛选提供了有效的方法。

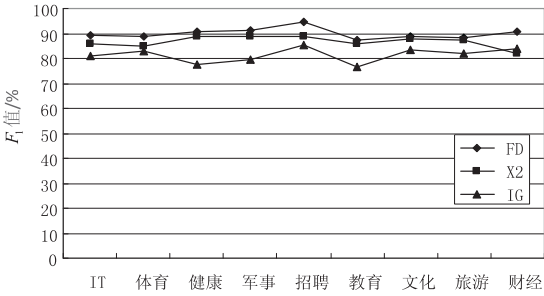


图 3 三种算法比值图

6 结束语

文中在针对应用 SimHash 进行文本相似检测时对文本特征降维,提出的关键词计算方法,综合考虑词项的词频分布、词项间的相关度以及相似度,提取的关键词能够有效地表达文本的核心内容,并能很好地区别文本。最后通过与 χ^2 、IG 算法进行对比,用实验验证了文中方法的可行性和有效性。下一步工作将引入语法分析和文本的语义相似度,进一步为文本特征提取降噪。

参考文献:

[1] 郭庆琳,李艳梅,唐琦. 基于 VSM 的文本相似度计算的研究[J]. 计算机应用研究,2008,25(11):3256-3258.
[2] 李连,朱爱红,苏涛. 一种改进的基于向量空间文本相似度算法的研究与实现[J]. 计算机应用与软件,2012,29(2):282-284.

(上接第 21 页)

[9] 王晶,杨煜. 基于边缘方向直方图的 Adaboost 人脸检测[J]. 计算机技术与发展,2007,17(12):5-7.
[10] 严云洋,郭志波,杨静宇. 基于特征空间划分的 AdaBoost 人脸检测算法[J]. 小型微型计算机系统,2007,28(11):2106-2109.
[11] Leibe B,Leonardis A,Schiele B. Combined object categorization and segmentation with an implicit shape model[C]//Proc of workshop on statistical learning in computer vision. [s. l.]:[s. n.],2004:17-32.
[12] Felzenszwalb P F,Girshick R B,McAllester D,et al. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(9):1627-1645.
[13] Shen Xiaohui. Object retrieval and localization with spatially-

[3] 和亚丽,陈立潮. Web 文本挖掘中的特征选取方法研究[J]. 计算机工程,2005,31(5):181-182.
[4] 代六玲,黄河燕,陈肇雄. 中文文本分类中特征抽取方法的比较研究[J]. 中文信息学报,2004,18(1):26-32.
[5] 程传鹏,苏安婕. 一种短文本特征词提取的方法[J]. 计算机应用与软件,2014,31(6):162-164.
[6] 吴双,张文生,徐海瑞. 基于词间关系分析的文本特征选择算法[J]. 计算机工程与科学,2012,34(6):140-145.
[7] 黄承慧,印鉴,侯昉. 一种结合词项语义信息和 TF-IDF 方法的文本相似度度量方法[J]. 计算机学报,2011,34(5):856-864.
[8] 翟延冬,王康平,张东娜,等. 一种基于 WordNet 的短文本语义相似性算法[J]. 电子学报,2012,40(3):617-620.
[9] 王立霞,淮晓永. 基于语义的中文文本关键词提取算法[J]. 计算机工程,2012,38(1):1-4.
[10] 耿焕同,蔡庆生,于琨,等. 一种基于词共现图的文档主题词自动抽取方法[J]. 南京大学学报:自然科学版,2006,42(2):156-162.
[11] 刘群,李素建. 基于《知网》的词汇语义相似度的计算[C]//第三届汉语词汇语义学研讨会. 中国台北:出版者不详,2002.
[12] 江敏,肖诗斌,王弘蔚,等. 一种改进的基于《知网》的词语语义相似度计算[J]. 中文信息学报,2008,22(5):84-89.
[13] Charikar M. Similarity estimation techniques from rounding algorithms[C]//Proc of 34th annual symposium on theory of computing. [s. l.]:[s. n.],2002.
[14] Manku G S,Jain A,Sarma A D. Detecting near-duplicates for web crawling[C]//Proc of international world wide web conference. [s. l.]:[s. n.],2007.
constrained similarity measure and k-nn re-ranking[C]//Proc of CVPR. [s. l.]:IEEE,2012.
[14] Shen Xiaohui. Detecting and aligning faces by image retrieval[C]//Proc of CVPR. [s. l.]:IEEE,2013.
[15] Ojala T,Matti P,Topi M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24(7):971-987.
[16] Liao Shengcai. Learning multi-scale block local binary patterns for face recognition[M]//Advances in biometrics. Berlin:Springer,2007:828-837.
[17] Friedman J,Tibshirani R,Hastie T. Additive logistic regression;a statistical view of boosting (with discussion and a rejoinder by the authors)[J]. The Annals of Statistics,2000,28(2):337-407.

基于级联Adaboost与示例投票的人脸检测

作者：[陈骁](#)，[金鑫](#)，[CHEN Xiao](#)，[JIN Xin](#)

作者单位：[南京航空航天大学 计算机科学与技术学院](#), 江苏 南京, 210000

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2015, 25(12)

引用本文格式：[陈骁](#). [金鑫](#). [CHEN Xiao](#). [JIN Xin](#) [基于级联Adaboost与示例投票的人脸检测](#)[期刊论文]-[计算机技术与发展](#) 2015(12)