摘　要

人脸检测肩负着在一幅输入图像中搜索并确定人脸是否存在以及如果存在，确定人脸位置和大小的任务。本文针对当前主流的基于统计学习和瀑布型分类器检测框架的人脸检测方法，在收集到更大规模样本集的基础上，重新对当前主流的瀑布型分类器检测和训练技术进行了系统分析。为了构建面对更加复杂样本时速度更快、精度更准的检测器，对主流瀑布型分类器训练和使用方法进行一系列的改进。这些方法可以概括为三个方面：降低瀑布型分类器的训练负担并提高检测精度，提高瀑布型分类器的检测速度，优化瀑布型分类器的速度\精度的折中调节算法。

首先，为了解决存在大规模训练人脸样本集和非人脸样本集的情况下，瀑布型分类器训练所需要的时间代价过高的问题，提出了训练过程中对正例样本和反例样本交叉自举的矩阵式学习算法。简单说来，对于瀑布式分类器，矩阵式学习并不仅在相邻的子分类器间对巨大的反例样本集进行“自举”，在每个子分类器的学习过程中，对大规模正例样本集也进行自举以降低学习中直接使用的训练正例样本集大小。当然，矩阵式学习并不只是简单的对正例样本和反例样本进行自举，针对矩阵式学习中的交叉自举，进一步提出了累加式样本自举方法和特征继承技术。在自举过程中，相邻的子分类器的共享信息被用来进行更加快速的学习，这些共享信息包括两个方面，分类器中已经学到的特征和已经自举得到的样本。

其次，提出一种更加高效的非人脸窗口预排除方法，进一步提高了人脸检测的速度，这包括了一种新的特征——局部组合二值特征，以及相应的以特征为中心的瀑布型检测算法。在提出的非人脸窗口预排除方法中，特征计算方式更加简单，避免可能重复的特征计算。局部组合二值特征的基本单元是二值特征，二值特征是图像上相邻区域间灰度差异的二值信息，即像 Haar 特征那样的多个区域间对比的更“黑”更“白”关系。局部组合二值特征是组合二值特征的一个有效子集，而组合二值特征表示的是多个不同二值特征值共生的信息。在局部组合二值特征中，组合的方式被限制在类似局部二值模式组合像素点时的方式，形成局部组合二值特征。局部组合二值特征数目少，描述能力强，被用来进行预分类器的学习。预分类器采用了一种类似“注意机制”的以特征为中心的瀑布型分类器运行技术，以特征为中心的运行机制首先针对某个显著特征模板对图像每个位置计算出特征值，构成特征图像，然后在穷举窗口上应用学习得到的本特征模板在各位置的统计信息，并进一步判断窗口为人脸或者非人脸。无论是赖以计算的特征模板还是窗口内特征模板的统计信息都使用学习的方法得到，具体来说，最终采用的模板是多个候选得到的最优模板，统计信息是 Boosting 学习得到的瀑布型分类器。

最后，提出一种基于搜索的瀑布型分类器精度\速度的折中算法。基于一个已有的 Boosting 瀑布型分类器，提出的方法能够构建多个速度标准下的瀑布型分类器，使得人

脸检测算法在应用中能够根据不同速度标准获得更优的分类性能。构建算法对构成瀑布型分类器的强分类器阈值空间进行搜索，以获得计算代价更优的瀑布型分类器。为了降低搜索的计算代价，搜索开始于正例分类正确率为百分之百的瀑布型分类器，以后每一轮的运行在当前瀑布型分类器基础上，搜索进一步损失小部分正例分类正确率条件下所有可能的瀑布型分类器，并获得当前计算代价最小的瀑布型分类器。然后，在当前获得的瀑布型分类器基础上，迭代运行，直到获得的瀑布型分类器正例分类正确率达到用户指定的正例分类正确率。

本文提出的方法在包括正面人脸检测和常用的多姿态范围内的人脸检测问题进行了实验，与别的已经公开发表过的相关工作相比，在已公开的标准人脸测试集上的实验结果表明，本文算法在精度方面具有优势，同时，在一幅 320×240 的图像上，多姿态人脸检测可以达到近实时的应用需求。

关键词：人脸检测；瀑布型分类器；二值特征；大规模样本集；自举

Research on Face Detection Based on Cascaded Classifier

Yan Shengye (Computer Application)
Directed By Gao Wen

Face detection task includes automatically searching in an input image and deciding whether or not there are faces, if the answer is YES, giving the size and location of each faces further. Provided with an enormous face dataset collected, aiming at finding more statistical results, we revisit the most popular Boosting Cascade method proposed by Viola and Jones. Under new conditions, both detection accuracy and speed need be improved. Seres of methods are proposed to solve the problem. In detail, these solutions are described as follow.

Firstly, aiming at promoting the classification performance of the face detector, we introduce enormous face training set into the learning process of cascaded face detector. To solving the difficulty of very high time cost of training, we proposed a novel training method named Matrix-Structural Learning (MSL) as an extension to Viola and Jones' cascade learning method for object detection. Briefly speaking, unlike Viola and Jones' method that learn linearly by bootstrapping only negative samples, the proposed MSL method bootstraps both positive and negative samples in a matrix-like structure. Moreover, an accumulative way is further presented to improve the training efficiency of MSL by inheriting features learned previously and training sample set bootstrapped previously during training procedure.

Secondly, we describe a novel fast and accurate face detection method. The whole detection procedure is partitioned into two phases. Firstly, a fast pre-classifier with high detection rate is run to reject a large number of simple non-faces. Then a more complex classifier with high detection rate such as the one learned from MSL is run to reject more complex non-faces. The novel techniques include the training and the running of pre- classifier. In the pre-classifier, firstly, a novel type of feature called the binary feature is proposed. The binary features modified Haar features to keep only the ordinal relationship rather than the absolute difference between the accumulated intensities. Further, several neighboring binary Haar features are assembled to capture their co-occurrence similar to the coding of local binary pattern (LBP). The assembled binary feature is called a Locally Assembled Binary (LAB) feature which is used to train the pre-classifier. Secondly, an efficient detection method called feature-centric method is adopted to build an efficient pre-classifier. Different from the feature-centric method, we further cascade the classifier learned by feature-centric method to improve the computation efficiency.

Finally, based on the ready-made Boosting cascade, we study how to build series of faster cascades with different computational cost for practical use. The new cascade is allowed to lose some detection rate. We propose to search the latent cascaded classifiers for the one with minimal computation cost. To unload the computation burden, the search algorithm is run sequentially forward. The search starts from the cascaded classifier with true positive rate 1. In each round of run, the best cascaded classifier is found out from all latent cascaded classifiers which are built with the same true positive rate a little smaller than the true positive rate of the selected cascaded classifier of the last round. The search algorithm is ended when target true positive rate is attained. Our contributions include three aspects. Firstly, a search-based method is proposed to build cascaded classifier. Secondly, we consider all possible cascaded classifier in the search-based method. Finally, different computational costs of different type of features in the cascaded classifier are considered in the evaluation of the computational costs of the cascaded classifiers.

To investigate the proposed method, experiments are committed on both frontal face detection and multi-view face detection. The experimental results on several public standard face test sets show that the classification accuracy of the proposed methods are comparable or better than the methods reported by others as to our knowledge. At the same time, the face detection system based on the proposed method has a nearly real-time detection speed on multi-view face detection.

Keywords: face detection, cascaded classifier, binary feature, enormous training set, bootstrap