

Fast Face Detection Technology Based on Adaboost Algorithm

Cheng Cui^{1,2a#}, Qichuan Tian^{1,2b}

1. Electrical Information Engineering Institute, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

2. Beijing Engineering Research Center of Monitoring for Construction Safety, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

Email: ^{a#} cuichengcc@126.com; ^b tianqichuan@163.com

Abstract

The principles of Adaboost algorithm are analyzed in detail, And, in turn, illustrating the Harr-like feature extraction method of Adaboost algorithm and realizing Harr-like characteristics extraction rapidly, greatly improving the human face feature extraction speed. Then illustrates the face and non-face threshold extraction method. Writing a vc++ program to implement Adaboost algorithm. Using MITEx face library. Through several rounds of training, getting a cascade-classifier. Finally making a experiment and the experimental results were analyzed.

Keywords: Adaboost Algorithm; Harr-like Features; Integral Figure; Face Detection

基于 Adaboost 算法的快速人脸检测技术*

崔成^{1,2}, 田启川^{1,2}

1. 北京建筑大学 电气与信息工程学院, 北京 100044

2. 北京建筑大学 北京市建筑安全监测工程技术研究中心, 北京 100044

摘要: 详细分析了 Adaboost 算法的原理, 依次说明了 Adaboost 算法的 Harr-like 特征提取方法, 并通过积分图的方法实现了 Harlike 特征的快速提取, 大大提高了人脸的特征提取速度, 然后说明了 Adaboost 算法的人脸与非人脸的阈值提取方法。并编写 c++ 程序对基于 Adaboost 算法进行了实现, 采用的人脸库为 MITEx 人脸库, 通过多轮的训练得到了一个级联的人脸检测分类器, 最后进行了实验并对实验结果进行了分析。

关键词: Adaboost 算法; Harr-like 特征; 积分图; 人脸检测

引言

人脸检测技术是一个复杂的极具挑战性的模式识别问题, 让计算机从一个复杂的变化多样的背景中快速准确的提取出我们所想要寻找的人脸是一个关键的技术难题, 基于特征匹配和模板匹配的方法具有简单直观等特点, 但在速度上并不能满足人们的需求, 相比其它的人脸检测算法, Adaboost 算法不仅在速度上有很大的优势, 而且在精度上达到很好的效果。本文利用 Adaboost 算法得出了一种快速的人脸检测技术。

1 Adaboost 算法的原理¹

1.1 Adaboost 算法概述

Adaboost 即 Adaptive Boosting, Adaboost 算法是由 Freund and Schapire (1995) 于 1995 提出的, 该算法

*基金资助: 该成果得到北京市教育委员会科技发展计划面上项目 (KM201410016016) 和北京建筑大学科学研究基金 (No.00331614021) 资助

不需要提前知道所有关于弱学习算法的先验知识，同时运算效率与 Freund 在 1991 年提出的 Boosting 算法几乎相同^[1]。取这个名字的原因是它能自适应的调整弱学习算法的错误率，经过若干次迭代后错误率能达到预定的期望值。另一方面，它不需要精确知道样本空间分布，在每次弱学习后调整样本空间分布，更新所有训练样本的权重，把样本空间中被正确分类的样本权重降低，而增加被错误分类的样本权重，这样下次弱学习时就更能更关注这些被错误分类的样本。

1.2 Adaboost 算法的理论分析

假设一个样本集合 $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$ ，共有 n 个样本，其中 x_1 到 x_n 为样本， y_1 到 y_n 对应每个样本的标号，正样本的标号为 1，负样本的标号为 0，假设需要进行 T 轮的训练，每一轮训练都会得到一个弱分类器，经过 T 轮的训练会得到 T 个弱分类器并构成一个强的分类器^[2]。

(1) 首先在训练的初始阶段需要初始化每个样本的权重，样本集合中一共有 n 个样本，假设正样本的个数为 p ，负样本的个数为 q ($p+q=n$)。

初始化每个正样本的权重为 $w_{1,i} = \frac{1}{2p}$ ，每个负样本的权重为 $w_{1,j} = \frac{1}{2q}$ 。在 T 轮训练中每一轮都要对本权重进行归一化处理设 $w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$ 。

(2) 每一轮训练结束后都会得到一个最佳的弱分类器设为 h_t (即为第 t 轮训练后的到的最佳弱分类器)；该弱分类器包括一个阈值 θ_t ，以及一个偏置 P_t 。 $h_t(x_i) = \begin{cases} 1, & P_t g_t(x_i) < P_t \theta_t \\ 0, & \text{其它} \end{cases}$ ， $g_t(x_i)$ 为样本的特征值，当 x_i 为正样本时 $h_t(x_i) = 1$ ，反之为 0，然后用这个弱的分类器对样本进行分类，计算相对于当前权重的分类误差：

$$\varepsilon_t = \sum_{i=1}^n w_{t,i} |h_t(x_i) - y_i|。$$

(3) 根据分类误差 ε_t ，更新每个样本的权重：

$$w_{t+1,i} = w_{t,i} \alpha_t^{\eta_i}，\text{其中 } \alpha_t = \frac{\varepsilon_t}{1 - \varepsilon_t}，\eta_i = |h_t(x_i) - y_i|$$

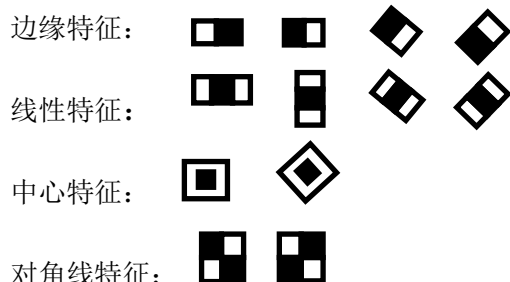
(4) 最后根据前 T 轮得到的 T 个弱的分类器，组合成一个强的分类器 $H(x)$ ：

$$H(x) = \begin{cases} 1, & \sum_{t=1}^T \beta_t h_t(x) > \frac{1}{2} \sum_{t=1}^T \beta_t \\ 0, & \text{其它} \end{cases} \quad \beta_t = \lg \frac{1}{\alpha_t}。$$

2 人脸检测中的 Harr-like 特征的提取

2.1 Harr-like 特征的概念

Harr-like 特征是各种不同尺寸大小的矩形，其主要分为四类，分别为边缘特征，线性特征，以及中心特征和对角线特征。下面举几个基本的 Harr-like 特征模板^[3]。



如上图所示，每个 Harr-like 特征内包含有白色的矩形区域和黑色的矩形区域，每个确定的 Harr-like 特

征的特征值等于白色矩形区域的像素和与黑色矩形区域像素和的差值，在确定了特征形式后 Harr-like 特征的数量就取决于训练样本图像矩阵的大小，特征模板在子窗口内任意放置，大小也会不断的变换，每一种形态称为一种特征，比方说在一个 20*20 的图像中，有 78460 个 Harr-like 特征。Harr-like 特征的数量为找出所有子窗口的特征是进行弱分类训练的基础。

2.2 Harr 特征的意义

不同形状的矩形特征对人脸的边缘，线性区域比较敏感，比如两只眼睛要比两眼中间的区域颜色更深一点，嘴巴要比其上下侧区域颜色要深，以及鼻子要比其两侧颜色更浅。这些特征都可以作为判断人脸的依据。但由于在一张很低的分辨率的图片（如下图我们所用到的训练样本为分辨率为 20*20）中包含有数量巨大的 Harr-like 特征，我们不可能把每一个特征都作为一个分类器，从以下图中我们可以清晰的看到所标出的 Harr-like 特征非常的典型，它们可以反映出来人脸某些区域的特征，我们的任务就是通过训练从这些大量的特征中筛选出来最能区分人脸与非人脸的 Harr-like 特征从而组成一个强的分类器^[4]。



图 1 典型的 Harr-like 特征

2.3 Harr-like 特征值的计算（积分图）

前文已经讲过 Harr-like 特征值的计算是由白色的矩形区域减去黑色的矩形区域的差值，但是一副 20*20 的分辨率的图像里都有将八万个 Harr-like 特征，如果把每个矩形特征里的像素值都一一相加然后白色矩形与黑色矩形彼此相减可想而知计算量会非常的巨大，为了提高特征值的计算速度，我们引进了积分图的概念。

2.3.1 什么是积分图

假设图像中的一个像素点(x,y)的像素值为 $p(x,y)$ ，则该点对应积分图上的值为该点到原点所组成的矩形中所有像素值的和，设为 $I(x,y)$ ，用公式表示即为：
$$I(x,y) = \sum_{x' \leq x, y' \leq y} p(x',y')。$$

也可以用迭代的方法计算出积分图，公式如下：

$$I(x,y)=I(x-1,y)+sc(x,y) \quad sc(x,y)=sc(x,y-1)+p(x,y)$$

其中 $sc(x,y)$ 为图像中 y 方向上所有像素之和，假设一副图像分辨率为 $m*n$ ，则用以上迭代方法只需要迭代 $m*n*2$ 次即可遍历完一幅图像^[5]。

2.3.2 利用积分图计算 Harr-like 特征的

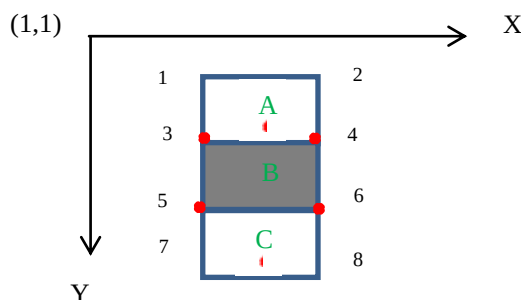


图 2 积分图像

上图中三个矩形框组成了一个线性的 Harr-like 特征，下面以此例子来计算该特征的特征值。

矩形区域 A 的像素和 $S_A = I_4 + I_1 - I_2 - I_3$

矩形区域 B 的像素和 $S_B = I_3 + I_6 - I_4 - I_5$

矩形区域 C 的像素和 $S_C = I_5 + I_8 - I_6 - I_7$

则该线性 Harr-like 特征值 $= S_A + S_C - S_B$

依次类推其它各类的 Harr-like 特征的特征值也是照此方法算出^[6]。

3 样本的训练

样本的训练是人脸检测中最重要的一步，它决定分类性能的高低，样本训练的目的就是为了从大量的训练样本中训练出一个强的分类器，这个强的分类器是通过多轮训练所得到的弱分类器组合而成的，每一个弱的分类器根据其分类性能的高低，都会赋予其相应的权值，每一个弱的分类器都是由阈值 θ ，以偏置项 p 组成（ p 是一个方向标识符取值为 1 或 -1）。

3.1 弱分类器的训练

假设我们需要进行 T 轮的训练，每一轮训练都要从大量的弱分类器中选出一个最优的弱分类器，首先在 t 训练中我们针对某个特征 f_t ，将所有样本的该特征的特征值进行排序，得到一个特征值由小到大的序列^[7]。我们要遍历整个序列中的每一个点，从其中找到一个最佳的点的值作为阈值使分类加权错误率最低，假设 S^+ 为所有正样本的权值之和， S^- 为所有负样本的权值之和， P^+ 为该序列某个点之前的所有正样本的权值之和， P^- 为该点之前的所有的负样本的权值之和。

为了找到那个最佳的阈值，我们通过以下图例进行分析。

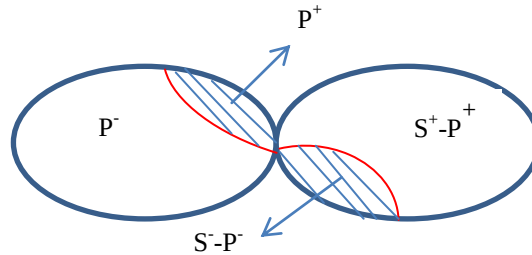


图3

我们将样本的加权和形象的用面积来表示，从上图我们可以看出要想使得分类的加权错误率最低，我们需要使得阴影部分的样本权值之和最小即阴影的面积最小，上图阴影左半部分为在某个点前的正样本加权之和 P^+ ，它也可以为负样本的加权和 P^- 。即为 $P^+ + (S^- - P^-)$ 或者 $P^- + (S^+ - P^+)$ 最小者， $e = \min\{P^+ + (S^- - P^-), P^- + (S^+ - P^+)\}$ 。然后将特征 f_t 的特征值序列从头到尾遍历一次，找到一个最小的 e 值记为 e_t ，对应序列中的那个点即为该特征的阈值，我们假设大于该阈值的点为正样本，反之则为负样本，如果该阈值点前所有正样本加权和大于 0.5，则偏置项 $p_t = 1$ ，反之则 $p_t = -1$ ^[8]。然后其它的所有特征也依次根据此方法进行计算，每一个特征都得到一个最优的错误率，然后从所有特征得到的 e 中找到最小的即为 $e_t = \min(e_1, e_2, \dots, e_i, \dots, e_N)$ （其中 N 为特征的数量）并记录 e_t 所对应的特征为 f_t ，以及该特征的分类阈值 θ_t ，以及偏置项 p_t 。

3.2 强分类器

经过 T 轮训练，每一轮都会得到一个弱的分类器 $h_t(f_t, \theta_t, p_t)$ 。

最终形成的强分类器 $C(x) = \begin{cases} 1, & \sum_{t=1}^T \beta_t h_t(f_t, \theta_t, p_t) \geq \frac{1}{2} \sum_{t=1}^T \beta_t \\ 0, & \text{其它} \end{cases}$ ，其中 $\beta_t = \lg \frac{1-e_t}{e_t}$ 。

3.3 级联分类器

为了能够进一步的提高检测速度，我们将训练好的多个强分类器组成一个级联分类器，在级联分类器中靠前的强分类器一般由十几个弱分类器组成，越靠后的强分类器越复杂，当检测人脸时，检测的大量的子窗口都会是非人脸，所以大量的非人脸区域会在级联分类器的前几个强分类器检测的过程中被滤除掉，可以想象一下，由于前面几个强分类器的所包含的弱分类器比较少，所以它们的检测速度会非常的快，而且每个强的分类器的检测率是非常高的，完全不用担心大量的人脸会被漏检，最后当所有的子窗口都被判定为人脸区域的时候则检测结束，通过这样一个级联的分类器最终进一步的提高了检测速度^[9]。

4 人脸检测

本文采用的人脸库为 MITEx 人脸库，其中人脸样本有 2706 个，非人脸样本有 4381 个，每张图片的分辨率都为 20*20。以下为该人脸库的部分人脸与非人脸。



图 4 部分人脸图像



图 5 部分非人脸样本

4.1 Harr-like 特征模板的选取

本文采用以下的 Harr-like 特征模板

$\text{HarrLike}\{1\}=[1 \ -1]$; $\text{HarrLike}\{2\}=[1 \ -1]$; $\text{HarrLike}\{3\}=[1 \ -1 \ 1]$; $\text{HarrLike}\{4\}=[1 \ -1 \ 1]$; $\text{HarrLike}\{5\}=[0 \ 1;-1 \ 0]$; $\text{HarrLike}\{6\}=[1 \ 0;0 \ -1]$; $\text{HarrLike}\{7\}=[1 \ -1;-1 \ 1]$; $\text{HarrLike}\{8\}=[1 \ 1;1 \ -1 \ 1 \ 1 \ 1]$;
 $\text{HarrLike}\{9\}=[0 \ 0 \ 1;0 \ -1 \ 0;1 \ 0 \ 0]$; $\text{HarrLike}\{10\}=[1 \ 0 \ 0;0 \ -1 \ 0;0 \ 0 \ 1]$;

例如 $\text{HarrLike}\{3\}=[1 \ -1 \ 1]$ ，其样式为：



其中 1 代表 Harr-like 特征的白色区域，其像素之和为正值，-1 代表 Harr-like 特征的黑色区域，其像素之和为负值，0 既不是白色区域也不是黑色区域，即为不参与 Harr-like 特征值计算的区域。通过对白色区域像素之和取正值，对黑色区域像素之和取负值，便可以求得整个 Harr-like 特征的特征值，即为白色区域像素之和减去黑色区域像素之和。从以上给出的模板我们可以看到，包含了线性特征，边缘特征，对角特征以及中心特征。每个 Harr-like 特征都包含有若干个小的黑色或白色的矩形，本文我们将这些基本的矩形边长最初设为两个像素，记为 $\text{basesize}=2$ ，然后边长依次以两个像素递增。

4.2 检测窗口的设定

训练时我们采用的训练样本图像的大小为 20*20，所以基准的检测窗口尺寸大小即为 20*20，一副图像的分辨率通常会远大于 20*20，所以检测窗口也要不断的变大，实验证明检测窗口每次放大 1.25 倍是最佳的。对于放大过的检测窗口，要想提取窗口中的特征我们需要把检测窗口归一化为 20*20，然后提取该窗口中的 T 个弱的特征，并输入强分类器进行判定。

每一个确定的检测窗口都要在图像中进行一次扫描，对图像上的每一个位置都要进行检测，所以我们需要设定一个确定的移动步长，移动步长的选择非常重要，它既影响检测的精度，也影响检测的速度。如果移动步长的选取过大，则检测速度会非常快，但是同时很多人脸由于检测器的移动步长较大最终被忽略；反之，如果移动步长选取的很小，虽然会提高检测的精度，但是检测速度会大大的下降。本文中，假设检测窗口的长度为 m，宽度为 n，本文采用在水平方向上的移动步长为 $m/5$ ，垂直方向上的移动步长为 $n/3$ 。每检测到一个人脸图像我们都会用一个矩形框框起来，在检测完一幅图片的时候我们会发现检测出的人脸上边会有很多的矩形窗口重叠在一起，这时我们只需要将所有重叠的矩形框的四个顶点取其平均值即可，

然后会得到一个简洁的单个矩形框。

4.3 实验结果及其分析

本文的实验平台是在 vs2010 上进行实验，用训练好的级联分类器对含有人脸的照片进行检测，所选取的照片都是平时所拍摄的随机的照片，照片拍摄的环境，地点，光照，以及人脸的姿态都有所不同，结果如下：



图 6 多人脸检测结果

图 6 中第一张图为 1033*700 分辨率的图片，可以看到所有的人脸都被检测出来了，第二张是分辨率为 400*271，但是有三张人脸漏检，从中我们可以发现当图片的分辨率较高的时候检测率会比较高，但是当图片的分辨率比较低的时候，其检测率会低于分辨率高的图片。第三张多人脸的图片，其分辨率为 962*665，其分辨率稍微高一点，其检测率还是很高的，漏检的人脸有两个。



图 7 检测结果

以上图片可以发现当检测人脸的侧脸的时候无法检测出人脸的，当人脸时横着的时候是无法检测出人脸的。这主要是因为我们的正样本中都是正脸，没有加入侧脸以及倾斜度很大的人脸样本，所以在检测的时候这类的人脸是无法检测出来的。

从以上实验结果可以看出，检测结果的好坏与图片的质量，以及人脸的姿态和图片摆放角度密切相关，但是从以上图片中可以看到几乎没有误检的人脸，这主要归功于级联分类器，经过多层的级联分类器，几乎所有的非人脸子窗口都被滤除去掉了，但是在实验中，由于图片大小以及背景的不同，也经常会出现误检的情况，但是误检率是很低的。

5 结论

本文利用 Adaboost 算法实现了人脸检测，其中在特征提取时积分图的引入大大加快了训练的速度，在正负人脸样本的训练时，Harlike 特征的阈值选取方法是关键的一步，每个 Harlike 特征阈值决定了最终形成强分类器的分类性能，最后为了能够提高分类器的检测速度，引入了级联分类器，使得检测速度有了明显的提升。实验结果表明在人脸上有遮挡物或则人脸姿态有明显倾斜，或侧脸的时候检测效果不佳，对于这一方面的问题，我们会在以后的实验中加入不同姿势以及各种带有遮挡物人脸的样本进一步做完善。

REFERENCES

- [1] Liang Luhong, Ai Haizhou. Review of Face detection research [J]. Journal of computer. 2002, 25 (5): 449-458
- [2] Viola P, Jones M. Robust real time object detection Technical Report[R].CRL 2001/01,Compaq Cambridge Reasearch Laboratory,

February, 2001

- [3] Liu Xiaoke, Sun Xiehua, Zhou Yongxia. Face detection from multiple perspectives based on the new characteristics of Haar - like [J]. Computer Engineering, 2009, 10
- [4] Bian Zhaoqi, Zhang Xuegong, Pattern Recognition[M].Beijing: Tsinghua university press, 2001, 01
- [5] Liu Xinping, Wei Qien. Probability and Mathematical Statistics [M].Xian: Northwestern polytechnical university press, 1997
- [6] Viola P. Rapid object detection using a Boosted cascade of simple features[J].In Proc. Audio and Video-based Biometric Person Authentication, 2003
- [7] Hjemas E. Face detection: A survey. Computer Vision and Image Understanding,2001,83(3): 236-274
- [8] Zhao Nan, Face detection based on Adaboost algorithm [Dissertation],Peking University, 2005.06.01
- [9] Lienhart R, Maydt J. An Extended Set of Harr-like Features for Rapid Object Detection[J].IEEE ICIP 2002,2002,1:900-903

【作者简介】



¹ 崔成（1989.08-），男，汉族，学士学位，人脸识别，2013 年考入北京建筑大学师从导师田启川研究人脸识别。

Email: cuichengcc@126.com



² 田启川（1971.3-），男，汉族，博士后，副教授，硕士研究生导师，研究领域包括图像处理、模式识别和智能控制等。山西孝义人，西北工业大学控制理论与控制工程专业博士毕业，获工学博士学位。现在北京建筑大学电信学院从

事教学、科研工作。Email: tianqichuan@163.com