

基于 AdaBoost 人脸检测算法中的特征筛选方法

钱力思

(西南大学信息与计算机科学学院,重庆 400715)

摘要 从机器视觉的发展历程出发,首先介绍了人脸检测的概念原理。在引入 AdaBoost 算法后,详细阐述了算法中的关于积分图、特征值的计算方法以及强分类器的具体训练过程。在对训练效率进行科学分析后,重点介绍了一种基于降低错误率的贡献度的特征筛选方法,以减少垃圾特征对系统的资源的消耗和不良的影响。最后给出人脸检测的实验结果并得出结论。

关键词 人脸检测 AdaBoost 算法 特征筛选 级联检测器

中图法分类号 TP301.6; **文献标志码** A

很早以前,人们为了让计算机能主动从真实世界中寻找信息,开始了对机器视觉的研究。迄今为止,计算机视觉的发展已经经历了一个漫长的过程,取得了不少的成果,为人们的生活带来了很多方便。而计算机视觉要处理的一个重要内容,就是对人脸的视觉处理,即人脸分析。

人脸分析,主要包括人脸检测(Face Detection)和人脸识别(Face Recognition)两部分。早期的人脸识别算法都是在假设已经得到了一个正面人脸或者假设人脸很容易获得的前提下进行的,但是随着人脸分析应用范围的不断扩大和开发实际系统需求的不断提高,这种假设下的研究不能满足需求。人脸检测开始作为独立的研究内容发展起来了。人脸检测是人脸分析的第一步,指对于任意给定的一幅图像,采用一定的策略对其进行搜索以确定其中是否含有人脸,如果含有则返回人脸的位置、大小和姿态^[1]。

1997 年,Freund 和 Schapire 提出了 AdaBoost 算法。AdaBoost 算法是在一系列矩形特征的基础上,通过 Boosting 算法由一系列“弱”分类器经过线性组

合而形成的“强”分类器。这些简单的分类器是由提取的矩形特征组成的,脸部的特征就能够由多个矩形特征简单地描绘并采用一定算法综合表示出来。现采用 5 种基本的矩形特征模板(图 1)。虽然这些特征的可控方向只是垂直或者水平的,但是这些以图像灰度值为自变量的矩形特征在边缘检测方面有着很好的效果,对脸部的各种特征也有较好的提取和编码能力。而它在计算效率上与可调控特征相比有着绝对的优势,减少了系统训练中的资源消耗。

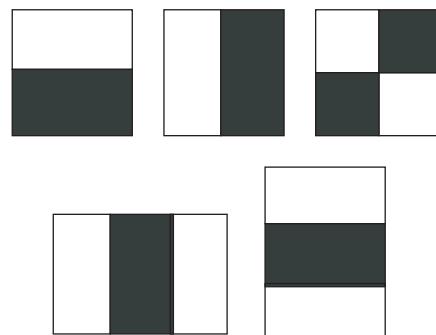


图 1 基本矩形特征模板

1 算法模型

1.1 积分图与矩形特征值的计算

Viola 在 2001 年的文章里首次引入了“积分

2010 年 8 月 17 日收到

作者简介:钱力思(1985—),男,重庆人,西南大学硕士研究生,研究方向:模式检测与识别,Agent 计算与电子商务。E-mail:
123513122@qq.com;Tplayerbaoyu@siha.com。

图^[1,2]的概念,大大加快了特征的计算速度。如图 2 所示,任意一坐标点 (x, y) 的积分图定义为图中位于该点左上方所有点的像素之和 $i_2(x, y) =$

$\sum_{x' \leq x, y' \leq y} i(x', y')$, 其中 $i_2(x, y)$ 表示像素点 (x, y) 的积分图, $i(x, y)$ 表示原始图像。ii(x, y) 可通过下面的迭代式进行计算

$$s(x, y) = s(x, y - 1) + i(x, y);$$

$$i_2(x, y) = i_2(x - 1, y) + s(x, y).$$

其中 $s(x, y)$ 表示行的积分和,且 $s(x, -1) = 0$, $i_2(-1, y) = 0$ 。设图像大小为 $m \times n$,则遍历一次图像需迭代 $m \times n \times 2$ 次,即可得到图像每个像素点的积分图。

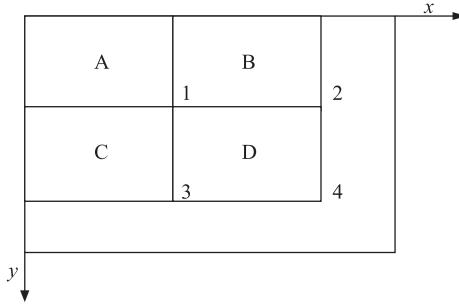


图 2 积分图的计算

如图 2 所示,区域 D 的四个顶点为 1、2、3、4,设区域 D 内的像素值合为 I_D ,则 $I_D = i_{24} + i_{21} - (i_2 + i_3)$ 。同理可以算得区域 A、B、C 内的各个像素合。以图 1 中的第一种特征模板为例,在图 2 中将区域 B 和 D 合成为一个区域作为一个特征,此特征的特征值即为:区域 B 的像素值 - 区域 D 的像素值,而 B、D 区域的像素值已经求得。而且可以推知,对每幅图像,只需遍历一次求得积分图矩阵,就很容易求得图像中所有矩形特征的特征值,因此特征计算的效率大大提高了。

1.2 AdaBoost 强分类器训练算法

(1) 给定一系列训练样本 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 其中 $y_i = 0$ 表示其为负样本(非人脸), $y_i = 1$ 表示其为正样本(人脸)。 n 为总的训练样本数量。

(2) 初始化权重 $w_{1,j} = D(i)$ 。

(3) 对 $t = 1, 2, \dots, T$ 。

$$\textcircled{1} \text{ 归一化权重 } q_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}} \quad (1)$$

\textcircled{2} 对于每个特征 f ,训练一个弱分类器 $h(x, f, p, \theta)$;计算对应所有特征的弱分类器的加权错误率 ε_f

$$\varepsilon_f = \sum_i q_i |h(x_i, f, p, \theta) - y_i| \quad (2)$$

\textcircled{3} 选取最佳的弱分类器 $h_t(x)$ 。

先计算拥有最小错误率的 ε_t

$$\varepsilon_t = \min_{f,p,\theta} \left\{ \sum_i q_i |h(x_i, f, p, \theta) - y_i| \right\} = \sum_i q_i |h(x_i, f_t, p_t, \theta_t) - y_i| \quad (3)$$

则最佳弱分类器为

$$h_t(x) = h(x, f_t, p_t, \theta_t) \quad (4)$$

\textcircled{4} 按照这个最佳弱分类器,调整权重

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad (5)$$

其中 $e_i = 0$ 表示 x_i 被正确地分类, $e_i = 1$ 表示 x_i 被错误地分类,

$$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t} \quad (6)$$

应说明 e_i 的意义。

(4) 最后合成的强分类器为^[3]

$$C(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{其他} \end{cases} \quad (7)$$

其中 $\alpha_t = \lg \frac{1}{\beta_t}$ 。

1.3 级联检测器

训练出的强分类器后,需要在级联的检测器中不断增加更多的强分类器^[4]。过程可以采用由粗到精的判断规则,这样可以很快地排除背景区域,从而集中精力对那些更像人脸的区域进行计算。

1.4 动态阈值特征筛选法

在分类器级联过程中,也存在一些问题。虽然实时系统重点关注的是实际检测的时间,但有时因为特征与样本空间过大而面临训练时间过长、对磁盘要求过大等问题而影响训练的可行性。仅以训练一个弱分类器为例。对于 5 种基本矩形特征模

板,大小为 20×20 像素的样本图像中特征总数量为 78 460 个(参照表 1),假设在每轮 AdaBoost 训练中,都要计算所有样本对所有特征的特征值,需要运算: $78\ 460 \times 7\ 087 = 556\ 046\ 020$ 次。在 PIII 850 MHz、512 Mb 的计算机上实测可得上述运算需花时 675 s。

因而,若要训练 100 个这样的弱分类器共需要 18.75 h,再加上每轮运算中进行的排序、遍历等运算,最终运算时间会更长。而实际系统中人脸和非人脸的样本空间远远超过 100 个,时间的开销是非常巨大的。

表 1 对应窗口大小的特征总数量

窗口数	36×36	30×30	24×24	20×20	16×16
特征数	816 264	394 725	162 336	78 460	32 384

若将每轮得到的特征值结果都保存到硬盘上以减少时间开销,采用 4 字节 unsigned long 存储,需要的磁盘空间为

$$556\ 046\ 020 \times 4 = 2\ 224\ 184\ 080 \text{ Bytes} = 2.07 \text{ GB}$$

空间开销也是巨大的。对于一般的 PC 机,IDE 接口的硬盘的峰值读取速度为 133 Mb/s,则读取上述内容的时间开销仍有 16 s。

为了提高训练速度,可以从很多方面考虑。如何采用多阈值划分^[5]来加速算法的收敛速度,还可以想到减少训练集数量和特征数量,而训练集数量不可能大量减少,这样会对训练强分类器的效果有很大影响,而特征在初始设计和选择时都有很大的自由度空间。因此可以采用一定方法减少特征数量。

现提出一种在训练过程中动态减少特征的方法,其基本思想是:在初始化特征时,可以多设计和选择一些可能对训练有益的特征,而在训练中不断地对其进行筛选。对错误率降低的贡献度较小的部分特征,被认为是对人脸和非人脸的判断力很低的,可以通过设定一个阈值范围,该阈值可以是随着训练集的变化而动态调整的,用于在每轮训练中动态地排除那些垃圾特征。在级联分类器中可以用以下的规则。

1.4.1 特征多,样本少

为尽可能避免将好的人脸检测特征排除,初始特征的选取应广泛,而样本空间不必太大,以减少运算开销。在后面的训练中,再不断的加入大量样本。

1.4.2 特征减少,样本增加

在每轮训练中,计算每个特征对降低错误率的贡献,并设一个定阈值范围,将贡献度过低的特征排除,达到特征筛选的目的。这样比开始仅凭经验筛选特征更有科学依据。其筛选方法如下。

设该轮训练中 K 个特征所进行的 T 轮训练中,每轮最小错误率为 $\varepsilon_{t,k}$ (对应前文算法中的(3)式),则相对错误率降低量

$$\Delta\varepsilon_{t+1,k} = \frac{\varepsilon_{t,k} - \varepsilon_{t+1,k}}{\varepsilon_{t,k}} \quad (8)$$

平均错误率为

$$\bar{\varepsilon}_k = \frac{1}{T} \sum_{t=1}^T \varepsilon_{t,k} \quad (9)$$

平均错误率降低量

$$\Delta\bar{\varepsilon}_k = \frac{1}{T} \sum_{t=1}^T \Delta\varepsilon_{t+1,k} \quad (10)$$

定义每个特征对于降低错误率的贡献度为

$$R_k = \frac{\Delta\bar{\varepsilon}_k}{\bar{\varepsilon}_k} \quad (11)$$

则平均贡献度为

$$\bar{R} = \frac{1}{K} \sum_{k=1}^K R_k \quad (12)$$

当一个特征的贡献度满足 $R_k < \bar{R}$ 时,可以说该特征在本轮的表现处于中下水平,但不一定被淘汰。所要淘汰的特征只是其中一部分。

采用动态阈值筛选法,设定阈值 φ 为: $\varphi = \lambda(\bar{R} - R_{\min}) + R_{\min}$;其中, R_{\min} 为本轮的最低贡献度, λ 是阈值的动态调整因子,设训练样本总数量为 N (包含正负样本),可取为 $\lambda = a \frac{\lg N}{N} + b$ 。 b 值作为最低淘汰率调整值,一般可取 0.05(即最低淘汰 5% 的特征)或零, a 为筛选严苛度, a 值越大特征代表筛选越严格,一般可取 0.6、0.8 或 1.0。通过计算贡献度,一个特征被淘汰掉的条件为

$$R_k < \varphi \quad (13)$$

最后,设定筛选系统的一个终止条件,即当最低贡献度已经达到很可观的水平时,允许暂停特征筛选。即设定阈值 φ 为最低必接受贡献度,当 $R_k > \varphi$ 时,让该特征通过筛选。而 φ 值按照不同的系统应有不同的选择。

1.4.3 达到稳定

在每轮训练后样本数量是不断增加的,因此阈值是动态变化的,筛选特征严格程度随样本数量的增加而降低,这样既降低了特征数量,节约了资源开销,也保证在后期能有一定数量的特征供参考。

根据实际需要,特征的数量不能过分减少,即使有的特征只有很低的贡献度。系统的稳定性可以通过设置特征筛选终止条件来保证,即特征数量必须保持在某一水平,当数量低于某值时停止筛选算法。一种人脸检测器训练与级联流程如图 3 所示。

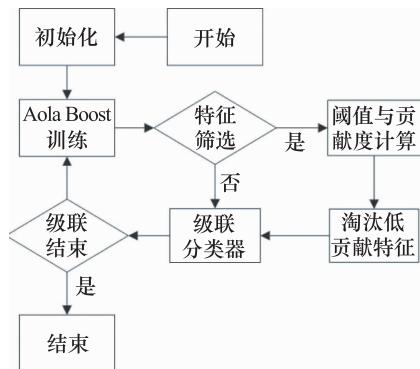


图 3 人脸检测器训练流程图

2 实验结果及分析

实际用于检测的图像较 20×20 的子窗口而言一般都具有较大的分辨率,对于这样不同尺度的检测,选择对检测器窗口进行缩放而不改变图像本身的大小。实验结果表明,一种较为高效的规则为检测器每次放大 1.25 倍。

此外,检测器平移和放大是密切相关的。设每一次检测平移 m 个像素,若放大的倍数为 k ,则检测窗口每次平移量为 km 。检测中 k 的取值按照人脸

结构的“三庭五眼”^[3]准则,选择 $k=4$ 。

对于重复检测的融合问题,采用消减法^[3,6]。即对于边界有相交区域的检测结果,对它们的四个顶点分别取平均值作为新的顶点,从而合成一个检测结果。实验结果由图 4—图 6 给出。

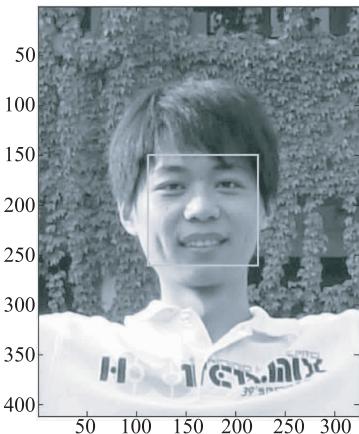


图 4 单人检测结果:总数 1;漏检 0;错检 0

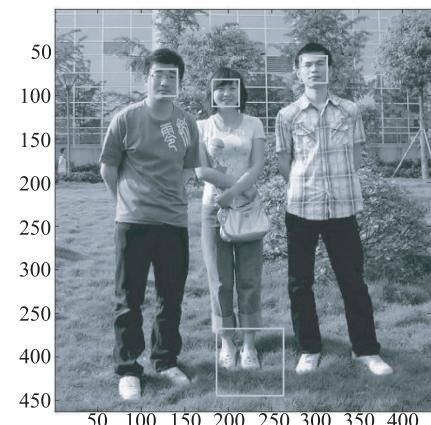


图 5 多人检测结果:总数 3;漏检 0;错检 1

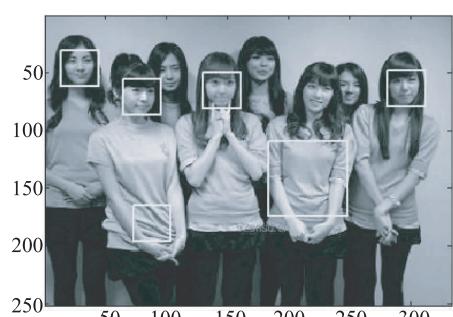


图 6 多人检测结果图:总数 8;漏检 4;错检 2

3 结语

重点是在理论框架上的研究,提出的方法在特征数量减少的作用上是明显的。而且文章的改进点位于样本训练中,可以和现今各种算法兼容使用,可能取得更好的检测效果。但由于条件所限,难以收集到足够人脸和非人脸样本进行大量的强分类器训练与级联,因此很难与 Viola 经典算法的训练效率进行比较。而对于较小的样本空间而言,在训练中本文设计特征的方法按照尽量广泛的原则,只是去除了宽带为一个像素的特征以及部分边缘特征,用于训练的初始特征数量为 46 953 个。后续的训练过程中,利用特征筛选方法,文章取 $a = 1$, $b = 0.05$, 总量为 1 000 的样本空间中,设定的样本初始数量及动态增加量均为 200。最后算法动态减少的特征数量为 28 580 个,而在减少了大量特征计算的基础上,本文的算法对于 20 万以下像素的图像,检测的漏检与误检率较低,偶尔出现的误检与

错检现象是因为训练集中的人脸与非人脸样本不足,特别是非人脸样本难以较完备地收集。因为实际中的非人脸样式的数量远远多于人脸样式。若拥有足够多的非人脸样本能使检测器对非人脸图像的判断能力大大增强。

参 考 文 献

- 1 Viola P. Rapid object detection using a boosted cascade of simple features. In: Proc IEEE Conference on Computer Vision and Pattern Recognition, 2001; 1(1): 511—518
- 2 Viola P, Jones M. Robust real-time face detection. International Journal of Computer Vision, 2004; 57(2): 137—154
- 3 赵楠, 查红彬. 基于 AdaBoost 算法的人脸检测. 北京: 北京大学物理学院视觉与听觉信息处理国家重点实验室, 2005
- 4 Viola P. Fast and robust classification using asymmetric adaboost and a detector cascade. In: Advances in Neural Information Processing Systems, 2002: 1311—1318
- 5 孙士明, 潘青, 纪友芳, 等. 多阈值划分的连续 AdaBoost 人脸识别. 计算机应用, 2009; 29(8): 2099—2014
- 6 边肇祺, 张学工. 模式识别(第二版). 北京: 清华大学出版社, 2001

Feature Selection Method in Face Detection Based on AdaBoost Algorithm

QIAN Li-si

(Southwest China Normal University, Chongqing 400715, P. R. China)

[Abstract] Begin with the machine vision development process, the concept of face detection is introduced at first. Then, after the introduction of AdaBoost algorithm is given. The calculation of points chart, eigenvalues and the strong classifier specific training process are described in detail. After the scientific analysis of the efficiency in the training, the article focused on reducing error rates based on the contribution of feature selection methods to reduce the rubbish features of the system resource consumption and adverse effects. Finally, the experimental results of face detection and a conclusion were reached.

[Key words] face detection adaBoost algorithm feature selection cascade detector