

文章编号:1001-9081(2006)03-0619-03

Adaboost 人脸检测方法的改进

魏冬生, 李林青

(中国科学院 研究生院, 北京 100040)

(wdongsheng@126.com)

摘要:针对 Adaboost 人脸检测训练非常耗时的问题,从训练中直接求解目标函数和弱分类器使用双阈值判决构造强分类器两个方面对人脸检测系统进行了改进。实验结果表明,改进后的系统使用的弱分类器数目大大减少,并且训练速度比传统方法高 11 倍左右。

关键词:人脸检测; 模式识别; 阈值; Adaboost 算法

中图分类号: TP391.41; TP181 **文献标识码:** A

Improvement of Adaboost face detection

WEI Dong-sheng, LI Lin-qing

(Graduate School, Chinese Academy of Sciences, Beijing 100040, China)

Abstract: Aiming at the problem that training time of Adaboost face detection is extremely long, two improvement methods were proposed: One method was to directly solve the parameter of single weaker classifier, the other was to introduce a double threshold decision to make stronger classifier. The experiment results show that the number of weaker classifiers needed in Adaboost face detection system updated is dramatically reduced and its training speed is about 11 times higher than that of the traditional method.

Key words: face detection; pattern recognition; threshold; Adaboost algorithm

0 引言

对人脸的检测识别是人类视觉最杰出的能力之一,在人机自然交互(Human-computer Nature Interaction, HCNI)、身份认证和安全防范等系统中有着重要的应用^[1,2]。在众多的检测方法中,Viola 和 Jones^[3]提出的 Adaboost 人脸检测方法,从根本上解决了检测的速度问题,同时有较好的识别效果。它利用一个只有 200 个关键特征的集合,就能达到 95% 以上检测率和仅 1:14 084 的错误率,检测一幅 380 × 280 像素的图像需要时间不到 0.7s,这样的性能使它成为人脸检测最好的方法。但是 Adaboost 算法本身训练非常耗时,其系统在高级工作站上完成最后训练花了数周时间。

1 Adaboost 人脸检测方法

Adaboost 人脸检测是一种基于积分图、级联检测器和 Adaboost 算法^[4]的方法,而 Adaboost 算法是它的核心内容。这个算法的基本思想就是将大量的分类能力一般的弱分类器通过一定方法叠加起来,构成一个分类能力很强的强分类器。理论证明,只要每个弱分类器分类能力比随机猜测要好,当弱分类器个数趋向于无穷时,强分类器的错误率将趋于零。Adaboost 用于人脸检测时,从人脸中抽取大量的一维简单特征。这些简单特征都有一定的人脸和非人脸区分性,最终系统使用数千个一维简单特征,组合起来达到很好的分类效果。

1.1 Adaboost 算法

设一个有 n 个训练样本 $(x_1, y_1), \dots, (x_n, y_n)$ 的训练集,其中 $y_i = \{0, 1\}$ ($i = 1, 2, \dots, n$) 对应样本的假和真,样

本中有 u 个假样本, v 个真样本;待分类物体有 k 个简单特征表示为 $f_j(\cdot)$, 其中 $1 \leq j \leq k$; 对于第 i 个训练样本 x_i , 它的特征为 $f_j(x_i)$ 。第 j 个特征的弱分类器 $h_j(x)$ 由一个特征 f_j , 一个阈值 θ_j 和一个指示不等式方向的偏置值 p_j (只有 ± 1 两种情况) 构成:

$$h_j(x) = \begin{cases} 1, & p_j f_j < p_j \theta_j \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

训练目标:通过对正反例进行分析,选择错误率最低的 T 个弱分类器,最终优化组合成一个强分类器。

训练过程如下:

1) 设 $\omega_{t,i}$ 为第 t 次循环中第 i 个样本的误差权重。对训练样本中的误差权重按如下公式初始化:对于 $y_i = 0$ 的样本, $\omega_{1,i} = \frac{1}{2u}$; 对于 $y_i = 1$ 的样本, $\omega_{1,i} = \frac{1}{2v}$ 。

2) For $t = 1, \dots, T$

权重归一化,使得 ω_t 为 $\omega_{t,i} \leftarrow \omega_{t,i} / \sum_{j=1}^n \omega_{t,j}$;

对于每个特征 j , 训练出其弱分类器 h_j , 也就是确定阈值 θ_j 和偏置 p_j , 使其目标函数 $\varepsilon_j = \sum_{i=1}^n \omega_{t,i} |h_j(x_i) - y_i|$ 达到最小;

从上述弱分类器中,找出一个具有最小的错误 ε_t 的弱分类器 h_t ;

对所有样本的权重进行更新 $\omega_{t+1,i} = \omega_{t,i} \beta_i^{1-e_i}$, e_i 确定方法:如果刚才 x_i 被 h_t 正确分类, $e_i = 0$; 反之 $e_i = 1$ 。 $\beta_i = \frac{\varepsilon_t}{1 - \varepsilon_t}$, 这样保证后面训练提取的弱分类器会更加强化对这

收稿日期:2005-09-18 收稿日期:2005-12-15

作者简介:魏冬生(1976-),男,湖南衡南人,硕士研究生,主要研究方向:人工智能、模式识别;李林青(1972-),男,江西兴国人,硕士研究生,主要研究方向:人工智能、图像处理。

些分类错误样本训练。

3) 最后求得的强分类器是:

$$h(x) = \begin{cases} 1, & \sum_{i=1}^T \alpha_i h_i(x) \geq 0.5 \sum_{i=1}^T \alpha_i \\ 0, & \text{otherwise} \end{cases}, \text{ 其中, } \alpha_i = \log \frac{1}{\beta_i}$$

1.2 Adaboost 人脸检测训练耗时原因

Adaboost 人脸检测训练耗时主要原因在于组成人脸检测系统的每个弱分类器要采用大量样本训练,并且系统使用了大量的弱分类器。从目前 Adaboost 人脸检测系统应用的情况上看,每个弱分类器一般都有上万个训练样本,训练弱分类器阈值和偏置非常耗时间。另外,为了得到一个有效的弱分类器,需要从所有的弱分类器中进行搜索,并且完成一次最优弱分类器提取后,样本的权重更新了,相当于训练样本概率分布完全变化了,因此再次训练弱分类器时,所有的弱分类器必须完全重新训练。假设人脸检测中需要 100 个弱分类器,有 50000 个特征,10000 个训练样本,训练每个弱分类器需要 0.6s,那么训练时间为 $50000 \times 100 \times 0.6s$,将近 35d 时间。

2 Adaboost 人脸检测系统改进

针对以上问题,本文对 Adaboost 人脸检测系统进行改进,着力解决人脸检测训练时间太长的的问题。其一在每个弱分类器训练中,本文的方法直接求解 Adaboost 训练中的目标函数,避免了其他训练方法中统计概率分布和迭代运算的耗时问题。其二利用弱分类器快速训练求解上的优势,提出使用双阈值的弱分类器来构造 Adaboost 强分类器的结构。采用双阈值弱分类器构造的系统,可以使用更少的弱分类器个数而达到同样的分类效果。

2.1 直接求解弱分类器错误率

求解一个特征 f_j 的弱分类器 h_j ,要确定阈值 θ_j 和偏置 p_j 。根据定义,弱分类器的错误率 ε_j 是 θ_j 和 p_j 的函数,记为 $\varepsilon_j(\theta_j, p_j)$ 。由于偏置 p_j 只有 ± 1 两种情况,可以分两种情况讨论:

1) 当偏置值 $p_j = 1$,根据式(1),特征 f_j 的值小于阈值 θ_j 时为真,则有:

$$\varepsilon_j(\theta_j, 1) = \sum_{i: h_j(x_i) \neq y_i} \omega_i = \sum_{i: f_j(x_i) > \theta_j, y_i=1} \omega_i + \sum_{i: f_j(x_i) \leq \theta_j, y_i=0} \omega_i \quad (2)$$

$$\varepsilon_j(\theta_j + \Delta, 1) = \varepsilon_j(\theta_j, 1) - \sum_{i: 1 < f_j(x_i) \leq \theta_j + \Delta, y_i=1} \omega_i + \sum_{i: \theta_j < f_j(x_i) \leq \theta_j + \Delta, y_i=0} \omega_i \quad (3)$$

将所有训练样本的特征 f_j 的值从小到大进行排列,得到一个次序表 s 。 $s(1)$ 表示训练样本中特征 f_j 的值最小样本的序号, $f_j(x(s(1)))$ 为它相应的特征所得的值。如以该值为阈值那么有:

$$\varepsilon_j(f_j(x(s(1))), 1) = \sum_{y_i=1} \omega_i \quad (4)$$

对于特征 f_j 的值次小的训练样本 $x(s(2))$,根据式(3),有:

$$\varepsilon_j(f_j(x(s(2))), 1) = \varepsilon_j(f_j(x(s(1))), 1) + \delta \omega(s(2)) \quad (5)$$

根据(3)式可得,当 $y(s(2)) = 1$ 时, $\delta = -1$; 当 $y(s(2)) = 0$ 时, $\delta = 1$,以此类推得:

$$\varepsilon_j(f_j(x(s(N+1))), 1) = \varepsilon_j(f_j(x(s(N))), 1) + \delta \omega(s(N+1)) \quad (6)$$

利用递推公式(6),可以求出所有不同阈值下的 $\varepsilon_j(\cdot, 1)$,从中获得最小的 $\varepsilon_{j \min}(\theta, 1)$,此时的 θ 值定为 θ_1 ,即 $\theta_1 = \theta_j |_{\varepsilon_j = \varepsilon_{j \min}(\theta, 1)}$,为相应特征 j 分类器的最优阈值。

2) 当偏置 $p_j = -1$,那么可以求得同样的一组 $\varepsilon_j(\cdot, 1)$,其值和 $p_j = 1$ 的一组值是互补的,即 $\varepsilon_j(\cdot, -1) = 1 - \varepsilon_j(\cdot, 1)$,于是有 $\varepsilon_{j \min}(\theta, -1) = 1 - \varepsilon_{j \max}(\theta, 1)$,此时的 θ 值定为 θ_2 ,即 $\theta_2 = \theta_j |_{\varepsilon_j = \varepsilon_{j \max}(\theta, 1)}$ 。因此只需要使用一次(6)式累加,然后比较 $\varepsilon_{j \min}(\theta, 1)$ 和 $1 - \varepsilon_{j \max}(\theta, 1)$,就可以求出相应的该弱分类器最佳的阈值和偏置,同时也可以得到相应的 ε_j 。

$$p_j = \begin{cases} +1, & \varepsilon_{j \min}(\theta, -1) < 1 - \varepsilon_{j \max}(\theta, 1) \\ -1, & \text{otherwise} \end{cases} \quad (7)$$

从本文所给出的求 ε_j 过程可以看到:改进后的 Adaboost 人脸检测系统,训练弱分类器的方法主要是一个权重累加过程,时间消耗小,算法中主要的耗时是在排序上面,如果使用快速排序算法^[5],训练时间会减小。并且改进后的算法中,训练样本自身的各个特征的值不会随着训练改变,改变的仅仅是样本的权重,因此针对样本自身特征的值排序只需要在算法初始化时排序一次,得到次序表 S ,以后可以反复使用。同时也可以看到,算法的时间复杂性只和样本的数量成线性增加关系。

2.2 弱分类器使用双阈值判决

在 Adaboost 人脸检测所采用的一些特征中,如果真样本具有对称性,而假样本具有非对称性,那么真样本的特征分布在绝对值较小的区域;当真样本具有非对称性,假样本具有对称性时,其分布呈相反状态,如图 1 所示。在这样的分布下,使用双阈值的弱分类器的分类效果更好。在以上算法基础上,可以方便地扩展到双阈值判决器。具体方法如下:

假设利用(6)式,得到 $\varepsilon_j(\cdot, 1)$,并利用所求得的 θ_1 和 θ_2 ,那么当 $\theta_1 \geq \theta_2$,对应于图 1(a),有:

$$h_j(x) = \begin{cases} 1, & \theta_2 \leq f_i(x) \leq \theta_1 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

当 $\theta_1 < \theta_2$,对应于图 1(b),有:

$$h_j(x) = \begin{cases} 1, & f_i(x) < \theta_1 \text{ or } f_i(x) > \theta_2 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

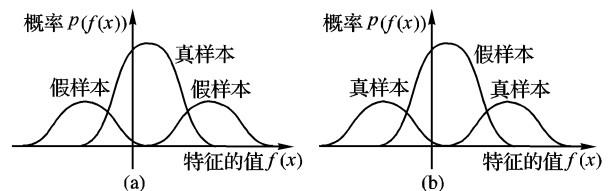


图 1 真样本和假样本特征的两种可能分布

(8) 式中 θ_1 为将真样本分类在坐标轴负方向的最佳阈值; θ_2 为将真样本分类在坐标轴正方向的最佳阈值,(8)式将两者结合起来,减少了错误率。以此类推(9)式,分类器错误率:

$$\varepsilon_j(\cdot) = \varepsilon(\theta_1, 1) - \varepsilon(\theta_2, 1) + 1 - \sum_{i: y_i=1} \omega_i \quad (10)$$

(10)式的每项在单阈值弱分类器中已经求出,因此将单阈值弱分类器扩展到双阈值判决,计算量没有增加。由于每个双阈值的弱分类器的错误率下降,分类能力更强,在 Adaboost 人脸检测中可以用较少的弱分类器组成强分类器,从而简化系统结构。这样一方面可以减少训练时间,另一方面也可以提高强分类器的检测(分类)速度。

3 仿真结果

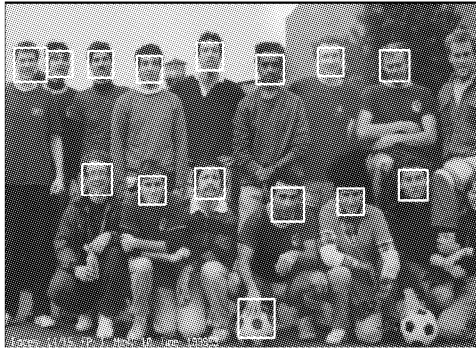


图2 使用100个弱分类器构造的人脸检测器的检测结果

表1 单阈值和双阈值分类器构造的强分类器分类能力对比

弱分类器个数	单阈值错图数目	双阈值错图数目
10	904	758
20	586	365
30	458	237
40	301	148
50	240	72
100	108	37

本文在人脸检测器训练中,将 Feret 人脸库^[6]作为人脸资料库,训练样本共有 4 520 幅人脸图像,6 305 幅非人脸图像。每幅图像为 24×24 像素,每幅图像提取 45 936 个矩形特征。本文使用快速算法提取 100 个单阈值特征,在 PM1.6 PC 机上,Matlab 7.0.1 环境一共耗时不到 72h,其训练时速度是改进前 Adaboost 人脸检测训练速度的 11 倍左右,且性能也有较大的改善。

在相同条件下,分别使用单阈值弱分类器训练和双阈值弱分类器训练。表 1 是算法在所有训练样本集上的收敛结果,其中的错误图像样本数是人脸和非人脸错误图像数目的总和。可以看到,采用双阈值弱分类器的强分类器在训练样本上收敛速度明显快于采用单阈值弱分类器的系统。譬如当错误样本数目达到 100 时,采用双阈值弱分类器的系统只需要不到 50 个弱分类器,而使用单阈值分类器需要 100 个以上

(上接第 606 页)

从表 1 可以看出:CEM 对大噪声的失真图像和解码图像的判断要优于 PSNR 的结果,对块噪声的失真图像的判断结果与 PSNR 的判断结果基本一致;而 FI 对前两类噪声图像的判断基本失效,它们的值基本上都停留在 0.035 和 0.0076 左右,对解码图像的判断较好,但是从相关系数上看不如 CEM 好。上述结果说明 CEM 的判断结果基本与主观判断相符合。

3 结语

本文使用图像的自相似性特征作为对失真图像的质量进行度量的依据,并结合人眼在观察图片时的一种方向效应,得到了一种新的、有效的图像度量方法。与峰值信噪比和模糊积分方法的三组对比实验表明,本文方法对失真图像质量的判断结果基本上是与人的视觉相符合的,且优于前两者。本文方法使用拼贴误差作为局部自相似性的表现形式,因为它最直观,而选用其余的编码参数如对比度伸缩因子或几个编码参数的结合来计算图像的局部自相似性可能会更好;此外,

的弱分类器。这样就大大简化了 Adaboost 人脸检测系统结构,训练速度提高了 11 倍,检测速度可以提高约 1.5 倍,总的训练速度比传统方法提高约 17 倍。

本文从因特网上随机抽取图像,使用 100 个弱分类器构成的强分类器对其检测,一张包含各种检测结果如图 2。

从图 2 中可以看到该检测器具有比较好的效果,图 2 右上角有一人脸的漏检情况,中下部出现一个将足球归类为人脸的情况。这是因为本文主要着重于快速的训练算法,仅仅提取了 100 个弱分类器,而实用时往往需要提取上千个弱分类器。

4 结语

本文从以下两个方面对 Adaboost 人脸检测系统进行改进。通过直接求解 Adaboost 弱分类器中目标函数,避免了传统方法中的概率统计和迭代训练过程,提高了单个弱分类器训练速度。同时,在单个弱分类器快速训练算法基础上,将单阈值判决扩展到双阈值判决,提高了每个弱分类器的分类能力,可以使用更少的弱分类器构造强分类器。这样较好地简化系统结构,减少了训练时间,提高了检测速度。

参考文献:

- [1] WECHSLER H, PHILLIPS PJ, BRUCE V, *et al.* Face Recognition from Theory to Applications[M]. Springer; NATO Scientific Affairs Division, 1998.
- [2] GONG S, MCKENNA SJ, PSARROU A. Dynamic Vision: From Images to Face Recognition[M]. Imperial College Press, 2000.
- [3] VIOLA P, JONES M. Robust Real Time Object Detection[A]. Proceedings of 2nd International Workshop on Statistical and Computational Theories of Vision[C], 2001.
- [4] FREUND Y, SCHAPIRE R. A Short Introduction to Boosting[J]. Journal of Japanese Society for Artificial Intelligence, 1999, 14(5): 771 - 780.
- [5] CORMEN TH, LEISERSON CE, RIVEST RL, *et al.* Introduction to Algorithms(2nd Edition)[M]. USA: MIT Press; 北京: 高等教育出版社, 2002. 145 - 164.
- [6] PHILLIPS P, WECHSLER H, JUANG J, *et al.* The feret database and evaluation procedure for face-recognition algorithms[J]. Image and Vision Computing Journal, 1998, 16(5): 295 - 306.

我们还可以引入多分辨率分析,结合图像在不同分辨率下的局部自相似性提出更合理的度量方法。

参考文献:

- [1] LI J, CHEN G, CHI Z. A Fuzzy Image Metric With Application to Fractal Coding[J]. IEEE Transactions on Image Processing, 2002, 11(6): 636 - 643.
- [2] DUNG LP, BALA S, SALAHADIN M, *et al.* A Measure for image quality[A]. Symposium on Applied Computing Proceedings of the 1998 ACM Symposium on Applied Computing[C]. Atlanta Georgia, United States, 1998. 513 - 519.
- [3] DALY S. The visible difference predictor: An algorithm for the assessment of image quality[A]. Digital Image and Human Vision [C]. MIT Press, 1993. 179 - 206.
- [4] 曹圣群,黄普明,鞠德航. HVS 模型及其在静止图像压缩质量评价中的应用[J]. 中国图像图形学报, 2003, 8A(4).
- [5] FISHER Y. Fractal Image Compression - Theory and Application [M]. New York: Springer, 1994.