

基于沃尔什特征的增强型 AdaBoost 人脸快速检测算法

郭志波^{1, 2}, 杨静宇¹, 刘华军¹, 严云洋¹

(1. 南京理工大学计算机科学与技术学院, 南京 210094; 2. 扬州大学信息工程学院, 江苏扬州 225009)

[摘要] 提出一种基于沃尔什特征的增强型 AdaBoost 人脸快速检测算法, 不仅具有很快的训练速度, 而且利用较少的非人脸样本进行训练就可以达到较好的检测效果。首先, 提出用较少的沃尔什特征来代替大量的 Haar-Like 特征可以较大幅度的降低特征之间的冗余。然后提出一种双阈值增强型 AdaBoost 算法, 其中双阈值的快速搜索方法大大节约了训练时间, 并且在训练 Cascaded 检测器过程中, 前层分类器的训练结果对后层分类器的训练具有指导作用, 加强了总体检测器的性能, 另外通过各层分类器阈值的调节, 能够将人脸和非人脸的训练结果尽量分离。最后, 使用该算法训练的检测器对 MIT+CMU 人脸测试库进行了测试, 结果表明该方法在训练速度、测试精度、检测时间等方面都优于相应的方法。

[关键词] 沃尔什特征; 增强型 AdaBoost; Cascaded 型检测器; 人脸检测

[中图分类号] TP391.4 **[文献标识码]** A **[文章编号]** 1009-1742(2008)07-0029-07

1 前言

人脸检测是模式识别和机器视觉领域的一个研究热点, 由于具备很高的实际应用和理论研究价值, 使得人脸检测成为一个非常具有吸引力的课题。人脸检测问题一直备受关注在很大程度上源于其自身的难度, 主要包括: 脸型、肤色、表情、姿态的多样性、人脸局部遮挡、光照影响、人脸图像的成像质量等。人脸检测的现有方法很多, 有模板匹配方法^[1]、镶嵌图方法^[2]、几何结构方法^[3]、差影方法^[4]、活动轮廓模型^[5]、可变形模板^[6]、肤色特征方法^[7, 8]等, 当前统计学习的方法在人脸检测中应用非常广泛, 如主分量分析^[9]、人工神经网络^[10, 11]、支持向量机^[12]、贝叶斯特征的方法^[13]等。统计学习的方法虽然在一定程度上提高了人脸检测的精度, 但是并不能达到很快的检测速度。

P. Viola 等提出了一种基于 Haar-Like 型特征的 AdaBoost 人脸检测算法^[14], 并综合 Cascade 算法实现了实时的人脸检测, 该方法是人脸检测速度提

高的一个转折点, 使得人脸检测从真正意义上走向实用。但是该方法也存在一些缺点, 比如说一张人脸图像 Haar-Like 型特征数目非常巨大, 利用过多的特征进行训练消耗了大量训练时间。另外使用 AdaBoost 算法进行训练时, 对每个弱分类器阈值的搜索需要消耗很多时间, 最为突出的一点是该方法需要非常多的非人脸样本进行训练才能达到较好的人脸检测效果, 而庞大的训练样本也大大限制训练速度, 并且得到的 Cascaded 检测器中的弱分类器的数目也变得相对较多。

针对这些问题, 笔者提出一种基于沃尔什特征的增强型 AdaBoost 人脸快速检测方法。首先使用沃尔什特征代替 Haar-Like 特征可以较大幅度的降低特征之间的冗余, 也就是说可以使用较少的沃尔什特征来代替大量 Haar-Like 特征。然后提出一种双阈值增强型 AdaBoost 算法, 其中双阈值的快速搜索方法大大节约了训练时间, 并且在训练 Cascaded 检测器过程中, 由于前层分类器的训练结果对后层分类器的训练具有指导作用, 因此加强了总体检测器

[收稿日期] 2006-07-24; **[修回日期]** 2006-10-26

[基金项目] 国家自然科学基金资助项目(60472060, 60473039); 江苏省高技术研究基金资助项目(BG2005008)

[作者简介] 郭志波(1975-), 男, 江苏江都市人, 南京理工大学博士生, 研究方向为模式识别、图像分析、人脸检测、人脸识别等

的性能。另外由于该方法使人脸和非人脸训练结果尽量分离,所以使用较少的非人脸样本进行训练就可以取得较好的实际测试结果。在实验当中,定义了3 610个沃尔什特征,使用6 058个人脸样本和9 083个非人脸样本进行训练,消耗较少的训练时间,就可以得出最终的 Cascaded 检测器,总共包含16层强分类器(530个弱分类器)。最后在 MIT+CMU 人脸测试集上检验了检测器的性能,该方法在训练速度、测试精度、检测时间等方面都优于 P. Viola 等的 AdaBoost 人脸检测方法。

2 基于 Haar-Like 特征的 AdaBoost 人脸检测

2.1 Harr-Like 特征与积分图

Harr-Like 特征是 P. Viola 等提出的一种简单矩形特征,因类似于 Haar 小波而得名,如图 1 所示。

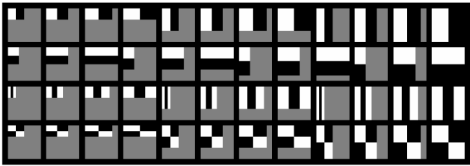


图 1 Harr-Like 部分特征
Fig.1 Some Harr-Like features

Harr-Like 特征定义为黑色矩形和白色矩形在图像子窗口中对应的区域的灰度级总和之差,可见,它反映了图像局部的灰度变化。P. Viola 等用到的 Harr-Like 特征共有 3 类,图 1 中显示了部分在图像子窗口起点位置处的 Harr-Like 特征,其中第一、第二行属于二矩形特征,第三行属于三矩形特征,第四行属于四矩形特征。在实际使用时,必须将每一特征在图像子窗口中进行滑动计算,从而获得各个位置的多个 Harr-Like 特征,如果选用的训练图像分辨率为 24×24 ,每个图像得到用于训练的 Harr-Like 特征超过 18×10^4 个。

在实际检测过程中,为了加快 Harr-Like 特征的计算,P. Viola 等提出积分图像的定义,灰度图像 I 的积分图像 \hat{I} 定义为

$$\hat{I}(x, y) = \sum_{i,j=1}^{x,y} I(i, j) \quad (1)$$

使用积分图像,只是使用少量的加减法运算就可以计算出 Harr-Like 特征,如图 2 所示。右边是积分图,要计算左边图像中某个位置的某种特征,只需要采用积分图中的对应位置 $P_1 \sim P_6$ 点的值进行加减法运算,右图特征数值 $d = P_4 + P_1 - P_2 - P_3 - (P_5$

$+ P_6 - P_4 - P_5)$ 。这样可以大大节省计算时间。

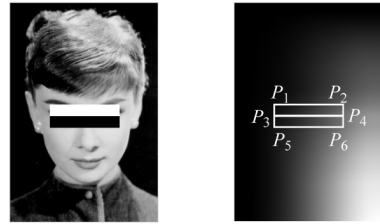


图 2 使用积分图进行特征提取
Fig.2 Features extraction using integral image

2.2 AdaBoost 算法

通过训练寻找多个弱分类器 h_i ,然后将多个弱分类器组合成一个强分类器 f ,即

$$f(x) = \sum_{i=1}^T \alpha_i h_i(x) \quad (2)$$

具体算法如下:

1) 对于训练集合 $(x_1, y_1), \dots, (x_i, y_i)$, $g_j(x_i)$ 代表第 i 个训练图像的第 j 个 Harr-Like 特征, $y_i \in \{1, 0\}$ 分别表示真假样本。

2) 初始化权重 $w_{1,i} = 1/2m, 1/2n$, 其中 m, n 分别是真样本、假样本的数目。

3) 寻找 T 个弱分类器 $h_t (t=1, 2, \dots, T)$ 。

a. 对所有样本的权重进行归一化:

$$w_{t,i} = w_{t,i} / \sum_{i=1}^l w_{t,i} \quad (3)$$

b. 对于每个样本中第 j 个特征,可以得到一个弱分类器 h_j ,也就是确定阈值 θ 和偏置 p_j ,使得误差

$\epsilon_j = \sum_{i=1}^l w_{t,i} |h_j(x_i) - y_i|$ 达到最小,而

$$h_j(x) = \begin{cases} 1 & p_j g_j(x) < p_j \theta \\ 0 & \text{其他} \end{cases} \quad (4)$$

偏置 p_j 决定不等式方向,只有 ± 1 两种情况。

c. 在确定的弱分类器中,找出一个具有最小的误差 ϵ_t 的弱分类器 h_t 。

d. 对所有样本的权重进行更新: $w_{t+1,i} = w_{t,i} \beta^{-\epsilon_i}$, 其中 $\beta = \epsilon_t / (1 - \epsilon_t)$, 如果 x_i 被 h_t 正确分类,则 $\epsilon_i = 0$,反之 $\epsilon_i = 1$ 。

4) 最后得到的强分类器为

$$H(x) = \begin{cases} 1 & \sum_{i=1}^T \alpha_i h_i(x) \geq 0.5 \sum_{i=1}^T \alpha_i \\ 0 & \text{其他} \end{cases} \quad (5)$$

其中 $\alpha_i = \ln(1/\beta)$ 。

2.3 Cascaded 检测器

对于一幅图像,所有子窗口中,非人脸的数目比人脸的数目要大很多,采用逐级抛弃非人脸的 Cascaded 检测器可以大量的节约检测时间。如图 3 所示,将所有的图像子窗口通过各层分类器,在每层分类过程中,将不符合条件的子窗口认为是非人脸图像并抛弃,这样越到后面,需要判断的图像子窗口数目越少,而位置越靠后的层分类器越复杂,即包含越多的弱分类器,因而也具有更强的分类能力。这样做是因为非人脸样本通过的层数越多就越像人脸,因而越接近分类边界,各层分类器通过 AdaBoost 算法训练得到。具体的 Cascaded 算法如下:

1) 首先设定每层最大误报率 f (被检测成真样本的假样本的百分率)、每层最小检测率(对真样本的检测通过率),设定最终的误报率 F_{target} 。

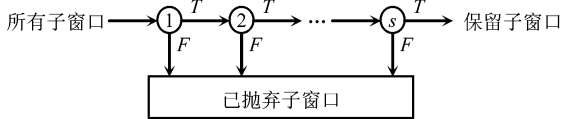


图 3 Cascaded 检测器的结构

Fig.3 The structure of cascaded detector

2) P, N 分别表示真样本(人脸)集合和假样本(非人脸)集合, n_i 表示第 i 层分类器中弱分类器的数目, F_i, D_i 分别表示前 i 层的总体误报率和总体检测率。

3) 初始化 $F_0 = 1.0, D_0 = 1.0, i = 0$ 。

4) 如果 $F_i > F_{\text{target}}$, 进行如下循环:

a. $i++$, $n = 0, F_i = F_{i-1}$;

b. 如果 $F_i > fF_{i-1}$, 进行如下循环

n_i++ ;

使用 AdaBoost 训练出包含 n_i 个弱分类器的第 i 层强分类器;

降低阈值使得目前 Cascaded 分类器对真样本的检测通过率 D_i 不小于 $d \times D_{i-1}$, 也就是说当前层分类器的检测通过率 $d_i \geq d$, 并计算 F_i ;

c. 进行假样本集合的调整, 在 N 集合中, 只保留错误检测的假样本。

3 沃尔什特征

3.1 沃尔什特征的构造

离散沃尔什变换是在数字信号处理方面经常采用的一种正交变换, 并且该变换同离散傅立叶变换

一样, 可以很好地反映数字信号的频率变化情况, 并且只由 ± 1 组成, 可以较大程度减少计算量。如何构造沃尔什特征, 下面给出一个简单的算法:

1) 设沃尔什特征的高度、宽度分别为 m, n (必须是 2 的整数次方)。

2) 分别构造 $m \times m, n \times n$ 维的沃尔什变换矩阵(沃尔什排列) $H_{w, m}, H_{w, n}$ 。

3) 得到的沃尔什特征 $W_{i, j} = H_{w, m}^T(i), H_{w, n}^T(j)$, 其中 $H_{w, k}(t)$ 表示 $H_{w, k}$ 第 t 行向量。对于沃尔什变换矩阵(沃尔什排列) $H_{w, k} = (h_{i, j})$ 的构造方法, 可以按照

$$h_{i, j} = wal_w(i, j) = \prod_{t=0}^{k-1} [R(t+1, j)]_i^{g(i)} \quad (6)$$

来构造, 其中 $R(t+1, j)$ 是拉德梅克函数, $g(i)_i$ 是 i 的格雷码的第 t 位数字。图 4 是 $k=2, 4, 8, 16$ 的沃尔什变换矩阵的图像, 图 5 是所有 4×8 维的沃尔什特征图像, 黑代表 +1, 白色代表 -1。



图 4 沃尔什变换矩阵 ($k=2, 4, 8, 16$)

Fig.4 Walsh transform matrices ($k=2, 4, 8, 16$)

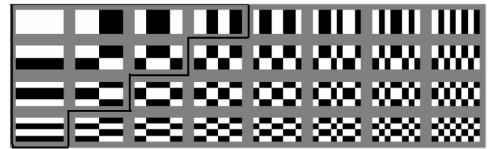


图 5 一组沃尔什特征 (4×8)

Fig.5 A group of Walsh features (4×8)

为了在检测过程能够使用积分图计算方法进行加速, 所以一般选用左上角的中低频特征。实验当中, 采用图 5 黑线框以内的 10 个形态(黑白位置分布), 其中有几种特征形态是和 Harr-Like 特征相似或一样。

3.2 沃尔什特征的优越性

由于所有 $m \times n$ 维的沃尔什特征都是相互正交, 这意味着它对图像某个局部特征的提取不产生冗余, 并且特征比 Harr-Like 特征丰富。另外, m, n 都要求是 2 的整数次方, 所以对一幅训练图像而言, 它的全部沃尔什特征数是有限的, 而实际 Harr-Like 特征提取过程中, 很多特征都非常相近, 冗余很大, 采用沃尔什特征来代替 Harr-Like 特征可以大

大降低特征之间的冗余,并取得较好的检测结果。由于沃尔什特征少于 Harr-Like 特征,所以在训练中加快了训练速度。

4 增强型 AdaBoost 算法

在 AdaBoost 算法中,第 2 步非常耗时间,要求出所有 $h_j(j=1:k)$,使用 Hare-Like 特征, k 非常大,而每个 h_j 也要通过搜索的方式找到,由于样本数也非常多,所以单纯一个 h_j 的搜寻就比较费时间,那求出所有 h_j 就更加费时间,而这只是得到一个 h_i 所要消耗的搜索时间,如果需要 500 个弱分类器,每个 h_j 的平均搜索时间是 0.1 s,有 10^5 种特征,大概需要消耗的时间为 $500 \times 10^5 \times 0.1 = 5 \times 10^6$ s = 1 389 h。要减少训练时间,必须在保证测试精度的基础上,降低消耗在 h_j 上的时间,减少冗余特征,使用较少样本。

4.1 增强型 AdaBoost 原理

在实际训练过程中,找出的弱分类器对应的局部特征主要是两种分布形状,图 6 给出这两种典型非人脸和人脸的局部特征在各个数值上的百分比图(简称分布图)。

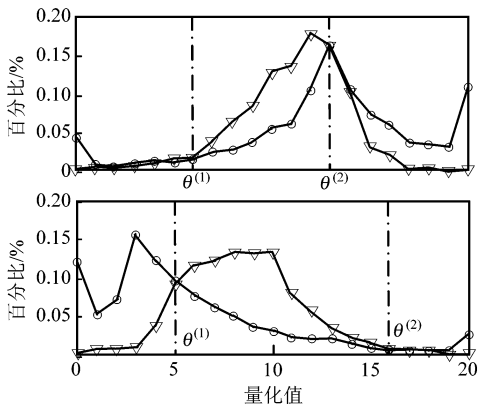


图 6 两个典型局部特征的分布图

Fig. 6 The distribution curve of typical local features

图 6 纵向轴表示所占总数的百分比,横轴是量化后的数值。通过分布图可以快速的求出每种特征的阈值,也就是分布图上,人脸(“○”线)高于非人脸(“△”线)的区域,如图 6 中的 $\theta^{(1)}$, $\theta^{(2)}$,具体方法是先找出分布图上真样本减去假样本分布曲线的最高点,然后向两边在该曲线上寻找 2 个“+,-”过渡点,如没有就取边界点,这种方法虽然是近似方法,但是与穷搜索相比较,性能几乎是一样,而它大大提高了阈值寻找的速度。并且采用双阈值代替单阈

值,还使得对阈值的寻找更加精确,有助于加快训练过程。当确定哪种特征作为当前的弱分类器的特征时,针对该特征进行双阈值的调整,以求弱分类器 h_i 满足要求。上述方法可以大大加快每个弱分类器的寻找速度,如果有 10 000 种特征,使用该方法消耗的时间不超过穷搜索方法的 1/500。

由于非人脸图像中像素分布几乎没有规律,即使使用上百万幅非人脸图像也不能完全代表世界上所有非人脸图像。当然,使用的非人脸样本越多,误报率就越低,但同时使得训练过程中内存和时间的消耗也越多。在 Cascaded 型检测器当中,越往后的层分类器需要处理的图像子窗口越少,很小的误差也会造成总体检测效果的下降,所以要求越往后的层分类器的分类性能越强。这里给出一种方法,在每层分类器训练过程中,都利用了前面各层分类器的训练结果,也就是越往后的层分类器包含越多的整体信息,从而加强了整体 Cascaded 检测器的性能。另外通过各层分类器阈值的调节,能够将人脸和非人脸的训练结果尽量分离。在实际测试过程中可以根据情况选择合适的阈值作为各层分类器的阈值,这样给非人脸图像留下的较多的泛化空间,可以去除更多的非人脸图像。

4.2 增强型 AdaBoost 具体算法

现在要训练第 k 层强分类器, $f_k(x)$, $H_k(x)$ 表示第 k 层强分类器函数和强分类器,并设保留样本集合前 m 个是真样本(人脸),后 n 个是假样本(非人脸),样本总数为 l 。

1) 对于训练集合 $(x_1, y_1), \dots, (x_l, y_l)$, $g_j(x_i)$ 代表第 i 个训练图像的第 j 个沃尔什特征, $y_i \in \{1, -1\}$ 分别表示真假样本。

2) 保留样本的权重 $w_{1,i}$,采用的是训练第 $k-1$ 层强分类器对应样本最后的权重,如果 $k=1$,则需要初始化权重,真假样本的权重为

$$w_{1,i} = \begin{cases} 0.5/m & i \leq m, \\ 0.5/n & \text{其他} \end{cases} \quad (7)$$

3) 通过真假样本的每种特征分布图,确定每种特征的双阈值 $q^{(1)}$, $q^{(2)}$ 。

4) 寻找 T 个弱分类器 $h_t(t=1, 2, \dots, T)$

a. 对真假样本的权重进行归一化:

$$w_{t,i} = \begin{cases} w_{t,i} / 2 \sum_{i=1}^m w_{t,i} & i \leq m, \\ w_{t,i} / 2 \sum_{i=m+1}^l w_{t,i} & \text{其他} \end{cases} \quad (8)$$

b. 对于每个样本中第 j 个特征, 根据双阈值 $\vartheta^{(1)}$, $\vartheta^{(2)}$ 得到一个弱分类器 h_j , 求出误差 $\varepsilon = \sum_{i=1}^l w_{i,j} h_j(x_i) \text{sgn}(y_i)$, 在确定的弱分类器中, 找出一个具有最小误差 ε 的弱分类器 h_i , 计算系数 $\alpha^{(0)} = (\ln((1 - \varepsilon)/\varepsilon))/2$ 。

c. 此时得到的强分类器函数为

$$f_k(x) = f_{k-1}(x) + \sum_{i=1}^{l-1} \alpha_i h_i(x) + \alpha^{(0)} h_i(x) \quad (9)$$

此时的强分类器为

$$H_k(x) = \begin{cases} 1 & f_k(x) \geq 0, \\ -1 & \text{其他} \end{cases} \quad (10)$$

d. 不断增大 $\Delta\theta$ ($\Delta\theta > 0$), 并调整阈值 $\vartheta^{(1)} \leftarrow \vartheta^{(1)} - \Delta\theta$, $\vartheta^{(2)} \leftarrow \vartheta^{(2)} - \Delta\theta$, 在新阈值下重新计算 ε , $\alpha = \mu\alpha^{(0)} + (\ln((1 - \varepsilon)/\varepsilon))/2$, 其中 μ 为抑振因子, 合理调节可以加快训练进程, 一般取 $\mu = (0.5 \sim 1)$, 最终使得 $f_k(x)$ 对真样本的测试达到设定的要求 $D_k \geq d \times D_{k-1}$ 。

e. 使用获得的 $f_k(x)$ 对假样本进行测试, 如达到要求 $F_i \leq f \times F_{i-1}$, 结束循环。

f. 对所有样本的权重进行更新: $w_{i+1,i} = w_{i,i} e^{-\alpha h_i}$, 其中 $\alpha = (\ln((1 - \varepsilon)/\varepsilon))/2$ 。

5) 最后得到的强分类器函数为

$$f_k(x) = f_{k-1}(x) + \sum_{i=1}^r \alpha_i h_i(x) \quad (11)$$

也就得到该级的强分类器

$$H_k(x) = \begin{cases} 1 & f_k(x) \geq \lambda\beta, \\ -1 & \text{其他} \end{cases} \quad (12)$$

其中 $\beta = \min_{i=1,m} (f_k(x_i))$, λ 是阈值调整系数, 增大 λ 会降低误报率和检测率, 降低 λ 会增大误报率和检测率。

在每层强分类器训练之前, 需要进行假样本集合的调整, 使用该层的强分类器对假样本进行测试, 将结果为 -1 的样本保留。对于训练过程中, 各层分类器的阈值调整系数 λ 取不一样的数值, 目的有 2 个: 一方面使得前几层分类器包含弱分类器的数目呈递增趋势, 因为越在前面的分类器要处理图像子窗口数就越多, 为了提高速度, 所以在前几层分类器中, 弱分类器的数目需要设计成图 7(a) 的分布形态; 另外一方面, 要使得各层分类器函数对真假样本的测试值尽量间隔大一点。也就是, 使得 $\Delta\beta = \min_{i=1,m} (f_k(x_i)) - \max_{i=m+1,l} (f_k(x_i))$ 越往后越大, 如图 7

(b) 所示, “*” 线表示 $\Delta\beta$, “○” 线表示 $\max_{i=m+1,l} (f_k(x_i))$, “▽” 线表示 $\min_{i=1,m} (f_k(x_i))$ 。所以 λ 在训练过程中, 一般采用递减的方式, 从 $\lambda=1$ 开始, 逐层递减到 $\lambda=0$, 然后保持不变。

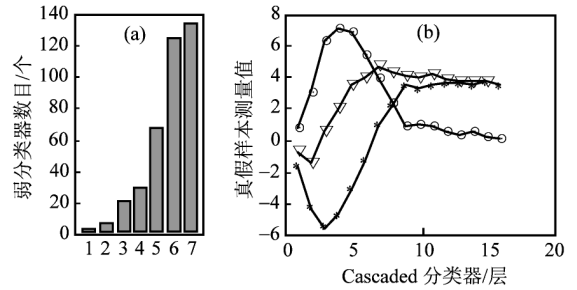


图 7 层分类器的两个特性

Fig.7 Two characters of layer classifier

5 样本选择及预处理

5.1 人脸训练样本的选择

人脸样本选择能够具有一定的代表性的人脸, 也就是说, 人脸图像训练集, 应该考虑到肤色、遮挡、姿态、光照等情况, 另外作为训练的人脸图像尽量去除人脸上多变性的地方(比如头发)。这里选用的人脸训练集, 一个是 MIT-CBCL 人脸训练库, 一个是 FERET 人脸库。MIT-CBCL 人脸训练库是专门用于人脸检测的一个具有较强代表性的人脸训练库, 总共有 2 429 个分辨率为 19×19 人脸样本, 包含各种肤色、遮挡、姿态、光照等情况, 实验中将分辨率缩小到 16×16 , 并根据每个样本平均值, 作了高斯分布的亮度调整, 调整后的样本均值呈高斯分布, 范围在 $25 \sim 235$, 最高点到最低点的样本数目比例为 $5:1$ 。图 8(a) 给出部分样本示范。FERET 人脸库主要用于人脸识别, 包括 14 051 个 256 色灰度人脸图像, 包含从正面到侧面多个头部姿态。这里选用了编号为 bd, be, bf, bg 的人脸图像, bd, be, bf, bg 分别表示照相机在头部正前方的左面 25° 、左面 15° 、右面 15° 、右面 25° 拍到的人脸图像, 总共 800 幅, 然后按照 MIT-CBCL 的人脸进行剪切、缩小到 16×16 , 从中选择了 600 幅进行样本均值高斯分布调整, 范围在 $25 \sim 235$, 示范图如图 8(b) 所示。FERET 人脸的选择主要是补充 MIT-CBCL 人脸训练库中这一类姿态不足的缺点, 这样总共使用的人脸样本总数为 3 029 个, 加上其镜像总共为 6 058 个。

5.2 非人脸样本的选择

非人脸样本的选择较随意, 只要其与人脸的相

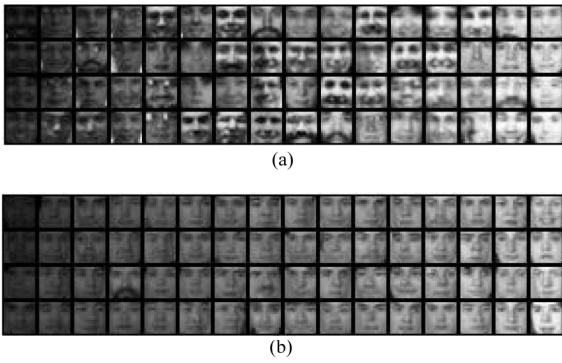


图8 人脸训练集中的部分样本
Fig.8 Some samples in face training set

似度的范围宽,并且在每一段范围内都有较多的非人脸样本。这里非人脸的获取有2个来源,一个是MIT-CBCL非人脸训练库,总共有4 548个分辨率为 19×19 非人脸样本,但对于使用AdaBoost算法来说,显得少一些。所以,又从40幅自然图像中,选取了4 535幅非人脸图像,其中包含几百个和人脸非常相似的非人脸样本,非人脸样本同样要缩放到 16×16 ,部分样本如图9所示,其中图9(b)显示的是和人脸较相似的部分非人脸样本。

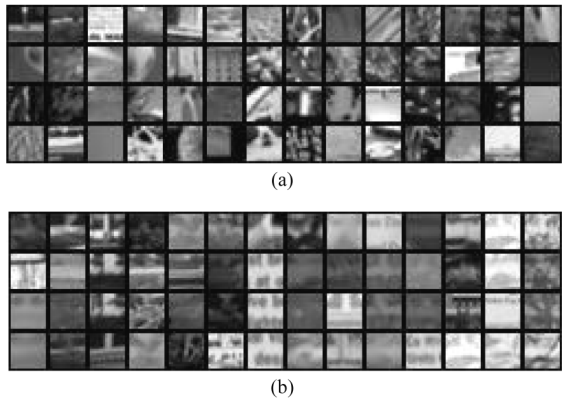


图9 非人脸训练集中的部分样本
Fig.9 Some samples in nonface training set

6 实验结果

实验中,总共选用3 610个沃尔什特征,其中最小特征滤波器在 16×16 的图像窗口中所占的区域最小为 4×4 ,最大为 16×16 ,每一种尺寸的特征形态,为前面提过的10种。人脸样本6 058个,非人脸样本9 083个,计算机P4-2.8 GB, 1 GB内存。测试集MIT+CMU人脸检测库,该库包含130幅不同背景、光照、尺寸的256色灰度图像,其中包含了507

个不同种族、姿态、光照、遮挡等情况的人脸。

设定每一层非人脸样本抛弃率大于35%,使用笔者提出的算法,花费训练时间25 min(Matlab 7),得到包含16层分类器的Cascaded检测器(530个弱分类器),每层分类器的所含弱分类器数目和阈值分布图,如图10所示,图10(c)是2种对应各层分类器的阈值,“ Δ ”线是训练用的阈值,“ \square ”线是一个测试用的阈值,也就是保证设定人脸样本(这里前两层99.5%,后面都是100%)通过的阈值。图11(a)是使用训练阈值检测的结果,误报较多,而使用测试阈值不仅可以去除误报,而且保证了检测结果,如图11(b)所示。也就是说采用较少的非人脸样本进行训练也能得到较好的结果。

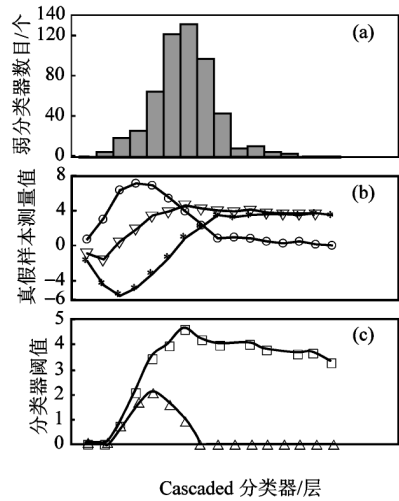


图10 分类器的几个性质
Fig.10 Several characters of classifier

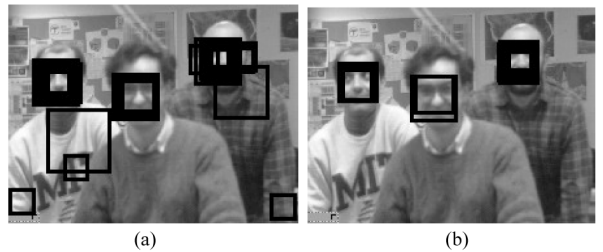


图11 不同阈值下的检测结果
Fig.11 The detection results using different thresholds

在进行测试过程中,图像缩放系数0.85,采用了15级缩放,为了加快检测速度,采用的是交叉点采样。本文方法对一幅 320×240 的图像进行检测消耗的时间大约在50 ms(VC++6)。选择多种层阈值调整系数,对MIT+CMU的测试结果如表1所

示,可以看到笔者的方法总体上优于 P. Viola 等的方法。部分检测结果如图 12 所示,图中黑色框线包含的区域为检测结果,为了保留原始检测结果,没有进行框线合并,其中有一个漏检和一个误报,漏检原因是脸部偏转太大造成,而误报是因为孩子窗口很像一个人脸。

表 1 对 MIT+CMU 测试集的测试结果对比

Table 1 The comparison of results on the MIT+CMU test set

	P. Viola 等			笔者方法		
误报数	50	78	167	53	84	146
检测率/%	91.4	92.0	93.9	91.7	92.5	94.3

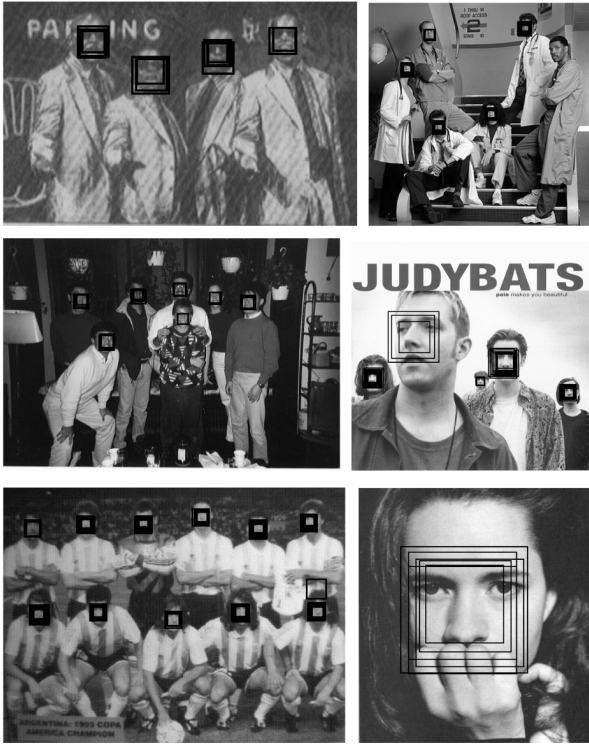


图 12 部分检测结果

Fig.12 Some detection results on MIT+CMU test set

7 结语

笔者提出了一种基于沃尔什特征的增强型 AdaBoost 人脸快速检测方法,定义了一种新的特征(沃尔什特征)并提出了一种基于双阈值快速搜索的增强型 AdaBoost 算法,使用了较少的非人脸样本,在较少的特征条件下进行训练就可以取得较好的实际测试结果,在 MIT+CMU 人脸测试集上检验了使用该算法获得的人脸检测器的性能,该方法在训练速

度、测试精度、检测时间等方面都优于 P. Viola 等的 AdaBoost 人脸检测方法。以后的工作将应用该算法对多视角人脸进行检测并采用该方法对其他一些目标进行检测。

参考文献

- [1] Brunelli R, Poggio T. Face recognition: Features versus templates [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993, 15 (10): 1042-1052
- [2] Yang G Z, Huang T S. Human face detection in a complex background [J]. Pattern Recognition, 1994, 27 (1): 53-63
- [3] Jeng S H, Liao H Y. An efficient approach for facial feature detection using geometrical face model [A]. Proc of 13th International Conference on Pattern Recognition [C]. Vienna, Austria: IEEE Computer Society Press, 1996: 426-430
- [4] Jain R. Difference and accumulative difference pictures in dynamic scene analysis [J]. Image and Vision Computing, 1984, 2(2): 99-108
- [5] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models [A]. Proc of the 1st International Conference on Computer Vision [C]. London, UK: IEEE Computer Society Press, 1987: 259-268
- [6] Yuille A, Hallinan P, Cohen D. Feature extraction from faces using deformable templates [A]. Proc of IEEE Computer Society Conference on CVPR '89 [C]. San Diego, California, USA, IEEE Computer Society Press, 1989: 104-109
- [7] Jones M J, Rehg J M. Statistical color models with application to skin detection [R]. Technical report, Cambridge Res Lab, Compaq Computer Corp, 1998
- [8] Martinkauppi B. Face colour under varying illumination—analysis and applications [D]. PhD Thesis, University of Oulu, 2002
- [9] Moghaddam B, Pentlan A. Beyond linear eigenspaces: Bayesian matching for face recognition [A]. Face Recognition: from Theory to Application [M]. Springer-Verlag, 1998: 230-243
- [10] Rowley H A. Neural network-based human face detection [D]. PhD Thesis, Carnegie Mellon University, USA, May 1999
- [11] Feraud R, Bernier O J, Viallet J E, et al. A fast and accurate face detector based on neural networks [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001, 23(1): 42-53
- [12] Osuna E, Freund R, Girosi F. Training support vector machines: an application to face detection [A]. IEEE Conference on CVPR '97 [C]. San Juan, Puerto Rico, USA, IEEE Computer Society Press, 1997: 130-136
- [13] Liu C J. A bayesian discriminating features method for face detection [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(6): 725-740
- [14] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [A]. IEEE Conference on CVPR '2001 [C]. Lihue, Kauai, Hawaii, USA, IEEE Computer Society Press, 2001: 511-518

(下转 136 页)