

# 基于多通道图判别投影 HAAR 特征的多视角人脸检测

沈继锋<sup>1</sup> 时士伟<sup>1</sup> 左欣<sup>2</sup> 徐丹<sup>2</sup>

(1. 江苏大学电气信息工程学院, 镇江, 212013; 2. 江苏科技大学计算机科学与工程学院, 镇江, 212013)

**摘要:** 提出了一种基于多通道图判别投影 HAAR 特征的多视角人脸检测算法。该算法首先将人脸图像提取多通道图, 降低图像中的光照和噪声影响; 其次基于正负训练样本集利用线性判别投影学习增强 HAAR 特征, 提高特征判别能力; 然后计算训练样本的增强 HAAR 特征在多通道图中的响应, 并利用非对称 GentleBoost 算法进行特征选择生成一组弱分类器; 最后利用线性非对称分类器重新调整强分类器的权重和阈值。该方法不仅提高了特征的判别能力, 而且实现了非平衡正负样本空间的合理划分。实验结果表明: 该方法与当前经典方法相比具有更快的检测速度和更高的检测精度。

**关键词:** 多视角人脸检测; 多通道图; Fisher 判别投影; 非对称线性分类器

**中图分类号:** TP391      **文献标志码:** A

## Multi View Face Detection Based on Multi-channel Discriminative Projection HAAR Features

Shen Jifeng<sup>1</sup>, Shi Shiwei<sup>1</sup>, Zuo Xin<sup>2</sup>, Xu Dan<sup>2</sup>

(1. School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, 212013, China; 2. School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang, 212013, China)

**Abstract:** A multi view face detection algorithm based on multi-channel map discriminant projection HAAR feature is proposed. Firstly, the multi-channel map is extracted from the face image by the algorithm, which can reduce the influence of illumination and noise in the image. Secondly, based on the positive and negative training samples, the enhanced HAAR feature is learned by the linear discriminant projection, which can improve the distinguishing ability of the feature. Then the response in multi-channel of the augmented HAAR feature in the training sample is calculated, and the non symmetric GentleBoost algorithm is used to generate a set of weak classifiers. Finally, the weight and threshold of the strong classifier are adjusted by the linear non symmetric classifier. This method not only improves the distinguishing ability of the feature, but also realizes the reasonable division of the non balanced positive and negative sample space. Experimental results show that the proposed method has a faster detection speed and the higher detection accuracy compared with the classical methods.

**Key words:** multi view face detection; multi-channel map; Fisher projection feature; asymmetric linear classifier

## 引 言

作为人脸信息处理中的一项关键技术,人脸检测的研究近年来已成为模式识别与计算机视觉领域内一项备受关注和重视的课题。

人脸检测源于人脸识别,且是自动人脸识别系统中的一个关键环节。作为近几年最有潜力的生物身份验证手段,人脸自动识别系统要求对一般的图像有着更强的适应性,这也催生了人脸检测作为一项独立的课题被研究。如今,该课题的应用背景已远远超出了人脸识别系统的范畴,在基于内容的检索、数字视频处理、视频检测和人机交互等方面都有着广泛的应用前景<sup>[1-7]</sup>。虽然如此,由于人脸的非刚性结构,表情、背景、姿态、光照及遮挡等都会造成巨大的表象变化,影响人脸检测的效果。所以如何选择有着极强判别力的特征和能够快速构造分类器的算法是当前人脸检测技术研究的重点。

人脸检测的方法主要分为 4 大类<sup>[1]</sup>:基于知识的方法,基于特征的方法,基于模板的方法和基于表观的方法。其中基于表观的方法利用了统计和机器学习技术,效果较好<sup>[2]</sup>。基于表观方法的早期研究主要集中于单分类器方面的研究,包括神经网络<sup>[3]</sup>,支持向量机<sup>[4]</sup>,稀疏网络<sup>[8]</sup>,贝叶斯分类器<sup>[9]</sup>等。Viola 和 Jones 于 2001 年提出了基于 Adaboost 的集成分类器人脸检测算法<sup>[10]</sup>,此后该算法在计算机视觉领域受到了极大地关注,并成为了人脸检测的主流算法。该算法也是目前最快的人脸检测方法,它首先通过积分图快速计算特征,然后采用 Adaboost 学习算法从庞大的过完备特征集中选择具有强判别能力的特征子集,最后使用级联技术,使得分类器能在早期就将 90% 以上的非人脸区域快速排除。这种基于 Boosting 算法的检测方法第一次实现了实时的人脸检测,成为了具有里程碑意义的研究。在此基础上后续的改进方法主要存在于以下几个方面:基于训练样本的选择<sup>[11]</sup>,基于 Haar-like 特征(以下简称 HF)的表示<sup>[12]</sup>,基于 Boosting 算法的改进<sup>[13]</sup>及基于集成学习框架的优化<sup>[14-15]</sup>。其中,45° 旋转 HF<sup>[12]</sup>,分离 HF<sup>[16]</sup>,疏粒度特征<sup>[17]</sup>,局部二值模式(Local binary pattern, LBP)特征<sup>[18]</sup>等特征表示方法都是对 HF<sup>[19]</sup>进行了改进,然而这些方法都未体现正负样本的分布信息,因此判别能力相对较弱;改进的 Boosting 方法主要有 RealAdaboost<sup>[10]</sup>,LogitBoost<sup>[10]</sup>和 GentleBoost<sup>[10]</sup>等,这些方法利用连续置信阈值代替离散阈值,提高了 Boosting 算法的分类能力,但都存在忽略正负样本的分布非平衡性等问题。最近有学者提出了积分通道特征(Integral channel feature, ICF)<sup>[20]</sup>用于检测行人,该特征通过提取不同的颜色通道和形状通道信息来描述行人,提高了检测精度。在 ICF 特征的基础上,该学者又提出了聚合通道特征(Aggregate channel features, ACF)<sup>[21]</sup>,虽然和 ICF 使用相同的通道特征,但是检测器的候选特征提取方法有所不同,ACF 通过缩小多通道图的尺寸提高了对噪声的鲁棒性。ACF 最近又被用于多视角人脸检测<sup>[22]</sup>,取得了较好的结果。

本文在此基础上,提出一种基于多通道图增强 HAAR 特征的多视角人脸检测算法。该算法首先在多通道图中利用正负样本学习判别投影的增强 HAAR 特征,提高了特征的判别能力。然后将该特征在正负样本的响应输入非对称线性分类器,利用正样本分布近似高斯分布的特性和正负样本分布的不平衡性,得到最优的非对称线性分类器。实验结果表明该方法有较好的检测效果和较快的检测速度。

## 1 相关特征

### 1.1 多通道图

假设输入图像为  $I$ ,通道函数为  $\Omega$ ,则通道图可以表示为  $C_i = \Omega_i(I)$ ,  $i=1, 2, \dots, n$ 。其中  $C_i$  为图像  $I$  的第  $i$  个通道,  $\Omega_i$  为通道图  $i$  的生成函数。通道函数可以是线性函数,如图像  $I$  的灰度图;也可以是非

线性函数,如图像  $I$  的梯度图。常见的通道图,包括灰度图和各种颜色空间图、边缘图、梯度直方图和 LBP 图等。图 1 显示了常见的 11 种通道特征图像。

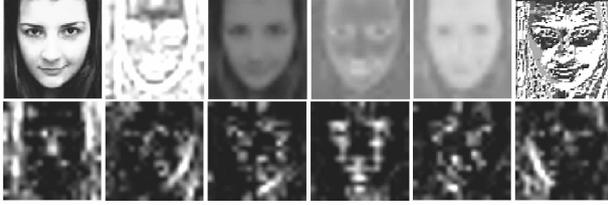


图 1 人脸图像的多通道图

(第 1 行从左至右分别为:原图、梯度幅值、LUV、LBP;第 2 行:6 个量化角度的梯度图)

Fig. 1 Multi-channel maps of human face image

(First line from left to right: Original, gradient amplitude, LUV, and LBP; Second line: Six-quantitative-angle gradient maps)

而在实际应用中,通常选择 LUV 色彩通道、量化梯度角度图通道和梯度幅值通道组成 10 通道特征。这些通道计算量较小,且具有较好的特征表达能力。从图 1 可以看出,不同角度的梯度图反映了图像中的方向信息,与 Gabor 滤波器相比有类似之处,LBP 图像代表人脸的纹理信息,而不同的颜色空间可以体现着人脸区域的颜色一致性信息。

## 1.2 聚合通道特征

ACF<sup>[21]</sup>是多通道图的改进版本,同样需要计算 10 个通道特征,但是 ACF 对多通道图进行了降采样处理。首先输入人脸彩色图像,然后通过图像平滑去噪声,再提取多通道图;其次对多通道图进行降采样。例如 80 像素 $\times$ 80 像素的人脸多通道图,每隔 4 像素进行降采样,把原有的 10 个通道图缩放为原来的 1/4。降采样操作具有两个优点:(1)降低了图像标注的歧义性;(2)提高了特征对噪声的抗干扰能力。最后将降采样后多通道图中的所有通道连接成特征向量。具体流程如图 2 所示。

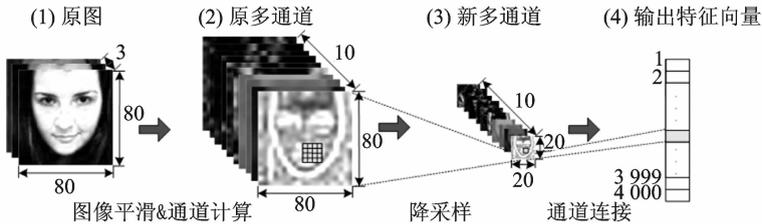


图 2 ACF 的特征提取过程

Fig. 2 Feature extraction process of ACF

## 1.3 HAAR 特征

传统的 HAAR 特征(如图 3(a)所示)分别定义了垂直、水平与对角方向相邻矩形区域内像素平均灰度值的差异,并使用该差异作为特征来区分人脸和非人脸。文献[12]将 HF 进行 45° 旋转,提出了一种扩展的 HF,应用到人脸检测方面,提高了检测率。文献[16]提出了一种不邻接的 HF(如图 3(b)所示),在检测侧面人脸方面取得了较好的效果。HAAR 特征具有速度快、简单易实现的特点,然而 HAAR 特征具有较弱的判别能力,尤其在 Boosting 训练的后期,层级分类器需要使用几百个 HAAR 特征来达到特定的检测率。

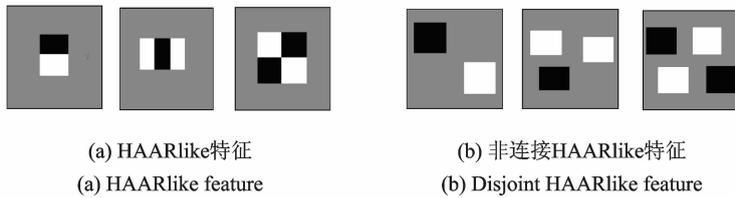


图 3 基于 HAARlike 特征的代表

Fig. 3 Representation based on HAARlike features

## 2 判别投影 HAARlike 特征

通过计算人脸样本多通道图的平均脸(图 4),不难发现每个通道具有明显的结构信息,比如第 4 个通道和第 8 个通道的结构信息尤为明显,而 HAAR 特征能够刻画这种亮度变化信息。

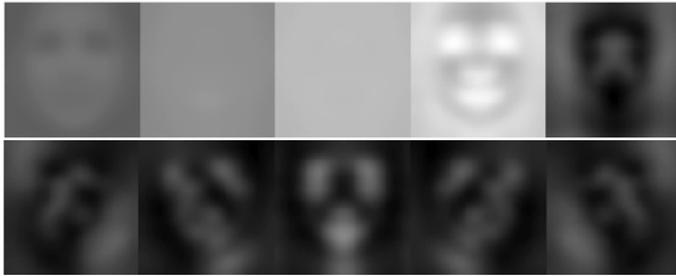


图 4 10 通道正面视角平均脸

Fig. 4 Average face from front view of ten channels

假设从另一角度来看待 HAAR 特征,它可以被视为一个二维特征向量在特定方向的投影过程,该过程可以定义为

$$f = \sum_{i=1}^k w_i \mu_i = \mathbf{w}^T \mathbf{u} \quad k=2,3,4 \quad (1)$$

式中: $w_i$  表示第  $i$  个矩形区域的权值, $\mu_i$  表示第  $i$  个矩形区域内的平均特征值, $k$  表示矩形区域总数。如图 3(a)所示,当  $k=2$ ,  $\mathbf{w}=(1,-1)$  时, $f$  就退化为二维 HAAR 特征。图 3(a)中的 HF 可以视作投影方向为  $(-1,1)^T$ ,  $(1,-2,1)^T$ ,  $(1,-1,1,-1)^T$  的 HF。

为了研究不同投影方向对应 HAAR 特征判别能力的影响,图 5 展示了 5 000 个人脸样本和 20 000 非人脸样本在某 HF 下的散点分布图,其中横坐标  $\mu_1$  和纵坐标  $\mu_2$ ,分别表示 HAAR 特征的黑色区域与白色区域的归一化灰度平均值。从图 5(a)中可以看出正样本分布集中呈椭圆球状,位于直线  $u_2 = u_1$  的左侧,而图 5(b)中的负样本分布比较分散,位于直线  $u_2 = u_1$  的两侧附近。

从图(5)可以看出,特征投影会直接影响 HAAR 特征的判别能力,而提高判别能力的关键在于寻找一个最优的投影方向,使其具有最佳的分类效果和泛化能力,即寻找具有最大 Margin 的投影方向。针对该问题,本文提出一种判别投影 HAAR 特征(Pooling haar feature, PHF),该特征使用 Fisher 判别分析(Fisher discriminative analysis, FDA)方法学习最优投影向量  $\mathbf{w}_{opt}$ ,该投影向量可以通过求解目标函数  $\mathbf{w}_{opt}$  得到

$$\mathbf{w}_{\text{opt}} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}} \quad (2)$$

式中  $\mathbf{S}_w$  与  $\mathbf{S}_b$  分别为类内和类间离散度矩阵。式(2)利用正负样本特征分布信息,寻找最优投影方向,使得类内差异最小化,类间差异最大化,从而提高了 HF 的判别能力。根据 Fisher 准则,易得  $\mathbf{w}_{\text{opt}} = \mathbf{S}_w^{-1}(\mathbf{m}_1 - \mathbf{m}_2)$ 。实验结果表明利用该判别投影特征,更容易区分正负样本,能够有效地降低误判率,提高检测率。

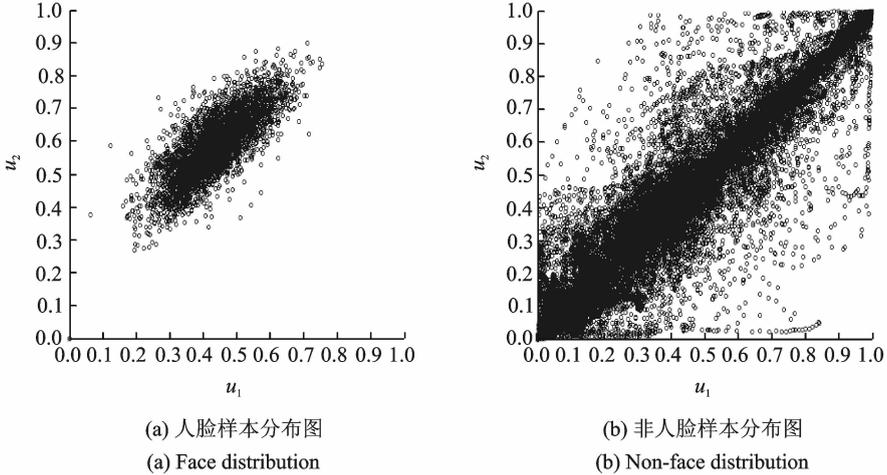


图 5 人脸和非人脸样本分布散点图

Fig. 5 Distribution scatter plot of face and non face samples

假设用  $N, W, H$  分别表示样本总个数,图像宽度和高度。 $\mathbf{S}_b$  为类间离散度矩阵,计算时间复杂度为  $O(10 \times N \times W \times H)$ 。 $\mathbf{S}_w$  为类内离散度矩阵,计算时间复杂度为  $O(10 \times N \times (N \times W \times H))$ 。本文中, $N$  取 50 000, $W$  和  $H$  相等,都取值 80。空间复杂度一般是用训练样本和测试样本所占的内存来表示的。每个像素是浮点型,占 4 个字节,一张多通道训练样本图像是  $80 \times 80 \times 10$ ,约为 256 KB,而训练样本为 50 000 张图像,大概占 12.8 GB 内存,测试样本约占 3.14 GB 内存,在普通电脑上完全可以运行检测器。

### 3 线性非对称分类器

基于 Adaboost 的人脸检测算法在训练强分类器过程中忽略了正负样本的高度非对称性,难以对正负样本正确分类,并且随着迭代次数增加权重会不断增大,导致最后的分类器发生性能退化。基于此缺点文献[23]提出了一种线性非对称分类器(Linear asymmetric classifier, LAC)方法,该方法以非对称性为目标训练强分类器,明显提高了强分类器的性能。它首先假设正负样本分别服从于  $x \sim (\bar{x}, \Sigma_x)$ ,  $y \sim (\bar{y}, \Sigma_y)$  的某分布,然后构建线性分类器  $H(z)$ ,使得在固定误判率条件下,检测率最大化。算法中优化目标函数为

$$\begin{aligned} & \max_{a \neq 0, b} \Pr_{x \sim (\bar{x}, \Sigma_x)} \{ \mathbf{a}^T x \geq b \} \\ & \text{st. } \Pr_{y \sim (\bar{y}, \Sigma_y)} \{ \mathbf{a}^T y < b \} = \beta \end{aligned} \quad (3)$$

式中:  $H(z) = \delta(\mathbf{a}^T x \geq b)$ ,  $\delta$  为狄拉克函数,  $\mathbf{a}$  为线性分类器系数向量,  $b$  为阈值标量,  $\beta$  为最小误判率阈值。求解目标函数式(3)首先对  $b$  消元,转化为

$$\min_{\psi_{x,a}} \left( \frac{\mathbf{a}^T(\bar{y} - \bar{x}) + \psi_{y,a}^{-1}(\beta) \sqrt{\mathbf{a}^T \boldsymbol{\Sigma}_y \mathbf{a}}}{\sqrt{\mathbf{a}^T \boldsymbol{\Sigma}_x \mathbf{a}}} \right) \quad \psi_{x,a}(b) = \Pr\{x_a \leq b\} \quad (4)$$

最后利用样本近似服从高斯分布假设, 概率的累积分布函数的单调性和样本分布中值接近均值的性质, 最终将式(4)转化为求解如下等价问题

$$\max_{\mathbf{a} \neq 0} \frac{\mathbf{a}^T(\bar{y} - \bar{x})}{\sqrt{\mathbf{a}^T \boldsymbol{\Sigma}_x \mathbf{a}}} \quad (5)$$

由式(5)优化形式和 FDA 的相似性, 可以得出封闭解

$$\mathbf{a}^* = \boldsymbol{\Sigma}_x^{-1}(\bar{x} - \bar{y}) \quad b^* = \mathbf{a}^{*T} \bar{y} \quad (6)$$

式中  $\boldsymbol{\Sigma}_x$  为半正定矩阵, 然而为了保证  $\boldsymbol{\Sigma}_x$  正定, 一般加上一个小的正偏置  $\lambda$ , 即  $\boldsymbol{\Sigma}_x + \lambda I$ 。具体推导过程参见文献[23]。

由此可见, LAC 方法利用正样本分布的近似高斯特性和负样本分布的对称性, 通过优化目标函数式(3)来构建最优非对称线性分类器。该方法直接利用样本特征的均值向量和协方差矩阵来计算分类器系数, 计算复杂度低, 对训练分类器速度影响较小, 同时又有效地提高了分类器性能。

#### 4 基于多通道 PHF 的非对称线性级联分类器

针对基于 HF 特征的分类器判别能力较弱及相应的强分类器误判率较高, 正负样本高度不平衡的缺点, 本文提出了一种改进的线性非对称分类器用于多视角人脸检测框架。利用 PHF 代替 HF 特征, 可以提高弱分类器对负样本的判别能力。利用 LAC 方法可以构建在给定目标误判率下, 最优的线性分类器。LAC 算法训练分类器分为 4 个步骤, 首先利用正负训练样本分布信息计算底层 ACF 特征, 其次计算所有的 PHF 特征, 再利用特征选择器选择特征, 然后利用训练样本的分布信息计算最优弱分类器组合系数, 最后输出最优判别强分类器。本文方法的流程如图 6 所示。

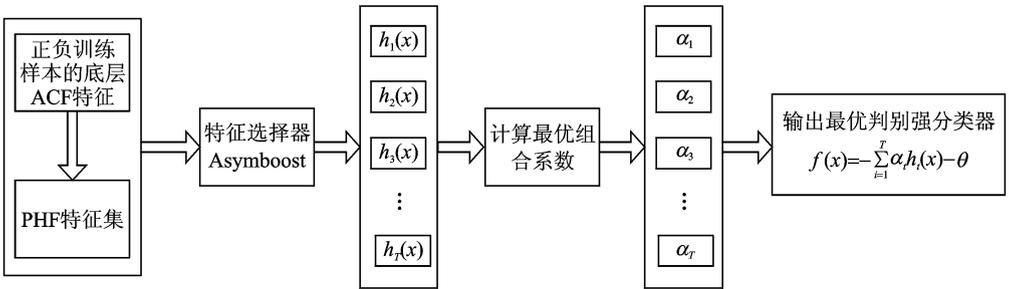


图 6 非对称线性分类器流程图

Fig. 6 Flow chart of ALC

本文使用非对称 GentleBoost 算法<sup>[24]</sup>作为特征选择器。使分类器训练过程中偏向正样本, 从而提高检测率。因为非对称 Adaboost 偏向于正确分类正样本, 该方法将正负样本错分的代价因子引入了损失函数

$$J(F) = E[I(y=1)e^{-C_1 F(x)} + I(y=-1)e^{-C_2 F(x)}] \quad (7)$$

式中:  $F(x)$  为强分类器;  $y$  为样本类别标签;  $C_1, C_2$  为代价因子。

本文提出的改进算法如图 7 所示, 该算法首先利用样本分布信息计算底层 ACF 特征, 然后在此基础上计算所有的 PHF 特征, 再利用非对称 Adaboost 算法选择  $T$  个弱分类器  $\{h_t\}, t=1, \dots, T$ ; 同时对每个样本计算  $T$  维特征向量  $\mathbf{h}(x_t)$ ; 接着计算正负样本的特征向量均值和协方差矩阵, 并利用式(6)计算系数向量  $\mathbf{a}$  和阈值  $b$ 。最后输出强分类器  $F(x)$ , 当  $F(x) > 0$  时判别  $x$  为正样本, 否则为负样本。

本文最后使用文献[10]中的软级联技术,将训练好的弱分类器进行重新学习得到每级的阈值,构成最终的强分类器。

输入:

正样本训练集  $S^+ = \{x_i, y_i\}_{i=1}^{N^+}$ , 负样本训练集  $S^- = \{x_i, y_i\}_{i=1}^{N^-}$ ,  $N^+$ ,  $N^-$  分布代表正负样本数。 $N = N^+ + N^-$  为样本总数,  $T$  为弱分类器个数,  $y_i \in \{-1, +1\}$ 。

流程:

- (1) 计算正负样本的 ACF 特征图;
- (2) 在 ACF 特征图基础上计算 PHF 特征;
- (3) 使用非对称 GentleBoost 算法选择  $T$  个弱分类器  $\{h_t\}, t=1, \dots, T$ ;
- (4) for  $i=1, \dots, N$ ;
- (5) 计算每个样本在所有弱分类器上的响应  $h(x_i) = (h_1(x_i), h_2(x_i), \dots, h_T(x_i))$ ;
- (6) 计算正负样本的均值和方差

$$\bar{x}^+ = \left( \sum_{i=1}^{N^+} h(x_i) \right) / N^+, \bar{x}^- = \left( \sum_{i=1}^{N^-} h(x_i) \right) / N^-$$

$$\Sigma^+ = \sum_{i=1}^{N^+} (h(x_i) - \bar{x}^+)(h(x_i) - \bar{x}^+)^T / N^+$$

$$\Sigma^- = \sum_{i=1}^{N^-} (h(x_i) - \bar{x}^-)(h(x_i) - \bar{x}^-)^T / N^-$$

- (7) 计算强分器组合系数与阈值,  $\mathbf{a} = (\Sigma^+)^{-1}(\bar{x}^+ - \bar{x}^-)$ ,  $b = \mathbf{a}^T \bar{x}^-$ ;
- (8) 输出强分类器  $F(x) = \text{sign}(\mathbf{a}^T h(x) - b)$ 。

图 7 基于多通道 PHF 的非对称线性分类器算法  
Fig. 7 ALC algorithm based on multi channel PHF

## 5 实验结果及分析

### 5.1 实验设置

本实验使用训练样本来源于 AFLW 数据库<sup>[25]</sup>, 训练图片尺寸是 80 像素  $\times$  80 像素, 根据人脸平面外旋转角划分为  $[-90^\circ, -60^\circ]$ ,  $[-60^\circ, -30^\circ]$ ,  $[-30^\circ, 0^\circ]$ ,  $[0^\circ, 30^\circ]$ ,  $[30^\circ, 60^\circ]$ ,  $[60^\circ, 90^\circ]$  6 个视角, 平面内旋转角 (Pitch) 和俯仰角 (Roll) 限制在  $[-22.5^\circ, 22.5^\circ]$  区间。每个视角的正样本数目分别为 3 726, 4 024, 4 636, 5 069, 4 024 和 3 726。强分类器包含的弱分类器总数为 2 048 个, 一共进行 3 轮的 Bootstrap。Bootstrap 数据通过网络随机收集了 20 000 多张不包含人脸的大图片, 用于训练过程中后期负样本的补充。本文使用图 3 中的基本特征用于生成 PHF, 采用特征数目降采样方法后, 最终包含 35 409 个特征。本实验运行环境为 DELL T7610 服务器, 双 CPU 16 核, 2.6 GHz, 64 GB 内存, 采用 MATLAB + VS2012 实现本文实验的人脸检测器。检测器训练过程中采用文献中的预计算策略后, 消耗时间约为 4 h。

为了验证本文方法的有效性, 实验使用 Fddb 公开标准测试集进行测试。该测试集包含了 2 845 张图片, 共 5 171 个正面人脸。该数据集包含各种复杂的环境因素, 如低分辨率、模糊、遮挡和光照变换等, 也是当前规模最大的人脸检测测试数据集。

### 5.2 实验结果及分析

本文实验结果的 ROC 曲线如图 8 所示, 横坐标 False positive (负样本被错误检测为正样本的个数) 代表误检数目, 纵坐标 True positive rate (正样本被正确检测到的概率) 代表检测率。调整最后一级节点的阈值得到不同检测率与相应的误检数目。图 8(a) 为本文方法和当前已有方法在连续数据上的比较。从 ROC 曲线可看出多通道特征 (ACF-ours) 比原始 haar 特征 (Viola-Jones)<sup>[10]</sup> 检测效果好很多。而在多通道基础上加上原始 HAAR 特征 (ACF-HF-ours) 和 ACF-ours 相比, 检测率高了 2% 左右, 说明了多通道特征加中间层滤波器的有效性。图中还可以看到本文提出的多通道增强型 HAAR 特征 (ACF-PHF-

ours) 的检测率相比 ACF-HF-ours 高了近 4%，这说明改进后的 HAAR 特征检测效果更好。而和当前主流的 NPD 特征<sup>[26]</sup>相比, ACF-PHF-ours 也高了 2% 左右的检测率。而诸如 Boosted Exemplar<sup>[27]</sup>、SURF-frontal/multiview<sup>[28]</sup> 和 PEP-Adapt<sup>[29]</sup> 等方法的检测率已明显低于本文方法。同时在图 8(b) 的离散数据 ROC 曲线上, 可看到在误检数目少于 250 时, 本文方法仍有着最好的检测效果。

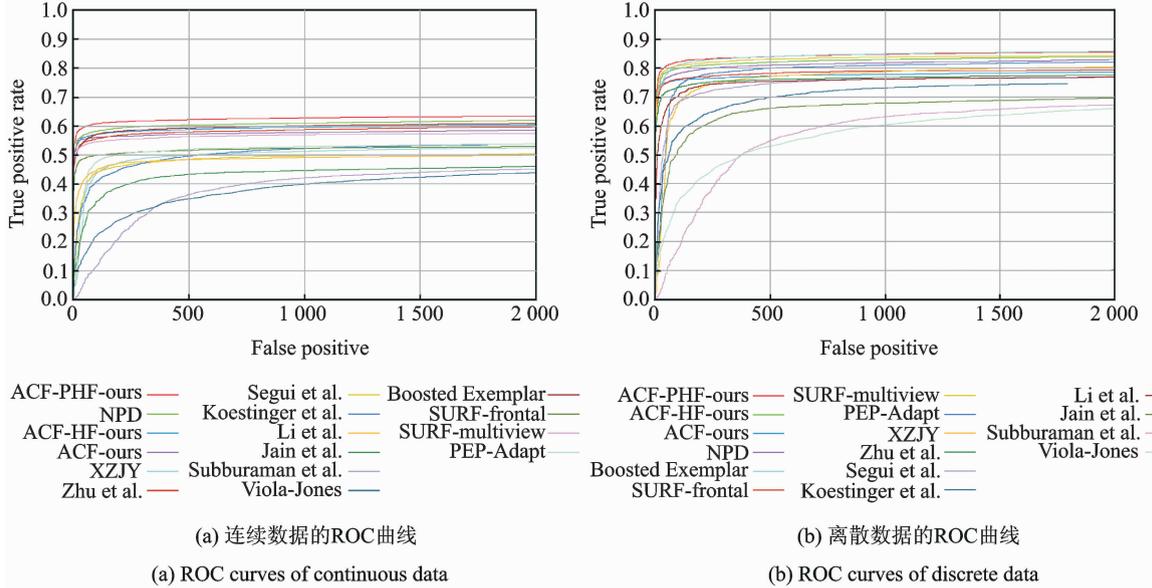


图 8 本文方法与相关算法对比的 ROC 曲线

Fig. 8 ROC curves comparison

本文在上面用了直观易懂的 ROC 曲线来表示实验结果, 这里把曲线数字化, 以表格的形式表示各个方法的平均检测精度, 如表 1 所示。从表中可以看到, 误检个数为 1 时, ACF-HF-ours 的正检率最高, 而当误检个数变大, 可以很清楚看到本文提出的 ACF-PHF-ours 有着相比之下最高的检测精度, 和 ROC 曲线一样, 说明了本文方法的有效性。

表 1 不同算法和误检数目下的正检率

Tab. 1 True positive rate of different algorithms under different false positive numbers %

Algorithm	False positive number				
	1	10	50	100	200
ACF-PHF-ours	50.30	<b>73.80</b>	<b>80.66</b>	<b>82.01</b>	<b>82.87</b>
ACF-HF-ours	<b>60.74</b>	73.00	80.33	81.90	82.75
ACF-ours	21.50	53.30	72.11	73.72	74.63
NPD	60.55	72.31	76.50	77.97	79.40
XZJY	0.80	7.80	51.54	67.51	73.08
Boosted Exemplar	57.00	69.29	77.66	80.81	82.36
SURF-multiview	49.49	69.43	79.21	80.60	81.63

由表 1 可知本文提出的基于多通道增强型 HAAR 特征的多视角人脸检测方法精度高, 检测效果好。图 9 为本文方法在 FDDB 库上的部分检测结果, 椭圆显示的是标注信息, 方框为本文的检测结果。可看出漏检和误检的情况仍少量存在。具体分析, 前者是由于人脸平面内旋转角度过大、遮挡以及超低分辨率或因拍摄角度等因素造成; 而误检主要发生在人脸附近, 其原因在于负样本中不包含人脸, 因此会在部分人脸局部区域产生误判。



图9 FDDB库上的部分检测结果

Fig. 9 Part of the test results on the FDDB dataset

## 6 结束语

本文提出了一种基于多通道图的判别投影 HAAR 特征,这是以多通道图为底层滤波器,在此基础上加了一种改进型的 HAAR 特征,该特征充分利用了正负样本在通道特征空间的脸部结构信息,提高了 HAAR 特征判别能力。而在分类器算法的选择上,由于传统的 Adaboost 分类算法在训练强分类器过程中忽略了正负样本的高度非对称性,继而难以正确分类正负样本,而且权重会随着迭代次数增加而增大,大大降低了分类器性能。因此本文采用非对称 GentleBoost 算法进行特征选择并结合 LAC 方法进行分类器二次训练,充分考虑了样本分布的不平衡特性,获得了最优线性强分类器。实验结果表明:本文提出的方法具有高检测率、低误判率,和当前比较流行的算法相比是一种性能更优越的人脸检测方法。此外,本文提出的方法具有很好的泛化性能,还可以用于刚体或半刚体的目标检测中,如行人检测,车辆检测等领域,是一种推广性比较强的方法。

## 参考文献:

- [1] Yang M H, Kriegman D J, Ahuja N. Detecting faces in images: A survey[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, 24(1): 34-58.
- [2] 李武军, 王崇骏, 张伟, 等. 人脸识别研究综述[J]. *模式识别与人工智能*, 2006, 19(1): 58-66.  
Li Wujun, Wang Chongjun, Zhang Wei, et al. Review of face recognition research[J]. *Pattern Recognition and Artificial Intelligence*, 2006, 19(1): 58-66.
- [3] Rowley H A, Baluja S, Kanade T. Neural network-based face detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1996, 20(1): 203-208.
- [4] Osuna E, Freund R, Girosi F. Training support vector machines: An application to face detection[C]// *Computer Vision and Pattern Recognition*, 1997, Proceedings. [S. l.]: IEEE, 1997: 130-136.
- [5] 胡昭华, 赵孝磊, 徐玉伟, 等. 一种新的低秩分块稀疏表示的人脸识别算法[J]. *数据采集与处理*, 2015, 30(5): 1109-1120.  
Hu Zhaohua, Zhao Xiaolei, Xu Yuwei, et al. Face recognition algorithm based on novel low-rank and block-based sparse representation[J]. *Journal of Data Acquisition and Processing*, 2015, 30(5): 1109-1120.
- [6] 杨军, 袁红照, 刘妍丽. 基于样本扩张和双子空间决策融合的单样本人脸识别算法[J]. *数据采集与处理*, 2015, 30(1): 148-154.  
Yang Jun, Yuan Hongzhao, Liu Yanli. Algorithm for single sample face recognition based on sample augments and double subspace decision fusion[J]. *Journal of Data Acquisition and Processing*, 2015, 30(1): 148-154.
- [7] 叶继华, 万叶晶, 刘长红, 等. 基于多子空间直和特征融合的人脸识别算法[J]. *数据采集与处理*, 2016, 31(1): 102-107.  
Ye Jihua, Wan Yejin, Liu Changhong, et al. Face recognition algorithm of feature fusion based on multi-subspaces direct sum [J]. *Journal of Data Acquisition and Processing*, 2016, 31(1): 102-107.
- [8] Yang M H, Roth D, Ahuja N. A SNoW-based face detector[J]. *Advances in Neural Information Processing Systems*, 1999, 17(3): 862-868.
- [9] Schneiderman H, Kanade T. A statistical model for 3D object detection applied to faces and cars[J]. *Proc IEEE Conf Computer Vision & Pattern Recognition*, 2005, 1: 1-6.
- [10] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[J]. *Proc CVPR*, 2001, 1: 511.

- [11] 陈杰,陈熙霖,高文. 基于遗传算法重采样的人脸样本扩张[J]. 软件学报,2005,16(11):1894-1901.  
Chen Jie, Chen Xilin, Gao Wen. Face sample expansion based on genetic algorithm resampling[J]. Journal of Software, 2005, 16(11):1894-1901.
- [12] Lienhart R, Maydt J. An extended set of Haar-like features for rapid object detection[C]// International Conference on Image Processing. [S. l.]:IEEE, 2002:1-900-I-903.
- [13] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: A statistical view of boosting (With discussion and a rejoinder by the authors)[J]. Annals of Statistics, 2000, 28(2):374-376.
- [14] Li Stan Z, Zhang Zhenqiu. FloatBoost learning and statistical face detection[J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 2004, 26(9):1112-1123.
- [15] Huang C, Ai H, Wu B, et al. Boosting nested cascade detector for multi-view face detection[J]. Proc of ICPR, 2004, 2(2):415-418.
- [16] Li Stan Z, Zhu Long, Zhang Zhenqiu, et al. Statistical learning of multi-view face detection[C]// European Conference on Computer Vision. Copenhagen, Denmark:IEEE, 2002:67-81.
- [17] Huang C, Ai H, Li Y, et al. Learning sparse features in granular space for multi-view face detection[C]// International Conference on Automatic Face and Gesture Recognition. [S. l.]:IEEE, 2006:401-406.
- [18] Zhang Lun, Chu Rufeng, Xiang Shiming, et al. Face detection based on multi-block LBP representation[J]. Lecture Notes in Computer Science, 2007, 4642:11-18.
- [19] Papageorgiou C P, Oren M, Poggio T. A general framework for object detection[C]// International Conference on Computer Vision. Bombay, India: IEEE, 1998:555-562.
- [20] Dollár P, Tu Z, Perona P, et al. Integral channel features[C]// British Machine Vision Conference, BMVC 2009. London, UK:[s. n.], 2009.
- [21] Dollar P, Appel R, Belongie S, et al. Fast feature pyramids for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(8):1532-1545.
- [22] Yang Bin, Yan Junjie, Lei Zhen, et al. Aggregate channel features for multi-view face detection[J]. IEEE International Joint Conference on Biometrics, 2014:1-8.
- [23] Wu Jianxin, Brubaker S C, Mullin M D, et al. Fast asymmetric learning for cascade face detection[J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 2008, 30(3):369-382.
- [24] Masnadi-Shirazi H, Vasconcelos N. Asymmetric boosting[C]// International Conference on Machine Learning. [S. l.]:ACM, 2007:609-619.
- [25] Kostinger M, Wohlhart P, Roth P M, et al. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization[C]// IEEE International Conference on Computer Vision Workshops, ICCV 2011 Workshops. Barcelona, Spain:[s. n.], 2011:2144-2151.
- [26] Liao Shengcai, Jain A K, Li Stan Z. A fast and accurate unconstrained face detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2):211-223.
- [27] Li Haoxiang, Lin Zhe, Brandt J, et al. Efficient boosted exemplar-based face detection[C]// Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA :IEEE, 2014:1843-1850.
- [28] Li J, Zhang Y. Learning SURF cascade for fast and accurate object detection[J]. Computer Vision and Pattern Recognition, 2013, 9(4):3468-3475.
- [29] Li H, Hua G, Lin Z, et al. Probabilistic elastic part model for unsupervised face detector adaptation[C]// IEEE International Conference on Computer Vision. Sydney, NSW, Australia :IEEE, 2013:793-800.

## 作者简介:



沈继锋(1980-),男,博士,研究方向:目标检测, E-mail: shenjifeng@uj.s.edu.cn.



时士伟(1992-),男,硕士研究生,研究方向:人脸检测。



左欣(1980-),女,博士,研究方向:图像检索。



徐丹(1981-),女,博士,研究方向:目标检测。

