

基于级联框架下的多特征融合人脸检测算法

鲁 鹏, 陈毅松, 陈文广

(北京大学机器感知与智能教育部重点实验室, 北京 100871)

摘 要: 针对 AdaBoost 人脸检测方法在高分辨率彩色图像上定位速度慢和误检率高的问题, 提出一种多特征融合的人脸检测方法。该方法使用级联策略将多种特征分类器有效地组合起来, 高效地利用各种特征之间的互补性, 形成一种新型的高性能分类器。实验结果显示, 该方法提高了检测速度、降低了误检率。

关键词: 人脸检测; 级联结构; 特征分类器

Cascade Based Multi-feature Fusion Method Algorithm for Face Detection

LU Peng, CHEN Yi-song, CHEN Wen-guang

(Key Lab of Machine Perception and Intelligent, MOE, Peking University, Beijing 100871, China)

【Abstract】 In order to improve the detection speed and decrease the false alarm rate of the AdaBoost algorithm in high resolution color image, a multi-feature fusion method is proposed. It adopts cascade strategy to combine multi-fuse classifiers. Benefiting from the complementary of multi-feature, the proposed method can achieve high performance. Experimental result demonstrates that the proposed method has better global precision and is less time consuming.

【Key words】 face detection; cascade structure; feature classifier

DOI: 10.3969/j.issn.1000-3428.2011.02.003

1 概述

人脸检测是指对于任意一幅给定的图像, 采用一定的策略对其进行搜索以确定其中是否含有人脸, 如果是则返回人脸的位置、大小和姿态。目前的研究^[1-2]大体可以分为 3 类: 基于特征的方法, 基于肤色模型的方法^[3], 基于统计理论的方法^[4]。文献[5]在达到较高检测率的同时, 使得人脸检测速度有了实质性的提高, 因此, 逐渐成为一种主流的研究方向。为了提高人脸检测系统的速度, Viola^[5]使用一种级联结构的分类器设计方式, 这种设计的主要思想是逐级提高检测精度。首先使用结构较简单(弱分类器数目较少)的强分类器进行非人脸窗口的排除, 后续的强分类器结构越来越复杂(弱分类器个数多), 检测精度越来越高, 但需要检测的子窗口越来越少, 从而达到提高检测速度的目的。另一方面, Viola 在进行人脸描述时采用了类 Haar 特征, 即利用人脸的灰度信息来形成弱分类器; 为了加快计算速度, Viola 同时设计了一种基于积分图的快速特征计算方法。

随着计算机图形硬件性能的快速提升以及多媒体、网络技术的发展, 高分辨率彩色图像的使用越来越广泛。当将 Viola 的 AdaBoost 算法应用于高分辨率彩色图像时, 由于待检测窗口数量巨大, 此时前端结构较简单的强分类器的总运算量显著加大, 同时由于前端分类器能排除的非人脸样本能力未变, 导致后端运算量较大的分类器需要进行更多的子窗口判断, 最终导致算法在高分辨率图像上定位速度较慢, 误检数量增加。另一方面, 虽然采用了基于积分图的快速特征计算方式, 但当待检窗口较多时仍然无法解决检测速度慢的问题。

级联分类器的结构事实上就是一个退化的决策树, 如果能利用多种特征信息设计一些结构简单且分类能力较强的强分类器, 并将其级联在最前端, 将能有效地滤除更多的非人

脸窗口、减少后端强分类器的待检测的窗口数量。为此, 本文提出了一种级联框架下的多特征融合人脸检测算法, 算法中主要采用了类 Haar 特征、肤色特征以及边缘特征。

2 分类器设计

2.1 类 Haar 特征分类器

本文采用的基于类 Haar 特征分类器主要基于 AdaBoost 算法, 采用级联方式将若干强分类器串联起来, 其中每一个强分类器均由 AdaBoost 算法学习出的若干类 Haar 特征构成。

2.1.1 类 Haar 特征

类 Haar 特征是由 Haar 小波变换而来的, 它通过相邻区域的灰度差, 也就是亮度关系来描述目标。类 Haar 特征在描述人脸目标时, 其物理意义十分明确, 如图 1 所示, 一张人脸 2 只眉毛区域肯定比其中间的区域暗。那么只要通过 AdaBoost 算法找出足够多的这样的特征, 就能将人脸和非人脸分开。本文采用 5 类 Haar 特征, 如图 2 所示。



图 1 类 Haar 特征描述目标

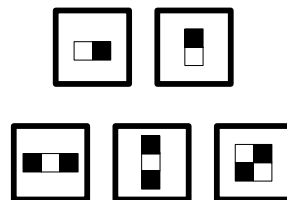


图 2 5 类 Haar 特征

基金项目: 国家“863”计划基金资助项目(2009AA01Z324, 2009AA01Z105, 2007AA01Z159, 2007AA01Z318); 中国博士后科学基金资助项目(20090450233)

作者简介: 鲁 鹏(1978 -), 男, 博士后, 主研方向: 计算机视觉, 人机交互; 陈毅松、陈文广, 副教授

收稿日期: 2010-07-17 **E-mail:** luchenpeng@msn.com

2.1.2 AdaBoost 算法流程

给定 n 个样本, $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 其中包含 m 个负样本且 $y_i = 0$, l 个正样本且 $y_i = 1$, $n = m + l$ 。初始时将负样本和正样本的权值分别设为 $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ 。

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

(1) 归一化权 $w_{t,j} \leftarrow \frac{w_{t,j}}{\sum_{j=1}^n w_{t,j}}$ 。

(2) 对于每个特征 j , 训练出一个弱分类器 h_j , 计算出与之对应的错误率: $\zeta_j = \sum w_i |h_j(x) - y_i|$ 。

(3) 选择错误率 ζ_j 最小的 h_j 。

(4) 更新权 $w_{t+1,i} = w_{t,i} \beta_i^{1-e_i}$, 其中, $\beta_i = \frac{\zeta_i}{1-\zeta_i}$; 如分类正确则 $e_i = 1$, 否则 $e_i = 0$ 。

(5) 最终 T 个弱分类器组成一个强分类器:

$$h(x) = \begin{cases} 1 & \text{当 } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \text{ 时} \\ 0 & \text{当 } \sum_{t=1}^T \alpha_t h_t(x) < \frac{1}{2} \sum_{t=1}^T \alpha_t \text{ 时} \end{cases}$$

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

2.2 肤色分类器

肤色作为人脸最显著的特征之一, 在图像领域的应用是相当广泛的, 从人脸检测到人体跟踪, 大部分算法都利用了肤色信息。已有研究^[6]表明, HSV(色度 H、饱和度 S 和亮度 V)颜色模型更加接近于人对颜色的感知, 且肤色对 HSV 色彩空间的 H 色调具有恒常性。因此, 本文使用 HSV 色彩空间中的 H 分量作为肤色的特征空间, 并利用一维高斯分布来表示人脸肤色的 H 分布。 μ 为肤色分布的均值, σ 为肤色分布的方差。那么, 图像上任意点是肤色的概率可表示为:

$$P_c(i, j) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(h(i, j) - \mu)^2}{2\sigma^2}\right] \quad (1)$$

其中, $P_c(i, j)$ 为图像点 (i, j) 是肤色的概率值, 其范围为 $[0, 1]$; $h(i, j)$ 为图像点 (i, j) 的 RGB 值转换到 HSV 空间后的 H 值; μ 、 σ 的实际值从人工标记的样本中统计得到。

2.2.1 肤色分类器的定义

利用上述肤色模型, 定义肤色分类器如下:

$$h_c = \begin{cases} 1 & \sum_{(i,j) \in R} p_c(i, j) > T_c \\ 0 & \sum_{(i,j) \in R} p_c(i, j) < T_c \end{cases} \quad (2)$$

其中, R 为子窗口内全部像素点集合; T_c 为预先设置的门限值。 h_c 为肤色分类器, 当 h_c 为 0 时, 当前子窗口为非人脸区域; 当 h_c 为 1 时, 当前子窗口可通过肤色检测器。

2.2.2 肤色分类器的快速计算

肤色分类器的计算量由两部分组成, 即 RGB 图像到肤色概率图像的变换以及肤色分类器的计算。

(1) RGB 图像到肤色概率图像的快速计算。RGB 图像到肤色概率图的转换包含 2 个步骤, 第 1 步对输入图像每个像素进行颜色空间转换, 第 2 步针对转换结果计算其概率值。显然, 这样的计算方法所需的计算量较大, 因此, 本文提出了一种基于概率查找表的快速计算方法。概率查找表的建立分为 3 步: 1) 建立 RGB 值到 H 值的三维转换查找表, 即在系统初始化阶段预先将 RGB 空间所有可能 RGB 值对应的 H 值

计算出来保存于内存中; 2) 将查找表内每个点的 H 值代入式(1)中计算其概率值, 并将计算结果替换原有的 H 值; 3) 为了计算方便, 将所有的概率值由 $[0, 1]$ 区间映射为 0~255 级灰度区间。有了肤色概率查找表, 仅通过寻址操作就可以将输入的 RGB 图像转换成 8 位的肤色概率灰度图像。

(2) 肤色分类器的快速计算。肤色分类器的计算量主要来自于肤色概率灰度图中的子窗口区域的像素值求和运算。本文采用积分图像的方法来加速矩形内像素灰度值求和, 基于积分图像的方法仅通过 2 次加法 1 次减法即可计算出肤色概率灰度图像上任意一个矩形子窗口内的像素值之和。因此, 肤色分类器的判定过程仅需 2 次加法 1 次减法操作, 积分图像的具体计算过程详见参考文献[5]。

2.3 边缘分类器

除了肤色信息人脸区域还存在着五官特征, 因此具有丰富的边缘信息。不同于类 Haar 特征与肤色特征边缘特征具有很强的局部描述性, 它能和肤色特征以及类 Haar 特征形成有效的互补。如能将边缘信息合理地利用起来必将能有效地提高人脸检测器的运行效率。而边缘分类器的设计首先要考虑边缘提取, 如何找到一种快速的边缘提取方法直接关系到边缘分类器的性能。本文中采用 SUSAN(Smallest Univalued Segment Assimilating Nucleus)边缘检测算子来实现快速的边缘检测算法。主要原因有 2 点: (1) 边缘检测效果好, 无论对直线, 还是曲线边缘, SUSAN 算法基本上可以检测出所有的边缘。(2) 算法运算量小, 速度快。由于不涉及梯度的运算, SUSAN 算子对每个像素点仅需 8 次加法运算, 而 Sobel 算子即使采用 2 个 3×3 的模板, 对每一点仍需要做 9 次加法、6 次乘法以及 1 次开方运算。其他的算子如 Gauss-Laplace 算子、Prewitt 算子以及 Canny 算子相对于 SUSAN 算子计算量就更大。

SUSAN 算子提取边缘可由图 3 来简单描述。用一个圆形模板在图像上移动, 若模板内像素的灰度与模板中心像素(称为核 Nucleus)灰度的差值小于一定阈值, 则认为该点与核具有相同(或相近)的灰度, 由满足这样条件的像素组成的区域称为 USAN(Univalued Segment Assimilating Nucleus)。当圆形模板完全处在图像或背景中时, USAN 区域面积最大(如图 3 中的 a 和 e 所示); 当模板移向图像边缘时, USAN 区域逐渐变小(如图 3 中的 d 所示); 当模板中心处于边缘时, USAN 区域很小(如图 3 中的 c 所示); 当模板中心处于角点时, USAN 区域最小(如图 3 中的 b 所示)。可以看出, 在边缘处像素的 USAN 值都小于或等于其最大值的一半。因此, 计算图像中每一个像素的 USAN 值, 通过设定一个 USAN 阈值, 查找小于阈值的像素点, 即可确定为边缘点。

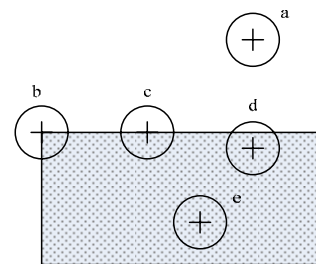


图 3 SUSAN 算子边缘提取示意图

对整幅图像中的所有像素, 用圆形模板进行扫描, 比较模板内每一像素与中心像素的灰度值, 通过与给定的阈值比较来判别该像素是否属于 USAN 区域, 如下式所示:

$$c(r, r_0) = \begin{cases} 1 & |I(r) - I(r_0)| \leq t \\ 0 & |I(r) - I(r_0)| > t \end{cases} \quad (3)$$

其中, $c(r, r_0)$ 为模板内属于 USAN 区域的像素的判别函数; $I(r_0)$ 是模板中心像素(核)的灰度值; $I(r)$ 为模板内其他任意像素的灰度值; t 是灰度差门限。图像中每一点的 USAN 区域大小为:

$$n(r_0) = \sum_{r \in D(r_0)} c(r, r_0) \quad (4)$$

其中, $D(r_0)$ 为以 r_0 为中心的圆形模板区域。

得到每个像素的 USAN 值 $n(r_0)$ 以后,再与预先设定的门限 g 进行比较,当 $n(r_0) < g$ 时,所检测到像素位置 r_0 可以认为是一个边缘点:

$$e(r_0) = \begin{cases} 1 & n(r_0) < g \\ 0 & n(r_0) \geq g \end{cases} \quad (5)$$

2.3.1 边缘分类器的定义

利用上述边缘检测算法,定义边缘分类器如下:

$$h_e = \begin{cases} 1 & \sum_{r_0 \in R} e(r_0) > T_e \\ 0 & \sum_{r_0 \in R} e(r_0) < T_e \end{cases} \quad (6)$$

其中, R 为子窗口内全部像素集合; T_e 为预先设置的门限值。 h_e 为边缘分类器,当 h_e 为 0 时,当前子窗口为非人脸区域;当 h_e 为 1 时,当前子窗口可通过边缘检测器。

2.3.2 边缘分类器的计算

与利用肤色特征进行判别的计算量类似,边缘特征的计算量由两部分组成,即边缘提取以及分类器计算。在边缘提取阶段,SUSAN 算子检测方法对每个像素点仅需 8 次加法运算,满足系统对速度的要求。在边缘分类器的计算上,依然采用基于积分图像的快速计算方法,此时每个子窗口的判定过程仅须 2 次加法 1 次减法。

3 级联框架下的多特征融合人脸检测算法

在进行特征分类器级联时,本文的准则如下:当计算量相差较大时,将各个分类器按照计算量由小到大的顺序级联;当分类器计算量相同或接近时,按照分类器的分类能力由大到小的顺序级联。

基于上述原则,本文提出的级联框架下的多特征融合人脸检测算法流程如图 4 所示。

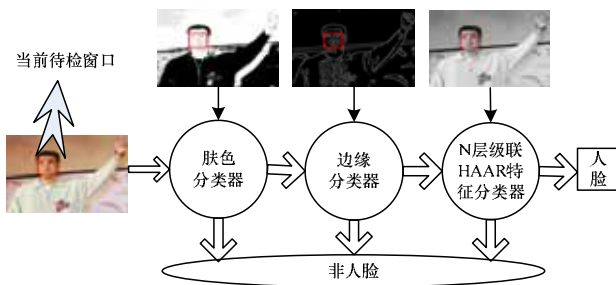


图 4 算法流程

算法分为预处理阶段与分类阶段。在预处理阶段,算法利用肤色概率查找表以及 SUSAN 边缘检测算子生成肤色概率灰度图与边缘灰度图。在分类阶段,算法首先利用肤色分类器对待检测窗口进行判定;如果判定结果为“是”,则立即启用边缘分类器继续对待检测窗口进行判定;如果结果仍然为“是”,则启用级联结构的类 Haar 特征分类器对其进行最终判定,如最终判定结果仍然为“是”,则判定当前窗口为人脸区域。在判定过程中,只要有一个分类器的判定结果为

“否”,则立即停止对当前窗口的判定,并将当前窗口判定为非人脸区域。

4 实验与分析

4.1 分类器主要参数设定

(1)N 层级联类 Haar 特征分类器。本文中采用 AdaBoost 算法进行类 Haar 特征分类器训练。使用了 2 000 张大小为 20×20 黑白人脸图像作为正样本,每层使用的负样本数为 3 000 张,它们分别从 200 张不包含人脸、大小为 $1\,024 \times 768$ 的图像中随机抽取。算法每层正样本错误率设置为 0.1%,负样本错误率为 50%。最终得到的类 Haar 特征分类器由 22 层组成,一共包含 1 290 个特征,其中第 1 层分类器包含 7 个特征,2 块结构的类 Haar 特征 2 个,3 块结构的类 Haar 特征 5 个。

(2)肤色分类器设计中肤色模型的参数 μ 、 σ 来自于对 50 张彩色照片中 317 张预先标定好的人脸区域的统计学习得到。

4.2 算法测试与分析

本文测试样本集包含 200 张彩色照片,其中部分来自于个人真实数码照片,部分来自于互联网。经过简单的缩放和剪切,每张图像大小归一化为 $1\,280 \times 1\,024$,其中包含完整人脸 831 张。算法参数设置为:人脸检测步长 1.0,窗口缩放参数初始值 1.0,步长 1.1。硬件环境为:P4 1.86 GHz CPU,1 GB 内存计算机。

为了更好地评估本文算法,在同一测试集上分别运行了类 Haar 特征分类器人脸检测算法、级联肤色特征分类器与类 Haar 特征分类器的人脸检测算法以及级联肤色特征分类器、边缘特征分类器与类 Haar 特征分类器的人脸检测算法,运行结果如表 1 所示。

表 1 测试结果比较

测试信息	人脸分类器组成		
	类 Haar 特征分类器	级联肤色特征分类器与类 Haar 特征分类器	级联肤色特征分类器、边缘特征分类器与类 Haar 特征分类器
实际人脸数量	831	831	831
正确检测人脸数量	792	789	789
检测率/(%)	95.31	94.95	94.95
误检人脸数量	93	37	24
误检率/(%)	11.19	4.45	2.89
每帧平均耗时/s	5.41	1.11	0.89

从表 1 中可以看到,级联肤色特征分类器与类 Haar 特征分类器的人脸分类器,检测速度明显提高。对每个待检窗口,肤色分类器仅包含 3 次加法运算,远低于第 1 级类 Haar 特征分类器的计算量。因此,肤色区域在图片中所占面积越小时,肤色分类器的分类能力就越强,后端类 Haar 特征分类器须判断的窗口就越少,系统性能也就提升越高。另一方面,算法的误检率明显地降低,这归功于肤色特征与类 Haar 特征的互补性。但算法在检测率上有所下降,主要原因在于本文给出的肤色模型无法适用于所有人种。当检测图片中含有黑人,且光线不均匀时,导致肤色分类器分类错误,直接将其识别为负样本。在类 Haar 特征分类器前端与肤色分类器后端级联边缘分类器后,由于边缘信息与肤色信息的互补性,算法速度再次有了小幅度的提升,误检率进一步降低。

通过以上实验可以看到,级联肤色分类器、边缘分类器与类 Haar 特征分类器的人脸检测算法在与原始的类 Haar 特征分类器检测率接近的前提下,速度上有了较大的提高,同时误检率有了明显的降低。(下转第 12 页)