

基于两级分类器的人脸检测系统设计

张 永 薛芝茂

(兰州理工大学 兰州 730000)

摘 要 人脸识别技术拥有广泛的应用前景,但是目前不少实现方式存在一些不尽人意之处。在对 OPENCV 与 SVM 分类器进行分析的基础上,阐述了基于两级分类器的人脸检测方法的原理和实现过程,首先分析了两级分类器的构建,引入人脸图像的矩形特征向量,将图像的矩形特征作为分类的依据,随后论述了系统设计与实现,包括灰度变换过程、直方图均衡过程、图像平滑过程以及金字塔序列化的实现。这种检测模式能够加快处理速度,提升效率。

关键词 人脸检测,OPENCV,SVM 分类器

Face Detection System Design Based on Two Classifiers

ZHANG Yong XUE Zhi-mao

(Lanzhou University of Technology, Lanzhou 730000, China)

Abstract The face recognition technology has broad application prospects, but right now there are some failings in many implementations. Based on the analysis of OPENCV with the SVM classifier, this paper elaborated a strategy based on two-stage classifier of face detection principle and the implementation process, first, analysed two classifier construction, introduced a rectangular face image feature vector, used the image of the rectangular features as the basis for classification, and then discussed the system design and implementation, including the gray-scale transformation process, the process of histogram equalization, image smoothing serialization process as well as the realization of the pyramid. This test model can speed up the processing speed and has higher efficiency.

Keywords Human face detection, OPENCV, SVM classifier

1 引言

人脸识别技术结合被识别者的面孔特点进行具体身份的确认,拥有广泛的应用前景。当前,已经存在不少对人脸模式特征进行综合判断的具体方法和实现模式,单个的方法由于自身的局限性,往往难以达到十分理想的效果。目前不少人脸检测实现方式都是基于几种方法的改进或者结合,这在一定程度上提升了检测效果,但是仍然存在一些不尽人意之处。另外,在人脸的样本采集方面尚存在一些问题,由于人脸具有一定的不稳定性(例如表情的变化、饰物的佩戴等),使得单一的标准模版往往难以实现理想的匹配。本文采用的基于两级分类器的人脸检测方法,以积分图矩形特征为特征向量,构建一个粗筛选分类器,利用此分类器对检测对象粗筛选,从而使大量可能造成干扰的背景部分被滤掉,达到加快处理速度、避免对非人脸部分的样本采集的目的。

2 两级分类器的构建

2.1 OPENCV 与 SVM 分类器

2.1.1 OPENCV

OPENCV 由 C++ 代码编写,具有平台无关性和程序运行的实时性以及方便灵活的用户接口。OPENCV(开放源代

码的计算机视觉类库)的主要用途是对图像进行高级的操作,具有很高的处理速度和效率。OPENCV 包含了以下模块:(1) CxCore: 为 OPENCV 的运行提供基本数据结构的支持,如错误处理的函数、矩阵和数组的操作等。(2) HighGUI: 在建立快速软件原型时,为用户提供 GUI。(3) CV: 存储着 OPENCV 所使用的主要函数。(4) ML: 支持系统进行机器学习,如回归分析、模式分类等。(5) CVAUX: 存储 OPENCV 的辅助函数。

在以上模块的支持下,利用 OPENCV 可以实现对图像的输入输出、矩阵和向量的操作、数字图像处理以及分析各类结构与运动等。

2.1.2 SVM 分类器

本文采用 SVM 分类器进行人脸检测系统的设计。SVM 分类器具有很强的泛化能力,可以较好地解决小样本问题,具有不错的非线性分类能力,也摒弃了神经网络的过学习和欠学习问题。SVM 的纹理特征向量共生矩阵计算方法为:

$$h(i, j) = \# \{ [(x, y), (x + \Delta x, y + \Delta y)] \mid \text{If } (x, y) = i \& f(x + \Delta x, y + \Delta y) = j \} \\ \# s$$

SVM 分类器的训练是一个关键问题,作为 SVM 的核心部分,选取高斯核函数作为支持向量机的核心:

到稿日期:2009-07-05 返修日期:2009-09-02 本文受甘肃省自然科学基金(0809RJZA015,智能化的混合元搜索引擎研究)资助。

张 永(1963—),男,硕士,教授,硕士生导师,主要研究方向为人工智能、智能信息处理技术、数据库技术;薛芝茂(1979—),男,硕士生,主要研究方向为模式识别。

$$K(u, v) = \exp\{-\gamma \|u - v\|^2\}$$

2.2 粗筛选分类器的构建

在进行实际的人脸检测时,人脸与非人脸的出现几率有着很大的差别,前者远远小于后者。在不存在人脸的区域中,背景的属性也千差万别。因此,本文提出的方法是:首先采用较为简单的线性分类器将大量不相关的背景滤除,随后,对于简单的分类器难以处理的区域,再使用 SVM 分类器做进一步的操作,这也是粗筛选分类器的实现原理。其流程如图 1 所示。

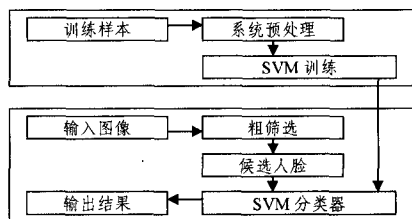


图1 粗筛选分类器的实现原理流程图

首先,引入人脸图像的矩形特征向量,将图像的矩形特征作为分类的依据,矩形特征的含义是人脸图像大小相同的矩形内部像素灰度值之和的差值。对图像进行一次双重积分:

$$f(x, y) = \int f(x_1, y_1) dx_1 dy_1$$

其中, $f(x_1, y_1)$ 是原始的人脸图像, $f(x, y)$ 是先按行积分, 然后按列积分的结果值。此值的含义是原始的人脸图像中, 以某点为基准, 其左上方所围成的像素值的和。对于原始的人脸图像点坐标 (x, y) 而言, 其积分图像的值可以表示为:

$$U(x, y) = \sum [u(x_1, y_1)]$$

图2为其几何意义。

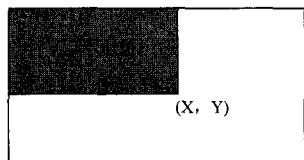


图2 积分图像值的几何意义

由此可知,在引入积分图后,可以通过较少的操作过程得到人脸图形的矩形特征。在 OPENCV 函数中, `cvIntegral` 函数可以很好地支持积分图:

```
void cvIntegral(const Cv* image, CvArr* sum, CvArr* sqsum, CvArr* tilted)
```

其中,参数的含义为:

Image 参数:单通道模式输入图像;

Sliiri 参数:单通道模式积分图像;

Sqsum:对象值平方的积分图像;

Tilted:45 度角积分图像。

图3中的人脸被分为 A, B, C, D, E, F 6 个区域, 3 行 4 列。其中,每一行的两部分基本对称,区域灰度值相近。对于区域 A 来讲,可以算出其像素值的和为:

$$S_A = u(x_4, y_4) - u(x_2, y_2) - u(x_3, y_3) + u(x_1, y_1)$$

定义 $N = S_A / T$ 表示区域 A 的平均灰度值,其中 T 为区域 A 的像素数目。 $(x_4, y_4), (x_2, y_2), (x_3, y_3), (x_1, y_1)$ 分别表示左上角、右上角、左下角、右下角 4 点坐标。随后,将 $(S_A, S_B, S_E, S_F, N_A, N_C, N_F, N_D)$ 作为特征向量输入粗筛选分类器进行训练,从而将非人脸的大部分区域排除掉,留下人

脸区域。本文针对上文所述提取的矩形特征向量 $(S_A, S_{BF}, N_A, N_C, N_D)$ 构建分类器分类函数为:

$$Q(x) = \begin{cases} 1, & S_A - S_B < S_B * 0.1, S_E - S_F < S_E * 0.1, M_A < M_C, \\ & M_A < M_D, M_E < M_D \\ 0, & \text{其他} \end{cases}$$

$$X = S_A, S_B, S_E, S_F, M_A, M_C, M_E, M_D$$



图3 人脸图像的矩形特征向量提取

通过上文所构建的粗筛选分类器,可以在最大程度上去除非人脸的背景图像,从而实现人脸区域的精确化。

3 系统设计与实现

3.1 系统框架分析

人脸的检测方法有两大类,分别是基于图像的方法与基于特征的方法。基于图像的方法结合了模式识别理论,根据事先存储的图像模型去确定样本图像里所包含的窗口图形是否为人脸图像,其优势是可以得到比较高的检测率,应用广泛,缺点是检测速度不够快;基于特征的方法则结合人类脸部的形状特点、肤色和运动特性对人脸进行检测,优点是检测速度快,缺点是适用范围狭窄,检测率不够高。本文综合两种方法的优势,尽可能提升检测速度和检测率,具体做法为:首先构建一个采用特征基方法的弱分类器,对可能存在人脸的区域进行粗筛选,将人脸候选区域快速确定下来,然后使用 SVM 分类器对上一步的检测结果加以确认,确认成功后,在图像中的人脸区域进行标记。图4为本文所构建的人脸检测系统框架。

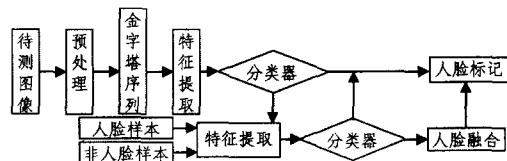


图4 基于两级分类器的人脸检测系统框架

3.2 人脸图像预处理过程

3.2.1 灰度变换过程

对于所采集的样本图像,第一个处理步骤是进行灰度变换,使其由彩色图像转换为灰度图像,为后续的处理做好准备。可以使用以下的方式进行灰度变换过程。

$$(1) grey = \frac{(R+G+B)}{3}。R, G, B \text{ 分别代表彩色的样本图像中某一点的红色分量、绿色分量和蓝色分量。将这3种分量的平均值作为图像中一个像素的灰度值。}$$

(2) $Gray = \text{MAX}(R, G, B)$ 。意即:在彩色图像中,选取某一像素的红色、绿色和蓝色分量的最高值作为此像素的灰度值。

(2) $Gray = \text{MAX}(R, G, B)$ 。意即:在彩色图像中,选取某一像素的红色、绿色和蓝色分量的最高值作为此像素的灰度值。

(3) $Gray = R * 0.3 + G * 0.59 + B * 0.11$ 。意即:将红色、绿色和蓝色分量的加权平均值作为一个像素的灰度值。

在此引入 OpenCV 提供的功能函数来实现输入图像的转换:

```
void cvCvtColor(const CvArr * src, CvArr * dst, int code)
```

以实现灰度值的变换,其中:

参数 src 表示所输入的具体图像,可以是 8 位整数格式,也可以是浮点数据格式;

参数 dst 表示所输出的灰度变换结果,可以是 8 位整数格式,也可以是浮点数据格式;

参数 code 表示色彩空间转换,一般为一个实现定义的常数值。

3.2.2 直方图均衡过程

直方图可以表示出样本图像的像素中,每一像素的灰度值与此像素出现的频率之间的具体联系。直方图可以对样本图像进行总体描述,比如其灰度范围、灰度的具体分布以及每一个灰度级别的出现频率,还包括样本图像的整体明暗以及对比度,这些都是对图像进行下一步处理所需的。直方图均衡过程的作用是使图像细节更清晰,由于所采集的样本图像灰度值分布区间往往比较小,容易导致图像细节模糊,经过直方图均衡过程之后,能够使灰度分布均匀,灰度间距拉开,反差变大。

具体实现方法为:首先,假设样本图像像素 $f(x, y)$ 的最大取值为 $f_{\max}$,最小值为 $f_{\min}$;对直方图进行归一化;设样本图像像素灰度值取值范围为 L ,设 r 的取值范围为 $[0, 1]$,

$$r = \frac{f - f_{\min}}{f_{\max} - f_{\min}} = \frac{f}{L}$$

则可将灰度直方图表示为:

$$P_k r_k = \frac{n_k}{n}$$

其中, n 表示样本图像中的像素数目之和, k 的取值为 0 到 L , n_k 的含义是灰度值为 Y_k 的像素的数目之和。引入函数 $s = T(r)$ 进行灰度映射,其目的是将样本图像的直方图 $P_r(r)$ 转换为目标直方图 $P_s(s)$ 。

首先有:

$$P_s(s) = P_r(r) \frac{dr}{ds}$$

考虑到直方图均衡化的目标,将样本图像改变为均匀分布的直方图,使图像信息量所占区域内的比重增加。由直方图均衡化可知: $P_s(s) = 1$ 。由此可得:

$$ds = \frac{p_s(r) dr}{p_s(s)} = p_s(r) dr$$

所以

$$s = T(r) = \int_0^r P_r(w) dw$$

在图像处理时,直方图均衡化遵循公式:

$$s_k = T(r_k) = \sum_{j=0}^k P_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}$$

OpenCV 提供的主要函数为:

对多维直方图进行定义的函数: `CvHistogram`;

对多维直方图进行归一化的函数: `CvNormalize`;

对多维直方图进行灰度均衡的函数: `cvEqualizeHist`。

具体算法为:

(1)对样本图像进行计算,求得其直方图 H ;

(2)对样本图像直方图 H 进行归一化;

(3)对归一化直方图进行求和,即:

$$H'(i) = \sum_{0 \leq j < i} H(j)$$

(4)对样本图像进行变换:

$$dst(x, y) = H'[src(x, y)]$$

得到的结果,就是已经对原始图像亮度进行了归一化,并增强了其对比度的图像。

3.2.3 人脸图像平滑过程

可以采取如下的方式进行人脸图像的平滑操作:(1)局部平均法,(2)中值滤波法,(3)频域平滑法。本文选取中值滤波法,这种方法的优势在于:它对图像的局部进行平均平滑,不但可以降低噪声,也可以尽量减轻图像在边沿处的模糊。中值滤波法是非线性的,其原理为:针对某一滑动窗口的众多灰度值进行排序,用灰度的中值去替换窗口中心像素的灰度。这种方法可以有效地去除图像中孤立的点元素与线元素,维持图像边缘的清晰性,去除二值噪声。

在 OpenCV 中,使用以下功能函数来实现上述算法:

```
void cvSmooth (const CvArr * src, CvArr * dst, int
smoothtype=CV_GAUSSIAN, int param1=3, int param2=
0, double param3=0, double param4=0)
```

其中的部分参数说明:

src 表示输入图像;

dst 表示输出图像;

smoothtype 表示平滑模式;

param1 及 param2 表示平滑操作的参数;

param3 表示高斯参数的标准差。

3.3 金字塔序列化的实现

金字塔序列化是为了克服检测图像中的人脸尺寸不统一的情况,SVM 分类器所扫描的窗口是固定值($19 * 19$),针对多尺度人脸的情况,采取以下方法进行处理:先把一个确定的候选区域转换为灰度图像,并进行直方图均衡操作,然后,依次以某个倍率进行不断的采样操作,本文将此倍率值选为 1.2,从而形成一个金字塔序列。

通过在矩形窗口的滑动来实现候选人脸区域的选取,因为涉及到积分运算,所以对于特征值的求取耗费差别比较小。针对多尺度人脸的检测,不断增加滑动窗口,窗口的初值为支持向量机分类器所支持的 $19 * 19$,增加滑动窗口的步长,设置为 6 像素,滑动方向为向右、向下,并以倍率值 1.2 不断扩大,直至超过样本图像为止。

OpenCV 的相关函数为:

```
void cvPyramidDown (const CvArr * src, CvArr * dst, int
filter=CV_GAUSSIAN 5 * S)
```

其中的部分参数:

src 表示输入图像;

dst 表示输出图像;

filter 表示卷积滤波器类型。

结束语 人脸检测技术是信息技术应用领域的热点,可以方便地进行人脸识别,具有较高的开发和应用潜力。本文针对已有检测技术的不足之处,引入基于特征的检测方法,使用两级分类器,使得检测速度和检测准确性均获得了较大的提高。

(下转第 298 页)

$\max\{R_{kp}, R_{\min}\}$ 。

情况 2: $p > p_{thresh}$ 。此时视频质量已严重受影响,需降低码率,同时又要考虑平滑性,调整:

$$R_{i+1} = \max\{\min\{R_{kp}, (1 - \sqrt{p - p_{thresh}}) \times R_i\}, R_{\min}\}$$

并记下此刻的 R_i 为 R_{con} 。

情况 3: $p = 0$ 。经过拥塞调整后,视频流码率降低,丢包率减小。丢包率又为 0 时,服务器端又将增大编码码率。调整:

$$R_{i+1} = \begin{cases} \frac{R_{con} + R_i}{2} & (R_i < 0.9R_{con}) \\ \min\{R_i + \alpha_2 \times (R_{\max} - R_{\min}), R_{\max}\} & (R_i \geq 0.9R_{con}) \end{cases}$$

其中, $0 < \alpha_2 \leq 1$ 为常量。算法思想是先将码率增长至接近 R_{con} ,若此时丢包率仍然为 0,则保持一定速度缓慢增长。若又出现丢包率大于 0 的情况,则采用前两种情况处理。

3 系统测试结果与分析

本流媒体服务器在嵌入式 Linux 环境下开发,客户端可使用 VLC 播放器来进行测试,在实验室局域网内建立连接,网络状况采用开源软件 nistnet^[10] 进行模拟。本系统中, $R_{\max}$ 为 2Mbps, $R_{\min}$ 为 64kbps, α_1 和 α_2 分别取值 0.1 和 0.01, p_{thresh} 为 0.04。客户端与服务器建立连接后,不断接收到视频流并解码播放,并且服务器端每隔 2 秒收到一个来自客户端的 RR 包。通过 nistnet 模拟网络状况,同时观察客户端视频,发现画面流畅,且画面质量随网络带宽变化及时变化。服务器端通过串口打印出每次改变。取其中一段数据加以分析,如图 3 所示。

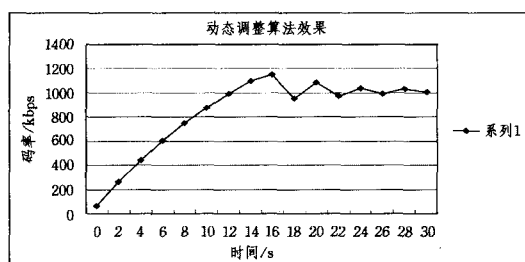


图 3 动态调整算法效果图

从图 3 可以看出,本系统能够实时地监测网络的状况,进而有效调整视频流的码率,提高客户端观看视频的质量。

结束语 本文针对网络视频监控系统,在现有算法 AIMD 和 TFRC 的基础上,提出了适用于实时流媒体的基于 RTCP 反馈信息的自适应动态 QoS 机制,实现了流媒体服务器根据网络状况自动调整编码码率进行传输,使客户端观看到实时流畅的画面。

参考文献

- [1] Tsai Chengfa, Tsai Chunwei. A new approach for solving large traveling salesman problem using evolutionary ant rules[C]// Proc. of International Joint Conference on Neural Network. New Jersey: IEEE Press, 2002: 1540-1545
- [2] 梁丰, 翁格奇. AIMD 拥塞控制算法分析[J]. 计算机工程与应用, 2004, 1: 173-174
- [3] 王光阳, 徐昌彪, 陈前斌. TFRC 与 TCP 流数之比对协议间公平性影响的研究[J]. 计算机应用研究, 2005, 22(11): 16-18
- [4] Handley M, Floyd S, Padhye J, et al. TCP Friendly Rate Control (TFRC); Protocol Specification[S]. IETF RFC 3448. 2003
- [5] Floyd S, Handley M, et al. Equation-based Congestion Control for Unicast Applications; The Extended Version[C]. 2000
- [6] Padhye J, Firoiu V, Towsley D, et al. Modeling TCP Throughput: A Simple Model and Its Empirical Validation[J]. IEEE/ACM Transactions on Networking, 2000, 8(2): 133-145
- [7] Schulzrinne H, Rao A, Lanphier R. Real Time Streaming Protocol[S]. RFC 2326. 1998
- [8] RTP: A Transport Protocol for Real-Time Application[S]. RFC 3550-2003
- [9] Sisalem D, Schulzrinne H, Emanuel F. The Direct Adjustment Algorithm: A TCP-Friendly Adaptation Scheme[R]. Berlin: Fraunhofer, GMD-FOKUS, 1997
- [10] Carson M, Santay D. Nistnet, a Linux-based network emulation tool[J]. Computer Communication Review (ACM SIGCOMM), 2003, 33(3): 111-126
- [11] 肖卫初, 陈伟宏. 监控系统中的多摄像头协同算法[J]. 重庆工学院学报: 自然科学版, 2008, 22(4): 117-123

(上接第 288 页)

- [4] Fisher S P B, Walker A, Wolfart E. Adaptive thresholding[M]. Department of Artificial Intelligence, University of Edinburgh, 1994
- [5] 苏金明, 王永利. MATLAB 图形图像(上)[M]. 北京: 电子工业出版社, 2005
- [6] Weszka J S. A survey of threshold selection techniques[J]. Computer Graphics and Image Processing, 1978, 7(2)
- [7] Haralick R M. Digital step edges from zero crossing of second

directional derivatives[M]. Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, 1987

- [8] Nalwa V S, Binford T O. On detecting edges[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6)
- [9] Otsu N. A Threshold Selection Method from Gray-Level Histograms[J]. Automatica, 1975(11)
- [10] 罗南超, 向昌成. 基于低频边缘特征和能量的多聚焦图像融合方法[J]. 重庆工学院学报: 自然科学版, 2009, 23(8): 168-171

(上接第 295 页)

参考文献

- [1] 叶俊勇. 人脸检测与识别方法研究[D]. 重庆: 重庆大学, 2002
- [2] Sobottka K, Pitas I. A novel method for automatic face segmentation, facial feature extraction and tracking[J]. Signal Processing: Image Communication, 2008, 12(3): 263-281
- [3] Craw I, Ellis H, Lehman J R. Automatic extraction of face-features[J]. Pattern Recognition Letters, 1987, 5(2): 183-187
- [4] Govindaraju V. Locating human faces in photographs[J]. International Journal of Computer Vision, 2006, 19(2): 129-146
- [5] Wang J, Tan T. A new face detection method based on shape information[J]. Pattern Recognition Letters, 21(6/7): 463-471
- [6] 闫文秀, 裴建岗, 孙颖, 等. 基于 Gabor 滤波器和改进 BP 神经网络的人脸检测方法[J]. 重庆工学院学报: 自然科学版, 2009, 23(4): 98-102