

一种基于知识的快速人脸检测方法

姜 军 张桂林

(华中科技大学图像识别与人工智能研究所图像信息处理与智能控制国家教育部开放实验室, 武汉 430074)

摘 要 为了提高人脸检测的速度,提出了一种基于知识的快速人脸检测方法.该方法是采用符合人脸生理结构特征的人脸镶嵌图模型,并在分析了足够多的人脸图象样本的基础上,针对人脸图象的灰度和边缘信息,建立了一种较为完备的知识库;为加快检测速度,该方法采用多级检测步骤,即在粗检测中,根据眼睛的特征,提出并采用一种新的人脸粗检测算法——广义几何投影方法,从而明显地提高了检测速度.实验结果表明,该方法具有较强的鲁棒性,能够很好地解决复杂背景下的多人脸检测问题.由于该方法实现的简单性,并容易由硬件实现,使得检测速度进一步提高,因此该方法在可视电话等领域有着广阔的应用前景.

关键词 人脸检测 人脸识别 基于知识的模式识别 镶嵌图 广义几何投影方法

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2002)01-0006-05

A Knowledge-Based Approach for Fast Human Face Detection

JIANG Jun, ZHANG Gui-lin

(Institute for Pattern Recognition and Artificial Intelligence Key Laboratory of Education Ministry
for Image Processing and Intelligent Control, Wuhan 430074)

Abstract For the speed of detecting human faces in single grayscale scene image is slowly, an approach of knowledge-based fast face detection is proposed. In the method a tri-part mosaic image model is supplied, which is consistent with physical structure features of human face. Basing on this model, knowledge base, including information of edge and gray in images, is established by analyzing fairly enough human face images. Moreover, in order to enhance the speed of detecting face, hierarchical detection processes is adopted. In the process of rough face detection, a new coarse detection method——extended projection which is according with the physical structure features of eyes is presented and applied, which makes it possible to get different scale face candidate region, thus reduces the computation time of face detection greatly. Our experimental results indicate that the approach is high robust, and it is fit to solve the problem of multi-face detection in complex background. Because of the algorithm's simpleness, it can be easily achieved by hardware, and this will shorten the time consuming on face detection. This made it possible to detect faces in real time. So it has wide application perspective at fields of visual telephone, intelligent monitor etc.

Keywords Face detection, Face recognition, Knowledge-based pattern recognition, Mosaic image, Extended projection

0 引 言

人脸是一种极为复杂的、多维的模式,也是最为典型的非刚性模式之一.众所周知,一套完整的人脸识别系统包含人脸检测和人脸鉴别两大模块.国内外不少研究人员都在进行这方面的工作,目前,人脸

的自动鉴别已有一些较为成功的方法,而相对来讲,人脸检测还需要做更多的工作.目前人脸检测方法已有很多,如有模板匹配方法^[1,2]、镶嵌图方法^[3]、几何结构方法^[4]、神经网络方法^[5]等等,其中,镶嵌图方法是目前公认的一种比较好的方法.本文对原有的镶嵌图方法进行了改进,即建立了较为完备的知识库,并提出一种新的粗检测方法,从而显著地提高

了人脸的检测速度。

1 人脸自动检测系统概述

如图 1 所示,本文人脸检测系统是由粗检测、多尺度空间遍历搜索、候选区域归并等 3 个模块构成。检测过程是,首先粗检测部分采用了一种广义的水平几何投影法来检测出人眼的候选位置,经过该次筛选,则大大减少了可能的候选人脸位置数,使最终人脸的快速检测得以实现;接着,在粗检测的基础上,对可能的人脸位置进行多尺度空间的遍历,以解决人脸的尺寸未知问题;最后,将检测出的人脸区域进行归并。这样做的原因有二:其一,使每一个人脸只有一个输出,这样应用镶嵌图模型,同一个人脸就可能在几个邻近尺度或一个尺度的几个邻近位置同时被检测出来,然后再采用求平均的方法将这些区域归并成一个;其二,降低虚警率,只要设定一个阈值,就可以去掉一些“假人脸区域”,因为这些区域周围通常不会连续有几个人脸。

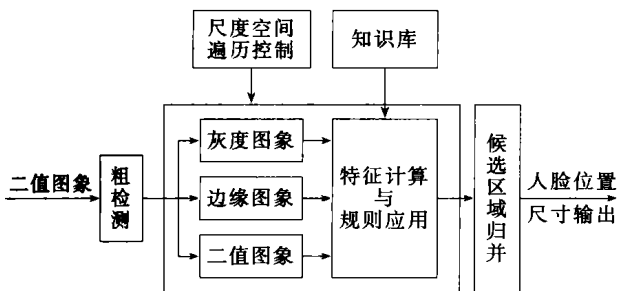


图1 人脸检测系统方框图

2 快速粗检测算法

因为检测的对象是多人脸及不同尺寸的场景图像,所以一个完善的人脸检测算法必须能同时检测尺寸不同的人脸,如果采用变大小窗口的方法来逐行逐列对图像进行扫描,当然可以解决这个问题,但速度将难以忍受,因此为提高检测速度,必须采用分级检测方法。这种分级检测方法一般首先检测人的眼睛,这一步称为粗检测过程;然后再检测整个人脸,如文献[6]就是采用 DoG 算子来先检测出人眼的候选位置,并通过局部搜索方法来进一步减少这些候选人眼的数目;文献[2]则是先将多个人眼的灰度图像作平均,然后进行压缩和放大,以得到不同尺寸大小的人眼模板,最后利用这些模板进行匹配,以解决不同尺寸人脸匹配的问题。

实际上,这些方法都不能完全解决不同人脸尺寸的问题,DoG算子对边缘提取的效果与参数的选择密切相关,且对不同大小的人眼难以兼顾;而模板匹配的方法就更有局限性,因为不可能生成所有尺寸的人眼模板。本文提出了一种新的广义几何投影法,它能够很快地筛选出所有可能的不同尺寸的人眼区域,由于这些可能的人眼区域的像素数目相对于整个图像的像素数目而言,少了很多,因此大大减轻了后续基于知识的多尺度空间遍历搜索的工作量,从而实现人脸的快速检测。

传统的几何投影法是根据图象的水平垂直投影中的波峰或波谷来确定目标的位置,虽然这种方法已得到了广泛的应用,但主要集中于单个目标的使用,在单个人脸的粗定位当中,也有许多人采用了该方法。本文对这种方法作了一些改进,并将其使用在本文的多人脸检测中,已取得了极为有效的候选人眼筛选效果。

本文的投影更加简单,因为它仅仅将图象的二值化边缘作水平“投影”,这种以模式 $f(x, y)$ 为客体,以 X 轴为投影轴的投影公式如下:

$$f(x) = \bigcup_{y=y_{\min}}^{y=y_{\max}} f(x, y)$$

由于上述图象各像素之间是一种求或的关系,而不是传统的求和关系,因此若将双眼的边缘二值图像作这种“或投影”运算,则将产生两条水平短线,且短线的宽度基本与眼睛等宽度(图 2)。

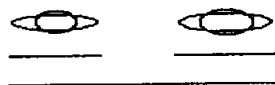


图 2 双眼的水平投影

本文的粗检测过程就采用广义几何投影方法来筛选候选人眼,再通过候选人眼得到候选人脸区域,具体描述如下:

- (1) 对二值化边缘图像作行扫描,在一定高度范围进行“或投影”;
- (2) 对所有的投影线段进行滤波,去掉一部分由干扰所产生的投影线段;
- (3) 检查相邻两条投影线段是否符合如下规则:
 - ① 两条投影线段的长度都在一个范围之内
 - ② 两条投影线段的距离在一定范围之内
 - ③ 两条投影线段的长度差在一定范围之内
- (4) 将所有符合上述条件的线段归并成一条线段(即图 3(b)最下面的一条线),并作为候选位置,

输入下一步的多尺度空间遍历搜索,以检测不同大小的人脸;

图 3(a)和图 4(a)是从一盘 VCD 影碟中得到的两幅图象,用上述方法检测时,首先生成二值图象(图 3(b)和图 4(b)),然后在 4 个象素范围内进行“或投影”,接着再进行滤波,最后依据规则 (3)来生成图 3(c)和图 4(c)的投影图.由于该投影图中只有

很少的线段,而且下一步的多尺度空间遍历搜索只需在这些位置进行,从而节省了大量的时间,实验结果表明,该方法使整个人脸检测速度比原始的逐行逐列、变尺度扫描的速度提高了几十倍乃至上百倍.从图 3(c)和图 4(c)中可以发现,一个人脸有多条投影线,它将增加人脸检测的时间,这个问题则有待进一步解决.



图 3 单幅人脸图象粗检测获得的结果

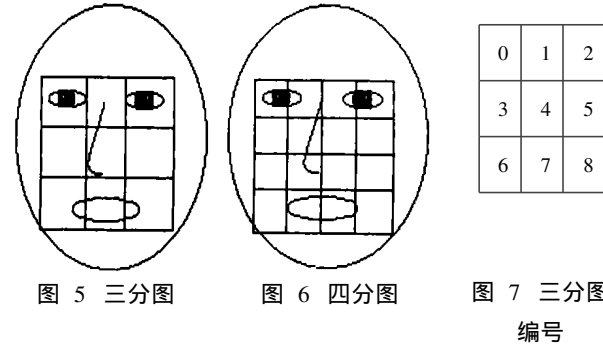


图 4 多幅人脸图象粗检测获得的结果

3 人脸的特征提取及知识库的建立

仅依靠 2D 单帧灰度图象来识别 3D 物体,尤其是识别非刚体,没有知识的加入是非常困难的,而要建立知识库,首先,就必须建立一个合理的模型,本文采用的知识库是基于三分图模型的^[6](图 5),与文献 [3]提出的四分图模型(图 6)相比,三分图的划分更加合理,因为它充分利用了人脸器官的自然分布,即从人脸的生理结构上讲,两眼睛的内眼角之间的距离约为一个眼睛长度,根据图 7 所示的三分图子块划分,从图 5 中可以看出,子块 0 对应于左眼区域,子块 2 对应于右眼区域,子块 4 对应于鼻子区域,子块 7 对应于嘴巴区域,子块 (3, 6, 5, 8)对应于两侧面颊区域,当然,对于不同的脸形,会存在一点偏差,例如,左眼并不完全在子块 0 位置,嘴巴不完

全在子块 7 位置,但这并不会影响规则的作用,因为提取的特征是子块区域的整体特征.



其次,要从模型中抽取恰当的特征统计量.众所周知,人脸是一个很特殊的生物模式,在人脸的照片中,人们可以发现人脸具有明显的灰度特征,如两眼睛的灰度、鼻孔的灰度、嘴巴中间的灰度值都比周围低;同时,人脸还具有很的梯度信息,如眼睛、嘴

巴、鼻子等部位的边缘信息比人脸的其他部位更丰富.可是,文献 [3] 采用的知识库完全是基于子块的灰度平均值,而文献 [6] 则采用完全基于梯度的知识库,由于这两种方法都没有充分利用人脸所具备的丰富信息,因此会造成较高的虚警率.本文则综合了人脸各种有用的信息,提出了一种比较完备的知识库,可使检测虚警率降得很低,同时,人脸的正确检测率几乎不变.根据人脸的灰度和梯度信息,可以得到下面 3 个知识规则库:

- (1) 灰度规则库 ($\bar{g}[i]$ 表示 i 子块平均灰度值)
- 规则 1 $\bar{g}[0]<\bar{g}[3]$,且 $\bar{g}[2]<\bar{g}[5]$
- 规则 2 $\bar{g}[0]<\bar{g}[1]$,且 $\bar{g}[2]<\bar{g}[1]$
- (2) 梯度规则库 ($\bar{T}[i]$ 表示 i 子块平均垂直梯度值)
- 规则 1 $\bar{T}[0]>\bar{T}[3]$,且 $\bar{T}[0]>\bar{T}[6]$
- 规则 2 $\bar{T}[2]>\bar{T}[5]$,且 $\bar{T}[2]>\bar{T}[8]$
- 规则 3 $\bar{T}[0]>\bar{T}[1]$,且 $\bar{T}[2]>\bar{T}[1]$

表 1 知识规则库测试结果

	规则库 (1)	规则库 (2)	规则库 (3)	规则库 (1)和 (2)	规则库 (1)和 (3)	规则库 (2)和 (3)	全部规则库
漏检数	0	0	0	2	2	1	3
漏检率 (%)	0	0	0	3.7	3.7	1.8	5.5
虚警数	161	143	124	27	17	74	8
虚警率 (%)	298	265	230	50	31	137	14.8

4 实验结果

本文构建了一个包含不同质量、不同来源 (主要为 VCD 影碟和照片)、不同背景共计 65 张图象 157 幅人脸的图象库,并利用本文检测方法对该图象库中的图象进行检测,共检测到真实人脸 148 个 (漏检率为 5.7%),检测到虚假人脸 22 个,虚警率为 14.0%.同文献 [3] [6] 所报道的检测结果相比,虚警率、漏检率有了明显的下降 (见表 2).

表 2 实验结果比较

	文献 [3]	文献 [6]	本文
漏检率 (%)	17 (10/60)	13.6 (15/110)	5.7 (9/157)
虚警率 (%)	46 (28/60)	17 (19/110)	14.0 (22/157)

图 8 图 9 图 10 为本实验中的几个样本图象的检测结果,其中,图 8 和图 9 的图象尺寸为 352×288pixels,图 8 检测耗时 6s,图 9 耗时 8s,图 10 尺寸为 1053×764pixels,检测耗时 81s.本实验平台是:PII-333,内存为 64M 的 PC 机.由图 8~

- 规则 4 $\bar{T}[3]<\bar{T}[4]$,且 $\bar{T}[5]<\bar{T}[4]$
- 规则 5 $\bar{T}[6]<\bar{T}[7]$,且 $\bar{T}[8]<\bar{T}[7]$
- (3) 二值规则库 (这里的二值图来源于梯度边缘图, $B[i]$ 表示的是 i 子块中灰度值为 1 的像素数)
- 规则 1 $B[0]>B[1]$,且 $B[2]>B[1]$
- 规则 2 $B[0]>B[3]$,且 $B[2]>B[5]$
- 规则 3 $B[7]>B[3]$,且 $B[7]>B[5]$
- 规则 4 $B[4]>B[3]$,且 $B[4]>B[5]$
- 规则 5 区域 0 和 2 中边缘像素点所占的比例大于 30%.

为了建立恰当的知识库,本文设计了一组实验,即用不同知识库,对 20 幅图象,54 个人脸进行测试,实验结果见表 1.从表 1 可以看到,当同时采用上述 3 个知识库时,虚警率有较大幅度的下降;同时,由于判定人脸的规则更加严格,使漏检率有极小幅度的增加,但这种付出是值得的,因此,本文采用 3 个知识库的并集来作为本次人脸检测的知识库.

图 1 可见,绝大部分人脸均被检出,图 9 图 10 中各检测出一个虚假人脸,可见,检测效果是很好的 (白色方框内为检测出的人脸).



图 8 单幅人脸检测结果



图 9 两幅人脸检测结果



图 10 多人脸图象检测结果

5 总 结

依据人脸的生理结构特征,本文采用了较为合理的分块方案(三分图),在此基础上,又制定了较为完备的知识规则集合,从而拓宽了检测算法的使用范围.实验结果表明,本检测算法适合于复杂背景下的多人、不同尺寸、人脸像基本为正面、表情变化不是很大的人脸检测情况.另外,还提出了一种新的广义几何投影方法,从而大大加快了人脸检测的速度,而且,这种投影方法,非常简单,且适合于用硬件实现.

参 考 文 献

- 1 Tsukamoto A, Lee C W, Tsuji S. Detection and tracking of human face with synthesized templates [A]. In Proc. Asian Conf. on Computer Vision [C], Osaka, Japan Kokichi SUGIHARA, 1993 183~ 186.
- 2 梁路宏,艾海舟,何克忠.基于多模板匹配的单人脸检测[J].中国图象图形学报,1999,4(10): 825~ 830.
- 3 Yang Guang-zheng, Thomas S H. Human face detection in a complex background[J]. Pattern Recognition, 1994, 27(1): 53~ 63.
- 4 Shihong Jeng, Hongyuan Mark Liao, Chinchuan Han *et al.* An efficient approach for facial feature detection using geometrical face model[A]. In Proc. 13th Int. Conf. on Pattern Recognition

[C], Vienna, Austria IEEE Computer Society Press, 1996 426 ~ 430.

- 5 Henry A R, Takeo Kanade. Neural network-based face detection [J]. IEEE Transactions on Pattern analysis and Machine Intelligence, 1998, 20(1): 23~ 38.
- 6 卢春雨,张长水,闻芳等.基于区域特征的快速人脸检测法[J].清华大学学报(自然科学版),1999,39(1): 101~ 105.



姜 军 1971年生,1992年获华中理工大学学士学位,2000年获华中科技大学(原华中理工大学)模式识别与智能系统专业硕士学位,现任华中科技大学讲师.主要从事图象处理、智能识别方面的研究,同时从事与人脸自动识别、检测方面相关的工程开发工作.



张桂林 1944年生,博士生导师,1967年毕业于北京邮电学院无线电通信广播专业,现任华中科技大学图像识别与人工智能研究所教授,图像信息处理与智能控制教育部开放实验室副主任.主要从事图象处理、自动目标识别、人体生物统计特征识别方面的研究.