

# 基于差分图象的人脸检测

艾海舟 王 栓 何克忠

(清华大学计算机科学与技术系 智能技术与系统国家重点实验室, 北京 100084)

**摘要** 提出了两种基于差分图象的人脸检测算法:扫描法和投影法,实验表明它们是准实时的、非常有效的人脸检测算法,可以作为人脸自动输入的有效途径,用于构造诸如安全检查、保安管理等应用视觉系统。

**关键词** 差分图象,人脸检测,人脸识别,目标跟踪,运动检测

## 1 引言

人脸识别的研究最近几年再度受到普遍重视,它与指纹识别、视网膜识别等同属于生物特征识别技术范畴,在诸如安全检查、保安管理等方面有着十分重要的应用价值。从各种文献中可以发现,对于人脸识别算法的研究已经相当深入,由一篇较近的综述文章可见<sup>[1]</sup>,但对于如何检测人脸或输入人脸的研究却不多见。在已发现的文章中,多数的研究是一般环境图象的人脸检测,如基于神经网络的人脸检测方法<sup>[2]</sup>,子空间投影方法<sup>[3]</sup>,镶嵌图方法<sup>[4]</sup>等。这些研究主要是针对静态图象的基于人脸特征的统计与结构分析方法,具有一般性,但缺点是计算量大、对噪音敏感,性能不稳定、难以实时。我们认为在构造实际的应用视觉系统时,为了提高可靠性、加快速度,对环境适当地加以约束是必要的、合理的。比如,对于安全检测、保安管理等场合,可以限定图象采集的背景,利用差分图象方法有效地检测人脸、输入人脸。这是本论文的出发点。可以说,这是一种观念的转变,适合于构造实际的应用系统。

基于差分图象处理的方法<sup>[5~7]</sup>是建立应用视觉系统的常规技术,主要用于印刷电路板的检验、交通及保安监测等。近年来,应用视觉研究在西方发达国家得到普遍重视,这是由于微电子与计算机技术的

高速发展,使得低成本的视觉系统成为可能,以摄像机作为传感器能够超越传统传感器所固有的局限性,有望研究开发出新一代的高技术产品,如交通流量检测器、智能闭路监视系统等。就我国的国情而言,随着经济的高速发展,流动人口激增,城市的交通问题、治安问题等日益严重,急需现代化的监测手段。在这方面,应用视觉研究将起到十分重要的作用。

## 2 人脸检测问题的背景及意义

从研究背景来讲,人脸检测不仅为人脸识别问题所需要,而且它也是人机交流智能化所需要依赖的手段之一。众所周知,人与人之间的交流除了听和说之外,另一个重要的途径就是看。信息的来源主要是头部的运动以及面部表情,当然也包括手势等。为了使计算机与人交流具有看的功能,就必须能够跟踪人头部的运动,进而分析、理解面部表情等。首要的问题是人脸的定位,即用一矩形框标注出人脸在图象中的位置,我们称之为人脸检测问题。我们需要的是一个快速、尽可能实时的人脸检测算法。

如果对环境不加限制,人脸检测问题是十分复杂的,这方面的研究可见文献<sup>[2~4]</sup>,现阶段还难以突破,不能实用。我们认为,在构造实际的应用系统时,可以限定一定的条件如固定背景,这样可以简化

• 受清华大学青年科学基金及国家教委基金资助

收稿日期:1998-03-09;收到修改稿日期:1998-04-24

问题的复杂性,从而更有效地解决问题。只要这种限定条件是合理的、可行的、不苛刻的就行,通常这在工程上是可以做到的。

本论文研究了在背景固定的环境中,对桌前电脑操作者的人脸定位问题。在一般情况下,可以调整摄像机,使得桌前环境中电脑操作者在图象中占中心位置,而且大部分时间内图象中只有操作者,因此主要是一个单人的人脸检测问题。这样的限定问题是有一般性的,同样适用于许多需要人脸自动输入的问题如安全管理等,因为在这些场合下,可以布置一定的环境使得摄像机获取的图象适应需要。方案是预先摄取一幅无人存在的背景图象储存起来(如图1(a)),然后根据摄像头摄取的当前图象与背景图象的差别,确定出人脸可能出现的部分,再根据启

发性的知识检测出人脸的位置(如图2(a))。

### 3 算法描述及实验

算法由预处理和人脸定位两部分组成。预处理过程如图1所示,包括以下几个步骤:

(1)将当前图象与背景图象相减,得到差分图象。

(2)选择适当的阈值,将差分图象二值化,得到二值图象。阈值可以自动选取。

(3)对二值图象作形态学处理:利用圆形结构元素作膨胀变换,抹掉细节,使图象平滑。



图1 对图象进行的预处理

经过预处理后,人脸定位是要得到一个人脸的外接矩形,即确定该矩形的上下左右。我们提出两种方法:扫描法和投影法,分别在下两节说明。我们的实验环境是:PC机(166MHz主频,32M内存),图象技术公司的IC-PCI图象采集卡,CCD摄像头,16mm镜头。软件是用MS-VC++4.2编制的Windows应用程序,论文中的图片大小是 $128 \times 128$ 个像素。

#### 3.1 扫描法

根据预处理后的二值图象的特点,当整个图象中只有一个人比较端正地坐在摄像头前时,可以认

为人脸在二值图象中非零部分的上部。算法由以下几个步骤组成:

(1)对图1中的图象(e)自上而下( $y$ 坐标增加)、从左到右( $x$ 坐标增加)进行扫描,直至找到一个非零点,作为候选人脸的上界。

(2)搜索以确定候选人脸的下左右界,根据:人脸从头顶到下颚的结构特征是宽度先增加后减小,最后到肩部附近时又增加。反映在二值图象(e)中就是, $y$ 向下逐行移动时,每行连续非零线的长度也是先增加后减小然后再增加。在增加到减小的转变点处的非零线的左右 $x$ 坐标就对应着人头部的左右界限。而减小到增加的转变点附近的 $y$ 坐标就对应着

人的下颚。

(3)如果所检测出的候选人脸区域(由头顶,下颚和头部的左右确定)的面积小于某经验值,或者该

区域内的非零像素数与总象素数之比小于某经验值则舍弃之,再转(1)接着向下扫描;否则对结果进行检验、修正,结束。



图 2 扫描法检测结果

图 2 给出了一些扫描法的检测结果。在实验中,对于单人较端正地坐在摄像头前这样的限定条件,允许人在限定范围内运动,我们每次连续运行 5 分钟,共计 1 445 帧  $128 \times 128$  图象或 2 914 帧  $64 \times 64$  图象,检测结果正确率大于 98%,而错误发生在头倾斜接近  $45^\circ$  时,可见扫描法的结果是令人满意的。所谓“较端正”即“限定范围”的含义是指头部左右倾斜小于  $45^\circ$ ,上下倾斜应保证头的上半部分高于肩部,左右手举起的高度不应高于头顶,手臂与头部在图象中不应有接触。这对于许多潜在的应用是可以接受的条件。这种方法的速度对于  $128 \times 128$  是 5 帧/秒,对于  $64 \times 64$  是 10 帧/秒。

### 3.2 投影法

投影法是根据图象在一些方向上的投影的分布特征来进行检测的,这种方法在本质上是一种统计

方法。首先采用纵向投影确定头部的左右两界: *left*, *right*。然后采用由左右两界限定的列的横向投影确定头部的上下两界: *vertex*, *chin*。即统计出图象中每列及每行某区间中非零象素点的数目。设二值图象为  $f(x, y)$ , 大小为  $M \times N$ , 非零象素值为  $T$ , 记第  $i$  列的非零象素点的数目为  $xs[i]$ , 设第  $j$  行在由 *left* 和 *right* 两界限定的区间的非零象素点的数目为  $ys[j]$ , 则:

$$xs[i] = \sum_{j=0}^{M-1} f(i, j) / T, i = 0, 1, \dots, N-1$$

$$ys[j] = \sum_{i=left}^{right-1} f(i, j) / T, j = 0, 1, \dots, M-1$$

对图 1(e)所示的二值图象进行计算,得到的结果如图 3 所示。

由图 3 可见当人较端正地坐在摄像头前时,在 *left* 和 *right* 两处各列非零象素点的数目的变化比

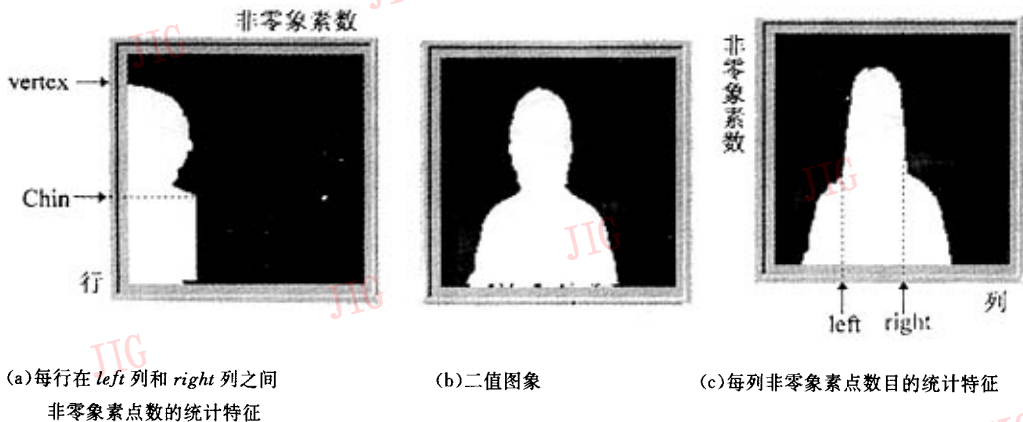


图 3 二值图象及其行、列非零象素数目的统计特征

较大,具体到  $xs[i]$  的值来说,在  $left$  的左边缓慢上升,在  $left$  的右边则急剧上升,而在  $right$  的左边是急剧下降,在  $right$  的右边是缓慢下降。据此就可以初步定出  $left$  和  $right$  的位置。方法是首先计算出  $xs$  在各点的斜率  $xslope$ (为了减小噪声的影响作平均处理):

$$xslope[i] = \frac{1}{3} \sum_{j=1}^3 (xs[i+j] - xs[i-j]) / (2 \cdot j), i=3, 4, \dots, N-4$$

然后确定  $left$  和  $right$ :

$$left = \max_i (\sum_{j=1}^3 (xslope[i+j] - xslope[i-j]) / (2 \cdot j))$$

$$i \text{ 满足 } 0 < xslope[i-1] < 2, xslope[i] > 0, xslope[i+1] - xslope[i-1] > 0, i=6, \dots, N-7$$

$$right = \max_i (\sum_{j=1}^3 (xslope[i+j] - xslope[i-j]) / (2 \cdot j))$$

$$i \text{ 满足 } -2 < xslope[i+1] < 0, xslope[i] < 0, xslope[i+1] - xslope[i-1] > 0, i=left+3, \dots, N-7$$

这样得到的  $left$  和  $right$ , 需要考虑其可信程度,方法是查看它们对应的  $\max$  的数值,如果该值大于某一设定值,则认为它们是可信的,否则不可信。因为如果  $\max$  小于设定值,则表示头部的位置很不端正,所得出的  $left$  和  $right$  的可信度低,这时可根据在  $left$  和  $right$  处的  $xslope$  的值很大这一特征,计算:

$$left = \max_i (\sum_{j=1}^1 (xslope[i+j] - xslope[i-j]) / (2 \cdot j))$$

$$i \text{ 满足 } xslope[i] > 0, xslope[i+1] - xslope[i] > 0, i=4, \dots, N-5$$

$$right = \max_i (\sum_{j=1}^1 xslope[i+j] - xslope[i-j])$$

$$i \text{ 满足 } xslope[i+1] < 0, xslope[i+1] - xslope[i] > 0, i=left+1, \dots, N-5$$

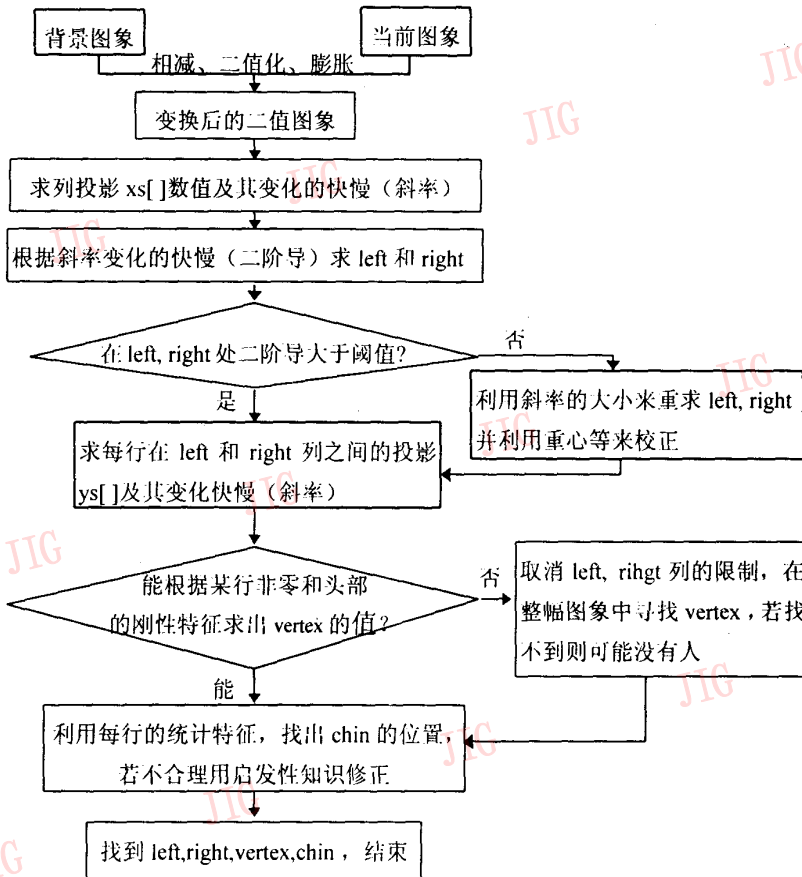


图4 投影法的流程图



(a)背景比较稳定



(b)背景有干扰

图 5 利用统计方法所得出的检测结果

在得到 *left* 和 *right* 之后,我们利用重心位置来判断其可靠性并加以修正。而且在寻找头顶的时候还要再根据头部特征来验证它们的值。

头顶(*vertex*)和下颚(*chin*)的位置需根据每行非零元素的统计特征  $ys[i]$  来确定。当将  $i$  从小到大增加时,若  $ys[k]$  非零,则  $k$  有可能就是头顶的位置。根据人头的刚性特征,如果  $k$  是头顶,并且 *left* 和 *right* 的位置是正确的,那么在矩形 ( $left, k, right, k+(right-left)$ ) 之内非零象素数目应占总象素数的 75% 以上。(由于尚未确定下颚的位置,暂时用  $k+(right-left)$  代替,因为通常头的长度大于宽度)。如果  $k$  点满足上述条件,则认为  $k$  是头顶 *vertex* 的位置。否则继续搜索。如果一直没有找到  $k$ ,有可能 *right* 和 *left* 出错,或者图中没有人。实验中发现 *right* 和 *left* 只会有一个出错,此时撇开它们,直接在整幅图象中寻找 *vertex* 和 *chin*,找出后再利用重心方法校正 *left* 和 *right*,当然此时结果的可信度会下降。

下颚位置的确定是最困难的,当人体相当端正地坐在摄像头前时(如图 3),可以利用  $ys$  值先变化到最小,然后不变,可以将不变点当作 *chin*。但头的前后左右的倾斜对这一特征有很大影响。因此,通常可利用启发式知识,在确定了 *left*、*right* 和 *vertex* 之后,*chin* 可用一定的公式计算出来,即假定人头部的长宽比大致为某个值,根据这个值求出人的下颚位

置,实验表明其效果比较好。

投影法的流程图如图 4 所示。在实验中,对单人的限定条件同扫描法,我们每次连续运行 5 分钟,共计 1 123 帧  $128 \times 128$  的图象或 2 610 帧  $64 \times 64$  的图象,检测结果正确率大于 97%,而错误发生在头倾斜接近  $45^\circ$  时,可见投影法的结果是令人满意的,如图 5(a) 所示。而且与扫描法相比,投影法更具一般性,对于条件的限制可以放松些,如待检测者占中心位置条件下有其它干扰存在时仍然有效,如图 5(b) 所示干扰源处于次要地位时仍能正确检测出人脸。这主要源于其统计特点。这种方法的速度对于  $128 \times 128$  的图象是 4 帧/秒,对于  $64 \times 64$  的图象是 9 帧/秒。

## 4 总 结

本论文从应用视觉的角度,研究了两种快速有效的人脸检测算法:扫描法和投影法。它们是基于差分图象的人脸检测算法,具有快速、可靠的特点,实验表明它们是十分有效的。可以作为人脸自动输入的有效途径,用于构造诸如安全检查、保安管理等应用视觉系统。尽管这两种算法是针对单个人脸的检测问题提出的,不难看出在限定条件下,它们同样适用于多个人脸的检测问题。

## 参考文献

- 1 Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: A survey, Proc. Of The TEEE, 1995, 83(5): 705~740.
- 2 Agui T, Kokubo Y, et al. Extraction of face regions from monochromatic photographs using neural networks, ICARCV'92, 1992.
- 3 Sugiyama Y, Arikawa Y. Facial region tracking and recognition by subspace projection, Proc. Of Virtual Sys. And MultiMedia, Gifu, Japan, 1996:18~20.
- 4 杨光正等. 镶嵌图在人面定位中的应用. 模式识别与人工智能, 1996, 9(3).
- 5 Ching W S. A novel change detection algorithm using adaptive threshold. Image and Vision Computing, 1994, 12(7): 459~463.
- 6 Fathy M, Siyal M Y. An image detection technique based on morphological edge detection and background differencing for real-time traffic analysis. Pattern Recognition Letters 16, 1995: 1321~1330.
- 7 Jain R. Difference and accumulative difference pictures in dynamic scene analysis. Image and Vision Computing, 1984, 2(2): 99~108.



**艾海舟** 1985、1988、1991 年获得清华大学计算机应用专业学士、硕士、博士学位,毕业后留校任教,1994. 9-1996. 8 比利时布鲁塞尔自由大学博士后,现为清华大学计算机系副教授。主要从事计算机视觉、模式识别、移动机器人技术的研究,发表论文 20 余篇。

**王 栓** 1996 年获得清华大学现代应用物理系学士学位,现为清华大学计算机系研究生,研究方向是计算机视觉。

**何克忠** 清华大学计算机系教授。1962 年毕业于清华大学自动控制系,主要从事计算机控制、智能移动机器人技术的研究,先后主持研制了智能移动机器人 THMR- I、THMR- II、THMR- III,多次获部委科技进步奖,出版专著 3 部,发表论文 40 余篇。

## Difference Picture Based Face Detection

Ai Haizhou, Wang Shuan, He Kezhong

(Computer Science and Technology Department, Tsinghua University

State Key Laboratory of Intelligent Technology and Systems, Beijing 100084)

**Abstract** This paper deals with the problem of face detection for applied vision. We introduce two difference picture based algorithms: scanning algorithm and profile projection algorithm. They are effective, reliable, and quasi-real-time algorithms. Experiment results show that they can be promising solutions to the face detection problem for applied vision systems as automatic face input procedure.

**Keywords** Difference picture, Face detection, Face recognition, Target tracking, Motion detection

## 电信业采用泰克彩色打印机

由“联想科技发展有限公司”代理经销的泰克彩色打印机以其优异的性能价格比赢得了电信用户的青睐。近日辽宁省电信局购买一批泰克彩色打印机做为总部及地方分支机构彩色打印输出的专用设备,此前泰克彩打已在银行,金融,石油等行业崭露头角。

邮电通信业的特点是信息,数据处理量大,因而其对外设输出设备性能要求高。泰克打印机因连网能力强,打印速度快,色彩打印标准,输出质量高的特色,满足了电信用户的需求,在其它多种彩色打印机的竞争中脱颖而出,一举成功。辽宁省电信局此次购买的是泰克彩色激光打印机 Phaser 560,此款打印机是目前市场流行的多种彩色激光打印机中性能价格比最高的,在国内,国外多次权威专家评测中屡获殊荣,除此款机型外泰克系列彩色打印机中还有 Phaser 350,360 固体热喷蜡打印机。相信随着经济的发展,广大用户对彩色输出要求的进一步提高,泰克彩色打印机会在彩色世界中大显身手。(汪虹)