

人脸特征的定位与提取

严超 苏光大

(清华大学电子工程系, 北京 100084)

摘要 人脸的特征定位与特征提取的好坏对于人脸图象识别效果有直接的影响. 本文根据人脸图象的灰度特性以及特征的几何形状提出了一种简便有效的特征定位和特征提取的方法. 本文所提出的方法能够比较准确而快速的提取较多的特征, 因而能够适用于在大量人像中进行特征提取和识别.

关键词 特征定位与提取, Hough 变换, 能量函数

1 引言

在图象处理领域, 有很多分类任务, 特别是目标识别, 仍然是没有解决的难题. 人脸识别则是其中之一. 目标识别如果能够取得成功, 可以使许多实用技术从中受益, 比如机器人视觉、军用目标自动跟踪等. 此外, 单就人脸识别来说, 在公安、防盗、来访控制、人工智能、自动摄影、交通流量控制等方面都有实用价值.

人们研究计算机人像识别已经有 20 多年的历史, 最早它和指纹识别一道作为识别罪犯身份的手段. 但是由于人脸的结构远比指纹来得复杂, 它给人们的视觉感受不是简单的通过计算就能得到的, 所以至今人脸识别尚未能取得像指纹识别那样令人满意的结果. 目前国内外研究用于识别人脸的方法层出不穷, 但总的说来不外乎以下 2 种: 一是基于灰度的方法, 二是基于结构的方法. 前者将整个人脸看成一个二维灰度变化的模板, 试图从整体上捕捉和描绘人脸的特征, 发现其中潜在的统计特性, 它所用的方法主要是一些标准的数理统计的技巧^[1,2]. 而后者通过分析人脸的各个组成部分, 将特征分成一个个单独的小模板, 然后对每个模板分别进行特征提取^[3,4]. 由于人脸的表情特征十分丰富, 使得识别过程中能建立起比较大的特征集合, 取得令人满意的识别效果. 本文所采用的就是这一

种方法. 具体过程是先根据人脸图象的灰度特性用直方图匹配的方法确定人脸各组成部分的位置, 然后根据 Hough 变换的原理构造能量函数进行局部特征的提取.

2 特征定位

在特征定位中采用的是一种直方图匹配的方法. 由于人眼的灰度特征与人脸其它部位有明显不同, 我们可以用直方图匹配的方法很容易地得到眼睛的大致位置.

处理的图象采用 $333 \times 480 \times 8$ bit 的灰度图象. 首先通过平滑及二次微分将其转化为一个二值图象, (如图 1 所示) 设点 x, y 处的灰度值为 $g(x, y)$, 二次微分的表达式为:

$$d(x, y) = \sum_{j=-1, 0, 1} g(x+i, y+j) - 9g(x, y)$$
$$\text{if } d(x, y) > D \quad d(x, y) = 255$$
$$\text{else} \quad d(x, y) = 0 \quad (1)$$

其中 D 是预先设定好的门限值. 由二值图象可以大致求出头部的顶点位置 H , 头顶的位置不要求很精确, 因为它对接下来的眼睛定位影响不大. 求到 H 点后, 由 H 点向下移动一定距离, 在这里开始依次向下移动, 每移动一次都抽取一横条进行投影直方图分析, 直到判断出眼睛的大致位置为止. 正确检测出人眼位置时的直方图如图 2 所示:



(a)原图 (b)二次微分后的图象
图1 原图与二次微分后的图象



(a) (b) (c)
图4 特征定位的结果



图2 人眼的正确投影直方图

由于眉毛与眼睛挨得比较近,在检测眼睛位置的时候有可能会受到眉毛的影响而产生偏差。因此在检测出眼睛的大致位置后需要进一步检测原图中眼球的位置。根据眼球处的灰度特性,对眼球的检测可以采用圆面匹配的方法。由于眼球的灰度特征比较明显,对眼球的检测正确率能够达到100%,从而也保证了眼睛位置的准确率。

为了判断鼻尖和嘴唇的位置,在脸轴处纵向抽取一个竖条,从其特有的投影输出推断出鼻、嘴、颚的位置。抽取的结果如图3所示。(脸轴右侧为投影直方图,左侧为投影直方图的微分。)由于我们是对投影直方图的微分进行分析,胡子附近的投影微分值比鼻尖和唇裂位置的微分值明显小了许多,这样就能够排除胡子所带来的干扰,对嘴巴位置检测正确率可达到100%,图4是检测的结果。另外,为了提高识别的速度,还须检测一下上唇的位置。很容易的发现在上唇位置的纵向直方图有一个小的突起,从唇裂位置往上搜索,就能够找到突起的位置,这就是上唇的位置。

得到人脸各部分的位置后,就可以进一步识别了。



(a) (b) (c)
图3 脸轴处的投影直方图

3 脸型特征

脸型特征包括脸颊和下巴。

在人脸的正面像中,脸颊基本上是一条垂直方向上的直线段。(如图5所示)一般检测直线的方法都采用Hough变换,但是由于边缘点比较多,计算比较繁琐,在这儿采用的是直线匹配搜索的方法。由于能大致确定脸颊的位置以及直线参数的范围,这种本来比较繁琐的方法在此显得简单而有效。脸颊直线的极坐标方程为:

$$r = x \cos A + y \sin A \quad (2)$$

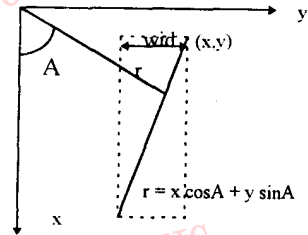


图5 脸颊模板示意图

在前面我们得到了眼睛的准确位置,现在可以由此计算出脸颊的位置。(如图6所示,其中 $\lambda_1 = 68\%$, $\lambda_2 = 40\%$, $\lambda_3 = 39\%$, $\lambda_4 = 69\%$,以上数据及计算方法详细情况请参考文献[5])具体方法是构造一个能量函数 $Eedge(wid, x)$:

$$Eedge(wid, x) = \sum edge[i] / number \quad (3)$$

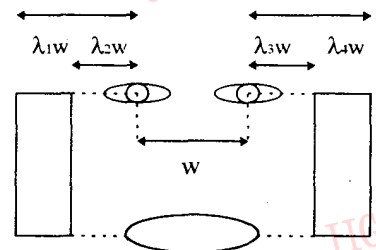


图6 脸颊位置的计算

其中 \sum 对直线上的所有点的灰度值梯度求和, 其中梯度算子采用二次微分算子, $number$ 为直线上点的个数. 不断改变 wid 和 x 的值并求出相应的 $Eedge$ 的值, 当 $Eedge$ 取最小值时候的 wid 、 x 就是我们所要求的脸颊直线的参数. 考虑到有很多人的脸颊接近一种折线形, 所以建议在检测脸颊时分两段检测, 结果将更为精确. 图 7 是采用折线近似检测到的脸颊曲线.



图 7 脸型特征的检测结果

在下巴的检测中采用的是抛物线的模板, (如图 8 所示) 抛物线的方程为:

$$y = y_0 + height \times 4 \times x / length.$$

被检测区域如图 9 确定.

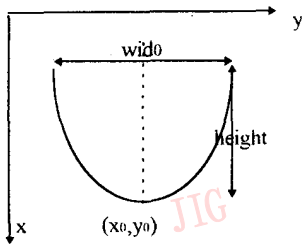


图 8 下巴模板

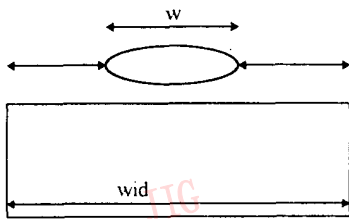


图 9 下巴检测的相关子图

同样构造能量函数

$Eedge(wid, x_0, y_0)$:

$$Eedge(wid, x_0, y_0) = - \sum greyn[i] / number \quad (4)$$

其中 \sum 对抛物线上所有点的灰度值纵向梯度求和, $number$ 为抛物线上点的个数. 不断改变 wid 和 x_0, y_0 的值并求出相应的 $Eedge$ 的值, 当

$Eedge$ 取最小值时相应的 wid 、 x_0 、 y_0 就是所检测的下巴曲线的参数值. 由于在前面求出了 2 个眼球的中心位置, 取其中点位置为 x_0 的位置, 这样可以大大加快检测的速度. 图 7 是检测结果.

4 眼睛特征的检测

人脸中最重要也最难检测的特征是眼睛的特征. 一般人提取眼睛特征都使用一种如图 10 所示的简化模板, 即用两条端点重合的抛物线分别近似上下眼皮, 有时还会在上下眼皮之间加上眼球的模板, 即一个小圆球, 然后用动态模板 (dynamic template) 匹配的方法进行特征提取^[5-7]. 但是我们知道人眼眼皮的特征曲线并不都是象抛物线那么规则, 如果在此采用这种简单模板, 所产生的必然结果就是误差太大以致无法正确描述眼睛的特征. 此外这种方法结果的好坏在很大程度上还取决于眼角位置检测的准确与否. 针对以上问题, 在此采取了一种边缘搜索的算法, 以期获得对人眼更为精细的描述.



图 10 入眼的简化模型

具体的搜索方法如下: 首先, 用二次微分算子计算出眼位区域的灰度值梯度, 这个过程包括灰度平滑消噪、梯度计算、二值化再次消噪, 这样得到的就是眼位图的二值图象. 这里的平滑采用平均值的计算方法. 图 11 是二次微分后的结果. (threshold = 8) 其次, 由二次微分后的图象搜索出上眼皮的特征曲线. 这一过程是由以下几个步骤来完成的. (以左眼为例)

(1) 确定起始点, 以眼位区域内最左边的黑点为起始点. 但是由于图象不可避免的会带来噪声, 再加上眼睛周围并不光滑, 经常会出现检测错误, 需要对起始点进行重复判定. 这里有 2 个标准: 第一, 这点必须是边界点. 第二, 以这点为起始点搜索出来的边界, 其像素点的个数不可能太少. 有这 2 个标准就可以保证得到正确的起始点.

(2) 由起始点确定下一点. 由于图象质量的好坏不同, 有时微分得到的图象中间会出现不连续的现象, 所以不能用简单的链



图 11 对人眼二次微分后的结果

码搜索,而必须根据上眼皮的特征重新设置链码。图12采用的是一种链码方式。其中'o'表示起点,搜索顺序为:00→01→02→03→10→...→16→20→...→29,可以得到一条完整的上眼皮曲线。

			20
		10	21
	00	11	22
0	01	12	23
03	02	13	24
16	15	14	25
29	28	27	26

图12 链码的设置

(3) 依次搜索下去直至被搜索范围内无边界点。

(4) 将得到的边界线进行优化。优化这一步是必不可缺的,它不仅仅是为了使得到的曲线变得光滑,更主要是为进一步将眼睛分类做一下预处理。图13是检测的结果。我们可以看到,这样提取出来的曲线与眼睛的特征吻合得非常好,它明显要比那些用椭圆作为模板检测出来的曲线精确得多,有效得多。这对于眼睛的分类识别是非常有用的。



图13 上眼皮检测结果

得到上眼皮的曲线后,就可以将两个端点看成眼角,再由眼球的位置可以得到下眼皮的曲线。

5 嘴巴特征的检测

人脸还有一个比较重要的特征是嘴巴的特征。嘴巴可以看成是由上嘴唇,唇裂和下嘴唇三部分组成。由于拍照的角度不同,嘴唇附近有时会产生阴影,给检测带来困难。因此在检测中必须先检测出嘴角的位置,然后在进一步检测上、中、下唇的形状参数。

同样我们对图象进行平滑、二次微分、二值化,得到嘴巴及其附近区域中所有梯度较大的点。而唇裂(含嘴角)则是由连续的梯度较大的点组成的。我们可以根据这一特点得到嘴角的位置。具体的搜索方法可以按照前面用过的链码方式。图14显示了

3幅正确检测的图象。

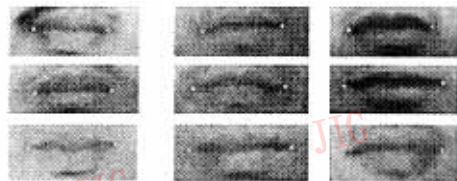


图14 嘴角检测结果

有了嘴角的位置,再进行动态模板(dynamic template)匹配,上唇及唇裂的形状就能检测得相当准确。上唇及唇裂明显是抛物线的形状,而且抛物线的顶点垂直方向上大致的位置在前面特征定位时就已经检测出来了,水平方向的位置则由人脸的中轴线确定,这2个位置的确定使模板匹配的速度提高了不少。

检测时需要选择合适的能量函数,从图中可以看到上嘴唇有2个重要的灰度特征:

(1) 正向梯度较大。正向是这么定义的,即用上面的灰度值减去下面的灰度值,若得到负值就定为零。

(2) 灰度值较大。这样我们就可以构造这样的能量函数:

$$E_{edge} = - \sum grad[i] \times (int)(grey[i]/num1)$$

对每一点都计算纵向梯度,并用灰度值进行加权,(对所有抛物线上的点求和,不断改变顶点的位置同时计算相应的 E_{edge} ,当 E_{edge} 取得最小值时,顶点就确定了,抛物线也就随之确定了。中唇有相反的灰度特征:反向梯度大;灰度值小。相应的能量函数构造如下:

$$E_{edge} = - \sum grad[i] \times (int)(num2/grey[i])$$

同样当 E_{edge} 取最小值时,抛物线匹配得最好。

下唇的特征相对而言比较模糊,这给检测带来了很大的难度。后来我们观察了几百幅的人像,发现基本上所有人像下唇的正下方都有比较明显的阴影。于是作了如下定义:从唇裂正中间开始向下搜索,当搜索到灰度值明显减小的位置时,将这一点定义为下唇的顶点。至于怎样才算灰度值明显减小,可以根据实际检测结果的好坏加以调整。图15是几幅检测后的结果。

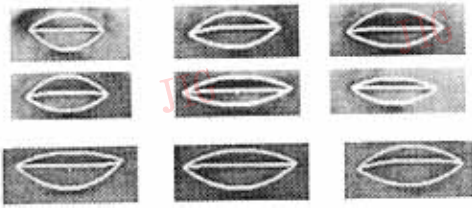


图 15 基于嘴角位置的嘴巴检测的结果

6 小 结

人脸的结构非常复杂,人脸特征也是千变万化的,再加上拍摄角度的不同,因此要准确地提取每一幅图象的特征是相当困难的。我们这里提出来的方法对绝大多数图象的特征都能准确提取,这需要图象拍摄时的角度要正,不能有太大的偏差,否则对脸颊及下唇的检测会产生影响。我们曾经对随机抽取的 100 幅人脸正面像进行过处理,其中眼睛、嘴巴、下巴的特征检测得近乎完美,只有脸颊的正确率较低,大概只有 80%左右,主要是受到耳朵的干扰。

另外还有一点必须明确指出的是,虽然本文所做的处理是针对某一固定大小的图象(即 $333 \times$

480)进行的,但是方法本身并不受图象大小的约束,即使人像有一定角度的旋转,对检测结果也不会有影响。

参 考 文 献

- 1 Turk M, Pentland A. Face processing: models for recognition. SPIE Intell. Robots Comput. Vision VIII Algorithms Techniq. 1192, 1989, 22~32.
- 2 Sirobich L, Kirby M. Low-dimensional procedure for the characterization of human faces, Optic. Soc. Am. J. 3, 519~524.
- 3 Shackleton M A, Welsh W J. Classification of facial features for recognition. Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVIP-91), 1991: 573~579.
- 4 Craw I, Ellis H Lishman J R, Automatic extraction of face features, Pattern Recognition Lett. 5, 1987: 183~187.
- 5 Xiaobo Li, Nicholas Roeder. Face Contour Extraction From Front-View Images, Pattern Recognition, 1995, 28(8): 1167~1179.
- 6 Kin-Man Lam, Hong Yan, Locating and extracting the eye in human face images. Pattern Recognition, 1996, 29(5): 771~779.
- 7 Xie X, Sudhakar R Zhuang H, On improving eye feature extraction using deformable templates. Patten Recogtion, 1994, 27(6): 791~799.



严超 95年本科毕业于清华大学电子工程系,被推荐免试直读该系图象专业研究生。本科期间完成了线阵B超的单片机系统的设计和调试工作。在研究生期间开始进行人象识别方面的研究。



苏光大 毕业于清华大学电子系无线电技术专业,留校任教,副教授。讲授《图象处理系统》课程,长期从事图象处理系统的科研,其多项科技成果获部级科技成果奖,主持研制成功的计算机人象组合系统已广泛用于公安刑侦部门,公开发表 10 余篇关于图象处理系统方面的学术论文。

Facial Feature Location and Extraction From Front-View Images

Yan Chao , Su Guangda

(Department of Electronic Engineering, Tsinghua Univ., Beijing 100084)

Abstract Facial feature detection is crucial for visual recognition of human faces. This paper reports the result on a study in detecting face contours from front-view ID-type pictures. A simple and effective way is proposed to detect the eye and mouth positions. Based on the relevant subimage defined in Ref. 5, the location and curvature of the chin line, together with the cheek lines can be detected and the shape of the face was characterized. Especially in this paper, the eye and mouth contours extractions become easier and more accurate.

Keywords Facial feature location and extraction, Hough transformation, Energy function

M4 Data 公司最新推出 M490F 磁带机

M4 Data 有限公司是一个有 25 年历史的磁带机生产厂商,总部设在英国,在美国、德国、印度、中国设有分支机构。最近它推出的 Twin Peak(双峰) M490F 使主机类后备的可靠性和性能在合理价位下满足了联网的多用户环境。本产品被誉为世界上最小的高性能 18/36 轨 IBM ANSI/ECMA 兼容的磁带驱动器,它代表着 M4 Data 又一个世界之最佳的出现。

采用 8"格式的 M490F 可以直接放置在桌上使用,其驱动器的磁头与磁带间采用的真空装置,可使磁带与磁头始终处于非接触状态,同时全对称的磁带路径最大程度地确保磁带运行时的平稳。

M490F 独有的 DUAL FIPS CHANNEL 接口,或 SCSI-2 接口(两者可选)可与用户现有的 AS400 或 RS/6000 IOP 连接。(# 2323-bus&tag 或 # 6501-SCSI)

M490F 加上自动装载机构,可构成具有 15 盘盒带的带库-M490L, M490L 比 IBM 3490E-F11 可多装 5 盘盒带,并可完全兼容,无需更换驱动软

件。

M490F 完成了与多种石油仪器的联机测试,并可支持 18 轨读/写。

M490F 的连续读/写速度高达 6 兆字节/秒,且具有一个可选用的 SCSI 接口,它所提供的速率为 10 兆字节/秒或 20 兆字节/秒(宽的 SCSI)。将可选式 15 盘盒带自动装载的库系统加到 M490F 上,可实现快速高容量(带数据压缩时大于 30 千兆字节),是摆脱人工管理后备系统的理想选择。它利用 36 轨技术以台式装配在合理价位上为用户提供主机级的可靠性。

后备系统的速度与高容量(2.4G 字节)盒带相结合,用户可以从使用单一的驱动器开始,通过简单自动装载库系统使之升级为全自动化的后备系统。其原始性投资可使用户直接扩充规模。

许多年来, M4 Data 在设计 9 轨磁带驱动器方面一直处于世界领先地位。现有产品中的 9914(四密度)和 9906(双密度)9 轨磁带驱动器,均由 M4 Data 有限公司设计,在英国制造并销往世界各地。