

允许姿态变化的快速人脸特征检测*

彭振云 游素亚 徐光祐

(清华大学计算机科学与工程系, 北京 100084)

摘要 本文提出一种在允许姿态显著变化的人脸图象上快速检测眼睛和嘴角的方法。该方法仅仅合理地假设人脸图象可以通过阈值分割将头发和眼睛与脸部明显分开。首先采用Hough变换并利用人脸的相似性和对称性约束确定两个瞳孔的位置, 然后根据人体测量关系粗略估计嘴部区域, 最后利用积分投影方法得到嘴角的精确位置。我们用此方法对250多幅具有不同光照条件和头部姿态的图象进行了实验, 结果表明其在速度和精度两方面都具有良好的性能。

关键词 人脸特征检测, 姿态可变, Hough变换, 积分投影

1 引言

在过去二十多年中, 人脸自动识别吸引了众多研究者, 并在商业和执法部门得到了广泛的应用^[1,2]。现有的人脸识别方法基本上分成两类: 基于特征和基于图象。前者将人脸用脸部结构特征(如眼睛和嘴)描述。一幅未知面孔的特征描述若和数据库中的某一幅样本图象相同或相似, 则认为被成功识别。基于图象的方法将人脸描述成灰度阵列, 通常用模板匹配策略进行识别。

尽管上述两种方法在很多方面不同, 但脸部特征抽取对二者来说都是非常关键和重要的一步。对于基于特征的方法来说, 这不言自明。对于后者, 往往要用脸部特征来估计姿态或引导匹配过程。正因为如此, 这一问题长期以来得到了广泛的研究。

基于变形匹配的脸部特征抽取是一个很成功的方法。该方法由Yille等人^[3]提出, Xie等^[4]对其进行了改进, 并被很多人^[5,6]采用。其主要思路是, 定义一个可变形的参数模板和一个能量函数来描述特征, 通过一个非线性最优化算法求得使能量函数最小的参数模板, 此模板即被认为是对所求特征的描述。这种方法稳定可靠, 而且与姿态和光照条件无关, 但有

两个主要问题。一是能量函数的系数很难适应一般情况, 二是计算量很大。

Reisfeld等^[7]提出了另一个基于广义对称性的方法。它可以可靠地检测出眼睛和嘴, 而且对图象的唯一要求是人脸占15%~60%的幅面。遗憾的是由于该方法没有利用关于脸部位置的任何先验知识, 因此计算量依然很大。其他典型的方法分别基于Hough变换^[8]、KL变换^[9]和SVD分解^[10]等。

现有脸部特征检测方法大都需要较大的计算量(许多文献都称在工作站上需要几分钟时间)和一组与特定环境有关的实验参数, 而且往往对头部姿态、脸部大小或光照条件有一定要求。我们试图在这三方面进行改进。本文所述方法基于一个常常被忽视的简单事实: 对于几乎所有的人脸图象, 头发和眼睛(至少是瞳孔)都很容易和脸颊明显分开。因此, 任何人脸图象都能在被很好地分割后将眼睛可靠地检测出来, 进而得到其他特征。该方法只假定: ①图象中只有一个人脸, 且占20%~60%的幅面; ②双眼和两个嘴角可见; ③头部在图象平面上的倾斜不超过50度。其主要特点是: ①与头部姿态、脸部大小和光照条件无关; ②不依赖于实验参数; ③速度快(在486/66微机上每幅图象约需3秒钟), 正确率高(85%以上)

* 张永越先生为本研究提供了测试样图和图象处理源程序, 谨致谢忱。

收稿日期: 1996-03。

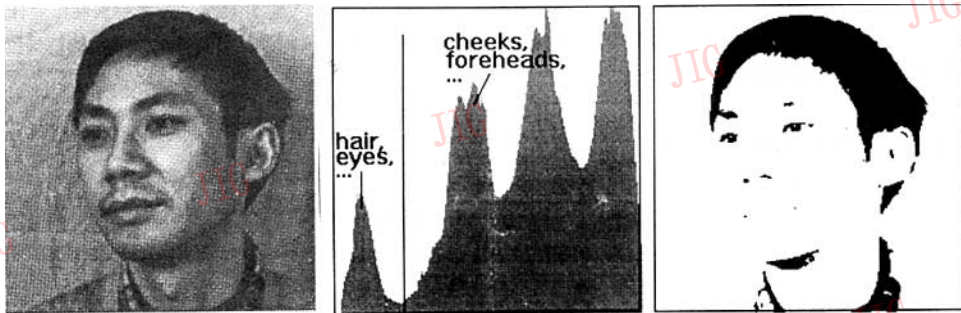
2 阈值分割

首先对原始图象 $I(x, y)$ 进行阈值分割, 以使头发和眼睛能与脸颊明显分开。虽然一般的阈值分割很难, 但在前述假设下此问题变得很简单。在此情况下, 人脸图象的典型直方图如图 1(b) 所示。其中两个明显的峰值分别对应于脸的暗部(头发、眼睛、眉毛、嘴等)和亮部(脸颊、额头等)。而且, 由于背景一般比脸的暗部要亮, 所以第一个峰值对应于脸的暗部。这样我们只须找到前两个峰值, 然后在其中

中间取一个灰度级作为阈值即可。为了避免光照不均匀, 在检测峰值前先对直方图 $h(x)$ 进行 Gaussian 滤波:

$$H(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^{N-1} h(u) \exp\left[-\frac{(x-u)^2}{2\sigma^2}\right] du \quad (1)$$

其中, N 为灰度级数, 在我们的实验中为 256。因为有多组可选阈值, 所以二值化过程非常可靠。结果图象 $B(x, y)$ 是一个 0-1 阵列, 0 和 1 分别代表脸的暗部和亮部。



(a) 原始图象 $I(x, y)$

(b) $I(x, y)$ 的直方图

(c) 二值化图象 $B(x, y)$

图 1 原始图象二值化

Fig. 1 Thresholding of original images

3 眼睛定位

在定位中, 眼睛用瞳孔代表, 后者又用圆来描述。圆半径的范围 $R_{\min} \sim R_{\max}$ 根据对脸的大小所作的假设来估计。在我们的实验中, 图象幅面为 256×256 , 半径范围为 $3 \sim 5$ 。首先用传统的 Hough 变换在 $B(x, y)$ 中检测所有合法的圆。为了避免使用边缘阈值, 我们对边缘点的定义是, 其值为 0, 且其 8 个相邻象素中有一个的值为 1。如果一个边缘点 (x, y) 满足式(1), 则圆 C_i 的“票数”加 1。

$$|\sqrt{(x - A_i)^2 + (y - B_i)^2} - R_i| \leq \delta \quad (2)$$

其中, $R_i \in [R_{\min}, R_{\max}]$, (A_i, B_i) 和 R_i 分别是圆 C_i 的中心和半径, $\delta = 0.2R_i$ 为容许误差。圆 C_i 的“得分”用文献[3]中所建议的公式求得:

$$S_i = N_i \frac{N_i}{R_i} \quad (3)$$

其中, N_i 是圆 C_i 的“票数”。

得到所有合格的圆后, 按“得分”从高到低的顺

序对其进行扫描。

$$\sum_{(x, y) \in Area(C_i)} (1 - B(x, y)) \geq 0.5\pi R_i^2 \quad (4)$$

$$\sum_{x=A_i-R_i}^{A_i+R_i} \sum_{y=B_i+2R_i}^{B_i+10R_i} (1 - B(x, y)) < 3R_i \quad (5)$$

$$\sum_{x=A_i-R_i}^{A_i+R_i} \sum_{y=B_i-4R_i}^{B_i-2R_i} (1 - B(x, y)) < 10R_i \quad (6)$$

式(4)说明瞳孔必须足够暗; 式(5)表明眼睛应在脸颊之上; 式(6)则意味着眼睛应在额头下面。

前 20 个满足式(4)到式(6)的圆被选作候选瞳孔, 表示为 C_1, C_2, \dots, C_{20} 。

设 C_i, C_j, \dots, C_{20} 是一个完全图 G 的节点。对图中 C_i 和 C_j 之间的边定义一个利润函数 $BF(i, j)$ 如下:

$$BF(i, j) = (Sim(i, j) + Sym(i, j)) * DC(i, j) * AC(i, j) \quad (7)$$

其中,

$$Sim(i, j) = \frac{E(C_i C_j) - E(C_i)E(C_j)}{\sigma(C_i)\sigma(C_j)} \quad (8)$$

$$Sym(i, j) = \frac{E(NC_i NC_j) - E(NC_i)E(NC_j)}{\sigma(NC_i)\sigma(NC_j)} \quad (9)$$

$$DC(i, j) = \text{sgn}(10 * (R_i + R_j) - D(i, j)) * \text{sgn}(D(i, j) - 3 * (R_i + R_j)) \quad (10)$$

$$AC(i, j) = \text{sgn}(0.25\pi - \text{abs}(\text{tg}^{-1}((B_i - B_j)/(A_i - A_j)))) \quad (11)$$

$E(R)$ 和 $\sigma(R)$ 分别是区域 R 内灰度分布的期望值和标准差。 NC_i 和 NC_j 是关于 C_i 和 C_j 连线中点对称的两个区域。 $Sim(i, j)$ 和 $Sym(i, j)$ 分别是 C_i 和 C_j 之间的相似性和对称性度量。 $DC(i, j)$ 和 $AC(i, j)$ 是对两眼间距离和角度的约束。通过在完全图 G 中搜索使 $BF(l, r)$ 最大的边可以确定瞳孔对 (C_l, C_r) 。

最后,通过在 C_l 和 C_r 的邻域内使利润函数最大来对眼睛的位置进行微调。

4 嘴角定位

嘴的位置由嘴角确定。给定眼睛位置后,此问题在容许有1~2个像素的误差时较容易解决。

首先,根据眼睛的位置和人体测量标准可以初略限定嘴部区域。如图2所示,假设 C_l 和 C_r 是两个瞳孔,则嘴可以认为位于平行四边形 $ABCD$ 内。



图 2 限定嘴部区域
Fig. 2 Mouth region binding

文献[11]中介绍的积分投影对于定位嘴角很有效。在前述情况下, $I(x, y)$ 在 $ABCD$ 内的水平投影可定义为:

$$H(y) = \sum_{x=AD(y)}^{BC(y)} I(x, y) \quad (12)$$

类似地,可定义垂直投影为:

$$V(x) = \sum_{y=AB(x)}^{DC(x)} I(x, y) \quad (13)$$

其中, $y=AB(x)$ 和 $y=DC(x)$ 分别是 AB 和 DC 的

直线方程, $x=BC(y)$ 和 $x=AD(y)$ 分别是 BC 和 AD 的直线方程。 $H(y)$ 在源图象上计算, $V(x)$ 则在垂直梯度图和源图象的乘积图上计算。

嘴的垂直位置通过先检测 $H(y)$ 的局部谷值,再根据其周围的灰度分布进行检验得到。假设得到的水平线是 L ,则 $V(x)$ 只须在 $A_1B_1C_1D_1$ 内计算。最后,通过寻找 $V(x)$ 的两个跳变点可以确定两个嘴角的水平位置。

5 实验和讨论

我们对由86个人拍摄的252幅具有各种姿态和光照条件的图象进行了实验。实验统计数据列在表1中,部分结果如图3所示。

令人惊讶的是,当将半径范围固定为 $R_{\min} = R_{\max} = 3 * (\text{image-size}/256)$ 时,定位结果几乎不受影响,而计算量大大降低。这进一步证明此方法对关于原始图象先验知识的依赖性很小。

表 1 实验统计

Table 1 Experimental statistics

结果	正确	全错	只有嘴错
数量	217	10	25
比例(%)	86.0	4.0	10.0

6 结论和下一步工作

前面介绍了一个在对背景稍作限制的图象上定位人脸特征的方案。此方案已在我们最近开发的一套人脸识别系统^[12]中采用。该方法原理简单,有效实用,更突出的是不依赖于实验参数,因而适用的范围相当广泛。不过,和所有其他串行算法一样,该方法在任何中间结果出错时都会失败,而且不能并行实现。后者在需要精确描述特征时显得很关键。

除在上述方面进行改进外,我们还打算在下列两方面对其进行扩展:①取消对背景的限制,以适应更复杂的场景;②根据位置进一步抽取特征的细节,如眼帘和嘴唇,最后达到描述整个人脸的目的。



图 3 部分实验结果

Fig. 3 Some experimental results

和刚体不同,人脸是变形体,更适合用弹性模型^[13]来描述。因此,任何基于刚体特性的特征抽取方法都很难达到满意的效果。其次,人脸识别被认为是人类视觉中独特的过程,因此对心理生理学结论是很有帮助的。遗憾的是我们对人类视觉知之甚少。可以预见,人类视觉和变形体两方面的成就将有助于找到抽取和描述人脸特征的最终答案。

参考文献

- 1 Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: a survey. Proc. of the IEEE, 1995, 83(5): 705~740.
- 2 Samal A, Iyengar P A. Automatic recognition and analysis of human faces and facial expressions: a survey. Pattern Recognition, 1992, 25(1): 65~77.
- 3 Yuille A L, Cohen D S, Halinan P W. Feature extraction from faces using deformable templates. Proc. IEEE Computer Soc. Conf. on computer Vision and Patt Recog, 1989: 104~109.
- 4 Xie X, Sudhakar R, Zhuang H. On improving eye feature extraction using deformable templates. Pattern Recognition, 1994, 27(6): 791~799.
- 5 Chow G, Li X. Towards a system for automatic facial feature detection. Patter Recognition, 1993, 26(12): 1739~1755.
- 6 Huang C L, Chen C W. Human facial feature extraction for face interpretation and recognition, Pattern Recognition, 1992, 25(12): 1435~1444.
- 7 Reisfeld D, Yeshuran Y. Robust detection of facial features by generalized symmetry. Proc. 11th Int. Conf. on Patt. Recog., 1992: 117~120.
- 8 Nixon M. Eye spacing measurements for facial recognition. SPIE Proc. 575, Applications of Digital Image Processing VIII, 1985: 279~285.
- 9 Kirby M, Sirovich L. Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Trans. Patt. Anal. and Mach. Intell., 1990, 12: 103~108.
- 10 Hong Z. Algebraic feature extraction of image for recognition. Pattern Recognition, 1991, 24: 211~219.
- 11 Brunelli R, Poggio T. Face recognition: features versus templates. IEEE Trans. Patt. Anal. and Mach. Intell. Oct., 1993, 15(10): 1042~1052.
- 12 Zhang Y, You S, Xu G. A view-based automatic face recognition system. Proc. 1st Int. Conf. on Multimodal Interface, 1996.
- 13 Terzopoulos D, et al. Elastically deformable models. Computer Graphics, 1987, 21(4): 205~214.



彭振云,1984 年和 1988 年于北京清华大学先后获工学学士和工学硕士学位,1993 年起任高级工程师,现在清华大学计算机系攻读博士学位。主要研究领域包括计算机视觉、模式识别和计算机图形处理。

A Fast Method for Detecting Facial Features under Varied Poses

Peng Zhenyun, You Suyu, Xu Guangyou

(Computer Dept., Tsinghua University, Beijing 100084)

Abstract This paper presents a fast method for detecting the eyes and the mouth corners on a facial image with significantly varied head poses. The described scheme only reasonably assumes that the facial image should be thresholded well enough that the hair and eyes can be distinctly separated from the cheeks. It first detects the two irises by using the Hough Transform together with the constraints on the facial symmetry and similarity then roughly binds the mouth region according to the anthropometric standards, and finally utilizes the integral projections to locate the two mouth corners. This method has been tested on more than 250 face pictures with different illuminating conditions and varied head poses, and showed a quite satisfactory performance both on speed and accuracy.

Keywords Facial feature detection, Varied poses, Hough Transform, Integral projection