

# 使用关键点信息改进弹性匹配人脸识别算法

丁 嵘, 苏光大, 林行刚  
(清华大学电子工程系, 北京 100084)

**摘要:** 本文提出了一种增加经典弹性匹配算法对强背景干扰鲁棒性的算法. 人脸的自动识别有着广泛的应用前景, 弹性匹配是其中的一种重要算法, 并有着相对于其他方法更高的鲁棒性, 但其并不能有效地防止强背景干扰的影响. 本文的主要思路是将局部信息和整体信息相结合, 利用没有受到背景干扰影响的局部关键点的位置信息来对基于整体信息的经典弹性匹配方法进行改进. 实验结果表明, 本文提出的算法有效地提高了弹性匹配识别算法的鲁棒性.

**关键词:** 人脸识别; 弹性匹配; 关键点

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112 (2002) 09-1292-03

## Using Key Points to Improve Elastic Matching in Face Recognition

DING Rong, SU Guang-da, LIN Xing-gang

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

**Abstract:** An algorithm to improve the robustness of elastic matching in face recognition is proposed when it is strongly disturbed by the background. Automatic human face recognition has various applications. Elastic matching is an important face recognition algorithm, and it has higher robustness compared to other algorithms. But it also cannot deal with the disturbance from background effectively. The basic point of our algorithm is combining local information and global information together and using the local information undisturbed by the background to improve elastic matching algorithm. Experimental results demonstrate that the proposed method has improved the robustness of elastic matching.

**Key words:** face recognition; elastic matching; key points

### 1 引言

计算机人脸识别是由计算机分析人脸图像, 从人脸图像中提取出有效的识别信息, 用来“辨认”身份的一门技术. 人脸识别技术有着很广泛的应用背景, 可以应用于公安系统的罪犯身份识别、驾驶执照及护照等证件照片与实际持证人的核对、银行及海关的监控系统及一些保密单位的自动门卫系统等.

人脸识别技术的研究开始于六十年代末七十年代初, 早期的识别方法是基于部件的<sup>[1]</sup>, 其主要思想是找出人脸中的各个部件(如眼睛、嘴巴、下巴等), 然后利用各部件的参数和各部件关键点间的相对距离进行人脸识别. 这类方法比较简单, 但很容易丢失人脸的灰度和纹理信息, 从而对视角、表情等变化的鲁棒性比较差. 随着计算机软硬件能力的提高, 90年代以来, 更多的研究集中在基于整体的方法, 如特征脸(eigenface)方法<sup>[2]</sup>、SVD分解的方法<sup>[3]</sup>、人脸等密度线分析匹配方法<sup>[4]</sup>、弹性匹配(elastic matching)方法<sup>[5]</sup>、隐马尔可夫模型(Hidden Markov Model)方法<sup>[6]</sup>以及神经网络的方法等<sup>[7]</sup>. 这类方法更多地考虑了模式的整体属性, 因此保留了更多的信

息. 其中的弹性匹配方法具有较高的识别率, 而且对一些背景、视角、表情的小的变化鲁棒性比较强, 因此受到了广泛的重视.

通常的研究是对裸脸进行的, 但在实际应用(如人脸识别用于公安系统)中, 往往会出现所采集的人脸周围背景比较复杂的情况, 而如果背景干扰比较强则会较大地影响弹性匹配的性能. 为此本文提出了一种利用关键点的位置信息对弹性匹配进行约束, 以减少背景干扰影响的算法. 下面首先简单介绍经典的弹性匹配算法, 然后介绍本文设计的使用关键点位置信息的改进算法, 最后是该算法的实验结果、性能分析及总结展望.

### 2 经典弹性匹配算法的基本原理<sup>[5]</sup>

弹性匹配法在二维的空间中定义了一种对于通常的人脸变形具有一定的不变性的距离, 并采用属性拓扑图来代表人脸(图1采用的是规则的二维网格图, 实际中也可以采用其他形状的网格), 拓扑图的任一顶点均包含一特征向量, 用来记录人脸在该顶点位置附近的信息, 如图2所示.

收稿日期: 2001-06-18; 修回日期: 2002-01-29

基金项目: “人脸识别查询技术”国家重点攻关项目(No. 2001BA801B07)



图 1 定义在人脸上的二维拓扑图

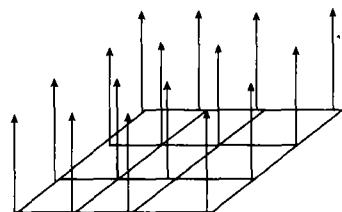


图 2 表达人脸特征的二维向量场

假设  $S_1$  是定义在已知库中某人脸像上的人脸模板的二维网格,则该人脸像上网格上节点  $i$  附近的信息可以用特征向量  $C_i$  来代表。 $C_i$  可以有多种选择,最常用的是 Gabor 特征:定义  $G = [g_1, g_2, g_3, \dots, g_m]^T$  的成员是中心频率不同,带宽不同,方向不同的二维 Gabor 滤波器,那么  $C_i$  就是  $G$  和人脸图像的卷积在节点  $i$  处的值。同样,在待识别人脸图像上,也定义一个二维网格上的向量场。其中  $X_j$  是和  $C_i$  一样类型的特征向量,只不过是定义在更大且更细密的二维网格  $S$  上的。

在弹性匹配中,库中人脸和待识别人脸间的匹配转化为  $S_1$  和  $S$  之间的匹配,也就是寻找  $S_1$  中的各节点在  $S$  中的最佳匹配节点。最佳的匹配应该同时考虑到特征的匹配和局部几何位置的匹配。如果  $S_1$  上的节点  $i$  在  $S$  中对应的匹配节点为  $j = M(i)$ ,那么特征匹配就是说向量  $C_i$  和向量  $X_j$  差别不大;如果  $S$  上  $i_1, i_2$  两节点在  $S_1$  上对应的匹配节点分别为  $j_1 = M(i_1), j_2 = M(i_2)$ ,那么局部几何位置的匹配就是说如果  $i_1$  和  $i_2$  位置很近,那么  $j_1$  和  $j_2$  位置也应该很近。文[5]中用了如下的匹配度函数来评价待识别人脸图像的向量场和库中已知人脸的向量场之间的匹配程度:

$$E(M) = \sum_i \frac{\langle C_i, X_j \rangle}{\|C_i\| \|X_j\|} - \lambda \sum_{i_1, i_2} \cdot \|(P(i_1) - P(i_2)) - (Q(j_1) - Q(j_2))\|^2 \quad (1)$$

其中的  $P(i)$  表示网格  $S_1$  中节点  $i$  在库中人脸上的坐标,  $Q(j)$  表示网格  $S$  中节点  $j$  在待识别人脸上的坐标。表示式中的  $\sum_i \frac{\langle C_i, X_j \rangle}{\|C_i\| \|X_j\|}$  是在  $S_1$  中对  $i$  求和,而  $j$  为  $S_1$  中的  $i$  节点在  $S$  中的对应匹配节点,当  $X_j$  和  $C_i$  相近的时候,该项就会接近于 1。而  $\sum_{i_1, i_2} \|(P(i_1) - P(i_2)) - (Q(j_1) - Q(j_2))\|^2$  是对  $S_1$  中所有相邻的两节点求和,  $j_1$  为  $S_1$  中的  $i_1$  节点在  $S$  中的对应匹配节点,  $j_2$  为  $S_1$  中的  $i_2$  节点在  $S$  中的对应匹配节点,当  $P(i_1) - P(i_2)$  约等于  $Q(j_1) - Q(j_2)$ ,即保留了局部距离和匹配次序的时候,该项就会接近于 0。 $M$  是从  $S_1$  到  $S$  的单射,最佳匹配就是选取合适的  $M$  使得  $E(M)$  最大时的匹配,此时的匹配值就是此两人脸间的匹配度。

在经典算法的基础上,弹性匹配还有一些改进算法,如文献[8~10]。

### 3 弹性匹配算法在面对强背景干扰时的缺陷及使用关键点信息的改进算法

理论上,弹性匹配算法对各种小的变形有比较大的鲁棒

性,因此得到了广泛的研究,并被认为是具有较高的识别率<sup>[11]</sup>。但在出现比较强的背景干扰时却很有可能出现错误匹配,例如在图 3 中,人脸  $a$  和人脸  $b$  属于同一个人,而  $c$  则是另一个人的脸,可在使用经典的弹性匹配算法进行判别时反而会认为  $c$  和  $b$  的匹配度比  $a$  和  $b$  的匹配度更高。

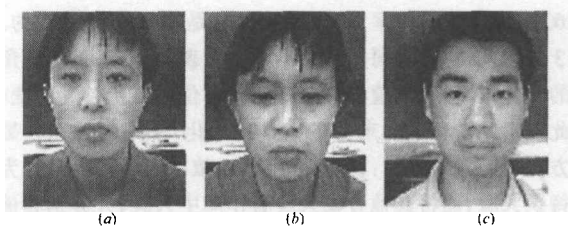


图 3 存在较大背景干扰的人脸

出现这种错误的原因是:在图 3 中有一个键盘,其靠近人脸的那一部分很难和人脸分开,因而会对识别产生影响,而键盘又具有较高的 Gabor 特征值,因此较大地影响了匹配度  $E(M)$ ,使人脸和键盘的相对位置与图  $b$  更接近的图  $c$  获得了比图  $a$  更高的匹配度。

经过仔细分析,认为这种错误的出现和弹性匹配是一种基于整体的方法有很大关系。虽然通常<sup>[1]</sup>认为基于整个人脸的分析要优于基于部件的分析,理由是前者保留了更多的信息,但是在另一方面,由于把整个人脸图像作为一个模式,那么各种整体和局部的干扰都会对人脸识别有很大的影响,如何有效地去掉这些干扰因素是个很头疼的问题,处理得不好,会导致有用的识别信息被淹没在大量的冗余信息中而失去识别的意义。相反基于部件的方法就能比较好地解决这个问题,因为部件在人脸内部,因此不太容易受到背景的影响。即使受到影响,各个部件也不太可能同时受到影响。为此本文考虑将基于整体的方法和基于部件的方法相结合:首先用基于部件的方法定位出部分部件上的关键点,然后利用能够成功定位的那些关键点的位置对匹配度函数  $E(M)$  进行矫正(是否能够正确定位由定位算法自动判别)。新的匹配度函数的一种可能形式是

$$E(M) = \sum_i \frac{\langle C_i, X_j \rangle}{\|C_i\| \|X_j\|} - \lambda_1 \sum_{i_1, i_2} \cdot \|(P(i_1) - P(i_2)) - (Q(j_1) - Q(j_2))\|^2 - \lambda_2 \sum_i \cdot \|(P(i) - Q(j)) - \frac{1}{K} \sum_{k=1}^K (P(I_k) - Q(J_k))\| \quad (2)$$

相对于式(1),式(2)增加了矫正项  $\lambda_2 \sum_i \|(P(i) - Q(j)) - \frac{1}{K} \sum_{k=1}^K (P(I_k) - Q(J_k))\|$ ,其更具一般性的形式为  $\lambda_2 \sum_i D\left((P(i) - Q(j)), \frac{1}{K} \sum_{k=1}^K (P(I_k) - Q(J_k))\right)$ 。其中  $K$  表示能够定位出的关键点的数目,  $P(I_k)$  表示定位得到的第  $k$  个关键点在库中人脸上的坐标,而  $Q(J_k)$  则表示定位得到的第  $k$  个关键点在待识别人脸上的坐标,  $\lambda_2$  则是加权系数,而  $D(\cdot)$  则是距离量度。这里采用了欧氏距离  $D_e(p, q) = [(p_x - q_x)^2 + (p_y - q_y)^2]^{1/2}$  作为距离量度,也可以考虑采用其他距离

量度,例如城区(city-block)距离  $D_4(p, q) = |p_x - q_x| + |p_y - q_y|$  等。

#### 4 实验结果

为检测改进算法的性能利用了一个包含了来自 12 个人的 60 幅人脸的图象库,该库中的人脸是用摄像机拍摄的.与图 3 中的人脸类似,同一人的不同人脸表情、姿态和拍摄角度等的变化很小,但人脸在背景中的相对位置有比较大的变化,因此识别的错误主要来自于背景的变化.采用一种比较简单的方法来定义“错误”:在识别过程中,依次取各幅人脸作为当前待识人脸,而将其他各幅人脸看作库中人脸,并分别和该人脸比较,得到匹配度,并将这些人脸按匹配度排序,如果在该序列中有一幅与当前待识人脸属于同一人的脸与当前人脸的匹配度排在不属于同一人的其他人脸之后,就认为出现了一次错误,如果有两幅这样的人脸就认为出现了两次错误,以此类推.使用两眼的瞳孔中心和鼻尖作为关键点,对其中 58 幅人脸,能够同时正确定位出两眼的瞳孔中心和鼻尖的位置,因此  $K$  取 3,而对另外 2 幅人脸,只能正确定位出两眼瞳孔中心的位置,因此  $K$  取 2(如果所有关键点均无法正确定位,则  $K$  只能取 0,也就是说无法采用本文的新算法).表 1 中显示了错误次数  $n$  随权值变化的情况

表 1 权值和错误出现次数的关系

权值 $\lambda_2$	0	$10^{-6}$	$5 \times 10^{-6}$	$10^{-5}$	$5 \times 10^{-5}$	$10^{-4}$	$5 \times 10^{-4}$	$10^{-3}$	$5 \times 10^{-3}$	$10^{-2}$
错误次数 $n$	42	42	36	32	12	12	12	0	0	0

可以看到,  $n$  随  $\lambda_2$  的增大而减小,当  $\lambda_2$  大于  $10^{-3}$  时就已经没有错误了.但并不是越大越好,对  $\lambda_2$  的调整其实就是对通过部件定位得到的关键点信息和通过经典弹性匹配得到的全局信息在识别中所占比重的调整.  $\lambda_2$  取 0 时,相当于经典的弹性匹配方法,  $\lambda_2$  越大,表明局部特征在算法中占的比例也越大,而当其取一个很大的值(例如 1)时,算法其实已经退化为简单的基于关键点位置信息的人脸识别方法.虽然在本实验中看不出其不良影响,但却会降低对表情和视角变化的鲁棒性,为此希望能够找到一个比较小的不出现错误的  $\lambda_2$  值.对上述人脸库的更精确的分析表明,错误次数减少为 0 的临界值为  $6.2867 \times 10^{-4}$ .取其他的人脸库进行试验,  $\lambda_2$  的临界值均在  $10^{-4}$  和  $10^{-3}$  之间,因此取  $10^{-3}$  就能比较好地满足要求.

#### 5 总结和展望

本文提出了将基于整体的方法和基于局部的方法相结合,利用局部的关键点信息来改进弹性匹配算法的思想,并研究了权值的选取.实验表明本文的方法确实能减少背景对弹性匹配产生的干扰,权值的选取也有一定的普遍性.

本算法不仅可用于背景干扰比较大的场合,也可考虑用于部分人脸受到遮挡的场合.

在定位得到关键点时,必须首先判别该部件是否受到了背景的影响,本文实际采用的是将定位算法的阈值设得对虚

警特别敏感的方法,但这样显然就降低了能成功定位的部件数,以后可以考虑如何在这两者间实现权衡.

本文对各个部件采用了相同的权值,也可以考虑根据其重要性给以不同的权值

#### 参考文献:

- [1] R Berto, T Poggio. Face recognition: feature versus templates [J]. IEEE Trans on PAMI, 1993, 15(10): 1042 - 1052.
- [2] M Turk, A Pentland. Face recognition using eigenfaces [A]. Proc of IEEE Conf on CVPR [C]. USA: Maui HI, 1991. 586 - 591.
- [3] Ziquan Hong. Algebraic feature extraction of image for recognition [J]. Pattern Recognition, 1991, 24(3): 211 - 219.
- [4] O Nakamura, et al. Identification of human faces based on isodensity maps [J]. Pattern Recognition, 1991, 24(3): 263 - 272.
- [5] M Lades, et al. Distortion invariant object recognition in the dynamic link architecture [J]. IEEE Trans on Computers, 1993, 42(3): 300 - 311.
- [6] F Samaria, S Young. HMM-based architecture for face identification [J]. Image and Vision Computing, 1994, 12(8): 537 - 543.
- [7] D Valentin, et al. Connectionist models of face processing: a survey [J]. 1994, 27(9): 1209 - 1230.
- [8] L Wiskott, et al. Face recognition by elastic bunch graph matching [J]. IEEE Trans on PAMI, 1997, 19(7): 775 - 779.
- [9] M A Grudin. Compact multi-level representation of human faces for recognition [D]. Liverpool, UK: Liverpool John Moores University, 1997.
- [10] C L Kotropoulos, et al. Frontal face authentication using discriminating grids with morphological feature vectors [J]. IEEE Trans. on Multimedia, 2000, 2(7): 14 - 26.
- [11] Jun Zhang, et al. Face recognition: eigenface, elastic matching, and neural nets [A]. Proc. of the IEEE [C]. UK: 1997, 85(9): 1422 - 1435.

#### 作者简介:



丁 彦 男, 1975 年 4 月出生于江苏省泰兴市, 1995 年和 1998 年于东南大学无线电工程系分别获学士学位和硕士学位, 其间主要研究多媒体数据的压缩, 在国际核心期刊和国内期刊上发表多篇学术论文, 现在清华大学电子工程系攻读博士学位, 目前主要研究方向为人脸识别.



苏光大 男, 1948 年 6 月出生于重庆, 毕业于清华大学电子系, 副教授, 长期从事图象处理系统的科学研究, 研制成功通用图象处理系统、γ 照相机图象处理系统、B 超图象处理系统、模糊图象复原系统、计算机人像组合系统、人脸识别系统, 先后五次获部委级科技成果奖, 公开发表 40 余篇学术论文, 出版一部著作, 目前主要研究方向为人脸识别和高速图象处理.