

# 基于层次模型和融合决策的多姿态人脸识别技术<sup>1</sup>

朱长仁 王润生

(国防科技大学电子科学与工程学院 ATR 国家重点实验室 长沙 410073)

**摘 要** 该文提出了基于层次模型和融合决策的多姿态人脸识别技术。它首先把各人脸按姿态分成几个大类,并且对各大类按人脸个体分成相应子类,然后对各个大类分别进行基于特征脸的人脸识别,最后对各个姿态的人脸识别中间结果进行融合决策得到真正的人脸识别结果。该算法同时也提供了其姿态识别结果,并且大大减小了耗时。该文算法对 ORL, UMIST, Stirling 数据和一些自拍数据共 1200 幅人脸图像进行了识别测试实验,其结果令人鼓舞。

**关键词** 人脸识别,多姿态,特征脸,多视图,层次模型,融合决策

**中图分类号** TN911.73, TP391.41

## 1 引言

人脸识别是计算机视觉、模式识别中一个重要内容,在电视监视、罪犯查询识别、安全门卫系统等领域中有着广泛的应用<sup>[1-3]</sup>。人脸识别大致可以分为 3 步:人脸检测定位、特征提取、分类识别。作为人脸识别的第 3 步——人脸分类识别虽然已经有广泛深入的研究<sup>[1-8]</sup>,但是其中仍然存在许多问题没有解决,这也是导致目前人脸识别技术没有得到足够推广的主要原因之一。

人脸识别方法很多,它们大致可分为 4 类:基于人脸几何特征匹配方法、基于弹性模板方法、基于灰度匹配的 PCA(特征脸)方法<sup>[7]</sup>、基于连接性模型方法。其中基于灰度匹配的特征脸方法是人脸识别中较为重要的方法,它的识别正确率较高而倍受关注。该方法的实质是进行图像之间的灰度快速匹配。但是经典的基于特征脸的人脸识别方法对人脸姿态变化等比较敏感,其识别算法基本思路是:不管何种姿态的人脸图像,都是基于某一姿态的 PCA 特征脸的投影特征基础上进行人脸识别。实际上,基于特征脸方法是基于待识别图像与训练图像人脸之间具有很好的相关性而建立起来的,这对于同一姿态下的归一化后人脸图像是合适的假设,但是多种姿态的人脸图像之间相关性明显下降,这也正是经典 PCA(特征脸)人脸识别方法在识别多姿态人脸图像时正确率迅速下降的主要原因之一。文献 [8] 提出了基于多视图和局部特征空间的方法来提提高识别正确率,但是该方法对输入的人脸图像有较高的分辨率要求,而实际遇到的人脸图像可能是千变万化的,不可能总满足此要求。为此,本文提出层次模型思想,即先把所有人脸训练样本按姿态划分若几个大类,再将各大类按人脸个体分成许多子类,这样待识别人脸图像与相应个体相应姿态的典型训练人脸图像之间相关性最强,这样各姿态大类的人脸识别则可以采用经典的人脸识别方法,而最终的人脸识别结果则可以通过对各个姿态的人脸识别结果进行融合决策求得。并且该算法还同时得到相应人脸的姿态识别结果。

另外,在非并行计算环境中为减少计算量,我们提出在进行各姿态的人脸识别之前先对姿态进行粗检测,提取候选的可能姿态,然后只对候选的姿态进行识别、融合决策,而不需要对每一个姿态都做。

由于本文算法中各姿态的人脸识别中搜索范围局限于本姿态的训练样本,所以大大减小了耗时。

下面首先介绍基于层次模型和融合决策的多姿态人脸识别算法原理,然后详细论述基于层次模型和融合决策的多姿态人脸识别算法 1,其次详细论述在非并行计算环境中为减少计算量的基于层次模型和融合决策的多姿态人脸识别算法 2。最后给出本文算法对 ORL, UMIST,

<sup>1</sup> 2001-03-30 收到, 2001-08-17 定稿

Stirling 数据和一些自拍数据共 1200 幅人脸图像数据的测试实验结果, 其人脸识别结果令人鼓舞。

## 2 基于层次模型和融合决策的多姿态人脸识别技术原理

由于随着姿态变化差异增大, 同一个体的各人脸图像之间的相关性明显下降, 为此, 该技术首先把人脸按姿态分成几个大类。人脸在 3D 空间活动有 3 个自由度 ( $X, Y, Z$ ), 如图 1 所示, 设  $XYZ$  直角坐标系的原点为检测后的人脸中心点, 人脸在 3D 空间中的姿态变化由 6 个因素引起: 沿  $X, Y, Z$  轴的平移、旋转。而沿  $X, Y$  轴的平移、沿  $Z$  轴平移引起的人脸尺度放缩、沿  $Z$  轴旋转引起的人脸图像倾斜, 都可以通过检测人眼的瞳孔中心两点, 并通过仿射变换使两个瞳孔中心两点在图像中处在相应固定位置即几何归一化来克服, 这样对于几何归一化后的人脸图像, 我们只需要分析沿  $X, Y$  轴的旋转引起的变化因素。

由此, 我们把人脸姿态变化的空间沿  $X, Y$  轴旋转的角度进行量化划分为  $M \times N$  等分子空间:  $K_{11}, \dots, K_{1N}, \dots, K_{M1}, \dots, K_{MN}$ , 具体如图 2 所示。考虑到相对于正前方姿态对称, 一般  $M, N$  都应设置为奇数。

把人脸按姿态划分若干个大类后, 对各个体在各姿态都建立相应视图的典型训练样本, 这样各类姿态的人脸识别就可以采用经典的人脸识别方法 (即层次模型思想), 最后对各个姿态视图的识别结果进行融合决策, 求得正确的人脸, 并输出相应的姿态识别结果。

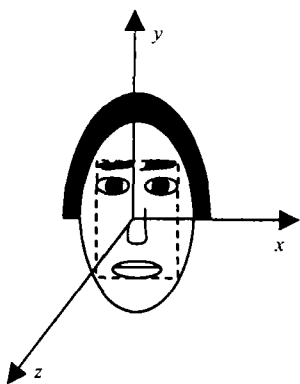


图 1 人脸姿态变化的自由度

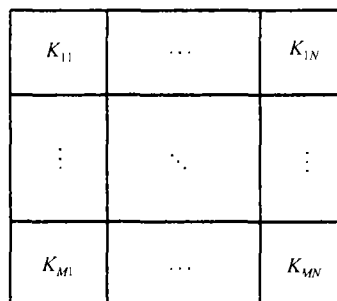


图 2 人脸姿态划分图

## 3 基于层次模型和融合决策的多姿态人脸识别技术算法 1

首先把归一化后的人脸按上述的姿态空间划分为  $M \times N$  个大类 ( $K_{11}, \dots, K_{1N}, \dots, K_{M1}, \dots, K_{MN}$ ),  $M, N$  都为奇数, 然后在各类中又按各个体人脸分成若干个子类。对于一幅人脸图像经过归一化后进行基于各个姿态的人脸识别, 然后利用基于各个姿态通道输出的人脸识别结果和相应的置信度进行融合决策得到最终的人脸识别结果和相应的置信度, 并得到相应的姿态识别结果, 具体算法框架如图 3 所示。

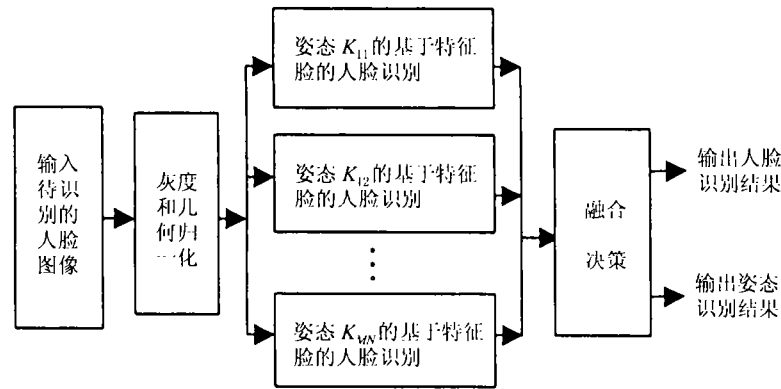


图 3 基于层次模型和融合决策的多姿态人脸识别技术算法 1 框图

(1) 人脸图像的灰度和几何归一化: 这是基于特征脸的人脸识别算法必须的, 主要目的是提高各个人脸图像之间的相关性。人脸图像的灰度归一化通过灰度直方图的规则化实现。人脸几何归一化是通过检测人眼的瞳孔中心两点, 并利用仿射变换使两个瞳孔中心点在图像中处在相应固定位置来实现。两瞳孔中心点的寻找算法目前已有一些<sup>[9,10]</sup>, 但是它们都太耗时, 我们此处人脸向上的(即眼睛在人脸区域的上半部分)即绕  $Z$  轴的旋转在  $(-90^\circ, 90^\circ)$  之间的假设下, 提出基于组合利用灰度、梯度、人脸器官的结构分布特性的瞳孔定位算法: (a) 图像分割提取人脸区域; (b) 人脸在  $Z$  旋转方向归一化。由于在人脸竖直向上时人脸器官的边缘方向大都为接近水平方向, 所以可以通过对人脸区域中的边缘方向分布进行统计分析, 并旋转图像使最大边缘方向为水平方向, 以消除图像平面内旋转因素, 即使人脸竖直向上; (c) 基于人脸认知模型粗略提取人脸中心区域(即图 1 虚线框内部分), 利用灰度的水平投影和垂直投影确定瞳孔中心: 在人脸中心区域中上  $1/2$  区域作水平投影  $H_x$ , 并求  $H_x$  满足一定条件的谷点作为瞳孔中心垂直坐标值, 并求相应区域垂直投影  $H_y$  的左右谷点分别作为两瞳孔中心水平坐标值。

(2) 各个姿态的人脸识别算法: 利用基于特征脸的经典算法实现。由于该方法是通过各人脸图像在特征脸上的投影权值矢量之间的最小距离匹配实现, 所以可以在提供人脸识别结果的同时, 也可以提供基于矢量之间匹配距离的置信度。这样, 我们就可以利用各姿态的识别结果进行融合决策确定最终的识别结果。

设参加训练的人脸个体数为  $S$ , 则把训练的人脸图像按个体和姿态差异分为  $(M \times N) \times S$  个子类, 把第  $j$  个个体的第  $i$  个姿态的第  $l$  个人脸图像记为  $C_{ijl}$ , 其中  $0 < i \leq M \times N, 0 < j \leq S, 0 < l \leq N_{ij}, N_{ij}$  为第  $j$  个个体在第  $i$  个姿态中的训练人脸图像数目。 $C_{ijl}$  在第  $i$  个姿态通道中第  $k$  个特征脸矢量上的投影权值记为  $w_{ijlk}$ , 那么第  $j$  个个体在第  $i$  个姿态上的第  $l$  个人脸图像的投影权值矢量为  $W_{ijl} = (w_{ijl1}, w_{ijl2}, \dots, w_{ijlq})$ ,  $q$  为投影矢量维数。待识别人脸图像在第  $i$  个姿态中的投影权值矢量为  $W_i^0 = (w_{i1}^0, w_{i2}^0, \dots, w_{iq}^0)$ 。这样待识别人脸图像与第  $j$  个个体在第  $i$  个姿态的第  $l$  个人脸图像的投影权值矢量之间的欧氏距离  $d_{ijl} = \|W_i^0 - W_{ijl}\| = \sqrt{\sum_{k=1}^q (w_{ik}^0 - w_{ijlk})^2}$ , 该欧氏距离越小, 两图像越相关, 由此可以定义待识别人脸图像是第  $j$  个个体在第  $i$  个姿态上的第  $l$  个人脸图像的基于欧氏距离的人脸置信度为  $h_{ijl} = (A - d_{ijl})/A$ , 其中  $A$  为一个很大的常数, 满足  $A \geq \text{Max}_{0 < i \leq M \times N, 0 < j \leq S, 0 < l \leq N_{ij}} (d_{ijl})$ 。

(3) 融合决策: 各个姿态通道人脸识别的结果需要识别决策提取最后的人脸识别结果。融合识别方法较多, 此处我们利用加权求和与求最大值的组合融合识别方法。由于随着姿态变化增大, 各个姿态人脸图像之间的相关性明显减弱。一般情况下, 待识别人脸图像应该与本姿态的

相应训练人脸图像相关性最强, 这样我们可以首先在各个姿态的人脸识别结果中寻找置信度最大值, 由此确定相应的姿态识别结果。另外, 某个体在同一个姿态中的各个人脸图像之间具有一定的相关性, 使得该个体在同一姿态的各人脸图像识别结果有一定的连动性, 即假如某个体在某个姿态的某个人脸图像识别置信度很高, 则该个体在同一姿态的其它人脸图像识别置信度也较高。这样我们可以对已识别姿态的各个人脸识别置信度进行加权识别, 即对它们的置信度进行融合加权求和产生该个体人脸的总的置信度, 然后根据各个体人脸总的置信度进行判决最终的人脸识别结果。

设待识别人脸图像是第  $j$  个个体在第  $i$  个姿态上的第  $l$  个人脸图像的人脸置信度为  $h_{ijl}$ ,  $0 < i \leq M \times N, 0 < j \leq S, 0 < l \leq N_{ij}$ , 则人脸姿态识别结果为  $Pose = i^* = \text{Arg} \max_{\substack{0 < i \leq M \times N \\ 0 < j \leq S, 0 < l \leq N_{ij}}} (h_{ijl})$ 。在姿态  $Pose$  上, 设第  $j$  个个体的各个训练人脸图像的人脸识别置信度从大到小排序为:  $h_{i^*j1} \geq h_{i^*j2} \geq \dots \geq h_{i^*jN_{ij}}$ , 则待识别人脸图像是第  $j$  个个体人脸的总置信度为  $h_j = h_{i^*j} = \left(1 / \sum_{l=1}^{N_{i^*j}} \beta_l\right) \sum_{l=1}^{N_{i^*j}} \beta_l h_{i^*jl}$ ,  $\beta_l$  为权系数, 根据实验经验取值。这样最终人脸识别结果是融合加权后的置信度取最大值所对应的个体, 即  $Person = j^* = \text{ArgMax}_{0 < j \leq S} (h_j)$ 。

有关  $\beta_l$  的取值: 在利用上面定义的置信度对经典算法进行的实验分析中, 我们发现识别错误往往出现在待识别图像对应个体的几个图像的识别置信度大都仅比第一大识别置信度低一点, 而第一大识别置信度图像对应个体的其它图像的置信度往往较低, 这样若能以最大限度提高识别正确率为原则, 对  $\beta_l$  根据实验经验在一定范围内取合适值, 就可以控制风险, 有效提高识别正确率。

算法 1 中各姿态识别中搜索范围局限于本姿态的训练样本, 所以在并行计算环境中其耗时大约为基于 PCA 特征脸的经典算法的  $1/(M \times N)$  多一点, 大大提高了算法速度。

#### 4 基于层次模型和融合决策的多姿态人脸识别技术算法 2

由于算法 1 中待识别人脸图像对每个姿态都要做经典人脸识别过程, 所以它较适合于并行计算, 对于非并行计算时算法 1 就较为耗时, 为此我们提出非并行计算时的算法 2(如图 4)。算法 2 是在算法 1 的基础上建立起来的, 主要是通过减少各姿态人脸识别的通道数目来减少计算量。我们采用的方法是: 在各姿态人脸识别之前先确定可能的几个候选姿态, 然后各姿态人脸识别、融合决策都只对候选姿态进行。

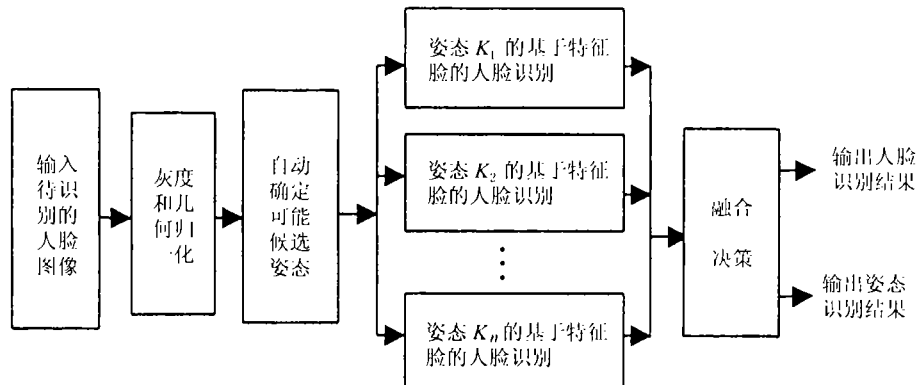


图 4 基于层次模型和融合决策的多姿态人脸识别技术算法 2 框图

由于待识别的多姿态人脸图像与本姿态的人脸图像灰度相关性最强, 并且随着姿态的变化逐渐增大, 这种相关性逐渐减弱。所以, 可以首先建立各个姿态的标准人脸图像, 然后根据待识别人脸图像与各个姿态标准人脸图像相关性强弱来确定候选可能的姿态, 即相关性由强到弱排序并选取前  $n$  种可能候选的姿态 (设为  $K_1 \sim K_n$ )。

各个姿态的标准人脸图像可以由几个同一姿态的灰度和几何归一化后的人脸图像平均得到。待识别图像与各姿态的标准图像之间的灰度相关计算方法有许多, 本文采用计算它们在特征脸上投影权值矢量之间距离实现, 其步骤为: 在训练时分别计算各个姿态的标准人脸图像在某一个姿态  $i$  (一般选择标准正前姿态) 的特征脸上的投影权值矢量  $W_1^1 \sim W_{MN}^1$ ; 在识别阶段, 首先计算待识别的灰度和几何归一化后的人脸图像在同一姿态的特征脸上的投影权值矢量  $W_i^0$ , 然后分别计算  $W_i^0$  与  $W_1^1 \sim W_{MN}^1$  矢量之间距离, 并取距离最小的前  $n$  个姿态作为可能的候选姿态。

算法 2 中各姿态识别中搜索范围局限于本姿态的训练样本, 所以在并行计算环境中其耗时稍大于基于 PCA 特征脸的经典算法的  $1/(M \times N)$ , 在非并行计算环境中其耗时约为基于 PCA 特征脸的经典算法的  $n/(M \times N)$ , 大大提高了速度。

## 5 实验结果和总结

本文算法对 ORL, UMIST, Stirling 和一些自拍数据共 100 人、1200 幅不同姿态的人脸图像数据进行了测试实验。实验中由于实验的姿态数据有限, 我们把姿态空间划分为  $1 \times 3$  个子空间, 权系数  $\beta_{l'} = 0.1^{(l'-1)}$ ,  $0 < l'$ , 算法 2 选择可能候选姿态数目为  $n = 2$ 。图 5 给出实验的部分样本。实验中我们选择人脸各个姿态的典型图像 480 幅作为训练集, 剩下的图像全作为测试集, 识别的实验结果见表 1。表中数据均是在 CPU 为赛扬 400、内存 64M 的微机得到的。

表 1 人脸识别实验结果对比表

方法	本文算法 1		本文算法 2	
	人脸识别	姿态识别	人脸识别	姿态识别
正确率	96.33%	98.62%	95.26%	97.86%
平均识别一幅人脸图像耗时 (s)	0.72		0.66	

表 1 数据表明基于本文两个算法的人脸识别正确率都较高, 而且本文算法在得到人脸识别结果的同时, 也能得到相应的姿态识别结果。关于耗时方面, 由于本文算法中各姿态识别中搜索范围局限于本姿态的训练样本, 所以在并行计算环境中算法 1 和算法 2 耗时均大约为基于 PCA 特征脸的经典算法的  $1/(M \times N)$ , 在非并行计算环境中本文算法 2 耗时大约为基于 PCA 特征脸的经典算法的  $n/(M \times N)$ , 算法 1 耗时和经典算法相当。若采用硬件 DSP + FPGA 并用流水作业和并行处理的体系结构实现本文算法, 则有望满足实时识别要求, 所以其耗时并不是问题。

另外需要指出的是, 本文算法适合于双眼可视并在人脸区域的上半部分, 即人脸绕  $Z$  轴旋转的角度在  $(-90^\circ, 90^\circ)$  之间的各种姿态人脸识别, 若采用其它更为准确可靠的人眼瞳孔计算方法, 则本文算法可以适合于双眼可视的任意姿态人脸识别。另一方面, 若人脸总是正直向上并已归一化的假设条件成立时, 即没有沿  $Z$  轴旋转和平移引起的变化时 (对应实际情况为: 相机与人脸相对距离固定, 且人脸是竖直向上), 那么本文算法还可以正确识别直至完全侧向的人脸, 如图 6 所示。总之, 本文算法具有实际应用潜力。



图 5 部分人脸样本图像



图 6 完全侧面的部分人脸样本图像

**致谢** 我们谨在此对为本次实验提供图像样本的 Olivetti 实验室、UMIST 的 Daniel B. Graham and Nigel M. Allinson., Stirling 大学以及为本次实验做模特的各位同仁表示感谢!

### 参 考 文 献

- [1] R. Chellappa, C. L. Wilson, S. Sirohey, Human and machine recognition of faces: A survey, Proc. IEEE., 1995, 83(5), 705-740.
- [2] A. Samal, P. A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: A survey, Pattern Recognition, 1992, 25(1), 65-77.
- [3] D. Valentin, H. Abdi, A. J. O'Toole, G. W. Cottrel, Connectionist model of face processing: A survey, Pattern Recognition, 1994, 27(9), 1209-1230.
- [4] R. Brunelli, T. Poggio, Face recognition: Feature versus templates, IEEE Trans. on PAMI., 1993, PAMI-15(10), 1042-1052.
- [5] S. Ranganath, Face recognition using transform features and neural networks, Pattern Recognition, 1997, 30(10), 1615-1622.
- [6] Lam Kin-Man, Yan Hong, An analytical-to-holistic approach for face recognition based on a single frontal views, IEEE Trans. on PAMI., 1998, PAMI-20(7), 673-686.
- [7] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience, 1991, 3(1), 71-86.
- [8] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, IEEE Conf. CVPR'94, Seattle, 1994. 6, 84-91.
- [9] K. M. Lan, H. Yan, An improved method for locating and extracting the eye in human face images, ICPR'96, Vienna, Austria, 1996. 8, C411-C415.

- [10] X. Xie, R. Sudhaker, H. Zhuang, On improving eye feature extraction using deformable templates, *Pattern Recognition*, 1994, 27(6), 791-799.

## MULTI-POSE FACE RECOGNITION BASED ON A HIERARCHICAL MODEL AND FUSION DECISION

Zhu Changren      Wang Runsheng

(*ATR National Lab., National University of Defense Technology, Changsha 410073, China*)

**Abstract** A multi-pose face recognition algorithm based on a hierarchical model and fusion decision is proposed in the paper. First, face images are divided into several classes based on their poses and every class into many sub-classes based on their subjects. Then, traditional face recognition technology based on eigenface is applied in the process of face recognition in every class. Finally, fusion decision is adopted to get the final result of face recognition based on the intermediate result of every class. It, also, can get the result of its pose recognition and its computation time is far less than that of the traditional methods. The algorithm discussed above has been tested in the experiments of a multi-pose face database with total 100 subjects and 1200 face images, which consists of ORL, UMIST, Stirling and face database of authors' lab. The results are encouraging.

**Key words** Face recognition, Multi-pose, Eigenface, Multi-view, Hierarchical model, Fusion decision

朱长仁: 男, 1968 年生, 讲师, 博士, 主要研究领域为图像分析理解、模式识别.

王润生: 男, 1941 年生, 教授, 博士生导师, 中国电子学会高级会员、中国图像图形学会常务理事, 主要研究领域为图像分析理解、模式识别、信息融合等.