

文章编号:1671-6833(2017)05-0023-05

基于发育网络的人脸朝向识别研究

王东署, 谭达佩, 韦晓琴

(郑州大学 电气工程学院, 河南 郑州 450001)

摘 要: 针对人脸识别中人脸的朝向、位置以及背景光线不固定的特点, 提出了一种基于发育网络的人脸朝向识别新方法. 对图像进行处理后发现人眼的特征非常突出, 故选择眼睛的位置作为人脸朝向的特征向量, 利用发育网络模型对不同背景光线图像中人脸的朝向进行识别. 通过和其它方法的测试结果对比, 该方法可以有效地解决不同光照条件下人脸朝向识别问题, 并具有快速、稳定、高效的特点, 且识别率高达 100%.

关键词: 发育网络; 人脸朝向识别; 光线背景; 特征向量; 识别率

中图分类号: TP18 **文献标志码:** A doi:10.13705/j.issn.1671-6833.2017.02.023

0 引言

在人脸识别研究中, 鉴于现实情况的要求, 主要是研究人脸正面模式, 人脸正面的意思是指在拍摄过程中人脸一直正对着摄像机. 但在采集样本图像时, 由于人脸的朝向是变化的, 光线背景也是各式各样的, 所以其最终获得的特征向量以及图像所包含的信息也会有变化, 这给人脸识别算法提高了难度. 在人脸检测时可以对不同朝向的人脸图像进行分类, 这样做可以缩小人脸识别算法中的搜索空间^[1].

在人脸朝向识别的研究中, 已有各种不同的方法. 文献[2]提出一种基于中心轮廓的 3D 人脸姿态估计算法, 定义了一个用于评价脸部轮廓的客观函数, 并利用 Hough 变换来搜索参数空间中的对称平面, 可有效地应用于人脸朝向估计. 文献[3]提出一种由粗到细的人脸姿态估计框架, 在粗细层上分别利用单位圆和三维球体建立流形拓扑模型, 该方法在人脸朝向识别上表现优越. 文献[4]提出一个分层时态图形化模型, 可以估计连续头部姿态角的人脸, 适应任何脸部分类任务. 文献[5]提出一个新的基于图像的连续人脸姿态估计嵌入方案, 利用稀疏编码方案对脸部特征和姿态标签进行线性嵌入.

目前国内研究中较有代表性的有: 陈锐等^[6]提出一种在相对自然的环境条件下进行人脸朝向方向分析的算法, 并基于摄像头来粗略和精细地分析朝向. 张彤等^[7]利用 BP 神经网络对人脸的朝向进行识别, 能达到一定的识别率. 谭乐平等^[8]提出基于 LVQ 神经网络的人脸朝向识别方法, 利用图像的二值信息作为网络的输入, 识别率高达 90% 以上. 朱宇鑫等^[9]将概率神经网络(PNN)用于人脸朝向识别, 结果表明 PNN 在人脸朝向识别方面表现优异. 然而以上利用神经网络来进行人脸朝向识别的研究都未考虑在不同光照条件下对人脸朝向进行识别. 目前仅有朱宇鑫^[10]研究了基于神经网络的不同光线背景下人脸朝向识别, 识别率不理想, 鲁棒性以及对环境适应能力较弱.

为了有效区分人脸的朝向, 笔者提出一种基于发育网络的人脸朝向识别新方法. 该方法旨在创建发育网络, 选择眼睛的位置信息作为特征向量输入网络, 对人脸朝向进行识别.

1 发育网络

1.1 发育网络模型

2001 年首次提出自主心智发育的概念^[11]. 笔者采用发育网络^[12]展开研究. 模型主要由 X、

Y、Z 3 个区域构成,具体结构如图 1 所示.

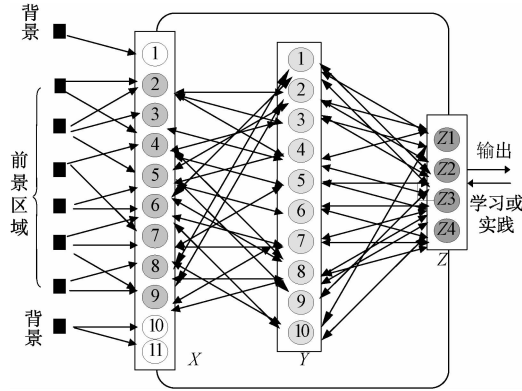


图 1 发育网络模型结构图

Fig.1 The structure diagram of developmental network model

1.2 发育网络算法

发育网络算法^[13]是以叶成分分析方法(lobe component analysis, LCA)为基础的.

(1)当 $t=0$ 时,对于网络 A 中所有的区域,初始化网络的自适应部分 N 和神经元的响应向量 r ,其中: $A = \{X, Y, Z\}$; $N = (V, G)$, V 和 G 分别是神经元的突触权值和年龄.

(2)当 $t=1, 2, \dots$ 时,对于网络 A 中所有的区域,反复地执行如下两个步骤:

①利用算法函数 f 计算并更新 N 和 r

$$(r', N') = f(b, t, N), \quad (1)$$

式中: b (bottom-up) 是自下而上的输入;而 t (top-down) 是自上而下的输入.

②对于网络 $A = \{X, Y, Z\}$ 中所有的区域,进行如下的更新:

$$N \leftarrow N' r \leftarrow r'.$$

对于网络 $A = \{X, Y, Z\}$ 中所有区域的神经元都有自身的权值向量 $v = (v_b, v_t)$, 可用于计算响应值,计算公式如下:

$$r(v_b, b, v_t, t) = \frac{v_b}{\|v_b\|} \cdot \frac{b}{\|b\|} + \frac{v_t}{\|v_t\|} \cdot \frac{t}{\|t\|} = \dot{v} \cdot \dot{p}, \quad (2)$$

式中: v_b 是自下而上的权值; v_t 是自上而下的权值; \dot{v} 是权值向量 $v = (v_b, v_t)$ 经过归一化后的单位向量; \dot{p} 是输入向量 $p = (b, t)$ 经过归一化后的单位向量. 内积可以衡量 \dot{p} 和 \dot{v} 相似的程度,也就是匹配程度,内积得到的数值位于 $[-1, 1]$ 之间.

获胜神经元 j 的位置可按下式计算:

$$j = \arg \max_{1 \leq i \leq c} r(v_{bi}, b, v_{ti}, t), \quad (3)$$

式中: j 表示被激活的神经元,其连接权值按如下公式更新:

$$v_j \leftarrow \omega_1(n_j) v_j + \omega_2(n_j) y_j \dot{p}. \quad (4)$$

式中: $\omega_2(n_j)$ 和 $\omega_1(n_j)$ 分别是学习率和保持率; y_j 是神经元 j 被激活后重置的响应向量; n 是神经元的激活年龄; $\omega_1(n_j) + \omega_2(n_j) = 1$; $\omega_2(n_j) = 1/n_j$. \dot{p} 的采样均值由如下公式计算:

$$v_j = \frac{1}{n} \sum_{i=1}^{n_j} \dot{p}(t_i), \quad (5)$$

式中: t_i 为激活时间,获胜神经元执行 $n_j \leftarrow (n_j + 1)$.

2 特征值提取

2.1 试验样本选择

笔者选择了 5 种不同朝向人脸图像样本, 5 种朝向分别表示为左边、左前、正前、右前和右边. 这些图像样本有各种不同的光线背景,如图 2 所示.

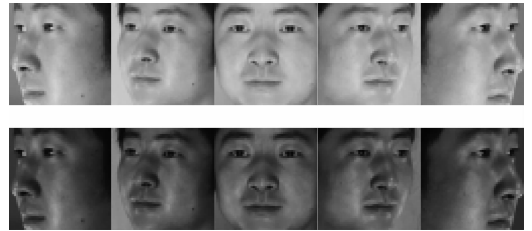


图 2 5 种不同朝向图像

Fig.2 Five kinds of different orientation images

2.2 特征值提取

利用 Matlab 的函数将图像转化为 $0 \sim 1$ 矩阵,通过观察发现人眼的特征非常突出,其数据是最多的,所以选取眼睛的位置作为特征向量. 首先将图像横向划分为 6 个部分,纵向划分为 8 个部分,如图 3 所示. 然后将划分好的图像转化为 6×8 的 $0 \sim 1$ 矩阵,对其做边缘检测可发现眼睛的位置就在矩阵横向的第二部分,并将该部分 8 个子矩阵的数据进行求和得到一个 8 维的特征向量输入网络^[14]. 考虑到某些图像不规范,眼睛位置不在横向第二部分,故而先采用人眼检测方法^[15]找出眼睛所在位置,再用离散程度分析法^[16]计算该位置眼睛数据的平均值和离散程度,根据这两个数值的大小选择眼睛所在位置最接近的横向部分再按前文方法计算特征值.

3 试验和结果分析

3.1 试验设计

通过上文对发育网络和特征值提取的介绍,笔者设计了相应的试验过程,如图 4 所示.

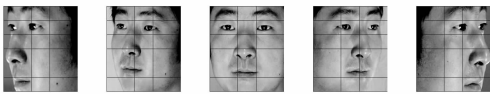


图 3 图像的划分

Fig.3 The division of images

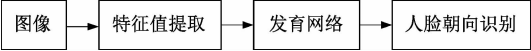


图 4 试验结构框图

Fig.4 The diagram of structure

3.1.1 网络的输入和输出模式设计

对于给出的人脸朝向试验样本,图像中人脸的朝向分别是左边、左前、正前、右前和右边.发育网络的输入就是提取的特征值,即 X 层的值为特征值,输出到 Y 层.由上文可知特征向量是一个 8×1 的行向量,即 X 层的维数为 8.图像样本只有 5 种朝向,可以设置输出的形式为 1、2、3、4、5,分别与左边、左前、正前、右前和右边相对应.

3.1.2 网络设置

(1)用 Matlab 编程实现发育网络的创建, X 层作为输入结点.

(2) Y 、 Z 层神经元的数量分别设置为 16 个和 5 个. Y 层神经元的数量可以任意改变,通过改变 Y 层神经元的数量找出最优的网络参数.因为该试验的图像样本中只有 5 种朝向,所以 Z 层神经元的数量固定为 5 个.

(3)初始化设置: X 、 Y 、 Z 层的输入都设置为 0,各层之间的连接权值都设置为 0,网络中所有神经元的年龄全部设置为 0,令匹配度的门限值为 0.9.可以改变匹配度的门限值来找出最优的网络参数.

(4)试验图库:笔者采集 10 个人不同朝向、光照的 150 张图像,选取其中 50 张光照相同而朝向不同的图像作为训练样本,其余 100 张图像作为测试样本.

3.1.3 网络训练和测试

训练过程中 Z 层一直受外界指导和监督,如图 5 所示.首先对图像进行特征提取,得到特征值作为 X 层的值输出到 Y 层. X 层输出 b 到 Y 层,而 Y 层输出 t 到 Z 层.由公式(2), b 和 t 分别与 X 、 Y 层和 Y 、 Z 层之间的连接权值相乘再相加可得响应向量 r ,然后找出响应值最大的位置 j ,激活相应位置的神经元,即 $r_j = 1$,其它位置的神经元响应值都置为 0,并更新被激活神经元 j 的连接权值,年龄增加一个单位值.

测试过程中各层之间的连接权值固定不变,

网络处于被“冻结”状态,如图 6 所示.

图像经过特征提取后得到特征向量赋予 X , X 层的值分别与 Y 层神经元的权值相乘得到响应值,响应值最大的神经元获胜,权值不更新. Y 到 Z 类似于 X 到 Y 的做法, Z 层中所有获胜神经元的位置与 1、2、3、4、5 对应,通过输出与目标的匹配结果来判断识别率.

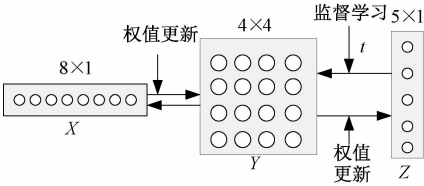


图 5 训练图

Fig.5 The diagram of training

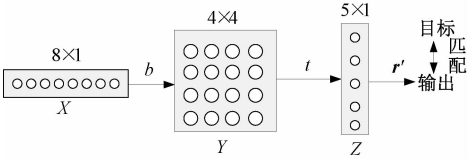


图 6 测试图

Fig.6 The diagram of testv

3.2 结果分析

3.2.1 试验结果

笔者通过设置网络参数,进行了大量的试验,以下列出其中一个测试结果.利用上文笔者设置的试验图库选取训练集和测试集,图像为 i_j 形式, i 表示人, j 表示人脸的 5 种朝向.部分测试集为:2_4 1_1 3_2 8_5 7_2 4_3 9_3 2_5 2_1 5_4 6_5 9_2 8_1 2_3 6_2 9_1 1_4 10_5 8_2 10_1 5_3 6_4 8_3 3_5 7_4. 试验设置的网络参数是: Y 层神经元个数为 16,匹配度门限值为 0.9,网络训练次数为 5.试验得到的结果为:ans = 4 1 2 5 2 3 3 5 1 4 5 2 1 3 2 1 4 5 2 1 3 4 3 5 4.

由结果可知,识别没有出现错误,识别率为 100%. 因此在网络参数选择合适的情况下,任意运行程序,每次运行得到的识别率都能达到 100%. 算法模型中的参数主要有匹配度门限值和 Y 层神经元的数目, Y 层不同的神经元个数和匹配度门限值对识别结果的影响如表 1 所示.

由表 1 可以发现,当 Y 层神经元个数较少以及匹配度门限值较低时,识别率偏低.当 Y 层神经元个数少于朝向个数,训练时某个神经元要兼顾学习至少两种以上的朝向特征,结果两种以上的朝向特征就会共同占有该神经元,测试时两种以上的朝向特征使该神经元被激活,然而该神经元未能识别是哪一种朝向,只能按训练时的先后

顺序指定输出的朝向,从而导致识别出现错误.当匹配度门限值偏低时,两个朝向特征本不相似,但是很容易越过门限值的门槛,使得Y层神经元判断错误,致使识别率偏低.门限值为1,神经元个数为5时,由于门槛高,神经元数目少,以致识别出现混乱.可见,合理设置神经元个数和匹配度门限值,才能取得最佳结果.笔者通过试验验证了训练次数对笔者识别率没有影响.

表 1 识别结果

Tab.1 The results of recognition %

Y 层神经 元个数	匹配度门限值				
	0.6	0.7	0.8	0.9	1.0
1×1	20	20	20	20	20
2×2	60	80	80	80	60
1×5	60	100	100	100	60
3×3	60	100	100	100	100
4×4	60	100	100	100	100

为了验证图像形状对特征的稳定性以及笔者方法对其它数据库的识别效果,笔者选用 CMU PIE 数据库和 MIT 人脸库的部分图像进行测试. PIE 人脸库采集 68 个人不同姿态、光照的4万多张图像,而 MIT 数据库包含 16 个人不同姿态、光照和大小的 2 593 张图像. 从这 2 个数据库中各选择 680 张 5 种不同朝向的图像作为测试集,部分图像分别如图 7(a)、(b)所示. 通过试验可得X到Y的权重,如图8所示.



图 7 不同姿态和光照的图像

Fig.7 Images of different pose and illumination



图 8 X 到 Y 的权重

Fig.8 Weights from X to Y

从图 8 可知,由两组不同规范的图片得到的特征只有 5 种,这说明人眼的位置在水平方向上倾斜较小角度不会影响结果,因为眼睛部位约在图片的 1/6 到 1/3 处. 试验显示,对 PIE 和 MIT 数据库测试所得结果相同,识别率为 100%,与笔者采集图像进行试验所得结果一致.

3.2.2 4 种方法识别效果的对比

为了更好地体现识别的效果,选择笔者采集的 150 张图像作为训练集,从 CMU PIE 和 MIT 人

脸库中各选择 680 张 5 种不同朝向的图像作为测试集. 使用笔者特征提取方法,利用 BP、LVQ 和 PNN 神经网络与发育网络的识别效果进行对比,图 9 为各网络参数对识别率的影响示意图.

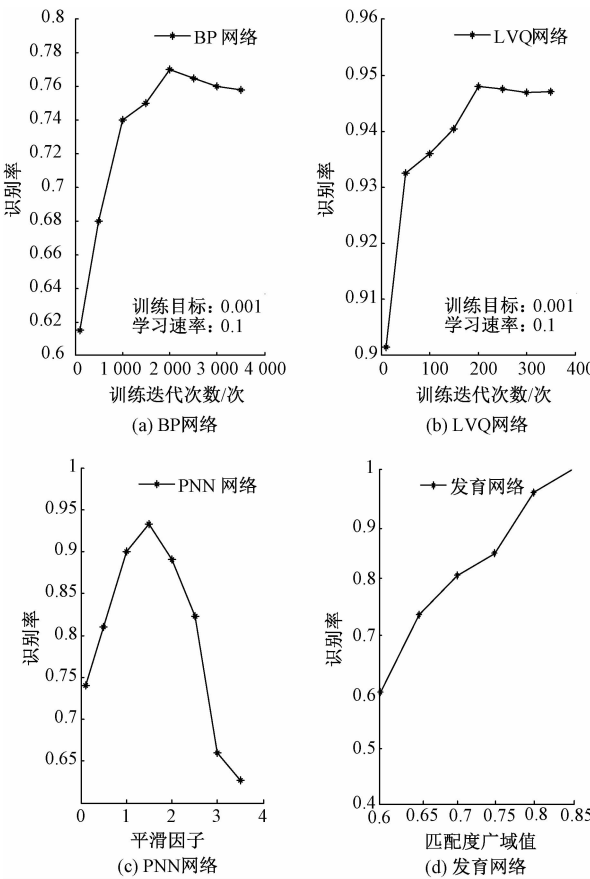


图 9 网络参数对识别率的影响

Fig.9 Effect of network parameter on recognition rate

由图 9 可得出各个网络的最优参数:BP 网络训练迭代次数为 2 000;LVQ 网络训练迭代次数为 200;BP、LVQ 的训练目标和学习速率都分别为 0.001、0.1. 而 PNN 只有一个参数—平滑因子,其最优值为 1.5. 每个网络都选择最优参数进行试验,每种方法都运行 15 次,结果见表 2.

表 2 4 种方法识别率的对比

Tab.2 The comparison of recognition rate for four methods %

方法	最小值	最大值	平均值
BP 神经网络	70.15	85.59	76.32
LVQ 神经网络	92.35	97.94	94.12
PNN 神经网络	90.44	96.00	93.09
发育网络	100.00	100.00	100.00

发育网络通过训练学习将人脸朝向的特征信息存储在神经元的权值,是对知识的记忆;测试的时候,根据输入与已学特征的匹配度来判断正误,输出无需期望值作比较,在线学习. 相对于以比

较、聚类、统计方法为原理的神经网络,发育网络的准确性高,鲁棒性好,容错力强.

4 结论

笔者模拟了人类大脑发育的功能以及大脑对物体的识别能力,提出了基于发育网络的人脸朝向识别新方法. 试验结果表明,发育网络不仅能很好地学习物体的特征,而且还能很好地识别不同光照条件下的人脸朝向. 网络的鲁棒性以及对环境适应能力非常强,是其它网络无法比拟的. 本文中人脸朝向的分类较少,增加不同人脸朝向将是未来的研究方向,应引起研究者的关注.

参考文献:

[1] 王蒙军. 人脸朝向识别的脉冲耦合神经网络分析 [C]//天津市生物医学工程学会第三十二届学术年会论文集. 天津: 生物医学工程与临床出版社, 2012: 40 – 40.

[2] LI D Q, PEDRYCZ W. A central profile-based 3D face pose estimation [J]. Pattern recognition, 2014, 47(2): 525 – 534.

[3] PENG X, HUANG J Z, HU Q, et al. From circle to 3-sphere: head pose estimation by instance parameterization [J]. Computer vision and image understanding, 2015(136): 92 – 102.

[4] DEMIRKUS M, PRECUP D, CLARK J J, et al. Hierarchical temporal graphical model for head pose estimation and subsequent attribute classification in real-world videos [J]. Computer vision and image understanding, 2015(136): 128 – 145.

[5] DORNAIKA F, CHAHLA C, KHATTAR F, et al. Discriminant sparse label-sensitive embedding: appli-

cation to image-based face pose estimation [J]. Engineering applications of artificial intelligence, 2016 (50): 168 – 176.

[6] 陈锐,李辉,侯义斌,等. 由人脸朝向驱动的多方向投影交互系统 [J]. 小型微型计算机系统, 2007, 28(4): 706 – 709.

[7] 张彤,卢雯雯,肖南峰. 基于 BP 网络的人脸朝向识别方法 [J]. 重庆理工大学学报 (自然科学版), 2010, 24(6): 62 – 65.

[8] 谭乐平,文军. 基于 LVQ 神经网络的人脸朝向识别方法 [J]. 湖北民族学院学报 (自然科学版), 2012, 30(4): 387 – 389.

[9] 朱宇鑫,顾军,刘晨思. 基于概率神经网络 (PNN) 的人脸朝向识别 [J]. 科技风, 2014, 63(7): 80 – 81.

[10] 朱宇鑫. 基于神经网络的人脸朝向识别问题研究 [D]. 苏州: 苏州大学机电学院, 2014.

[11] WENG J Y, MCMLELLAND J, PENTLAND A, et al. Autonomous mental development by robots and animals [J]. Science, 2001, 291(5504): 599 – 600.

[12] WENG J Y. Natural and Artificial Intelligence: introduction to computation brain-mind [M]. Michigan: BMI Press, 2012.

[13] 蔡金收,陈铁军,郭丽. 一种基于投票极限学习机的人脸识别混合算法 [J]. 郑州大学学报 (工学版), 2016, 37(2): 37 – 41.

[14] 史峰,王小川,郁磊,等. MTALAB 神经网络 30 个案例分析 [M]. 北京: 北京航空航天大学出版社, 2010.

[15] 尹方平. 复杂背景下的人眼检测方法 [J]. 计算机仿真, 2009, 26(10): 225 – 228.

[16] 刘昊,方雯逸. 基于 BP 神经网络的人脸朝向分类的新思路 [J]. 计算机科学, 2012, 39(11A): 366 – 369.

Study of Face Orientation Recognition Based on Development Network

WANG Dongshu, TAN Dapei, WEI Xiaoqin

(School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Based on the characteristic of face orientation, position and the light background in face recognition, a new method of face orientation recognition based on development network is proposed. The characteristic of human’s eye was very prominent, so the position of eyes was chosen as the face orientation feature vector. And the development network model was used to recognize human’s face orientation in the different light background images. The result showed that this method could effectively solve the difficult problem of face orientation recognition under varying illumination conditions by comparing with the test results of other methods, which was fast, stable and effective. The recognition rate was as high as 100%.

Key words: development network; face orientation recognition; light background; feature vector; recognition rate