

基于仿生模式识别的多镜头人脸身份确认系统研究

王守觉^{1,2}, 徐 健¹, 王宪保², 覃 鸿¹

(1. 中国科学院半导体研究所神经网络实验室, 北京 100083; 2. 浙江工业大学智能信息系统研究所, 杭州 310014)

摘 要: 提出了一种利用仿生模式识别原理以多镜头信息融合的人脸身份确认方法. 讨论了以多镜头与多次采样建立仿生模式识别多权值神经网络的基本理论. 并介绍了以三个镜头五次采样作训练样本的实验系统, 以及人脸预处理过程步骤. 实验结果表明在确保无误认的前提下, 正确确认率达 96%, 漏认率(即拒认率)为 4%. 实验中并对比显示了增加神经网络复杂度提高识别效果的作用.

关键词: 仿生模式识别; 人脸识别; 身份确认; 数据融合; 神经网络

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2003) 01-0001-03

Multi-Camera Human-Face Personal Identification System Based on the BIOMIMETIC PATTERN RECOGNITION

WANG Shou-jue^{1,2}, XU Jian¹, WANG Xian-Bao², QIN Hong¹

(1. Lab of Artificial Neural Networks, Institute of Semiconductors, CAS, Beijing 100083, China;

2. Research Institute of Intelligent Information System, Zhejiang University of Technology, Hangzhou, 310014)

Abstract: A Multi-camera Human-face Personal Identification method, based on the BIOMIMETIC PATTERN RECOGNITION, had been proposed. The basic theory of neural networks with MULTIPLE-WEIGHT NEURON, in Multi-camera Human-face Personal Identification System Based on the BIOMIMETIC PATTERN RECOGNITION, had been discussed. A three-camera system takes five combined pictures as learning samples in experiments. The basic course of human-face picture preprocessing had been described. The results of experiments showed the correct identification rate as 96%, rejection rate as 4%, and no mistake in identification. Experiments showed that increasing the complexity of neural networks makes better identification result.

Key words: pattern recognition; human-face identification; personal verification; data fusion; neural networks

1 引言

自美国世贸大厦被毁以来, 国际反对恐怖活动的安全措施日益显得重要. 其中对人员身份的确认, 防止不法人员混入重要会议、重要场地也成为安全措施的重要组成部分^[1]. 与视网膜、指纹、声音等识别途径相比, 人脸自动识别被认为是身份确认途径中最方便, 但却错误率太高的一种手段^[2]. 而人们在社会活动中, 古今中外最常用最可靠的确认相识熟人的方法就是靠用眼睛看脸. 因而, 从信息含量来看人脸自动识别用以确认身份是完全可能提高到有使用价值高度正确率的.

作者提出的“仿生模式识别”(拓扑模式识别)就是强调用“认识”模式取代传统的模式“分类”与划分(关于仿生模式识别的研究已在另文发表^[3], 本文仅作简要介绍), 本文的目的将是利用“仿生模式识别”基本原理研究一种确认正确率较高的三镜头摄入的神经网络人脸识别身份确认系统.

2 身份确认与人脸识别

传统说的“识别”是多类事物的区分. 如果针对人脸则就

是甲、乙、丙、丁……等若干人中识别出谁是甲、谁是乙……因而, 它要求判别被识别人最像甲、乙、丙、丁……中的哪一个. 而身份确认则是要求判别被识别人是否确是某甲, 回答只有“是”或“非”. 而“仿生模式识别”正是强调“认识”而不是“区分”实物, 因而它更适合用于身份确认.

3 仿生模式识别简介

传统模式识别中把不同类样本在特征空间中的最佳划分作为目标, 而仿生模式识别则以一类样本在特征空间的分布的最佳覆盖作为目标, 以二维空间的情况示意如图 1.

图 1 中, 三角形为要识别的样本, 小圆圈和十字形为与

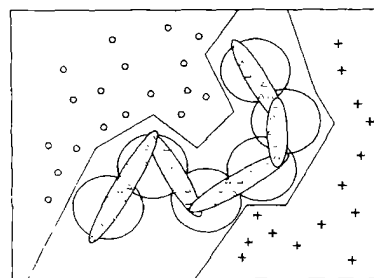


图 1

收稿日期: 2002-09-23; 修回日期: 2002-10-23

基金项目: 国家自然科学基金(No. 60135010)

三角形不同类的两类样本,折线为传统 BP 网络模式识别的划分方式,大圆为 RBF 网络的划分方式(等同于以模板匹配的识别方式),细长椭圆形构成的曲线代表仿生模式识别的“认识”方式.因而,仿生模式识别把分析特征空间训练样本点间的关系作为基点,而特征空间中样本分布的连续性规律为此提供了可能性.仿生模式识别用的方法是“高维空间复杂几何形体覆盖识别方法”,是在特征空间中研究某类样本的分布状况而加以合理覆盖,从而来“认识”某类样本.

4 用神经网络的多镜头人脸信息融合及人脸确认

针对同一目标(同一人脸)多镜头采集的信息经预处理后映射到特征空间的“像”是多个样本点(样本矢量).这多个样本点的分布,正部分地反映了该人脸在特征空间中的“像”的分布状况.根据这多个样本点来决定特征空间中此类样本的覆盖范围,是人脸确认的有效步骤.而一个多权值神经元正是可根据多个权值矢量作成特征空间中复杂形状的覆盖区域^[4],一个多权值神经元性能的通用表达式应为:

$$Y=f[(W_1, W_2, \dots, W_m, X)]$$

式中 W_1, W_2, \dots, W_m 为 m 个权值矢量; X 为输入矢量; 为多权值矢量神经元^[5]决定的计算函数(多个矢量输入,一个标量输出); f 为非线性转移函数;

在身份确认的场合,回答问题的结果是“是”或“非”,因而可用阶跃函数,即:

$$f(x)=\begin{cases} 1 & \text{当 } x \leq k \\ 0 & \text{当 } x > k \end{cases}$$

设特征空间是 n 维实数空间 R^n , 矢量函数方程

$$(W_1, W_2, \dots, W_m, X) = k,$$

式中 k 为常数可视为由 W_1, W_2, \dots, W_m 等 m 个权值矢量所决定的在特征空间 R^n 中 X 矢量的一种轨迹,此轨迹为 R^n 中的 $(n-1)$ 维超曲面(或超平面),它把 R^n 分成两个部分^[3].在它一侧的部分 $(W_1, W_2, \dots, W_m, X) < k, Y=1$,在它另一侧的部分 $(W_1, W_2, \dots, W_m, X) > k, Y=0$.

如果使 $(W_1, W_2, \dots, W_m, X) = k$ 是一个封闭超曲面,则就在特征空间中形成了一个有限覆盖区域.输入矢量 X 若在此覆盖区内则神经元输出 $Y=1$,在此覆盖区外则 $Y=0$.

设多镜头(m 个镜头)采集到的样本映射到特征空间的“像”为矢量 S_1, S_2, \dots, S_m ; 而权值矢量 W_1, W_2, \dots, W_m 是由 S_1, S_2, \dots, S_m 来决定的.亦即:

$$W_1 = w_1(S_1, S_2, \dots, S_m)$$

$$W_2 = w_2(S_1, S_2, \dots, S_m)$$

$$\dots \dots \dots$$

$$W_m = w_m(S_1, S_2, \dots, S_m)$$

因而 $(W_1, W_2, \dots, W_m, X)$ 必可改写成

$$(S_1, S_2, \dots, S_m, X) = k$$

用仿生模式识别原则^[3]和高维空间几何分析方法^[6],找出合适的(针对具体识别问题的)函数使 $X = S_1$ 或 $X = S_2, \dots$ 或 $X = S_m$ 时上式都能成立,(参考文献[3]中针对全方位旋转实物目标识别的神经网络算法).

如此,该神经元覆盖的特征空间状况是由样本 S_1, S_2, \dots, S_m 所共同确定,因而它本身就起到了将 m 个镜头所采集

的 S_1, S_2, \dots, S_m 样本所提供的人脸信息相融合的作用.

上述是一个神经元的情况.以下假设有 p 个神经元,每个神经元的权值各由一次采集的样本群确定,则各个神经元的输出分别是 Y_1, Y_2, \dots, Y_p .

$$Y_1 = f[(S_{11}, S_{12}, \dots, S_{1m}, X)]$$

$$Y_2 = f[(S_{21}, S_{22}, \dots, S_{2m}, X)]$$

$$\dots \dots \dots$$

$$Y_p = f[(S_{p1}, S_{p2}, \dots, S_{pm}, X)]$$

今输入矢量也为多镜头所采集,因而 X 也有 m 个矢量,以 X_1, X_2, \dots, X_m 代表,则

$$Y_{11} = f[(S_{11}, S_{12}, \dots, S_{1m}, X_1)]$$

$$Y_{12} = f[(S_{21}, S_{22}, \dots, S_{2m}, X_1)]$$

$$\dots \dots \dots$$

$$Y_{1p} = f[(S_{p1}, S_{p2}, \dots, S_{pm}, X_1)]$$

$$Y_{21} = f[(S_{11}, S_{12}, \dots, S_{1m}, X_2)]$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$Y_{2p} = f[(S_{p1}, S_{p2}, \dots, S_{pm}, X_2)]$$

$$Y_{31} = f[(S_{11}, S_{12}, \dots, S_{1m}, X_3)]$$

$$\dots \dots \dots$$

$$Y_{mp} = f[(S_{p1}, S_{p2}, \dots, S_{pm}, X_m)]$$

而人脸确认时的判别函数为:

$$R = \sum_{j=1}^m \sum_{i=1}^p Y_{ji}$$

式 R 代表了判别输入矢量 X_1, X_2, \dots, X_m 中有无一个落在特征空间中 p 个神经元所覆盖范围的并的范围中,这 p 个函数神经元覆盖范围的并,可形成特征空间中非常复杂的几何形体,能满足人脸确认的要求.

5 人脸确认实验系统的实验过程和结果

神经网络多镜头人脸确认实验系统采用了三个镜头(即 $m=3$),确认一个人脸采用五个神经元($p=5$). (在用神经网络硬件 CASSANN- 神经计算机实现中用了 5 个神经元合成一个函数神经元,总共每确认一个人脸实用 25 个神经元)实验系统的框图如图 2.

由三个摄像头组成一组摄像采集头,采集的人脸以 bmp

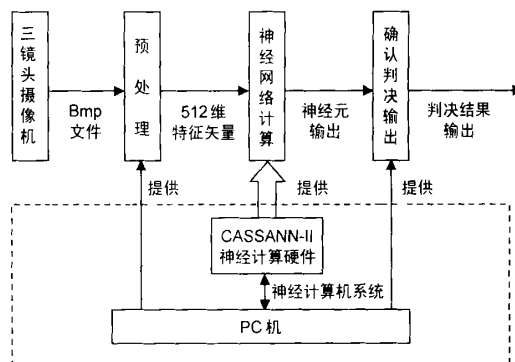


图 2

文件输入 PC 机,该 PC 机同时作为神经计算机硬件的宿主机,同时提供预处理和确认判决及输出等后处理用。PC 机与神经计算硬件和软件组成神经计算机系统。以下对预处理过程作简要说明。各个摄像头各个时间(包括神经网络训练时期、确认时期)所采集的图像均通过同一预处理过程映射到特征空间中(512 维每维 16 位浮点数的特征矢量)。其具体计算步骤为:

- (1) 确定人脸所在位置;
- (2) 用左右匹配的方法将倾斜人脸转正;
- (3) 定出眼睛精确位置;
- (4) 缩放人脸使眼睛间距为预定常数;
- (5) 定出数个关键点位置;
- (6) 以各关键点为核心向四面八方放射方向各点灰度的微商作为特征矢量的各维,以一定的次序排列成 512 维矢量;
- (7) 将 512 维矢量归一化,以减少灰度总体直流分量的影响(即光线强弱的影响)。

初步实验是由 5 个人进行,每个人先在组合摄像头前采集五次像,共得五组,每组 3 个头像。以多权值神经网络算法(多权值神经元已被受理申请专利,申请号 02122638.5)训练构筑神经网络(对应于 5 个 函数神经元, 函数神经元的算法,在 CASSANN-II 神经计算机仿生模式识别系统软件中配套提供),其中以 k 值为可调距离常数。(参见文献[3]中距离常数 k)。然后以本人和非本人作多次交叉确认试验。计应确认为本人的试验次数为 25 次,应确认为非本人的试验 100 次。在不同距离常数 k 值下所得到的正确确认率、误认率和漏认率分别如表 1 所示。

表 1 用 5 个 函数神经元的实验结果

k^2	正确确认率	误认率	漏认率
350	20/25 = 80 %	0	5/25 = 20 %
375	22/25 = 88 %	0	3/25 = 12 %
400	24/25 = 96 %	0	1/25 = 4 %
425	24/25 = 96 %	1 %	1/25 = 4 %
450	24/25 = 96 %	4 %	1/25 = 4 %
475	24/25 = 96 %	7 %	1/25 = 4 %
500	25/25 = 100 %	12 %	0
525	25/25 = 100 %	21 %	0

表 2 只用 1 个 函数神经元的实验结果

k^2	正确确认率	误认率	漏认率
400	62 %	0	38 %
425	66 %	1/500 = 0.2 %	34 %
450	70 %	5/500 = 1 %	30 %
475	71 %	11/500 = 2.2 %	29 %
500	76 %	23/500 = 4.6 %	24 %
525	80 %	46/500 = 9.2 %	20 %
550	81 %	63/500 = 12.6 %	19 %
575	86 %	90/500 = 18 %	14 %

为说明用多个 函数神经元增加覆盖区复杂度对提高确认率效果的作用,同一实验方法而简化神经网络覆盖空间的复杂度,确认一个人脸只用一个 函数神经元(即 $P =$

1);作本人和非本人多次交叉确认实验,共计应确认为本人的试验次数为 100 次,应确认为非本人试验次数为 500 次。实验结果如表 2。

6 实验结果讨论与结论

由上述实验结果可以看到:

(1) 如果选用 $m = 3, p = 5, k^2 = 400$,则具有 96 % 的正确确认率而错误确认次数为零,虽然有 4 % 的漏认,则可以再确认一次来弥补,因而,是离使用价值不远的结果,这说明该技术方向是可取的。

(2) 增加 函数神经元的数量(即 p 的值),即增加神经元覆盖范围的复杂度,可以改善确认效果。

(3) 在增加 p 的同时应适当减小距离常数 k 值,以确保误认率为零。

本文所介绍的实验中的前处理过程是基于对人脸各关键点周围部分作等同处理的简单办法,如果突出重点并加入部位相互间关系的考虑将可进一步提高效果,特别是提高对不同表情的泛化能力。所以本方向前景很好但还有待于在前处理方法上作进一步深入研究。

参考文献:

- [1] 冀复生. 关于美国信息技术发展情况的一些看法[N]. 科技日报, 2002-07-18(特别关注)。
- [2] A K Jain, R Bolle, S Pankanti. Biometrics: Personal Identification in Network Society [M]. Kluwer Academic Publishers, 1999.
- [3] 王守觉. 仿生模式识别(拓扑模式识别)——一种模式识别新模型的理论与应用[J]. 电子学报, 2002, 30(10): 1417 - 1420.
- [4] 王守觉等. 通用神经网络硬件中神经元基本数学模型的讨论[J]. 电子学报, 2001, 29(5): 577 - 580.
- [5] 专利申请文件, 申请号 02122638.5.
- [6] 王守觉等. 人工神经网络的多维空间几何分析及其理论[J]. 电子学报, 2002, 30(1): 1 - 4

作者简介:

王守觉 男, 1926 年生于江苏, 历任中国科学院半导体研究所室主任、副所长、所长等职, 1980 年当选中国科学院院士, 现为半导体神经网络实验室负责人, 他是我国半导体器件与微电子奠基人之一, 为研究解决我国发展两弹一星所需高速计算机的半导体器件与集成电路作出过重要贡献, 在国家“八五”、“九五”科技攻关中, 他承担了半导体神经网络的实现和应用技术的攻关工作, 研制了我国最早可商品化的神经计算机 CASSANN-I、CASSANN-II 等, 提出了新的“仿生(拓扑)模式识别”理论和“高维空间复杂几何形体覆盖”识别方法, 使平面上各方向的钢体实物模型正确识别率达到了前所未有的高度, 经数千次自动测试结果统计, 全方位不同角度的正确识别率达到 99.75 %, 而无 1 次错误识别。为模式识别开辟了一个崭新的理论与实现的新途径, 他是《电子学报》和《CTE》主编, 中国计算机学会 CAD 与图形学专业委员会名誉主任。e-mail: wsjue@red.semi.ac.cn

徐 健 中国科学院半导体研究所神经网络实验室博士研究生, 从事神经网络人脸识别硬件与算法研究。

王宪保 浙江工业大学智能信息系统研究所硕士研究生, 从事神经网络人脸识别研究。

覃 鸿 中国科学院半导体研究所神经网络实验室在职研究生, 从事神经网络软、硬件研究工作。