

基于神经网络的不变性模式识别的人脸识别系统

袁友伟 金可音 湛含辉
(株洲工学院计算机系, 株洲 412008)
E-Mail: YuanYouwei@21cn.com

摘要 文章提出了一种优化的多层神经网络, 能完成对图像不变性模式识别。基于此多层神经网络开发了人脸识别系统。仿真实验表明该人脸识别系统, 具有较高的识别速度、准确率、容错性和鲁棒性, 并且基本解决了开发实用化的人脸识别系统所面临的问题, 即模式识别不变性的问题。

关键词 人脸识别 模式识别 多层神经网络

文章编号 1002-8331-(2002)16-0095-02 文献标识码 A 中图分类号 TP183

Face Recognition System Based on Multilayer Neural Networks for Motion Invariant Pattern Recognition

Yuan Youwei Jin Keyin Zhan Hanhui

(Department of Computer Science & Technology Zhuzhou Institute of Engineering Zhuzhou 412008)

Abstract: The paper brings forward a new multilayer neural networks for motion invariant pattern recognition. A new face recognition system based on multiplayer neural network is also presented. Simulations show that it not only has upper recognition rapidity accuracy tolerance and rubustness but also meet most of the demands for developing practical face recognition systems basically.

Keywords: Face recognition, Pattern recognition, Multilayer neural networks

1 引言

不变性模式识别研究的目标是建立一个灵活的图象识别系统, 所谓灵活, 即系统不受模式图像位置、大小和方向的约束。此外, 这种系统应具有一定程度的容错性, 即允许轻度的形状变化、扭曲和噪声的存在。在模式识别中, 使识别对运动具有不变性是一个极其重要的课题。

人脸识别是一个跨学科富有挑战性的前沿课题, 它跨越了图像处理、模式识别、计算机视觉和神经网络等研究领域。人脸识别难度较大, 主要难在于人脸具有各种变化的相式刚体, 由于人脸不仅存在各种变形, 而且和皮肤之间是平缓过渡。人脸表情丰富; 人脸随年龄增长而变化; 人脸所成图像受光照、成像角度及成像距离等影响。随着神经网络研究的发展, 国内外提出了许多基于神经网络的人脸识别方法, 在识别效果的准确率、容错性、鲁棒性等方面取得了一定的进展。已有学者提出用神经网络来建立运动具有不变性的模式识别问题, 但其所得到的网络或是识别的精度不高, 或是学习时间过长, 有时还不能完全用神经网络来作研究。现有的人脸识别方法存在当人脸图像平移、旋转、缩放和形状变化、扭曲和噪声存在时, 识别准确率不高或“拒识”。

高阶神经网络, 通过约束其权值以产生各种变换下的大致不变性输出, 就克服了上述缺点, 可以有效地进行不变性模式识别和分类, 即使各种变换同时发生也可以成功地识别模式。为此该文构造了一种优化的多层神经网络, 当人脸图像平移、

旋转、缩放和形状变化、扭曲和噪声存在时, 均能准确地识别。该系统不仅具有较高的识别速度、准确率、容错性和鲁棒性, 并且基本解决了开发实用化的人脸识别系统所面临的问题, 既模式识别不变性的问题。

2 最优化神经网络

笔者构造的能保持模式不变性的多层神经网络结构如图1所示, 使 S_i 代表模式 $i=1, 2, \dots, N$ 分别表示 N 个像素屏幕上的每一个像素。像素 i 是网格或规则方框网格内的一点, 其位置由向量 $r(i)$ 表示。现在考虑有这样的一个自然的网络, 它具有单层权值, 输入是 $S_{i1}, S_{i2}, \dots, S_{iP}$ (P 是网络阶数)。

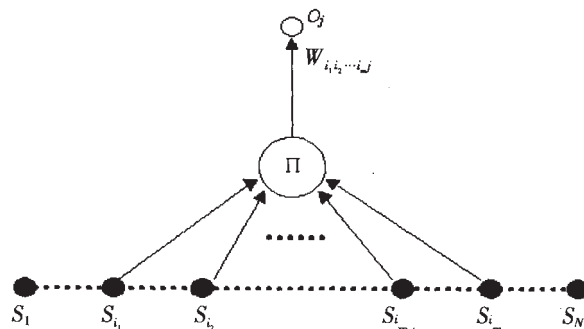


图1 单层高阶网络的连接图

经点 j 的输出:

基金项目: 湖南省自然科学基金项目 (编号 01JJY3015)

作者简介: 袁友伟 (1966-) 男, 现为株洲工学院副教授, 主要从事图像处理、神经网络研究, 发表论文 30 篇。金可音, 博士, 副教授, 主要从事神经网络研究。湛含辉, 教授, 博导, 主要从事图像处理研究。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

计算机工程与应用 2002.16 95

$$O_j = f(\sum_{i_1} \sum_{i_2} \dots \sum_{i_p} w_{i_1 i_2 \dots i_p j} S_{i_1} S_{i_2} \dots S_{i_p})$$

其中 f 是正式神经元的激励函数 (这里将使用逻辑函数 $f(x) = 1/[1 + \exp(-x)]$)。

设 ξ 是像素网格上的一组坐标变换, 在这组变换下希望 O_j 不变。经过 ξ 的变换 O_j 将保持不变, 前提是网络权值必须满足以下方程:

$$w_{i_1 i_2 \dots i_p j} = w_{k_1 k_2 \dots k_p j} \quad (1)$$

经网格向量平移变换后欲获得输出不变的二阶网络, 必须设置 $w_{i_1 i_2 j} = w_{k_1 k_2 j}$, 即这些线段平行且长度相同。

高阶网络顶层结点的输出可记为:

$$O_j = f(\sum_{h_j} W_{h_j} \xi_{h_j} \dots \xi_{h_p} S_{h_j} \dots S_{h_p}) \quad (2)$$

这样就可以用权 w_{hj} 来有效地处理单层前馈网络 (感知器) 输入是:

$$I_h = \sum_{\xi_1 \dots \xi_p} S_{i_1} S_{i_2} \dots S_{i_p} \quad (3)$$

它代表图像的不变性特征, 网络使用这种特征加以分类。

3 人脸识别系统的基本结构

该文在研究过程中, 以 Delphi 为开发工具, 结合关系型数据库技术, 在 Windows 98+ Windows NT LAN 环境下, 开发了一套基于多层神经网络的人脸识别实验系统。该系统主要由预处理、特征抽取、基于神经网络的分类器和数据库四个部分组成。如图 2 所示。其中特征抽取与分类器是解决人脸识别问题的关键。

预处理: 完成对人脸图像的大小归一化、消除噪声、灰度归一化等, 可视具体应用而选用, 以便在同一条件下完成训练和识别。

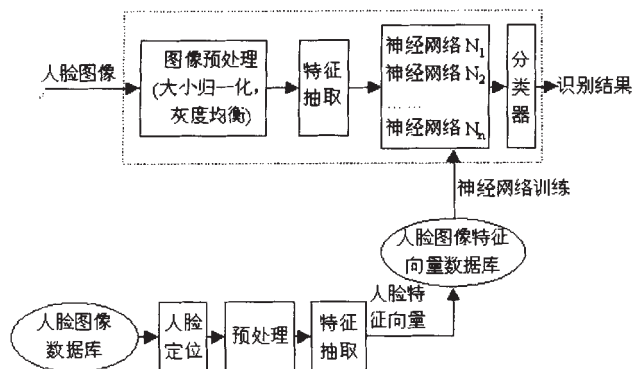


图 2 基于多层神经网络的人脸识别系统

特征抽取: 完成提取人脸特征, 如何提取稳定和有效的特征是识别系统成败的关键, 构造人脸特征的向量有: 人脸的虹膜中心、内眼角点、外眼角点、鼻尖点、鼻孔点、耳屏点、耳下点、口角点、头顶点、眉内点、眉外点等标志点。

人脸是一个很特殊的生物模式, 在人脸的照片中, 人们可以发现人脸具有明显的灰度特征, 如眼睛的灰度、鼻孔的灰度、嘴巴中间的灰度值都比周围低; 同时, 人脸还具有很强的梯度信息, 如眼睛、嘴巴、鼻子等部位的边缘信息比人脸的其他部位更丰富。当原始特征样本的维数很大 (比如人脸图像矢量为 $92 \times 112 = 10304$ 维) 时^[6], 在如此高维的原始特征空间内, 用传统的扰动法和补空间法直接求解最优鉴别矢量集几乎是不可行的。

Karhunen—Loeve 提出了主成分分析法 (Principal component analysis 简称 PCA), 降低原始特征样本的维数以达到减少计算的复杂度。

首先根据图像的复杂程度, 将输入图像 N , 划分为 $S_x \times S_y$ 区域块, 表面点图像光亮度即像素点 $I_p(x, y)$ 为:

$$I_p(x, y) \approx \bar{I}_p(x, y) + \sum_{i=1}^{PC_1} \alpha_i \phi_i(x, y) \quad (4)$$

其中 $\bar{I}_p(x, y) = 1/N \sum_{i=1}^N I_p^i(x, y)$ 是该处图像光亮度的平均值, 平均值可以减少误差, $\phi_i(x, y)$ 是数据集的主成分, 且 $PC_1 \leq N$ 主成分的系数为:

$$\alpha_i = \sum_x \sum_y I_p(x, y) - \bar{I}_p(x, y) \phi_i(x, y) \quad (5)$$

数据集的特征向量由下式可得,

$$R_1 \phi_i = \lambda_i \phi_i \quad (6)$$

其中 R_1 是图像 N 的自动协方差, λ 是特征值, ϕ 是主成分的特征变量。

基于神经网络的分类器设计:

在抽取有效的特征后, 关键的问题是设计出具有良好区分度和鲁棒性的识别器。该文设计了基于多层神经网络的分类器。对于任意一个人脸测试图像, 可根据网络输出层输出矢量的最大分量分类。人脸图像矢量的维数 N 通常比较大, 而训练样本数 K 通常比较小。用此神经网络分类器, 可以从特征压缩着手, 压缩输入矢量的维数, 并适当地选择隐含层的神经元。为了加快网络训练的收敛速度, 可以输入矢量作标准化处理, 并给各连接权值适当地赋予初值。

KL 变换是在最小均方差准则意义下获得数据压缩的最佳变换, KL 变换的产生矩阵常取为训练样本集的总体工作面矩阵或类间散布矩阵, 该文将类间相关矩阵取作为 KL 变换的产生矩阵:

$$R_b = \sum_{i=1}^p m_i m_i^t \quad (7)$$

这里 $m_i (i=1, \dots, p)$ 是训练样本集中各类模式样本的均值矢量。

产生矩阵 R_b 的秩一般为 p , 由奇异值分解定理, 可以得到 R_b 的 p 个正交归一特征向量, 对于任意 N 维的人脸图像矢量, 通过向特征向量投影可将人脸图像的特征维数从 N 维压缩到 p 维。大大降低了计算复杂度, 消除了数据冗余。

三阶网络可用于识别平移、旋转和比例缩放后的模式。如果设置任意两个权 $w_{i_1 i_2 i_3 j}$ 和 $w_{k_1 k_2 k_3 j}$ 相等, 那么就可以获得相对于所有三种变换的近似不变输出。三阶网络 (TON), 角度变化有一定限制 (即 $\omega \in [\pi/144, \pi/6]$), 网络具有一或二层权, 训练过程中, 当输入第 i 个数字时, 其对应的第 i 个输出为 1, 其余输出结点为 0。这里用提取器——分类器模式实现此网络, 用 GBP 训练网络。

人脸图像数据库: 实验测试所使用的图像数据库包括 400 副图像。

4 仿真实验

实验测试所使用的图像数据库包括 400 副图像 (共 80 人, 每人 5 副图像, 这些图像在训练样本里都没有出现过), 得到

(下转 249 页)

符合强度条件。

4.3 设计可靠性分析

经典有限元是在力学模型基础上进行近似计算的方法。近似计算通过结构的离散化模型,通过单元的划分,单元分析,整体分析,从而求出各节点的位移和应力。这种方法能处理确定量,不能处理概率量。随机有限元是 80 年代初发展起来的处理随机现象的分析工具,它指的是引进相应的确定性分析方法,显式处理任意不定参数,并求其对所求物理量的响应。此法的特点在于将随机参数空间离散化,并将随机输入变为一个方差矩阵依赖于有限元网格划分的随机变量。具体的随机有限元理论和可靠性理论参阅其它有关文献。

车身实际变形中包括了前后悬架的弹性变形引起的整车结构的刚体位移。在计算中,剔除了这一变形。把悬架处理成大刚性杆,放松在铰接点处的自由度。后悬架钢板弹簧处采用铰支,前悬架对不同工况采用不同的支撑方式处理。由于本身的侧倾角很小,这样作对内力的影响不大。其它简化方法同上述的有限元分析,建立模型与上述有限元计算模型一致。经计算,车身结构的最小可靠度为 0.986453。

4.4 优化设计

具体的优化理论参阅其它有关文献。在这里,只考虑在扭转工况下进行截面尺寸的优化问题,选用截面空间的尺寸为优化设计变量,因而对车身作了部分简化处理。简化处理措施同上述的有限元分析和可靠性分析,后悬架钢板弹簧处采用铰支,前悬架对不同工况采用不同的支撑方式处理。由于本身的

侧倾角很小,这样作对内力的影响不大。在靠的很近的接头处,采用节点合并。同时,在刚架系统中,用等截面替换变截面,用直杆代替曲杆,用刚接头代替螺栓、铆接结构。通过优化计算,车身质量为 1870kg。质量减轻 10.9%,效果较好。

5 结论

文章研究了有限元法在并行产品设计中的应用,提出了典型产品的并行设计过程与方法。典型产品设计过程中,产品的静力学、动力学有限元分析,有限元优化计算、可靠性计算是很重要的过程。在应用实例中,研究了汽车车身并行虚拟样机方法设计和分析方法,加快了开发周期,节约了高额开发费用,降低了研发成本。(收稿日期:2002 年 6 月)

参考文献

1.李伯虎,柴旭东.航天 CIMS 研究、实践与发展[J].制造业自动化,2001-01
2.吴祚宝,肖田元.虚拟制造环境下有限元分析的应用[J].清华大学学报,2000-07
3.(德)约瑟夫·萧塔纳著.祁国宁译.制造企业的产品数据管理[M].机械工业出版社,2000
4.李少波,谢庆生,楚甲良.汽车车身设计方法探讨[J].贵州工业大学学报,2000-01
5.楚甲良,谢庆生.基于随机有限元的汽车车身可靠性分析[J].贵州工业大学学报,2000-09

(上接 96 页)

400 组人脸特征向量,这些特征向量经过模糊编码后作为多层神经网络的输入,用于训练网络的长期记忆权和映射权。人脸图像经灰度化、分割处理后用双线性插值方法归一化到 24×28 像素大小。图 3 是测试数据库里 8 个人,这些人脸图像分别在不同时间、不同光照、不同头部角度、不同人脸表情(闭/睁眼,笑/严肃)和不同于测试系统识别成功率的 8 个人的 18 幅不同表情、细节、方位的人脸图像,人脸细节(有无眼睛),且图像平移和旋转一定角度的条件下摄制得到。运用该文的基于多层神经网络的人脸识别系统识别成功率为 95%,而应用 BP 网络的识别成功率为 80%。



图 3 用于测试系统识别成功率的 8 个人的 18 幅不同表情、细节、方位的人脸图像

5 结论

该文构造了基于多层神经网络的人脸识别系统,当人脸图像平移、旋转、缩放和形状变化、扭曲和噪声存在时,均能准确地识别。而且使用方便,在普通 PC 机上或 PC LAN 平台上用软件加以实现,而且人脸数据库的容量大、易于扩充。实验表明,该系统具有较高的识别速度、准确率、容错性和鲁棒性等优点,具有推广的价值。(收稿日期:2002 年 7 月)

参考文献

1.Dibyendu Nandy, Jezekiel Ben-Arie. Shape from recognition: A novel approach for face shape from shading[J]. IEEE Transactions on image processing, 2006~217
2.金忠,胡钟山,杨静宇.基于 BP 神经网络的人脸识别方法[J].计算机研究与发展,1999,36(3):274~277
3.张林,胡波,凌雯亭.基于主元分析和 FUZZY ART 模型的人脸识别算法[J].电路与系统学报,1999,4(3):9~17
4.陈刚,戚飞虎.实用人脸识别系统的本征脸法实现[J].计算机研究与发展,2001,38(2):170~174
5.刘增良.模糊技术与神经网络技术选编(4)[M].北京:北京航空航天大学出版社,1999
6.P. G. J. Lisboa. 邢春颖,阳影等译.现代神经网络应用[M].电子工业出版社,1996