

# 基于 BP 神经网络的人脸表情识别

胡继胜<sup>1</sup>, 赵力<sup>2</sup>

(1.安徽职业技术学院 电气工程系,安徽 合肥 230051;

2.东南大学 无线电工程系,江苏 南京 210096)

**摘要:** 在 BP 神经网络及算法的基础上,提出了将其用于人脸表情识别的方法,不对人脸表情图像作特征提取,直接把图像数据作为神经网络识别器的输入。利用 Matlab 完成了网络的训练与测试。结果表明,本方案简单、快速、正确率高、识别时间短、泛化能力较强,可满足实时处理要求。

**关键词:** 人脸表情识别;BP 神经网络;Matlab

中图分类号: TP389.1

文献标识码: A

文章编号: 1674-7720(2010)11-0033-03

## Face expression recognition based on BP neural network

HU Ji Sheng<sup>1</sup>, ZHAO Li<sup>2</sup>

(1.Electrical Engineering Department, Anhui Vocational Technical College, Hefei 230051, China;

2.Radio Engineering Department, Southeast University, Nanjing 210096, China)

**Abstract:** Based on BP neural network and its algorithm, the paper proposes method of face expression recognition without feature extraction. The image data are served as the network input. This paper concludes the network training and testing which are presented by Matlab language. The experiment result shows that algorithm in this approach is easy and fast, has high recognition ratio, and the generalization is very good and content with realtime processing.

**Key words:** face expression recognition; BP neural network; Matlab

人类的表情在人类交流过程中扮演着很重要的角色。关于面部表情识别的研究能帮助人们建立更加智能化和交互性良好的系统。1971年,Ekman和Friesen研究了6种基本表情:高兴、悲伤、惊讶、恐惧、愤怒和厌恶,并系统地建立了上千幅不同的人脸表情图像库<sup>[1]</sup>。随着人机交互研究的深入和巨大的应用前景,人脸面部表情识别已经成为当前模式识别和人工智能领域的一个研究热点。在过去的几十年内,研究者已经提出了很多用于面部表情识别的方法。按照识别时所用的特征不同分类,可以将这些方法分成3类:基于灰度分布特征的方法、基于运动特征的方法和基于频率特征的方法<sup>[2]</sup>。

本文利用 BP 神经网络的高度非线性映射能力,提出了一种人脸表情的识别方法。一般的神经网络图像识别系统是由预处理、特征提取和神经网络分类器组成,对于有特征提取部分的系统实际上利用人的经验来获取模式特征以及神经网络分类能力来识别目标图像。但是在表情识别系统中,很容易丢失表征表情特征的细节

信息,导致它的抗干扰能力不强。所以本实验在预处理的基础上,采用无特征提取方法来对 BP 网络进行训练。

### 1 BP 神经网络

人工神经网络 ANN(Artificial Neural Networks)系统从 20 世纪 40 年代诞生至今仅半个多世纪,但由于它具有信息的分布存储、并行处理以及自学习等优点,已经在信息处理、模式识别、智能控制及系统建模等领域得到越来越广泛的应用。BP 算法是目前最广泛应用神经网络学习算法之一,尤其是基于误差反向传播(Error Back Propagation)算法的多层前馈网络(Multiple-Layer Feedforward Network)(简称 BP 网络),由 Rumelhart 于 1986 年提出。这种算法可以对网络中各层的权系数进行修正,可以以任意精度逼近任意的连续函数,故适用于多层网络的学习,广泛应用于非线性建模、函数逼近、模式分类、数据压缩等方面。80%~90%的人工神经网络模型采用 BP 网络或它的变化形式,它是前馈网络的核心部分,也是人工神经网络最精华的部分。

BP 网络是 Widrow\_hoffi 学习算法和非线性可微转移函数的多层网络,该名是由于其权值的调整采用反向传播(Back Propagation)的学习算法而得。一个典型的 BP 网络采用的是梯度下降算法。实际应用中 BP 网络一般为多层结构,既有多输入层,又有多个输出层,还包括若干隐层。典型的 BP 网络结构如图 1 所示<sup>[3]</sup>。

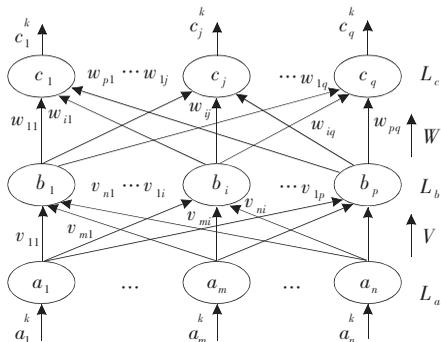


图 1 典型 BP 神经网络结构

隐含层单元  $i$  的加权输入为:  $net\ b_i = \sum_{m=1}^n w_{mi} a_m$ 。

其中,  $w_{mi}$  表示输入神经元  $m$  到隐含神经元  $i$  的权值,  $a_m$  表示当前输入样本,  $n$  表示输入样本个数。输出层单元  $j$

的加权输入为:  $net\ c_j = \sum_{i=1}^p w_{ji} b_i$ 。隐含层单元  $i$  的实际输出为:  $b_i = f(net\ b_i)$ 。输出层单元  $j$  的实际输出为:  $c_j = f(net\ c_j)$ 。

函数  $f$  为非线性可微分递减函数:  $f(x) = \frac{1}{1+e^{-x}}$ 。样本的实际输出与理想输出误差:  $E = \frac{1}{2} \sum_j (c_j - c_j^k)^2$ 。

在网络进行训练时,根据输出与输入样本的误差不断调整每个层神经元之间的权值,输出神经元与隐含神经元之间的权值调整为:  $\Delta w_{ji} = -\eta \delta_{nj} c_j$ 。

第  $m$  个输出与第  $j$  个样本间误差梯度:  $\delta_{nj} = \partial E / \partial net\ c_{nj}$ 。隐含神经元和输入神经元之间的权值的调整为:  $\Delta w_{mi} = -\eta \delta_{ni} b_i$ 。

第  $m$  个输出与第  $j$  个样本间误差梯度:  $\delta_{nj} = \partial E / \partial net\ c_{nj}$ 。隐含神经元和输入神经元之间的权值的调整为:  $\Delta w_{mi} = -\eta \delta_{ni} b_i$ 。

隐含神经元和输入神经元之间的权值的调整为:  $\Delta w_{mi} = -\eta \delta_{ni} b_i$ 。

$$\delta_{mi} = \sum_{j=1}^q \delta_{nj} w_{ji} f'(net\ b_{mi})$$

在实际修正时,为了加快收敛速度,往往在权修正量中加上前一次的权修正量:

$$\Delta w_{ji}(t) = -\eta \delta_{nj} c_j + \alpha \Delta w_{ji}(t-1)$$

经过反复调整网络的权值,使网络的实际输出和理想输出之间的误差达到要求,这样一个好的网络就训练成功。

## 2 BP 神经网络表情识别

本文应用 BP 神经网络进行人脸表情识别时,首先对待识别图像进行预处理,然后确定好网络结构(关键的一步),再选择适当的学习算法,最后由网络分类器输出识别结果。

《微型机与应用》2010 年第 11 期

## 2.1 图像预处理

本实验采用日本标准表情库 JAFFE,其规范化标准脸像如图 2 所示。库中包含 10 个人的 7 种表情(6 种基本表情和 1 种是中性感),平均每种表情每人 3 张,共有 210 张,大小为 256×256,格式为 BMP。首先进行平滑滤波处理消除噪声,然后利用类间方差判断分析方法确定二值化阈值,接着采用最大梯度差和灰度投影方法定位人眼瞳孔和嘴巴,使图像的分割和五官的定位相对容易得多,最后对图像进行切割和归一化处理。



图 2 规范化标准脸像

## 2.2 BP 网络隐层个数的选择<sup>[4]</sup>

对于含有一个隐层的 3 层 BP 网络可以实现输入到输出的任何非线性映射。增加网络隐层数可以降低误差、提高精度,但同时也使网络复杂化,增加了网络的训练时间,所以本实验只选择包含一个隐层的 3 层 BP 网络结构。

## 2.3 神经元节点数的选择<sup>[5]</sup>

### (1) 输入节点选择

输入层起缓冲存储器的作用,它接收外部的输入数据,因此其节点数取决于输入矢量的维数,本实验的人脸表情图像为 BMP 格式的目标图像,大小为 256×256,经归一化处理后为 120×120,若直接作为输入矢量,节点数为 14 400,显然其计算量是巨大的。虽然双 3 次插值效果好,但计算量大,本实验中并没有采用。实验中采用双线性插值缩放使其转换到 32×32 大小,其输入节点数为 1 024 个。

### (2) 输出节点数的选择

由于实验中要识别出 7 种人脸表情,因此正常情况下输出节点数为 7 个,以二进制形式来表示模式分类输出的结果。但由于输出层采用 S 型传输函数,其输出值只能在 (0, 1) 的范围内。若采用传输函数,输出结果可以是任意值,则可以给 7 种表情赋予不同的值来加以识别(如 7 种表情分别对应数字 1、2、3、4、5、6、7),只需要 1 个输出节点。本次实验对两种情况都进行了测试。

### (3) 隐含层神经元个数的选择

隐含层神经元个数直接影响神经网络的学习能力和归纳能力。隐含层神经元个数较少时,网络每次学习的时间较短,但有可能因为学习不足导致网络无法记住全部学习内容;个数较多时,学习能力增强,网络每次学习的时间较长,网络的存储容量随之变大,导致网络对未知输入的归纳能力下降。隐含层神经元个数的选择至今仍得不到统一的规范,一般可由以下 2 个公式中的 1 个来确定。设隐含层神经元个数为  $q$ ,输入节点数为  $m$ ,

欢迎网上投稿 www.pcachina.com 39

输出节点数为  $n$ , 则:

$$q = \sqrt{m+n} + a \quad (a \text{ 为 } 0 \sim 10 \text{ 的整数})$$

$$\text{或 } q = \sqrt{0.43mn + 0.12n^2 + 2.54m + 0.77n + 0.35 + 0.51}$$

由上式本实验的隐含层神经元个数选为 40 个。

### 3 实验及结果分析

实验中进行 2 组安排, 第 1 组与人有关的表情方案: 参加训练与测试的表情来自相同的人, 一部分表情为训练样本, 一部分表情为测试样本, 即从 210 张图像中选取 140 张作为训练样本, 每人每个表情 2 张, 剩下 70 张做测试样本。第 1 组做了输出节点为 1 个和 7 个的实验; 第 2 组与人无关的表情方案: 参加训练与测试的表情来自不同的人, 选取其中 8 个人共 168 张作为训练样本, 测试样本为其余 2 个人的 42 张表情样本。测试结果如表 1~表 3 所示。

表 1 与人有关的表情脸方案(输出为 1 个节点)

表情	训练样本数	测试样本数	正确判别数	错误数	识别率/%
愤怒(angry)	20	10	9	1	90
厌恶(disgust)	20	10	8	2	80
恐惧(fear)	20	10	9	1	90
高兴(happy)	20	10	10	0	100
悲伤(sad)	20	10	9	1	90
惊讶(surprise)	20	10	9	1	90
中性(neuter)	20	10	10	0	100
平均识别率: 91.43%					
平均识别时间: 0.44~0.47 s					

表 2 与人有关的表情脸方案(输出为 7 个节点)

表情	训练样本数	测试样本数	正确判别数	错误数	识别率/%
愤怒(angry)	21	9	9	0	100
厌恶(disgust)	21	9	7	2	77.8
恐惧(fear)	21	9	8	1	88.9
高兴(happy)	21	9	9	0	100
悲伤(sad)	21	9	8	1	88.9
惊讶(surprise)	21	9	8	1	88.9
中性(neuter)	21	9	9	0	100
平均识别率: 92.06%					
平均识别时间: 0.17~0.18 s					

由表 1、表 2、表 3 可以看出, 与人有关的表情脸识别率达到 90% 以上, 与人无关的表情脸识别率也达到 85.7%, 而在识别时间方面, 输出节点数取 7 个时, 识别

表 3 与人无关的表情脸方案(输出为 7 个节点)

表情	训练样本数	测试样本数	正确判别数	错误数	识别率/%
愤怒(angry)	24	6	5	1	88.9
厌恶(disgust)	24	6	4	2	66.7
恐惧(fear)	24	6	5	1	83.3
高兴(happy)	24	6	6	0	100
悲伤(sad)	24	6	4	2	66.7
惊讶(surprise)	24	6	6	0	100
中性(neuter)	24	6	5	1	83.3
平均识别率: 85.7%					
平均识别时间: 0.17~0.18 s					

时间比输出节点为 1 个时缩短约 2/3。结合网络训练的收敛情况, 可以得出结论: 不作特征提取的神经网络表情识别系统具有很强的抗干扰能力, 本方案简单、快速、正确率高、识别时间短、泛化能力较强, 有满足实时处理要求的能力。由于本次识别方案的预处理方法是针对 JAFFE 这个特定的表情库设计的, 先验知识十分充足。但是, 如果更换表情库或者直接拿日常拍摄的图像来识别, 预处理效果不会太好, 正确率不会有该实验中这么高。对于在复杂背景中的人脸检测以及优化 BP 神经网络在人脸表情识别中的应用将是以后研究的方向。

#### 参考文献

- [1] FABIO L, ROBERTO P. An efficient use of MPGE-4 FAP interpolation for facial animation at 70 bits/Frame [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(10).
- [2] 何良华, 邹采荣, 赵力, 等. 人脸面部表情识别的研究进展[J]. 电路与系统学报, 2005, 10(1): 70-75.
- [3] 彭淑敏, 王军宁. 基于神经网络的图像识别方法[J]. 电子科技, 2005(1): 38-41.
- [4] 飞思科技产品研发中心. MATLAB 6.5 辅助神经网络分析与设计[M]. 北京: 电子工业出版社, 2003.
- [5] 丛爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 合肥: 中国科学技术大学出版社, 2003.

(收稿日期: 2009-01-07)

#### 作者简介:

胡继胜, 男, 1970 年生, 副教授, 硕士, 主要研究方向: 信号与信息处理。

赵力, 男, 1958 年生, 教授, 博导, 主要研究方向: 语音信号处理, 自然语言处理, 情感处理。