

# 基于 BP 神经网络的嘴型分类算法

魏 爽

(西安电子科技大学 工程学院, 陕西 西安 710071)

**摘 要** 针对传统嘴型分类中医生肉眼观察主观性大且无法量化等问题,文中提出一种嘴型分类算法。该算法在人脸特征点标定 ASM 的基础上,提取嘴型模板,经过一系列旋转、二值化以及重采样构建特征向量,然后利用 BP 神经网络进行训练完成嘴型分类。实验结果表明,该算法可以实现较为准确的嘴型分类,对临床医疗整形具有一定的参考价值。

**关键词** ASM;嘴型分类;BP 神经网络

中图分类号 TP391.41 文献标识码 A 文章编号 1007-7820(2016)08-089-04

## Research on Mouth Type Classification Algorithm Based on BP Neural Network

WEI Shuang

(School of Electronic Engineering, Xidian University, Xi'an 710071, China)

**Abstract** Aiming at traditional mouth type classification mainly depends the observation of doctor, which is easily influenced by subjective factors and can't be quantified, This paper proposes a kind of mouth type classification algorithm. On the basis of ASM, the algorithm extracts the mouth template, constructs the feature vector with a series of rotation, binarization and heavy sampling, and then using BP neural network training feature vector to complete the mouth type classification. Experimental results show that the proposed algorithm can achieve accurate classification of mouth, which has a certain reference value for clinical medical treatment.

**Keywords** ASM; mouth type classification; BP neural network

近年来,与人脸相关的处理技术受到了越来越广泛的关注,并取得了一定的研究成果<sup>[1-4]</sup>,所以利用计算机人脸技术辅助医疗整形是具备可行性的,而且具有可观的应用前景。对此,本文提出了一种嘴型分类算法,该算法的主要思想是利用 ASM (Active Shape Model)<sup>[5-6]</sup>算法标定人脸特征点,接着在此基础上构建嘴型模板,经过一系列旋转、二值化以及重采样构建特征向量。然后利用 BP 神经网络算法进行训练,最终完成不同嘴型的分类。实验结果表明,该算法对嘴型实现了较准确的分类。

### 1 人脸特征点标定 ASM

在进行嘴型分类之前需标定人脸的特征点,本文采用人脸特征点标定的经典算法 ASM 来解决该问题。ASM 算法主要分为 ASM 建模和 ASM 匹配两部分。ASM 建模是利用样本图片进行训练探索出目标形状变化的规律,其主要分为形状建模和局部纹理

建模。它首先利用人脸样本库标定的特征点获取人脸形状信息,通过形状对齐去除形状变化的各类姿态的影响,并用 PCA (Principal Component Analysis)<sup>[7]</sup>统计形状模型。通过形状模型可以获知形状特征点的动态变化,从而将最终得到的形状约束到一定合理范围内;其次,在特征点法线方向上获取像素信息,建立复合局部纹理模型,即平均灰度向量和协方差矩阵;接着利用局部纹理信息与特征点的线性关系,迭代搜索最佳特征点;最后,经约束后的最佳形状即为人脸特征点坐标。图 1 所示为经 ASM 算法标定的人脸图片。



图 1 ASM 标定人脸特征点

收稿日期:2015-11-24

作者简介:魏爽(1991-),女,硕士研究生。研究方向:计算机视觉,模式识别。

2 BP 神经网络

得到标定的人脸特征点后,需进一步完成嘴型分类。嘴型分类本质是一个分类问题,对于一个分类问题,目标就是学习决策函数,机器学习中的典型代表,如 SVM、Adaboost<sup>[8-9]</sup>、BP<sup>[10]</sup> 等算法都擅长解决分类问题。实际仿真处理时,BP 神经网络的分类效果最准确,所以本文主要介绍 BP 神经网络相关原理。

2.1 BP 算法原理

BP 神经网络算法主要包括信息正向传播和误差反向传播两个过程,其主要思想是输入层接受外界信息,将信息传递给隐层,并进一步正向传递直至输出;计算输出结果与实际结果的误差然后反向传播调节各层权值。如图 2 所示是一个典型的 3 层网络拓扑图,3 层分别指输入层、隐含层和输出层。

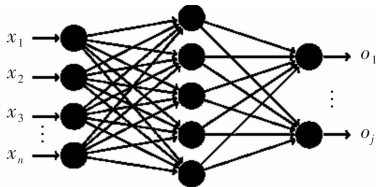


图 2 3 层 BP 网络拓扑图

算法的具体流程如下:

(1)网络的初始化。设输入层、隐含层、输出层的节点分别为  $n, l, m$ 。  $w_{ij}$  代表第  $i$  个输入单元到第  $j$  个隐层单元的权值,  $w_{jk}$  代表第  $j$  个隐层单元到第  $k$  个输出单元的权值,训练前各权值都初始化为随机值;

(2)隐含层的输出。隐含层的输出  $H_j$  为

$$H_j = \sigma(\sum_{i=1}^n w_{ij}x_i + a_j) \tag{1}$$

其中,  $a_j$  为输入层到隐含层的偏置;  $\sigma(y)$  代表激励函数。

$$\sigma(y) = \frac{1}{1 + e^{-y}} \tag{2}$$

(3)输出层的输出

$$O_k = \sum_{j=1}^l H_jw_{jk} + b_k \tag{3}$$

其中,  $b_k$  为隐含层到输出层的偏置;

(4)误差的计算。取误差公式为

$$E = \frac{1}{2} \sum_{k=1}^m (Y_k - O_k)^2 \tag{4}$$

其中,  $Y_k$  为期望输出,记  $Y_k - O_k = e_k$ ,则  $E$  可以表示为

$$E = \frac{1}{2} \sum_{k=1}^m e_k^2 \tag{5}$$

其中,  $i = 1, 2, \dots, n, j = 1, 2, \dots, l, k = 1, 2, \dots, m$ 。

(5)权值的更新。权值的更新公式为

$$\begin{cases} w_{ij} = w_{ij} + \eta H_j(1 - H_j)x_i \sum_{k=1}^m w_{jk}e_k \\ w_{jk} = w_{jk} + \eta H_j e_k \end{cases} \tag{6}$$

其中,  $\eta$  表示学习速率,用于控制下降的步长。

1)更新隐藏层到输出层的权值

$$\frac{\partial E}{\partial w_{jk}} \sum_{k=1}^m (Y_k - O_k) \left( \frac{\partial O_k}{\partial w_{jk}} \right) = -e_k H_j \tag{7}$$

则权重的更新公式为

$$w_{jk} = w_{jk} + \eta H_j e_k \tag{8}$$

2)更新输入层到隐藏层的权值

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial H_j} \cdot \frac{\partial H_j}{\partial w_{ij}} \tag{9}$$

其中

$$\begin{aligned} \frac{\partial E}{\partial H_j} &= (Y_1 - O_1) \left( -\frac{\partial O_1}{\partial H_j} \right) + L + \\ & (Y_m - O_m) \left( -\frac{\partial O_m}{\partial H_j} \right) = -\sum_{k=1}^m w_{jk}e_k \end{aligned} \tag{10}$$

$$\frac{\partial H_j}{\partial w_{ij}} = \frac{\partial \sigma(\sum_{i=1}^n w_{ij}x_i + a_j)}{\partial w_{ij}} = H_j(1 - H_j)x_i \tag{11}$$

则权重的更新公式为

$$w_{ij} = w_{ij} + \eta H_j(1 - H_j)x_i \sum_{k=1}^m w_{jk}e_k \tag{12}$$

(6)偏置的更新。偏置的更新公式为

$$\begin{cases} a_j = a_j + \eta H_j(1 - H_j) \sum_{k=1}^m w_{jk}e_k \\ b_k = b_k + \eta e_k \end{cases} \tag{13}$$

1)更新隐藏层到输出层的偏置

$$\frac{\partial E}{\partial b_k} = (Y_k - O_k) \left( -\frac{\partial O_k}{\partial b_k} \right) = -e_k \tag{14}$$

则偏置的更新公式为

$$b_k = b_k + \eta e_k \tag{15}$$

2)更新输入层到隐藏层的偏置

$$\frac{\partial E}{\partial a_j} = \frac{\partial E}{\partial H_j} \cdot \frac{\partial H_j}{\partial a_j} \tag{16}$$

其中

$$\frac{\partial H_j}{\partial a_j} = \frac{\partial \sigma(\sum_{i=1}^n w_{ij}x_i + a_j)}{\partial a_j} = H_j(1 - H_j) \tag{17}$$

$$\begin{aligned} \frac{\partial E}{\partial H_j} &= (Y_1 - O_1) \left( -\frac{\partial O_1}{\partial H_j} \right) + L + \\ & (Y_m - O_m) \left( -\frac{\partial O_m}{\partial H_j} \right) = -\sum_{k=1}^m w_{jk}e_k \end{aligned} \tag{18}$$

则偏置的更新公式为

$$a_k = a_k + \eta H_j(1 - H_j) \sum_{k=1}^m w_{jk}e_k \tag{19}$$

(7)迭代停止条件。本算法在处理时制定的停止条件为:当算法迭代至一定次数或相邻两次误差之差小于一定值时,将停止迭代。

2.2 避免训练中的局部极小值

BP 神经网络利用梯度下降法<sup>[11]</sup> 计算全局最小

值,但实际处理时,误差函数曲面可能存在多个局部极小值,而误差函数沿着梯度方向下降的过程中可能陷入这些局部极小值。通过增加冲量项,可尽量避免陷入局部极值,而且可以加快收敛速度。

将前文得到的权值更新法则 $\Delta w_{ji}(n)$ 进行修改,使得上一轮的更新结果将影响本轮更新权值,表达式为

$$\Delta w_{ji}(n)' = \Delta w_{ji}(n) + \alpha \Delta w_{ji}(n-1) \tag{20}$$

其中, $\Delta w_{ji}(n-1)$ 和 $\Delta w_{ji}(n)$ 分别代表算法未加冲量项第 $n-1$ 次和第 $n$ 次迭代时权值结果; $\Delta w_{ji}(n)'$ 代表加冲量项后第 $n$ 次权值结果; $\alpha$ 代表冲量,满足 $0 \leq \alpha < 1$ ,用于控制上一轮的权重更新对本轮影响程度。

3 算法实现





3.1 嘴型分类标准

在进行嘴型分类之前,需要制定嘴型分类标准。由于目前关于嘴型分类的相关研究较少,没有权威的嘴型分类标准。考虑到现有技术条件,并参考医疗整形中相关数据手册,本文将嘴型大致分为理想嘴形、厚嘴形、薄嘴形、口角下垂形4种,分类标准如表1所示。

3.2 特征设计

为识别出图片中的嘴型,需要用一种特征表征该图片的内容,从而构建成BP神经网络的特征输入。本算法在设计时采用嘴型模板的全部像素作为输入特征,这样做一方面可以省去采用原始图片需要降维的计算量,另一方面也可以为神经网络从样本中学习提供更大的自由度,其特征设计详细过程如下。

表1 嘴型分类标准

编号	嘴型类别	图片示例	类别描述
1	理想嘴型		嘴型线条流畅,唇峰唇谷明显,上下唇厚度适中,并分别在5~8 mm和10~13 mm范围内
2	厚嘴型		唇峰较高,上下唇过厚,上唇>8 mm,下唇>13mm,唇型有外翻感,常见于黑人嘴型
3	薄嘴型		唇形扁平,上下唇过厚,上唇>5 mm,下唇<10 mm,常见于白人嘴型
4	口角下垂型		唇型两侧有下垂角度,给人以严肃愁苦之感,这是需要调整的嘴型

(1)图像旋转。为避免因嘴型图片存在角度而影响结果,需要将人脸旋转至水平。由于人眼的角度基本已经决定图片的偏转角度,所以算法采用双眼特征点偏转计算旋转角度。图片旋转公式为

$$\begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix} = \begin{bmatrix} x_0 & y_0 & 1 \end{bmatrix} \begin{bmatrix} \cos a & -\sin a & 0 \\ \sin a & \cos a & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{21}$$

其中, $a$ 为旋转角度; $(x_0,y_0)$ 代表旋转前特征点; $(x_1,y_1)$ 代表旋转后特征点。经旋转后的效果如图3所示。



图3 经旋转后的效果

(2)嘴型图片二值化。根据模板训练的特征点,将嘴型区域分割出来,分割区域保证长宽一样,这样可以避免尺寸缩放导致嘴巴形状变化,然后通过射线法将嘴型内的像素点与嘴型外的像素点区分开来并二值

化,结果如图4所示。



图4 嘴型图片二值化

(3)将嘴型二值化图片归一化至指定大小,去除尺寸的影响。本算法设定为 $80 \times 80$ 。



图5 归一化嘴型模板

(4)模板重采样。直接采用嘴型模板作为特征向量输入会有6 400维,由于神经网络的输入单元数目和特征维数相等,维数过大会增加处理的复杂度,因此需要对原归一化嘴型模板图片进行重采样。原嘴型模板图片大小均为 $80 \times 80$ ,通过设置步长为5,可将图像统一降采样到 $16 \times 16$ 的分辨率。降采样后的图像是对原 $80 \times 80$ 像素图像的低分辨率描述,实际计算时,每个低分辨率像素是根据对应的局部高分辨率像素灰度

的均值计算得到的。通过对图片进行降采样可以有效的减少特征输入的维数和权值数量,从而降低运算量。并且该图像也保留了足够的大小以正确地识别目标图像。

4 实验结果与分析

实验环境为:Windows 7,CPU Intel(R) i3,基本频率 2.13 GHz,软件平台为 VS2008,OpenCV2.4.6。  
嘴型分为理想嘴型、厚嘴型、薄嘴型和口角下垂型 4 类,分别用 perfect、thick、thin、ptosis 来表示。图 6 是采用 MUCT<sup>[12]</sup>人脸库中部分图片的测试效果图。

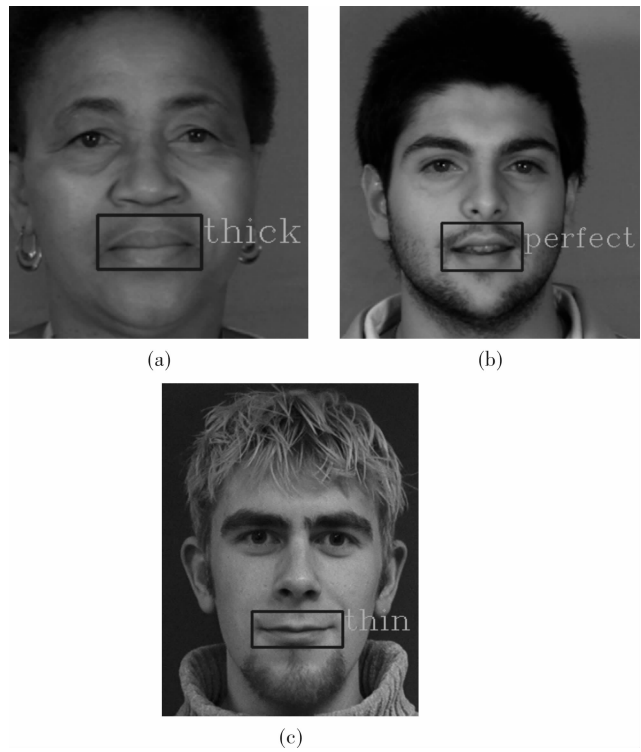


图 6 嘴型分类效果图

在实际处理时为比较算法性能,分别采用 SVM、Adaboost 和 BP 进行测试。由于嘴型分类对人脸图片会有一定要求,其需要人脸图片无表情且无较大偏转,所以训练图片采用 MUCT 库中符合要求的 200 张图片,测试图片也采用另外 100 张符合要求的图片,测试结果如表 2 所示。

表 2 嘴型分类测试结果

序号	方法	正确率/%
1	SVM(Liner)	61
2	SVM(RBF)	61
3	Adaboost	79
4	BP(隐藏层 10)	79
5	BP(隐藏层 10,加冲量项)	85
6	BP(隐藏层 64,加冲量项)	86

通过分析实验结果可知,Adaboost 和 BP 表现相对

较好。BP 加冲量项对提高正确率有一定的帮助,增加隐藏层对正确率帮助不大,反而会使训练时间变得相当长,综合考虑,效果较好的是 BP(隐藏层 10,加冲量项)。

5 结束语

本文提出了一种有效方法用于解决嘴型分类,研究工作主要包括嘴型分类标准制定和分类算法设计。该算法在人脸特征点定位 ASM 的基础上建立嘴型模板,经过旋转、归一化和重采样等一系列操作构建输入特征向量。实验结果表明,该算法实现了较好的嘴型分类效果。

参考文献

[1] Cootes T F, Taylar C D, Cooper D H, et al. Active shape models:their training and application[J]. Computer Vision and Image Understanding,1995,61(1):38-59.

[2] Cootes T F,Edwards G J,Taylor C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001,23(6): 681-685.

[3] Eliktutan O,Ulukaya S,Sankur B. A comparative study of face landmarking techniques[J]. Eurasip Journal on Image & Video Processing, 2013(1):1513-1517.

[4] Roberts M,Cootes T E,Pacheca E,et al. Quantitative vertebral fracture detection on DXA images using shape and appearance models[J]. Academic Radiology, 2007, 14(10): 1166-1178.

[5] Milborrow S,Nicolls F. Active shape models with SIFT descriptors and MARS[C]. Qubec:International Conference on Computer Vision Theory & Applications, 2014.

[6] Milborrow S,Tom E Bishop,Nicolls F. Multiview active shape models with SIFT descriptors for the 300-W face landmark challenge[C]. Porland: International Conference on Computer Vision,2013.

[7] Jolliffe I T. Principal component analysis[M]. Berlin:Springer Verlag,2002.

[8] Perter Harrington. 机器学习实战[M]. 李锐,李鹏,曲亚东,等,译. 北京:人民邮电出版社,2013.

[9] 李航. 统计学习方法[M]. 北京:清华大学出版社,2012.

[10] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [M]. MA, USA: MIT Press, 1988.

[11] Tom M Mitchel. 机器学习[M]. 曾华军,张银奎,译. 北京:机械工业出版社,2003.

[12] Milborrow S,Morkel J,Nicolls F. The MUCT landmarked face database[J]. Pattern Recognition Association of South Africa,2010(8):316-322.