

# 基于 SOM 网络的三维人脸表情识别

李淑红, 尹小娟, 句全

(河南财经政法大学 计算机与信息工程学院, 河南 郑州 450002)

**摘要:** 针对二维人脸表情数据所含信息量有限, 在光照、姿态变化的情况下识别性能较差等缺点, 提出了基于 SOM 网络的三维人脸表情识别方法. 该方法用均值和方差来描述人脸表面的凹凸情况, 以此作为进一步描述人脸表情变化的特征数据. 仿真实验结果表明, 采用 SOM 网络的分类效果和识别效果, 均优于 AdaBoost 算法.

**关键词:** 三维人脸表情识别; 形状描述; 自组织神经网络

**中图分类号:** TP391.41      **文献标志码:** A      **DOI:** 10.3969/j.issn.2095-476X.2013.05.017

## 3D facial expression recognition based on self organizing mapping network

LI Shu-hong, YIN Xiao-juan, JU Quan

(College of Computer and Information Engineering Henan University of Economics and Law Zhengzhou 450002, China)

**Abstract:** It is well known that the two-dimensional facial expression data contains limited information, and the poor performance of the facial expression recognition under the condition of changing illumination and posture. In order to overcome these shortcomings of the 2D facial expression, In this paper, we propose and explore a novel method to recognize human facial expression in 3D based on Self Organizing Map (SOM). In the method, the mean and variance are used to describe the convex and concave surface of the face which become the facial expression change characteristics datas. The simulation experimental results showed that the effect of using the classification and recognition of SOM network was superior to the Ada-Boost algorithm.

**Key words:** 3D facial expression recognition; shape description; self organizing mapping(SOM) network

## 0 引言

人的面部表情即人脸表情是人类在进行交流时所采用的最有效、最自然和最直接的一种途径,也是人机交互应用中很重要的一种信息资源<sup>[1]</sup>. 基于图像的人脸表情识别是以一种非接触式的方式来完成识别工作的,这一点对于实现自然、和谐的人机交互至关重要<sup>[2]</sup>. 目前,基于二维的人脸表情识别已经成为包括情感计算、人机交互、计算机视

觉、模式识别等在内的相关研究领域的一个重要发展方向. 但是在光照、姿态变化的情况下,二维人脸表情识别性能将急剧下降<sup>[3-5]</sup>.

三维图像获取设备的发展使得三维人脸表情识别成为可能,在已获得数据的前提下,三维人脸图像数据比二维图像数据信息量更大、鲁棒性更强,几乎不受光线影响,而且三维人脸数据具有的几何形体使三维人脸可以更好地克服姿态变化的影响,所以将三维人脸表情数据图像用于人脸表情

识别可以得到更好的效果<sup>[6-8]</sup>. 由于三维人脸数据维数较高, 而 SOM (self organizing mapping) 网络可以将任意维数的输入信号模式转变为低维的离散映射, 并以拓扑有序的方式自适应这个变换. 故本文拟在突破二维人脸表情识别数据的局限性的同时, 采用三维人脸表情数据, 并选用拓扑直观的 SOM 网络进行识别, 以期在提高表情识别率的同时有效降低数据复杂度.

### 1 三维人脸描述及表情特征提取

SOM 网络又被称为自组织映射神经网络. 它的自组织体现在它能够通过无监督学习和没有指导的训练改变权重, 将无序的系统状态变成有序系统状态<sup>[9]</sup>. SOM 网络适合用在那些数据关系不是很紧密的情况下, 它通过自身训练, 自动对输入模式进行聚类.

#### 1.1 三维人脸描述

由于人脸表情变化时会引起人脸表面的凹凸情况发生变化, 所以当表情不一样时, 人脸表面的凹凸情况也不一样. 基于此, 可以用人脸的凹凸变化情况来描述人脸表情.

三维面中的一点  $p$  与其邻近点  $p_i$  之间的关系见图 1.

图中  $p_i$  是满足  $p_i - p$  的长度约等于半径  $r$  的三维面上的所有点  $p$  与  $p_i$  组成一个环网. 那么  $p_i - p$  与点  $p$  处的法线  $N_p$  之间的夹角即为

$$\theta = \arccos\left(\frac{(p_i - p) \cdot N_p}{|p_i - p| \cdot |N_p|}\right)$$

其中  $N_p$  是点  $p$  处的法线,  $\theta$  是  $N_p$  与边  $p_i - p$  组成的角, 范围为  $0^\circ \sim 180^\circ$ .

$$mean(\theta) = \frac{1}{n} \sum_{i=1}^n \theta_i$$

$mean(\theta) > 90^\circ$  则表明半径为  $r$  的球面是凸面; 反之, 如果  $mean(\theta) < 90^\circ$ , 则表面这个面是凹面.

为了更好地描述三维图像的表面情形, 用公式  $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (\theta_i - \bar{\theta})^2$  求取角的方差, 来确定图像表面凹凸变化的差别.

#### 1.2 三维人脸表情特征提取

目前基于三维数据的人脸表情识别的算法大致可以分为 3 种, 即基于局部或整体或者是两者相结合的方法. 基于局部特征提取的人脸表情识别方法实际上就是利用人脸面部中重要部位的特征进行分类识别.

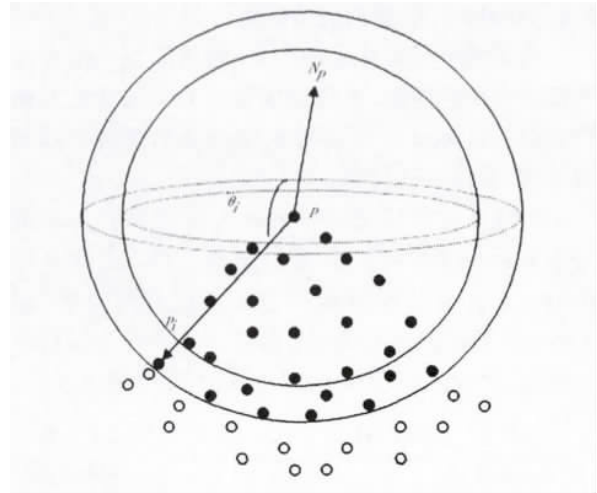


图 1 点  $p$  与其周围邻近点之间的关系

由于鼻子是人脸的主要特征之一, 又处于人脸的中心位置, 所以选取鼻尖为人脸特征提取的参照点. 进行三维人脸图像数据描述的步骤如下.

1) 对 FRGC 3D 人脸数据库中矫正过的三维人脸, 以鼻尖为基准点, 其中矫正过的三维人脸鼻尖点所处坐标为  $p_0(0, 0, 0)$  (如图 1 所示). 求距离鼻尖环半径为 5 mm 的第 1 个环面上其他数据点  $p_i$  与鼻尖点  $p$  的凹凸关系, 用  $\theta$  表示, 求第 1 个环面上的  $mean(\theta)$  和  $\sigma^2$ .

2) 用同样的方法求半径为 5 mm 的第 2 个环面上点  $p_i$  与鼻尖点  $p$  的凹凸关系, 以此类推. 共求 10 个环面上的点与  $p$  的凹凸关系, 最终得出一个 20 维的有关均值和方差的矩阵.

## 2 实验结果与分析

### 2.1 实验数据准备

FRGC 3D 人脸数据库是 2002—2004 年间通过激光扫描得到的, 是具有不同姿态与表情的人脸数据库. FRGC 3D 人脸库具体见表 1.

表 1 FRGC 3D 人脸表情数据库

类别	图像	人数	数据集
训练集	943	275	FRGCv1
测试集	4007	466	FRGCv2

本研究课题的仿真实验训练样本集, 是从 FRGC 3D 人脸表情数据库的训练部分中抽取的 136 张包括笑和平静 (平静也可以称为中性表情) 的图像数据, 用本文有关人脸描述和特征提取的方法进行特征提取.

### 2.2 AdaBoost 分类器表情识别

为了能更直观地说明 SOM 网络用在三维人脸表情识别中的性能,本仿真实验用相同的实验数据分别在 AdaBoost 分类器和 SOM 网络这 2 种不同的分类器上进行表情识别.

对比实验所用的 AdaBoost 分类器是将 3 个弱分类器最近邻(NN)、朴素贝叶斯(NB)及 LDA 用 AdaBoost 针对多态分类问题的算法进行分类.这 3 种弱分类器针对人脸模型特征识别的平均识别率分别为 84.6% 88.8% 和 87.8%.实验结果见表 2.

表 2 AdaBoost 对人脸表情的识别结果 %

分类算法	平静	笑
NN	90.2	91.2
NB	89.0	90.2
LDA	88.8	89.7

从表 2 可以看出 经过 AdoBoost 算法后 3 个弱分类器分类效果都有不同程度的加强.经过 AdaBoost 训练后 3 个弱分类器特征提取表情的识别率分别平均为 90.7% 89.6% 和 89.3%.

### 2.3 SOM 网络表情识别

SOM 网络是一种无监督的学习,通过自我训练对输入的数据进行自动聚类.用它进行三维人脸表情识别步骤如下:

1) 将 data 格式的三维人脸表情数据训练集加

载到 Matlab 下并且标准化;

2) 用标准化后的三维人脸表情训练数据集对 SOM 神经网络进行训练;

3) 记录 SOM 训练输出节点集及它们代表的人脸表情类别;

4) 将测试三维人脸数据集输入到训练好的 SOM 网络中,找出 SOM 网络的获胜节点  $M'$  和训练时的  $M$  比较后给出分类.

对 FRGC 3D 人脸表情数据库中的训练部分的 136 张图像数据进行特征提取,并进行 SOM 网络训练.为使在结果中查找 BMU 描述更方便,对 136 张图像数据进行编号,训练结果如图 2 所示.

图 2 中 a) 图是 u-matrix 显示方式,测度条上面高值代表的颜色表示聚类的边界,低值表示聚集在一起的一类.从 u-matrix 图中可以看到,通过训练, SOM 将输入数据聚为 2 类,在本实验中即是 将笑与平静 2 类人脸表情聚到不同的 2 类中; b) 图是通过标签标识成不同类别的,从图中可以看出,平静和笑被标出分别对应着不同的类别; c) 图是三维网格图,通过三维空间展示,可清楚地看出 SOM 将笑与平静分成 2 类.从 SOM 网络的实验结果可以看出, SOM 网络对训练样本的 136 张人脸表情数据中笑和平静的聚类结果很理想.

为了测试 SOM 网络对人脸表情的识别性能,上

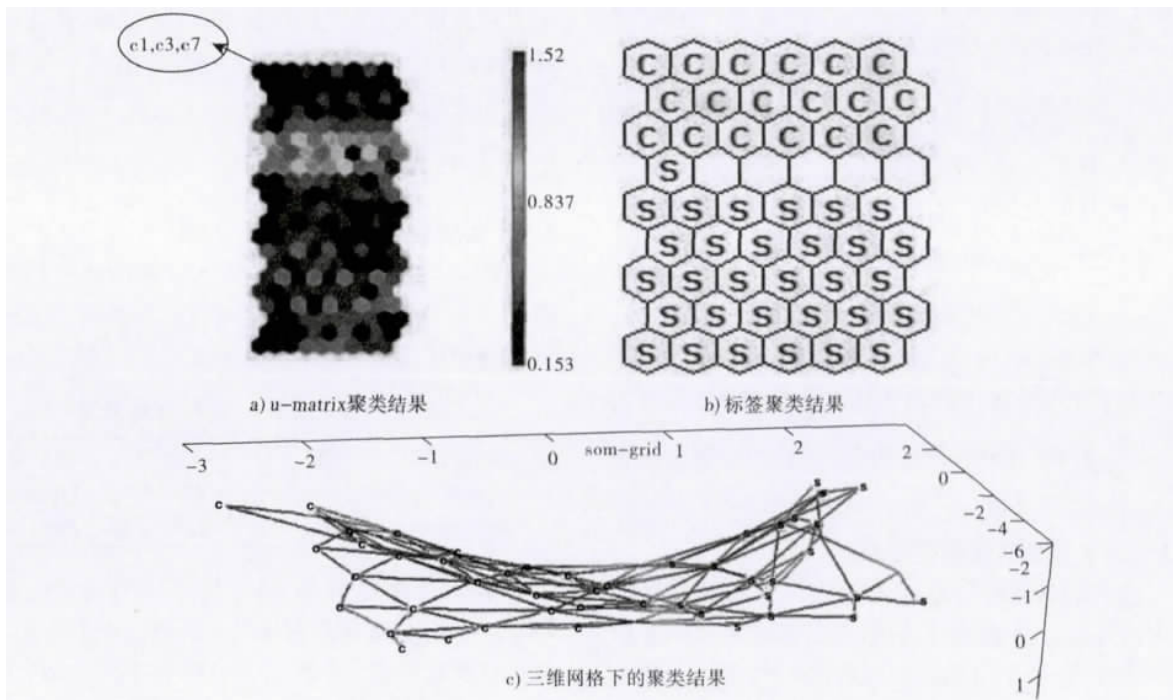


图 2 SOM 网络人脸表情聚类结果

述聚类过后还需要用测试样本集对其分类效果进行测试. 测试实际上就是找到其最佳匹配单元 BMU, 然后对比匹配结果.

通过 find() 函数找到每个神经节点所对应的获胜节点 BMU, 如图 2a) 所示, u-matrix 图的第 1 个胞体对应着平静表情下的序号为 (1 3 7) 的图像数据训练结果. 然后输入测试集进行测试, 通过找到其所对应的 BMU 得出测试集表情所属的类别.

本实验是对 FRGC 3D 人脸表情数据库中完全不同于训练集的人脸图像分别进行测试, 平静表情数据 50 个, 笑表情数据 50 个, 测试结果见表 3.

表 3 SOM 网络识别结果

测试集	笑	平静	识别率/%
笑	49	1	98
平静	3	47	94

## 2.4 实验结果分析

将 SOM 网络与 AdaBoost 算法相比较, 在对相同特征数据进行分类识别的情况下, AdaBoost 对人脸表情识别结果与 SOM 网络对人脸表情识别效果进行比较 (如图 3 所示). 从图 3 可以得出, 无论是从分类效果还是从最终的识别效果上看, SOM 网络都优于 AdaBoost 迭代分类.

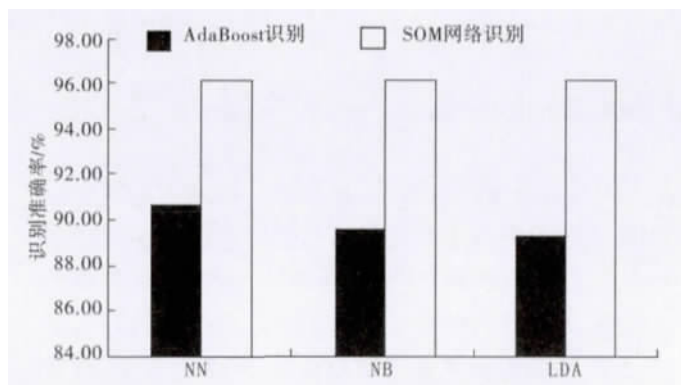


图 3 AdaBoost 与 SOM NB 网络性能比较

## 3 结语

本文将三维人脸表情数据用于表情识别, 将物体表面凹凸情况的三维物体描述方法用于描述三维人脸表情, 同时采用 SOM 网络这种拓扑直观的分

类器对人脸表情进行分类, 从而使人脸表情识别率得到提高.

虽然三维表情识别在精度方面还存在瑕疵, 但比起二维表情数据, 无论是在数据信息量方面, 还是在抗光照和姿态方面, 三维表情数据都有很强的鲁棒性, 值得研究者重视三维人脸表情的识别. 因此, 后续工作的重点应放在对三维人脸表情特征的描述以及 SOM 训练上.

## 参考文献:

- [1] Essa I A, Pentland A P. Coding, analysis, interpretation, and recognition of facial expressions [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 757.
- [2] Fukui K, Yamaguchi O. Facial feature point extraction method based on combination of shape extraction and pattern matching [J]. Systems and Computers in Japan, 1998, 29 (6): 49.
- [3] Ekman P, Friesen W V, Hager J C. Facial action coding system [J]. A Human Face, 2002, 23(5): 301.
- [4] Bartneck C. Integrating the occ model of emotions in embodied characters [C] // Workshop on Virtual Conversational Characters: Applications, Methods and Research Challenges, Melbourne, Citeseer, [s. n.], 2002.
- [5] Hui K, Wang C. Clustering-based locally linear embedding [C] // ICPR, [s. l.]: IEEE, 2008: 1-4.
- [6] Rish I. An empirical study of the naive Bayes classifier [C] // Proceedings of LICA Workshop on Empirical Methods in Artificial Intelligence, [s. l.]: IEEE, 2001: 41-46.
- [7] Chua C S, Han F, Ho Y K. 3D human face recognition using point signature [C] // International Conference on Automatic Face and Gesture Recognition, [s. l.]: IEEE, 2000: 233-238.
- [8] Darr H, Mayshar Y, Benvenisty N. Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features [J]. Development, 2006, 133(6): 1193.
- [9] Teuvo Kohonen. Self-organizing Maps [M]. New York: Springer Verlag, 1997.