

一种基于机器学习的人脸情绪识别方法研究

李广鹏¹, 刘 波¹, 李 坤², 黄思琦³

(1. 国防科技大学 计算机学院, 湖南 长沙 410073;

2. 国防大学政治学院 教研保障中心信息技术室, 上海 201602;

3. 大连交通大学 软件学院, 辽宁 大连 116000)

摘 要:文中对人脸情绪识别整个过程所使用的算法进行分析和总结,并参考了国内外的人脸情绪识别方面的论文和报告,提出一种传统的基于机器学习的人脸图像情绪识别方法。文中将重点放在人脸表情核心部位(眼睛、眉毛、鼻子和嘴巴),将这些部位的特征向量作为人脸表情主要特征进行处理。使用的人脸关键部位特征提取算法是基于回归的 Explicit Shape Regression 算法,该算法在人脸特征部位定位方面取得了很好的效果。因为 Gabor 小波有很好的仿生效果,可以很好地表达表情的变化,所以文中的特征提取算法使用的是 Gabor 小波。但是考虑到提取出的特征维数过高,算法复杂,耗时较大,在后期用 PCA 算法进行降维处理,降维之后进行表情分类。

关键词:情绪识别; Explicit Shape Regression; PCA; 支持向量机

中图分类号: TP301

文献标识码: A

文章编号: 1673-629X(2019)05-0027-05

doi:10.3969/j.issn.1673-629X.2019.05.006

Research on Face Emotion Recognition Based on Machine Learning

LI Guang-peng¹, LIU Bo¹, LI Kun², HUANG Si-qi³

(1. School of Computer, National University of Defense Technology, Changsha 410073, China;

2. Information Technology Office of Teaching and Research Security Center, Political

College of National Defense University, Shanghai 201602, China;

3. School of Software, Dalian Jiaotong University, Dalian 116000, China)

Abstract: We analyze and summarize the algorithms used in the whole process of face emotion recognition, and propose a traditional face image emotion recognition method based on machine learning by referring to domestic and foreign papers and reports on face emotion recognition. We focus on the core parts of facial expression (eyes, eyebrows, nose and mouth) to deal with the feature vectors of these parts as the main features of facial expression. The Explicit Shape Regression algorithm based on regression is used as the key facial feature extraction algorithm, which has been proved to be very effective in face feature location. Because Gabor wavelet has great bionic effect, it can express the change of expression very well, so the feature extraction algorithm in this paper uses Gabor wavelet. But considering that the extracted feature dimension is too high, the algorithm is complex and time-consuming, the PCA algorithm is used to reduce the dimension in the later stage, and then the facial expression is classified after dimensionality reduction.

Key words: facial expression recognition; ESR; PCA; SVM

0 引言

在日常生活中,人与人的交流主要是通过自然语言和形体语言,自然语言可以表达交流的内容,这种交流基本可以满足人与人之间大部分的信息交流,但是一些深层次的交流还是需要形体语言。形体语言则可以表达自己的情绪,也可以通过对方表达的情绪来辨认对方的态度和内心世界。形体语言交流主要体现在

人脸面部表情。人在说话交流时,面部表情随时在变,体现出内心的心理状态。自然语言加上形体语言才能使人與人之间的交流变得更加自然,关于形体语言交流的作用,心理学家 Mehrabian 给出了一个公式,感情表露=7%的言词+38%的声音+55%的面部表情^[1]。

如今的人工智能发展已经走上正轨,各项人工智能技术都已经适用于学习生活中的各个方面。人工智

收稿日期:2018-05-29

修回日期:2018-09-20

网络出版时间:2018-12-21

基金项目:国家自然科学基金(61572513)

作者简介:李广鹏(1996-),男,硕士,助理工程师,研究方向为社交网络数据分析。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20181221.1524.026.html>

能中最重要的一个方面就是人和计算机的信息交互,也称人机交互。人工智能要做的就是让人和计算机交互更加智能,更加方便,对此,要给计算机加入人的一些基本功能如视觉和听觉,这样人与计算机就消除了基本的隔阂^[2]。在此基础上,为了让人和计算机的交互更加常态化,需要给计算机加入情感理解和情感识别,让计算机能够很自然地和人交流,进一步消除人和计算机的隔阂。拥有情感理解和情感识别的计算机是人工智能未来的发展趋势,而人机交互最重要的一部分就是人脸情绪识别,人的情感主要表现在脸部,计算机要理解人的情感,第一步就是识别人脸的情绪。正因为这样,人脸情绪识别得到了很多科研机构和技术公司的重视,具有情感识别的计算机将彻底改变人机交互的现状^[3]。

1 相关工作

人脸检测识别的方法有很多种,可以划分为三大类。第一类方法是人脸局部特征作为重点研究的人脸检测方法,第二种是人脸整体特征作为重点研究的人脸检测方法,第三种方法是将前两种方法的优缺点结合的基于局部和整体的人脸检测方法。

人脸检测是要找出图像中人脸的大概区域,而人脸特征点定位则是要更加准确地定位出眼睛、眉毛、鼻子、嘴巴轮廓等一些人脸区域中的关键部位,以便对得到的人脸进行进一步的处理。研究方法主要分为两大类,一类是基于模型的方法,一类是基于回归的方法^[4]。

表情特征提取是 facial expression recognition 系统

中重要的一部分,高效的表情特征提取将大大提高表情识别率。表情特征提取算法主要有四种方法:基于统计的方法、基于频率特征提取的方法、基于运动特征的方法和基于神经网络的方法。

降维算法的本质是学习一个映射函数 $f: x \rightarrow y$,其中 x 是原始数据点的表达,目前最多使用向量表达形式; y 是数据点映射后的低维向量表达,通常 y 的维度小于 x 的维度(当然提高维度也是可以的); f 可能是显式的或隐式的、线性的或非线性的。常用的特征降维算法有主成分分析算法(PCA)和线性判别分析算法(LDA)^[5]。

人脸表情识别过程中的最后一步,也是直接影响结果的一步就是表情分类。表情分类是指通过设计分类机制来对表情进行分类,将表情归入相应的类别。常用的分类方法是按照感情分类,将表情分类到 Paul Ekman 归纳的 6 种基本感情。文中将所有表情分为四类:悲伤、开心、惊讶、害怕。人脸表情分类的方法有很多,主要是利用基于机器学习的方法如人工神经网络、贝叶斯分类、支持向量机以及一些其他的分类算法^[6]。

2 方法

算法的主要架构如图 1 所示。从图像的来源不同,针对视频,将视频图像序列提取关键帧,处理为单个图像,然后对图像进行预处理;对图像中的人脸进行检测,将人脸分离出来进行特征点定位并利用 Gabor 小波对人脸图像进行特征提取,提取出来的向量维数较高,接下来进行特征降维,最后对表情进行分类处理,得出测试图像所属的类。

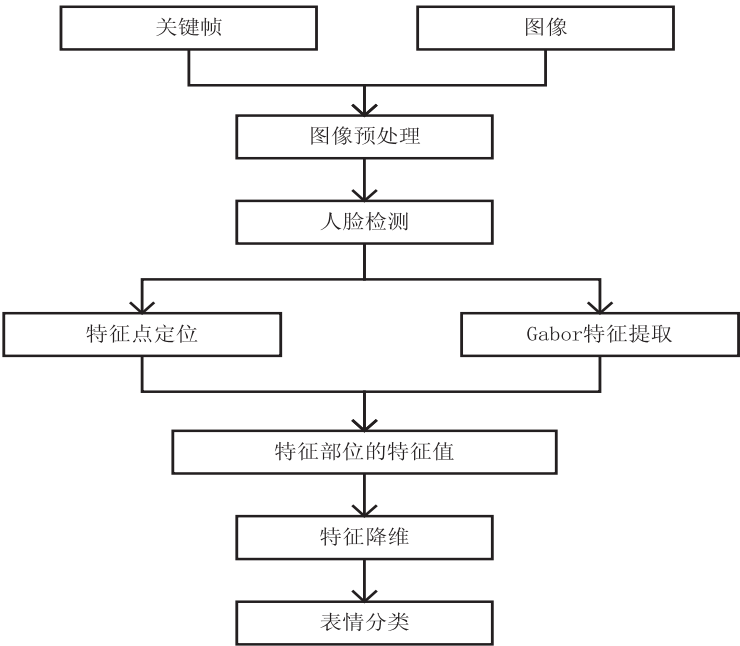


图 1 算法架构

2.1 图像预处理 (image preprocessing)

人脸检测是指对于任意一幅给定的图像,通过算法搜索图像,判断图像中是否存在人脸。如果算法检测出输入的图像中存在人脸图像,那么就将人脸从整个图像背景中分离出来^[7],将人脸的位置以坐标形式返回,同时返回人脸的大小和此时的姿态。人脸检测是人脸表情识别最基础的一步,能不能从图像中提取人脸,以及提取出来的人脸是否符合后期处理的标准等这些条件直接影响后期的识别率。人脸检测技术目前可供选择的方法很多,文中将使用 OpenCV 现有的 Haar 分类器进行人脸检测^[8]。

图像预处理是人脸情绪识别过程中最基础也是相当重要的一部分,在提取图像中的人脸位置和大小之后,对截取的人脸图像进行预处理是很重要的一步,因为输入图像的来源不同,图像的采集环境也不一样,特别容易受噪声影响。另外图像采集的时候人脸离采集器的距离,焦距的大小等因素也会使得人脸在整幅图像中的大小位置不确定,在这种情况下必须对图像进行预处理。图像的预处理分为点运算和几何处理^[9]。点运算可以改变图像的直方图分布,使图像更加清晰,利于后期的处理。除了点运算之外的图像预处理就是几何运算,几何运算可以得到人脸位置端正的人脸图像。通过点运算和几何运算可以得到尺寸、人脸位置统一的标准化人脸图像。

2.2 Explicit Shape Regression 实现人脸特征点定位 (face feature point positioning)

文中方法是基于模型的方法,使用的是微软研究

院孙剑提出的基于回归的 Explicit Shape Regression (ESR) 算法,该算法在人脸特征部位定位方面取得了很好的效果^[10]。通过 ESR 算法能够精确定位人脸关键点部位,给后期处理减少了大量的数据量。ESR 核心内容是使用了 cascade regression 框架以解决对齐问题。孙剑在 ESR 方法的基础上做了几个扩展,使其更为适合做人脸关键点定位,下面对 ESR 进行介绍^[11]。

主要思路:通过对输入训练集最小化对齐误差,可以得到一系列的回归函数,通过这些函数,可以进一步推断出人脸整体的形状。

训练过程:

(1) 两层级联回归模型,第一层 (R^1, R^2, R^3), 第二层 $R^i = (r^1, r^2, r^3)$ 。

(2) r^K 是由 2^F 个回归组成, F 为特征个数,通过阈值可以划分成 2^F 个 bin, 每个 bin 与一个回归输出 δS_b 相关。

(3) 通过优化目标函数 $\delta S_b = \arg \min_{\delta S_b \in \theta_b} \| S_i - (S_i) + \delta S \|$, 求解并加入收缩参数 (shrinkage parameter) 后有:

$$\delta S_{ikb} = \frac{1}{1 + \frac{\beta}{|\theta_b|}} \frac{\sum_{i \in \theta_b} (\hat{S}_i - S_i)}{|\theta_b|} \tag{1}$$

(4) 训练完成后,应该得到 $T * K * 2^F$ 个回归器: δS_{ikb} , $t = 1, 2, \dots, T, k = 1, 2, \dots, K, b = 1, 2, \dots, 2^F$ 。

特征部位识别如图 2 所示。

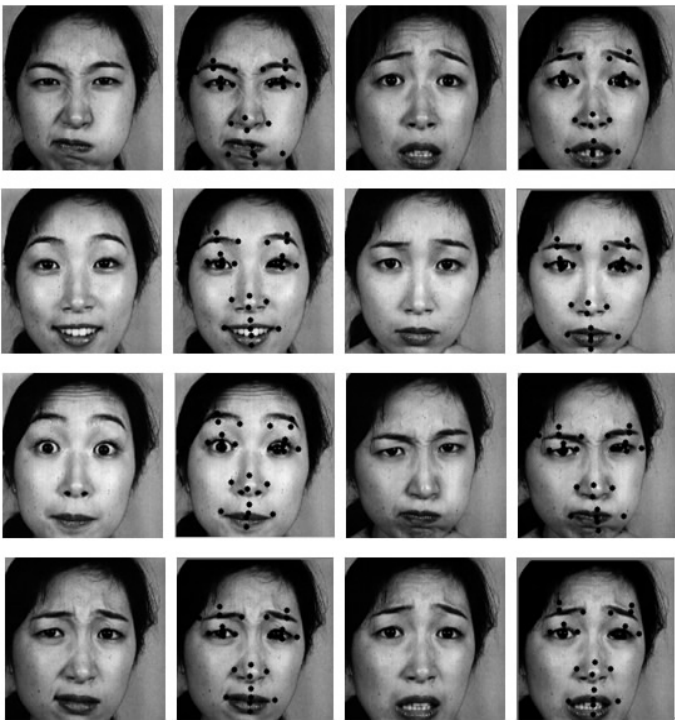


图 2 特征部位识别

2.3 Gabor 人脸特征提取 (Gabor face feature extraction)

在图像处理、模式识别以及计算机视觉等领域中, Gabor 滤波器应用广泛。生物学科的实验发现, Gabor 滤波器可以很好地近似单细胞的感受野函数。Gabor 滤波器也成为了少有的适用于计算机算法的生物仿生方面的数学模型。如图 3 所示, 第一行代表脊椎动物的视觉皮层感受野, 第二行是 Gabor 滤波器, 第三行是两者时间的差值, 可以看出差值非常小, 正是因为这种相似性, Gabor 滤波器被用于图像的分析 and 处理^[12]。

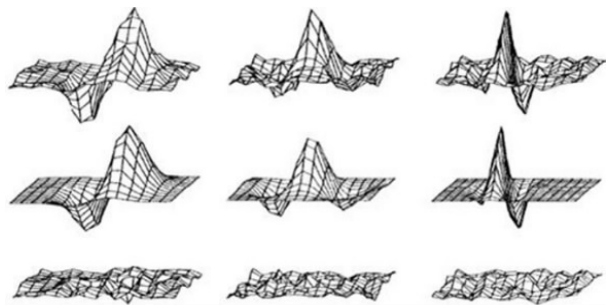


图 3 滤波器

Gabor 滤波器公式化定义如下:

复数表达:

$$g(x, y, \lambda, \theta, \phi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \exp\left[i\left(2\pi \frac{x'}{\lambda} + \phi\right)\right] \quad (2)$$

实数部分:

$$g(x, y, \lambda, \theta, \phi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \phi\right) \quad (3)$$

虚数部分:

$$g(x, y, \lambda, \theta, \phi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \phi\right) \quad (4)$$

其中, λ 为正弦函数的波长; θ 为 Gabor 核函数的方向; ϕ 为相位偏移; γ 为图像空间的宽高比; b 为 Gabor 滤波器的带宽值, σ 的值随着带宽 b 的变化而变化, 这里需要注意的是, b 的值必须是正实数, 否则将影响计算结果。 b 的值越小, 式子的标准差越大, Gabor 形状就会越大。

通过 40 个滤波器对图像进行滤波, 若图像宽高为 100×100 , 维数将达到 400 000 维, 这对后期的表情分类会带来很大的麻烦。图 4 中的特征向量选取了图像中 29 个特征点, 一共 40×29 , 1 160 维。图 4 是针对同一张脸, 选择不同的 Gabor 核函数生成的图, 并将特征点区域的像素值提取出来。

2.4 特征降维 (feature dimensionality reduction)

提取出来之后, 接下来就是对特征进行降维处理, 维度太高会影响系统的运行速度。文中使用主成分分

析算法 (principal component analysis, PCA)。PCA 算法主要用于数据的降维处理, 在很多数据分析的系统中, 数据以多维的方式处理, 而多维数据中, 某些数据本身没有区分性, 如果用这些数据进行区分, 那么结果代表性不大^[12]。这些数据也成为无用数据, 因此要做的是找到多维数据中那些变化比较大的数据, 变化大小可以通过数据的方差来辨别, 通过去除这些变化小的数据, 使得留下来的数据都能够最大限度地表达原始的数据^[13]。这样处理过的数据, 数据维度低了, 但是数据的原始含义并没有发生变化。

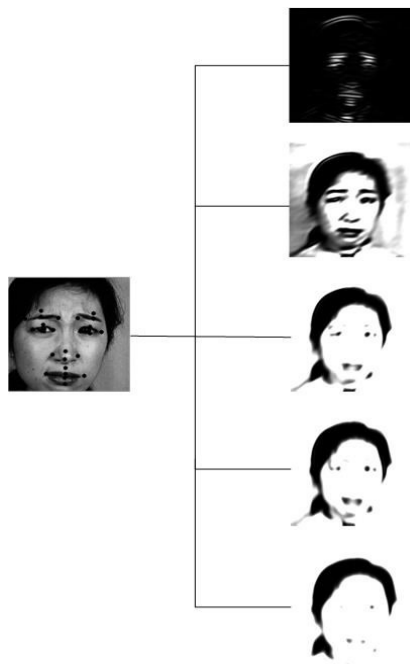


图 4 关键部位的特征值

PCA 是在尽可能好地表示原特征的情况下, 将原特征进行线性变换, 映射至低维度空间。其问题可以简单描述为对于 d 维空间中的 N 个样本 $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$, 在低维的空间进行表示。首先假设在 0 维空间中, 然后扩展到 1 维, 然后到 m 维, 其中 $m < d$ 。

2.5 表情分类 (expression classification)

提取人脸图像特征向量, 对高维向量进行降维之后, 接下来是对人脸表情进行识别分类。在前文的介绍中, 表情分类方法主要是基于机器学习的人工神经网络、贝叶斯分类、支持向量机等, 文中选用支持向量机 (SVM)。为了使得分类结果更加准确, 将使用 LibSVM 对表情进行分类。LibSVM 是林智仁先生编写的 SVM 软件, 该软件将 SVM 方法集合在一起, 在很大程度上便于分类时使用 SVM 的方法。

表情分类流程如图 5 所示。第一步准备训练数据, 将数据处理成 LibSVM 的标准格式; 第二步训练 model, 调整好参数, 载入数据进行训练; 第三步利用训练好的 model 对样本数据进行预测; 最后分析预测结

果,与标准结果进行比对分析。

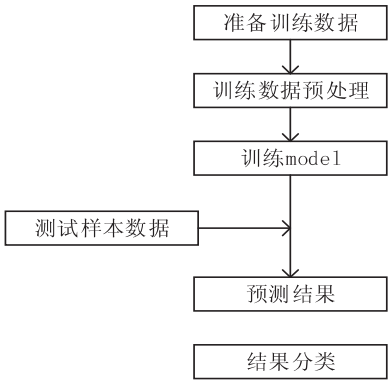


图5 表情分类过程

3 结束语

总结了人脸表情识别流程,在此基础上提出了文中的识别流程。和其他表情识别过程不一样,文中将重点放在表情关键部分。人脸表情主要体现在眼睛、鼻子、嘴巴、眉毛四个关键部分,通过对这四个关键部分的定位,可以将人脸特征选择缩小至这些部位中^[14]。减少了计算量,同时也提高了识别率。

阐述了人脸情绪识别研究的意义和背景,对人脸情绪识别整个过程所使用的算法进行分析和总结,并参考了国内外有关人脸情绪识别的论文和报告,对人脸情绪识别的算法发展和现状进行总结,并分析了这些算法的优缺点^[15]。在此基础上,提出了一种新的人脸情绪识别流程,大多数的人脸情绪识别算法将重点放在人脸整体上,将人脸整体进行分析处理,算法核心是降维处理。

文中使用的人脸关键部位特征提取算法是 ESR 算法,该算法在人脸特征部位定位方面取得了很好的效果^[16]。通过 ESR 算法能够精确定位人脸关键点部位,给后期处理减少了大量的数据量。

文中的人脸特征向量选择采用了 Gabor 小波,因为 Gabor 小波有很好的仿生效果,可以很好地表达表情的变化^[17]。但是考虑到提取出的特征维数过高,算法复杂,耗时较大,不对全图像进行特征向量求解,而是针对人脸关键部位^[18]。

参考文献:

[1] 王科俊,姚向辉. 人脸图像检测与识别方法综述[J]. 自动

化技术与应用,2004,23(12):5-8.

[2] 潘峥嵘,贺秀伟. 人脸表情识别在智能机器人中的应用研究[J]. 计算机技术与发展,2018,28(2):173-177.

[3] 崔凤焦. 表情识别算法研究进展与性能比较[J]. 计算机技术与发展,2018,28(2):145-149.

[4] 刘 焱. 基于 Gabor 特征和 Adaboost 算法的人脸表情识别研究[D]. 武汉:中南民族大学,2011.

[5] 李文革. 基于主成分分析的人脸识别[D]. 济南:山东大学,2008.

[6] 王 超. 特征提取技术在人脸识别中的研究与应用[D]. 无锡:江南大学,2009.

[7] 冯 翔. 基于人脸对齐和多特征融合的人脸识别方法研究[D]. 南京:南京理工大学,2015.

[8] 薛雨丽,毛 峡,郭 叶,等. 人机交互中的人脸表情识别研究进展[J]. 中国图象图形学报,2009,14(5):764-772.

[9] 陈秀端. 基于 Gabor 与 Pca 融合算法的人脸识别研究[D]. 厦门:厦门大学,2013.

[10] ESSAI A, PENTLAND A P. Coding, analysis, interpretation, recognition of facial expressions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7):757-763.

[11] 李彦璇,王 蓉,方银旺. 视频人脸检测与追踪方法研究[J]. 电脑知识与技术,2015,11(28):131-135.

[12] 李 康,尚 鹏,耿国华. 基于深度图像的人脸模型特征点自动标定[J]. 计算机科学,2014,41(4):287-291.

[13] 谢玉鹏,吴海燕. 基于 AAM 的人脸图像描述与编码[J]. 计算机仿真,2009,26(6):272-276.

[14] 高 文,金 辉. 面部表情图像的分析与识别[J]. 计算机学报,1997,20(9):782-789.

[15] 金 辉,高 文. 人脸面部混合表情识别系统[J]. 计算机学报,2000,23(6):602-608.

[16] PANTIC M, ROTHKRANTZ L J M. Expert system for automatic analysis of facial expressions[J]. Image and Vision Computing, 2000, 18(11):881-905.

[17] BARTLETT M S. Independent component representations for face recognition [C]//Face image analysis by unsupervised learning. [s. l.]:Springer, 2001:39-67.

[18] LYONS M J, BUDYNEK J, AKAMATSU S. Automatic classification of single facial images[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1999, 21(12):1357-1362.