

引用格式:郭文强,高文强,肖秦琨,等.基于小数据集下贝叶斯网络建模的面部表情识别[J].科学技术与工程,2018,18(35):179—183
Guo Wenqiang, Gao Wenqiang, Xiao Qinkun, et al. Facial expression recognition with small data sets based by Bayesian network modeling [J]. Science Technology and Engineering, 2018, 18(35): 179—183

基于小数据集下贝叶斯网络建模的面部表情识别

郭文强¹ 高文强¹ 肖秦琨² 徐成¹ 李梦然¹

(陕西科技大学电气与信息工程学院¹,西安工业大学电子信息工程学院²,西安 710021)

摘要 针对面部表情识别过程中获得的特征样本稀少的问题,提出了一种基于小数据集下贝叶斯网络(BN)建模的面部表情识别方法。首先提取面部表情图像的几何特征和HOG特征,经特征融合和归一化等处理构成动作单元(AU)标签样本集;其次提出了用于面部表情识别的BN结构,并将定性专家经验转化为BN条件概率之间的约束集合,随后引入凸优化最大化求解完成BN模型参数的估算;最后利用联合树推理算法识别出面部表情。实验结果表明:在小数据集条件下,与支持向量机(SVM)、Adaboost和卷积神经网络(CNN)等人脸表情分类方法相比,该方法能够取得更准确的面部表情识别结果。

关键词 面部表情识别 活动单元 小数据集 贝叶斯网络建模

中图分类号 TP301.6;

文献标志码 A

随着人们对人类行为分析、新的人机通信和多媒体检索技术的应用越来越感兴趣,面部表情识别成为一个快速增长的研究领域^[1]。

大多数面部表情识别方法是将面部行为分成活动单元(action unit, AU),它们代表着面部某部位肌肉的变化,通过识别AU实现对面部表情的识别。文献[2]提出了一种利用卷积神经网络(convolutional neural network, CNN)实现面部表情分类的方法,但是CNN的训练依赖大量的数据样本且训练时间较长;文献[3]提出了基于支持向量机(support vector machines, SVM)方法对面部表情进行分类,但SVM通常用来解决二值分类问题,在多分类问题应用上有一定的局限性。贝叶斯网络(Bayesian network, BN)在表示和推理不确定性知识方面拥有独特优势,文献[4]基于AU及它们之间的关系,提出了一种利用BN测量自发面部活动单元的强度的方法,获得了较好的识别效果。但是上述方法在模型训练和参数学习中均需要依赖大量的数据样本集。然而,从某些系统中获得充足的参数建模数据十分困难,例如医疗诊断系统中的患者病例表征数据、安保系统中识别出恐怖分子样本数据等,这样的数据集往往表现为体量较少,常被称为小数据集^[5]。

本文提出了一种基于小数据集下BN建模的面部表情识别方法。首先提取图像的几何特征和方向梯度直方图(histogram of oriented gradient, HOG),经过特征融合和归一化等处理,得到相应表情的AU标签样本集,再利用AU标签样本集、参数约束集合和凸优化成果训练得到BN模型;最后利用联合树(junction tree, JT)推理算法进行推理,实现面部表情的分类识别。

1 相关理论

1.1 贝叶斯网络

贝叶斯网络常用 $B(G, \theta)$ 来表示。其中, G 是一个具有多个节点和有向边的无环图; θ 代表与每个节点相关的条件概率表(conditional probability table, CPT)^[6]。BN参数学习是在网络拓扑结构已建立好的基础上,在某些条件下依照一定的准则,利用数据样本集来学习、计算BN的参数分布。

1.2 面部特征点定位

特征点定位是几何特征提取的关键,本文采用基于模型的条件局部神经网络方法(conditional local neural fields, CLNF)^[7],如图1中白色点群所示为面部根据图像选取的特征点位置。原图像来源于Cohn-Kanade表情数据库^[8]。已有算法可方便定位出眼、鼻、嘴、眉及脸周围的68个点作为特征选取的关键点。

1.3 AU组合

由面部特征点可进一步获取具有直观意义的AU组合表示^[9]。如图1所示为快乐表情面部特征

2018年8月16日收到

国家自然科学基金(61271362)、

陕西省科技厅自然科学基金(2017JM6057)和

陕西省教育厅2018年度服务地方科学研究计划(18JC003)资助
第一作者简介:郭文强(1971—),博士,副教授。研究方向:人工智能,模式识别。E-mail:guowenqiang@sust.edu.cn。

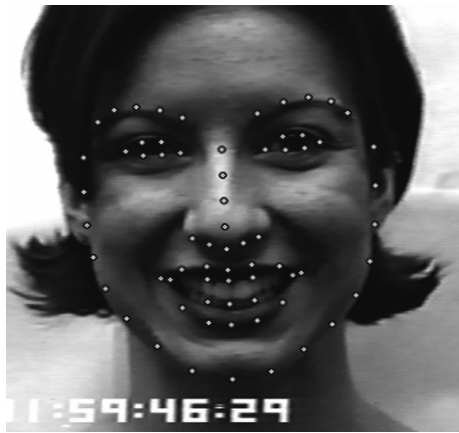


图1 关键特征点位置
Fig. 1 Key feature point location

点定位。由文献[9]可得到对应的 AU 组合。其中, AU6 + AU12 + AU25 表现为“下嘴轮廓上翘,左右眼眉角下降,双唇分开暴露牙齿”事件的组合,代表“快乐”的表情(详见文献[1])。六种常见的基本表情出现的 AU 组合如表 1 所示。

1.4 凸优化

凸优化问题可表示为形如式(1)、式(2)的问题^[10]。

maximize $f_0(x)$ (1)

subject to $f_r(x) \leq 0; \quad r = 1, 2, \dots, m$ (2)

式中, f_0 和 f_r 为凸函数, m 取自然数。式(2)对应的是不等式约束集合和等式约束集合构成的凸集。

表1 六种基本表情的 AU 组合
Table 1 AU combination of six basic expressions

组合	快乐	惊讶	害怕	愤怒	厌恶	悲伤
AU1	1	2	2	1	1	2
AU2	1	2	1	1	1	1
AU4	1	1	2	2	2	2
AU5	1	2	1	1	1	1
AU6	2	1	1	1	1	1
AU7	1	1	2	2	2	1
AU9	1	1	1	1	2	1
AU12	2	1	1	1	1	1
AU15	1	1	1	1	1	2
AU17	1	1	1	2	2	2
AU23	1	1	1	2	1	1
AU24	1	1	1	2	1	1
AU25	2	2	2	1	2	1
AU27	1	2	1	1	1	1

注:“1”表示“不发生”,“2”表示“发生”。

在 BN 建模过程中,定性的专家经验常可以表示为条件概率上的约束集合。因此,BN 参数若能表示为式(1)、式(2)的形式,则可利用凸优化求解实现 BN 参数计算。

2 基于小数据集下 BN 建模的面部表情识别算法

2.1 基于 HOG 的面部特征提取

如图 2 所示为 HOG 特征提取过程,本文选用的 HOG 特征具提取过程如下^[11]:

步骤一 将 image 分成 $N \times N$ 块 $M \times M$ 像素大小相连的 cell。

步骤二 计算 cell 各像素点梯度,统计梯度方向和边缘的直方图。

步骤三 把得到的直方图组合起来构成每个 cell 的特征。

步骤四 把 $L \times L$ 个 cell 组成 block,计算 block 中各直方图的密度。

步骤五 根据步骤四计算得到的密度,对每个 block 做归一化处理,得到 block 特征。

步骤六 组合所有 block 的特征向量构成最终的 HOG 特征。

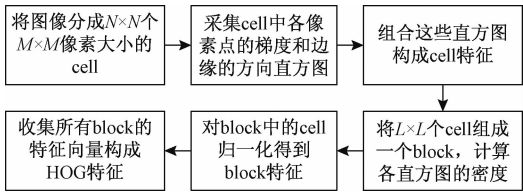


图2 HOG 特征提取过程
Fig. 2 HOG feature extraction process

2.2 构建贝叶斯网络的结构

采用的面部表情识别贝叶斯网络结构如图 3 所示。

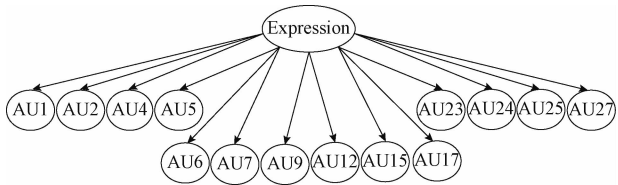


图3 贝叶斯网络结构
Fig. 3 Bayesian Network structure

图 3 中,Expression 节点代表面部表情的状态, AU 节点代表可能由表情触发的活动单元;显然,面部表情的状态可由 AU 状态所表征。

2.3 基于小数据集下 BN 建模的面部表情识别算法

领域专家往往可以获取一些类似如下的定性约束: $P(AU6 = 2 | Expression = \text{“快乐”}) > P(AU6 = 1 | Expression = \text{“快乐”})$ 。其含义为在“快乐”的表情条件下,“AU6 发生的概率大于 AU6 不发生的概率”。若将此类专家经验以约束的形式引入 BN 参

数学习机制,可以缩小 BN 参数学习、寻优的空间,从而提高 BN 参数估算精度。

基于以上分析,本文提出了一种在小数据集下利用 BN 建模算法进行面部表情识别的方法。具体步骤如下:

第一步 根据 2.1 节得到面部表情识别所需的 HOG 特征,再经过特征融合和归一化等处理构成相应的 AU 标签样本数据集 D。

第二步 判断 BN 是否已经结构建模。若是,则跳转至第三步;否则,根据 2.2 节描述得到 BN 模型结构 G。

第三步 判断 BN 是否已经参数建模。若已参数建模,则跳转至第七步;否则执行第四步至第六步,完成 BN 参数建模,得到模型参数 CPT。

第四步 根据样本数据集 D 统计样本量 N_{ijk} ,即样本数据中父节点状态为 j 、第 i 个节点取第 k 个状态的统计值,再根据式(3)计算得到 θ_{ijk} 。

$$\theta_{ijk} = \frac{N_{ijk}}{\sum_k N_{ijk}} \tag{3}$$

若 $N_{ijk} = 0$, 则令 $N_{ijk} = 0.002$ 。

第五步 将专家经验根据式(4)形成约束集合 ζ ; BN 节点参数可描述为类似于式(4)的等式和不等式约束集合,即:

$$\sum_{\theta_{ijk} \in \theta_A, \alpha_{ijk} \in \alpha_A} \alpha_{ijk} \theta_{ijk} \leq \alpha \tag{4}$$

式(4)中, θ_A 表示参数序列; α_A 是相应的常数序列; α 是一个常数且 $\alpha \geq 0$ 。

第六步 根据 θ_{ijk} , 约束集合 ζ , 即式(4)以及目标函数式(5)进行参数优化确定出 BN 参数 CPT。

$$\max_{\theta} - \sum_{i,j,k} \theta_{ijk} \ln \theta_{ijk} \tag{5}$$

显然式(5)为凸函数可利用文献[10]提出的凸优化求解工具得到 BN 参数 CPT。至此,用于面部表情识别的 BN 建模完成,返回第一步,获取表情识别样本数据集。

第七步 利用 BN 理论中的联合树(junction tree)推理算法进行 BN 推理,得到面部表情属性概率 Ψ' 。

第八步 判别面部表情。若面部表情属性概率 Ψ' 小于等于阈值 Ψ , 则返回第一步,重新获取新证据样本;否则输出面部表情属性,即面部表情识别结果。

3 仿真实验与分析

3.1 实验条件

实验所使用的数据库 Cohn-Kanade 表情数据

库^[8],选取其中的六种基本表情图像:1 = “快乐”, 2 = “惊讶”, 3 = “害怕”, 4 = “愤怒”, 5 = “厌恶”, 6 = “悲伤”。目标属性概率阈值 Ψ 取 0.78。仿真实验平台为 Windows10 系统,处理器为 Intel CPU 2.6 GHz,编程工具为 MATLAB R2014a。

3.2 BN 建模与识别分析

HOG 特征提取:将 112×112 像素的原图像,分解成 14×14 个 8×8 像素大小的 cell,由 2×2 个 cell 组成一个 block,于是得到 12×12 个 block,每个 block 可以提取 31 维特征,于是每幅图像可得 $12 \times 12 \times 31 = 4\,464$ 维特征向量。采用主成分分析(PCA)降到 1 391 维特征向量,再借助 OpenFace 工具^[12]构成实验所需的 AU 样本集。

根据本文方法得到的用于面部表情识别的 BN 结构,如图 3 所示。

根据领域专家经验得到不等式约束备选集合 ζ 。表 2 所示为以 AU25 为例得到 BN 参数模型的 6 组等式约束和 4 组不等式约束。

表 2 AU25 节点参数的约束集合

Table 2 Constraint set of parameters of AU25 node	
编号	约束集合
1	$P(\text{AU25} = 1 \text{Expression} = 1) + P(\text{AU25} = 2 \text{Expression} = 1) = 1$
2	$P(\text{AU25} = 1 \text{Expression} = 2) + P(\text{AU25} = 2 \text{Expression} = 2) = 1$
	$\vdots$
6	$P(\text{AU25} = 1 \text{Expression} = 6) + P(\text{AU25} = 2 \text{Expression} = 6) = 1$
7	$P(\text{AU25} = 1 \text{Expression} = 1) - P(\text{AU25} = 2 \text{Expression} = 1) < 0$
8	$P(\text{AU25} = 1 \text{Expression} = 2) - P(\text{AU25} = 2 \text{Expression} = 2) < 0$
9	$P(\text{AU25} = 1 \text{Expression} = 3) - P(\text{AU25} = 2 \text{Expression} = 3) < 0$
10	$P(\text{AU25} = 1 \text{Expression} = 5) - P(\text{AU25} = 2 \text{Expression} = 5) < 0$

根据文献[5]可知,在上述贝叶斯网络结构下,参数学习所需的最小样本数据量:

$$\frac{1}{2} \frac{(1 + \varepsilon^2)}{\varepsilon^2} \lg \frac{Nk^{d+1}}{\delta} \approx 202 \tag{6}$$

式(6)中,网络节点的数量 $N = 15$, 节点最大状态数 $k = 6$, 最大父节点数 $d = 1$, KL 误差 $\varepsilon = 0.1$, 置信度 $\delta = 0.05$ 。因此,数据量小于 202 的数据集称为小数据集。

Cohn-Kanade 库中每种表情可选取 262 张图像(图像序列的最后几帧图像)作实验数据集。在此数据集中,每种表情选取 140 组 AU 样本数据用于参数模型训练,因此是个典型的小数据集建模问题。另外的 122 组 AU 样本数据用于测试。

采用本文提出的算法进行 BN 建模与表情识别实验,实验结果如表 3 所示。在相同条件下,利用文献[13]中所述的 SVM 建模算法、Adaboost 算法和 CNN 算法实验结果同样也列于表 3 中。

表 3 不同建模算法识别率结果对比

Table 3 Recognition results comparison of different modeling algorithms

面部表情	建模方法识别率/%			
	SVM	Adaboost	CNN	本文法
快乐	68.31	72.89	69.31	97.54
惊讶	65.73	73.72	75.14	100
害怕	70.14	70.23	70.64	72.13
愤怒	71.45	65.18	80.15	94.26
厌恶	73.81	71.01	76.81	80.33
悲伤	65.89	71.25	75.83	80.80

对比实验结果表明:在小数据集条件下,本文方法的平均识别率以及每种表情的识别率均高于 SVM、Adaboost 和 CNN 方法的识别率,验证了本文所提方法的有效性。分析可知,本文方法充分利用建模参数间的约束关系,缩小了 BN 参数寻优的空间,有效弥补了建模训练数据样本集小的不足,从而改善了表情识别 BN 的建模性能,尤其是提高了 BN 参数估算精度。

4 结论

针对面部表情识别过程中获得的特征样本较少的建模难题,提出了一种基于小数据集下 BN 建模的面部表情识别方法。该方法根据面部表情图像的构成 AU 活动单元标签样本集,并引入定性专家经验形成 BN 参数条件概率之间的约束集合,随后引入凸优化求解完成 BN 模型的参数学习,最后利用 BN 推理算法识别出面部表情。实验结果表明:在小数据集条件下,本文提出的方法充分利用建模参数间的约束关系,缩小了 BN 参数的寻优空间,有效弥补了面部表情识别图像训练数据样本集小的不足,提高了 BN 建模的估算精度,取得了较好的面部表情识别结果。

参 考 文 献

1 Valstar M F, Almaev T, Girard J M, *et al.* FERA 2015-second facial expression recognition and analysis challenge. International Conference and Workshops on Automatic Face and Gesture Recognition.

New York: IEEE,2015:1—8

2 Ivanovsky L, Khryashchev V, Lebedev A, *et al.* Facial expression recognition algorithm based on deep convolution neural network. Open Innovations Association (FRUCT), 21st Conference. New York: IEEE, 2017: 141—147

3 孔英会,陈咨彤,车轱辘. 基于关键子区域及特征提取的表情识别. 科学技术与工程,2017;17(34):257—262

Kong Yinghui, Chen Zitong, Che Linlin. Facial expression recognition based on key sub-region and feature extraction. Science Technology and Engineering, 2017; 17(34):257—262

4 Li Y, Mavadati S M, Mahoor M H, *et al.* Measuring the intensity of spontaneous facial action units with dynamic Bayesian network. Pattern Recognition, 2015; 48(11): 3417—3427

5 Höffgen K U. Learning and robust learning of product distributions. Proceedings of the sixth annual conference on computational learning theory. New York:ACM, 1993: 77—83

6 Heckenman D, Dan G, Chickering D M. Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning, 2013; 20(3):197—243

7 Baltrusaitis T, Robinson P, Morency L P. Constrained local neural fields for robust facial landmark detection in the wild. International Conference on Computer Vision Workshops. New York: IEEE, 2013:354—361

8 Kanade T, Cohn J F, Tian Y. Comprehensive database for facial expression analysis. International Conference on Automatic Face and Gesture Recognition. New York:IEEE, 2000:46—53

9 Tong Y, Liao W, Ji Q. Facial action unit recognition by exploiting their dynamic and semantic relationships. Transactions on Pattern Analysis & Machine Intelligence,2007; 29(10):1683—1699

10 Boyd S, Vandenberghe L. Convex optimization. Cambridge: Cambridge University Press, 2004

11 Nigam S, Singh R, Misra A K. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. Multimedia Tools & Applications, 2018;(12):1—23

12 Baltrusaitis T, Robinson P, Morency L P. OpenFace: an open source facial behavior analysis toolkit. Winter Conference on Applications of Computer Vision. New York:IEEE, 2016:1—10

13 徐琳琳,张树美,赵俊莉. 基于图像的面部表情识别方法综述. 计算机应用,2017;37(12):3509—3516,3546

Xu Linlin, Zhang Shumei, Zhao Junli. Summary of facial expression recognition methods based on image. Journal of Computer Applications, 2017; 37(12): 3509—3516,3546

Facial Expression Recognition with Small Data Sets Based by Bayesian Network Modeling

GUO Wen-qiang¹, GAO Wen-qiang¹, XIAO Qin-kun², XU Cheng¹, LI Meng-ran¹
(School of Electrical and Information Engineering, Shaanxi University of Science and Technology¹,
College of Electronic Information Engineering, Xi'an Technological University², Xi'an 710021, China)

[**Abstract**] Aiming at the problems of facial expression recognition with the scarce feature samples, a Bayesian network (BN) modeling facial expression recognition method based on small data sets is proposed. Firstly, the geometric feature and HOG feature of facial expression image are extracted, the feature fusion and normalization are used to form the set of action unit label samples; Secondly, the BN structure is constructed according to the expert experience, and the qualitative expert experience is transformed into the constraint set of the BN conditional probability, and the BN model parameter is derived in the light of the convex optimization; Finally, junction tree inference algorithm was used to recognize facial expressions. The experimental results show that the method can obtain more accurate facial expression recognition results than support vector machine, Adaboost or convolution neural network method under the condition of small data sets.

[**Key words**] facial expression recognition action unit small data sets Bayesian network modeling