

文章编号: 2095-4980(2022)07-0718-04

雷声数据库的建立与研究

王珂, 张伟, 赵慧, 王运

(江苏开放大学 信息化建设处, 江苏 南京 210036)

摘要: 针对气象观测领域的研究与应用需要, 设计并建立了一种面向气象观测的雷声数据库, 其中包含闷雷 28 个、霹雳 33 个和干扰声 61 个, 共 122 个样本。首先对样本进行采集、制定命名规则并做标签, 再针对此数据库使用支持向量机(SVM)与组稀疏最小二乘回归(GSLSR)进行基准实验并对样本进行评价。仿真实验结果表明, 雷声数据库的建立为气象领域的研究提供了应用基础和前提。

关键词: 雷声数据库; 支持向量机; 组稀疏最小二乘回归; 交叉验证

中图分类号: TP183

文献标志码: A

doi: 10.11805/TKYDA2020366

Construction and research of thunder database

WANG Ke, ZHANG Wei, ZHAO Hui, WANG Yun

(Information Construction Office, Jiangsu Open University, Nanjing Jiangsu 210036, China)

Abstract: Aiming at the research and application needs in the field of meteorological observation, a thunder database for meteorological observations is designed and constructed, including stunned thunder, sputum and interference sounds, with 28, 33 and 61 samples, respectively, a total of 122 samples. Firstly, samples are collected, naming rules are formulated and labeled, and then the samples of the database are evaluated by using Support Vector Machine(SVM) and Group Sparse Least Squares Regression(GSLSR) model as a benchmark. The experimental results show that the establishment of the thunder database provides an important application basis and premise for the research of the meteorological field.

Keywords: thunder database; Support Vector Machine; Group Sparse Least Squares Regression; cross-validation

雷电灾害对人类生活造成较大的影响, 如在雷暴日, 房屋、建筑物等都会被严重损坏^[1], 造成严重的经济损失。日常生活中存在多种报警系统, 而雷声是多种声音报警系统的干扰源^[2]。

目前, 国内对于雷声的识别问题研究较少, 且公开可用雷声数据库并未建立, 因此建立一个数据一致性好、权威性高、接受度广的雷声数据库具有重要意义。为证明数据库的可靠性, 分别将闷雷与霹雳、闷雷与干扰声、霹雳与干扰声、雷声与干扰声进行分类识别。

对此数据库中样本研究, 将 SVM 与 GSLSR 作为评价数据库的基准方法。其中 SVM 分类算法^[3]常常作为一种有效的分类识别方法, 该方法所建立的模型比较客观; 同时为充分验证此数据库的可靠性, 对其音频数据的特征进行选择, 将 GSLSR 模型^[4]作为一种与 SVM 分类器的对比。

实验结果表明, 对于雷声与干扰声的分类识别, SVM 与 GSLSR 模型的分类效果都较好且对闷雷与霹雳的分类识别准确率均在 90% 以上, 而 GSLSR 模型的分类识别效果更好, 由此可见选取数据库中样本合适的特征可对雷声数据进行有效的识别。

1 雷声数据库的建立

1.1 雷声数据的采集

在数据库的建立过程中, 为保证数据一致性, 截选《英国广播公司(British Broadcasting Corporation, BBC)音

收稿日期: 2020-08-04; 修回日期: 2020-09-18

基金项目: 江苏高校哲学社会科学研究基金资助项目(2021SJA0763); 江苏开放大学“十三五”学校发展对策研究课题基金资助项目; 江苏开放大学智慧校园建设研究基金资助项目(19SSWZ-03); 江苏省教育科学“十四五”规划 2021 年度课题研究(C-B-2021-03-03)

效库合辑》中音频数据，这是一个声音效果数据库，拥有 16 016 种经典的英国广播公司档案库音效和现场录音，范围涵盖从空袭音效到各种雷声、斑马嘶鸣等音效。所有的音效都采用波形声音文件(Wave，WAV)格式，在遵循 RemArc 许可条款下下载使用。该音效是 BBC 的版权，可用于个人、教育或研究目的。此数据库共包含相关雷声音频 12 类，共含音频数据 122 条，其中包含雷声 61 个样本，对雷声识别干扰较大的非雷声样本 61 个。对于数据库中的每个音频数据，首先将其转换为标准的 WAV 格式、时间长度 3 s、采样频率 8 000 Hz、采样精度为 16 位、声道为单声道的样本，该数据库中部分音频数据的名称和所包含的文件数量见表 1。

1.2 数据库命名规则及标注

为便于数据库的研究，将其 12 类文件重新划分为 3 类，分别是闷雷、霹雳和其他，并对样本的命名规则^[5-10]进行重新制定。一个样本的名称由两部分组成，即类别名称缩写与样本编号，其命名规则见表 2。例如，“r2.wav”表示闷雷的第 2 个样本，每个样本都对应一个标注文件，用于区分样本的种类与其序号。

2 雷声识别

2.1 特征提取及选择

选取有效的特征对雷声的分类识别研究很重要，首先使用 openSMILE-2.1.0 对采集的 122 个音频样本进行特征提取，其中 openSMILE 为特征提取提供了各种特征集，而提取的 IS09 特征集能有效地对样本特征进行表示，其包含 384 个特征，主要包含均方误差(Mean Square Error，MSE)、美尔倒谱系数(Mel Frequency Cepstrum Coefficient，MFCC)、时间信号的过零率(基于帧)、发声概率和基音频率(F0)等特征。

2.2 支持向量机模型

支持向量机(SVM)是一种以统计学理论为基础的有监督的机器学习模型，常用来解决回归和分类问题。支持向量机属于一般化的线性分类器，具有对数据敏感、计算速度快、泛化性能好等优点，针对高维问题、非线性问题、小样本情况下的学习问题有着较好的效果。其原理是：将训练样本数据(输入向量)通过用内积函数定义的核函数(非线性映射函数)从低维空间映射到高维空间，然后在高维空间获取最大间隔的超平面。SVM 的目的是获取最大间隔的超平面，其通过找到一个平面，有效地把训练样本数据分隔在平面两侧，且能够保证分隔后的 2 类数据间的距离为极大值，可以保证分类结果的高可靠性。

SVM 分类器的工作流程^[11]如图 1 所示。由于线性核函数主要用于线性可分的情形，其有参数少、速度快等优点，对于雷声与非雷声的二分类识别其分类效果比较理想，在此实验中选用线性核函数。为估计模型分到每类的概率，设置不同的参数获取输出分类的概率。在训练过程中，实验中选取 IS09 特征集，然后对其进行归一化，在对模型进行训练后，测试时选用十折交叉验证的方法对其进行测试，最后得到实验结果。



Fig.1 Workflow of SVM classifier
图 1 SVM 分类器的工作流程

2.3 组稀疏最小二乘回归模型

GSLSR 模型的主要思想是将数据库样本特征投影到标签空间，并通过增加 l_2 范数来实现特征选择。令 $\mathbf{X}$ 为数据库中 $\mathbf{X} \in \mathbf{R}^{k \times d}$ 的样本特征矩阵，其中 k 表示样本个数， d 表示特征维度，则第 i 个样本特征可表示为 $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{ik}]^T$ ，令 $\mathbf{L}$ 为标签矩阵，其中第 i 列 $\mathbf{l}_i$ 是一个与第 i 个音频数据相关的 $n \times 1$ 类别标签向量， n 表示类别的

表 1 数据库结构

Table1 Database structure

number of documents		number of documents	
setting sound	4	international voice	1
adventure sports	3	livestock voice	1
explosion sound	26	mine	28
thunderbolt	33	the train	6
urban sound	1	acoustic	8
factory sound	1	weather sound	10

表 2 命名规则

Table2 Naming rules

type of document	abbreviations	sample number
thunderbolt	c	1~33
mine	r	1~28
other	o	1~61

数目, 则 $I_i = [l_{i1}, l_{i2}, \dots, l_{in}]^T$, 二进制值判定:

$$l_{ij} = \begin{cases} 1, x_i \text{ 属于第 } i \text{ 类} \\ 0, \text{其他类} \end{cases} \quad (1)$$

GSLSR 模型的目标函数:

$$\min_c \left\| L - \sum_{i=1}^l c_i^T x_i \right\|_F^2 + \lambda \sum_{i=1}^l \|c_i^T\|_2 \quad (2)$$

式中: 第 1 项为 F 范数的平方, 其目的是建立标签与特征之间的联系; 第 2 项为 l_2 范数, 其目的是对特征进行选择; λ 为调节参数 ($\lambda > 0$)。上述问题最优化求解可用增广拉格朗日乘子法 (Augmented Lagrange Method, ALM) 来求解^[12]。

3 实验结果与分析

3.1 实验设置

在此实验过程中, 采用十折交叉验证方法对数据库进行验证, 即将 122 个样本随机分为 10 份, 其中轮流将 9 份作为训练集, 1 份作为测试集对标签进行预测。为保证对算法准确性的估计, 进行多次十折交叉验证并求其平均值。其中使用 SVM 与 GSLSR 模型对雷声进行识别时, 综合准确率 (Comprehensive Accuracy, CPA) 与召回率 (Recall, REC) 的评估指标, 使用其调和平均值 (F1score) 作为综合反映整体的指标。通常在比较算法的性能时还将等错误率 (Equal Error Rate, EER) 作为一个评价指标^[13], 其通过调整阈值使得误拒率 (False Rejection Rate, FRR) 与误识率 (False Acceptance Rate, FAR) 相等, 此时其值即为等错误率。其相关公式定义如下:

$$REC = \frac{\text{正确分类的样本数}}{\text{识别的相关样本}} \times 100\% \quad (3)$$

$$F1score = \frac{2 \times \text{正确率} \times \text{召回率}}{\text{正确率} + \text{召回率}} \times 100\% \quad (4)$$

$$FRR = \frac{\text{本该匹配成功的判为失败}}{\text{总的匹配成功次数}} \times 100\% \quad (5)$$

$$FAR = \frac{\text{本该匹配失败的判为匹配成功}}{\text{总的匹配失败次数}} \times 100\% \quad (6)$$

3.2 实验结果与分析

对于雷声的分类识别, 其实验结果见表 3。表中: I 为雷声干扰实验, II 为霹雳与干扰实验, III 为闷雷与干扰实验。

实验结果表明, 2 种模型对此数据库的识别效果均较好。在以上 3 组实验中, 使用 SVM 对于霹雳与干扰声识别效果最好, 其次是雷声与干扰声的分类识别, 闷雷与干扰声的识别效果较差; 使用 GSLSR 模型对于霹雳与干扰声的分类识别效果最好, 其次是雷声与干扰声、闷雷与干扰声。由此可见, 对于霹雳与干扰声的分类识别, 使用传统的 SVM 模型与 GSLSR 模型都可以得到很好的分类效果, 相比 SVM 模型, GSLSR 模型的分类效果更好, 其原因在于 SVM 模型选用 IS09 特征集中的特征参数, 而 GSLSR 模型在 IS09 特征集的基础上对其特征进一步选择, 使得识别效果更好。

表 3 实验结果

Table 3 Experimental results

experimental	classification methods			
	SVM		GSLSR	
I	EER		EER($\lambda, 3, 4$)	
	0.528 5		0.500 0	
II	wac	meanF1	wac($\lambda, 2$)	meanF1($\lambda, 2$)
	0.926 9	0.916 8	0.933 3	0.920 9
III	wac	meanF1	wac($\lambda, 2, 3$)	meanF1($\lambda, 2, 3$)
	0.983 3	0.983 4	1	1

4 结论

为方便雷声识别的研究, 首先建立一个数据一致性好、权威性佳、接受度广的雷声数据库, 然后通过特征提取、特征选择以及特征归一化对数据进行处理, 最后使用 SVM 和 GSLSR 对数据库中雷声种类进行分类识别,

并对数据库做基准实验,实验结果表明,2种模型对数据库的识别效果较好,因此可将此数据库中数据样本用于实际问题的研究。此数据库虽然可取得较好的识别效果,但在雷声识别领域中相比自然环境采集的雷声样本仍有一定差距,如对真实环境中雷声的降噪等一系列问题,因此未来针对雷声样本的采集问题仍是研究的方向。

参考文献:

- [1] LI Lin,JI Hongbing. Signal feature extraction based on an improved EMD method[J]. Measurement, 2009,42(5):796–803.
- [2] UMATHY K,KRISHNAN S,RAO R K. Audio signal feature extraction and classification using local discriminant bases[J]. IEEE Transactions on Audio Speech & Language Processing, 2007,15(4):1236–1246.
- [3] LU N,WANG J,MCDERMOTT I,et al. Uterine electromyography signal feature extraction and classification[J]. International Journal of Modelling Identification & Control, 2009,6(2):136–146.
- [4] GUO Y,BENAMOUN M,SOHEL F,et al. A comprehensive performance evaluation of 3D local feature descriptors[J]. International Journal of Computer Vision, 2016,116(1):66–89.
- [5] YU M,LING S,ZHEN X,et al. Local feature discriminant projection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016,38(9):1908–1914.
- [6] HUANG D,TANG Y,WANG Y,et al. Hand-dorsa vein recognition by matching local features of multisource keypoints[J]. IEEE Trans Cybern, 2017,45(9):1823–1837.
- [7] ARMANFARD N,REILLY J P,KOMEILI M. Local feature selection for data classification[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016,38(6):1217–1227.
- [8] MATURILLI A,HELBERT J,MOROZ L. The Berlin Emissivity Database(BED)[J]. Planetary and Space Science, 2008,56(3–4): 420–425.
- [9] ALKHAWLANI M,ELMOGY M,ELBAKRY H. Content-based image retrieval using local features descriptors and bag-of-visual words[J]. International Journal of Advanced Computer Science & Applications, 2015,6(9):212–219.
- [10] ZHOU Weixun,SHAO Zhenfeng,HOU Jihu. Remote sensing imagery retrieval method based on visual attention model and local features[J]. Geomatics and Information Science of Wuhan University, 2015,40(1):46–52.
- [11] LIU Feng,SUN Dapeng,HUANG Yu,et al. Multi-component LFM signal feature extraction based on improved Wigner–Hough transform[J]. Transactions of Beijing Institute of Technology, 2008,46(4):1–4.
- [12] ZHENG Wenming,XIN Minghai,WANG Xiaolan,et al. A novel speech emotion recognition method via incomplete sparse least square regression[J]. IEEE Signal Processing Letters, 2014,21(5):569–572.
- [13] 刘默涵,侯巍. 基于光强变化场景的目标检测与跟踪新方法[J]. 太赫兹科学与电子信息学报, 2019,17(6):1000–1005. (LIU Mohan,HOU Wei. A new method of target detection and tracking based on light intensity changing scenes[J]. Journal of Terahertz Science and Electronic Information Technology, 2019,17(6):1000–1005.)

作者简介:

王珂(1993–),男,硕士,助理工程师,主要研究方向为人工智能、模式识别、图像处理, email: 459506810@qq.com.

赵慧(1995–),女,硕士,助理工程师,主要研究方向为音频处理、模式识别.

张伟(1995–),男,硕士,助理工程师,主要研究方向为信号处理、模式识别.

王运(1981–),男,硕士,高级工程师,主要研究方向为信息化教育、计算机科学与技术.