

蔡杏辉, 廖诗荣, 张燕明, 等. 基于支持向量机的地震事件类型自动识别及应用[J]. 华南地震, 2021, 41(2): 27-35. [CAI Xinghui, LIAO Shirong, ZHANG Yanming, et al. Automatic Identification of Earthquake Event Types Based on Support Vector Machine and Its Application[J]. South China journal of seismology, 2021, 41(2): 27-35]

基于支持向量机的地震事件类型自动识别及应用

蔡杏辉, 廖诗荣, 张燕明, 陈惠芳, 林彬华

(福建省地震局, 福州 350003)

摘要: 天然地震与非天然地震自动识别是地震自动编目系统的重要功能之一, 是监测数据产出智能化的基础应用。从福建天然地震和人工爆破事件中, 提取小波分析特征、P/S 震相振幅比、波形能量分布特征, 对以上特征组合联合支持向量机进行大批量数据测试分析, 研究得出识别效果较好的事件类型判别算法, 最优测试识别率为 94.5%; 采用最优算法研发基于支持向量机事件类型自动识别模块, 将研制的自动识别软件模块应用于福建台网的日常地震编目工作, 对 2019 年 8 月 1 日至 2020 年 2 月 29 日共计 1531 个日常触发事件进行准实时分类, 自动识别软件模块分类的正确率为 93.9%。另外, 采用 AMQ 消息中间件为信息中介, 从 redis 波形共享内存中获取事件波形, 实现自动识别模块与自动编目系统对接。

关键词: 事件类型; 特征组合; 自动识别模块; 支持向量机

中图分类号: P315

文献标志码: A

文章编号: 1001-8662(2021)02-0027-09

DOI: 10.13512/j.hndz.2021.02.04

Automatic Identification of Earthquake Event Types Based on Support Vector Machine and Its Application

CAI Xinghui, LIAO Shirong, ZHANG Yanming, CHENG Huifang, LIN Binhua

(Fujian Earthquake Agency, Fuzhou 350003, China)

Abstract: Automatic identification of natural and non-natural earthquakes is one of the important functions of earthquake automatic cataloging system, and it is the basic application of intelligent monitoring data output. The wavelet analysis features, P/S phase amplitude ratio, waveform energy distribution features and spectrum features were extracted from the natural earthquake and artificial blasting events in Fujian province to carry out large-scale data test and analysis on the combination of the above features and the Support Vector Machine, the algorithm of event type identification is proved to effective, and the optimal test recognition rate is 94.5%. An automatic event recognition module based on Support Vector Machine is developed by using the optimal algorithm, the automatic identification software module was applied to the daily earthquake cataloging work of Fujian Seismic Network, and a total of 1531 daily triggered events were quasi real time classified from August

收稿日期: 2020-08-10

基金项目: 监测预报科研三结合(3JH20200105); 中国地震局监测预报业务专项; 地震科技星火计划项目(XH20029) 联合资助

作者简介: 蔡杏辉(1976-), 男, 高级工程师, 主要从事地震监测工作。

E-mail: 13870511@sina.com

1st, 2019 to February 29th, 2020, the correct rate of automatic recognition software module classification was 93.9%. Using Amq message middleware as the information medium, the event waveform is obtained in Redis waveform shared memory, and the automatic recognition module is connected with automatic cataloging system.

Keywords: Event type; Feature combination; Automatic recognition module; Support Vector Machine

0 引言

地震事件类型分类是地震监测的必要工作,信息的迅速可靠对地方政府和公众至关重要。目前我国非天然地震的判别与分析主要靠人工进行,响应速度慢。虽然非天然地震事件直接破坏性普遍不大,但仍有一定数量的事件对社会产生重大影响,如2015年8月天津港爆炸,分别记录到 $M_L2.3$ 级与 $M_L2.9$ 级两次非天然地震,2019年3月江苏响水爆炸,记录到 $M_L2.2$ 级非天然地震等,这些突发事件对地震部门快速应急响应提出了新要求,非天然地震进行准实时处理,自动判定事件类型,建立规范化、标准化、自动化处理流程,可为政府的快速应急响应提供技术支撑。此外,由于事件波形复杂性,不能完全剔除非天然地震事件,会影响地震目录产出质量,导致天然地震活动性估计不准确,对疑似事件准确定性可提高地震编目产出质量和工作效率。

目前,已有许多学者对天然及非天然地震的识别进行研究,并取得相关研究成果。如BP神经网络、卷积神经网络、支持向量机、矩阵决策算法等方法被运用于天然地震与非天然地震自动识别分类^[1-6]。支持向量机(SVM)是人工智能算法的一个重要分支,是一类按监督学习方式对数据进行二元分类的广义线性分类器,在人像识别、文本分类、模式识别等许多领域中得到应用^[7]。已有研究者将支持向量机技术应用到地震事件类型的分类识别研究。国内,研究者^[2,3,5]利用波形小波特征,采用支持向量机对地震和爆破事件进行识别。国外,Kortstrom等(2016)^[7]提出了一种基于支持向量机的事件全自动分类方法,该方法通过SVM学习天然地震和人工地震在P、P波尾波、S、S波尾波等地震记录不同部分能量分布的差异来进行事件类型判别。Rabin(2016)^[8]应用机器学习算法中的扩散图方法,将地震图转换为标准化的超声波图,进行天然地震事件与核爆事件的自动识别。Saad等(2019)^[9]采用小波滤波器组联合支持向量机提出了天然地震事件与采石场人工爆破事件的自

动分类算法。研究结果表明,支持向量机用于事件分类具有很好的效果。

应用于事件自动识别的算法主要有神经网络和支持向量机,支持向量机在解决小样本、非线性及高维模式识别中表现出许多特有的优势,并能够推广应用到函数拟合等其他机器学习问题中^[10]。BP神经网络是包含多个隐含层的网络,具备处理线性不可分问题的能力,但BP神经网络也存在主要缺陷学习速度慢及容易陷入局部极小值;算法特点上支持向量机在样本一定的情况下运算量比BP要小,判别结果产出快;CNN卷积神经网络是深度学习的代表算法之一,其无需人为提取特定特征信息,但运算量大;非天然地震塌陷、矿震、滑坡为小样本事件,福建台网对塌陷、矿震等特殊地震动鲜有记录到;基于以上原因,因此本研究采用支持向量机作为自动识别分类器。本研究,在前人基础上充分利用福建台网记录的非天然地震、地震数字波形资料,采用波形能量分布、P/S振幅比、小波分析等多种方法,开展基于支持向量机事件类型自动识别研究及软件模块研发,以期提高地震事件类型识别的准确率和应用区域的广泛性。研发可在地震台网实时运行的软件模块,提高人工编目的效率和质量,实现地震事件类型产出自动化、智能化处理,为政府的快速应急响应提供技术支撑。

1 识别算法及基本原理

1.1 支持向量机原理

支持向量机方法的基本思想是:定义最优线性超平面,并把寻找最优线性超平面的算法归结为求解一个凸规划问题。进而基于核函数展开定理,通过非线性映射 φ ,把样本空间映射到一个高维乃至无穷维的特征空间,使在特征空间中可以应用线性学习机的方法解决样本空间中的高度非线性分类和回归等问题^[11]。简单地说就是升维和线性化。常见的核函数有多项式核、径向基函数核、拉普拉斯、Sigmoid核。基于RBF核可将样本

空间映射至无限维空间的考虑,采用径向基函数核作为支持向量机核函数,径向基函数核有两个参数 C 和 γ 。经过尝试多个 C 和 γ 配对组合,采用 C 为 100, γ 为 0.49 效果较好。对于测试样本,将提取的特征向量联合支持向量机构建自动识别算法,依其输出值 a 判定该事件的类型,判定标准为:当 $a=1$,判定为天然地震,当 $a=0$,判定为人工爆破(非天然地震事件)。

本文所应用的支持向量机算法(v-SVC) [12]:

(1) 设已知训练集, $T = \{(x_i, y_i) \cdots (x_m, y_m)\} \in (\chi \times \gamma)^m$ 其中 $x_i \in \chi = R^n$, n 为特征数, m 为样品数。

(2) 选取适当的核函数: $K(x_i, y_j)$ 和参数 ν , 构造并求解最优化问题。

$$\min \left[\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j a_i a_j K(x_i, x_j) \right]$$

$$s.t.: \sum_{i=1}^m y_i a_i = 0$$

$$0 \leq a_i \leq \frac{1}{m} (i=1, \cdots, m)$$

$$\sum_{i=1}^m a_i \geq \nu$$

得最优解 $a^* = (a_1^*, \cdots, a_m^*)^T$ 。

(3) 选取 $j \in S_+ = \{i | a_i^* \in (0, \frac{1}{m})\}$, $y_i = 1 | k \in S_- = \{i | a_i^* \in (0, \frac{1}{m}), y_i = -1\}$, 据此计算阈值

$$b^* = -\frac{1}{2} \sum_{i=1}^m y_i a_i^* (K(x_i, y_j) + K(x_i, y_j))$$

(4) 构造决策函数

$$f(x) = \text{sgn}(\sum_{i=1}^m y_i a_i^* K(x_i, y_j) + b^*)$$

1.2 分类识别器设计

自动识别系统关键算法有两个模块:一为特征提取模块,从数据中提取出识别所需特征;其二为识别判定模块,在特征空间中用模式识别方法把识别对象归为某一类别。在对事件类型的分类识别中,如何提取有效的识别特征是识别的关键。P/S 振幅比是研究最为深入的一个判别量,在地震与爆破识别中效果较好;小波分析算法被广泛应用于识别研究;Kortstrom(2016) [13] 方法是将地震信号通过数据处理,转换成“数值谱图”,在事件分类识别上取得了较好的效果。本文采用 P/S 振幅比、小波分析、波形能量分布特征等多种特征组合联合支持向量机(SVM)进行事件类型判定。对多种特征组合进行测试,得出识别效果较好的分类识别器。图 1 为分类识别器(SVM)设计示意图。

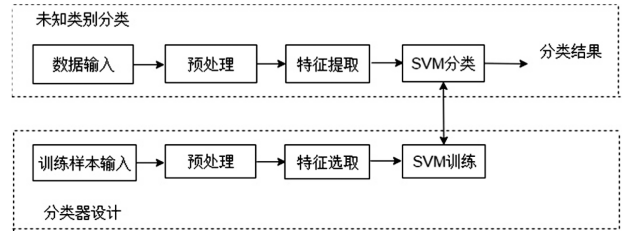


图 1 分类识别器设计示意图

Fig.1 Schematic diagram of classification recognizer design

1.3 小波分析

一维离散小波变换(DWT)能量比特征:

若 S 为原始信号,其长度为 J ,信号采样点序号为 j , S_i 为信号 S 分解后的第 i 个小波系数,其长度为 K , k 为其样点序号,则小波系数的能量比 (E_{wt}) 按(1)式定义 [3]。

$$E_{wt}(i) = \frac{100 \times \sum_{k=1}^K S_{i,k}^2}{\sum_{j=1}^J S_j^2} \quad (1)$$

小波包变换(WPT)小波包对数能量熵特征:

若 S 为原始信号,对它进行 n 层小波包分解后,得到第 n 层的小波包系数总共为 N 个。 S_i 为信号 S 分解后的第 i 个小波的系数,其长度为 J ,小波系数的结点序号为 j ,则从第 i 个小波的系数中提取出小波包对数能量熵 (E_{lg}) 特征按下式定义 [3]。

$$E_{lg} = \sum_{j=1}^J \lg[S_{i,j}^2] \quad (2)$$

1.4 波形能量分布

采用 20 个窄带通道滤波器过滤地震记录(滤波器为零相位二阶巴特沃斯滤波器,频带 1~41 Hz),并将事件波形分为四个阶段的窗口: P 和 P coda, S 和 S coda。然后,计算了每个滤波器通道和相位窗口的短期平均(STA)值,共 80 个判别参数 [7]。

短期平均(STA)值计算公式:

$$STA = \frac{1}{N} \sum_{i=1}^N y_i^2 \quad (3)$$

其中 N 为样本中 STA 窗口的长度, y_i 为第 i 个样本经过过滤的时间序列 y 。

2 数据资料及特征组合方式

2.1 数据资料

为使研究成果能应用于台网日常监测与编目

工作,需使学习样本泛化能力强。采用福建台网 2016 年 1 月至 2019 年 6 月所有触发事件,不进行震中区域、台站、震级的划分,事件数 6570 个,事件震级范围 M_L 1.1~4.0,包括极低震级单台事件,每个事件各提取出一个单台波形记录($\Delta < 100$ km),训练集和测试集按约 50%:50%的比例,随机将 6570 个事件中的 3253 个事件作为训练集,其余 3317 个事件作为测试集。训练集 907 个为天然地震记录,另 2346 个为人工爆破记录;测试集 2054 个为天然地震记录,另 1263 个为人工爆破记录。

2.2 特征提取及组合方式

对信号特征提取并截取有效波形,首先要确定 P 到时,对离线数据识别训练和测试,直接从 Jopens 震相数据表中获取台站震相 P 到时数据;处理波形窗长选择,选取 P 波到时前 0.5 s 至 P 波到时后 19.5 s,共 20 s 的单台 Z 分向和三分向事件波形(需包含完整的 P 与 S 波列)进行处理。

P/S 振幅比特征的提取:信号的 S 波和 P 波最大振幅比是天然地震和人工爆破波形信号中的一个重要特征,Colin 等^[14]研究表明 6~8 Hz 和 8~10 Hz 的滤波频带对爆破的识别效果较好,因此采用一阶 Butterworth 对原始波形进行滤波,共选取 3 个滤波频带,分别为 6.0~8.0、8.0~10.0、6.0~10.0,分别量取 P 与 S 波列的最大值,获取 3 个 P/S 振幅比特征值。小波特征的提取:对信号进行 4 层小波及小波包分解。对小波分解得到小波系数利用公式(1)提取出小波能量比(E_m)特征组成 5 维特征向量;对第 4 层的 16 个波包系数利用公式(2)提取小波包对数能量熵波形特征(E_{lg})分别组成 16 维特征向量。波形能量分布特征的提取:事件波形四个阶段的窗口,通过 20 个窄带通道过滤器滤波,利用公式(3)计算短期平均(STA)值,获得 80 个特征值。

将提取的特征进行组合,在特征组合方式选取上,采用以下 3 种组合方式:① P/S 振幅比+小波包对数能量熵;② P/S 振幅比+小波能量比+小波包对数能量熵;③ P/S 振幅比+小波包对数能量熵+波形能量分布(P, P code, S, S code)。分别组成 19 维、24 维、99 维特征向量,利用支持向量机检验各特征组合的分类能力。

3 识别能力检验及分析

3.1 各特征组合训练及测试结果

将测试集记录用上述方法进行特征提取并组合,用对应特征组合训练得到的支持向量机进行地震事件类型分类,检验各特征组合的分类能力。测试数据输入采用单台 Z 分向和三分向数据作为输入,首先采用单台 Z 分向数据测试各个特征组合分类效果,获得分类效果最好的特征组合,其次对该特征组合采用单台三分向数据测试分类效果,对比不同数据的输入方式对分类效果的影响。训练及测试结果见表 1。

表 1 各特征组合对训练集及测试集的分类效果

Table 1 Classification effect of each feature combination on training set and test set

特征指标序号	训练集准确率 /(%)(共 3253)	测试集准确率 /(%)(共 3308)
1	89.6	83.2
2	91.3	86.9
3-1	96.8	91.3
3-2	98.3	94.5

注:1. P/S 振幅比+小波包对数能量熵;2.P/S 振幅比+小波能量比+小波包对数能量熵;3.P/S 振幅比+小波包对数能量熵+波形能量分布(P, P code, S, S code)(单台 Z 分向);4. P/S 振幅比+小波包对数能量熵+波形能量分布(P, P code, S, S code)(单台 3 分向)。

由表 1 结果可知,3 种特征组合训练集和测试集准确率达到 80%以上,说明选用的特征指标可对事件进行有效分类,有效特征进行组合有助于识别率的提高;其中 P/S 振幅比+小波包对数能量熵+波形能量分布特征组合的识别率最高,自验证准确率达到 96.8%,分类准确率达到 91.3%;对该特征组合采用单台三分向数据输入,自验证准确率达到 98.3%,分类准确率达到 94.5%,分类识别率比单台单通道输入提高了 3.2%。

3.2 本文最优算法性能评估及事件实例

表 2 给出了本文最优算法性能评估指标结果,结果表明,灵敏度和特异度比较均衡,精确率达到 96.0%,说明算法比较好,对地震和爆破多数都能有效识别;灵敏度稍高于特异度,说明对地震

识别效果好于爆破;该算法对地震分类的判别准确率为 95.5%,对爆破分类的判别准确率为 93.4%。算法测试 3308 个事件,共消耗的时间 8270 s,平均每个事件全流程产出结果需耗时 2.5 s;测试环境:使用英特尔处理器核心 E5-26090@2.40 GHz,内存 8 GB, windows 10, 64 位操作系统。

图 2、3 分别是福建仙游爆破正确和错误识别事件仙游西苑台波形,表 3 是其波形特征计算参数,每个事件有 99 个计算参数。采用同一类型事件同一台站对正确和错误识别事件波形特征计算参数进行对比更具有可比性,对比发现特征序列的部分特征值存在较大的差别;对两者波形进行对比,错误识别事件波形背景噪声较大,可能事件波形背景噪声较大对特征计算参数结果有影响,导致事件错误识别。

表 2 本文最优算法性能评估指标
Table 2 The performance evaluation index of optimal algorithm given in this paper

评估指标	数值
TP	1962
TN	1180
FP	83
FN	92
灵敏度	95.5%
特效度	93.4%
精确率	96.0%
测试耗时	8270s

注:真正值(TP)代表地震的正确分类,真负值(TN)代表爆破的正确分类,假负值(FP)代表爆破的错误分类,假正值(FN)代表地震的错误分类,灵敏度= TP/(TP+FN),特异度= TN/(TN+FP),精确率=TP/(FP+TP)。

表 3 最优算法正确和错误事件波形特征计算参数
Table 3 The optimal algorithm calculates the parameters of the characteristic of the correct and wrong event waveform

事件时间/ 年-月-日	特征值								
2020-05-07	0.967	1.048	0.894	-0.229	-0.222	-0.218	-0.225	0.789	0.143
	0.190	0.090	0.074	0.073	0.069	-0.011	-0.291	-0.203	-0.097
	-0.240	-0.081	-0.106	-0.189	-0.210	0.003	0.001	0.001	0.002
	0.002	0.012	0.072	0.044	0.011	0.039	0.726	1.461	0.822
	2.556	3.760	0.973	0.982	1.267	0.002	0.004	0.003	0.003
	0.001	0.009	0.017	0.060	0.084	0.122	0.065	0.126	0.546
	1.131	0.316	0.616	1.447	1.026	1.708	2.831	0.001	0.004
	0.002	0.002	0.005	0.007	0.011	0.036	0.078	0.082	0.182
	0.193	0.353	0.372	0.746	1.693	11.481	4.794	0.831	0.585
	0.002	0.126	0.005	0.207	0.003	0.277	0.001	0.315	0.003
2019-05-24	1.788	0.009	1.451	0.007	8.272	0.038	15.085	0.116	5.871
	1.102	1.041	0.894	-0.225	-0.224	-0.220	-0.225	0.854	0.109
	0.193	0.061	-0.027	-0.034	0.0667	-0.075	-0.249	-0.162	-0.003
	-0.183	-0.118	-0.126	-0.115	-0.191	0.015	0.017	0.030	0.011
	0.005	0.008	0.018	0.006	0.011	0.023	0.045	0.262	0.956
	1.739	1.460	3.745	0.298	0.134	0.022	0.014	0.035	0.019
	0.015	0.008	0.013	0.003	0.004	0.003	0.006	0.005	0.011
	0.096	0.177	0.210	0.311	0.289	0.122	0.368	0.016	0.006
	0.021	0.023	0.012	0.013	0.019	0.010	0.003	0.009	0.022
	0.032	0.056	0.376	0.784	0.287	0.926	1.487	11.907	18.333
	0.011	0.007	0.005	0.009	0.005	0.005	0.000	0.028	0.020
	0.013	0.022	0.058	0.230	0.247	0.015	7.216	16.213	11.019

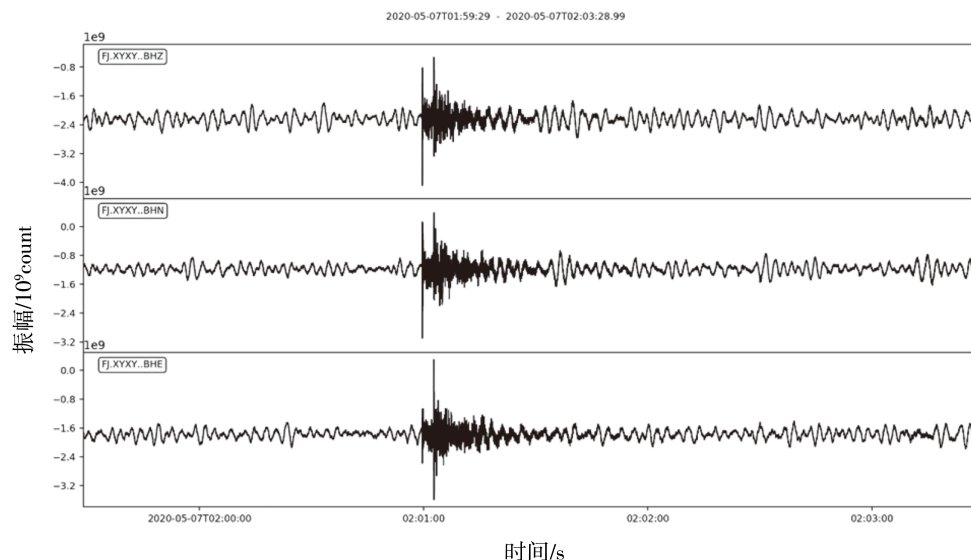


图2 福建仙游爆破正确识别事件波形(2020-05-07)

Fig.2 The waveform of the correct identification of the blasting event in Xianyou, Fujian

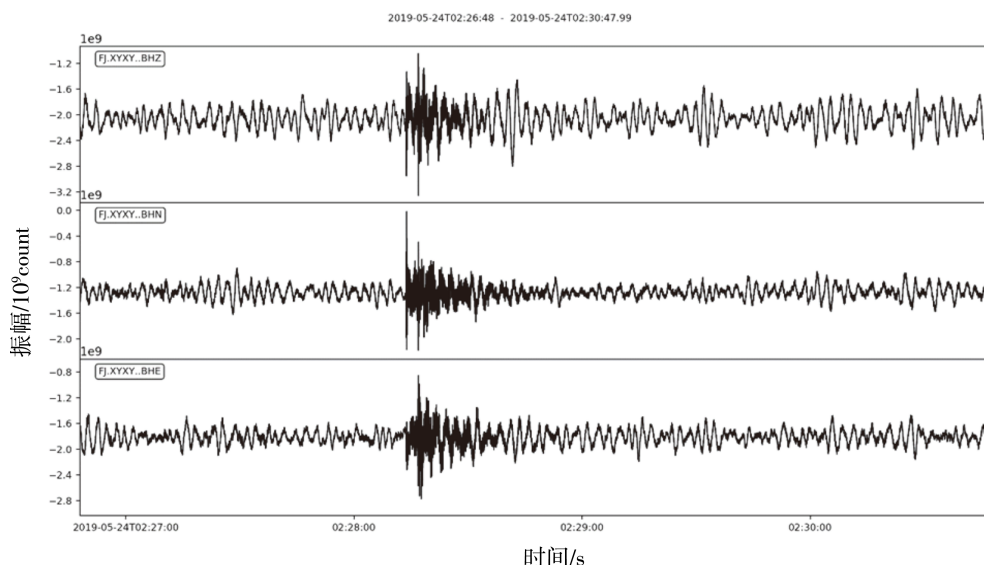


图3 福建仙游爆破错误识别事件波形(2019-05-24)

Fig.3 The waveform of the misidentification of the blasting event in Xianyou, Fujian

图4、图5为本文最优算法正确和错误识别事件震中分布图,正确识别的事件中共有1962个地震和1180个爆破事件;错误识别的事件中共有92个地震和83个爆破事件;爆破错识率高于地震错识率,分析错误识别的事件波形,发现个别事件波形背景噪声大和波形零漂的情况;错误识别的爆破事件分布较为分散,错误识别的地震事件则主要集中在台湾海峡区域,占错误识别的地震事件总数51%,本文研究训练集未将台湾海峡地震作为训练样本,说明学习样本覆盖,对识别效果存在影响;通过图4、图5比较,错误识别的爆破

和地震均有与正确识别的爆破和地震位置重叠,反映了算法模型具有一定程度的不一致性。

4 自动识别模块研制及应用

选择本文最优算法研发自动识别模块,自动识别模块的研发采用Python语言编程实现,采用obspy库进行地震数据处理和分析。模块的实现还需要应用到关键的第三方库有PyMysql、redis、stomp。自动识别模块有日常人工编目应用和与自动编目系统对接的在线应用。福建台网区域自动地震编目系统已投入测试运行,初步具备多台触

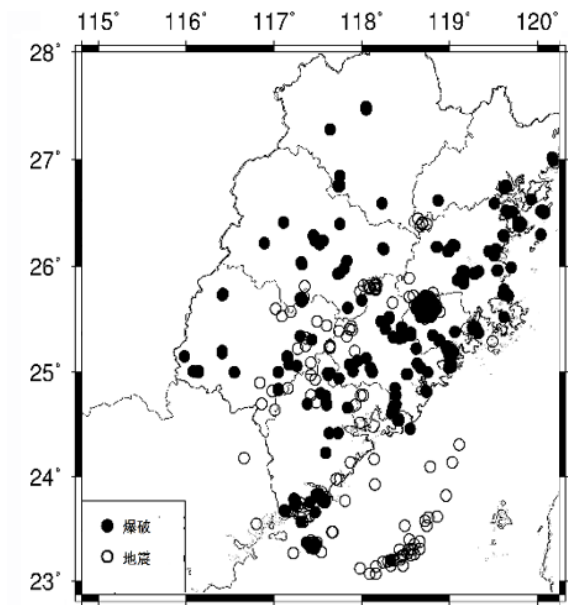


图 4 正确识别事件震中分布图

Fig.4 Epicentral distribution of correct identification events

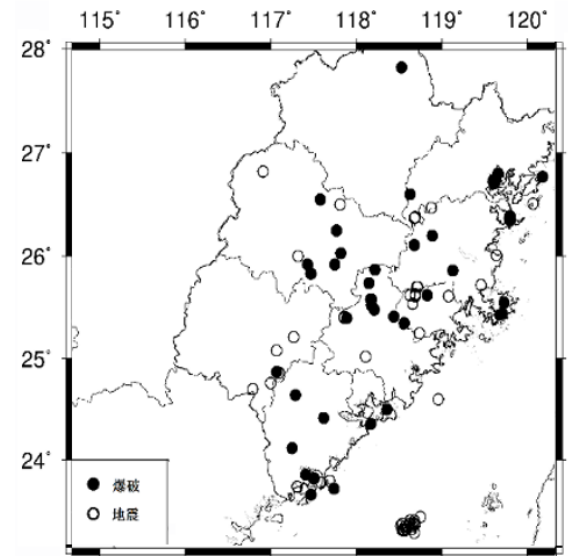


图 5 错误识别事件震中分布图

Fig.5 Epicentral distribution of misidentified events

发地震事件的自动编目能力，自动识别模块作为自动地震编目系统子功能模块，需研发与自动编目系统的接口，实现本台网事件类型自动识别日常化应用。

4.1 自动识别模块人工编目应用

4.1.1 实现方法

事件自动导入，将前一天 Jopens 数据库触发波形事件(00~24 h)自动导入指定文件夹；自动识别模块定时自启动，读入指定文件夹中事件波形，从震相数据表中获取台站震相 P 到时，截取有效波形，经特征值提取步骤，进行事件类型判断，并与数据库人工分析结果对比，产出类型不一致

结果报表文件。软件流程自动化完成，无需人为操作。图 6 为自动识别模块人工编目应用技术路线图。

4.1.2 人工编目应用

事件类型自动识别模块已应用于人工编目，软件自动对每天产出触发事件进行检测，产出报表文件供编目分析人员参考，极大减小人工识别事件类型的误判；软件运行以来，检测出多个人工编目误判事件及对可疑事件性质进行判识。2019 年 8 月 1 日至 2020 年 2 月 29 日，福建台网共实时触发 1531 个事件，事件震级范围 M_L 1.0~3.3，其中天然地震 911 个、爆破事件 620 个。从检测结果来看，事件类型识别准确率为 93.9%，实

际应用识别准确率接近测试集的测试准确率,共检测出 8 个事件类型人工识别错误,错误提交数据库;其中天然地震的识别准确率为 95.7%,错误识别的地震事件中有 20 个为台湾海峡地震,占地震错误总数 51.3%,有 3 个为人工难以判识事件,事件波形与爆破极为相似;2019 年 9 月 26 日

至 27 日共发生仙游小震群共 152 个地震事件,对该仙游小震群识别准确率达到 100%;爆破的识别准确率为 92.1%,错误识别的爆破事件中有 2 个信噪比低;地震识别效果优于爆破,与测试集的测试结果一致。表 4 为福建台网 2019 年 8 月至 2020 年 2 月实时触发事件识别情况统计表。

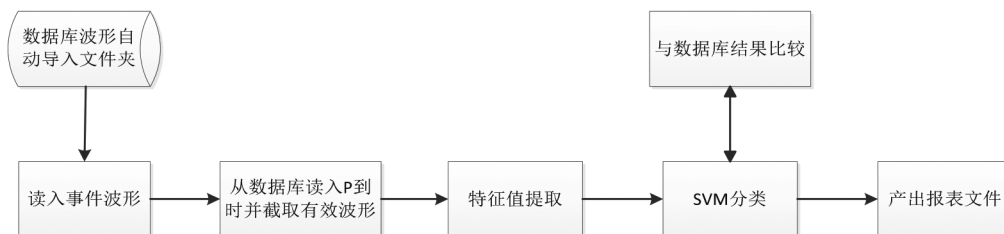


图 6 自动识别模块人工编目应用技术路线图

Fig.6 Technical roadmap for manual cataloging application of automatic identification module

表 4 2019-08-01—2020-02-29 福建台网实时触发事件识别情况统计表

Table 4 Statistical table of real-time triggered events recognition in Fujian Network from August 1st, 2019 to February 29th, 2020

事件类型	实时触发数	识别准确数	识别错	人工误识别	识别准确率/(%)
天然地震	911	872	39	5	95.7
爆破	620	571	49	3	92.1
合计	1531	1443	88	8	93.9

4.2 自动识别模块自动编目在线应用

4.2.1 实现方法

AMQ 消息中间件是 Apache 出品,最流行的,能力强劲的开源消息总线。AMQ 的消息传递模式有二种,一种为点对点模式,另一种为发布/订阅模式^[10];自动识别模块与自动编目接口开发应用第二种模式。redis 是完全开源免费的,遵守 BSD 协议,是一个高性能的 key-value 数据库^[15]。采用 AMQ 消息中间件为信息中介(需下载安装 AMQ 并启动 AMQ 服务),redis 波形共享内存中获取事件波形,研发了自动识别模块与自动编目系统接口,将自动识别模块嵌入自动编目系统。具体流程:从消息中间件接收事件参数信息,获取目标地震

记录的开始与结束信息,从 redis 波形共享内存中获取事件波形,经过数据预处理、特征值提取等步骤,进行事件类型判断;产出的事件类型信息,通过消息中间与其他软件模块共享。图 7 为自动识别模块与自动编目接口技术路线图。

4.2.2 在线应用

目前自动识别模块已具备在线运行的能力,初步实现事件类型的实时全自动判别。2019 年 11 月至 2020 年 3 月已在线试验运行 5 个月,其间检测出多个人工编目漏分析地震事件,提高目录完整性。对自动识别在线系统 7×24 运行稳定性,进行不间断测试,系统未出现死机、进程中断现象,测试运行性能稳定良好,功能和性能达到预期设计,验证了自动识别模块实用性和稳定性。

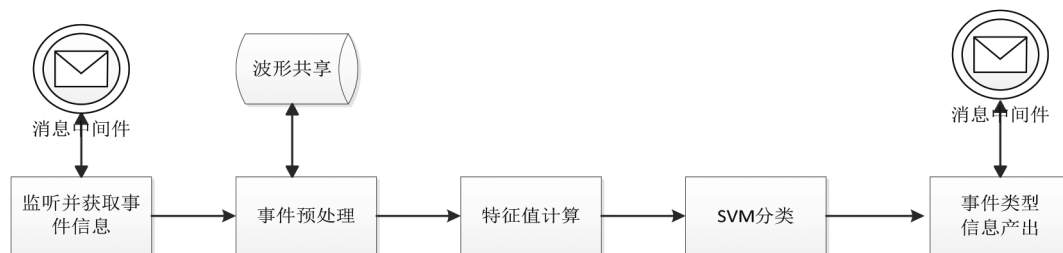


图 7 自动识别模块与自动编目接口技术路线图

Fig.7 Technical roadmap of interface between automatic identification module and automatic cataloging

5 讨论和结论

(1)本文通过对多种特征组合联合支持向量机进行大批量事件类型测试分析,结果显示特征指标可对事件进行有效分类;据此得出识别效果较好的事件类型识别算法,事件类型识别准确率达到94.5%,灵敏度和特异度比较均衡,灵敏度和特异度均达到93%以上,对多数地震和爆破都能有效识别,极低震级单台事件也能较好的分类,判识结果产出快,单个事件全流程识别仅需耗时2.5 s。

(2)算法数据输入采用单台Z分向和三分向数据作为输入;采用同样特征组合,三分向数据识别效果优于单分向;可能的解释是更多输入数据对于支持向量机的输入参数约束更好。此外,事件类型自动识别研究,训练集中含有识别错误的事件也可能影响支持向量机分类性能;通常情况下,区域台网地震目录很少是100%正确;在对训练集自验证,最初训练支持向量机已能准确识别出人工编目错误识别事件,其中地震训练集中有3个错误识别事件,人工爆破中有8个错误识别事件;在最终测试前,对此数据进行剔除。

(3)采用本文研究最优算法联合支持向量机研制事件类型自动识别模块,日常人工编目实际应用效果显示,识别准确率达到93.9%,可检测出人工识别错误事件,减少日常编目工作量和错误率。另外,实现与自动编目系统对接的在线自动识别系统接口,在线自动识别系统测试运行性能稳定良好,检测出多个人工漏分析地震事件,自动识别模块具有实用性和稳定性。

在后期工作中,增加其他区域事件进行完善和优化学习样本,并改进算法提高系统的泛化能力,自动识别模块将会更具有普适性,实现事件类型识别的多样性,塌陷、矿震、滑破等,可应用于其他区域事件类型自动判别。

参考文献

- [1] 边银菊. 遗传BP网络在地震和爆破识别中的应用[J]. 地震学报, 2002, 24(5): 516-524.
- [2] 卢世军, 黄汉明. 基于能量比的天然地震与人工爆破自动识别算法研究[J]. 华南地震, 2010, 30(1): 34-39.
- [3] 黄汉明, 边银菊, 卢世军, 等. 天然地震与人工爆破的波形小波特征研究[J]. 地震学报, 2010, 32(3): 270-276.
- [4] 赵明, 陈石, Dave Yuen. 基于深度学习卷积神经网络的地震波形自动分类与研究[J]. 地球物理学报, 2019, 62(1): 374-382.
- [5] 王军. 支持向量机在地震和爆破识别中的应用[J]. 地震地磁观测与研究, 2018, 39(3): 181-188.
- [6] 刘方斌, 曲军浩, 张志慧. 基于BP神经网络的天然与非天然地震识别[J]. 计算机应用与软件, 2020, 37(1): 109.
- [7] Jari Kortstrom. Automatic classification of seismic events within a regional seismograph network[J]. Computers Geosciences, 2016, 87: 22-30.
- [8] N. Rabin. Earthquake-explosion discrimination using diffusion maps[J]. Geophysical Journal International, 2016, (207): 1484-1492.
- [9] Omar M. Saad. Automatic discrimination of earthquakes and quarry blasts using wavelet filter bank and support vector machine[J]. Journal of Seismology, 2019(23): 357-371.
- [10] 郝红艳. 基于支持向量机的可靠性优化设计方法研究[D]. 武汉: 华中科技大学, 2013.
- [11] Cristianini N. 支持向量机导论[M]. 北京: 机械工业出版社, 2005.
- [12] Pai-Hsuen Chen, Chih-jen lin, Bernhard Scholkopf. A tutorial on v-Support Vector Machines [J]. Applied Stochastic Models in Business and Industry, 2005, 2(2): 111-136.
- [13] 倪炜. 分布式消息中间件实践[M]. 北京: 电子工业出版社, 2018.
- [14] Colin T O, G. Eli Baker, Anne F. Sheehan. Using P/S amplitude ratios for seismic discrimination at local distances[J]. Bulletin of the Seismological Society of America, 2016, 106(5): 1-12.
- [15] 付磊, 张益军. Redis 开发与运维[M]. 北京: 机械工业出版社, 2017.