

# 基于离散时间信号相关性的交通事件检测算法

孙 倩, 郭忠印

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

**摘要:** 提出了基于离散时间信号相关性的自动交通事件检测算法. 将交通信息数据转化为离散时间信号并进行相关性计算, 有效定位通过上、下游截面的同一组交通流. 解释了互相关系数的特征, 并采用仿真数据进行性能验证. 结果表明: 基于离散时间信号相关性的自动交通事件检测算法具有可视性且易于理解, 在低饱和交通环境下表现依然稳健, 具有很好的适应性.

**关键词:** 交通事件; 信号处理; 信号相关性

**中图分类号:** U491.3

**文献标志码:** A

## Automatic Incident Detection Algorithm Based on Discrete Time Signal Correlation

SUN Qian, GUO Zhongyin

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China)

**Abstract:** An automatic incident detection algorithm based on discrete time signal correlation was proposed. Traffic information data were converted to discrete-time signals and the correlation was calculated to locate the same traffic stream passing the upper and lower sections. The characteristics of correlation coefficients were explained. The results show that the algorithm is visual and easy to understand. The algorithm performs well under low-saturated traffic conditions and has better adaptability.

**Key words:** traffic incident; signal processing; signal correlation

交通事件包括的内容很多, 如交通事故、故障停车、道路施工和天气情况等, 它将直接导致某段道路通行能力减弱, 从而引起交通拥堵等现象. 据统计, 在美国交通事件引起的拥堵占城市快速路总拥堵事件的 50% ~ 75%<sup>[1]</sup>. FHWA (Federal Highway

Administration) 的报告指出, 由交通事故、故障停车等交通事件引起的公路延误占交通拥堵总延误的 1/3<sup>[2]</sup>.

目前已有许多自动交通事件检测 (AID) 算法. 随着智能交通系统的发展, 城市内间断交通流环境道路渐渐能够实现摄像头全覆盖, 因此本文主要关注高速公路和城市快速路等连续交通流环境道路的 AID 算法.

早期开发的 AID 算法主要有标准偏差算法<sup>[3]</sup>、加利福尼亚 (Ca) 算法<sup>[4]</sup> 和基于突变理论的 McMaster 算法<sup>[5]</sup>等. 随着交通流理论和机器学习的发展, 越来越多的算法涌现出来, 这些算法主要分为 2 个方向: 一个方向是基于机器学习, 通过输入不同参数训练分类模型以评价是否发生交通事件, 如神经网络模型<sup>[6-7]</sup>、二分类支持向量机模型<sup>[8-9]</sup>、多分类支持向量机模型<sup>[10]</sup>、决策树模型<sup>[11]</sup>、贝叶斯模型<sup>[12]</sup>和核极限学习机模型<sup>[13]</sup>等; 另一个方向是预测法, 即通过已有数据对未来状态进行预测并与实测数据进行比较, 以判断是否发生交通事件, 如基于突变理论的交通事故离群挖掘算法<sup>[14]</sup>、事件影响指数检测算法<sup>[15]</sup>和交通状态变异指数法<sup>[16]</sup>等. 此外, 不同的 AID 算法对不同条件的适应性不同, 因此一些算法的组合和集成类模型被提出, 如以因子分析与聚类分析为手段的基于多个 AID 算法的决策级融合方法<sup>[17]</sup>、基于时间序列方法和机器学习方法的组合 AID 算法<sup>[18]</sup>和多个朴素贝叶斯分类器的 AID 算法<sup>[19]</sup>等. 现有算法普遍存在计算复杂、泛化能力不强、容易出现过拟合等问题, 其表现方式没有传统算法直观, 并且需要具体路段大量事件与非事件历史数据的支持, 影响了模型的推广及应用.

信号处理技术中互相关常常应用于时延估计<sup>[20]</sup>和定位<sup>[21]</sup>, 其原理为利用声波在空气中直线传播和遇到障碍物反射抵达接收端的时间差进行时延

收稿日期: 2017-12-02

基金项目: 国家自然科学基金(71673201)

第一作者: 孙 倩(1990—), 女, 博士生, 主要研究方向为道路安全与环境. E-mail: sun\_sunqian@163.com

通讯作者: 郭忠印(1962—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为道路安全与环境. E-mail: zhongyin@tongji.edu.cn

估计和障碍物定位. 通过计算 2 次声音抵达接收端获得的信号互相关系数能够分析 2 组信号不同相位条件下的相似程度, 相关性最高点对应的相位即为 2 次声音抵达接收端的时差. 交通流有与声波类似的性质. 本文以此为基础, 论证了将上、下游采集到的交通信息等价于 2 组离散时间信号的合理性. 通过计算互相关系数分析交通流变化, 并据此提出基于离散数字信号相关性的 AID 算法. 通过仿真模拟不同交通流条件下的事件场景, 与加利福尼亚算法进行对比评价.

## 1 离散时间信号相关性算法

### 1.1 离散时间信号相关性理论

离散信号是离散自变量的函数, 当自变量为时间  $n$  时, 叫做离散时间信号, 它表现为在时间上按一定次序排列的不连续的一组数的集合, 故也称为时域离散信号, 用  $x(n)$  表示.

相关是研究 2 个信号之间, 或一个信号和其位移后的相关性, 前者叫做信号互相关, 后者叫做信号自相关, 本文主要讨论的是信号互相关. 离散时间信号  $x(n)$  和  $y(n)$  的互相关函数

$$R_{xy}(\tau) = \sum x(n)y(n+\tau) \quad (1)$$

根据广义瑞利公式和互能谱定义推导可知, 互相关函数  $R_{xy}(\tau)$  和互能谱  $W_{xy}(k)$  是一对傅里叶变换. 因此, 在计算  $R_{xy}(\tau)$  时, 通过对离散时间信号  $x(n)$  和  $y(n)$  进行离散傅里叶变换得到频谱密度  $X(k)$  和  $Y(k)$ , 计算公式如下所示:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}nk}$$

$$Y(k) = \sum_{n=0}^{N-1} y(n)e^{-j\frac{2\pi}{N}nk}$$

式中:  $N$  为信号长度.

对互能谱进行傅里叶变换得到互相关函数  $R_{xy}(\tau)$ , 如下所示:

$$W_{xy}(k) = X(k)Y^*(k)$$

$$R_{xy}(\tau) = f(W_{xy}(k))$$

式中:  $*$  表示共轭转置运算;  $f$  表示傅里叶变换运算. 最终得到互相关系数

$$\rho_{xy}(\tau) = \frac{R_{xy}(\tau) - \mu_x\mu_y}{\sigma_x\sigma_y}$$

$$\mu_x = \sqrt{\lim_{\tau \rightarrow \infty} R_{xx}(\tau)}, \mu_y = \sqrt{\lim_{\tau \rightarrow \infty} R_{yy}(\tau)}$$

$$\sigma_x = \sqrt{R_{xx}(0)}, \sigma_y = \sqrt{R_{yy}(0)}$$

互相关系数  $\rho_{xy}(\tau)$  为一个  $0 \sim 1$  的序列, 它表示

2 个离散时间信号相位差为  $\tau$  时的相关性, 越接近 1 表示越相关.

### 1.2 离散时间信号相关性应用

将上、下游采集到的交通信息数据转化为离散时间信号, 交通信息数据为通过采集区的车辆速度. 转化为离散时间信号的方法是: 选取固定周期, 当周期内有一个数据, 该数据的值即为离散时间信号该时间位置上的幅值; 当周期内有 2 个或以上数据, 该数据的平均值即为离散时间信号该时间位置上的幅值; 当周期内没有数据, 幅值为零. 现选取 G109 某两相邻截面的调研数据, 该数据采集时间内未发生交通事件, 两截面距离约 800 m. 由于交通量较少, 周期取 70 s. 离散时间信号分布如图 1 所示. 横坐标时间序列代表第  $n$  个周期, 纵坐标代表第  $n$  个周期内的速度平均值, 可见两截面的离散时间信号分布图存在一定的相似性, 但截面 2 的幅值普遍比截面 1 低.

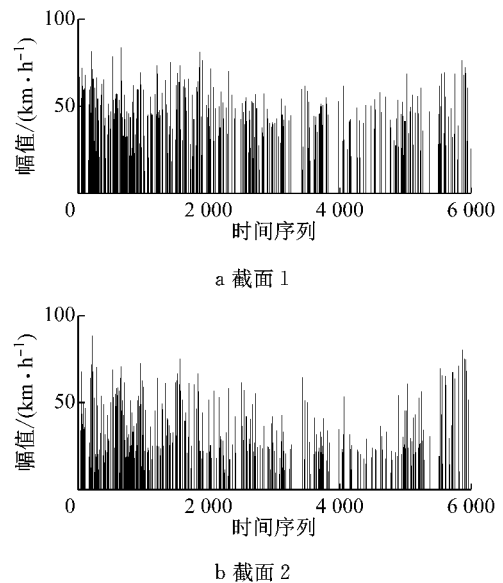


图 1 离散时间信号分布

Fig. 1 Distribution of discrete time signal

根据第 1.1 节的离散时间信号相关性理论, 分别计算 1 组离散时间信号的自相关系数和 2 组离散时间信号的互相关系数, 如图 2 所示. 横坐标为 2 段离散事件信号的相位差, 纵坐标为该相位差下信号的相关系数. 从图 2 可以看到非常明显的峰值. 图 2a 中, 离散时间序列相位差  $\tau=0$  时自相关系数为 1, 代表信号完全相同; 图 2b 中, 离散时间信号相位差  $\tau=1$  时 2 组信号存在最大相似性, 其所对应的数据为同一组交通流, 根据截面距离和交通流速可知结果可信.

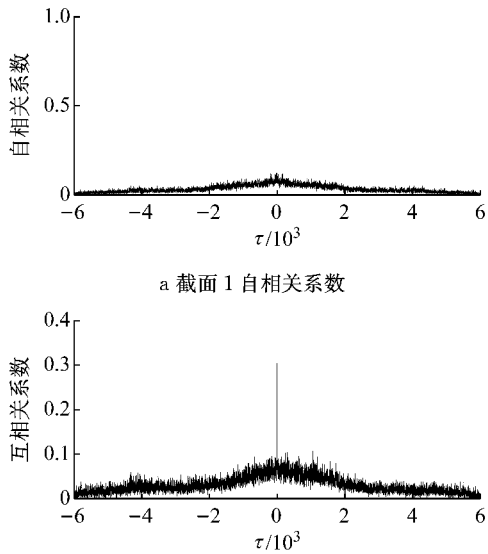


图 2 相关系数分布

Fig.2 Distribution of correlation coefficient

由此可见,将交通信息数据转化为离散时间信号并进行相关性计算,可以有效定位同一组交通流并计算其相似性.

## 2 交通离散时间信号相关性特征分析

### 2.1 互相关系数特征分析

取图 2b 中的峰值为互相关系数  $F$ ,根据信号处理理论,互相关系数  $F$  越接近 1 表示上、下游交通流越相似,即交通流在该路段越稳定,但是该系数本身没有实际物理意义.随着时间的发展,交通流相似性会越来越低,互相关系数  $F$  也会越来越小,最终成为背景噪声,因此本节研究互相关系数  $F$  的主要影响因素.

将不同的上、下游截面距离  $L$  和离散时间信号重叠长度  $N_0$  绘制成多条信号长度和互相关系数的关系线,如图 3 所示.可以发现,当  $L$  和  $N_0$  一定时,互相关系数  $F$  与离散时间信号长度  $N$  的倒数成线性相关,且通过横轴同一点,可用下式表示:

$$F = k(1/N - a)$$

式中: $k$  为斜率,其大小与  $L$  和  $N_0$  有关; $a$  为待定参数.

当互相关系数  $F$  过小时会成为背景噪声,因此互相关系数  $F$  的大值端更受关注,即斜率  $k$  与  $L$  和  $N_0$  的关系更受关注.取不同的截面距离和重叠信号长度绘制斜率的等值线图(见图 4),发现截面距离和重叠信号长度可以互补,以使斜率达到足够大,但并

不是没有极限.重叠信号长度不可能大于信号总长度,即当上、下游截面距离固定,重叠信号长度等于信号总长度时,互相关系数  $F$  达到最大.根据这个性质,以保证检测效果为目的,可以确定上、下游截面容许距离.

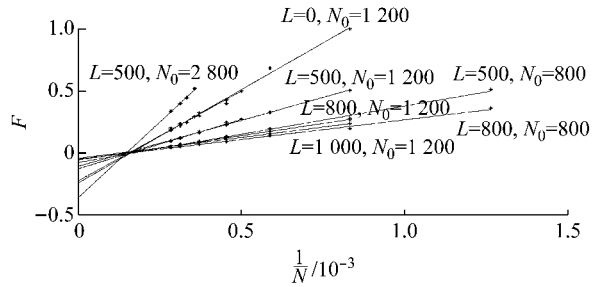


图 3 信号长度与互相关系数关系

Fig.3 Relationship between signal length and correlation coefficient

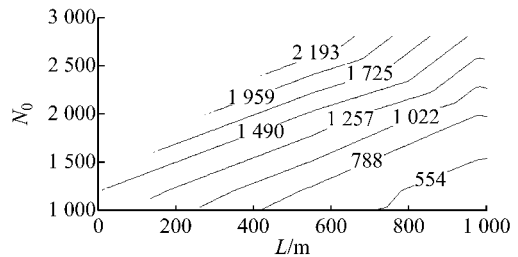


图 4 斜率等值线

Fig.4 Contour map of slope

### 2.2 事件与非事件条件下相关性特征分析

由于事件条件难以在实际交通环境下采集到,并且难以在完全相同交通环境下采集不同事件的交通状态,因此为便于比较,采用了仿真手段获取交通数据.利用 Vissim 进行仿真,时长 6 h,设定交通事件发生在第 4 h 内.交通事件将导致部分路段通行能力下降.事件持续时间分别为 1 h、30 min、10 min 和 0 min(非事件条件),采集点上、下游距离 1 km,根据第 2.1 节的分析结果可以通过控制  $N_0$  避免互相关系数  $F$  过小,保证检测效果.

以事件持续时间 30 min 为例,绘制整小时内上、下游交通离散时间信号的分布(见图 5).从图 5 可以看到,在 4 h 时间段上信号互相关系数峰值明显小于其他时间段,即该时间范围内交通流出现比较大的扰动,发生了交通事件.

从图 5 可知,1 h 检测显然过于粗糙,难以及时反映数据变化,因此采用持续更新的方式监测离散时间信号最大互相关系数.当交通流出现扰动时,互相关系数  $F$  所在的时间序列相位差可能出现扰动,

因此同时监测最大互相关系数所在的时间序列相位差变化,具体流程如图 6 所示. 根据流程绘制事件持续时间分别为 1 h、30 min、10 min 和 0 min 的相关性发展图和扰动发展图(见图 7、8). 相关性发展图代表离散时间信号互相关系数随时间的波动情况,扰动发展图代表互相关系数最高点对应的的时间序列相位差随时间的波动情况. 从图 7、8 可以非常明显地看到,发生交通事件后相关性迅速下降,时间序列相位差也发生扰动,并且事件持续时间越久,相关性发展得越低. 因此,当前期平稳状态下相关性开始持续走低或时间序列相位差产生明显变化时,判断交通事件发生.

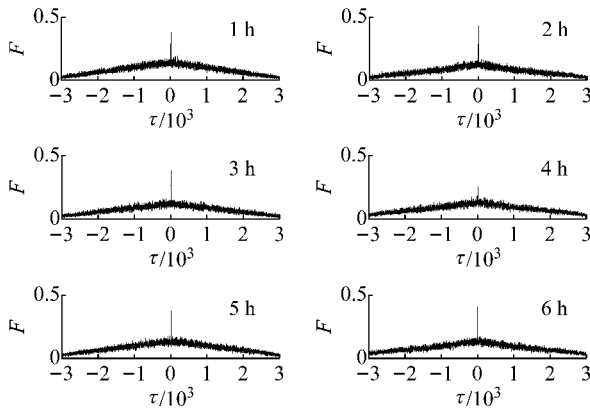


图 5 6 h 相关系数分布

Fig.5 Distribution of correlation coefficient in 6 h

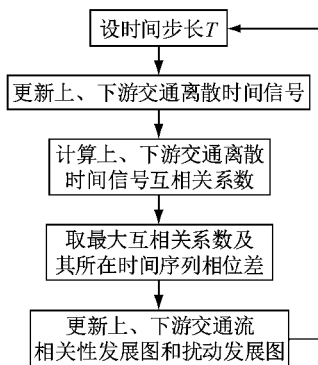


图 6 自动交通事件检测流程

Fig.6 Flow chart of AID

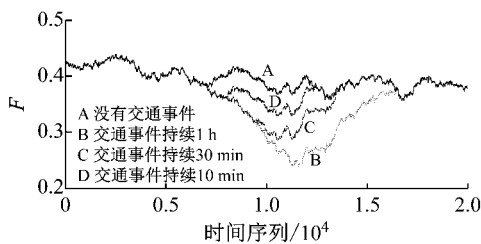


图 7 相关性发展图

Fig.7 Development of correlation

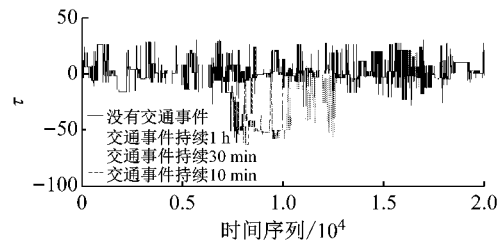


图 8 扰动发展图

Fig.8 Development of disturbance

### 3 交通事件检测试验

#### 3.1 检测试验

AID 算法性能的评价指标包括交通事件检测率 (DR)、误报率 (FAR) 和平均检测时间 (MTTR). 检测率和误报率用于评价 AID 算法的检测性能,平均检测时间能够评价算法的检测效率. 3 个指标存在互相制约的关系.

首先利用 Vissim 分别模拟单向两车道低饱和交通 ( $500 \text{ veh} \cdot \text{h}^{-1}$ )、中饱和交通 ( $2\ 500 \text{ veh} \cdot \text{h}^{-1}$ ) 和高饱和交通 ( $3\ 500 \text{ veh} \cdot \text{h}^{-1}$ ) 环境下非交通事件 50 次,以此作为基准,确定判别事件发生的互相关系数及相位差阈值. 阈值决定了检测率和误报率,交通事件检测率越高,误报率也越高(见图 9). 高饱和交通环境下交通事件容易造成堵车或连环交通事故,因此取较高的检测率牺牲了较低的误报率;反之,低饱和交通环境下取相对较低的检测率获得较低的误报率,节约人力. 经过对比,3 种环境下的互相关系数和相位差阈值分别确定为 (0.36, 0)、(0.57, -10) 和 (0.68, -22).

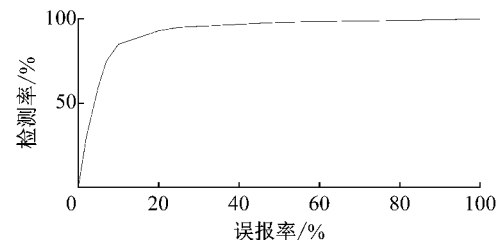


图 9 误报率和检测率关系

Fig.9 Relationship between FAR and DR

分别模拟 3 种交通环境下交通事件条件和非交通事件条件各 100 次,分别采用加利福尼亚算法和基于离散时间信号相关性检测算法进行评价,评价结果如表 1 所示.

#### 3.2 结果分析

在低饱和交通环境下,本文算法检测率达

97.1%，而加利福尼亚算法仅 84.1%；在中饱和交通环境下，本文算法检测率略微下降(96.5%)，而加利福尼亚算法升至 92.7%；在高饱和交通环境下，2 种算法检测率基本持平，在 96.0%左右。综合来看，本文算法略高于加利福尼亚算法，由此可以判断，本文算法比加利福尼亚算法具有更加良好的环境适应性。在平均检测时间方面，随着交通饱和度的上升，采集数据量的增加能够缩短时间步长，从而相关性发展图和扰动发展图更新速率提高，本文算法的检测效率也持续提升。

表 1 加利福尼亚算法和本文算法的检测效果  
Tab.1 Performance of Ca algorithm and the proposed algorithm

交通环境	检测算法	检测率/%	误报率/%	平均检测时间/min
低饱和	加利福尼亚算法	84.1	0.95	0.98
	本文算法	97.1	0.51	1.19
中饱和	加利福尼亚算法	92.7	1.23	1.02
	本文算法	96.5	0.67	1.05
高饱和	加利福尼亚算法	95.7	1.30	0.95
	本文算法	96.2	1.08	0.94

从表 1 可以看到，本文提出的算法在低饱和和交通环境下性能远好于加利福尼亚算法。随机抽取一个低饱和交通环境下发生交通事件的样本，为方便比较，将数据归一化处理，如图 10 所示。由图 10 可知：由于加利福尼亚算法过度依赖上、下游交通占有率，当交通饱和度很低时发生交通事件，即便降低了路段的通行能力，但较小的交通量并不足以造成占有率发生非常明显的变化；本文算法主要依赖于上、下游交通流分布形式，发生交通事件会导致换车道等行为比例升高，相当于使交通流重新分布，从而导致相关性急剧下降，交通流的相位差发生波动。

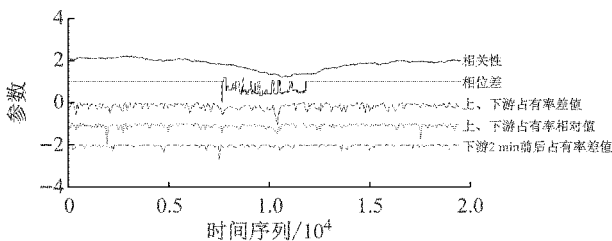


图 10 低饱和交通环境下参数变化

Fig.10 Variation of parameters under low saturated traffic environment

以同样方法绘制交通接近饱和时参数对比图，如图 11 所示。从图 11 可以看到，加利福尼亚算法在交通量较大的情况下明显比小交通量环境下表现好，说明交通事件引起的通行能力降低影响到了上、

下游占有率的分布。车辆受到空间限制，即使重分布，高饱和交通环境交通流分布形式也较低饱和和交通环境时选择性小，因此相关性发展图没有低饱和和交通环境时明显，但在时间序列上仍然有非常明显的差异，相位差的波动性则表现更加稳健。

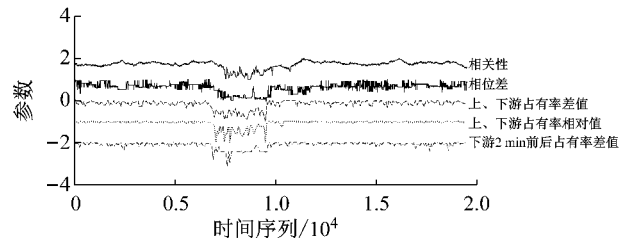


图 11 高饱和交通环境下参数变化

Fig.11 Variation of parameters under high saturated traffic environment

## 4 结语

根据信号处理理论，提出了基于离散时间信号相关性的 AID 算法。将交通信息数据转化为离散时间信号并进行相关性计算，可以有效定位同一组交通流并计算其相似性。通过对照分析解释了互相关系数的特征，并采用仿真数据进行性能验证。结果表明，采用基于离散时间信号相关性的 AID 算法在低饱和和交通环境下表现依然稳健，具有可视性且易于理解。在未来研究中将采集实际高速公路的交通事件条件数据对算法性能进行进一步验证。

## 参考文献:

- [1] LINDLEY J A. Urban freeway congestion: quantification of the problem and effectiveness of potential solutions [J]. ITE Journal, 1987, 57(1):27.
- [2] KEETER B C, SINGER N. Nation’s top highway official helps Launch New North Carolina Road and traffic information service[R]. Washington DC: Federal Highway Administration, 2004.
- [3] DUDEK C L, MESSER C J, NUCKLES N B. Incident detection on urban freeways [C]//Transportation Research Board. Washington DC: Transportation Research Board, 1974: 12-24.
- [4] PAYNE H J, HELFENBEIN E D, KNOBEL H C. Development and testing of incident detection algorithms[R]. Washington DC: Federal Highway Administration, 1976.
- [5] PERSAUD B N, HALL F L. Catastrophe theory and patterns in 30-second freeway traffic data: implication for incident detection[J]. Transportation Research: Part A, 1990, 23(2): 103.
- [6] SRINIV D, SHARMA V, TOH K A. Reduced multivariate polynomial-based neural network for automated traffic incident

- detection[J]. *Neural Networks*, 2008, 21(2/3): 484.
- [7] 刘庆华,丁文涛,涂娟娟,等. 优化 BP\_Adaboost 算法及其交通事件检测[J]. *同济大学学报(自然科学版)*, 2015, 43(12):1829.  
LIU Qinghua, DING Wentao, TU Juanjuan, *et al.* Improved BP\_Adaboost algorithm and its application in traffic incident detection[J]. *Journal of Tongji University (Natural Science)*, 2015, 43(12):1829.
- [8] 王武功,马荣国. 交通事件检测的加权支持向量机算法[J]. *长安大学学报(自然科学版)*, 2013, 33(6): 84.  
WANG Wugong, MA Rongguo. Weighed support vector machine for traffic incident detection[J]. *Journal of Chang'an University (Natural Science Edition)*, 2013, 33(6): 84.
- [9] 陈维荣,关佩,邹月娴. 基于 SVM 的交通事件检测技术[J]. *西南交通大学学报*, 2011, 46(1): 63.  
CHEN Weirong, GUAN Pei, ZOU Yuexian. Automatic incident detection technology based on SVM[J]. *Journal of Southwest Jiaotong University*, 2011, 46(1): 63.
- [10] 刘志强,吕学,张利. 基于多分类 GA-SVM 的高速公路 AID 模型[J]. *系统工程理论与实践*, 2013, 33(8):2110.  
LIU Zhiqiang, LÜ Xue, ZHANG Li. Highway automatic incident detection based on multi-class classification and GA-SVM[J]. *Systems Engineering: Theory & Practice*, 2013, 33(8):2110.
- [11] CHEN S, WANG W. Decision tree learning for freeway automatic incident detection [J]. *Expert Systems with Application*, 2009, 36(2):4101.
- [12] 张轮,杨文臣,刘拓. 基于朴素贝叶斯分类的高速公路交通事件检测[J]. *同济大学学报(自然科学版)*, 2014, 42(4):558.  
ZHANG Lun, YANG Wenchen, LIU Tuo. A Naive Bayesian classifier-based algorithm for freeway traffic incident detection [J]. *Journal of Tongji University (Natural Science)*, 2014, 42(4): 558.
- [13] 商强,林赐云,杨兆升,等. 基于变量选择和核极限学习机的交通事件检测[J]. *浙江大学学报(工学版)*, 2017, 51(7):1339.  
SHANG Qiang, LIN Ciyun, YANG Zhaosheng, *et al.* Traffic incident detection based on variable selection and kernel extreme learning machine[J]. *Journal of Zhejiang University (Engineering Science)*, 2017, 51(7):1339.
- [14] 龙琼,胡列格,张谨帆,等. 基于尖点突变理论模型的交通事故检测[J]. *土木工程学报*, 2015, 48(9):112.  
LONG Qiong, HU Liege, ZHANG Jinfan, *et al.* Traffic incident detection based on the cusp catastrophe theory model [J]. *China Civil Engineering Journal*, 2015, 48(9):112.
- [15] 李红伟,陆键,姜桂艳,等. 快速路交通事件检测方法[J]. *东南大学学报(自然科学版)*, 2013, 43(3):649.  
LI Hongwei, LU Jian, JIANG Guiyan, *et al.* Traffic incident detection algorithm for expressway[J]. *Journal of Southeast University(Natural Science Edition)*, 2013, 43(3):649.
- [16] 牛世峰,姜桂艳,李红伟,等. 基于纵向时间序列的快速路交通事件检测算法[J]. *哈尔滨工业大学学报*, 2011, 43(2):144.  
NIU Shifeng, JIANG Guiyan, LI Hongwei, *et al.* Automated detection algorithm for traffic incident in urban expressway based on lengthways time series of traffic parameters [J]. *Journal of Harbin Institute of Technology*, 2011, 43(2):144.
- [17] 李琦,姜桂艳,杨聚芬. 基于因子分析与聚类分析的交通事件自动检测算法融合[J]. *吉林大学学报(工学版)*, 2012, 42(5):1191.  
LI Qi, JIANG Guiyan, YANG Jufen. Automatic incident detection algorithms fusion method based on factor analysis and cluster analysis [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2012, 42(5):1191.
- [18] WANG J, LI X, LIAO S, *et al.* A hybrid approach for automatic incident detection[J]. *IEEE Transaction Intelligent Transportation Systems*, 2013, 14(3):1176.
- [19] LIU Q, LU J, CHEN S, *et al.* Multiple Naive Bayes classifiers ensemble for traffic incident detection [J]. *Mathematical Problems in Engineering*, 2014, 2014(4):1.
- [20] 窦慧晶,王千龙,张雪. 基于二次相关的广义互相关时差估计算法[J]. *北京工业大学学报*, 2016, 42(2):197.  
DOU Huijing, WANG Qianlong, ZHANG Xue. Generalized cross-correlation algorithm based on secondary relation [J]. *Journal of Beijing University of Technology*, 2016, 42(2):197.
- [21] 张一闻,刘建平,张世全. 利用多点互相关值均值的实时声源定位算法[J]. *西安电子科技大学学报*, 2015, 42(1):168.  
ZHANG Yiwen, LIU Jianping, ZHANG Shiquan. Real-time acoustic source localization algorithm based on the average of multiple cross correlation values [J]. *Journal of Xidian University*, 2015, 42(1):168.