2016 American Control Conference (ACC)
Boston Marriott Copley Place
July 6-8, 2016. Boston, MA, USA

Detection of Abrupt Changes in Autonomous System Fault Analysis Using Spatial Adaptive Estimation of Nonparametric Regression

Alexander Kalmuk, Oleg Granichin, Olga Granichina and Mingyue Ding

Abstract— **The paper deals with the detection of abrupt changes in autonomous systems. We consider this problem in the presence of Gaussian noise and solve it in two steps. At first, spatial adaptive estimation of nonparametric regression is used to estimate the observable data. Then Filtered Derivative Algorithm is used to detect abrupt changes in the obtained data using an adaptive threshold. The estimation of this adaptive threshold is presented. This approach is then applied to demonstrate the slowdown detection of a small autonomous vehicle.**

I. Introduction

The problem of fault detection is to determine changes in the characteristics of a dynamical system. It is assumed that the fault takes place inside the system. The problem of change point detection is similar to the fault detection, but in this case both faults inside and outside of the system are of interest.

The problem of abrupt changes detection is an important one, since any abrupt changes could potentially indicate that something goes wrong. This is highly important for autonomous systems that operate without human supervision. Numerous approaches exist that can be used to detect the abrupt change [1]–[3]. Some of them imply the dependence of the probability density function upon a scalar parameter, while others imply multidimensional probability density function. In this paper we consider only the first case.

We also assume that measurable data contains some Gaussian noise. Usually it is a very natural assumption – for example, for GPS/IMU systems. Furthermore, we assume complete absence of the information about the system except for the observable part. Under such conditions, the actual readings of measurable variables appear to be the most valuable information. Therefore, the first step is to filter noise from the observable data. For this purpose we use the spatial adaptive estimation of nonparametric regression [4] [5]. The main idea behind this approach is to estimate the input signal using Least Square Estimate within a sliding window of adaptive width. Once we obtain the value of the estimated signal, we can detect the abrupt changes. There are numerous ways of performing such estimation (see [2]). We employ the Filtered Derivative Algorithm, that is widely used in image processing for the discontinuity detection. This method is based on the likelihood ratio test. The principle of this approach lies in the calculation of the likelihood ratio gradient change rate. The occurrence of an abrupt change corresponds to the state when the value of the gradient exceeds a given threshold value.

Since the initial observable data contains noise, we cannot set the threshold to be an exact constant value. There is no guarantee that the selected threshold would be valid for the varying values of the input signal and the adaptive threshold should be considered. It can be used to reduce the delays associated with the constant threshold method [6] and for more accurate estimation in the presence of noise [7].

The main proposition of this paper is to combine the Filtered Derivative Algorithm with an idea of adaptive threshold to achieve more accurate and efficient method for the abrupt change detection. For example, if we set up a constant threshold that is needed to be satisfied in ideal conditions, then we can calculate an adaptive threshold for filtered data. So the threshold is adaptive in terms that it depends on the inputs and the initial constant threshold. We show in the paper how it can be calculated.

In order to demonstrate the efficiency of this approach in a real-world environment we performed a simulation considering a small autonomous vehicle model, that detects its own slowdown. The vehicle data system was based on STM32F3-Discovery circuit containing the accelerometer and gyroscope. The situation of steady motion along the road with a smooth surface was simulated. Upon tripping over a piece of the foam rubber a slowdown occurred, which was detected by the system and followed by issuing a signal for the engine shutdown.

The article is organized as follows. Section II presents the problem statement. In sections III and IV an overview of used methods is provided. The main result is presented in Section V. The results of numerical experiments are given in Section VI, followed by Conclusions and Future work discussion.

II. Problem statement

Consider now the dynamical system with the model of observations:

$$\begin{cases} X_{t+1} = F_t X_t + W_t, \\ Y_t = \Phi_t X_t + V_t \end{cases} \tag{1}$$

A. Kalmuk, O. Granichin and O. Granichina and are with the Saint Petersburg State University (Faculty of Mathematics and Mechanics, and Research Laboratory for Analysis and Modeling of Social Processes), 7-9, Universitetskaya Nab., St. Petersburg, 199034, Russia, O. Granichin is also with the Institute of Problems in Mechanical Engineering, Russian Academy of Sciences. O. Granichina is also with Herzen State Pedagogical University. M. Ding is with the Huazhong University of Science and Technology, Wuhan, Hubei 430074, China. alexkalmuk@gmail.com, o.granichin@spbu.ru, olga_granichina@mail.ru, myding@hust.edu.cn

This work was supported by Russian Science Foundation (project 16-19-00057).

where $X_t \in \mathbb{R}^n$ is known state vector, $Y_t \in \mathbb{R}^m$ is known measurement vector, that consists of the observable system parameters, $F_t \in \mathbb{R}^{n\times n}$, $\Phi_t \in \mathbb{R}^{n\times m}$ are specified matrices, $W_t \in \mathbb{R}^n$, $V_t \in \mathbb{R}^m$ are unknown vectors of an input disturbance.

Our goal is to determine the fault in the observable variables Y_t. Each of the variables $Y^i = (y_1, y_2, ..., y_m)$ has the probability density function p^i_θ depending only on single scalar parameter $\theta \in \mathbb{R}$. Until the unknown moment of time t_0, the value of θ parameter is equal to θ_0. The problem is to detect a change of the parameter θ.

The solution of this problem includes two steps:

1) Filter a noise out of Y_t.
2) Apply the Filtered Derivative Algorithm to the obtained data.

III. Spatial Adaptive Estimation of Nonparametric Regression

Let us consider the following problem. We want to restore a signal function from noisy observations. Suppose there are noisy observations $y(x)$ of a signal function $f(x)\colon [0,1] \to \mathbb{R}$ – along the regular grid $\Gamma_n = \{i/n, i = 0, \ldots, n\}$:

$$y(x) = f(x) + \xi(x), \tag{2}$$

where $\{\xi(x)\}_{x\in\Gamma_n}$ is a sequence of independent $\mathcal{N}(0,1)$ random variables defined on the underlying probability space (Ω, A, P).

In [4] the estimates by the least square method of $f(x)$ at a given point $u \in [0,1]$ are considered. Suppose that the degree of estimate is 1. So we get the following approximation at a given point x_0:

$$\hat{f}_\Delta(x_0) = \frac{1}{N_\Delta} \sum_{x\in M_\Delta} y(x), \tag{3}$$

where $\Delta \in [0,1]$ is some segment $[x_0 - \delta, x_0 + \delta]$ centered at x_0 and containing at least one observation point, M_Δ is the set of observation points in Δ, N_Δ is the cardinality of M_Δ.

The problem is to select the optimal window in absence of a priori information on $f(x)$. Let us introduce the following estimation:

$$|\hat{f}(x_0) - f(x_0)| \le \omega_f(x_0, \delta) + N_\Delta^{-1/2}|\zeta(\Delta)|, \tag{4}$$

where $\zeta(\Delta) = \frac{1}{N_\Delta}\sum_{x\in M_\Delta} \xi(x)$, $\omega_f(x,\delta) = \sup\limits_{x\in\Delta} |f(x) - f(x_0)|$. The right part of (4) is comprised of two terms – deterministic $\omega_f(x,\delta)$ and stochastic error $N_\Delta^{-1/2}|\zeta(\Delta)|$. Since $\zeta(\Delta)$ is $\mathcal{N}(0,1)$, the stochastic error typically is of order of $(n\delta)^{-1/2}$:

$$P\{N_\Delta^{-1/2}|\zeta(\Delta)| > \kappa(n\delta)^{-1/2}\} \le exp\{-c\kappa^2\}, \tag{5}$$

with certain absolute constant $c > 0$. Now, there are no more than n essentially different (resulting in different sets M_Δ) choices of Δ. Let these choices be

$$\Delta_1 \subset \Delta_2 \subset ... \subset \Delta_N,$$

and let $2\delta_1, 2\delta_2, \ldots, 2\delta_N$ be the length of the windows $\Delta_1, \Delta_2, \ldots, \Delta_N$. We obtain

$$\Omega_\kappa = \{\omega \in \Omega \mid N_{\Delta_i}^{-1/2}|\zeta(\Delta_i)| \le \kappa(n\delta_i)^{-1/2}, \quad where\ i = 1, \ldots, N\}. \tag{6}$$

Assuming that $\omega \in \Omega_k$, (4) can be strengthen as

$$|\hat{f}_{\Delta_i}(x_0) - f(x_0)| \le \omega_f(x_0, \delta_i) + \kappa(n\delta_i)^{-1/2}, \tag{7}$$

Notice that as i grows, the first term in the right hand side increases, and the second term decreases. Therefore, a reasonable choice of the window to be used is the one balancing the both terms, namely, the one related to the largest i with $\omega_f(x_0, \delta_i) \le \kappa(n\delta_i)^{-1/2}$.

IV. Filtered Derivative Algorithm

This method is based on the likelihood ratio test (see [8]). In a completely noiseless situation a change in the mean level of a sequence of observations is locally characterized by a great absolute value of the (discrete) derivative of the sample observations. Since the derivative operator starts to act in a very poor manner as soon as the noise appears, a more realistic detector should use a filtering operation before calculating the derivative. An simple way to increase the robustness of this detector is to count the number of threshold crossings within a fixed time interval before deciding whether the change actually occurred.

Using the derivation of the finite moving average control charts [9] we get

$$g_k = \sum_{i=0}^{N-1} \gamma_i \ln \frac{p_{\theta_1}(y_{k-i})}{p_{\theta_0}(y_{k-i})}, \tag{8}$$

where the weights γ_i are again any weights for causal filters. Suppose that we have already filtered noise out of the input data. Therefore, let $\gamma_i = 1$ for all $i = 0, \ldots, N-1$. We consider the discrete derivative of g_k:

$$\nabla g_k = g_k - g_{k-1}, \tag{9}$$

and the following stopping rule:

$$t_a = \min\{k : \sum_{i=0}^{N-1} I_{\{|\nabla g_{k-i}| \ge h\}} \ge \eta\}, \tag{10}$$

where I is an indicator function, h is a threshold for the derivative, and η is a threshold for the number of crossings of h.

V. The Estimate of a Signal

In this section we introduce an adaptive threshold and show its dependence on a constant threshold in the presence of standard Gaussian noise. In order to do this we apply (7) for the estimation f of the input signal (3). Then we employ this estimation within the Filtered Derivative Algorithm.

Theorem 1 in [4] describes a way to estimate $|\hat{f}(x_0) - f(x_0)|$. Now we will reformulate this theorem for the case of the degree of estimation being equal to 1 (i.e. (3)).

At first, let us define the ideal window $\Delta_k^*(x_0) = [x_0 - \delta_k^*, x_0 + \delta_k^*]$ for the new conditions - $p = l = 1$, $\theta_m = 1$:

$$\delta_\kappa^* = \max\Big\{\delta \le \delta_0 : \frac{\kappa}{\sqrt{2n\delta}} \ge 4\int_{x_0-\delta}^{x_0+\delta} |f(x)|dx\Big\}. \quad (11)$$

One can check that coefficients $\alpha_\Delta(x,u)$ (see 3.1 Construction in [4]) will be equal to $\frac{1}{N_\Delta}$. That means that the index of the window Δ defined as follows

$$r_{\Delta,u} = \Big(\sum_{x\in M_\Delta} \alpha_\Delta^2(x,u)\Big)^{1/2} = \frac{1}{\sqrt{N_\Delta}}. \quad (12)$$

So Ξ will be defined as follows:

$$\Xi = \Big\{\omega \in \Omega : \frac{1}{\sqrt{N}}|\sum_{x\in M_\Delta} \xi(x)| \le \kappa\Big\}. \quad (13)$$

In this case we get the following estimation using Theorem 1 in [4] with conditions (11), (12), (13):

$$|\hat{f}(x_0) - f(x_0)| \le \frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_0)}},\ when\ \omega \in \Xi. \quad (14)$$

Now we want to set up an adaptive threshold to apply it in Filtered Derivative Algorithm described above. Assuming that $\hat{f}(x)$ has Gaussian distribution $\mathcal{N}(\mu,\sigma)$, we can write out new expression for g_k:

$$\begin{aligned} g_k &= \sum_{i=0}^{N-1} \ln\frac{p_{\theta_1}(y_{k-i})}{p_{\theta_0}(y_{k-i})} = \sum_{i=0}^{N-1} \ln\frac{\frac{1}{\sigma\sqrt{2\pi}}\exp^{\frac{(y_{k-i}-\mu_1)^2}{2\sigma^2}}}{\frac{1}{\sigma\sqrt{2\pi}}\exp^{\frac{(y_{k-i}-\mu_0)^2}{2\sigma^2}}} \\ &= \sum_{i=0}^{N-1} \frac{(y_{k-i}-\mu_1)^2 - (y_{k-i}-\mu_0)^2}{2\sigma^2} \\ &= \sum_{i=0}^{N-1} \frac{2y_{k-i}(\mu_0-\mu_1) + \mu_1^2 - \mu_0^2}{2\sigma^2} \\ &= \sum_{i=0}^{N-1} \frac{y_{k-i}(\mu_0-\mu_1)}{\sigma^2} + N\frac{\mu_1^2-\mu_0^2}{2\sigma^2}. \end{aligned}$$

where μ_0 and μ_1 are the mean values of the distribution corresponding to probability density functions p_{θ_0} and p_{θ_1}.

Let us estimate the discrete derivative of g_k using (14):

$$\nabla g_k = g_k - g_{k-1} = S(y_k - y_{k-N}),$$

where $S := \frac{\mu_0-\mu_1}{\sigma^2}$. Suppose that $f(x_k)$ is the actual value of y_k, and $\hat{f}(x_k)$ is the estimated value as in (14). Using (14) we get

$$\begin{cases} |\hat{f}(x_k) - f(x_k)| \le \frac{6\kappa_1}{\sqrt{2n\delta_\kappa^*(x_k)}}, \\ |\hat{f}(x_{k-N}) - f(x_{k-N})| \le \frac{6\kappa_2}{\sqrt{2n\delta_\kappa^*(x_{k-N})}}. \end{cases} \quad (15)$$

Assume that $|S(f(x_k) - f(x_{k-N})| \ge h$, and $\kappa = max(\kappa_1,\kappa_2)$. Then we get

$$|S(\hat{f}(x_k) - \hat{f}(x_{k-N}))| \ge h - |S|\Big(\frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_k)}} + \frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_{k-N})}}\Big), \quad (16)$$

or

$$|S(\hat{f}(x_k) - \hat{f}(x_{k-N}))| \ge h - \frac{6\kappa|S|}{\sqrt{2n}}\Big(\frac{1}{\sqrt{\delta_\kappa^*(x_k)}} + \frac{1}{\sqrt{\delta_\kappa^*(x_{k-N})}}\Big). \quad (17)$$

Or

$$|S(\hat{f}(x_k) - \hat{f}(x_{k-N}))| \ge h - \frac{6\kappa|S|}{\sqrt{2n}}\Big(\frac{1}{\sqrt{\delta_\kappa^*(x_k)}} + \frac{1}{\sqrt{\delta_\kappa^*(x_{k-N})}}\Big). \quad (18)$$

Now we can formulate the following theorem.

Theorem *1:* Let y_k be the set of observations of the signal $f(x)$, and $g_k, \hat{f}, \kappa, n, \delta_\kappa, S$ be defined as above. Also let $\nabla g_k = S(f(x_k) - f(x_{k-N}))$ and $\nabla \hat{g}_k = S(\hat{f}(x_k) - \hat{f}(x_{k-N}))$. Then if $|\nabla g_k| \ge h$ we have:

$$|\nabla \hat{g}_k| \ge h - \frac{6\kappa|S|}{\sqrt{2n}}\Big(\frac{1}{\sqrt{\delta_\kappa^*(x_k)}} + \frac{1}{\sqrt{\delta_\kappa^*(x_{k-N})}}\Big). \quad (19)$$

The main objective of this theorem is to show how an adaptive threshold in the presence of Gaussian noise can be obtained using the constant threshold. For instance, in case of the slowdown detection described below, if we empirically select some threshold on the smooth road, then the Theorem 1 provides an estimation of the adaptive threshold on "the real world road with arbitrary potholes and bumps".

VI. Testing

A. Description of the Test Track

We have designed a small autonomous vehicle model (Fig.1), based on STM32F3-Discovery board containing the accelerometer and the gyroscope. The board was connected to a simple motor through the motor driver L293B. So we were able to control the direction of the vehicle movement. The power supply was provided by 4×1.2 V batteries or the USB connector.

We have implemented several software libraries in C: namely, the motor control, LED control, sensor, filter and fault detection algorithm libraries. One can use them to verify the algorithm his own [1]. Also we have implemented a simple program using this libraries to control the vehicle. Then we used OS Embox [2] (high-modular operating system for resource-constrained devices) to program the board.

One of the sufficient on practice properties of the proposed approach is very low requirements. For example, the board has only 48Kb RAM and ARM Cortex-M4 (with 3.40 CoreMark/MHz Performance Efficiency). And it follows that the algorithm could be applied for low-power and low-cost solutions.

We set the accelerometer to 1344 HZ and read the values of the acceleration along x-axis in the polling mode. During the movement we stored obtained values in the flash memory. Then we restored this values from the flash memory on PC.

[1] https://github.com/embox/embox/tree/master/platform/stm32f3_sensors

[2] https://github.com/embox/embox

Once obtained, these values were used to test the change detection algorithms. For this purpose a dedicated class library was implemented in Python.

This car should be able to recognize its own slowdown using the data obtained from the accelerometer. We made all experiments indoors, so we decided to replace a sand with a foam rubber (1 meter).

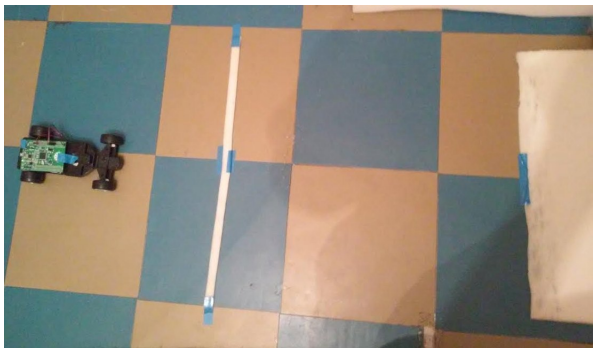

Fig. 1: Test track for the vehicle model. It begins with a smooth road. The tiny piece of a foam rubber is a random disturbance. The big piece of a foam rubber is a serious obstacle.

When the vehicle model tripped over a foam rubber a slowdown occurred, which we had to recognize. Also one can see the tiny strip of a foam rubber. It served to create a below threshold noise that should have been ignored by the system.

Our goal was to detect a foam rubber in real time, shut down the motor and turn on the red LED (Fig.2). This is demonstrated in this video [3].

B. The Experiment

On the Fig.3 we denoted with a red color the part that we intended to recognize. Also, one can see two areas marked by the green rectangles. They correspond to the moments when the front-wheel and the back-wheel of the vehicle tripped over the tiny piece of a foam rubber, accordingly.

At first, we should filter the noise out of the acceleration values. Let us suppose that the goal is to estimate the input

[3] https://www.youtube.com/watch?v=4rKXX11HOYE

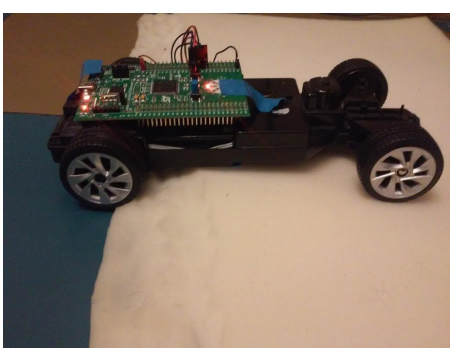

Fig. 2: The system detected a fault - motor has been shut down, the red LED has been turned on.

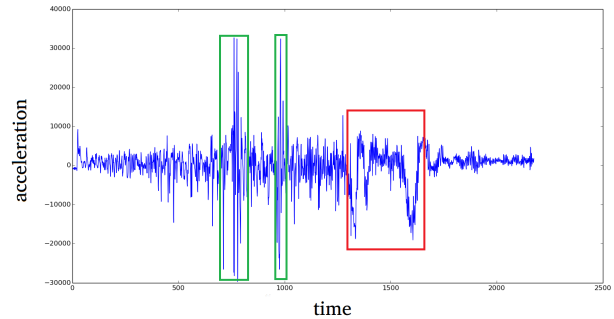

Fig. 3: Red rectangle denotes the moment when the fault was detected. Green rectangles denote the moments when the vehicle tripped over the tiny piece of a foam rubber that should have been ignored by the system.

signal at a given point x_0. We should use the spatial adaptive estimation, described above. Assuming $\kappa = 800$, we can calculate the ideal window Δ_k^* for every moment of time. Accordingly we use the method described in section "The idea" from [4]. The risk value is defined as follows

$$\rho_i = \kappa(n\delta_i)^{-1/2}.$$

And let us consider the following segments

$$D_i = [\hat{f}_i(x_0) - 2\rho_i, \hat{f}_i(x_0) + 2\rho_i].$$

Since the dynamic term $\omega_f(x_0, \delta_i)$ is dominated by the "stochastic" term ρ_i, one obtains that D_i, $i \leq i^*$ (i^* is an index of ideal window) have a point in common with $f(x_0)$. So we only need to construct D_i iteratively while they intersect each other.

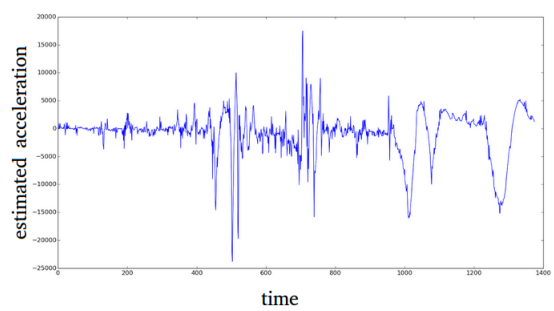

Fig. 4: Spatial adaptive estimation of the accelerometer data.

On the Fig.5 one can see how the window size (δ_i) changes during the time.

Now we need to detect a change in the mean value. It is clearly visible from the graphic, though, that for the most part the mean value is approximately equal to 0. Therefore, in this case the Filtered Derivative Algorithm is not appropriate.

If we take the absolute values of the signal, we get the following picture (Fig.7). One can see the abrupt change of the the mean value around $time = 1500$.

After obtaining the appropriate data (with a change in the mean value), we can apply the Filtered Derivative Algorithm

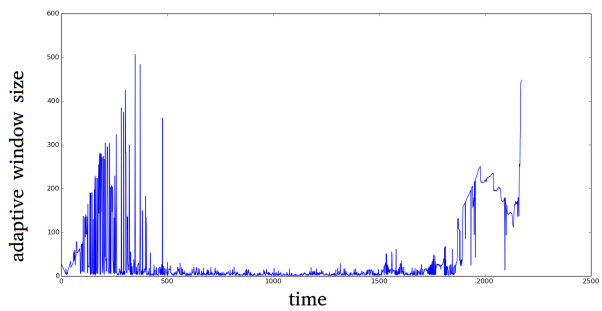

Fig. 5: Adaptive window size.

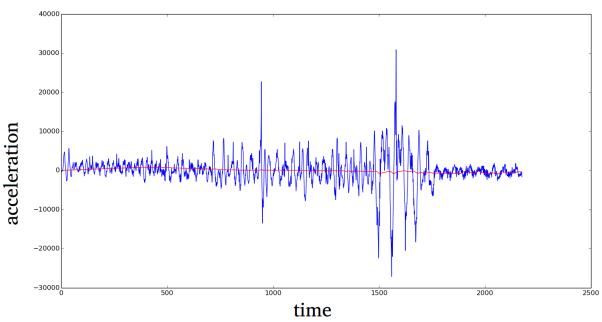

Fig. 6: Red line denotes the mean value of the accelerometer data (window size = 200).

to detect the change in the mean value. Let $f(x)$ be the real accelerometer data (without a noise), and $\hat{f}(x)$ – the estimation of $f(x)$ similar to the Theorem 1. Let $N = 2p$ and consider the p derivatives to be:

$$\begin{cases} |\nabla g_{k-1}| = |S(f(x_{k-1}) - f(x_{k-p-1}))| \geq h_1, \\ |\nabla g_{k-2}| = |S(f(x_{k-2}) - f(x_{k-p-2}))| \geq h_2, \\ \dots \\ |\nabla g_{k-p}| = |S(f(x_{k-p}) - f(x_{k-2p}))| \geq h_p. \end{cases}$$

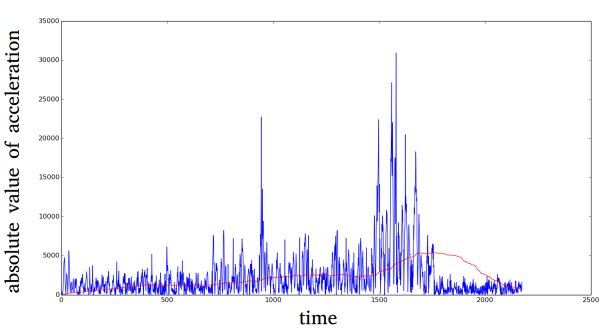

Fig. 7: Red line denotes the mean value of accelerometer data absolute values (window size = 200).

The same applies for $\hat{f}(x)$

$$\begin{cases} |\nabla \hat{g}_{k-1}| = |S(\hat{f}(x_{k-1}) - \hat{f}(x_{k-p-1}))| \geq \hat{h}_1, \\ |\nabla \hat{g}_{k-2}| = |S(\hat{f}(x_{k-2}) - \hat{f}(x_{k-p-2}))| \geq \hat{h}_2, \\ \dots \\ |\nabla \hat{g}_{k-p}| = |S(\hat{f}(x_{k-p}) - \hat{f}(x_{k-2p}))| \geq \hat{h}_p. \end{cases}$$

Let us define $G_k := \sum_{i=1}^{p} |\nabla g_{k-i}|$, $h := \sum_{i=1}^{p} h_i$, the alarm time $t_a := \min\{k : G_k \geq h\}$ and accordingly, in case of for $\hat{f}(x)$: $\hat{G}_k := \sum_{i=1}^{p} |\nabla \hat{g}_{k-i}|$, $\hat{h} := \sum_{i=1}^{p} \hat{h}_i$, and the alarm time $\hat{t}_a := \min\{k : \hat{G}_k \geq \hat{h}\}$. Our purpose is to estimate $\hat{h}$ using h. One can obtain this estimation by applying (19) to each of g_k and then sum the results up.

$$\hat{h} = h + \frac{6\kappa|S|}{\sqrt{2n}} \sum_{i=0}^{p-1} \left(\frac{1}{\delta^*_{k-i}(x_{k-i})} + \frac{1}{\delta^*_{k-i}(x_{k-p-i})} \right).$$

Assuming the window size of $N = 150$ ($p = 75$), we select the threshold h for the signal without noise. With reference to the range of accelerometer values, we chose the h value to be $h = 2000$. The calculated adaptive threshold is presented below.

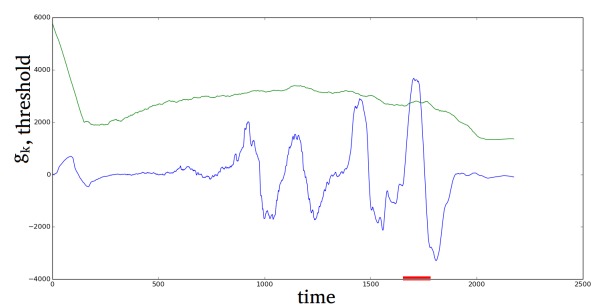

Fig. 8: Blue line is ∇g_k, green line is an adaptive threshold. The region where a fault occurred is depicted by red color.

One can see a moment when the fault occurred - intersection of the green and blue lines. Also, one should notice that if we used the constant value $h = 2000$, the intersection would have occurred in the wrong place.

We provide the results of the experiments with different values of κ using 10 tracks of the accelerometer data for the each value of κ. The best result was shown by $\kappa = 800,\ 900,\ 1000$, but it is difficult to calculate exactly which value of κ should be preferred in the concrete situation, since there is no way to obtain the signal-to-noise ratio except for the empirical estimate. The results are presented in Table 1

VII. Conclusions and Acknowledgments

The problem of fault detection in autonomous systems is considered and the new method of the adaptive threshold estimation is offered. The experiment performed with the small autonomous vehicle model demonstrates the increased efficiency of the proposed technique with adaptive threshold, compared to the conventional method using the constant threshold.

For the further investigations the estimation of the κ parameter in the conditions of the unknown value of the

κ	percent of fails
400	60%
600	40%
700	30%
800	10%
900	10%
1000	10%
1500	20%
1600	30%
2000	60%

TABLE 1: Fail counts according to the different values of κ.

signal-to-noise ratio seem to be of interest. Also, more detailed description of a surface type could be useful. This could possibly be performed by means of hypothesis testing, with each of the hypotheses describing the specific surface type.

Moreover, we want to consider the possible applications of this method for the medical diagnosis of biological processes in the human body.

The authors are grateful to Arkadi Nemirovski for the interesting and helpful discussion regarding the noise filtering with moving average using adaptive window size.

REFERENCES

[1] P. Akimov and A. Matasov, "Weight and time recursions in dynamic state estimation problem with mixed-norm cost functions," *Automatic Control, IEEE Transactions on*, vol. 60, no. 4, pp. 1050–1063, 2015.

[2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Inc, 11 1998.

[3] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*. Springer, 2006.

[4] A. Goldenshluger and A. Nemirovski, "On spatial adaptive estimation of nonparametric regression," *Mathematical methods of Statistics*, vol. 6, pp. 135–170, 1997.

[5] O. Lepskii, "On one problem of adaptive estimation in gaussian white noise," *SIAM Theory of Probability and Its Applications*, vol. 35, no. 3, pp. 454–466, 1990.

[6] M. Perhinschi, M. Napolitano, G. Campa, B. Seanor, J. Burken, and R. Larson, "An adaptive threshold approach for the design of an actuator failure detection and identification scheme," *Control Systems, IEEE*, vol. 14, no. 3, pp. 519–525, 2006.

[7] A. Davis, S. Nordholm, and R. Togneri, "Statistical voice activity detection using low-variance spectrum estimatio and an adaptive threshold," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 2, pp. 412–424, 2006.

[8] O. Granichin, V. V. Zeev, and D. Toledano-Kitai, *Randomized Algorithms in Automatic Control and Data Mining*, 2015.

[9] C. George and R. L. Berger, *Statistical Inference*, 2001.