

Competitive Detector of Changes with a Statistical Test

Leszek J Chmielewski, Konrad Furmańczyk, and Arkadiusz Orłowski

Faculty of Applied Informatics and Mathematics – WZIM,
Warsaw University of Life Sciences – SGGW
ul. Nowoursynowska 159, 02-775 Warsaw, Poland
{leszek_chmielewski,konrad_furmanczyk,arkadiusz_orlowski}@sggw.pl
<http://www.wzim.sggw.pl>

Abstract. The detector of jumps or changes in the function value and its derivative designed with the use of the concept of competing approximators is revisited. The previously defined condition for the existence of a jump in the function value is extended by introducing a statistical test of significance. This extension makes it possible to eliminate some false positive detections which appeared in the previously obtained results. The features of the extended detector are demonstrated on some artificial and real-life data.

Keywords: competitive detector, function change, statistical test

1 Introduction

The detector of changes in a one-dimensional signal which is our object of interest originated from the concept of a filter in two-dimensional images denoted as the *competitive filter* in [10,11]. The change detection ability of this filter was later noticed in [3]. The detector did not work well in two-dimensional images [4] but for one-dimensional signals it appeared useful and made it possible to detect the change in the function value as well as in its first derivative [5]. In the present paper we shall put aside the filtering effect and pay attention to the detection of the change of the function value only.

One of the problems noticed in [5] was that the detector made some false positive errors. In this paper we shall complement the basic detector with a statistical test which in our opinion will reduce the number of false positives.

The question of detecting changes in signals is one of the domains of intensive research. Within the domain of image processing it was surveyed in [1,2,9] and in motion detection in [7,12]. However, it seems that the concept of competitiveness understood as in the paper [3] was absent in the research reported.

Because the detector of our interest has its origin in the domain of image processing, the change will be sometimes called a jump or an edge, the change in the value of the function will be called a step and the change in the derivative of the function can be also called a roof.

The concept of the method does not need any assumption on the nature of the data analyzed. The only operation on the data is the approximation with polynomial functions, one at each side of the considered point, without the condition of continuity at this point. In the present paper we use the simple approximation with a linear function, which can be treated as the first approximation of a potentially more developed approach. The assumptions are introduced in the statistical test of the significance of the detected edge. At present, we make an assumption of the Gaussian distribution of noise, but this is also only the first, simplest example, which can easily be replaced with more advanced approaches.

This paper is organized as follows. In the next Sect. 2.1 the main concept of the detector, already described in [5], is reminded in its basic form. As a complement of the heuristic criterion of edge existence, described in Sect. 2.2, the statistical criterion proposed in this paper is introduced in Sect. 2.3. In Sect. 2.4 the functioning of the detector with both criteria is explained on simple data. Finally, the ability to reject false positives in noisy data is shown in Sect. 3.1 and the detection of changes in some real-life data is presented in Sect. 3.2. The results are discussed in Sect. 4 and the paper is closed in Sect. 5.

2 The method

2.1 General concept

A sequence of measurements $z(x) = y(x) + n(x)$, where the independent variable x is discrete and $n(x)$ is noise, will be considered. The filtering and detection is performed in the point x_0 called the *central point*. If x is time, the past measurements are considered to be known up to the point x_0 , and also the further measurements, up to $x_0 + D$, are considered as known. The competitive structure of the detector can be seen in that two approximators, referred to as the *Left* and the *Right* one, are used to find $y(x_0)$. The first one operates on the *past* data at the left side of x_0 , using $z(x), x \in [x_0 - s - \Delta, x_0 - \Delta]$ to find $\hat{y}_L(x_0)$. The second one operates on the *future* data, at the right side of x_0 , using $z(x), x \in [x_0 + \Delta, x_0 + s + \Delta]$ to find $\hat{y}_R(x_0)$. The parameter s is the scale of the filter. The parameter Δ is the gap between the central point x_0 and the estimators.

Each approximator makes its error. Errors $e_L(x_0)$ and $e_R(x_0)$, respectively, can be approximated from the differences between the data and the approximated values. The filtered value at the central point, $\hat{y}(x_0)$, is taken as the the output of that filter which has a smaller error. In this the competitiveness of the filter can be seen. As the value of the output, the value at $x_0 - \Delta$ from the left approximator, or for $x_0 + \Delta$ for the right one, is used, to avoid using the extrapolated values. This gives stabler results than in the case of values extrapolated to x_0 .

As in [3], linear least square approximators are used and their mean square errors are used as the approximation errors.

2.2 Heuristic criterion of jump

When the results from the two approximators are known, they can be used to find the jump in the function value, that is, the step edge intensity value E_0 , and jump in the first derivatives, that is, the roof edge intensity value E_1 , as the difference of outputs from the two approximators $\hat{y}_L(x_0)$, $\hat{y}_R(x_0)$, and their derivatives

$$\begin{aligned} E_0(x_0) &= \hat{y}_R(x_0 + \Delta) - \hat{y}_L(x_0 - \Delta), \\ E_1(x_0) &= \hat{y}'_R(x_0 + \Delta) - \hat{y}'_L(x_0 - \Delta), \end{aligned} \quad (1)$$

where $\hat{y}'(\cdot)$ can be found from the approximators due to that they are linear. The question remains, where is the jump.

The conditions for the existence of the jump is that the graphs of the approximation errors cross in such a way that for increasing x the error from the past increases and for decreasing x the error for the future increases. In the present paper it is assumed that at least one of the errors increases in this case. These conditions can be expressed as

$$\begin{aligned} e_R(x_0 - \delta) &> e_L(x_0 - \delta) \wedge e_R(x_0 + \delta) < e_L(x_0 + \delta), \\ e_R(x_0 - \delta) &> e_R(x_0 + \delta) \vee e_L(x_0 - \delta) < e_L(x_0 + \delta). \end{aligned} \quad (2)$$

Because the future error should be known for $x_0 + \delta$ then the measurements for $x_0 + D = x_0 + \Delta + s + 2\delta$ should be known. The parameter δ can be called the neighborhood parameter. For simplicity, it is assumed $\Delta = \delta = 1$, so $D = s + 3$.

The crossing of the graphs of errors around a jump is illustrated in Fig. 1. Let us imagine the process of filtering and edge detection in such a way that the central point, with the two approximators at its left and right side, move along the data from left to right. When a step is encountered, first the right approximator moves over it. The step enters the right approximator's support. Therefore, the error of the right approximator goes up, as in Fig. 1a. As the analyzed point moves forward, the step leaves the support of the right approximator, so its error goes down, and enters that of the left one, as in Fig. 1b. Hence, the error of the left approximator increases. When both approximators leave the region of the step, both errors go down.

It can be noticed that there are no separate conditions for the two edge types detected. In the case of detection, if one of the edge types is missing, then its intensity is zero (examples will be shown further in Fig. 2, where the roof edge is zero for $x = 10, 11$ or the step edge is zero for $x = 20$).

2.3 Statistical criterion of significance of a jump

The process of crossing the graphs of errors described in the previous Section goes on precisely in the described way provided that the edges are isolated, with respect to the scale s . However, it is not always so; therefore, sometimes the false

The graphs used in this paper as well as the software were developed in Matlab[®].

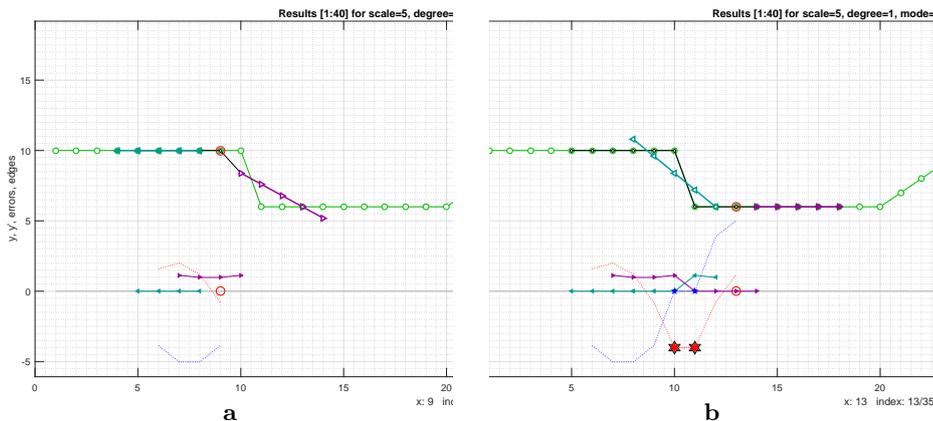


Fig. 1. Intermediate results for the two approximators for $x_0 = 9$, before the jump, and $x_0 = 13$, after the jump. Graphs of errors (thin magenta and cyan lines) cross between points 10 and 11. Pale green (■): function; dark cyan (■): left error, dark magenta (■): right error; red (★) star: jump of the function. Current central point x_0 marked with a red circle on the axis and on the graph. The left and right approximators around the central point shown with thicker cyan and magenta lines. The approximator with zero error has full triangular marks, the other one has marks filled with white. Other symbols will be explained further.

positive detections (as well as false negative ones) can occur. This is why we have introduced a simple mechanism of additionally testing the edge significance in a statistical way, to exclude false positive detections.

In the present Subsection some notations will contain a superscript s as *statistical* to underline the differences between these notations and those from the previous text. Finally the mutual relations of the relevant notations will be explained.

Let us assume that the sequence of measurements form a piecewise linear signal, not necessarily continuous, with additive Gaussian noise. For an isolated point x_0 it is observed $y(x) = a_L + b_Lx + \epsilon_x$ for $x < x_0$ and $y(x) = a_R + b_Rx + \epsilon_x$ for $x \geq x_0$, where the noise ϵ_x has a zero mean normal distribution. There is a jump at x_0 if $\theta = a_R - a_L \neq 0$. Let us verify a hypothesis $H_0 : \theta = 0$ – the jump is absent, against the alternative hypothesis $H_1 : \theta \neq 0$ – the jump is present. To verify this, the test statistics $|\hat{y}_R^s(x_0) - \hat{y}_L^s(x_0)|$ is used, where $\hat{y}_L^s(x_0)$ is a linear regression function of s points on the left of x_0 , without this point, that is, from the set $X_L = x \in [x_0 - s, x_0 - 1]$, and $\hat{y}_R^s(x_0)$ is a linear regression function of s points on the right of x_0 , with this point, that is, from the set $X_R = x \in [x_0, x_0 + s - 1]$. An isolated jump is detected if

$$P(|\hat{y}_R^s(x_0) - \hat{y}_L^s(x_0)| > t_\alpha) = \alpha, \tag{3}$$

where α – significance level (value $\alpha = 0.05$ will be assumed throughout). Provided the hypothesis H_0 holds, the distribution of $\hat{y}_R^s(x_0) - \hat{y}_L^s(x_0)$ is zero-mean normal with variance $\sigma_L^2 + \sigma_R^2$, where $\sigma_L^2 = \text{Var}(\hat{y}_L^s)$ and $\sigma_R^2 = \text{Var}(\hat{y}_R^s)$.

Therefore,

$$t_\alpha = \sigma \Phi^{-1}(1 - \alpha/2), \quad (4)$$

where $\sigma = \sqrt{\sigma_L^2 + \sigma_R^2}$. In practice, as σ , the standard residual error of the respective estimator \hat{y}_i^s is taken.

The step of the derivative has not been considered at present. The test used can be considered as a greatly simplified, basic version of tests described in [6,8].

It should be noticed that the normality of the noise distribution is assumed and this assumption will not be verified. Moreover, for some of the measurements studied with the method considered, such an assumption is merely a convenient model, but not an actual process in which the signal appears. Nevertheless, we shall use this model as a way to interpret the data in which unknown processes give rise to complex patterns and in which we seek an explanation in terms of simplified events.

2.4 Results for both criteria

In the heuristic detector, as the jump in x_0 the relation between measurements concerning points $x_0 - 1, x_0 + 1$ are considered, while in the statistical detector this concerns $x_0 - 1, x_0$. To use the condition (3) in the common setting together with the conditions (2) the following should be noted. Due to the structure of sets X_L, X_R around x_0 , the following relations between the approximations from the heuristic notation and the regressions from the statistical notation hold

$$\begin{aligned} \hat{y}_L^s(x_0) &= \hat{y}_L(x_0), \\ \hat{y}_R^s(x_0) &= \hat{y}_L(x_0 - 1), \end{aligned} \quad (5)$$

and similarly the error measures from the heuristic notation can be related to the standard residual errors from the statistical notation.

To come to a common meaning of a jump in x_0 it can be considered, in the statistical formulation, that a jump exists if there is a jump between $x_0 - 1, x_0$ or between $x_0, x_0 + 1$. If only one jump exists, its value θ is taken as the step of the function $y(x)$. If both are present, the one having a larger modulus is taken.

If there is an edge according to (2) and (3), then a statistically significant edge exists. If (2) holds and (3) does not, then the edge is statistically insignificant and it is dismissed. If (2) is false, then there is no need to check (3), although in the present paper both conditions are calculated independently to show the results in a detailed way.

In Fig. 2 the result is shown for data in which all the changes detectable by the heuristic algorithm are present: a step edge, a roof edge and a combined step and roof edge. The data are synthetic and clean. What is apparent is that the roof edge is detected in a single point, like this at $x = 20$, while the step edge, like this at $x = 10, 11$ is found at two points. This is correct due to that the jump of a discrete function appears between two points.

It can be noted that the statistical criterion tends to detect very small changes of the value if the error measures are small, due to that in (4) the threshold depends on the variance. This gave rise to a continuous edge between $x = 20$ and 26. This edge was not accepted by the heuristic condition, though.

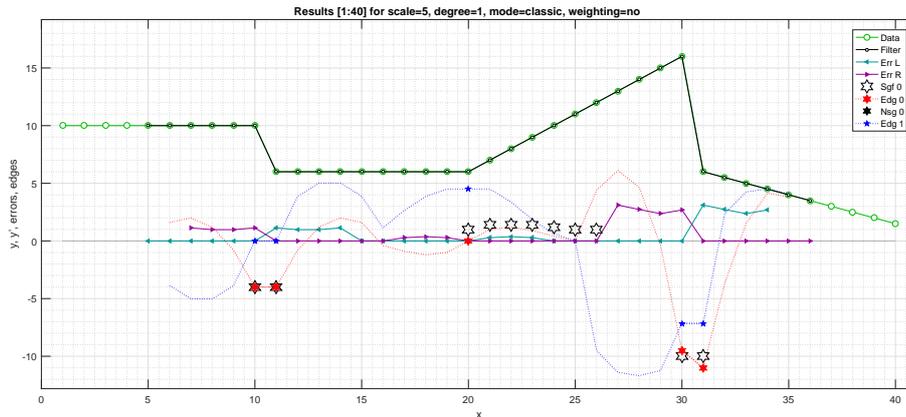


Fig. 2. Results for synthetic data with all the detectable changes represented: jump of the value at $x = 10, 11$, jump of the derivative at $x = 20$ and combined jump of value and derivative at $x = 30, 31$. The meaning of types and colors of lines and marks partly explained in the legend. **Err L**: left error, **Err R**: right error; **Edg 0**: jump of the function, marked with a red star; **Edg 1**: jump of its first derivative, marked with a blue star. **Sgf 0**: Statistically significant edge ((3) true) marked with an empty black star (slightly larger than other stars so that they do not obscure each other). **Nsg 0**: Statistically nonsignificant edges ((2) true and (3) false), marked with full black stars (there are no such points in this image). Angles shown in tens of degrees.

3 Examples

3.1 Rejection of false positives in noise

The ability to reject the less significant step changes was tested with data with noise which was actually Gaussian, zero-mean, with $\sigma = 10$. The results are shown in Fig. 3. It can be seen that indeed, some, but not all, false positive detections were successfully rejected, while the most significant, strong jump at $x = 30, 31$ was constantly maintained.

3.2 Real-life example

As an example let us consider the graph of processor load of some web server in 100 intervals of one minute each, measured on 04 July 2016 (Fig. 4). The load was averaged in each minute. The graph indicates an uneven load so there are important changes of the values. It is interesting to see that the competitive detector tends to find the steps but is not sensitive to the changes consisting in a steady increase or decrease of the value. The statistical detector, however, detects a steady increase, like in Fig. 4b at $x = 60, 65$. In this way, the criteria cooperate in forming the right decision and complement each other.

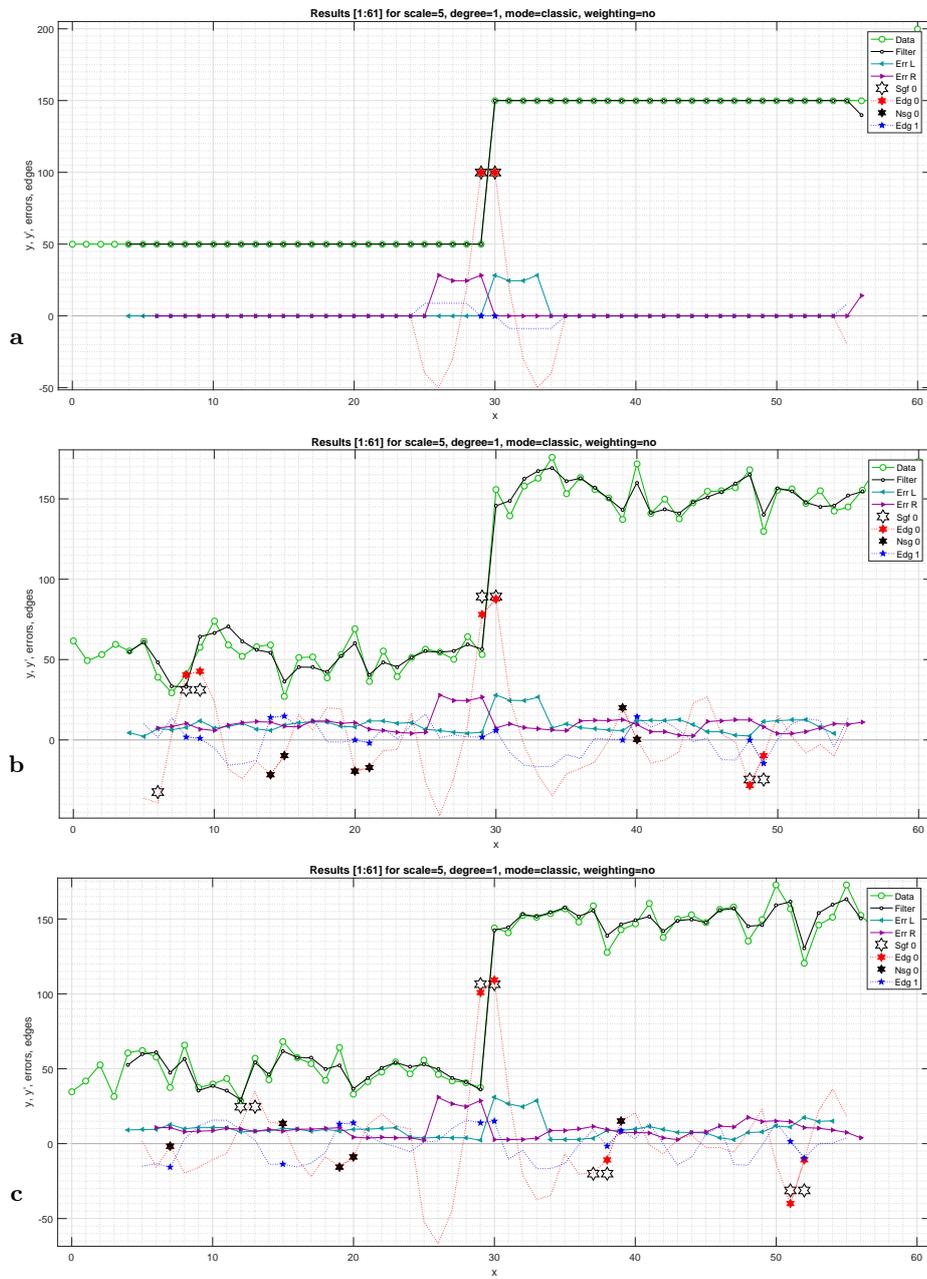


Fig. 3. A test step without noise (a) and with additive zero-mean noise with $\sigma = 10$ ((b) and (c), two realizations). It can be seen that some false positives were successfully rejected.

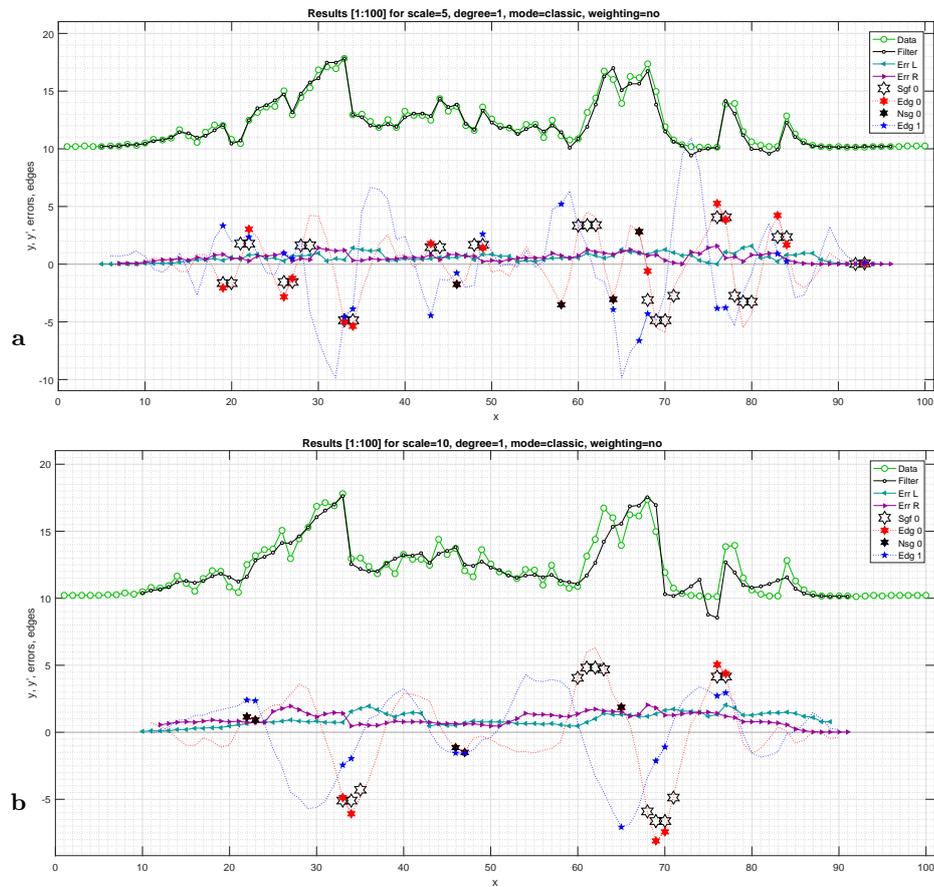


Fig. 4. Analysis of CPU load of a web server averaged for 1 min intervals in 100 min. Scale: (a) $s = 5$ measurements; (b) $s = 10$ measurements. At both scales it can be seen that some of the less significant jumps were dismissed by the statistical criterion. At the larger scale the minute details are neglected. Angles shown in tens of degrees. Results scaled and moved up by 10 units to make the error graphs visible.

4 Discussion

The use of the statistical test reduced the number of false positive detections. The number of false negatives sacrificed seems to be small, but this needs further analysis.

The considered algorithm is characterized by a set of advantages and drawbacks. As the advantages, the following features can be named. Two approximators are used so the jump can be directly modelled. The complexity of the algorithm with respect to the size of the data is linear due to that only a local neighborhood of a data point of a fixed size is considered. Coming to the drawbacks, it should be said that some data concerning the future with respect to

the considered data point should be known to perform the analysis, and that the method has some free parameters which should be selected, while the criteria for such selection are not self-explanatory.

5 Summary and prospects

The concept of the competitive filter was extended by adding the statistical test used to check the significance of the jump of the function value. In the test, the results available from the calculations already performed are used, so the computing load is small and the complexity of the algorithm remains linear with respect to the data size.

The introduction of the statistical test made the number of false positive detections smaller. The the test can be used as a post-processor of the detection results. The design of the test can be extended to the derivatives of the function and its form can be improved. Also the assumptions on the distribution of noise in the data can be changed and the criterion can be reformulated accordingly. This stage of research can be treated as the proof of concept only, but the idea of combining the statistical testing with the heuristics seems to be one of the promising directions of the development of the concept of competitive filtering and detection.

References

1. Basu, M.: Gaussian-based edge-detection methods-a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 32(3), 252–260 (Aug 2002), [doi:10.1109/TSMCC.2002.804448](https://doi.org/10.1109/TSMCC.2002.804448)
2. Bhardwaj, S., Mittal, A.: A survey on various edge detector techniques. *Procedia Technology* 4, 220–226 (2012), [doi:10.1016/j.protcy.2012.05.033](https://doi.org/10.1016/j.protcy.2012.05.033)
3. Chmielewski, L.: The concept of a competitive step and roof edge detector. *Machine Graphics & Vision* 5(1-2), 147–156 (1996),
4. Chmielewski, L.: Failure of the 2D version of the step and roof edge detector derived from a competitive filter. Report of the Division of Optical and Computer Methods in Mechanics, IFTR PAS (Dec 1997),
5. Chmielewski, L.J., Orłowski, A.: Detecting changes with the robust competitive detector. In: Alexandre, L., et al. (eds.) *Proc. 8th Iberian Conference on Pattern Recognition and Image Analysis IbPRIA 2017. Lecture Notes in Computer Science*, vol. 10255. Springer, Faro, Portugal (20-23 Jun 2017), [doi:10.1007/978-3-319-58838-4_39](https://doi.org/10.1007/978-3-319-58838-4_39)
6. Furmańczyk, K., Jaworski, S.: Large parametric change-point detection by a v-box control chart. *Sequential Analysis* 35(2), 254–264 (2016), [doi:10.1080/07474946.2016.1165548](https://doi.org/10.1080/07474946.2016.1165548)
7. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 34(3), 334–352 (Aug 2004), [doi:10.1109/TSMCC.2004.829274](https://doi.org/10.1109/TSMCC.2004.829274)

8. Jaworski, S., Furmańczyk, K.: On the choice of parameters of change-point detection with application to stock exchange data. *Quantitative Methods in Economics* XII(1), 87–96 (2011)
9. Maini, R., Aggarwal, H.: Study and comparison of various image edge detection techniques. *International Journal of Image Processing (IJIP)* 3(1), 1–11 (2009)
10. Niedźwiecki, M., Sethares, W.A.: New filtering algorithms based on the concept of competitive smoothing. In: *Proc. 23rd ISICIE Symp. on Stochastic Systems and its Applications SSS'91*. pp. 129–132. Osaka, Japan (1991)
11. Niedźwiecki, M., Suchomski, P.: On a new class of edge-preserving filters for noise rejection from images. *Machine Graphics & Vision* 1-2(3), 385–392 (1994)
12. Rätty, T.D.: Survey on contemporary remote surveillance systems for public safety. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40(5), 493–515 (Sept 2010), doi:[10.1109/TSMCC.2010.2042446](https://doi.org/10.1109/TSMCC.2010.2042446)