Segmentation of Experimental Curves Distorted by Noise

Vladimir Kalmykov[1*] and Anton Sharypanov[2]

[1]IMMSP, Prosp. Acad. Glushkova 42, 03680, Kiev 187, Ukraine
[2]ICYB, Prosp. Acad. Glushkova 40, 03680, Kiev, Ukraine

Abstract

A new segmentation method of signals distorted by noise is proposed. Unlike other known methods, for example, the Canny method, a priori data on interference and/or a signal (image) is not used. Segmentation of signals and halftone images distorted by interference is one of the oldest problems in computer vision. But human vision solves this task almost independently of our consciousness. It was discovered for vision neurons, that sizes of receptive fields' excitatory zones change during visual act, which eventually mean dynamical changes in visual system's resolution i.e., coarse-to-fine phenomenon in living organism. We assumed that "coarse-to-fine" phenomenon, i.e., several different resolutions, is used in human vision to segment images. A "coarse-to-fine" algorithm for segmentation of experimental graphs was developed. The main difference of algorithm mentioned above from others is that decision is made taking into the account all partial solutions for all resolutions being used. This ensures stability of final global solution. The algorithm verification results are presented. It is expected that the method can naturally be expanded to segmentation of halftone images.

Keywords: Experimental curves; Segmentation; Coarse-to-fine

Introduction

Most information systems accumulate information in the form of various graphs, various kinds of images, which are required to obtain the opinion of the expert, in other words, to take this or that decision. A large and constantly growing volume of data is forcing the development of tools that allow reducing the burden on the decision-makers through automatic and/or automated processing of raw data.

The initial data or experimental curves represent measurement results that are usually distorted by interference. Graphs, contours of objects on images seems to be the simplest and long used method for cognitive presentation of measurements in various areas of human activity, which allow to evaluate the qualitative properties of the process despite on the interference and measurement error. The most basic feature of a graph or contour is its shape, which reflects a function that generates a visible representation of the curve. It is the form of graphic curve that characterizes parameters of the reflected object or process. It is assumed that the measured data is the representation of some unknown function y=f(x), that is defined on measuring range [a, b]. Measurement result is a finite sequence of I pairs $\{x_i, y_i\}$; i=1, 2…I. In other words, there is a tabular implementation of a given function.

Because various representations related to that same object can vary in scale, in noise level, in number of measurements etc. it is not possible to use methods based on neural networks or statistical image recognition methods directly to address the problem of comparison graphs or experimental curves. In that case the unknown functions that define experimental curves should be approximated with functions that are invariant to affine transformations in order to automatically compare them on later processing stages.

Analytical description of a curve based on parametrically defined splines [1,2] is one of the suitable forms for further processing and analysis. However, the methods of presentation of the experimental curves by splines suggest that obtained experimental curves represent processes or phenomena that are determined by unknown smooth functions. At the same time, a large number of practical problems require processing of experimental curves that can be represented adequately only by unknown piecewise smooth functions. It is natural

to assume that the approximating function must also be piecewise smooth.

The curve defined by function y=f(x), (a ≤ x ≤ b) is piecewise smooth if the function y=f(x) has a finite number of discontinuities on [a, b] and the segment [a, b] can be divided by points into a finite number N of partial segments, so that function y=f(x) has continuous derivatives not equal to zero simultaneously on each segment.

If splines are selected as approximating function, this function takes the form of a sequence of polynomials

$$y = f^{(1)}(x)\,|\,(t_0 \le x < t_1),\ f^{(2)}(x)\,|\,(t_1 \le x < t_2)...,\ f^{(N-1)}(x)\,|\,(t_{N-1} \le x < t_N)$$

where

$$f(x)\,|\,(t_n \le x < t_{n+1}),\ a_0^{(n)}x^k + a_1^{(n)}x^{k-1} + a_{k-1}^{(n)}x + a_k^{(n)}$$

Thus, it is necessary to consider a set of boundary points $T=\{t_0, t_1, …, t_N\}$ and their number N+1 in order to segment the experimental curve (Figure 1).
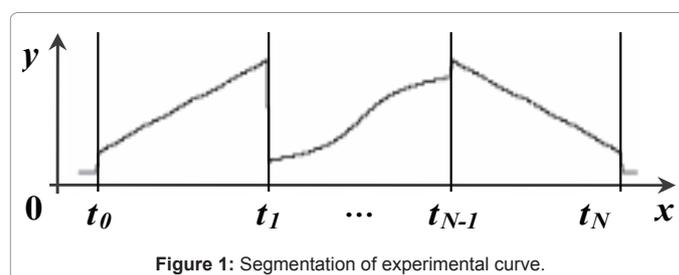

Figure 1: Segmentation of experimental curve.

*Corresponding author: Vladimir Kalmykov, IMMSP, Prosp. Acad. Glushkova 42, 03680, Kiev 187, Ukraine, Tel: +380502051153; E-mail: vl.kalmykov@gmail.com

Phenomenon of selecting the separate segments on a graph is an act of visual perception by its nature. The discontinuities of the curve or its gradient are identified visually and used for making decisions. Segmentation of images (i.e., object contour selection) presumably has the same nature as graph segmentation.

So, the mechanisms of visual perception and known methods of image segmentation should be taken into account when developing the method for segmentation of experimental curves.

The purpose of the article is to introduce a new curve segmentation algorithm necessary for automated signal processing and to present its results when applied to signals distorted by noise.

Necessary Additional Mathematical and Neurophysiological Information

Researches in human vision have shown that visual neurons process signals from a set of receptors that form receptive field of a neuron. The simplest receptive fields are circle in shape and discrete by their nature. Receptive fields of neighbouring neurons overlap [3]. Later it was discovered that the excitatory zone of the receptive field doesn't stay constant during a visual act (approx. 150 ms). It decreases from maximum to 1-2 receptors wide (Figure 2). After a while this phenomenon was carefully investigated in [4]. It was also shown in [4] that the number of neurons being activated decrease during stimulation of visual act. Eccentric stimulus spot presented outside minimum field center but inside maximum field center gave a fast-initial response that disappeared as center shrank toward minimum.

Thus, it is possible to assume that visual neuron could be matched to a point in some discrete two-dimensional space and the excitatory zone of the neuron could be matched to discrete representation of point neighbourhood. Then changes in receptive field excitatory zone could be matched to changes in point neighbourhood. During one visual act image processing of the object in the field of view proceeds in visual system with different resolution changing from minimal (coarse, blurred) to maximal (sharp image).

The definition of function continuity in ε-δ form states that if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all values of variable x from δ-neighbourhood of point c the function values f(x) belong to ε-neighbourhood of f(c) (Figure 3).

This definition is used successfully for function analysis but it could not be used in analysis of experimental curves. Experimental
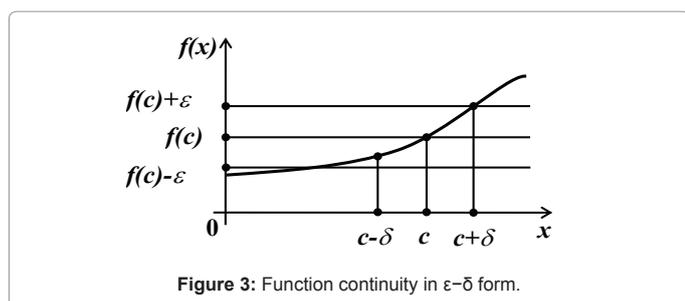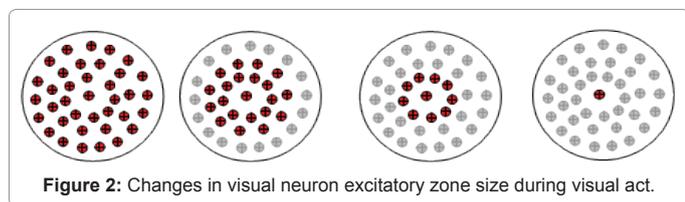
curves are representations of respective unknown functions presented as sequences of measurements. A sequence of measurements is a set of points in some discrete space. It should be emphasized that for checking the continuity condition of f(x) in a given point c the sequence of function values considers. Starting from a specific value $|x_1 - c|$ the neighbourhood of point c reduces ($|x_1 - c| > |x_2 - c|$, $|x_2 - c| > |x_3 - c|$, …) tending to 0. f(x) considers continuous in c if the neighbourhood of f(c) reduces to 0 at the same time ($|f(x_1) - f(c)| > |f(x_2) - f(c)|$, $|f(x_2) - f(c)| > |f(x_3) - f(c)|$, …). So, changing neighbourhood of c is used for analysis of function continuity in a given point.

Decreasing of receptive field excitatory zone sizes could be considered as decreasing of neighbourhood of a dot in the center of this receptive field. The process that is used in calculus for analyzing function continuity in a dot repeats in visual system of humans and animals at each visual act. The essential difference between changes of resolution in visual system during visual act and analysis of function continuity in a dot is that the elements of receptive field are the objects of discrete space. Similar to that, the definition of function continuity from calculus is not suitable for continuity analysis of experimental curves because they are reflections of unknown functions and presented as sequences of values that in turn are the sets of dots in some discrete space. But at the beginning of visual act the excitatory zones of receptive field consist of many dots (receptors). Until the sets of receptors in the receptive fields' excitatory zones are not empty the application of continuity definitions to brightness function defined in discrete space of receptors does not come in contradiction with classical theory of continuity of functions. Mentioned above phenomena being observed in visual system of living beings could be used to prototype a new method of signal processing based on variable resolution concept.

Image Processing Methods that Use Resolution Changes

Human vision solves the problem of object detection in the field of view and its shape identification seamlessly, on subconscious level, even for noisy signals (Figure 4). So, it seems reasonable to suggest that creation of technologies for processing visual representation of different signals and researches in neurophysiology of visual perception should be considered jointly because of the fact that they share the same subject of research-visual perception.

Usually image processing systems deal with an image distorted by noise. In the simplest case input image is convolved with (typically) Gaussian filter in order to remove unwanted details and processing algorithm is applied to blurred image. For example, Canny edge detector [5] relays on that processing procedure. The result depends on σ parameter of Gaussian filter (Figure 5). No recommendations were issued in original work on how to choose that parameter.

Other filters can also be used during signal pre-processing aiming at removing noise from the original signal. The result depends on the aperture size of the filter which is the unknown parameter.

It was shown in [6] that processing of artificially blurred images allows to solve problems that cannot be addressed at all by traditional image recognition methods but nothing was said about choosing the reasonable degree of that blur.

In some fields of science, it is not possible to get sharp image at one shot. In biological imaging with conventional light microscopy a problem of limited depth of focus exists. If the specimen is thicker than attainable focal depth portions of the object surface outside the focal plane will be defocused. To overcome this, multiple shots of specimen are taken with different focal depth along optical axis resulting in
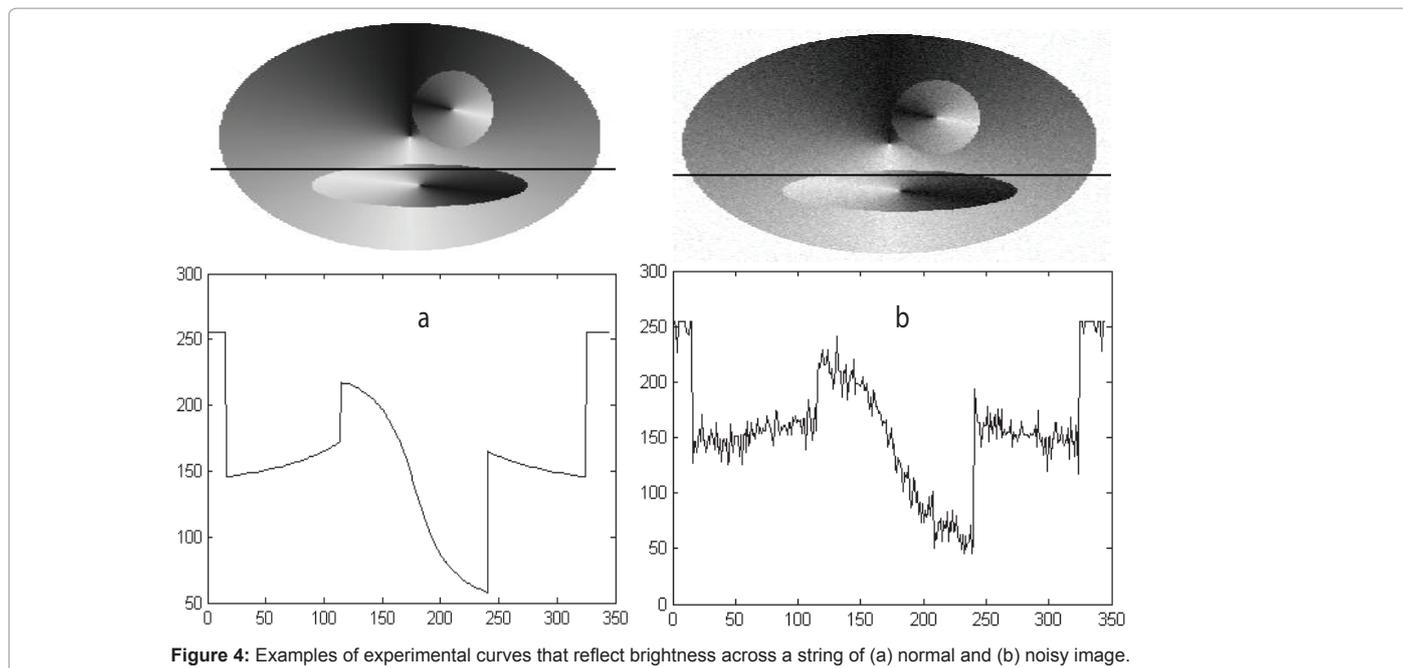

Figure 2: Changes in visual neuron excitatory zone size during visual act.

Figure 3: Function continuity in ε−δ form.

Figure 4: Examples of experimental curves that reflect brightness across a string of (a) normal and (b) noisy image.
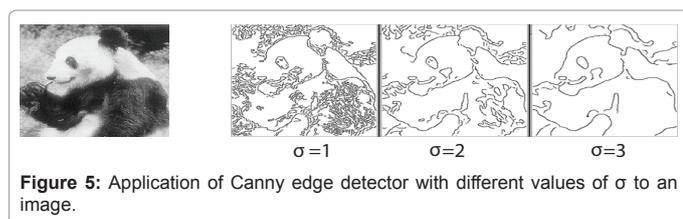

Figure 5: Application of Canny edge detector with different values of σ to an image.

series of images where the certain parts of specimen appear in and out of focus. After that a procedure of reconstruction from image set involving wavelets follows in order to obtain an image that is sharp everywhere [7].

However, blurring of an image may be considered as decreasing of physical resolution. At the same time decreasing of physical resolution is widely used to reduce the computational complexity of existing image processing algorithms in order to get performance gain.

In [8] proposed a multi-resolution part based model and a corresponding coarse-to-fine inference algorithm. It is based on the observation that matching of each part of the image is the most expensive computational operation in comparison to detection of significant parts and computation of their optimal configuration, so the minimization of number of part-to-image comparisons implies detection acceleration. Starting from matching the lowest resolution part the method selects only the best placement in each image neighbourhood. These locally optimal placements are then propagated recursively to the parts at higher resolution. By recursive elimination of unlikely part placements from the search space, the set of possible locations is narrowed so that the computation of only few part-to-image comparisons is performed. This method gives a ten-fold speed-up over the standard dynamic programming approach.

The task of establishing the correspondence between pixels in two images (finding a markup) with human faces, addressed in [9], is effectively solved by building "cascades" of markups. The resolution is decreased two times in both initial images per cascade and new markup

for them is built. After that the starting approximation for initial markup is defined based on the new markup and the field of motion is searched but with less quantity of markings. The algorithm that solve the task utilizing one "cascade" runs eight times faster while preserving accuracy in finding the field of motion for two images.

As we can see, several fixed decreased resolutions relative to resolution of original image are taken and transition rules are introduced but the best resolution for processing particular image or its part is never estimated.

Proposed Algorithm

Statement of the problem: there exists an unknown function $y=f(x)$ with domain bounded to $[a, b]$. The image of this function is observed on $[a, b]$. The resolution needed for analyzing the image of this function is unknown. Under the assumption that the given image of a function represents an unknown piecewise smooth function, the boundaries of partial segments $a=t_0<t_1<\ldots< t_N=b$ and their number $N+1$ should be found.

Analytical solution of the segmentation problem stated above should be considered as finding the points of discontinuity for unknown piecewise smooth function. The following discontinuities are of interest: jump discontinuities, when the ε-neighbourhood of the function is empty in a given point and removable discontinuities when the first order derivative of function does not exist in the given point (jump discontinuity of function gradient). However only the image of unknown piecewise smooth function is observed so we are allowed to consider only the discrete analog of discontinuities in the form of irregular points on experimental curve.

Preliminary stage consists of presenting the experimental data as I "reference-value" pairs $\{i, x_i\}$; $i=1, 2 \ldots I$, that corresponds to maximum resolution. Acquisition of coarse resolution signal is performed (as well as in visual system) using a source signal with maximum resolution. The following initial conditions are set in algorithm that implements this method:

1. Maximum neighbourhood size of an arbitrary reference i for the coarsest resolution is taken as $s_0 \sim I/10$.

2. List of M resolutions is used with neighbourhood sizes of s_0, s_1, s_2, ..., s_m, ..., s_M; $sm+1=ksm$. In this case the value of k is 0.67. The M-th list item corresponds to source sequence with maximum resolution. M can be calculated based on s_0 and k.

3. The total number of values in the sequence corresponding to resolution s_m is $N_m=3I/s_m$ considering the mutual intersection of adjacent samples neighbourhoods.

4. $g(s_m)$ is the curve for resolution sm. Its n-th value $g_n(s_m)$ is calculated as median of values x_i, x_s for the sequence of references i, i+1, ... , $i+s_m$ in experimental curve.

5. The points of $g(s_m)$ being calculated with overlapping neighbourhoods are considered. Irregular points $r_n(t_m)$, ($t_m=1$, 2, ..., T_m) are fixed based on analysis of curves $g(s_m)$:$|g_n(s_m) g_n 1(s_m)|>d$. Here d is some predefined threshold. If the experimental curve has jump discontinuity in some dot then starting from some resolution the discontinuity will be found that will not disappear while dot neighbourhood shrinks.

6. Application of rules from item 5 to curves g(s) for all M resolutions results in a list of irregular points sequences r(T). Each sequence includes all irregular points $r(T_m)$ of curves $g(s_m)$ for a given resolution sm.

7. Irregular point $r_n(t_m+1)$ from sequence m+1 corresponds to irregular point $r_n(t_m)$ from sequence m respectively if $r_n(t_m) \leq r_n(t_m+1) \leq r_n(t_m) + s_m$ is true.

8. Sequences m and m+1 are grouped together in a sublist if they have equal number of irregular points T_m and the correspondence condition for each pair of respective irregular points is fulfilled. If sequence m already belongs to another sublist then sequence m+1 is added to that same sublist.

The result of segmentation is considered as sequence of irregular points with the largest resolution number m being taken from the longest sublist.

Results

Proposed algorithm was experimentally checked in MATLAB development environment. The graphs of line brightness in grayscale images were selected as objects of concern (Figure 6a and 6b). X-axis on all graphs of Figure 6 contained references of original signal, Y-axis on Figure 6a and 6b contained values of brightness across corresponding lines of images while Y-axis on Figure 6a.1 and 6b.1 contained resolution numbers in a list of resolutions that were used for analysis of experimental curve. Each horizontal line on Figure 6a.1 and 6b.1 corresponded to intervals in the space of fine references where jump discontinuities were found. Solid lines marked jumps from lower to higher values, dashed lines marked jumps from higher to lower values. Bold vertical lines on both sides of Figure 6a.1 and 6b.1 marked the sublists formed on step 8 of algorithm. Black color for that lines is used to mark out the sublist that considered as the result of segmentation. The information about jumps of experimental curve obtained on low resolutions (Figure 6.b.1) allowed discarding the regions of experimental curve that contained jumps on fine resolution discovered because of noise.

In contrast to other known methods of segmentation proposed algorithm doesn't use any a priori information about noise level etc. for the graph been processed.

With some necessary modifications, the algorithm was also successfully implemented in the application for cardiac signal segmentation and was tested on over 100 samples. Results of segmentation for 90-seconds long cardiac signal are shown on Figures 7 and 8.

Usually a so-called R-wave is used for automatic separation (segmentation) of cardiac signal into cardiac cycles and usually its amplitude is larger than amplitudes of other waves in the cycle. This assumption can't be applied to example curve shown above due to interference that introduces drift of zero base from cycle to cycle. Proposed method based on variable resolution concept allowed solving the problem of cardio signal segmentation in this case.
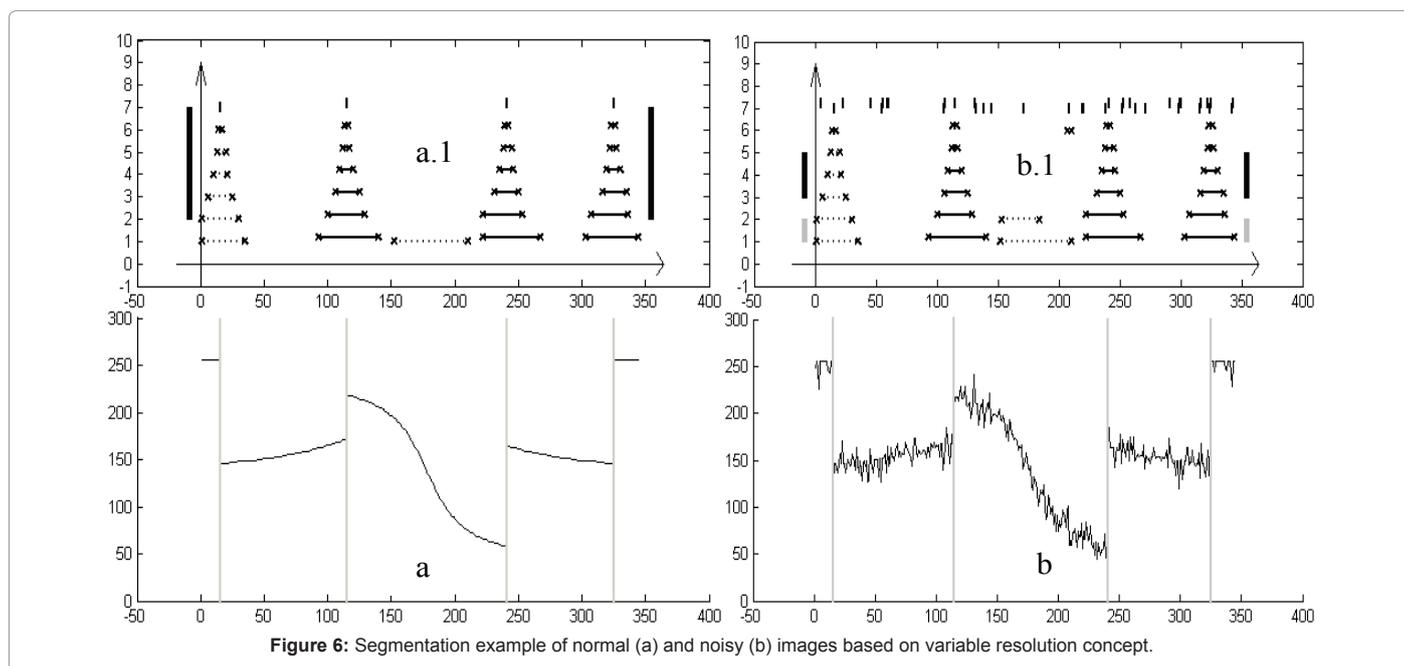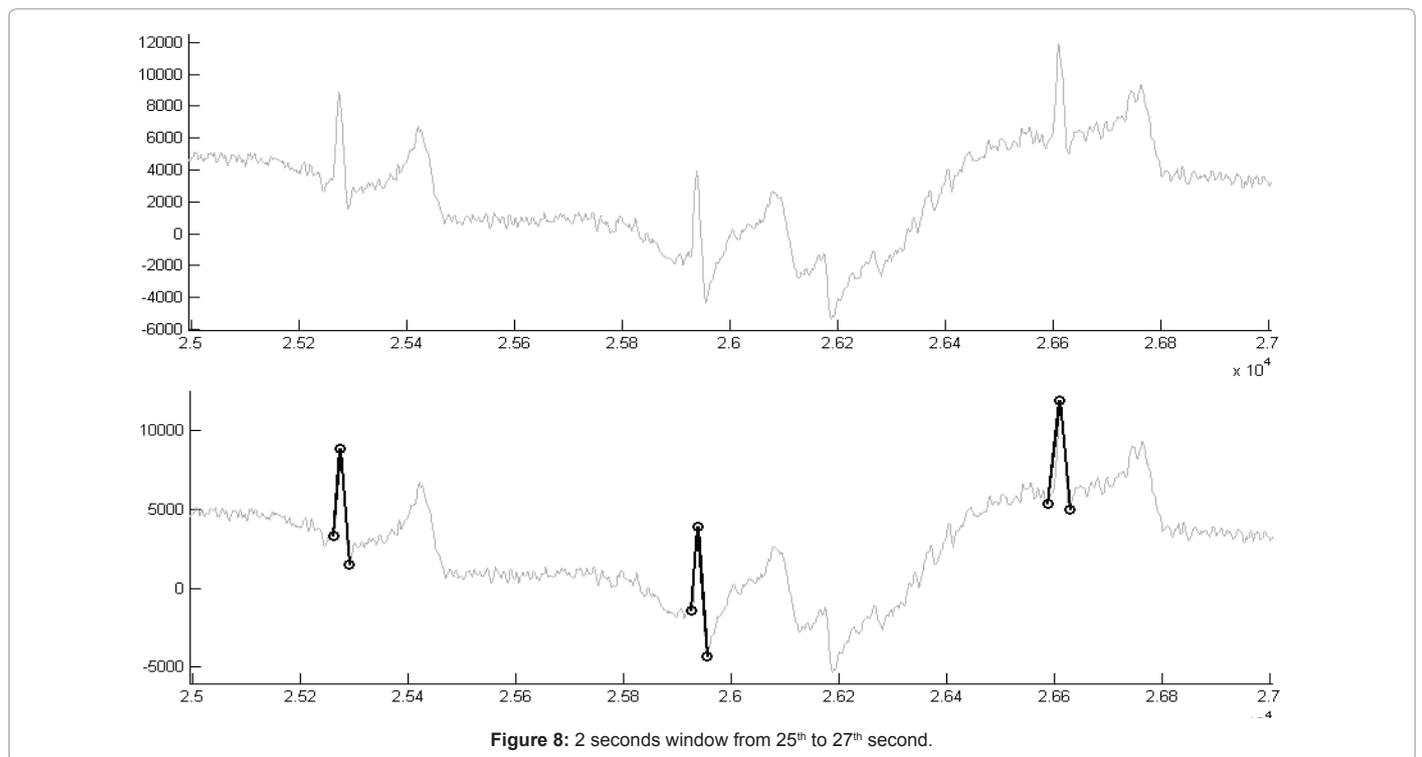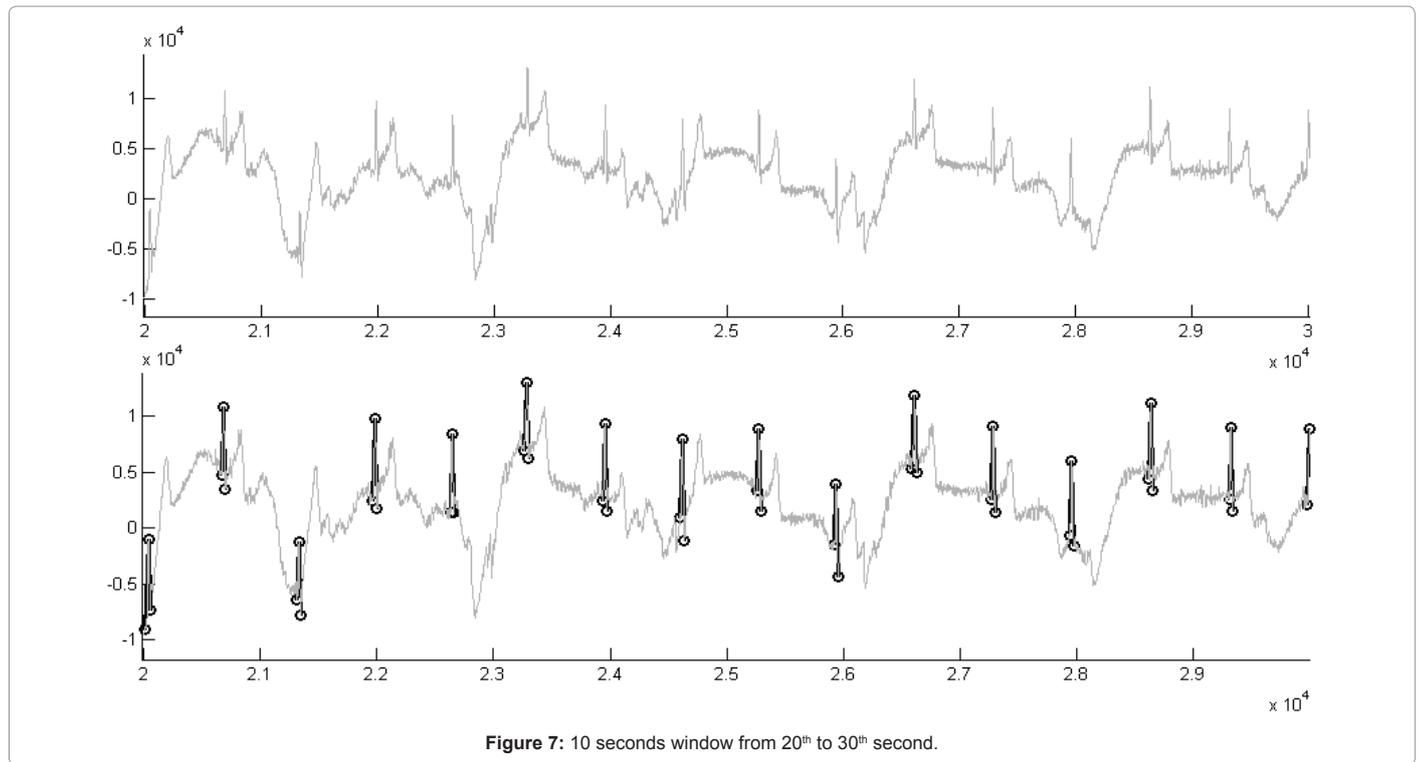

Figure 6: Segmentation example of normal (a) and noisy (b) images based on variable resolution concept.

Figure 7: 10 seconds window from 20th to 30th second.

Figure 8: 2 seconds window from 25th to 27th second.

Conclusion

For the first time, it was proposed and experimentally verified curve segmentation algorithm that used variable resolution (coarse-to-fine) in the decision-making process. Similar principle was found and studied in the visual system of animals [4,10]. Using the "coarse-to-fine" principle it is possible to successfully execute segmentation of experimental curves distorted by noise, which are the implementations of piecewise smooth functions.

It was shown that proposed algorithm was capable of obtaining results of segmentation based on artificially acquired information about

graph resolution. No additional a priori information about noise level was used while processing graphs distorted by noise.

These solutions will be used in the development of new methods for processing halftone images.

References

1. Vishnevskey V, Kalmykov V, Romanenko T (2008) Approximation of experimental data by Bezier curves. Inter J Info Theor Appl. 15:235-239.

2. Romanenko T, Vishnevskey V, Kalmykov V (2013) Analytical Representation of Graphs by Means of Parametrically Defined Splines. Proceedings of the international conference on applications of information and communication technology and statistics in economy and education ICAICTSEE. UNVE, Sofia, Bulgaria, pp: 536-542.

3. Hubel DH (1988) Eye, brain and vision. New York: Scientific American Library. Distributed by WH Freeman.

4. Ruksenas O, Bulatov A, Heggelund P (2007) Dynamics of Spatial Resolution of Single Units in the Lateral Geniculate Nucleus of Cat During Brief Visual Stimulation. J Neurophysiol 97: 1445-1456.

5. Canny JF (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Machine Intell 8: 679-698.

6. Sharypanov A, Antoniouk A, Kalmykov V (2014) Joint study of visual perception mechanism and computer vision systems that use coarse-to-fine approach for data processing. Inter J Info Theor Appl 1: 287-300.

7. Forster B, Van de VD, Berent J, Sage D, Unser M (2004) Complex Wavelets for Extended Depth-of-Field: A New Method for the Fusion of Multichannel Microscopy Images. Microsc Res Tech 65: 33-42.

8. Pedersoli M, Vedaldi A, Gonzalez J (2011) A Coarse-to-fine approach for fast deformable object detection, pp: 1353-1360.

9. Tyshchenko MA (2012) 3D reconstruction of human face in person identification problems. PhD Thesis. International Research and Training Center for Information Technologies and Systems.

10. NF (1979) Providing Dynamic properties of neural structures of the visual system. Leningrad, Nauka, p: 158.