Bibliography

[Леонтьев, 1972] Леонтьев В.В., Форд Д. Межотраслевой анализ воздействия структуры экономики на окружающую среду// Экономика и математические методы.- 1972.-Т.VII.-Вып.3.-С.370-400.

[Гасс, 1961] Гасс С. Линейное программирование. Физматгиз,-1961.

[Волошин, 1987] Волошин А.Ф. Метод локализации области оптимума в задачах математического программирования // Докл. АН СССР. - 1987. -293, N 3.- С. 549-553.

[Орловский, 1981] Орловский С.А Принятие решения при нечёткой исходной информации.- М:, Наука,- 1981,- 206с.

[Волошин,1993] Волошин А.Ф. Войналович В.М., Кудин В.И. Предоптимизационые и оптимизационные схемы сокращения размерности задачи линейного программирования // Автоматика,N4, 1993.

[Войналович, 1987] Волкович В.Л., Войналович В.М., Кудин В.И. Релаксационная схема строчного симплекс метода // Автоматика.- 1987. -N4.-С. 79-86.

[Войналович, 1988] Волкович В.Л., Войналович В.М., Кудин В.И. Релаксационная схема двойственного строчного симплекс метода // Автоматика.-1988. -N 1,-С.39-46.

[Кудин, 2002] Кудин В.И. Применение метода базисных матриц при исследовании свойств линейной системы // Вестник Киевского университета. Серия физ.-мат. науки. - 2002.-2., С. 56-61.

Authors' information

Kudin V.I. – National Taras Shevchenko University of Kiev, Cybernetics department, Ph. D. Kiev, Ukraine. E-mail: V_I_Kudin@mail.ru

Kudin G.I. – National Taras Shevchenko University of Kiev, Cybernetics department, Ph. D.i Kiev, Ukraine. E-mail: kuding@mail.univ.kiev.ua.

Voloshin O.F. – National Taras Shevchenko University of Kiev, Cybernetics department, professor. Kiev, Ukraine. E-mail: ovoloshin@unicyb.kiev.ua,

APPROXIMATION OF EXPERIMENTAL DATA BY BEZIER CURVES

Vitaliy Vishnevskey, Vladimir Kalmykov, Tatyana Romanenko

Abstract: Very often the experimental data are the realization of the process, fully determined by some unknown function, being distorted by hindrances. Treatment and experimental data analysis are substantially facilitated, if these data to represent as analytical expression. The experimental data processing algorithm and the example of using this algorithm for spectrographic analysis of oncologic preparations of blood is represented in this article.

Keywords: graphics, experimental data, Besie's curves

ACM Classification Keywords: I.4 Image processing and computer vision - Approximate methods

Introduction

The experimental data, as a rule, represent the distorted by hindrances certain process fully determined by some unknown function *y= f(x)*, and distorted by hindrances. In most cases experimental data are represented as the graphical curves. The graphs, i.e. graphical curves is, apparently, the simplest and a long ago in-use means of cognitive presentation of experimental data in the most different scopes of human activity which allow to estimate evidently the qualitative property of the process, in spite of hindrances, measurement errors. Graphs displaying the same process, description of certain object can substantially differ

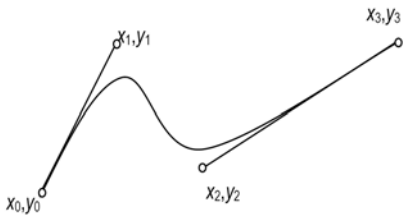

Fig. 1. Bezier curve

from each other by scales, amount of the used observations, level noises, and so on. At the same time the form feature of the graphical curve characterizes the parameters of the displayed object or process.

Automatic (automated) graph (graphical curve) processing supposes comparison of their forms to set, whether the different graphs characterize the same or different processes or objects. The solving of this task, using the neuronet methods or statistical visual patterns recognition methods, is complicated because of every graph can differ from other by scales, noises level, amount of the used measurements. At the same time the method [Kussul, 2004] of neuronet recognition of contours of the images represented with the Bezier curves is known. Using Bezier curves, it is possible to provide the invariance of descriptions of image contours in relation to position of images in eyeshot, scale, turn. So, the experimental data – graphical curves may be replaced with the analytical curves that are congruent enough to the initial graph form, and, at the same time, invariant relatively to the scale change, amount of measurements, noises level. Such curves may be chosen as the polynomial functions $x(t)$, $y(t)$ of n power of the parameter t :

$$\begin{aligned} x(t) &= a_0 t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n, \\ y(t) &= b_0 t^n + b_1 t^{n-1} + \ldots + b_{n-1} t + b_n. \end{aligned} \tag{1}$$

Choice of n – the power of polynomials depends on complication of the experimental curves to be approximated.

If such approach would be realized it is possible to hold all the substantial form features of the experimental curve and to eliminate the hindrances influencing. In addition, instead of the graph description in the space of initial signals it is possible to operate approximating polynomials in the space of their coefficients. The substantial advantage of such description is its invariance relatively the amount of the measured signals in every graph.

However, if the direct task to build the curve, having the set coefficients of polynomials is trivial, the reverse task – the set coefficients definition of approximating polynomials having the points of the graph – does not have the simple decision [Vishnevskey, 2004]. One of this task decisions is proposed at the article.

Formal problem definition

The graph can be represented by a polyline, the kinks of which form the sequence *of N* experimental values of some function *y(x):* $y_Э(x)=\{y(x_0),y(x_1),y(x_2),\ldots, y(x_N)\}$, at the predetermined values $x_0,x_1,x_2,\ldots, x_N$. Let us consider the Bezier curve as an approximating curve because of the most often using in practical applications, parametrically defined polynomial of the third degree.

Under the Bezier curve we understand the arc of flat curve of the third degree [Deniskin, 1999]:

$$Bz(t) = \sum_{j=0}^{3} Br_j(t) \cdot Q_j \tag{2}$$

where $Br_j(t)$ are base Bernstein polynomials of the third degree,

Q_j are the coefficients of curve

t is an real parameter which changes in an interval [0,1].

Bernstein polynomials are defined as:

$$Br_j(t) = C_3^j \cdot (1-t)^{3-j} t^j, \quad 0 \le j \le 3 \tag{3}$$

where $C_3^j, \ 0 \le j \le 3$ are binomials coefficients

$$C_3^j = \frac{3!}{j! \cdot (3-j)!} \tag{4}$$

Thus, the Bezier curve (fig. 1) can be set by two polynomials of the third degree, depending on the parameter t:

$$\begin{aligned} X(t) &= x_0 \cdot (1-t)^3 + x_1 \cdot (1-t)^2 \cdot t + x_2 \cdot (1-t) \cdot t^2 + x_3 \cdot t^3 \\ Y(t) &= y_0 \cdot (1-t)^3 + y_1 \cdot (1-t)^2 \cdot t + y_2 \cdot (1-t) \cdot t^2 + y_3 \cdot t^3 \end{aligned} \tag{5}$$

The experimental data approximating as Bezier curve has many of advantages. Describing the wide enough class of curves, Bezier curves use reasonable quantity of coefficients - 8. Coefficients (x_0,y_0), (x_3,y_3) are the co-ordinates of initial and eventual points of approximating curve arc, accordingly, and coefficients (x_1,y_1), (x_2,y_2) - co-ordinates of points which control its form (so-called "whiskers"). Changing coefficients or moving points proper

to them on the screen of monitor (both initial and eventual and controls), an user can edit the form of approximating curve in the interactive mode, improving the results of automatic approximation.

Thus, every graph may be represented as a vector $v = \{ x_0,y_0, x_1,y_1, x_2,y_2, x_3,y_3 \}$, and its components are the coefficients of Bezier curve equations. The graph description as the vector *v,* always has the same amount of components regardless of number of measured experimental points, that makes solving of different tasks, in particular, tasks of the graphs recognition more simple.

It was before noticed, that determination of coefficients of Bezier curve, most exactly approximating one or another graph, is not a trivial task. The iterative method to minimize the sum of distances squares between every point of the graph and the approximating curve is used for determination of Bezier curve coefficients, approximating some experimental sequence of points [2]. Such problem definition follows from presumption, that graphs under consideration are arbitrary, rather than just those which display some unknown function. It is assumed that the graph is the aggregate of unconnected, mutually independent points. To calculate the distance from every point of the graph to the approximating curve proper nearest point on the approximating curve must be found. This task also is iterative as the approximating curve is parametrically defined. When the quantity of points on the graph is large (more than 100), and the amount of iterations is not limited beforehand, such method of task solving resulted in substantial calculable difficulties.

It is proposed to accept the area of the $S(v,y(x))$ figure, limited by the contour, formed with the approximating and experimental curves as the measure of likeness of this experimental graph with the approximating curve. So the values of Bezier curve coefficients v_{opt}, which approximates certain experimental graph in the best way, correspond to the minimum value $S(v,y_Э(x))$:

$$v_{opt} = \arg\min_{v} S(v, y_Э(x)) \tag{6}$$

Algorithm

It is follows from formula 6, that computation of the area, limited by the contour $S(\boldsymbol{v},y_Э(x))$, is basic operation of algorithm, searching the optimum coefficient values v_{opt}. The direct calculation $S(\boldsymbol{v},y_Э(x))$ is not possible because of this contour is not simply connected.

Graph $y_Э(x)$ and Bezier curve intersect each other many times. It is difficult to compute the intersections co-ordinates because of Bezier curve is parametrically defined. The simplest decision seems to replace Bezier curve by the polyline *Vt*. The fracture points of the polyline (x_t,y_t) belong to that Bezier curve and are calculated for the $t = 0, \delta, 2\delta, 3\delta,...,1$. The value δ is chosen small enough, to represent initial curve by the polyline *Vt* exact enough for the practical calculations. Possibility to calculate for every value $y(x_n)$ the proper value $V_t(x_n)$ is the result of the accepted simplification. Then the area $S(v, y_Э(x))$ can be calculated as

$$S(v, y_Э(x)) = \sum_{n=0}^{N} \left| V_t(x_n) - y_Э(x_n) \right| \tag{7}$$

The value v_{opt}, proper to the minimum value $S(v,D(n))$, can be got using the gradient method.

Experimental verification of algorithm

The offered algorithm was experimentally tested more than on 1100 real graphs of spectrographs of the medical preparations of blood, got on the method of Oncotest [Vishnevskey, 2006]. The method of early diagnostics of oncology diseases allows defining patient's presence or absence of oncology disease by the use of spectrograms of medical preparations of blood (Fig.2). At present researches are conducted also and for determination of localization of disease when being it. On fig.3 the spectrograms of preparations of patient's blood are represented. Their diagnoses are the oncology disease of identical localization (stomach). In spite of obvious distinctions of spectrograms on the levels of signals, hindrances, form of curves have undoubted likeness. The form of curves is specific property of the resulted graphs, which can correspond to the diagnosis. The identification task to form a diagnosis by spectrograms supposes additional researches which must be based on treatment of plenty of experimental information. In the same queue, the processing of large data arrays is

impossible without the decision of tasks of their computer treatment, and including the tasks of automatic or automated classification of spectrograms on their form must be decided, using the methods of patterns recognition.

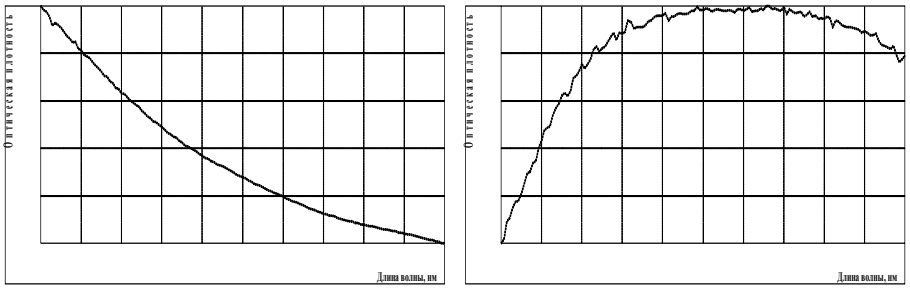

Fig.2. Examples of spectrograms of blood serum preparation:
left - spectrogram of healthy, right - sick man.

The direct use of statistical pattern recognition methods for the automated diagnosis forming in this case is not possible owing to large changeability of scale (dozens of one times), and, also, configuration of function of realizing the same appearance. Therefore it is appeared expedient to find some suitable Bezier curve for every graph, which polynomial coefficients, after adduction to one scale, can be used as pattern parameters, invariant to the changes of scale, amount of measures, noises level.

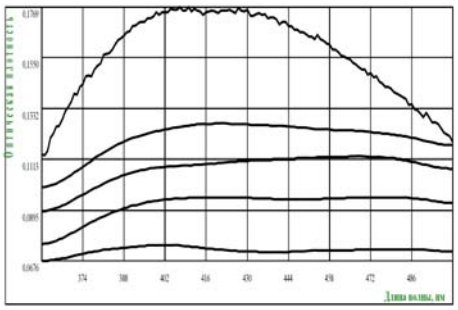

Fig.3 Examples of spectrograms of blood preparation of patients by the oncology disease of the same localization (stomach).

Examples of spectrograms and approximating them Bezier curves are represented on a fig. 4,5. A result was achieved in most cases for 15 - 40 iterations (in especially difficult cases - to 300). As it is obvious from figures, approximating the Bezier curves give the satisfactory approaching.

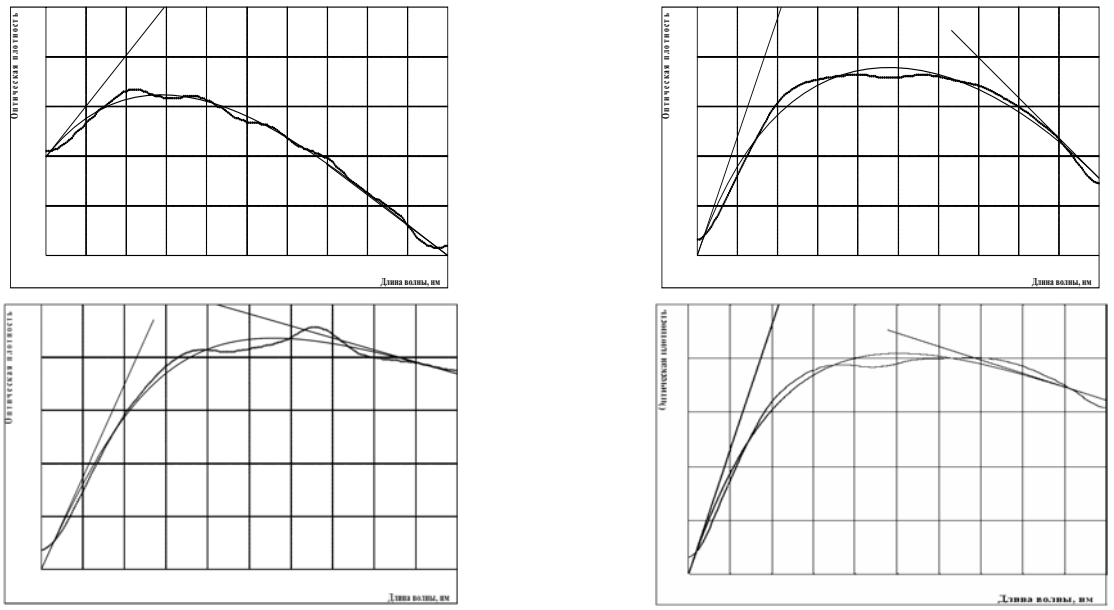

Fig.4. Spectrograms of blood preparations and their approximation by Bezier curves;
localization is a mammary gland

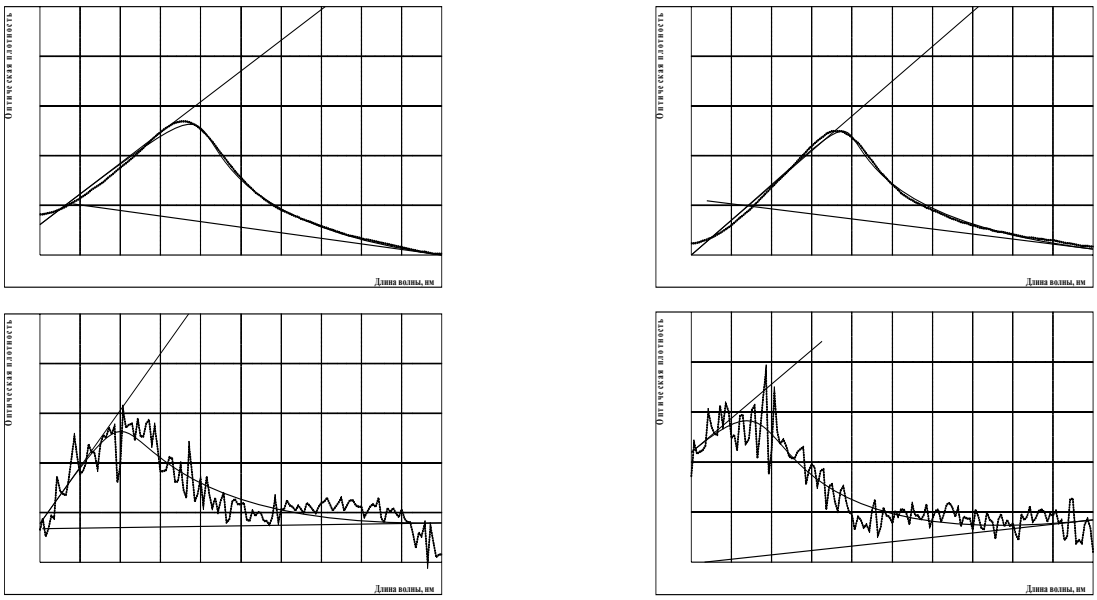

Fig.5. Spectrograms of blood preparations and their approximation by the Bezier curves; localization is a liver

Conclusion

The offered algorithm allows finding the parameters of Bezier curve, which approximates the experimental graph. The examples of approximation of spectrograms are resulted for implementation of the automated diagnostics. It will enable in the future to automatize the process of screening examination of population with the purpose of early detection of malignant tumors.

Bibliography

[Kussul, 2004] Куссуль М.Э. Кодирование контуров, представленных кривыми Безье, в задачах нейросетевой классификации / // Математические машины и системы. – 2004. – № 3. – С.17-30.

[[Vishnevskey, 2004] Вишневский В.В., Рысцов І.К., Волжева М.В Итерационный алгоритм построения кривой Безье по заданными точками //Математические машины и системы. - 2004 №4 - С. 108 – 116.

[Deniskin, 1999] Денискин Ю.И. Особенности аппроксимации обводов параметрическими полиномами в форме Бернштейна //Прикладная геометрия. - 1999. - вып. 2 № 2

[Vishnevskey, 2006] Вишневский В.В., Владимиров В.А., Романенко Т.Н. Программно-аппаратный комплекс "Онкотест-WM-01" // Український журнал телемедицини та медичної телематики. – 2006. – т. 4 №2 С.182-185

Authors' Information

Vitaliy Vishnevskey - the head of the department, senior research worker, doctor of sciences, Institute of Mathematical Mashines and Systems, 42 Academician Glushkov St., 03680, Kiev 187, Ukraine; e-mail: vit@immsp.kiev.ua

Vladimir Kalmykov - senior research worker, doctor of sciences, Institute of Mathematical Mashines and Systems, 42 Academician Glushkov St., 03680, Kiev 187, Ukraine; e-mail: kvg@immsp.kiev.ua

Tatyana Romanenko - junior research worker, Institute of Mathematical Mashines and Systems, 42 Academician Glushkov St., 03680, Kiev 187, Ukraine; e-mail: romanenko@immsp.kiev.ua