Information Sciences 179 (2009) 1483–1494

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Automatic knot adjustment using an artificial immune system for B-spline curve approximation

Erkan Ülker *, Ahmet Arslan

Selçuk University, Engineering and Architecture Faculty, Department of Computer Engineering, 42075 Konya, Turkey

ARTICLE INFO

Article history:
Received 27 July 2007
Accepted 25 November 2008

Keywords:
Artificial immune system
B-spline curve approximation
Clonal selection
Knot adjustment

ABSTRACT

Reverse engineering transforms real parts into engineering concepts or models. First, sampled points are mapped from the object's surface by using tools such as laser scanners or cameras. Then, the sampled points are fitted to a free-form surface or a standard shape by using one of the geometric modeling techniques. The curves on the surface have to be modeled before surface modeling. In order to obtain a good B-spline curve model from large data, the knots are usually respected as variables. A curve is then modeled as a continuous, nonlinear and multivariate optimization problem with many local optima. For this reason it is very difficult to reach a global optimum. In this paper, we convert the original problem into a discrete combinatorial optimization problem like in Yoshimoto et al. [F. Yoshimoto, M. Moriyama, T. Harada, Automatic knot placement by a genetic algorithm for data fitting with a spline, in: Proceedings of the International Conference on Shape Modeling and Applications, IEEE Computer Society Press, 1999, pp. 162–169] and Sarfraz et al. [M. Sarfraz, S.A. Raza, Capturing outline of fonts using genetic algorithm and splines, in: Fifth International Conference on Information Visualisation (IV'01), 2001, pp. 738–743]. Then, we suggest a new method that solves the converted problem by artificial immune systems. We think the candidates of the locations of knots as antibodies. We define the affinity measure benefit from Akaike's Information Criterion (AIC). The proposed method determines the appropriate location of knots automatically and simultaneously. Furthermore, we do not need any subjective parameter or good population of initial location of knots for a good iterative search. Some examples are also given to demonstrate the efficiency and effectiveness of our method.

1. Introduction

The problem of recovering the 2D/3D shape of a curve/surface, also known as *curve or surface reconstruction*, has received much attention in the last few years. In literature, there are two different orientations [24]. In first orientation the authors address the problem of obtaining the curve or surface model from a set of given cross-section or points. This is a typical problem in most research and application areas such as CAD/CAM, biomedical and medical science, in which usually an object (acquired from 3D laser scanning, ultrasound imaging, magnetic resonance imaging, computer tomography, etc.) is defined as a sequence of 2D cross-sections. The other different approaches, in the second orientation, include the reconstructing curves/surfaces from a given set of data points. Two different approaches are employed depend on the nature of these data points: interpolation and approximation. In the fitting form by interpolation, parametric curve is constrained to pass through all the given set of data points. The technique of fitting of data parametrically by interpolation is suitable when the data

* Corresponding author.
E-mail addresses: eulker@selcuk.edu.tr (E. Ülker), ahmetarslan@selcuk.edu.tr (A. Arslan).

0020-0255/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2008.11.037

points describing the curve are sufficiently accurate and smooth. In the data fitting form by approximation, parametric curve passes close to the given set of data points, but not necessarily through them. By the help of a few distance criteria, the curve is approximated but it is not ensuring the protection of whole shape of the real curve that is defined by given data points set. The distance (between fitting curve and given data points) is usually measured along a normal to fitting curve or along a coordinate. Approximation techniques are especially recommended when data is not exact or noisy. The other important reason for choosing the approximation technique is great computational effort required to obtain surfaces by interpolating an infinite number of data like curve forms. Approximated curve/function minimizes the total error.

A second problem emerges when the input data presented to the approximation method is irregularly spaced (randomly scattered set of points without any predefined sequence or order). Unfortunately, all standard interpolation and approximation techniques such as Hermit interpolation, Bezier curves or B-spline curves need a sequence of points. For this reason if you want to apply these methods to scattered data, you have to order the data. More research on ordering of dense and noisy points can be found in Cheng [5], Goshtaby [18], Cohen et al. [6], and Yu [41].

Furthermore, B-spline curve or surface fitting needs a very good parameterization model [26]. Although numerous work has been done on selection of parameters, especially when the points are irregularly spaced and lying on a complex base curve or surface, it is difficult to calculate better parameters with the contemporary approaches [26]. Determination of good knots is the key for successfully fit a spline to the data. For this reason it must be solved as a multivariable and multimodal nonlinear optimization problem [39]. A good model is the one with the smallest possible number of parameters and with the least difference between the model and underlying data function. For the large data with this problem, attending the optimization algorithms that are abstain from possible local optima and at the same time getting the desired solution in iterative way.

The obvious solution to these problems is to consider an approximation scheme that allows us a significant saving of time and memory. In this approach, the goal of the curve/surface reconstruction methods can be defined as follows: *given a set of sample points X assumed to lie on an unknown curve/surfaceU, construct a curve/surface model S that approximates U*. Of course, this implies giving a tolerance error without sacrificing the "quality of the shape". Where three criterias are important [28]: "smoothness" of the curve/surface, regularity of polygon and the distance from obtain curve/surface to given points set. This problem was analyzed from different point of views such as parametric methods, function reconstruction, implicit surfaces, etc. [24]. The latest work in literature that is parallel to our goal in this study is Li et al. [29] on adaptive knot placement considering the discrete curvature of the points.

One of the most attractive and promising artificial intelligence approaches to this problem is based on neural networks. The first study in curve design that uses neural networks is given in Chen et al. [4] and Hoffmann and Várady [23]. Later, these studies were extended to surface design with use of Kohonen networks [19,20,22,3], self-organized maps [2,27] and functional networks [24]. The last study about the artificial neural networks is about numerical control of Kohonen neural networks [21]. Also many evolutionary optimization techniques such as genetic algorithm [39,35,34,26], simulated annealing, and simulated evolution [36] are successfully implementing to this problem. Many solutions are searched for this problem with traditional optimizations techniques such as modified continuous reactive tabu search [40], multi-objective optimization [17] and an error-bounded method based on Dominant Point selection [30]. Goldenthal and Bercovier [16,33] adopted a strategy that is used a cost function usage approach to design the curve shape and it is proved that using a genetic algorithm is an efficient way to solve the cost function minimizing the cost function.

As result of the scientist are inspired by nature to solve the complex problems, at human systems modeling trend beginning with the artificial neural networks, after the evolutionary algorithms based on the Darwin's evolution theory, artificial immune systems are generated by the human immune systems modeling. Human immune system is very powerful, adaptive, distributed, provides memory, self-organized, strong ability to pattern recognition and has a evolutionary structure to give reaction against foreign factors [8]. Scientist, engineers, mathematicians, philosophers and other researches are interested in capabilities of this system. The result of this, research areas which implementing the immunity principles of human are begin to growing. This area is called as artificial immune systems or immunological computation.

GA and AIS have similar components, however there are proven advantages of AIS over GA [13]. In our view, AIS's most important quality is the convergence speed to the objective function which makes it a better choice than GA. To our best knowledge, at the time of writing, there is no study in the literature that addresses the above problem using AIS. In this paper, we aimed to automatically determine the locations and counts of interior knots. In order to fit the given set of 2D points by B-spline curves, artificial immune system methodology was applied.

The organization of the paper is as follows: in the next section the B-spline curve fitting problem is explained. Then, the details of artificial immune system are given. After explained the principles of our proposed automatic knot placement method with AIS, we show the efficiency of our proposed method with some numerical results. The paper is ending with conclusions and future works.

2. Description of the problem

2.1. B-spline curves

B-splines research started back in 1940s. But it did not become popular in the industry until de Boor [7] published his study. Given control points $\{P_i\}$ $(i = 0,1,2,\ldots,n)$, a kth-order (or degree is d, and $k = d + 1$) B-spline curve is defined by

$$P(u) = \sum_{i=0}^{n} p_i N_{i,d}(u). \tag{1}$$

Cox–de Boor recursive function is the core of B-spline curves and basis functions of B-spline curves are calculated by using Cox–de Boor functions. $T = \{t_0, t_1, \ldots, t_{r-1}, t_r\}$, is a nondecreasing sequence of real numbers known as *knots*. T is called the *knot vector*. The *ith B-spline basis function* $N_{i,k}(s)$ of order k can be defined as a recursive function as follows:

$$\begin{aligned} N_{i,1}(u) &= \begin{cases} 1, & t_i \leqslant u \leqslant t_{i+1}, \\ 0, & otherwise, \end{cases} \\ N_{i,k}(u) &= \frac{u - t_i}{t_{i+k-1} - t_i} N_{i,k-1}(u) + \frac{t_{i+k} - u}{t_{i+k} - t_{i+1}} N_{i+1,k-1}(u). \end{aligned} \tag{2}$$

A more detailed discussion about B-spline curves can be found in [32].

2.2. Data fitting with B-spline

In curve reconstruction problems, the input is a set of unorganized points. Therefore the order, the knots and the control points are all unknowns. For this, parametric curve/surface fitting to irregular data can be formulated as a nonlinear least-squares problem [25]. The overall aim of reconstruction methods is to determine these values where the basic strategy is the following [37]:

1. Order of curve (k), control points number (n) and knot values t_j are given constant. Fix the order (k), the number of control points (n) and the knot values t_j.
2. Assigned a parameter value $C(t_j)$ to each scattered $\mathbf{F_j}$ points.
3. System $C(t_j) = F_r$ solved or minimized $\sum_j \|C(t_j) - F_j\|^2$.

The crucial point of this strategy from the view of B-spline curves, frequently the second step frequently referred as the parametrization of the given data. Parameterization is the way of how to assign parameter values to each point. Here normally different restrictions and assumptions are introduced. Someones, can try to consider the assigned values as unknown parameters in an optimization problem, but for very large amount of data this approach, leads to a unknown complex nonlinear system with some unknowns. For detailed explanations about control point adjustment problem for parametric B-spline curve approximation problem to a sequence of ordered dense data, the reader is referred to Yang et al. [38].

Within the framework of the strategy above, we will explain our method for the B-spline curve approximation. In this paper, for converting the original problem into a discrete combinatorial optimization problem, we benefit from mathematical basics in Yoshimoto et al. [39], Safraz et al. [35] and Raza's studies [34]. Then we propose a new method that solves the converted problem by an artificial immune system.

Let us assume that the data to be fitting are given on the close interval $[a,b]$ in x-axis. We can write an equation like below

$$F_j = f(x_j) + \Delta_j, \quad (j = 1, 2, \ldots, N). \tag{3}$$

In this equation $f(x)$ is the underlying function of the data and Δ_j is a measurement error. Let t_i ($i = 1 - m, 2 - m, \ldots, n + m$) be knots of a spline for data fitting, where n is the number of knots t_i ($i = 1,2,\ldots,n$) settled in the interval (a,b), m is the order of a B-spline $N_{m,i}(x)$. At the ends of the interval $[a,b]$, knots are set to a and b

$$\left\{ \begin{aligned} a &= t_{1-m} = \ldots t_0, \\ b &= t_{n+1} = \ldots t_{n+m}. \end{aligned} \right\}. \tag{4}$$

At this adjustment a model function for $f(x)$ is given as follow

$$S(x) = \sum_{i=1}^{n+m} c_i N_{m,i}(x), \tag{5}$$

where c_i is a B-spline coefficient

B-splines in Eq. (3) are calculating with Eq. (2). In Eq. (2), the knots are not only at the enumerator of fraction at the same time it is at the denominator of fraction. For his reason a spline given by Eq. (5) is a nonlinear function of knots. Eq. (5) is fitting to data given by the Eq. (3) by least-square method. Then sum of the squares of remainders Q_1 is written as follows

$$Q_1 = \sum_{j=1}^{N} w_j \{S(x_j) - F_j\}^2, \tag{6}$$

where w_j is the weight of data and $N > n + m$. In our study all weight values are equal 1. The subscript of Q means the dimension of data. B-spline Coefficients c_i ($i = 1,2,\ldots,n + m$) are defined by the condition of minimizing Eq. (6). To find a good model, together the location and number of interior knots t_i ($i = 1,2,\ldots,n$) are defined as certain as possible. In this study,

centripetal method is used to assign a parameter value to each point in given points set and to determine the knots on the vector for selection of a suitable knots vector.

As explained above, objective function to be minimized is Eq. (6) and its variables are the B-spline coefficients and interior knots [39,35,34]. B-spline coefficients are linear parameters. Nonetheless the interior knots are nonlinear parameters, because the function $S(x)$ is a nonlinear knots function. This minimization problem is known as multimodal optimization problem [17].

3. Artificial immune systems

Human immune system can be regarded as a complex network structure protecting against countless different sickness causing microorganisms such as virus, bacteria, fungus and parasite. It is a parallel and distributed adaptive system that has the potential to perform very intelligent operations. In addition, the following properties of *AIS* have attracted the researchers [9]: Uniqueness, Detection of Foreign Effect, Finding Irregularity Areas, Distributed Detection, Noise Tolerance, Imperfect Detection, Reinforcement Learning and Memory. The modeling of such a structure can be used as a new problem solving approach. This approach is called artificial immune system (*AIS*). By the intelligent computation technique, artificial immune systems were used in very diverse areas such as pattern recognition, data analysis, classification, learning, error detection, optimization, memory allocation, robotic and computer security [10,11].

We need the following components to construct an artificial immune system: (i) the representation of system parts, (ii) a mechanism to compute the interaction of system parts with each other and with the environment, (iii) adoption procedures. More detailed explanations are given below. Further information about AIS can be found in [9–12].

3.1. The representation of system parts (antigen and antibody)

The most popularly accepted and used approach for antigen and antibody representation of system parts, is Perelson and Oster's [31] shape space approach proposed in 1979. Antibody–Antigen notation partially decides the distance measurement which is used for calculating interaction order. Mathematically, generalized shape of a molecule is an antigen or an antibody. We assume that; m represents an antigen or antibody, m is a vector ($m = \{m_1, m_1, \ldots, m_L\}$) represented in a space with L dimensions and $m \in S^L$, where S is shape space. Complementary feature of human immunity is modeled with distance concept in shape space. In the shape space notation, antigens are commonly shown after that they have been reversed. There is shape space notation of two antibodies and one antigen in Fig. 1a. In Fig. 1b, complementary notation of antigen A in Fig. 1a is shown. A name is given to shape space according to distance criteria used for interactions. When Euclidean distance is used, it is called Euclidian shape space; when Manhattan distance used, it is called Manhattan shape space and when Hamming distance used, it is called Hamming shape space [12].

3.2. The mechanism to compute the interaction (affinity measure)

A mechanism must be used to compute the interactions. The interactions among antigen and antibody can be modeled by an affinity measure. Affinity between an antigen and an antibody depends on relative distance which can be estimated by distance measure between two arrays (or vector). As the distance between two cells in the immune system represented as two points in the shape space gets larger, the complementary feature between these cells increases. For example, in Fig. 1, two kinds of physicochemistry interactions are accepted between antigen A and antibody B. So, there are two axes representing the interactions and they are depicted in this figure. There are two theorems in interactions; maximum distance for maximum interaction and minimum distance for maximum interaction. First theorem is valid for Fig. 1a, and second theorem is valid for Fig. 1b. Therefore, in Fig. 1a since the distance between antigen A and antibody C is longer than the distance between antigen A and antibody B, the intensity of interaction between antigen A and antibody C is more than the one between antigen A and antibody B. In Fig. 1b, since the distance between antigen A and antibody C is shorter than the distance between antigen A and antibody B, the intensity of interaction between antigen A and antibody C is more then the one between antigen A and antibody B. The distance between antigen and antibody can be calculated using different methods. If the vectors, m, signifying antigen and antibody are real values vectors then Manhattan or Euclidian distance measures can be used. The Euclidian distance between antibody $Ab = \{Ab_1, Ab_2, \ldots, Ab_L\}$ and antigen $Ag = \{Ag_1, Ag_2, \ldots, Ag_L\}$ calculated by Eq. (7), and Manhattan distance calculated by Eq. (8)

$$D = \sqrt{\sum_{i=1}^{L} (Ab_i - Ag_i)^2}, \tag{7}$$

$$D = \sqrt{\sum_{i=1}^{L} |(Ab_i - Ag_i)|}. \tag{8}$$

If antigens and antibodies are declared by binary symbols instead of real value vectors, then the Hamming distance is used. In this case, the following factors become important: The coverage of the antibody, recognition of all antigens for cal-

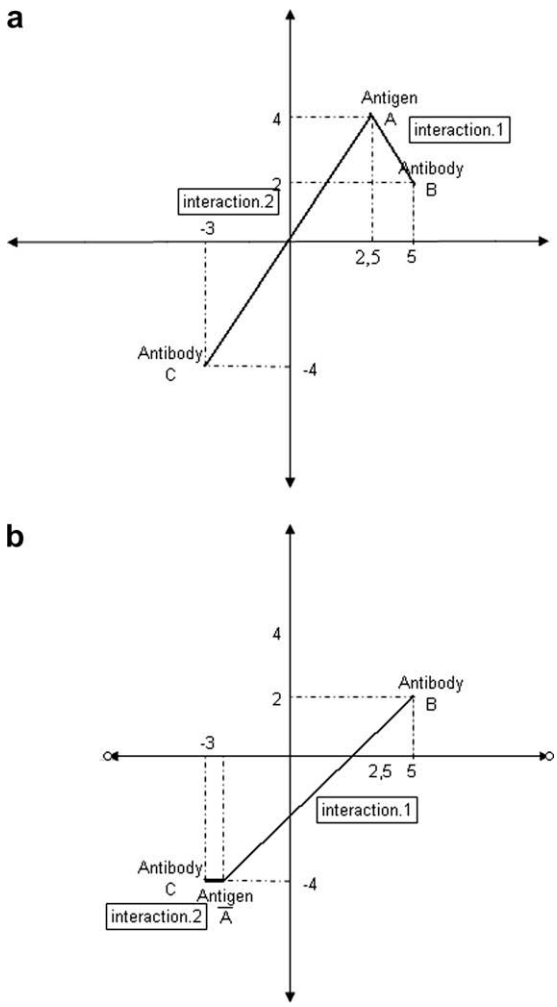

Fig. 1. (a) Shape space demonstration and (b) the complementary demonstration in shape space.

culation of the minimum number of antibody, binary distance estimation and evaluating the connection values. Detailed information on Hamming distance can be found in [12]. We have chosen the Euclidian distance in our study since it is widely used in the literature [9].

3.3. Adoption procedures (the Clonal Selection Algorithm)

Adaptation procedures model how system changes in time. These are constructed with an immune function, process or theory. Clonal Selection Algorithm [1], Negative Selection Algorithm [15] and immune nets [14] are the widely adopted ones. From among these, we decided that Clonal Selection Algorithm was the best for the purpose of our study. Therefore Clonal Selection Algorithm discussed in below.

In the immune system, many selection mechanisms are involved in affinity maturation process. De Castro et al. [10,11] proposed the Clonal Selection Algorithm based on the operations during affinity maturation process. They applied their algorithm on problems like character recognition and optimization and analyzed its performance. The algorithm has two basics: only the cells recognizing the antigen are selected for growing (mutation and cloning), the selected cell's affinity to the antigen is increased by affinity maturation process. The flow of this algorithm is given in Fig. 2.

Steps of the algorithm as in follows:

- A set (P) from candidate solutions is generated, subset of the memory cells (M) adding at the population.($P = P_R + M$)
- Determine (Select) the n best individuals of the population (P_n), based on an affinity measure;
- Reproduce (Clone) these n best individuals of the population, giving rise to a temporary population of clones (C). The clone size is an increasing function of the affinity with the antigen;

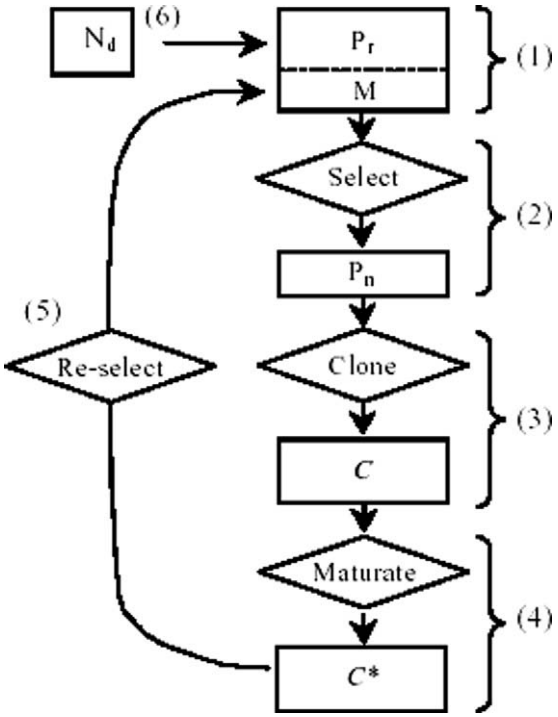

Fig. 2. The Clonal Selection Algorithm [10,11].

- Submit the population of clones to a hyper mutation scheme, where the hyper mutation is proportional to the affinity of the antibody with the antigen. A maturated antibody population is generated (C^*);
- Re-select the improved individuals from C^* to compose the memory set M. Some members of P can be replaced by other improved members of C^*;
- Replace d antibodies by novel ones (diversity introduction). The lower affinity cells have higher probabilities of being replaced.

This algorithm presents an evolutionary strategy suitable for complex machine learning tools. It usually improves the performance by decreasing the learning time and increasing the memory gain. The algorithm has successful applications in diverse areas such as binary character recognition, multimode and combinatorial optimization and immune network theory. It resembles the genetic algorithm but as De Castro et al. demonstrated, when compared with genetic algorithm, it can produce a set of many different local optimum solutions. Genetic algorithm on the other hand, tries to make the rest of the population similar to the best individual. The two algorithms are quite alike in their coding and computation method, but they are different by their search process, their notation and the order of operations.

4. Automatic knot adjustment by an artificial immune systems

The B-spline curve fitting problem is to produce a B-spline curve to approximate a target curve within a prespecified tolerance. We assume that the target curve is defined in 2D plane by a sequence of ordered dense data points. To estimate the curve, number of given points, N, assigned to L which is the length of antigen and antibody formed by a bit string. Every bit called a molecule of antibody and antigen, every bit corresponds to a data point. In this formulation, if a molecule is equal to 1, we put a knot on the partitioning point corresponding to the string position, and if the molecule is equal to 0 we do not (*see* Fig. 3). This representation is equivalent to terms of gene and chromosome in the genetic algorithm. The Clonal Selection Algorithm is different from genetic algorithm and its convergence is faster because individuals with much complementary feature are cloned more. If the given points lie in the interval $[a,b]$ then the appropriate number of knots are defined in this interval and called as interior knots. Initial population includes K antibodies, where each antibody is composed of L molecules. These molecules are randomly set to 0 or 1. Molecule set, which includes the place and number of interior knots that define the B-spline fitting to data points with minimum error, is the antigen to be recognized.

In the study, minimum distance theorem (for maximum interaction) is used for the interactions between antigen and antibody. To determine the level of antigen–antibody cell interactions, Euclidean distance as in Eq. (7) is used as a criterion. For Antibodies to produce a response to Antigens, they have to recognize the Antigens. In our study, the affinity of antibodies to antigens is determined in terms of (1) distance between the antigen and the antibody, and (2) Akaike's Information Cri-

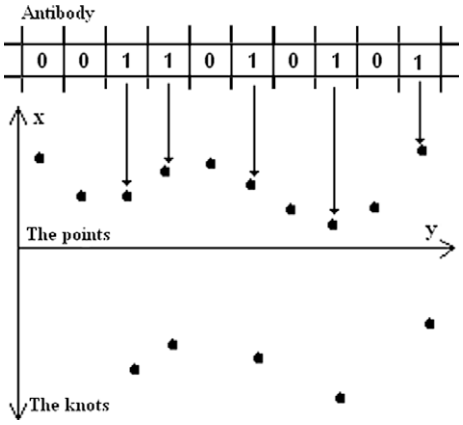

Fig. 3. The antigen–antibody coding method.

terion (AIC) which is preferred as a fitness measure in references Yoshimoto et al. [39] and Sarfraz et al. [35]. This formula is as follows

$$\text{Affinity} = 1 - \left(\frac{AIC}{Fitness_{avrg}}\right), \tag{9}$$

where $Fitness_{avrg}$ is the arithmetic mean of the AIC values of all antibodies in population and calculated as follows. If the AIC value of any individual is bigger then the $Fitness_{avrg}$ value, then we assumed that Affinity = 0 in Eq. (9)

$$\text{Fitness}_{avrg} = \frac{\sum_{i=1}^{K} AIC_i}{K}, \tag{10}$$

where K is the size of the population and AIC_i is the fitness measure of ith antibody in population. AIC is given as follows

$$AIC = N\log_e Q_1 + 2(2n + m), \tag{11}$$

where N is the number of data, n is the number of interior knots, m is the order of the spline to be fitted on the given data and Q_1 is calculated by Eq. (6). Beware that an antibody has the minimum error if it has the highest affinity among all antibodies. The affinity value of antibody that is the perfect complement of antigen which has the optimum solution and has to be recognized, is the one closest to 1 in population(actually in memory). The Euclidean distance between the ideal antibody and searched antigen equals to zero. In this case, the problem is not the curve approximation; it is the curve interpolation.

Before searching a solution to the problem, some control parameters should be given to program.These control parameters consist of the order of the B-spline curve, the size of the population, the size of the memory, the variety ratio, and the probability of mutation. Preferably, the memory size is double the amount of population size. In the memory inside, the best antibodies of all iterations are kept. The variety order of population also determines the parameter of variety ratio. This value is the ratio of the number of antibodies (whose molecules are randomly selected) to the population size. In the cloning phase, cloning is performed according to AIS guides and affinity value. Antibodies with high affinity are cloned more and antibodies with lower affinity are cloned less if not cloned at all. During the maturate application on cloned antibodies, usually double point crosswise operation is applied to an antibody which is selected randomly from memory or its molecule arrangement is changed randomly. If the clone of an antibody is more, usually both of them were applied to its clones. After running the AIS for a definite iteration number, the antibody which has the maximum sensitivity against the antigen is selected as solution.

In order to integrate the Clonal Selection Algorithm to this problem, some modifications are necessary in the original algorithm. Below we show how the modifications are applied to the above algorithm step by step.

1. Input the data point fitted.
2. Input the control parameters.
3. Create an initial antibody population by using random molecules.
4. If the population is created for the first time; create the memory array and store all antibodies in memory.
5. Otherwise, update the memory cells by using antibody population and improve the memory.
6. For each antibody, calculate the B-spline curve from Eq. (5) and fit it to data given by Eq. (3). Then, calculate sum of the squares of the residuals Q_1 (Eq. (6)).
7. Calculate the AIC value of each antibody in population (Eq. (11)) and calculate the average AIC value of the population (Eq. (10)).
8. Calculate the affinity value of each antibody (Eq. (9)).
9. Select the best antibodies by using the affinity and the interactions (between searched antigen and related antibody) of all antibodies in population.

10. Generate the maturity antibody population with the molecule deformation as proportional with the affinity values of clones.
11. Do mutation according to mutation rate.
12. Generate the new antibodies of the next generation according to variety rate.
13. If not reached to iteration limit or the antigen not recognized exactly go back to step 5.

5. Results

Fig. 4 shows an example of B-spline curve parameterization to evaluate the Automatic Knot Placement algorithm based on AIS proposed by us. Ten percent noise from uniform distribution was added artificially to a clean data for obtaining the 2D data (200 in number) shown in Fig. 4.a. The intended curve to he modeled is a nonuniform B-spline cubic curve with seven control points and $\{0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.75, 1.0, 1.0, 1.0, 1.0\}$ knot vectors. These figures also show the control polygon. In AIS architecture, because the antibodies which have the most sensitivity are saved in memory content, the most sensitive antibody of the memory is given in results for each generation. The root mean square (RMS) error between the modeled curve and point cloud based on the antibodies in memory population is reported in Table 1. The initial population was bred for 500 generations. Fig. 4b shows the best supplemental antibody in the initial population. Fig. 4c shows the convergence pattern after 500 generations. The curve now conforms better to the data points. The fitness is increases (the error decreases) while the generations are increasing. The slope of the curve shows that there is still possibility of converge in the further generations. Table 2, gives the set of parameters of the AIS optimization run.

To compare performance evaluation and speed of convergence, the compare were done between our algorithm based on AIS and the GA algorithm proposed from Sarfaz et al. [35,34]. In GA algorithm, knot ratio and manual fixing of important points at knot chromosomes are not taken into consideration. Input points are again the data points shown in Fig. 4a. To evaluate the accuracy, the RMS values obtained form GA method are given at (3rd column of) Table 1. Nonetheless, the values of parameters used for both of two algorithms are presented in Table 2. Our AIS based algorithm and GA based algorithm proposed by Safraz et al. are also compared according to speed of convergence. We are getting the program outputs of some

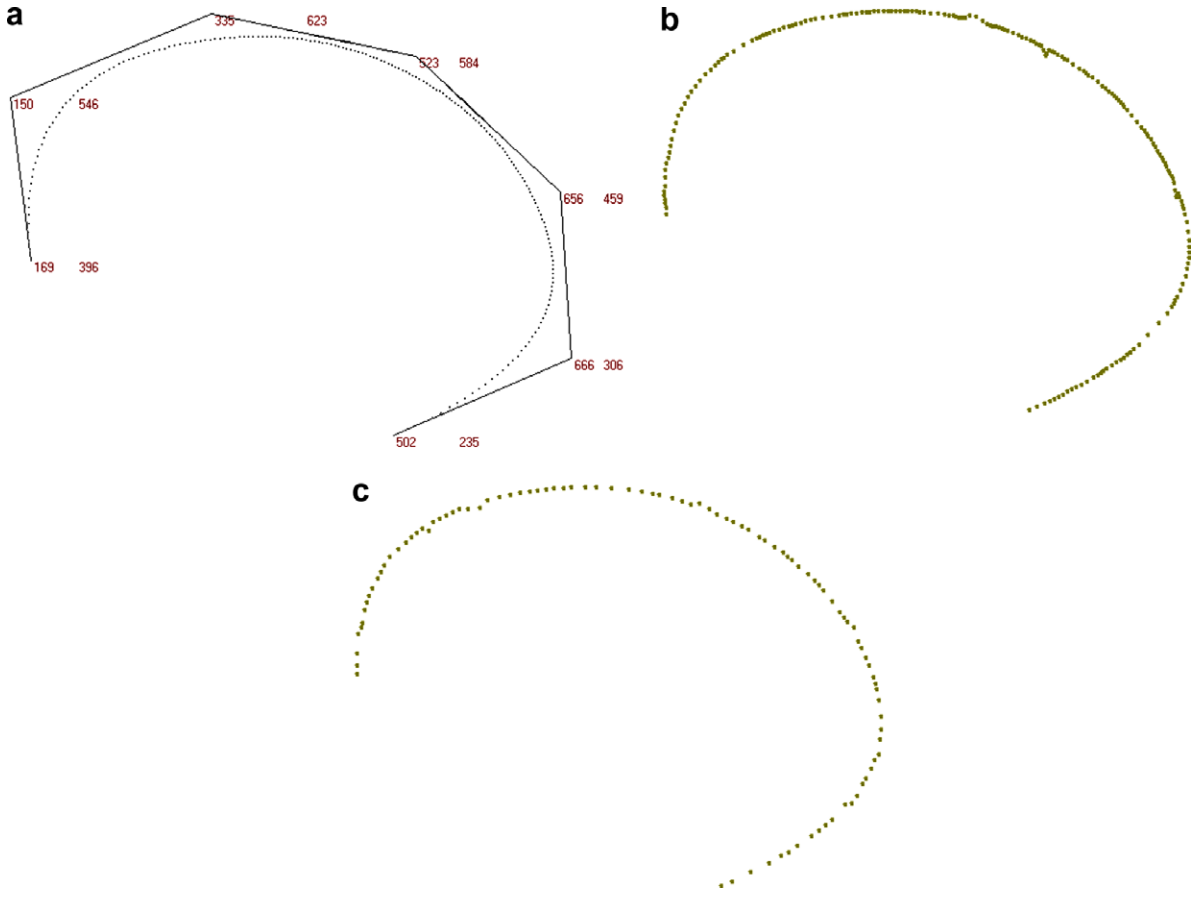

Fig. 4. A case study on knot adjustment for curves with noisy data: (a) The data points, B-spline with 7 control points, (b) the best complementary Antibody in initial population, and (c) the convergence Antigen after AIS optimization (500 generations).

Table 1
The RMS values of AIS and GA for the example shown in Fig. 4a.

Generation	BEST RMS (AIS)	BEST RMS (GA)
İnitial	3.3571	3.4517
10	3.3198	3.3578
25	3.3198	3.2348
50	3.2848	3.2348
100	3.2848	3.2348
200	3.2840	3.1016
300	3.2682	3.1016
400	3.2443	3.1016
500	3.2443	3.1016

Table 2
The set of parameters.

Parameter	AIS	GA
Population size	20	20
String length	200 (antibody cell length)	200 (chromosome gen-length)
Mutation rate	None	0.001
Crosover	None	0.7
Variety	6 (30%)	6 (30%)
Memory size	40	None

Table 3
Statistics of GA and AIS optimization shown in Fig. 5.

	GA				AIS			
	Max AIC (fitness)	Max RMS	Average AIC (fitness)	Average RMS	Max AIC (fitness)	Max RMS	Average AIC (fitness)	Average RMS
Initial	2188	6.498	2001	4.078	2396	9.693	2032	4.398
10	2028	4.403	1960	3.641	1913	3.749	1906	3.545
25	1959	3.802	1943	3.551	1978	3.604	1868	3.518
50	2015	3.933	1911	3.529	1840	3.500	1838	3.429
100	2069	4.370	1911	3.510	1819	3.511	1813	3.408
200	2017	4.456	1900	3.539	1801	3.390	1798	3.336
300	2068	4.228	1958	3.588	1797	3.349	1793	3.318
400	1994	4.558	1920	3.579	1798	3.346	1791	3.296
500	1948	3.864	1924	3.659	1798	3.346	1791	3.296

generations in training period. According to these outputs, according to population in that generation, maximum and average fitness (AIC) values of the individuals and antibodies are given in Table 3 (Also, maximum and average RMS values). According to all generations, the convergence diagrams of our proposed AIS approach and the GA approach are presented in Fig. 5. In this figure, bold lines are the maximum fitness values while normal lines are the minimum fitness values.

As a second case study, data to be fitted are generated by

$$F_j = 90/(1 + e^{-100(x_j - 0.4)}) \quad (j = 1, 2, \ldots, N), \tag{12}$$

where any error value is not used, the values of x_j is 0.0, 0.01, ..., 1.0 and the number of them is 101. The interval of the fitting was set to $[a,b] = [0,1]$. A cubic spline is used as an approximating function. Our method is not depends on the order of approximating function. Control parameters are the same with the values in Table 2.

Fig. 6 shows the fitness and number of knots against generations. In this figure, the black line shows the best fitness. From the black line, we can see that our computation converged at the 38th generation. The dotted line is the number of knots for the best fitness. The number of knots decreases from 54 at the first generation to 23 at the 38th generation. The GA solution for the same example is given in Fig. 6b. The best fitness value is 1436 obtained at 488th generation. The number of knots in this generation is 41. The generations, fitness values and knot numbers are given in Appendix A.

6. Conclusions and future works

In order to obtain a good B-spline curve model from large data, the knots are usually respected as variables. A curve is then modeled as a continuous, nonlinear and multivariate optimization problem with many local optima. For this reason it is very difficult to reach a global optimum. In this paper, we convert the original problem into a discrete combinatorial optimization problem like in Yoshimoto et al. [39] and Sarfraz et al. [35]. Then, we suggest a new method that solves the converted problem by artificial immune systems. We think the candidates of the locations of knots as antibodies. We define the affinity measure benefit from Akaike's Information Criterion (AIC). The proposed method determines the appropriate

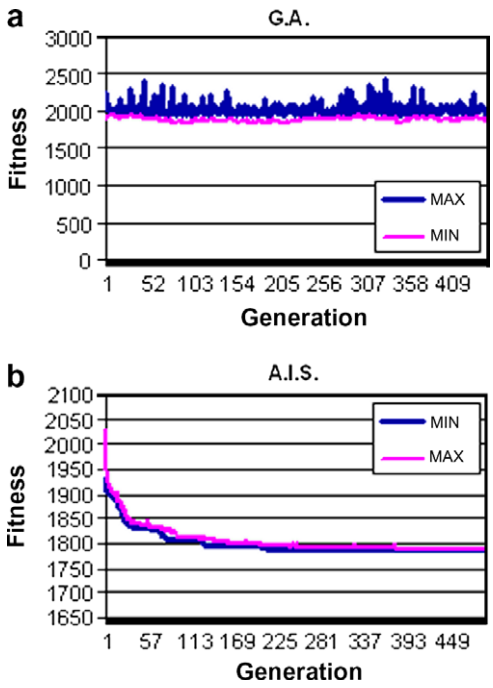

Fig. 5. GA and AIS based parameter optimization according to generations.

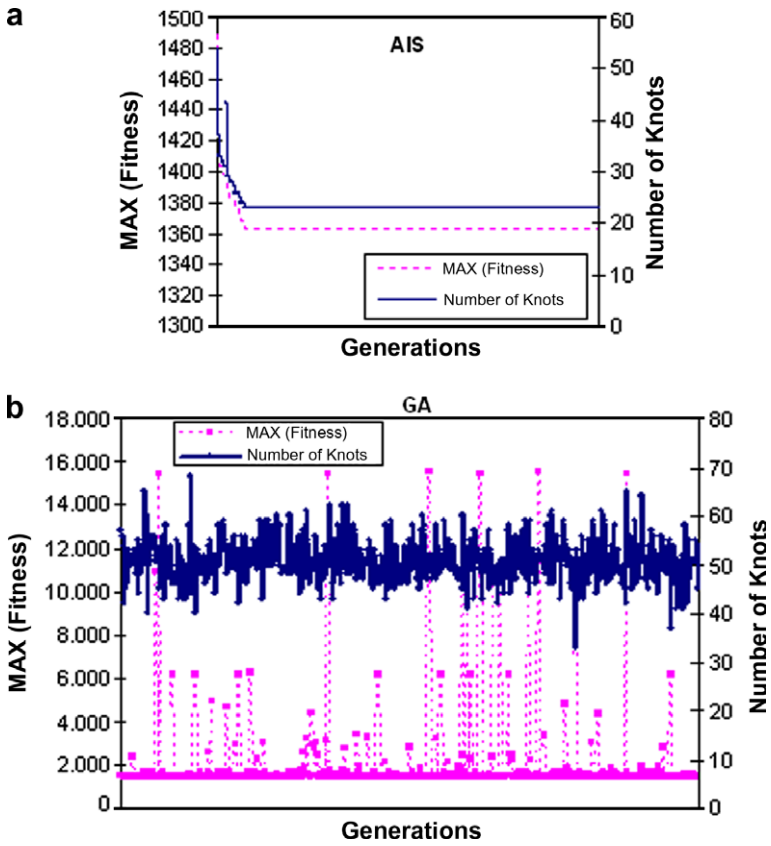

Fig. 6. Fitness and number of knots for second case study (a) A.I.S. and (b) G.A.

location of knots automatically and simultaneously. Furthermore, we do not need any subjective parameter or good population of initial location of knots for a good iterative search.

Genetic Algorithms have global perspective and robust. However, their convergence is slower and computation cost is more for generations which are much closer to optimal solution. Our algorithm has lower computation cost because of some AIS features such as capability of generating better antibodies from good antibodies, more cloning for good antibodies and having memories for antibodies. This approach is more promising due to its robustness and efficiency.

In the study, Centripetal method is used for assignment of a parameter value to each point in set of points and for selection of appropriate knot vector. Using of other methods (for example, equally spaced or chord length) instead of this method may provide more flexibility at adjusting relations between knot vector and parameterization and selection of operators. This flexibility may help us adjust the algorithm for better convergence to results. However, these methods have not been compared in this study.

Using of a cost function to design curve shape is a good strategy. Using of GA to solve for minimal cost function is proofed as a reliable way in literature. Using of AIS algorithm as Clonal Selection Algorithm which accelerates convergence speed of GA can provide much yield as presented in our study. An obligation in selection of appropriate cost function is a fundamental constraint. Another way is performance; algorithm is based on Visual Basic Programming Language. For better performance, neither a C implementation nor optimization of program code is expected. In addition, GA, ANN and AIS are not going to be appropriate for interactive curve design programming in near future.

Presented method is a curve approximation process. Then our method can be extended to multidimensional mesh data and B-spline surface approximation. For the extension of algorithm, NURBS are used with optimization processes of weights and control points in next steps. The extension is a challenge has to be solved because of complex structure of NURBS weights. Another expectation property for curves is to solve optimal design problem. We are developing an algorithm for minimization state in prediction of NURBS curve recently.

In addition to these, a framework is presented for constitution of gene pools which include initial population in GA in Kumar et al. [26]. To speed up our method, initial antibody population can be created to increase the convergence via this framework algorithm.

Parallel processors can also be used for reducing the running time of algorithm. However, the structure of our algorithm is not appropriate for this aim. With simple modifications, the algorithm can be adopted to a parallel algorithm.

Acknowledgements

This study has been supported by the Scientific Research Projects of Selçuk University (in Turkey).

Appendix A

See Table A1.

Table A1
Fitness and number of knots for 500 generations.

Generations	Fitness		Number of Knots		Generations	Fitness		Number of Knots	
	AIS	GA	AIS	GA		AIS	GA	AIS	GA
1	1488	1500	54	57	260	1364	1500	23	57
2	1488	1496	54	56	270	1364	1464	23	48
10	1400	1476	32	51	280	1364	1682	23	59
20	1384	1666	28	55	290	1364	1484	23	53
30	1376	1488	26	54	300	1364	1460	23	47
40	1364	1468	23	49	310	1364	1476	23	51
50	1364	1492	23	55	320	1364	1632	23	45
60	1364	1472	23	50	330	1364	10853	23	46
70	1364	1464	23	48	340	1364	2317	23	55
80	1364	1464	23	48	350	1364	1518	23	54
90	1364	1713	23	59	360	1364	1456	23	46
100	1364	2979	23	53	370	1364	1504	23	58
110	1364	1456	23	46	380	1364	1658	23	53
120	1364	1492	23	55	390	1364	1456	23	46
130	1364	1492	23	55	400	1364	1447	23	56
140	1364	1464	23	48	410	1364	1509	23	59
150	1364	1488	23	54	420	1364	1607	23	59
160	1364	1827	23	55	430	1364	1572	23	50
170	1364	3080	23	56	440	1364	1484	23	53
180	1364	15,521	23	50	450	1364	1488	23	54
190	1364	1464	23	48	460	1364	1674	23	57
200	1364	1464	23	48	470	1364	2839	23	55
210	1364	1460	23	47	480	1364	1456	23	46
220	1364	1484	23	53	490	1364	1448	23	44
230	1364	1486	23	43	499	1364	1493	23	55
240	1364	1456	23	46	500	1364	1452	23	45
250	1364	2817	23	51					

References

[1] G.L. Ada, G. Nossal, The clonal selection theory, Scientific American 257 (2) (1987) 50–57.
[2] J. Barhak, A. Fischer, Adaptive reconstruction of freeform objects with 3D SOM neural network grids, special issue on geometrical modeling and computer graphics, Journal of Computers and Graphics 26 (5) (2002) 97–105.
[3] F. Boudjemaï, P.B. Enberg, J.-G. Postaire, Surface modeling by using self organizing maps of Kohonen, in: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, vol. 3, Washington, DC (USA) 2003, pp. 2418–2423.
[4] D.S. Chen, R.C. Jain, B.G. Schunck, Surface reconstruction using neural networks, in: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'92, 1992, pp. 15–817.
[5] W.-K. Cheng, Data thinning for the NURBS surface in reverse engineering (CAD), Master Thesis, California State University, Long Beach, 1996.
[6] F.S. Cohen, W. Ibrahim, C. Pintavirooj, Ordering and parametrizing scattered 3D data for B-spline surface approximation, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (6) (2000) 642–648.
[7] C. De Boor, A Practical Guide to Splines, Springer-Verlag, 1978.
[8] L.N. De Castro, F.J. Von Zuben, Immune and neural network models: theoretical and empirical comparisons, International Journal of Computational Intelligence and Applications 1 (3) (2001) 239–257.
[9] L.N. De Castro, F.J. Von Zuben, Artificial Immune Systems: Part I – Basic Theory and Applications, Technical Report-RT DCA 01/99, 1999, p. 89.
[10] L.N. De Castro, F.J. Von Zuben, Artificial Immune Systems: Part II – A Survey of Applications, Technical Report-RT DCA 02/00, 2000, p. 65.
[11] L.N. De Castro, F.J. Von Zuben, The Clonal Selection Algorithm with Engineering Applications, in: Workshop Proceedings of GECCO'00, Workshop on Artificial Immune Systems and their Applications, Las Vegas, USA, 2000, pp. 36–37.
[12] L.N. De Castro, J. Timmis, Artificial Immune Systems: A New Computational Intelligence Approach, Springer-Verlag, 2002.
[13] O. Engin, A. Doyen, A new approach to solve hybrid flow shop scheduling problems by artificial immune system, Future Generation Computer Systems 20 (2004) 1083–1095.
[14] J.D. Farmer, N.H. Packard, A.S. Perelson, The immune system, adaptation, and machine learning, Physica D 22 (1986) 187–204.
[15] S. Forrest, A. Perelson, L. Allen, R. Cherukuri, Self-nonself discrimination in a computer, in: Proceedings of the IEEE Symposium on Research in Security and Privacy, 1994, pp. 202–212.
[16] R. Goldenthal, M. Bercovier, Spline Curve approximation and design by optimal control over the knots using genetic algorithms, in: Int. Cong. on Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems EUROGEN 2003, CIMNE, Barcelona, 2003, pp. 1–18.
[17] R. Goldenthal, M. Bercovier, Design of curves and surfaces by multi-objective optimization. Leibniz Report 2005-12, 2005.
[18] A.A. Goshtasby, Fitting parametric curves to dense and noisy points, in: 4th Int. Conf. Curves and Surfaces, St. Malo, France, 1999, pp. 227–237.
[19] M. Hoffmann, Modified Kohonen neural network for surface reconstruction, Publicationes Mathematicae Debrecen 54 (1999) 857–864.
[20] M. Hoffmann, Local update of B-spline surfaces by Kohonen neural network, in: Proceedings of the 5th International Conference on Computer Graphics and Artificial Intelligence, Limoges, 2002, pp. 103–112.
[21] M. Hoffmann, Numerical control of Kohonen neural network for scattered data approximation, Numerical Algorithms 39 (2005) 175–186.
[22] M. Hoffmann, E. Kovács, Developable surface modeling by neural network, Mathematical and Computer Modelling 38 (2003) 849–853.
[23] M. Hoffmann, L. Várady, Free-form curve design by neural networks, Acta Academiae Paedagogicae Agriensis, Sectio Mathematicae 24 (1997) 99–104.
[24] A. Iglesias, G. Echevarr'ýa, A. Galvez, Functional networks for B-spline surface reconstruction, Future Generation Computer Systems 20 (2004) 1337–1353.
[25] V. Krishnamurty, M. Levoy, Fitting smooth surfaces to dense polygon meshes, Computer Graphics (SIGGRAPH'96 Proceedings) (1996) 313–324.
[26] G.S. Kumar, P.K. Kalra, S.G. Dhande, Parameter optimization for B-spline curve fitting using genetic algorithms, The Congress on Evolutionary Computation 3 (2003) 1871–1878.
[27] S.G. Kumar, P.K. Kalra, S.G. Dhande, Curve and surface reconstruction from points: an approach based on self-organizing maps, Applied Soft Computing Journal 5 (5) (2004) 55–66.
[28] P. Laurent, M. Mekhilef, Optimization of a NURBS representation, Computer-Aided Design 25 (11) (1993) 699–710.
[29] W. Li, S. Xu, G. Zhao, L.P. Goh, Adaptive knot placement in B-spline curve approximation, Computer-Aided Design 37 (2005) 791–797.
[30] H. Park, J.-H. Lee, Error-bounded B-spline curve approximation based on dominant point selection, in: International Conference on Computer Graphics, Imaging and Visualization (CGIV'05), 2005, pp. 437–446.
[31] A.S. Perelsen, G.F. Oster, Theoretical studies of clonal selection: minimal antibody repertuarie size and reliability of self-nonself discrimination, Journal of Theoretical Biology 81 (1979) 645–670.
[32] L. Piegl, W. Tiller, The NURBS Book, second ed., Springer-Verlag, Berlin, Heidelberg, 1997.
[33] R. Goldenthal, M. Bercovier, Spline curve approximation and design by optimal control over the knots, Computing 72 (2004) 53–64.
[34] S.A. Raza, Visualization with spline using a genetic algorithm, Master Thesis, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, 2001, p. 126.
[35] M. Sarfraz, S.A. Raza, Capturing outline of fonts using genetic algorithm and splines, in: Fifth International Conference on Information Visualisation (IV'01), 2001, pp. 738–743.
[36] M. Sarfraz, S.A. Raza, M.H. Baig, Computing optimized curves with NURBS using evolutionary intelligence, Lecture Notes in Computer Science 3480 (2005) 806–815.
[37] V. Weis, L. Andor, G. Renner, T. Varady, Advanced surface fitting techniques, CAGD 19 (2002) 19–42.
[38] H. Yang, W. Wang, J. Sun, Control point adjustment for B-spline curve approximation, Computer-Aided Design 36 (2004) 639–652.
[39] F. Yoshimoto, M. Moriyama, T. Harada, Automatic knot placement by a genetic algorithm for data fitting with a spline, in: Proceedings of the International Conference on Shape Modeling and Applications, IEEE Computer Society Press, 1999, pp. 162–169.
[40] A.M. Youssef, Reverse engineering of geometric surfaces using Tabu search optimization technique, Master Thesis, Cairo University, Egypt, 2001.
[41] Z. Yu, Integrated NURBS surface reconstruction and boundary element analysis, Ph.D. Thesis, Universite Laval, 2003, p. 102.