Discrete NURBS-Surface Approximation
using an Evolutionary Strategy

Klaus Weinert, Jörn Mehnen

Department of Machining Technology, University of Dortmund, Germany

Abstract:
This article presents a study about the application of Computational Intelligence (CI) methods to the problem of optimal discrete surface approximation. The class of CI methods comprises subsymbolic algorithms like neural networks, fuzzy logic systems and evolutionary algorithms. Non-Uniform Rational B-Splines (NURBS) are flexible parametric functions which are commonly used in modern CAD/CAM systems. Here, a special CI method – the evolution strategy (ES) – will be used to approximate NURBS-surfaces to discrete 3D-point sets. The evolution strategy is a numeric optimization method that deals well with multi-modal optimization problems in real value vector spaces. The article focuses on the convergence behavior of ES regarding different parameterizations. The dependencies of the spline surface approximation algorithm on different surface structures are analyzed. The results are discussed from a mathematical and a practical point of view.

Keywords: NURBS surface approximation, surface reconstruction, evolution strategy.

Introduction

NURBS (Non-Uniform Rational B-Splines) are a flexible class for curve and surface descriptions that is commonly used in CAD/CAM systems. The advantages of this class lies in its easy and intuitive editing capabilities, the numerical stability, and the efficient computability. Further important features are their smoothness as well as their ability to form edges or peaks.

The application of NURBS in the field of surface reconstruction (sometimes called reverse engineering) yields CAD-adequate surfaces, i.e., surfaces that can easily be loaded, processed and manipulated in CAD/CAM systems. Discrete NURBS surface approximations imply the solution of multimodal nonlinear parameter optimization problems in the $\mathbb{R}^n$ (Laurent-Gengoux, 1993). Interpolations are only feasible in special cases and they require heuristic assumptions. Large equation systems have to be solved which generate surface descriptions that are not easy to be edited manually or processed by CAD systems. Here, discrete NURBS surface approximations of real world data are discussed.

The evolution strategy (ES) (Schwefel, 1994) is a probabilistic numerical optimization method that fits well with complex multimodal optimization problems in real value vector spaces. In certain design situations, when deterministic algorithms fail or are difficult to realize, evolutionary algorithms are very promising.

A deterministic approach using NURBS curve schemes has been studied by Piegl and Tiller (1997). This method was also extended for surface approximations. A method that uses genetic algorithms (GA) for Bézier-curve design was developed by Márkus, Renner and Vánca (1995). Mussa, Roy and Jared (1998)
integrate a Genetic Algorithm into a CAD system for spline curve optimization. These curves can be extended to ruled surfaces. Laurent-Gengoux and Mekhlief (1993) compare different deterministic strategies for knot and weight adaptation for NURBS curve and surface approximation. The approximation problem becomes even more difficult when the number of knots is dynamic or the NURBS surfaces are trimmed. Many authors stress the necessity of meaningful fitness criteria and robust optimization strategies. Evolution strategies can be used as an ideal test environment for the analysis of NURBS surface reconstructions because they do not use deterministic internal models and are independent from specific surface types or fitness functions.

In the following, an overview on the mathematical model and properties of NURBS will be given. A short description of the evolution strategy follows. The focus of this article lies on the analysis of the parameters of the strategy regarding the quality of the reconstructions as well as the influence of different surface types on the convergence of the ES. The conclusion will give insight into the capabilities and possible extensions of this NURBS surface approximation approach.

Expression for NURBS surfaces

NURBS surfaces are parametric bivariate tensor products that map the \( \mathbb{R}^2 \) to \( \mathbb{R}^3 \). The definition of a NURBS surface of order \((p, q)\) needs the following components:

- the control point net with \( n \times m \) vertices \( \mathbf{P} = \{ \mathbf{P}_{i,j} \in \mathbb{R}^3, i = 1, \ldots, n, j = 1, \ldots, m \} \),
- the knot vector \( U \) and \( V \), where \( U = (0, \ldots, 0, u_{q+1}, \ldots, u_{s-q-1}, \ldots, 1)^T \), \( V = (0, \ldots, 0, u_{p+1}, \ldots, u_{r-p-1}, \ldots, 1)^T \), and \( r = n + p + 1 \) and \( s = m + q + 1 \),
- the weight vector \( \mathbf{W} = \{ w_{i,j} \in \mathbb{R}^+, i = 1, \ldots, n, j = 1, \ldots, m \} \).

The parametric surface function is defined over the domain \((u, v) \in [0,1] \times [0,1]\) and expressed by

\[
\mathbf{S}(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} \mathbf{P}_{i,j} N_{i,p}(u) N_{j,q}(v) w_{i,j} / \sum_{i=0}^{n} \sum_{j=0}^{m} N_{i,p}(u) N_{j,q}(v) w_{i,j}.
\]

\(N_{i,p}(u) : \mathbb{R} \to \mathbb{R}\) are the \(i\)th basis functions of order \( p \) for the parameter \( u \in \mathbb{R}\) computed on a knot vector \((u_0, \ldots, u_n)^T\). The basis functions \(N_{i,p}(u)\) can be defined recursively by

\[
N_{i,0}(u) = \begin{cases} 1 & \text{if } u_i \leq u < u_{i+1} \\ 0 & \text{otherwise} \end{cases}
\]

\[
N_{i,p}(u) = \frac{(u-u_i)N_{i,p-1}(u)}{u_{i+p} - u_i} + \frac{(u_{i+p+1} - u)N_{i+1,p-1}(u)}{u_{i+p+1} - u_{i+1}}
\]

and evaluated efficiently by e.g. the Cox–de Boor algorithm.

The parameters of the control net vertices, knot vectors and the weight vectors can be joined to one vector \( \varphi = (\mathbf{P}, \mathbf{T}, \mathbf{W}) \) of the space \( \mathbb{R}^7 \) with dimension \( \gamma = 4nm + (n + p + m + q) \).

Depending on the parameter settings, NURBS are able to describe algebraic polynomials as well as Bézier descriptions or non-rational B-splines. Due to their rational characteristic, NURBS are a real superset of polynomials or piecewise polynomial non-rational models.

Formal problem definition

\( Q = \{ Q_1, \ldots, Q_k \} \) is a set of discrete points in \( \mathbb{R}^3 \) that come from digitizing systems. Without loss of generality, it is assumed that the set can be transformed into the interval \([0,1] \times [0,1]\). These points describe surfaces that do not have undercuts. The transformation parameters must be stored in order to expand the normalized reconstruction to its original size. It is assumed that the points are good representations of the real object surface. Due to the smoothing effect of NURBS, small noise in the data is allowed. The density and number of the points should be high enough to allow a NURBS-surface approximation that satisfies a given approximation error \( \epsilon \). The number of points should not be too high because this number has a direct effect on the performance of the reconstruction algorithm.

The approximation problem can be written as

\[
\Phi(\varphi) = \sum_i d(Q_i, S_{\varphi})^2 \leq \min_{\varphi'}.
\]

The cost function \( d \) describes the discrete difference between the NURBS surface and the set of digitizing points. \( d \) can be chosen arbitrarily to fulfill the expected properties of a reconstruction. The problem is to find the parameter setting \( \varphi \) of \( S \) that minimizes \( \Phi \).

In order to calculate the distance between the parametric NURBS surface \( S(u,v) \) and the discrete points \( Q \), a Newton scheme has been used to calculate the values \((u_k, v_k)\) of \( S \) that are closest to each \( Q_k \). In order to find good starting points for the Newton strategy, the two-dimensional intervals defined by the knot vectors \( U \times V \) are divided into \((2p+1) \times (2q+1) \) equally spaced sub-intervals.
$t \in [t, \infty]$ is a scaling factor that is increased as long as the Newton strategy does not converge. The result of the Newton algorithm is a point $(u_k, v_k)$ of the surface $S$ that can be used to calculate the discrete distance $d$ between $S$ and $Q$. The complexity of the evaluation of the fitness criterion using Newton’s scheme is determined by the number $(m+1-p)(2p+1)t+1)$ of evaluations and the number of digitized points $l$. The complexity of the algorithm can be estimated with $p = q$ and $n = m$ to $O(n^2 p^2 t^2 + l)$. The number of sampling points has a quadratic effect on the performance of a reconstruction algorithm. Therefore, the points have to be reduced in advance in an adequate way.

The central-limit theorem implies that the most reasonable choice for the distribution of points without $a - priori$ knowledge is Gaussian. This leads, via the maximum-likelihood approach, to the least-squares criterion as a measure of the approximation quality.

Due to the fact that any solution of the approximation interpolates the border of the point set (Pigg and Tiller, 1997), the control points on the border of the NURBS surface are kept fixed during the optimization process. They are defined by an interpolation function along the square border of the point set. This procedure does not only allow to find exact approximations but also reduces the dimension of the search space.

Evolution strategy

The evolution strategy (ES) developed by Rechenberg and Schwefel in the 1960s belongs to a class of numerical probabilistic optimization algorithms that also includes genetic algorithms (GA) (Holland, 1975) and evolutionary programming (EP) (Fogel, Owens and Walsh, 1966). This class is called evolutionary algorithms (EA).

Evolutionary algorithms follow Darwin’s idea of evolution. Evolution strategies use vectors $x \in \mathbb{R}^n$ to find extremal values of (usually) scalar functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The search process uses a set (population) of $\mu$ test points in the search space $\Pi^\mu$ (parent individuals), from which $\lambda$ new solutions $\Pi^\lambda$ (offspring) are generated by recombination. After recombination the offspring is changed by mutation in order to introduce new possible solutions. Depending on the selection strategy, $\mu$ new best parent individuals are selected either only from the offspring population $(\mu, \lambda)$-ES or from the offspring and the parent individuals $(\mu + \lambda)$-ES. The new $\mu$ individuals are the new parents for the next generations.

Each individual $I$ contains two components:

- object parameters $x \in \mathbb{R}^n$ that are evaluated by the fitness function,
- step-size parameters $\sigma \in \mathbb{R}^n$ that are used in the mutation operator.

Sometimes a covariance matrix $C \in \mathbb{R}^{2n}$ is also encoded by $I$ (Schwefel, 1994). In this article the covariance matrix is not used. The ES algorithm contains three main operators that are iterated as long as a stopping criterion does not hold:

- recombination $r$: $I^m \subseteq I^\mu \rightarrow I$. The operator chooses $m$ individuals $I$ by an equally distributed random process from the parent population. In the experiments intermediate recombination on both components $x$ and $\sigma$ with two ($m = 2$) parent individuals is used. A new individual is defined by the arithmetic mean of each component of the two parents.

- mutation $m$: $I \rightarrow I$ maps each recombined individual to a slightly changed offspring individual. The mutation follows the scheme:

$$\sigma_i^t = \sigma_i \cdot \exp(N(0, \tau_0) + N(0, \tau_1))$$

$$x_i^t = x_i + N(0, \sigma_i^t)$$  \hspace{1cm} (4)

$N(\nu, \sigma)$ denotes the normal distribution with expectation value $\nu$ and variance $\sigma$. Algorithms that generate a sequence of normally distributed pseudo random numbers (e.g. the polar method of Box and Muller) can be found in (Knuth, 1981).

Here, $\tau_0, \tau_1 \in \mathbb{R}$ are constant external parameters that control the step size variation. Their specification depends on the fitness function. Classically the learning rates are defined like this (Schwefel, 1994):

$$\tau_0 = \frac{1}{\sqrt{2m}} \hspace{1cm} \tau_i = \frac{1}{\sqrt{2\sqrt{n}}}$$  \hspace{1cm} (5)

Due to results from extensive experiments, Kursawe (1999) proposes for search spaces with dimension $n > 30$ depending on the fitness function several different $\tau_0$ and $\tau_i$. For $\tau_0$ his experiments agree to the classical heuristic $\tau_0 = 0.8n^{-0.32}$. The behavior of $\tau_i$ seems to follow the regression function $\ln(n \cdot 10^6) \cdot 0.035$. In order to improve the explorative character of an ES, a relation of $\tau_0/\tau_i \approx 0.2/0.1$ seems to be appropriate. For each individual the normal distribution using $\tau_0$ is calculated only once yielding a variation
factor that has a constant influence on the step size of an individual. Furthermore, the object variable vector of an individual is varied with factors from normal distributions using $\tau_i$ that are calculated for each element separately.

- selection $s$: $I^\lambda \to I^\mu$ is an operator that chooses individuals from a population that are most promising to improve the search results regarding the fitness criteria. In ES either the ($\mu$, $\lambda$)-strategy or the ($\mu$ + $\lambda$)-strategy (so called truncation selection schemes) is used. Other strategies like tournament selection, ranking selection or fitness-proportional selection are also possible. The choice of the best strategy depends on the problem. One of the advantages of truncation selection schemes is the scalability of the selection pressure $\lambda/\mu$ that determines the take over time of good solutions within a population over the generations (Bäck, 1996).

The fitness rating of the object variables allows an indirect rating of the step sizes and, thus, the self-adaptation of the step-size parameters. An external deterministic step-size adaptation operator is not necessary.

The stopping criterion can be chosen arbitrarily. Typically, an ES is stopped when the fitness value reaches a sufficient accuracy, the fitness values do not change over an externally defined number of generations or the total number of generations exceeds a given limit.

**Experimental setup**

Each individual represents a NURBS surface by a vector $I \in \mathbb{R}^7$. Here, $(p,q) = (3,3)$ and $m \times n = 8 \times 8$ vertices were used. The populations were initialized by deterministically computed approximations of the discrete point sets. The deterministic approximation algorithm follows the optimization scheme of Piegl and Tiller (1997).

The following table shows the parameters for the evolution strategy that have been tested using a one-factor-at-a-time scheme. A ($\mu$ + $\lambda$)-ES and a ($\mu$, $\lambda$)-ES were used. All runs have been stopped after 2,000 generations.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>1</th>
<th>10</th>
<th>20</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>75</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>0.1</td>
<td>0.15</td>
<td>0.2</td>
<td>0.25</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>0.475</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to describe the dependencies of the reconstruction process on the shape of the surfaces, the top of a piston and a segment of a cover plate of a brake were digitized. The piston and the cover plate are each represented by about 3,000 points. The tactile digitizing system used in the experiments has an accuracy of about $\pm 20\mu m$. In order to avoid errors via other surface estimation models, a touch probe radius compensation of the measured point data was not performed. The dimensions of the original object should be kept together with the sampling points and all transformation parameters (e.g. for scaling purposes) used during the reconstruction process.

The parameter tests were performed on a SGI ORIGIN 2000 machine. The software was developed in C++ on PCs (Intel Pentium III) under LINUX.

**Results**

The design of the piston top is artificial and shows a wave-shaped surface with a steeply bending ridge in the middle. The segment of the cover plate belongs to a sheet metal that was generated by a deep-drawing process. The plate segment shows deep cavities with smooth chamfers. Fig. 1 presents photos of the two original workpieces.

![Photo of the top of a piston (above) and a cover plate of a brake (below).](image)
In Fig. 2 the best solutions of the reconstruction processes are shown. The rendered NURBS-surfaces and the control nets are displayed. The different gray scales denote the deviations from the original point set. Dark areas show large approximation errors while light regions denote small or no errors.

This point distribution corresponds with the notion of human CAD design. Other evolutionary reconstructions do not always follow human intuition of good point distributions. Since reconstruction results are generally not unique also these automatically generated solutions are correct approximations of the given original surfaces indepentent.

The analysis of the parameter settings of the evolution strategy shows that large population sizes yield better results than small ones. Populations with many individuals support the evolutionary reconstruction process because recombination between individuals with very different control net structures often cause offspring individuals with bad fitness values. Large population sizes reduce this effect. This recombination effect hampers approaches with distributed or interleaved population structures. The increasing of the relative selection pressure \( \lambda / \mu \) causes a little bit less than linear improvement. In the case of \( \mu = 1 \) the fitness values become worse with increasing numbers of offspring. In the experiments, the parameter settings \( \mu = 50 \) and \( \lambda = 200 \) yielded the best results. This choice corresponds with the theory where typical selection pressure values of 5 to 7 are recommended.

The optimum of the parameter settings of \( \tau_0 \) and \( \tau_1 \) is not easy to find. The maxima of the solution qualities follow a small ridge that becomes even smaller for higher dimensions. Kursawe (1999) calls the values along this ridge \( F(\tau_1, \tau_0) = \log \frac{\mu}{\mu + \lambda} \) (measure of convergence) and Rechenberg (1994) labels them the 'window of evolution'. High values of \( \tau_1 \) and \( \tau_0 \) yield to stagnation of the \( (\mu + \lambda) \)-ES, while \( (\mu, \lambda) \)-ES may show divergence. Too low values of \( \tau_0 \) and \( \tau_1 \) imply low convergence velocities. In the experiments of the NURBS reconstruction, \( \tau_0 = 0.2 \) and \( \tau_1 = 0.475 \) yielded the best results.

In the experiments, the \( (\mu, \lambda) \)-ES usually yielded better results than the \( (\mu + \lambda) \)-ES. The global behavior of both strategies were quite similar, although the fitness function could not always be evaluated in a stable manner for the \( (\mu, \lambda) \)-ES. The better results of \( (\mu, \lambda) \)-ES is typical for multi-modal fitness functions, because the algorithm is able to 'forget'. This way, on the long term, the genetic diversity in panmictic populations can be kept high enough to avoid stagnation of the search process in suboptima. Kursawe (1999) shows more generally that for optimal parameter settings a \( (\mu + \lambda) \)-ES can also perform very good on special fitness functions.
Fig. 3 shows the fitness function values and the average step sizes of a (50 + 200)-ES used to reconstruct the surface of a piston.

Tests with digitized surfaces that show small stochastic errors in the sampling points (e.g., caused by the digitizing process) indicate that applications of ES to NURBS surface reconstructions are robust, i.e., no numeric instabilities appeared. The smoothing (low-pass filter) effect of NURBS reduced the random noise in the data. This smoothing effect may also hamper the exact representation of very detailed surfaces. A reconstruction of small details is only possible with higher numbers of control points. Alternatively, a complex surface has to be decomposed into smaller segments that can be reconstructed separately. The necessary number of basis functions for B-Spline curve approximations depends on the curvature (Lipschitz continuity), the expected tolerance and the distance metric. Experiments with self-adapting varying dimensions are still a matter of actual research activities.

The time consumed by the ES per parameter setting lies between 3 and 6 hours. In the experiments the memory consumption did not exceed 8MB. These values depend on the number of individuals, the control-point density and the number of sampling points. Experiments on the same surface with high \( \lambda \) values showed similar results in less time than tests with a higher grid resolution.

Conclusions and Outlook

Quality demands in CAD, e.g., the design of turbine blades, can be very high. Triangulations may cause problems in CAD systems when the number of vertices reaches a certain threshold (Weinert, 1998). NURBS are an efficient and intuitive way to represent smooth surfaces using only few control points. The complexity of the optimization problem excludes manual techniques and hampers deterministic as well as probabilistic methods. Although the processing times of evolutionary algorithms are quite long, the results are of high quality. The solutions of the reconstruction algorithm can directly be used in common CAD/CAM systems for manual or computer-aided processing.
It has been shown that evolution strategies are able to solve the complex optimization problem of NURBS surface approximation and can be better than deterministic strategies. Very large and composed objects will need high numbers of control points and, thus, the ‘curse of dimensionality’ will slow down deterministic as well as probabilistic strategies. Intelligent pattern recognition and segmentation schemes will be necessary to improve the performance of reconstruction systems.

First experiments with self-adapting numbers of control points show that the problem of variable dimensionality in ES raises many new questions. A property of NURBS is that the number of control points of a NURBS surface can be increased without changing the shape of a surface. This is not always true for the deleting operation. An increasing number of control points increases the dimension of the search problem but also simplifies the reconstruction process, because the NURBS surface becomes more flexible. Because deletion has a strong impact on the surface structure the deletion of control points often yields surfaces with fitness values that are worse than without deletion. Hence, these new individuals often will not survive selection. In order to avoid the introduction of too many new control points, an additional fitness function (multi-criterion optimization) or new evolutionary operators have to be introduced. The introduction of an ageing factor in the selection process may help to keep individuals with relatively low fitness values for a certain period of time.

Future research aspects will also focus on the combination of variable-dimensional ES/CP (Genetic Programming) systems using NURBS and CSG (Constructive Solid Geometries) (Keller, 1999). This way, two different modeling techniques will be joined in one hybrid system.

Acknowledgments

Many thanks to P. Drerup, R. Körner, R.E. Keller, and M. Kalveram. This research was supported by the Deutsche Forschungsgemeinschaft as part of the Collaborative Research Center “Computational Intelligence” (SFB 531).

References


