A SPLINE APPROXIMATION OF A LARGE SET OF POINTS

Jana Kostková, Radim Halíř
Department of Software Engineering, Charles University,
Malostranské nám. 2/25, 118 00 Prague, Czech Republic
jkos4217@ss1000.ms.mff.cuni.cz, *halir@ms.mff.cuni.cz*

ABSTRACT

This paper presents a spline approximation method for the representation of a large set of points. The representation should be smooth with preserving important shape characteristics given by the points. Because of a large size of the set, the standard spline interpolation cannot be used. The proposed method is based on a least squares minimization of the distances of the points from the spline function subject to the conditions of smoothness of the representation. The spline approximation produces accurate and suitable representation of the points. The proposed approach has been verified on both synthetic and real data sets of points.

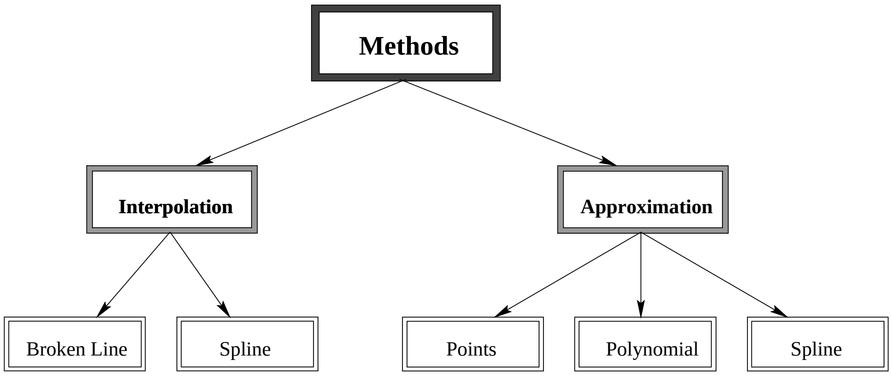
Keywords: spline approximation, fitting, least squares

1 INTRODUCTION AND MOTIVATION

An important problem of pattern recognition and computer vision is fitting the set of points by parts of geometric primitives, described by a curve or a function. The set of points is a typical result of many applications such as computer acquisition of various real objects. The representation of these results by the set of points is not convenient for two reasons: Firstly, the set is usually too large to be processed due to the computational time constraints. Therefore, the number of points needs to be reduced while all the essential characteristics expressed by the set have to be preserved. Secondly, the function representation is more appropriate for further processing and it better describes the given set of points. For that reasons it is advisable to replace the set of points by a function.

The proposed method was motivated by the requirement of a suitable representation of profiles of fragments of archaeological pottery. Each fragment needs to be measured and classified according to its shape and material characteristics [Rice87]. The manual classification is very time-consuming process. Thus, the computer based acquisition of the processing of fragments is highly required. The computer processing is performed in this manner: Each fragment is measured by a computer, based on the computer vision methods (stereo-aquisition or structure-light-aquisition) [Sabla93], what gives its 2.5-D object model. On this data, the axis of rotation of the fragment is estimated [Halíř99]. After the axis of rotation is known, the model of the fragment is transfered according to the rotation axis to the 2-D set of points which represents the profile of this fragment. These points are noisy caused by errors in aquisition the 2.5-D model, in estimation the rotation axis and in transfering the data to the 2-D set of points. Consequently, the set of points is not suitable for further processing and thus it needs to be replaced by better representation. This representation should be smooth and preserve the important shape characteristics of the given set. The spline approximation method represents the profile of a pottery fragments sufficiently.

In this paper, the spline approximation method for the replacing of the set of points is proposed. The approximation minimizes the average distance of the points from the fitted function. The proposed method is able to estimate the spline in accurate, fast and stable manner. The paper is organized as follows: in the next Section, various methods for the representation of a set of points are mentioned. The following Section describes the spline



(a)

Figure 1: An overview of the methods for the representation of the set of points

(a) (b)

Figure 2: Two possibilities of fitting the function to the set of points: (a) interpolation, (b) approximation

approximation method in details. In the next Section, the proposed method is verified on synthetic and real data. In the last Section, the whole paper is concluded and some future directions are outlined.

2 REPRESENTATION METHODS

It was already noted that the set of points is typically not convenient for further processing. Therefore, it needs to be replaced by another representation. There already exist various methods for the representation of the given set of points. The methods can be divided into two main groups, as depicted in Fig. 1: interpolation or approximation.

The basic difference between interpolation and approximation methods is demonstrated in Fig. 2. When the number of points is small, the interpolation method is convenient. But when the set is large and the points are noisy, the approximation method should be applied.

3 SPLINE APPROXIMATION

The method for the representation by the *spline approximation* replaces the set of points by a spline function. The spline is a piecewise polynomial function with conditions of smoothness on the endpoints of the polynomials. The mathematical theory of splines is known for a long time [Reins67, Reins71]. In spline approximation method, the given points are divided into sufficient number of intervals which are defined by the user. In each interval the points are approximated by a polynomial. The coefficients of the polynomials are computed with respect to the smoothness of the spline function. In typical applications the spline functions of the second or the third degree are used, because this degree is sufficient [Boor78]. The higher degree only adds the computational difficulty. In this approach the splines of the third degree are used.

The accuracy of the representation can be characterized by the approximation error. The approximation error is given as the average distance of the points from the approximating function in this approach. In [Punta98], another criterion of the accuracy was defined: the number of inflex points. However, that criterion does not characterize the smoothness and the accuracy of the function sufficiently, because of its discontinuity and irregularity.

In order to represent the given set of points by the spline approximation it is necessary to solve the following tasks:

- define the sufficient number of intervals
- compute the appropriate spline coefficients
- estimate the error of the provided approximation

3.1 Number of Intervals

The decision of the right number of intervals is crucial for the overall success of the method. More intervals bring more accurate approximation. However, in order to avoid unwanted side-effects (such as oscillations of the function) and also due to the performance, it is desirable to keep the number of intervals small. For many applications, the small number of intervals $(5 - 10)$ is sufficient.

3.2 Spline Coefficients

After the number of intervals is known, the coefficients of the appropriate spline function are estimated by the approximation of the points with respect to the function smoothness. The standard least squares method [Lawso95] is chosen for computing these coefficients, what formulated the following task:

$$\min_{\mathbf{c}} \|\mathbf{Ac} - \mathbf{b}\|^2 \quad \text{under} \;\; \mathbf{Bc} = \mathbf{0} \; . \tag{1}$$

The matrix $\mathbf{A}$ expresses the approximation of the points, the matrix $\mathbf{B}$ represents the smoothness conditions and $\mathbf{c}$ is the vector of spline coefficients. In the following equations, two constants are used: N is the number of points and M is the number of intervals the points are divided into.

The vector $\mathbf{c}$ of the length $(4 * M)$,

$$\mathbf{c} = (a_1, b_1, c_1, d_1, ..., a_M, b_M, c_M, d_M)^T \; , \tag{2}$$

represents the coefficients of the spline function on each interval. The coefficients $\{a_j, b_j, c_j, d_j\}$ define the polynomial in the j-th interval as follows:

$$\begin{array}{l} a_j * x_i^3 + \; b_j * 3x_i^2(1 - x_i) + \\ + \; c_j * 3x_i(1 - x_i)^2 + \; d_j * (1 - x_i)^3 \end{array} \; , \tag{3}$$

where $\{x_i^3, 3x_i^2(1-x_i), 3x_i(1-x_i)^2, (1-x_i)^3\}$ are so called Bernstein polynomials and $x_i, \; i = 1, ..., N_j$ are the x-coordinates of the points in the j-th interval, N_j is the number of points in the j-th interval.

The matrix $\mathbf{A}$ of the size $N \times (4 * M)$,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_1} & 0 & ... & 0 & 0 \\ 0 & \mathbf{A_2} & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & \mathbf{A_{M-1}} & 0 \\ 0 & 0 & ... & 0 & \mathbf{A_M} \end{pmatrix} , \tag{4}$$

represents the approximation of the points over all the intervals. It consists from the sub-matrices $\mathbf{A_j}$ for each interval. The matrices $\mathbf{A_j}$ of the size $N_j \times 4$,

$$\mathbf{A_j} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & \delta_1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{Nj} & \beta_{N_j} & \gamma_{N_j} & \delta_{N_j} \end{pmatrix} , \tag{5}$$

describes the approximation of the points from the j-th interval, $j = 1, ..., M$. The $\{\alpha_i, \beta_i, \gamma_i, \delta_i\}$ are the Bernstein polynomials defined in Eq. 3. Note that each interval can contain different number of points (and thus the matrix $\mathbf{A}$ is not block diagonal) and that

$$\sum_{j=1}^{M} N_j = N \; . \tag{6}$$

The column vector $\mathbf{b}$ of the length N,

$$\mathbf{b} = (y_1, y_2, ..., y_{N-1}, y_N)^T \; , \tag{7}$$

is created from the y-coordinates of the points.

The matrix $\mathbf{B}$ of the size $(3 * (M - 1)) \times (4 * M)$:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B_1} & \mathbf{B_2} & 0 & ... & 0 & 0 & 0 \\ 0 & \mathbf{B_1} & \mathbf{B_2} & ... & 0 & 0 & 0 \\ 0 & 0 & \mathbf{B_1} & ... & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & ... & \mathbf{B_2} & 0 & 0 \\ 0 & 0 & 0 & ... & \mathbf{B_1} & \mathbf{B_2} & 0 \\ 0 & 0 & 0 & ... & 0 & \mathbf{B_1} & \mathbf{B_2} \end{pmatrix} , \tag{8}$$

represents the smoothness conditions in the end-points. In this approach the splines of the third order are used, and for that reason is required the continuity of the function to the second derivative. The continuity conditions are expressed in the following matrices:

$$\mathbf{B_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 \end{pmatrix} , \tag{9}$$

and

$$\mathbf{B_2} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & -1 & 2 & -1 \end{pmatrix} . \tag{10}$$

To solve the problem described in Eq. 1, a standard method of Lagrange multipliers [Barre94] is used, which yields:

$$\begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{A} & \mathbf{B}^{\mathrm{T}} \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{c} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{b} \\ \mathbf{0} \end{pmatrix}, \qquad (11)$$

where λ is a vector of Lagrange multipliers. From Eq. 11 the coefficients of the spline function (the vector $\mathbf{c}$, defined in Eq. 2) can be computed directly by Gauss elimination.

3.3 Approximation Error

The aim of the sufficient representation is to replace the given set of points by the spline which minimizes the approximation error. In our approach, the approximation error is defined as the average distance of the points from the estimated spline function,

$$E = \frac{1}{N} \sum_{i=1}^{N} |f(x_i) - y_i| \ , \qquad (12)$$

where N is the number of points, $f(x_i)$ is the spline function defined in Eq. 3 and (x_i, y_i) are the coordinates of the points. It is clear that increasing the number of intervals decreases the approximation error. But due to the input noise, this decreasing is limited. Consequently, when the approximation error reaches the input error, the increasing number of intervals does not change the approximation error considerably.

4 PRACTICAL REALIZATION

The proposed spline approximation is performed in the following steps:

1. *Preprocessing of the set of points:* After the number of intervals is defined the given points are divided into the corresponding intervals.

2. *Construction of the linear system of equations:* The matrices $\mathbf{A}$ (Eq. 4) and $\mathbf{B}$ (Eq. 8) and the vector $\mathbf{b}$ (Eq. 7) are constructed with respect to the given points divided into relevant intervals. From these matrices the linear system of equations, defined in Eq. 1, is set up.

3. *Computation of the spline coefficients:* The coefficients of the spline, defined in Eq. 2, are computed by Gauss elimination method form the Eq. 11.

4. *Estimation of the approximation error:* The approximation error of the estimated spline representation is computed as given in Eq. 12.

5 EXPERIMENTAL RESULTS

The proposed method has been evaluated in many experiments. The experiments confirmed the accuracy and stability of the spline approximation used for the representation of the given sets of points. The approach was tested on both synthetic and real data. The method works fast, comparing with other methods, such as the spline interpolation.

The method has been used practically in the processing of the fragments of archaeological pottery. An automated processing of fragments is highly requested in archaeology now [Sabla93, Halíř97a, Halíř97b]. Every fragment needs to be measured and classified [Orton93]. An important part of the classification is the estimation of so called profile of a fragment. First, the fragment is measured by a computer, using the computer vision methods, what results in its 2.5-D model. After the axis of rotation of this fragment is known [Halíř99], the model is transformed to the set of 2-D points representing the profile of the investigated fragment. The set of points is typically not convenient for further processing and that is why it needs to be replaced by better representation. The ancient pottery were hand-made on the potter's wheel and therefore it is rotationaly symmetric and the shape of the pots is typically smooth. Consequently, the spline approximation is suitable for the representation of the profile of the fragments.

Various representations of the given set of points are compared in Fig. 3. This experiment was performed on the real data set which represents the profile of a pottery fragment (depicted in (a)). Five intervals were used for the estimation in all the techniques. In the broken line representation (b), the number of points is reduced first. In each interval, the points are replaced by their center of gravity. These centers are marked by a "+" in sub-figures. The broken line is fitted through these new points. This method is fast, but it does not represent the given set of points very well. The interpolation spline (c) works also with the reduced number of points (made by the same way). Then the spline function is interpolated through these points. This representation is smooth, but it is not accurate due to the systematic errors caused during the reduction of the number of points. The spline approximation (d) represents the set of points conveniently, it is smooth

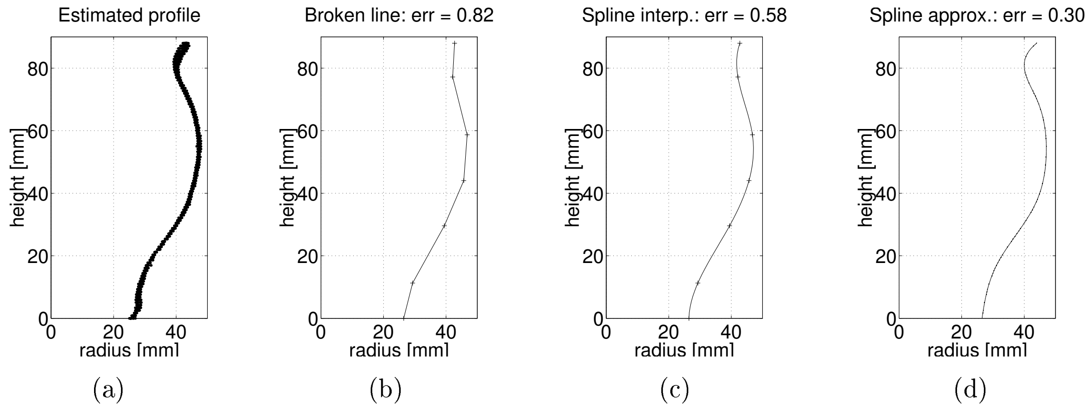


Figure 3: Different representations of the set of points: (a) the given set of points, which represents the profile of a fragment, (b) broken line, (c) spline interpolation and (d) spline approximation. Five intervals are used in all the representation. Compare the appropriate error of approximation given in each example.

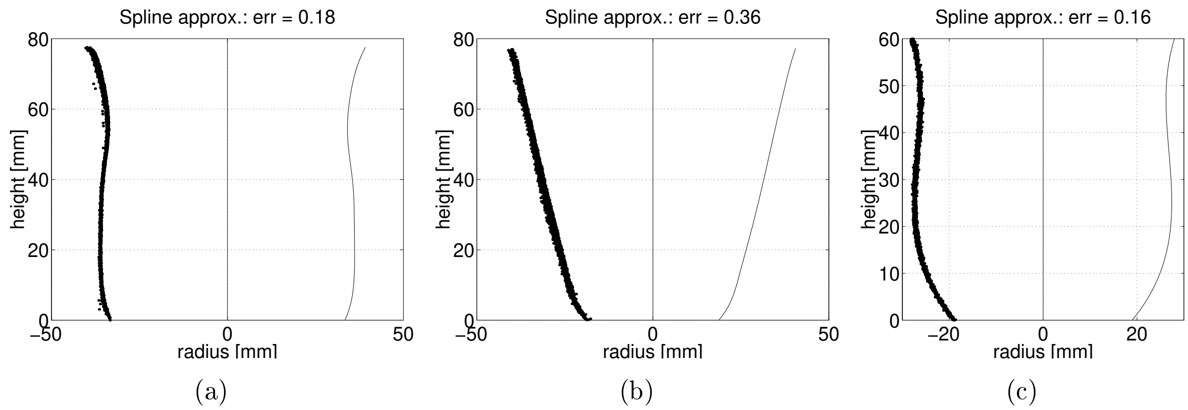


Figure 4: Results of fitting the spline approximation technique applied on various input data: (a) a real data set which represents the profile of a pottery fragment, (b) a real data set which represents the profile of another fragment, (c) a synthetic data set which represents the noised cubic function. The given sets are illustrate on the left sides of sub-figures, the estimated splines on the right sides. Compare the appropriate error of approximation given in each example.

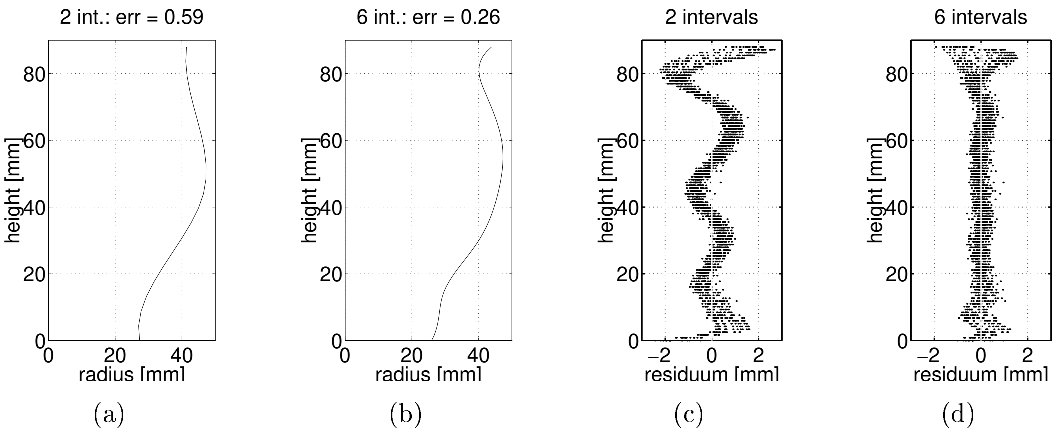
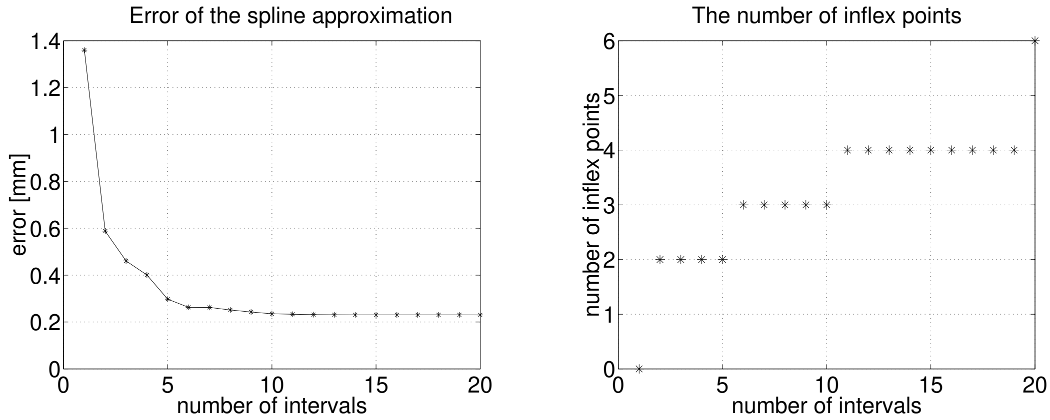


Figure 5: Two representations of the real data set from Fig. 3a using different number of intervals: (a) two intervals, (b) six intervals. Compare the appropriate error of approximation given in each example. The accuracy of representations is illustrated by the residua of points of function with: (c) two intervals and (d) six intervals. In can be seen that two intervals are insufficient, but six intervals represents the given set accurate.



(a) (b)

Figure 6: The influence of the number of intervals to the accuracy of the representation: (a) the approximation error, (b) the number of inflex points. The approximation error decreases continuously, but from six intervals it is not changing considerably. The number of inflex points changes slowly and non-continuously. Thus, it does not characterize the accuracy of the estimated representation conveniently.

and accurate.

The proposed spline approximation technique applied on various data sets is depicted in Fig. 4. Six intervals were used in all examples. The method applied on real data is demonstrated in Fig. 4 (a) and (b). These sets represent the profiles of fragments from different archaeological pots. It can be seen that the spline approximation is accurately, even with small number of intervals. The application on the synthetic data is illustrated in Fig. 4 (c). This set was generated by adding the Gaussian noise to the points which represent the cubic function:

$$50\tilde{x}^3 - 90\tilde{x}^2 + 49\tilde{x} + 18\ ,\ \tilde{x} = x/60\ ,\ x \in \langle 0, 60\rangle$$

This cubic function was chosen, because it simulates a profile of a real pot. For that reason it was possible to evaluate the proposed method on the data which reminds the data of the real application, but its equation is known. The experiments on this synthetic data verified the accuracy and stability of the method even on very noisy data.

The influence of the number of intervals on the accuracy of the proposed spline approximation is demonstrated in Fig. 5. Two spline functions with different number of intervals are fitted to the same data set, depicted in Fig. 3a: with two intervals (a) and with six intervals (b). It can be seen that it is possible to approximate the points even with small number of intervals. But with bigger number, the spline represents the points accurate. The accuracy of the representation from sub-figures (a) and (b) is illustrated by the residua of the given points in sub-figures (c) and (d). The residuum of the point is the distance of this point from the estimated function. The sub-figure (c) corresponds to the approximation depicted in (a). It can be seen, that the residua "oscillate" around the zero, what means that this representation is insufficient. But in sub-figure (d), corresponding to (b), the residua are only around the zero, what means that this number of intervals is sufficient. Only in the ends the residua are changed. It is caused by two reasons: First, there was bigger noise. Second, the bottom part of the fragment changes its form distinctly and thus, it should need more intervals in this part. The dispersion of the residua (± 1 mm) is the input noise of the data.

The dependence of the accuracy on the number of intervals is depicted in Fig. 6. The figure illustrates the progress of the approximation error (a) and the number of inflex points (b) with increasing number of intervals. For this experiment the same data as in Fig. 3a were used. The approximation error (a) decreases continuously by increasing the number of intervals. It can be seen, that from six intervals the approximation error does not decrease considerably, because it already reaches the input error. For that reason, the increasing number of intervals is inconvenient, because it does not change the error, but it increases the computational difficulty. The progress of the number of inflex points is demonstrated in (b). It should be noted that the number of inflex points need not change continuously not even regularly. Thus, it does not characterize the accuracy of the representation conveniently.

6 CONCLUSION AND OUTLOOK

In this paper, the spline approximation of the large set of points was presented. This approach started with a brief overview of possible representations of the set of points and then the spline approximation method was described in details. The approximation problem is based on a least squares minimization under the smoothness conditions. The method was verified on various data sets, which confirmed the accuracy of the method and its advisability for representation of the set of points.

Many other methods for representation of the set of points by a spline function already exist. But for the archaeological task, the priority is the quickness of the method and preserving the essential characteristics of the points. The spline approximation covers excellent this conditions.

There are many possibilities for improving of this approach. The same lengths of the intervals are used, but for some applications should be convenient the various interval lengths. The mentioned criterion of the representation, the number of inflex points, is not good enough to determine the quality of the representation, so another task can be to find the suitable criterion. The proposed method is general, and therefore, there are no problems with including these extensions.

7 REFERENCES

[Barre94] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* Society for Industrial and Applied Mathematics, 1994.

[Boor78] C. de Boor. *A practical guide to splines.* Springer-Verlag, 1978.

[Halíř97a] R. Halíř. Estimation of the axis of rotation of fragments of archaeological pottery. In Axel Pinz, editor, *Proc. of the 21th Workshop of the Austrian Association for Pattern Recognition (OEAGM'97)*, pages 175–184, 1997.

[Halíř99] R. Halíř. An automatic estimation of the axis of rotation of fragments of archaeological pottery: A multi-step model-based approach. In V. Skala, editor, *Proc. of the 7th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media (WSCG'99)*, 1999.

[Halíř97b] R. Halíř and J. Flusser. Estimation of profiles of sherds of archaeological pottery. In *Czech Pattern Recognition Workshop (CPRW'97)*, pages 126–130, Czech Republic, Mílovy, February 1997.

[Lawso95] Ch. L. Lawson and R. J. Hanson. *Solving Least Squares Problems.* Number 15 in Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1995.

[Orton93] C. Orton, P. Tyers, and A. Vince. *Pottery In Archaeology.* Cambridge University Press, 1993.

[Punta98] N. V. Puntambekar and A. G. Jablokow. Selection of the number of control points for spline surface approximation. In V. Skala, editor, *Proc. of the 6th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media (WSCG'98)*, 1998.

[Reins67] Ch. H. Reinsch. Smoothing by spline functions. *Numerische Matematik*, 10(3):177–183, February 1967.

[Reins71] Ch. H. Reinsch. Smoothing by spline functions II. *Numerische Matematik*, 16(5):451–454, March 1971.

[Rice87] P. M. Rice. *Pottery Analysis: A Sourcebook.* University of Chicago Press, 1987.

[Sabla93] R. Sablatnig, Ch. Menard, and P. Dintsis. A preliminary study on methods for a pictorial acquisition of archaeological finds. In P. M. Fisher, editor, *Archaeology and Natural Science*, volume 1, pages 143–151. Paul Astroems Foerlag, Gotenburg, 1993.