Automated Segmentation of Archaeological Profiles for Classification *

Martin Kampel and Robert Sablatnig
Vienna University of Technology
Institute of Computer Aided Automation
Pattern Recognition and Image Processing Group
Favoritenstr.9/183-2, A-1040 Vienna, Austria
{kampel,sab}@prip.tuwien.ac.at

Abstract

Classification and reconstruction of archaeological fragments is based on the profile, which is the cross-section of the fragment in the direction of the rotational axis of symmetry. In order to segment the profile into primitives like rim, wall, and base, rules based on expert knowledge are created. The input data for the estimation of the profile is a set of points produced by the acquisition system. A function fitting this set is constructed and later on processed to find the characteristic points necessary to classify the original fragment. The one we propose is based on B-splines or bell-shaped splines.

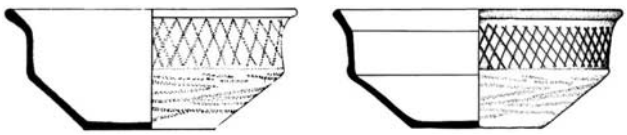
1 Introduction

A large number of ceramic fragments, called sherds, are found at excavations (see Figure 1). These fragments are documented by being photographed, measured, and drawn; then they are classified and stored in boxes and containers. The purpose of *classification* is to get a systematic view on the excavation finds. As the conventional method for documentation is often unsatisfactory [9], we are developing an automated archivation system with respect to archaeological requirements [6], that tries to combine the traditional archaeological classification with new techniques in order to get an objective classification scheme.

A graphic documentation deviced by hand additionally raises the possibility of errors. This leads to a lack of objectivity in the documentation of the material found. To give an example, a vessel was drawn by 2 different illustrators resulting in two different vessels as shown in Figure 2. Note the different shape and decoration, the rim and the thickness for instance are significantly different.

Manual drawings like Figure 2b indicate the beginning and end of defined shape features by horizontal lines. The left half of Figure 2b shows the shape features defined by archaeologists and a subdivision of the profile into intervals. They depict the borders of certain parts of the vessel like rim, wall, and base in this example. By classifying the parts of the profile, the vessel is classified, missing parts may be reconstructed with the expert knowledge of the archaeologist [13]. Segmentation of the profile is done for three reasons: to complete the archive drawing, to classify the vessel and to reconstruct missing profile parts.

Figure 1. Boxes filled with ceramics stored in archives.



(a) (b)

Figure 2. Same vessel drawn by two different illustrators.

Following this manual strategy, the profile should first be segmented into its parts, the so-called *primitives*, automatically. Our approach is a hierarchical segmentation of the profile into rim, wall, and base by creating segmentation rules based on expert knowledge of the archaeologists and the curvature of the profile. The segments of the curve are

*This work was supported in part by the Austrian Science Foundation (FWF), grant P13385-INF, the European Union, grant IST-1999-20273, and the Austrian Federal Ministry of Education, Science and Culture.

1051-4651/02 $17.00 (c) 2002 IEEE

divided by so called segmentation points. If there is a corner point, that means a point where the curvature changes significantly, the segmentation point is obvious. If there is no corner point, the segmentation point has to be determined mathematically.

The curve is characterized by several points. Figure 3 shows the segmentation scheme of an S-shaped vessel as an example. A set of points is defined like,
inflexion point (IP): point, where the curvature changes its sign;
local maximum (MA): point of vertical tangency;
local minimum (MI): point of vertical tangency;
orifice point (OP): outermost point, where the profile line touches the orifice plane;
base point (BP): outermost point, where the profile line touches the base plane;
point of the axis of rotation (RP): point, where the profile line touches the axis of rotation.

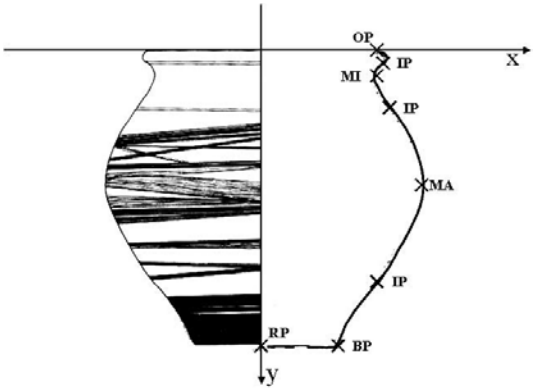


Figure 3. S-shaped vessel: profile segmentation scheme.

By means of these curve points several main segments of a vessel are distinguished: rim, upper part, lower part, neck, shoulder, belly and bottom. On the basis of the number and characteristics of these segments different kinds of vessels can be classified.

2 Automated Segmentation

The profile sections are achieved automatically by a 3D-measurement system based on structured light and a two laser-technique [7]. The profile determined has to be converted into a parameterized curve [12, 5] and the curvature has to be computed [2, 8]. Local changes in curvature [11] are the basis for rules required for segmenting the profile.

Our formalized approach uses mathematical curves to describe the shapes of the vessels and their parts. The profile is thus converted into one or more mathematical curves. We apply four methods for interpolation and four methods for approximation by B-splines on the reconstruction of the vessel profiles (i.e. the profiles are projected into the plane).

2.1 Interpolation by Cubic Splines

The following definitions were adopted from [4]. We suppose that the planar closed curve r to be fitted (interpolated or approximated) will be represented by parametric equations

$$\mathbf{r}(\mathbf{t}) = [\mathbf{x}(\mathbf{t}), \mathbf{y}(\mathbf{t})] \tag{1}$$

in an interval in the Cartesian coordinates of $\mathcal{R}^2$ and has continuous second derivatives. The curve is given by a set of points $P_i = [x(t), y(t)]$ together with the non decreasing sequence of knots $\{t_i, i = 1, \ldots, n+1\}$ of parameter t. Constructing a curve $S(t)$, which approximates the function given by the points can be done by a cubic spline with an adequate parametrization and external conditions. The curve must be initially divided into sub-intervals, where functional approximation and interpolation methods can be applied.

The support of a cubic spline is 5 intervals. Denote by B_i^4 an $k-th$ order spline ($k \leq 3$) whose support is $[t_i, t_{i+4}]$. Then, it is possible to normalize these splines so that for any $x \in [a, b]$

$$\sum_{i=-3}^{n+3} B_i^4(x) = 1 \tag{2}$$

Any cubic spline $S_n(x)$ with knots $t_0, \ldots, t_n$ and coefficients $a_{-3}, a_{-2}, \ldots, a_n$ can be written in the form

$$S_n(x) = \sum_{i=-3}^{n} a_i B_i^4(x) \tag{3}$$

There are $n+3$ coefficients a_i in representation (3) showing that the vector space of cubic splines has dimension $n+3$, so that the $n+1$ functional values will not determine $S_n(x)$ uniquely - two additional constraints must be supplied. Cardinality of the basis has been sacrificed for small support in the basis. Consequently, in evaluating $S(x)$ for any $x \in [a, b]$, only four terms at most in the sum (3) will be non-zero.

The basis cubic splines can be constructed by the following recurrent relationship:

$$B_i^n(x) = \frac{x - t_i}{t_{i+n-1} - t_i} B_i^{n-1}(x) + \frac{t_{i+n} - x}{t_{i+n} - t_{i+1}} B_{i+1}^{n-1}(x), \tag{4}$$

$i = -3, \ldots n-1$ and $n = 1, 2, 3, 4$. A useful convention is to define the first-order splines as *right-continuous* so that

$$B_i^1(x) = \delta_i \quad for \quad x \in [t_i, t_{i+1}), i = -3, -2, \ldots, n+3, \tag{5}$$

The method is of local character: the change of the position of one control vertex influences only 4 segments of the curve. The resulting curve is in particular coordinates a polynomial of $3 - rd$ degree for $t \in (t_j, t_{j+1})$ and has continuous all derivatives in these coordinates.

Since $B_i^n(x)$ is nonzero only on the interval $[t_i, t_{i+4}]$, the linear system for the B-spline coefficients of the spline to be determined, by interpolation or least-squares approximation, is banded, making the solving of that linear system particularly easy.

$$S^4(x_j) = \sum_{i=0}^{n} B_i^4(x_j) a_i = y_j, \qquad j = 0, \ldots, n \quad (6)$$

for the unknown $B-$ spline coefficients a_i in which each equation has at most 4 nonzero entries.

We selected four interpolation methods:

a) Cubic spline interpolation with Lagrange end-conditions ($cs1$) (i.e. it matches end slopes to the slope of the cubic that matches the first four data at the respective end);

b) Cubic spline interpolation with not-a-knot end-condition ($cs2$);

c) Spline interpolation with an acceptable knot sequence ($cs3$);

d) Spline interpolation with an optimal knot distribution ($cs4$). As 'optimal' knot sequence the optimal recovery theory of Micchelli, Rivlin and Winograd [3] is used for interpolation at data points $\tau(1), \ldots, \tau(n)$ by splines of order k;

All the discussed interpolation methods satisfy the Schoenberg-Whithey conditions, i.e. the achieved representation is for the method, the given data and knot sequences unique. These methods were applied to each of the intervals of the curve, and compared from the point of view of their approximation error (least mean square of the differences of the input value and the spline value) on the given data.

We made a surprising observation: Spline interpolation with an acceptable knot sequence in all intervals of all profiles approximated the data with a smaller error than spline interpolation with optimal knot distribution.

We select an 'optimal' method according to the following criteria: The first criterion for selection of the most appropriate interpolation method is the minimal approximation error on the data in the corresponding interval. To exclude ambiguity, the second criterion is applied: minimal length of the knot sequence corresponding to the method. To further exclude ambiguity, the third criterion is applied: the priority of the interpolation method based on the statistical observations. The priority of the methods was achieved experimentally on profiles and their particular intervals and expresses a 'statistical' ordering according to the smallest approximation error over all intervals of the tested profiles.

2.2 Approximation by Cubic B-Splines

Since in the task being solved, the amount of data pairs acquired to describe a vessel or its parts do not always suffice to represent the shape of the vessel reliably, interpolation does not have to be always the appropriate method. From this reason, we compared the approximation methods on representing the overall shape of the whole curve with respect to the interpolation methods.

The following approximation methods were applied and compared:

a) Cubic smoothing spline with the smoothing parameter $p > 0$ ($cs5$);

b) Smoothing spline with the smoothing parameter $tol > 0$ ($cs6$);

c) Least squares spline approximation with the number of knots equal to a half of the amount of the data ($cs7$);

d) Least squares approximation with the number of knots equal to the number of data - degree of the spline in the particular interval, ($cs8$);

3 Results

When the most appropriate interpolation and approximation methods are computed and selected for each of the intervals of the curve, the method with a smaller error (in case of ambiguity, the interpolation method is preferred) is selected for the interval. The approximation error of the representation over the whole curve is computed. This representation is unique and optimal with respect to the above-mentioned criteria. The method was tested on profiles like shown in Figure 4.

Figure 4. Profiles of different fragments.

All interpolation and approximation methods are applied for every sub-interval of the curve after each run of the program. While the curve is generated gradually for each sub-interval of the curve, the overall approximation error is computed. As a result the profile is constructed from the selected methods and is compared to the data set. Table 1 displays the approximation errors for all methods in all intervals of the leftmost profile in Figure 4, including the selected interpolation and approximation methods for the corresponding interval and the selected overall method for the whole profile. The whole data sets contained approximately 350 data points and the length of the whole curve was approximately 400 points.

The most frequently selected interpolation method was $cs1$ and the most frequently selected approximation method

was $cs6$ in our experiments. An interpolation method was preferred in the intervals, where a sufficient number of data with respect to the length of the interval was given. An approximation method was preferred in the intervals, where there was a lack of data. Figure 5 right half shows one example of an automatically segmented pot with the characteristic points detected and the appropriate manual segmentation on the left of Figure 5 .

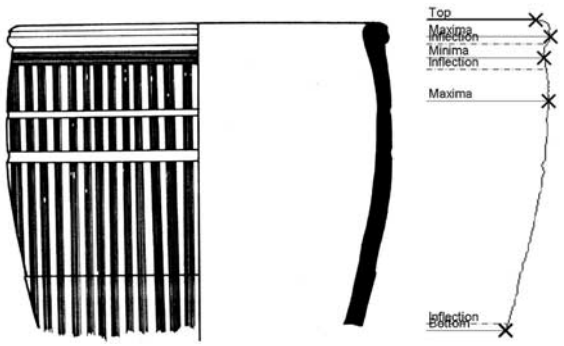
method / interv.	1	2	3	4
cs1	0.2163	0	0.6047	0.0781
cs2	0.2163	0	0.5994	0.0782
cs3	0.2163	0	0.5994	0.0782
cs4	0.2163	0.6169	2.1080	0.0877
cs5 ($tol = 5$)	0.2163	2.3114	0.5994	1.1816
cs6 ($p = 1$)	0.1350	0	0.6229	0.07812
cs7	0.2163	5.9470	5.5298	0.5015
cs8	0.2163	0.0032	0.6014	0.1308
select. intp.	1	1	2	1
select. appr.	6	6	5	6
overall select.	6	1	2	1
method / interv.	5	6	7	8
cs1	1.1685	2.2497	1.1424	0.0884
cs2	1.1686	2.2514	0.1433	0.0884
cs3	1.1686	2.2514	0.1430	0.0883
cs4	1.4510	2.3485	0.1615	0.0991
cs5 ($tol = 5$)	2.9430	2.2514	2.2073	0.0884
cs6 ($p = 1$)	1.1687	2.2496	0.1646	0.0884
cs7	6.9127	6.2323	0.8617	1.0675
cs8	1.1850	3.8347	0.1430	0.2551
select. intp.	1	1	1	1
select. appr.	6	6	8	6
overall select.	1	6	1	6

Table 1. Approximation errors for all methods in all intervals.

4 Conclusion and Outlook

The method presented for selection of an 'optimal' representation (optimal with respect to the considered methods and selection criteria) of a 2-dim profile of an archaeological fragment computes and displays a unique solution. The achieved fragment representations, the first part of an automated system for classification of archaeological fragments, are the input of the second part, the classification.

The profile parts, the so-called profile primitives, are used to perform the classification. The segmentation (division) into primitives depends on the orientation of the fragment. In order to achieve a unique representation, it is important to set a unique orientation for all fragments. The classification will be solved in the high dimensional real space and therefore the uniqueness and the high precision of the profile representation are very important.



(a) (b)

Figure 5. classified pot, (a) manual drawing, (b) detected characteristic points for primitive classification.

The method has been tested on synthetic and real data with good results. The current task is to meet the archaeological requirements as for the achieved representation.

References

[1] W. Adams and E. Adams. *Archaeological Typology and Practical Reality. A Dialectical Approach to Artifact Classification and Sorting*. Cambridge, 1991.

[2] J. Bennett and J. MacDonald. On the Measurement of Curvature in a Quantized Environment. *IEEE Trans. Computer*, 24:803–820, 1975.

[3] T. R. C.A.Micchelli and S. Winograd. The Optimal Recovery of Smooth Functions. *Numerische Mathematik*, 26:191–200, 1976.

[4] R. DeVore and G. Lorentz. *Constructive Approximation*. Springer, 1993.

[5] Z. Hu and S. Ma. The Three Conditions of a Good Line Parameterization. *PRL*, 16:385–388, 1995.

[6] M. Kampel and R. Sablatnig. On 3d Modelling of Archaeological Sherds. In *Proceedings of International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, Santorini, Greece*, pages 95–98, 1999.

[7] C. Liska and R. Sablatnig. Estimating the Next Sensor Position based on Surface Characteristics. In *15th. ICPR*, pages 538–541, September 2000.

[8] J. Matas, Z. Shao, and J. Kittler. Estimation of Curvature and Tangent Direction by Median Filtered Differencing. In *8th International Conference on Image Analysis and Processing*, pages 83–88, 1995.

[9] C. Orton, P. Tyers, and A. Vince. *Pottery in Archaeology*, 1993.

[10] P. Rice. Pottery Analysis: A Sourcebook, 1987.

[11] A. Rosenfeld and A. Nakamura. Local Deformations of Digital Curves. *PRL*, 18(7):613–620, July 1997.

[12] A. Shoukry and A. Amin. Topological and Statistical Analysis of Line Drawings. *PRL*, 1:365–374, 1983.

[13] C. Sinopoli. *Approaches to Archaeological Ceramics*. New York, 1991.