SemanticArchaeo: A Symbolic Approach of Pottery Classification

C. Maiza and V. Gaildrat
Department of Computer Science, IRIT
UPS, 118 route de Narbonne
31062 Toulouse Cedex 4, France
E-mail: {maiza,gaildrat}@irit.fr

Abstract
During archaeological excavations, one of the most time consuming stages is the treatment of the great number of pottery fragments found on the site. A very complex treatment is the matching of every potsherd with the pottery model it belongs to. This step is based on archaeologists knowledge and usage-based nomenclatures. Studied ceramics are revolution shapes so we first obtain a 2D profile and then we segment every pottery model in characteristic elements by detecting 2D geometric attributes and according to the characteristics defined by the archaeologists. A label is then associated with each characteristic element of models and potsherds so as to ease and speed up the process of matching. In our approach, we have organized the complete objects database in a different way than the archaeologists' classical style so as to facilitate the treatment of the potsherds. In order to perform a matching between a potsherd and a model, the pottery models present into the classification are analysed and picked so as to provide a first level of matching: a symbolic matching. Thus, one can achieve matching using first the symbolic description and in a second way the exact geometry of the objects.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications

1. Introduction
During archaeological excavations, a small number of entire potteries and a great quantity of potsherds are found. The archaeologists have to label every sherd, represent them by a two-dimensional drawing and take different measures (height, diameter, thickness, etc.). Then, they have to classify each sherd in order to find the shape it comes from, consulting voluminous paper catalogs which reference the identified shape models. This research stage is not documented because it is only performed from the gained knowledge of archaeologists, thanks to their field experience. This is why it takes at least one to six hours to match a fragment.

As digitization techniques become affordable, new computerised solutions can help the archaeologists to solve the fragment matching problem. Different approaches have been presented for achieving geometrical matching between fragments and database shape models in [MG05, SMK98]. The major problem is the time needed to perform this matching over a big database, because all the objects must be tested. This is due to the fact that pottery databases are organized according to the archaeologists usage i.e. the objects are in classes that depend on the objects usage.

In the framework of the SIAMA (Système d’Imagerie et d’Analyse du Mobilier Archéologique) [SMM’05] project done in collaboration with archaeologists and historians from the UTAH research group in the UTM (Université Toulouse le Mirail) university, we study well standardized ceramics, called “sigillées” potteries. These objects were mass-produced by molding or turning and are assimilated to revolution shapes. They were produced in different French sites during the first centuries of our era, and were sold all over the Roman world. We focused on a subgroup issued from the site of “La Graufesenque” in the south of France near the town of Millau. These potteries have been produced in a relatively standardized manner, in ovens that contained over 40,000 pieces [BJ86, Mar96]. So, retrieving shape models from “sigillées” fragments is very useful for sites dating.

We dispose of a manifold classification of these vessels...
already done by archaeologists [Dra95, Dö64, Kno19]. And a precise description of all these vessels has been carried out. This description is driven by the study of the pottery profiles, see Figure 1. A profile is segmented into the three principal pottery feature i.e. the base, the wall and the rim. Each principal feature being itself segmented into parts representing curves.

Figure 1: The relevant description of the Dragendorff 27b (done by archaeologists).

In order to speed up the matching stage, we propose to build a new database structure based on the descriptions and the geometrical characteristics of the potteries instead of their usage. One can then realize high level shape-based searches on the database using the descriptions of the potteries. This method consists in removing the classes of objects that cannot correspond to the searched fragment, before achieving geometrical matching between a sherd and the shape models.

We take advantage of studying shapes of revolution to present a segmentation algorithm that works in 2D. We segment the object’s profiles into parts that are labeled. This segmentation must be as conform as possible to the archaeologists’ description. Then, we transpose this segmentation in 3D to the object’s meshes. The new database structure is then inferred from the segmentation of the shape models.

Once that the database is available, the matching process between a complete vessel or a fragment and database objects is divided in two stages:

- **A high-level search** based on a description of some part’s characteristics of the searched object. It eliminates the vessels that cannot correspond to the queried object.
- **A geometrical search** based on the profiles. It finds the shape models that present the best probabilities of matching with the searched object.

In the following part, we are going to present some background methods which tried to solve similar problems. Then in section 3, we will present and detail our approach, followed by our experiments in section 4. Finally, we will conclude and give future ways of research in section 5.

2. State of the Art

Many projects have already tried to make easier the work of archaeologists by providing computerized solutions to some of their faced problems.

In order to manage the great amount of potsherds, some researchers are interested in the estimation of the main characteristics of these rotational shapes, namely the axis of rotation and the 2D profile. Different approaches have been used: an algebraic model of the surface [WOC03], the spheres of curvatures [CM02], a Hough-inspired transformation [YM97, KS03], a multi-step optimization technique using notably M-estimators, circle and line fitting [Hal99, HF97]. Also, two approaches try to imitate the archaeologists’ work by taking advantage of the potteries concentric circular rills [KSM05], or by using a semi-automatic system using genetic algorithms to treat rim-fragments [MTL03].

Once that the profile and the axis of rotation are computed, the fragments are stored into a database. They are used even for reconstruction of potteries (by associating two fragments at a time and aligning their curves [KS04], or by using a Bayesian approach to reconstitute the entire object [WC04a, WC04b]), or for studying the standardization of hand-made potteries [Sim02] and the uniformity of wheel produced potteries [MSKS04].

Sablatnig et al. represent their profiles database as a graph [KSC01], where each profile is segmented into the three principal pottery features (base, wall and rim). They carry out matchings between profiles by applying a similarity measure in this graph.

The 3D Knowledge Project is the only one that permits to search the vessels database by sketching a 2D profile [RLB’01]. But their stored profile curves are excessively simplified: they only use the external profiles and the curves do not contain a lot of points (less than 15 per profile while our profiles contain about 2300 points). This makes their data less realistic. Thereafter, they also segment the profiles into base, wall and rim.

We have previously presented an algorithm that realizes matchings between fragments and model shapes [MG05]. This approach was based on both the use of Implicit Surfaces to obtain a distance metric and Genetic Algorithms in order to find the best possible position relatively to the previous distance measure. We faced a speed problem while browsing the whole database in a dummy manner. The issue is that the database of digitalized objects must not be managed like the archaeologists equivalent one and we have to organize the data in a different manner. Taking advantage from the available high-level description of our potteries we have
considered a database where objects have to be segmented in parts that can be easily classified.

3. Contributions

We dispose of a database that contains twenty-two 3D objects that represents eleven vessels in different scales and from different period (i.e. with significant differences), see Figure 2. These objects have been digitized with a Minolta VI-910 laser scanner in the museum of Millau in south of France.

Each object is represented as a 3D mesh (Figure 3), a 2D inner and outer double-profile (Figure 4) and a textual description of the profile made by an archaeologist (Figure 1).

Preliminary remarks:

- First remark: a classification made by archaeologists always contains pottery descriptions, measures and ratios for complete objects and for fragments. These informations are sometimes stored in digitalized objects databases but never used in a matching process. This is the basic idea of our work: using these descriptions to achieve efficient matchings. We then focused on the way to automatically obtain these descriptions. We need to segment these shapes into meaningful parts that we classify.

- Second remark: most of the matching operations that have to be done in an excavation site are matchings between fragments and complete objects rather then matchings between two complete objects or two fragments. This is why detecting a part as a rim or a base on a fragment and identifying it allow us to discard all the database objects that are not composed the same parts. This could be considered as partial matching, i.e. matching between a complete object and a piece of an object. This research domain is very poorly documented even for non revolution objects. Funkhouser et al. have presented an approach to enable weighted comparisons between two objects and giving a bigger weight to the searched part of an object in [FKS’04]. Suzuki et al. divide every shape into a huge number of parts based on the angles of the normal vectors, then similarity comparisons are carried out between a part and all parts of the database objects [MS05].

- Last remark: 3D vessels are considered to be shapes of revolution: objects that are completely defined by a profile and an axis of revolution. This means that shape segmentation can be done on the profile and reported directly to the 3D object. This avoids us the use of a 3D segmentation method like watersheds [MW99, PRF02], especially knowing the drawbacks of such methods (over-segmentation, noise sensitivity and need of dynamic tolerance when dealing with large data sets).

We segment profiles by studying their curvature plots. And once that the profile is segmented, we find the endpoints of the parts on the 3D mesh, and we segment the 3D object according to horizontal planes passing through these points. We use the inner and outer double profile in order to decrease noise influence.

3.1. Segmentation

One of the important characteristics of a curve is the curvature. The curvature is very useful for analysis and classification of vessel shapes. In 3D space the curvature of curves is nonnegative by definition. However, we can obtain signed curvatures $\kappa(u)$ for planar curves using:

$$\kappa(u) = \frac{x(u)y'(u) - y(u)x'(u)}{[(x(u))^2 + (y(u))^2]^{3/2}}$$

where, dots denote derivatives with respect to the given parameter $u$.

We first tried to use this curvature to segment the profile, see Figure 5. But, our curves were too noisy. Thus, we used
a B-spline curve to smooth the profile curvature as shown in [Far96]. Then, we segment the profile by detecting the inflection points and the extrema of the curvature.

The user is lastly allowed to accept the segmentation as it is; to stick again some parts in case of over-segmentation; to re-segment under-segmented parts; or, to completely define a manual segmentation if the automatic one does not meet his needs. See Figure 6 for an example of a relevant segmentation result for an archaeologist.

3.2. Labeling

Once that a profile has been segmented into parts we can group them into the three principal features, namely the base, the wall and the rim. This allow us to take advantage of the description conventions of the "sigillées" potteries, see Figure 7, then, we label each part with its corresponding name.

There are two steps in the labeling algorithm:

1. Describing each part detected at the segmentation stage, i.e. a curve description like the ones of Figure 1.
2. Merging parts (curves) together to form the base, the wall and the rim. Matching each of the principal features with those provided in the description conventions and labeling it with the corresponding name.

For now, the automatic analysis of the segmentation in order to generate the corresponding labeling is not fully implemented. So, we still have to do it manually.

3.3. From 2D to 3D

We obtain the 3D segmentation of the potteries into their principal features using the labeled segmentation previously computed. All the potteries are placed in the same pose, their bases are parallel to the \( y = 0 \) horizontal plane just as if they were put on a table.

From the base part, we compute two planes. One plane that passes by the two end points of the outer profile curve and is parallel to the \( (y = 0) \) plane. And an other one that passes by the two end points of the internal profile curve. All the triangles that are below the lower plane are grouped to form the object’s base. This separates the base from the...
Figure 7: Some possible rims for the “sigillées” potteries.

Figure 8: A Dragendorff 27 mesh segmentation.

rest of the vessel and creates in the same time a transition part that belongs to the two parts.

We repeat this algorithm for the wall part to obtain the segmentation into wall and rim, as shown in Figure 8.

3.4. Matching
For retrieving a complete pottery in the database, we segment its profile to achieve a first high-level search form its labels. The selected objects are then used in a geometrical search step using the tool presented in [MG05].

For retrieving the location of a fragment relatively to the database model shapes, we first extract the profile curve using a technique from the section 2. Then we apply the same algorithm as for the complete objects (a high-level search, followed by a geometrical one).

4. Experiments
The approach presented in section 3 was implemented in a tool named SemanticArchaeo using Java/Java3d.

We first have segmented the available potteries in 2D and in 3D. We have labeled the segmentations (with a manual verification since it’s not fully operational). Then, we have generated a database that contains all the segmentations with the associated labels for each object. This database should be available at the http://semanticarchaeo.online.fr, that is a PHP website associated with a mySql database. So one can browse the database and try to find matching objects based on their descriptions.

This database is extensible: new potteries have to be segmented and labeled with SemanticArchaeo before they are added to the database.

5. Conclusion
We have proposed in this paper a symbolic approach of pottery classification by taking advantage of the ceramic descriptions established by archaeologists. Taking advantage of shapes of revolution, we presented a 2D segmentation algorithm that works on the potteries profiles. Processing these segmentations with a part-labeling method, we have obtained two levels of potteries descriptions: a simple curve segmentation based on the curvature changes, and a decomposition into the three principal pottery features (base, wall and rim) with a description conform to the archaeologists one. These data are then used to build a database of archaeological vessels that can be queried from our two steps matching algorithm: a high-level search using SemanticArchaeo and a geometrical search using CLAPS.

© The Eurographics Association 2006.
The interactivity on the segmentation of the profile is a powerful characteristic because it allows the user to avoid possible errors due to noise or smoothing.

We have presented an easy way of segmenting the pottery’s three dimensional mesh using the two dimensional segmentation. This allows us to achieve fragment matchings directly in 3D by comparing a sherd to the three principal pottery features using a shape descriptor.

Soon, we will have to improve the robustness and the quality of our segmentation process, especially our labeling algorithm. And as an outlook for further research, we plan to develop a complete pottery classification system by adding methods of profile extraction like those from section 2. Then, we will have to test this system on data issued from an excavation site.

We also plan to build a greater database (currently containing twenty-two objects) with objects that are not only “sigillées” potteries in order to test the efficiency of our segmentation pipeline (to recognise and classify such objects).

References


© The Eurographics Association 2006.