Computer Tools for Archaeological Reference Collections: The Case of the Ceramics of the Iberian Period from Andalusia (Spain)

A.L. Martínez-Carrillo[1], A. Ruiz[1], M.J. Lucena[2], and J.M. Fuertes[2]

[1] Centro Andaluz de Arqueología Ibérica
[2] Departamento de Informática, Escuela Politécnica Superior
Universidad de Jaén
Campus de las Lagunillas, 23071. Jaén, Spain

Abstract. The use of new media in the service of cultural heritage is a fast growing field. The development of dynamic web has changed the concept for sharing information, allowing quick access to data and enabling the contents update through the active participation of users. Building digital heritage requires substantial resources in materials, expertise, tools and cost. Also there is a necessity of reflection to promote forms of electronic publication adapted to the needs of archaeologists. This contribution describes an approach and it main strategic choices followed in the construction of an open system through Internet to access and share archaeological information concerning to pottery shapes.

Keywords: Archaeological pottery, integrated technologies, web system.

1 Introduction

The study and analysis of archaeological ceramics constitutes one of the most frequent activities of the archaeological work, which consists habitually of classifying the thousands of ceramic fragments gathered in the interventions and selecting those that contribute to deduce forms, functions and chronology [7].

The different criteria used in the elaboration of classifications do not contribute to homogenize the analysis of the pottery shapes, since the election of criteria depends on each researcher and moment [9]. Shepard saw three phases in the election of criteria: the study of whole vessels as culture objects; the study of sherds as dating evidence for stratigraphic sequences; the study of pottery technology as a way of relating more closely to the potter, but she did not try to put dates to them.

Chronologically, the most used criteria have been artistically, typological, functional, technological, statistical, and contextual. In the last years there is a growing interest in integrating ceramics into a wide analysis of finds assemblages. This must be the next step in ceramic analysis: having integrated the various aspects of ceramics investigations, paying special interest in the spatial and temporal context.

C. Grana and R. Cucchiara (Eds.): MM4CH 2011, CCIS 247, pp. 51–62, 2012.

In this contribution an on line system for storage, analysis, query and visualization of archaeological pottery is shown. Also this system is created as aggregating tool of pottery shapes (complete vessels or fragments from the rim or the base). This system is a useful tool for integrating all the information concerning to pottery sherds.

This paper is structured as it follows: first of all the spatial and temporal context is defined, then the methodology carried out for the systematization of the semantic and graphical information concerning to archaeology pottery is exposed, next the computer tools used for the realization of the ceramic reference collection are explained and finally the conclusions will be exposed.

2 A Computerized Method for Documentation of Ceramics: The Case of the CATA Project

2.1 The CATA Project: An Introduction

The CATA project (Archaeological Wheel Pottery of Andalusia) is an on line system for classifying and categorizing archaeological pottery.

The main goal of this project is to create a useful working tool in Internet, which would help to solve usual problems that archaeologists normally deal with. As has been mentioned, one of their more normal tasks is to classify thousand of pottery fragments retrieved in each archaeological work and select those ones, which give enough information for inferring shapes functions and chronologies.

The objectives of this virtual Corpus are:

- To achieve general valid protocols for the studying of the wheel-made pottery of Andalusia based on a selected reference array.
- To obtain a database with the pottery collections and the 2D drawings done until the moment and published or stored in archives.
- To develop useful communication channels on Internet to arrange a reference collections for the archaeologists and to make a database where professional archaeologists would be able of adding the results of their archaeological works.

The model is designed to be used as a professional tool in Archaeology; any archaeologist would have the opportunity of comparing or testing the materials of his/her excavation or surveys and obtain levels of similarity concerning the series include in this project.

2.2 Area of Study

The selected ceramic material comes from different archaeological sites located in the East area of Andalusia, specifically from the provinces of Jaén, Granada and Córdoba. Most of the ceramic vessels have been documented in the province of Jaén, since they pertaining to Iberian period (S. V B.C. I A.C.) In this area there is an expanded tradition with respect to the study of ceramic typologies of

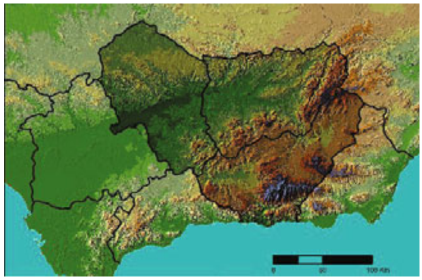

Fig. 1. Studied Area

the Iberian period, emphasizing the work of Pereira [8] for the Iberian ceramics painted of the valley of Guadalquivir, see Fig. 1.

The combination of diverse archaeological sites, with different chronologies makes the accomplishment of a diachronic and synchronous study possible. (Table 1).

Table 1. N^o settlements, chronology and number of studied vessels

N^o of Settlements	**Chronology**	**N^o of Vessels**
2	VII BC - V BC	90
9	IV BC - III BC	829
4	II BC - II AC	255
1	III AC	85
TOTAL		1259

The sample for the analysis has been made in basis of 1259 complete forms whose chronology goes from the VII B.C to the S. II A.C., belonging to 16 archaeological sites from the previously mentioned area of Andalusia.

2.3 The Standardization of the Data

The first step in creating the CATA system is the definition of the protocols required for the analysis and study of pottery vessels found in Andalusia. These protocols are valid and applicable to different historical periods.

The second step is to define the key variables which synthesis the description and definition of vessels and sherds. The four most important variables are based on the work of Orton, Tyers and Vinci [7] date, distribution, functionality and state of preservation.

CATA adds an additional variable to the above:

Measurement variables. Measurements are based on a raster or vector image that is the drawing of the pottery vessel. The following subclassification is used:

- Basic measurements: vessel diameter, height, volume and weight.
- Complementary measurements: these define and numerically specify the most significant parts of the morphology of a vessel (rim, handle and base).

Qualitative variables are related to the manufacturing process of the vessel. Therefore inside this range of variables are included aspects regarding the shaped of the vessel, type of oven treatment, chemical composition of the clay and additives added to the clay. Also is considered the description of the morphology of the vessel, distinguishing rims, handles and bases; surface treatment and chemical analysis.

Preservation variables. The preservation variables are related to the physical state of the vessel (complete or fragmented), alterations suffered by the artifact, and the manufacturing treatments used to create the vessel (used propose to restoration and conservation of a pottery vessel).

Contextual variables. Finally, variables have been added to describe the context in which the artifact is found. A pottery vessel or fragment is associated with a temporal and a spatial context. The identification of the context allows correct dating, deduction of functionality, and the application of geographical significance.

Together with the systematization of the semantic information regarding to the archaeological pottery, graphical information has been standardized through a process of digitalization. This process can be divided into the following steps:

- Digitalization of drawings from archives.
- Vectorization of the contours.
- Separation of the profiles for its later computer processing.
- Export of the drawings of the profiles to a raster format without compression (PNG) for their subsequent comparison.

The sample of the reference collection is formed by drawings of complete vessels, understanding as such drawings of complete profiles or fragment drawings which had enough graphical information to reconstruct the complete section of the vessel, see Fig. 2.

The drawings of the complete vessels come mainly from archives of scientific literature. The documentation available has been compiled to homogenize the

Fig. 2. a) Digitized drawing from a publication; b) Previous image vectorization; c) Vessel's profile

graphical information, which is not standardized and does not follow canons at the time of its study and publication.

Once compiled all the publications in which appear drawings of ceramic vessels it carries out a task of digitalization of these drawings to homogenize the visual reconnaissance of all the drawings, to convert images into vectors and to compress the space of each image for its later computer processing.

3 Computer Tools for on Line Reference Collection

In this section is exposed the main computer tools for the construction of the ceramic reference collection developed in this project.

3.1 Databases

Data archaeology is a skilled human task, in which the knowledge sought depends on the goals of the analyst, cannot be specified in advance, and emerges only through an interactive process of data segmentation and analysis.

Data from archaeological excavation is suitable for computerization although they bring challenges typical of working in non-scientific subjective areas. Meaning and significance within data are established on-site and afterwards by a heuristic process of discussion and contestation, a process at odds with the rigorous demands of database design.

A common and powerful method for organizing data for computerization is the relational data model. Relational databases have a very well-known and proven underlying mathematical theory, which makes possible automatic query optimization, schema generation from high-level models and many other features that are now vital for mission-critical Information Systems development and operations [2].

The database engine selected for CATA system is MySQL due to the facilities given for Web design and the easy implementation with programming language such PHP.

Once the above mentioned variables are defined and clarified, the next step is to create the table schemas using these variables and defining the relation amongst them, see Fig. 3.

In this case has been differentiated between two main levels: a high level, the archaeological site; and an artifact level represented by the pottery vessels. The systematization of the documentation is made using different types of fields for storing and searching numerical values, text, images and videos.

3.2 Image Comparison

There are different methodologies for the estimation of the similarity of vessels profiles. Maiza and Gaildrat [3] propose the use of an automated methodology to establish the relationship of relating a sherd to a known pottery vessel model in basis of 3D models forms and semantic information.

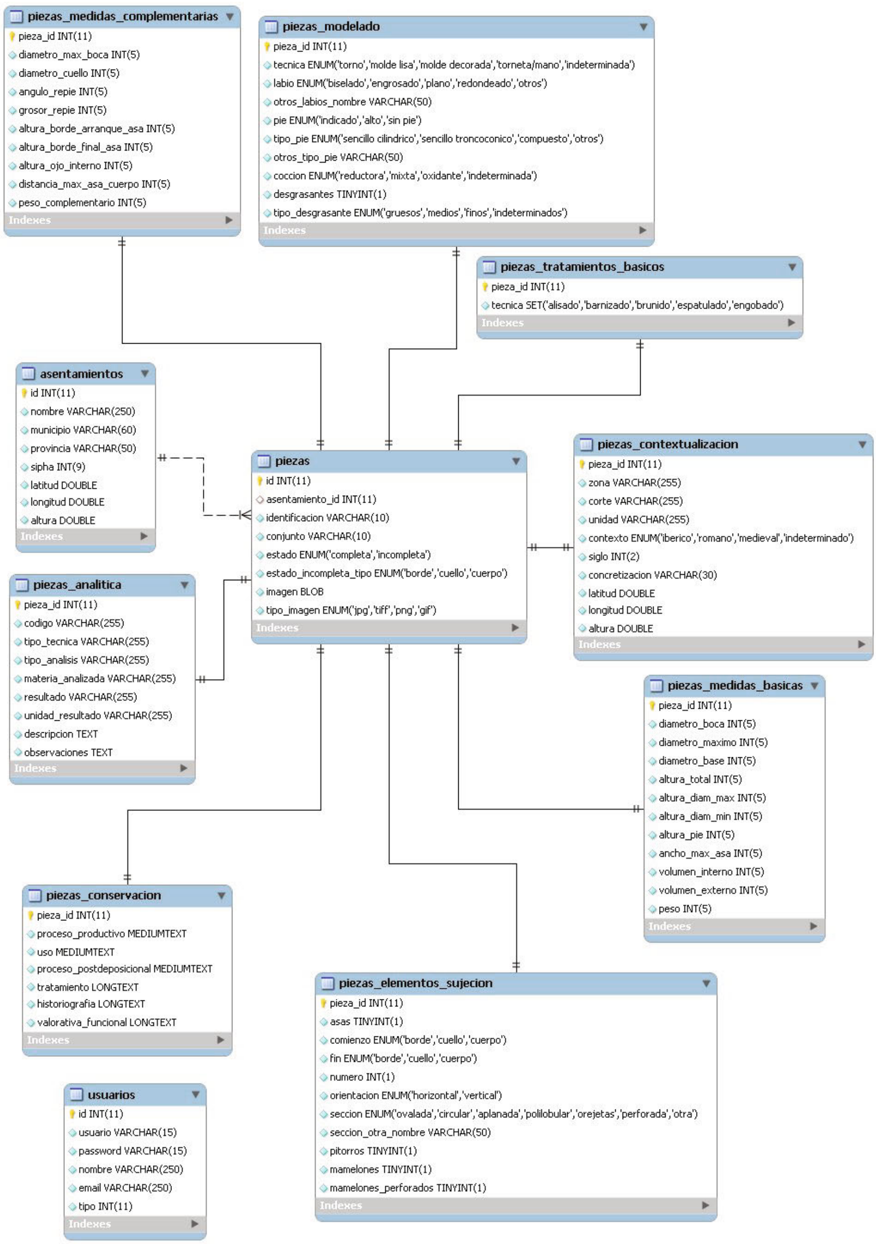

Fig. 3. Table schemas for the relational database used in CATA system

Bishop, Cha and Tappert [1] combine information of shapes forms, textures and colours for obtaining degrees of similarity between ceramics shapes.

In order to estimate the similarity between two profiles, a comparison technique based on non-rigid deformation analysis is designed and developed in CATA project [4].

First of all, a measure to evaluate the effort or deformation energy needed to apply to a given contour in order to adapt it to another is defined. The deformable model given by Nastar [5][6] is used, which is described briefly below.

This model was first used for analyzing the non-rigid motion of structures in temporal sequences of 2D and 3D biomedical images. The mechanical formulation of the problem consists in assuming that the contour is made up of a set of points (or nodes) with mass, joined together by springs. These elastic springs provide a polygonal approach of the contour and are supposedly identical, without mass, with stiffness k and length l_0. These springs modelize the surface elasticity of the object.

For the comparison of the complete profiles, being that is already known the number of points of the simple, a reference prototype is used. Each profile can be classified in relation to its deformation spectrum to the prototype. This way, similarity between two profiles can be computed as the Euclidean distance between their associated deformation spectra.

$$d(D_1, D_2) = \frac{1}{p}\sqrt{\sum_{i=1}^{p}(\tilde{u}_i(D_1) - \tilde{u}_i(D_2))^2} \qquad (1)$$

The prototype C (Figure 4) is the circumference centred in (0.5, 0.5) and that passes through the point (0, 0) subsampled uniformly in N points. All the profiles should be scaled in relationship to the prototype. This makes our measure of deformation invariant to scale.

First the profile P is aligned over the profile to calculate the spectrum, the lowest point of the axis of rotation is aligned with the point (0, 0). Next, P is scaled uniformly so that its highest point corresponds with the edge of the piece that belongs to the circumference C.

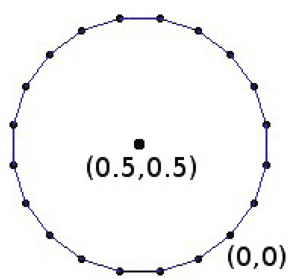

Fig. 4. Circumference C used as prototype figure

Different techniques are designed to establish the correspondence between the points (nodes) of the profile and the points of the prototype. The best results have been achieved when the profile is divided in exterior and interior halves that connect the origin with the edge of the piece. Both curves have been subsampled uniformly in N/2 points (Figure 5).

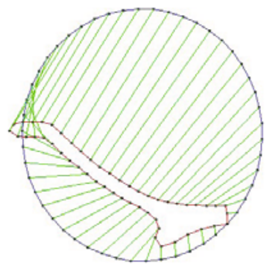

Fig. 5. Correspondence between profile and prototype nodes

Finally some results are shown (Figures 6 and 7).

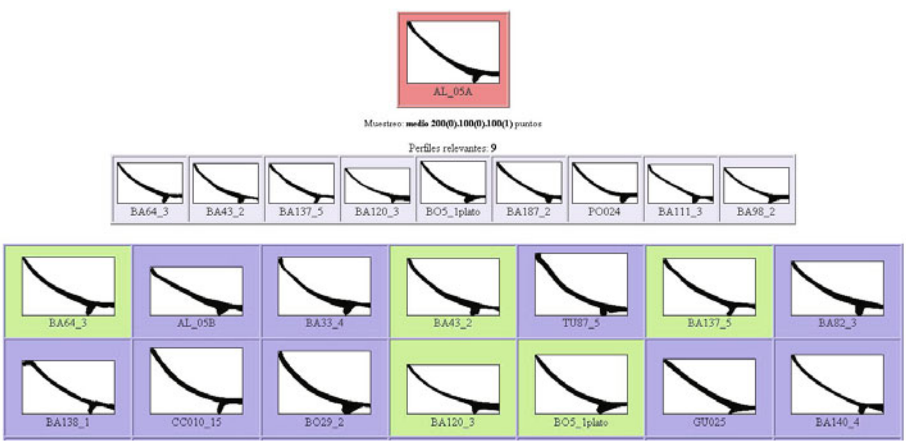

Fig. 6. Output obtained from the system. First row: query profile. Second row: related profiles (ground truth) provided by an expert. Third an fourth rows: first fourteen profiles returned by the system, with the ones belonging to the related profiles set marked in light gray.

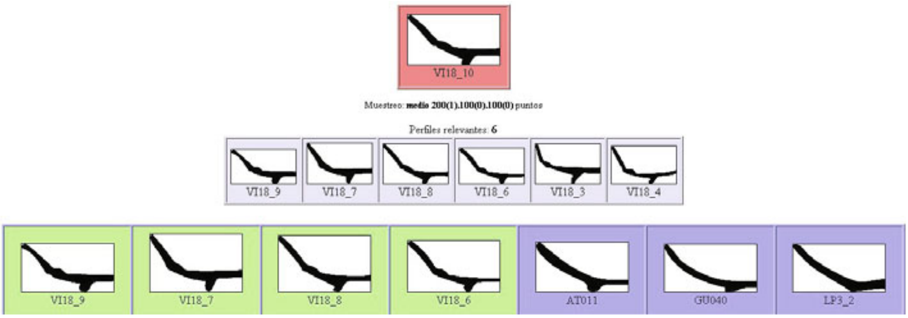

Fig. 7. Output obtained from the system. First row: query profile. Second row: related profiles (ground truth) provided by an expert. Third row: first seven profiles returned by the system, with the ones belonging to the related profiles set marked in light gray.

3.3 User Interface

The Internet interface is oriented on the one hand to implement the computerization tools to storage, retrieval and compare ceramic shapes and by another one to offer to the users an interactive access to the system. The results and the access on line to the application of this project are exposed in the following URL: http://cata.cica.es/.

First of all, different categories of users are defined:

- Invited: they can consult all the information available in the reference collection.
- Registered: they can add and edit information in the system with data of fragments or complete vessels from their own archaeological interventions.

The data uploaded in the system is controlled and validated by an administrator to guarantee the quality of the contents (Figure 8).

Once the user has accessed into the application, the information concerning to the archaeological sites is showed. In this section are exposed the administrative data and their geographical situation through Google maps. In another screen the information of the complete vessels and the fragments is available. The first level of information of the pieces is a list with the identification of each one, their historical period and the site in which they are documented. Also, more specific information within the descriptions of the variables above mentioned and the graphical materials (drawings and videos) is exposed (Figure 9)and (Figure 10).

Besides, it is possible to search information on line. There are two systems for retrieving information:

- Search by predetermined data (chronology, site, type of rim, base or handle and type of surface treatment).

- Search by image profile. The above mentioned module for image profile comparison is implemented in the system.

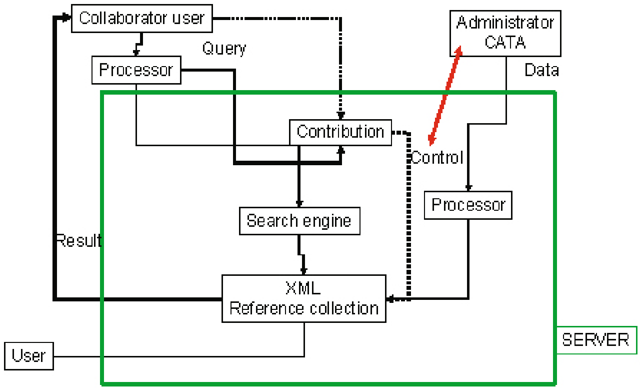

Fig. 8. Schema of CATA

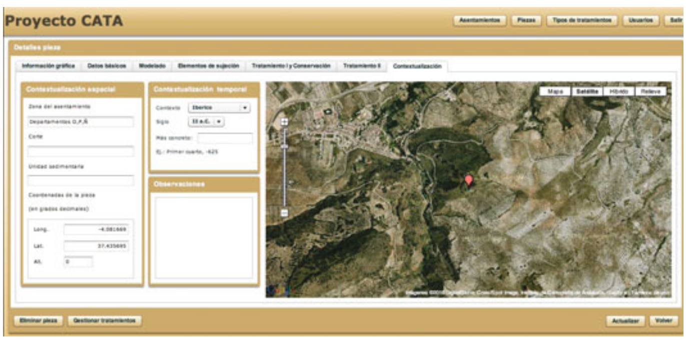

Fig. 9. CATA web application screenshot

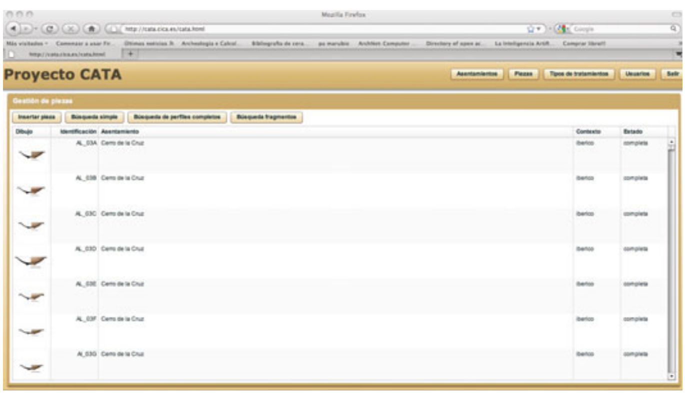

Fig. 10. CATA web application screenshot

Also, this system can be considered a 2.0 on line application, since this type of systems is defined as *all those utilities and services of Internet that contain a database, which can be modified by the users (adding, editing or deleting information or associating data to the existing information* [10].

4 Conclusions

To resume, the investigation of this project is focused on the development of a methodology for the classification, storage and management of archaeological pottery sets. In this sense, the collaboration between archaeologist and computer scientists permits the development of useful applications. There is, however, still a need for further development of information systems specifically targeted at using the full range of applied computation mathematics in archaeological pottery analysis.

This application can be considered an on line reference collection of archaeological ceramic. One of the advantages of the use of this systems is that make possible the unification and standardization of different criteria.

Nevertheless, there must be agreement on common standards for sharing information and the use of controlled vocabularies. This is the reason why standardized data interchange formats should be used and enforced for Internet knowledge transactions. In this direction, it is important to achieve full interoperability though the contents translation into multiple languages and to develop multi-lingual thesauri.

Other challenge is to ensure that e-reference collections must be developed in ways that are suitable for long-term digital preservation.

The work involved in re-purposing reference collections for multiple audiences is not trivial. Reference collections are generally developing by specialists for specialists and may required layers of supporting information to render them comprehensible to general users [10]. In this way it is necessary make available contents on line taking into account the different profiles of all the possible users that can access.

Acknowledgment. This work was has been supported by the Excellent Projects Program of CICE (regional government) and the European Union ERDF funds under research projects P07-TIC-02773 and Project HUM-890 *Corpus Virtual de Cerámica Arqueológica*.

References

1. Bishop, G., Cha, S.-H., Tappert, C.: A Greek Pottery Shape and School Identification and Classification System Using Image Retrieval Techniques. In: Proceedings of Student/Faculty Research Day CSIS. Pace University (2005)
2. Kadar, M.: Data connection and manipulation of archaeological database created in visual environment. In: Proceedings of theInternational Conference on Theory and Applications of Mathematics and Informatics - ICTAMI 2004, Thessaloniki, Greece (2004)

3. Maiza, C., Gaildrat, V.: Automatic Classification of Archaeological Potsherds. In: The 8th International Conference on Computer Graphics and Artificial Intelligence (2005)
4. Martínez-Carrillo, A.L., Lucena, M.J., Fuertes, J.M., Ruiz, A.: Morphometric Analysis Applied to the Archaeological Pottery of the Valley of Guadalquivir. Lecture Notes in Earth Sciences, vol. 124, pp. 307–323 (2010)
5. Nastar, C.: Modèles physiques déformables et modes vibratoires pour l'analyse du mouvement non-rigide dans les images multidimensionelles. PhD thesis, INRIA (1994)
6. Nastar, C., Ayache, N.: Frequency-based nonrigid motion analysis: application to four dimensional medical images. IEEE Trans. on Pattern Analysis and Machine Intelligence 18(11), 1067–1079 (1996)
7. Orton, C., Tyers, P., Vinci, A.: Pottery in archaeology. Cambrige University Press, Cambridge (1993)
8. Pereira, J.: La Cerámica ibérica de la Cuenca del Guadalquivir. I, Propuesta de clasificación. Trabajos de prehistoria, 45 (1989)
9. Shepard, A.: Ceramics for the archaeologist. Carnegie Institution of Washington, Washington D.C. (1956)
10. Yehuda, E., Kvan, T., Affleck, J.: New heritage: new media and cultural heritage. Routledge (2007)