Towards the automatic classification of pottery sherds: two complementary approaches

Chiara Piccoli*
Faculty of Archaeology, Leiden University, Netherlands c.b.m.piccoli@arch.leidenuniv.nl
Prashant Aparajeya
Department of Computing, Goldsmiths, University of London, UK p.aparajeya@gold.ac.uk
Georgios Th. Papadopoulos
Information Technologies Institute, Centre for Research & Technology – Hellas, Greece papad@iti.gr
John Bintliff
Faculty of Archaeology, Leiden University, Netherlands j.l.bintliff@arch.leidenuniv.nl
Frederic Fol Leymarie
Department of Computing, Goldsmiths, University of London, UK ffl@gold.ac.uk
Philip Bes
Faculty of Arts, University of Leuven, Belgium philipmbes@gmail.com
Mark van der Enden
School of Archaeology and Ancient History, University of Leicester mv92@leicester.ac.uk
Jeroen Poblome
Faculty of Arts, University of Leuven, Belgium jeroen.poblome@arts.kuleuven.be
Petros Daras
Information Technologies Institute, Centre for Research & Technology – Hellas, Greece daras@iti.gr

* corresponding author

Abstract
This paper presents two complementary approaches to automatically classify pottery sherds: one that focuses on the sherd's profile and the other that examines visual features of the sherd's surface. The methods are validated using a set of pottery sherds that were collected during surveys at the ancient site of Koroneia (Greece), which were carried out by the 'Ancient Cities of Boeotia' team (under the directorship of Professor J. Bintliff). Both automatic classification techniques produce good results using different sherd classification criteria, such as shape, production technique and chronology.

Keywords
Sherd classification, profile matching, visual feature extraction.

1. Introduction

The automatic classification of pottery sherds has stimulated great interest among archaeologists and computer scientists in recent years (e.g. Makridis 2012, Karasik 2011, Martinez-Carrillo 2011). Pottery is the most common archaeological evidence that is found during fieldwork and pottery specialists are confronted with the time consuming task of processing tens of thousands of sherds to be able to make hypotheses on the chronology, the functional zoning and trade exchanges of the site under study. Comparisons between different archaeological sites and regions are then made possible, thus highlighting differences and commonalities in the economy of the ancient world[1]. This paper presents two automated complementary approaches that compensate and complement one another in the difficult task of pottery classification: one is based on the matching of the sherd's profile and the other one

[1] See the recent contribution of Poblome *et al*. in comparing the Late Roman phases of the Boeotian cities of Tanagra, Thespiae and Koroneia (Poblome et al. 2012), and Reynolds 2010.

exploits the extraction of visual features (such as colour and texture[2] information) from the pottery sherds. The final goals that we aim to achieve are to (1) reduce the time of pottery classification and (2) improve the consistency of the results. Besides being an aid for the pottery specialists, these two approaches can also be used as a training tool for archaeology students in the process of learning how to classify pottery, since these methods are based on the same procedure that archaeologists follow to process sherds.

2. The archaeological context

The pottery dataset that was used for the development of the matching algorithms and techniques that we present in this paper comes from the survey of the ancient city of Koroneia in Boeotia, situated on a hill surrounded by the ridges of Mount Helicon in Central Greece. The study of this site started in 2006 as part of a regional survey, which is carried out since the late 1970s on the whole Boeotia region, under the directorship of Professor John Bintliff. The archaeological traces on the hill suggest occupation phases from Prehistory up to the 14th century when the site was abandoned (Bintliff 2011). The aims of the survey of Koroneia are to (1) map the extent of the ancient city and to identify spatial changes in the settlement over time, and (2) recognise different functional zones in the city, such as domestic, public, and production units. A variety of non-destructive methods are being applied to study Koroneia, such as geophysical analysis, recording of architectural remains and surface collection of pottery sherds (Bintliff et al. 2011).

Pottery, as is the case for other Graeco-Roman sites, is the most abundant source for reconstructing the nature and history of ancient Koroneia. The amount of sherds that is present on the hill is estimated at about 2 million, of which around 100.000 were systematically collected by the end of the survey. The majority of the material that comes from Koroneia is made of fragments of pottery, whose edges have been worn by exposure on the surface and agricultural damage. The problems related with collection and identification of survey pottery have been already discussed elsewhere (Schon 2011; Rutter 1983), and this material constitutes a particularly interesting and challenging dataset to train the matching algorithms presented in this work.

In order to develop the classification algorithms, we selected 203 sherds that had been classified by the pottery specialists. In detail, the distribution of the sherds in the classes assigned by the experts was as follows: regarding shape types, 24,6% were bases, 17,7% were body sherds, 3,5% were handles 53,7% were rims and 0,05% were unclassified; regarding the production technique used, 95,1% were wheel made and 4,9% were hand made; finally, regarding chronology, the sherds have been grouped into Classical-Hellenistic (1,5%), Hellenistic (36,9%), Roman (36,5%), Hellenistic-Roman (7,4%) and unclassified (17,3%). Pictures of the sherds were taken from the front and back side, and from the profile view with an HD camera. The following sections present the two approaches, along with the results of the pilot stage and suggestions for future improvements.

3. Profile matching approach

Our first approach tackles the pottery classification task by taking into consideration the sherd's profile, on the classical assumption that two pots can be differentiated from each other on the basis of their profile structures. This approach uses as ground truth archaeological taxonomic books, in which the profiles of the various pottery shapes are published along with their description. In our case, we used the publications by Susan Rotroff (1997, 2006) on the pottery from the Athenian Agora for the Hellenistic period, and the publication by John Hayes (1972) for the Roman period. The algorithm aims to match the profile of a newly collected sherd to the profiles that are published in the pottery catalogues. To develop the method, we

---

[2] The term 'texture' will be used throughout the article as intended in computer vision, i.e. any visual characteristics of a picture such as colour, visual patterns and shapes. This in contrast to the more specific meaning that texture has in archaeology which mostly refers to the fabric of a sherd.

selected 9 sherds that the pottery specialists had classified by indicating exact matches in the reference books. By matching these sherds from Koroneia with corresponding profiles that are published in the chosen reference books, we hypothesise that it is possible to highlight interesting similarities and differences among pottery productions that can be quantified in a mathematical way.

Recent works propose automatic ways to classify artefacts based on their profile. For example, Durham et al. (1995) used the generalized Hough transform to perform artefact retrieval and matching by edge detection and thresholding. A reference point at the top left corner of an image is chosen to create the feature vectors. The matching process is done in two different ways: the whole shape and part of it is matched, in the latter case with a manual intervention. In a similar fashion, Mara et al. (2002, 2003) developed a system for sherd classification based on a Hough-inspired method where the curvature properties of the object were used. Mom and Paijmans (2005, 2008) designed a tool named SECANTO that considers the sum of squared distances between the contour of vessels to perform the comparison by measuring dissimilarities and finding “look-alikes”. Gilboa et al. (2004) developed a mathematical and computational tool for morphological description, classification and analysis of archaeological artefacts. In their approach, first a curvature function is defined for each fragment and then compared by measuring their relative distance. Hristov and Agre (2013), Karasik (2010), Kampel & Sablatnig (2007), and Maiza & Gaildrat (2005) present other recent profile-based automated pottery classification approaches. Our new method is designed to be applicable to large datasets and be robust to noise, deformations and even deal with partial shape matching, all topics which prove very challenging for these other previous works.

The strategy we adopted is inspired by the Scale Invariant Feature Transform (SIFT) method of Lowe (1999, 2004) and is based on the development of a shape matching algorithm that is invariant to translation, scale and rotation. Importantly, we are not directly using the SIFT method but, instead, apply our method to the shape of 2D objects rather than to the intensity fields of images as is traditionally done with SIFT. In addition, our shape representation is derived from the region based medial point description of shapes, proposed by Kovács et al. (1998). This model is based on human visual perception and how human attention is driven to certain shape characteristics such as corners and salient symmetries. In this paper we propose a possible implementation of Kovács’ model, by developing a method that performs shape matching to pottery sherds’ profiles, but that can be generalized to all kinds of 2D objects and their contours. In order to match the sherd’s profile with the profiles in the reference books, the algorithm exploits feature extraction through a top-hat filter (Vincent 1993) (from mathematical morphology applied to 2D images) and dominant feature points analysis.

3.1 Feature extraction and matching

3.1.1 Medialness measurement

Our feature extraction process is based on the medialness measurement of the pottery sherds’ profile. The purpose of performing the medialness measurement of the object is to provide an effective description of the image which is local and compact and can be easily applied at different spatial scales (Kovács 1998). The goal is to extract the most informative description of the object (or shape), cumulatively, in order to have sufficient information on the object with which to classify it. A medial point is well defined by computing the D_{eps} function, which is based on equidistance. The D_{eps} value at any point in space represents the degree to which this point is associated with a percentage of bounding contour pixels of the object within a tolerance of value *eps* (Kovács 1998) (see Figure 1).

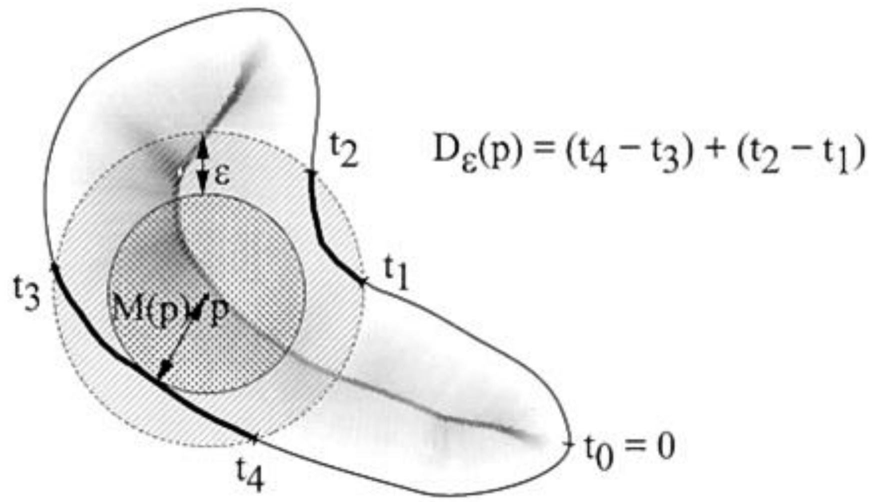

Figure 1: The D_{eps} function for a simple shape (after Kovács 1998, 2325).

The D_{eps} function is defined as the sum of the curve segments falling inside the *eps* neighbourhood (represented as the thick boundary segments within the grey ring) of the *M(p)* radius circle around *p* (Kovács 1998). The mathematical definition of D_{eps} is:

$$D_{eps} = \frac{1}{T} \int_{|p-b|-M(p) \le \varepsilon} db \qquad (1)$$

where *p* = $[x_p, y_p]$, i.e. a point in the image space; *b(t)* = $|x(t), y(t)|$ is a vector in the 2D space of the image describing the bounding contour *B* of the object; *M(p)* is the smallest distance between pixel *p* and the bounding contour. Mathematically:

$$M(p) = \min_{0 \le t \le 1} |p - b(t)| \qquad (2)$$

and *T* is defined as:

$$T = \int_{b \in B} db \qquad (3)$$

which gives the total measure of the set. The medialness measurement is performed for both the internal and external regions of the object. For the exterior medialness measurement, a region restriction has been applied which depends on the parameter R_{MaxMin} (see Figure 2).

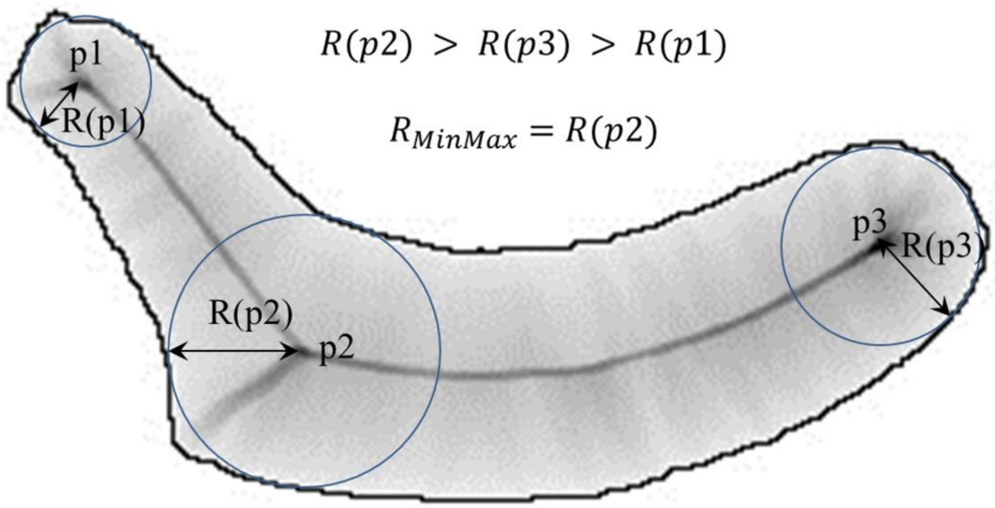

Figure 2: Illustration of maximum of the minimum radial distance (R_{MaxMin}).

The interior medialness reflects in fact the sherd's medial description, while the exterior medialness shows the nature of the concavities that are present on the sherds. Figure 3 shows the internal and external medialness of a sherd.

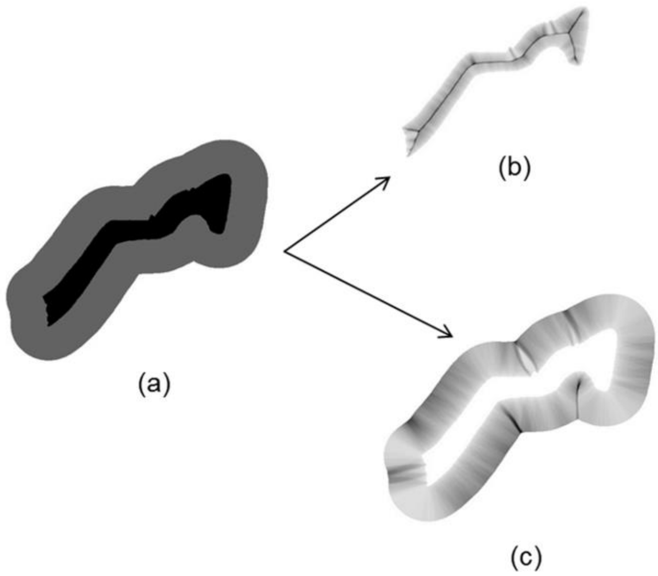

Figure 3: The black region of (a) indicates where the internal medialness will be performed, while the grey region is the restricted region for the exterior medialness measurement. (b) represents the internal medialness and (c) the exterior medialness of the sherd.

3.1.2 Dominant point extraction

The dominant point extraction is a process for identifying the most informative feature points from the medialness image. Dominancy is decided by how many boundary pixels are in the vicinity of a medial point. If a point (p) represents a large amount of edge information (under the tolerance *eps*), then it will be considered as a candidate dominant point. To extract automatically such dominant points, a top-hat transform is used.

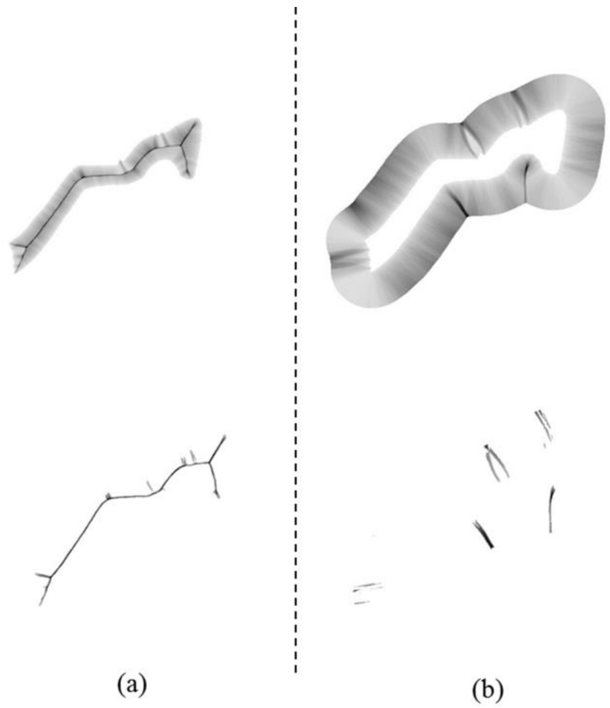

Figure 4: Top-hat transform on the image showing (a) internal medialness and (b) external medialness. The images at the bottom display the candidate feature points.

The top-hat transform is a well-known morphological processes in the field of image processing. Between the two types of top-hat transform (black and white), here (as shown in Figure 4) we chose to use the black top-hat transform to extract the most informative feature points from the medialness image. In order to correctly detect the sherd's profile, the image of its medialness measurement is first filtered by using 'image closing', a morphological transformation that fills the gaps in the image's contours. The top-hat transform is then applied, as the difference of image closing and the original image (of medialness) itself, followed by thresholding the peak values (i.e. discarding the lower peaks considered as noise). Upon application of the top-hat transform, sequences of peaks are generated, which are called dominant points.

3.1.3 Matching process

The matching process uses the dominant points to match the external and internal features of the test case, respectively, to pre-computed external and internal dominant points of the target image. In order to make the matching invariant to scale and rotation, first it is required to find the scale and rotation of the test image with respect to the target image. Scale (β) is defined as the ratio of the minimum radial distance (as defined in medialness measurement) and rotation is the difference of orientations (response direction) of two matching feature points with respect to the positive x-axis. A rotation invariant feature spread has been used to perform the matching task efficiently. If the feature points (p_1, p_2) of the test image match to (q_1, q_2) of the target image, respectively, then the scale and rotation of the image are defined as:

$$\beta = \frac{R\min_p}{R\min_q} \tag{4}$$

[illegible] (5)

The total matching performance is evaluated as the ratio of the total number of dominant points matched (both internal and external) to the total number of dominant points in the test case.

4. Visual features extraction approach

Recently, several approaches for the automatic classification of archaeological sherds have been presented that consider colour and texture information from pictures. In Kampel and Sablatnig (2000), the sherd colour is used, while Smith et al. (2010) employ both colour and texture characteristics. Additionally, texture-based features are used and a profile morphological analysis is performed by Li-Ying Qi and Wang Ke-Gang (2010) and Karasik and Smilansky (2011), respectively. A common characteristic of the above methods is that they are designed to take into account the particular characteristics that are present in the sherd databases that are used for experimentation.

In this section, a novel technique for automatic archaeological sherd classification is presented, which is based on the extraction of colour and textural local features from the sherd surface and the subsequent estimation of a global sherd descriptor vector, using a new 'Bag-of-Words' technique. The method takes into account information from both the front and the back view of the sherd for computing a more complete description. Additionally, a feature selection algorithm is applied in order to maintain the most discriminating features.

Regarding the feature extraction procedure, a combination of relatively simple low-level visual features focusing mainly on the colour properties of the sherds, is used. These were selected after extensive experimentation and aim at handling also low-textured sherds or fragments that present extensive deterioration on their surface. The features employed are:

- Colour components RGB, HSV and YIQ, which are typically present in most visual classification frameworks and also exhibit low computational complexity.
- Standard deviation (Guo et al. 2010), whose histogram is proven to be very efficient in a variety of classification tasks.
- Michelson contrast (Michelson 1927), which is appropriate for modeling faint areas on the sherd surface that are due to deterioration caused by subsoil substances and time.
- Kirsch edge map (Kirsch 1971), whose statistical analysis is used to reveal the degradation degree or rills on the sherd surface.
- Local binary patterns (Ojala et al. 2002) are efficient descriptors designed for texture classification and have been widely used due to their simplicity and rotation-invariance characteristic.

All the above features were selected after extensive experimentation, as they have been shown to outperform typical well performing descriptors that have been proposed in the literature for the analysis of general purpose images, e.g. Scale Invariant Feature Transform (SIFT). To produce a global descriptor vector for every examined sherd image, the 'Bag-of-Words' (BoW) (Csurka et al. 2004) methodology is followed. In more detail, this methodology initially requires a clustering of the computed descriptors, where the estimated clusters include a fixed-size vocabulary of so-called visual words. Subsequently, histograms of the estimated words, that are computed using the constructed vocabulary of visual words and the original descriptors, are used for representing the image content. Typical techniques of this category employ the K-means algorithm for clustering (Sheng et al. 2010; Kandasamy and Rodrigo 2010; Hotta 2009; Chimlek et al. 2010), mainly due to its ease of implementation. However, K-means has increased sensitivity to its initialization and local search strategy. In order to overcome the aforementioned problems, a new technique for creating 'Bag-of-Words' is proposed, using the Reddi multi-thresholding concept (Reddi et al. 1984). In particular, visual 'words' are created by applying multiple thresholding on each local feature's histogram of values. The adopted method of Reddi has the advantage of maximizing the interclass variance between histogram peaks and locating thresholds on histogram valleys. By locating thresholds on histogram valleys, the possible loss of information due to histogram clustering is minimized. More specifically, if we create an image *I* of dimensions *K*x*L* and that each pixel's value belongs in the interval *[0, 255]*, where *0* corresponds to black and *255* to white colour (i.e. grayscale values). The proposed BoW model can then be described by the following steps: initially, histogram extraction for each local feature f_s takes place according to the following equation:

$$h_{f_s}(x)=\sum_{i,j=0}^{K,L} d\left(f_s(i,j)\right),\ x\in[0,255], \quad (6)$$

where $f(i,j)$ is the feature value at pixel (i,j) and *d* is the delta function, which is defined as follows:

$$d\left(f_s(i,j)\right)=\begin{cases}1 & if\ f_s(i,j)=x\\ 0 & otherwise\end{cases} \quad (7)$$

Then, the accumulated histogram AH_{f_s} is estimated for every feature, taking into account all sherds in the used dataset according to the following equation:

$$AH_{f_s}=\sum_{i=0}^{N} h_{f_s}(x), \quad (8)$$

where *N* is the total number of sherd's images. Finally, (Reddi et al. 1984) multi-thresholding is applied to each feature's accumulated histogram AH_f. Visual "words" are created according to the features' values and the estimated thresholds, as described above. Using this transformation, the dimensionality of the final global feature vector is efficiently decreased from its original value (i.e. the number of histogram bins) to the number of the utilized thresholds. After creating the BoW, all local features are concatenated forming a global descriptor vector that describes the whole sherd image.
Having computed the low-level visual description of a given sherd image, a feature selection step is applied for maintaining the most distinctive features and improving the time performance of the overall classification framework. The feature selection techniques that were comparatively evaluated are: a) Correlation-based Feature Selection (CFS) (Hall 2000), b) Chi-Square attribute selection (Jonhson et al. 1994), c) Consistency-based (Liu and Setiono 1996), d) Principal Component Analysis (PCA) (Jolliffe 1986), e) Relief Attribute Selection (Robnik-Sikonja and Kononenko 1997) and e) Support Vector Machines (SVM)-based (Guyon et al. 2002). From the above mentioned methods, the PCA technique led to the best classification performance. To this end, the reported experimental results are computed using PCA for realizing feature selection.
At the final stage of the proposed algorithm, every examined sherd is associated with one of a set of predefined classes/types. To perform this task, a series of different classification techniques were experimentally evaluated. The latter included the application of the following classifiers: a) K-Nearest Neighborhood (KNN) (Aha and Kibler 1991), b) SVM (Burges 1998), c) Naïve Bayes (John and Langley 1995), d) Sequential Minimal Optimization (SMO) (Platt 1999) and e) Simple Logic (Sumner et al. 2005). From the aforementioned classification schemes, the KNN algorithm led to the best overall classification results.

4.1 Reddi multi-thresholding

In this section, the multi-thresholding technique which is used in the 'Bag-of-Words' creation is described in more detail In the literature, there are several histogram-based multi-thresholding techniques which are used mainly for image segmentation or image binarization (single threshold). The method of Reddi et al. (1984) extends the global (binary) threshold method of Otsu (1979), which is one of the most efficient methods for global thresholding, to the multi-thresholding case. The considered criterion consists of the selection of the thresholds so that the interclass variance between dark and bright regions is maximized. The Reddi multi-thresholding technique, which is applied to all selected visual features, except for LPB (LBP histogram has only 24 bins), can be summarized in the following steps: Initially, the number of thresholds *N* is defined. Then, the threshold values are initialized according to the following equation:

$$k_i = \frac{256}{N+1}, \quad (9)$$

where *256* is the range of possible pixel values in a gray scale image. When this algorithm is applied to normalized features, the range of possible values can change from *[0,255]* to *[0,1]*. Subsequently, the following error values are calculated for each threshold k_i :

$$e_1(0,k_1) = \frac{[m(0,k_1)+m(k_1,k_2)]}{2} - k_1$$
$$e_2(k_1,k_2) = \frac{[m(k_1,k_2)+m(k_2,k_3)]}{2} - k_2$$
$$\ldots$$
$$e_n(k_{n-1},k_n) = \frac{[m(k_{n-1},k_n)+m(k_n,255)]}{2} - k_n$$

(10)

where $m(k_i,k_j) = \frac{\sum_{k_i}^{k_j} x p_x}{\sum_{k}^{k_j} p_x}$, k_i and k_j are neighboring thresholds, *x* is the position in the histogram and p_x is the value in this position. Then, new threshold values are calculated according to the following equation:

$$k_i = k_i + \mathrm{round}(e_i) \; [illegible] \quad (11)$$

The overall procedure is repeated until $\max(|e_1|, |e_2|, \ldots, |e_n|)$ [illegible]

A more detailed description of the presented approach can be found in Makridis and Daras (2012), where extensive experiments as well as a comparative evaluation are also given.

5. Results

The profile matching approach based on medialness measurement performs well when the sherd has reasonably recognisable features such as ridges, concavities, and variations of the object's thickness. Figure 5 shows the matching of an actual sherd profile with the profiles labelled in the reference books. If the sherd has less revealing medialness or it is flattened (e.g. body sherds), then the algorithm gives multiple matching locations with a similar matching percentage as shown in figure 6. Although the results shown in figure 6 are promising and demonstrate that in principle it is possible to find good matches between sherds and reference profiles, at this stage manual inspection of the results is essential to verify the archaeological meaningfulness of the matches.

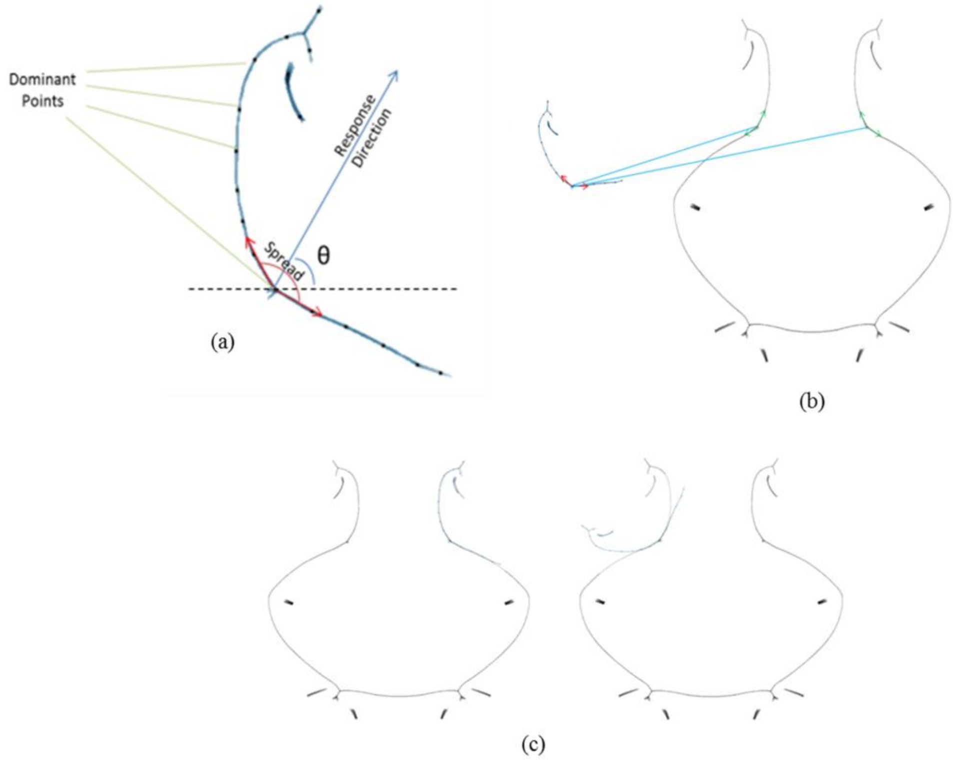

Figure 5: (a) displays the description of features that are extracted using the medialness measurement. The algorithm relies on recognising the dominant points that describe the geometry of the profile. For this reason, there can be two locations in the target image where a dominant point in the test case finds a match, as shown in (b). A following step must be performed that finds the correct rotation, scale and orientation of the test image, in order to find the best matching location, as shown in (c).

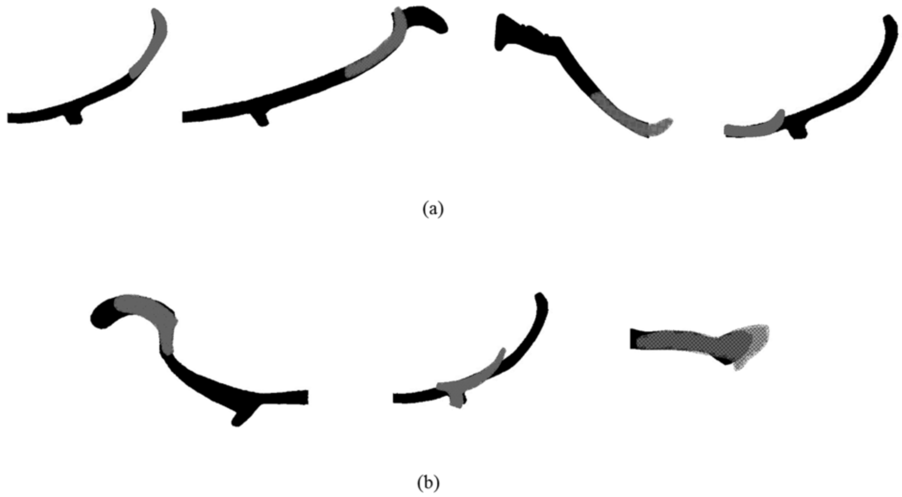

Figure 6: This image shows the matching between the profiles of a number of sherds as they are found in the reference books (in black) and those that were found during the survey at Koroneia (in grey). The first match to the left shows a good correspondence obtained between the sherd and the reference profiles. However, since the sherd has a high degree of flatness, multiple matches are found (a). In contrast, when the sherd has more prominent morphological features, only a single match per sherd is found.

The visual features extraction approach was experimentally evaluated using the already described Koroneia pottery dataset that contains images of approximately 200 sherds, where 25% of the sherds were used for training and the remaining 75% were used for evaluation. Detailed experimental results are given in Tables I-III for three different classification criteria, namely sherd type, production technique and chronology. The obtained classification results are given in the form of the calculated confusion matrices, while the overall classification accuracy (i.e. the percentage of the sherds that were classified correctly) is also given for every case.

Table I: Experimental results for 'Sherd type' criterion

Confusion matrix					
	Associated class				
Actual class		body	base	handle	rim
	body	63,89%	13,89%	0,00%	22,22%
	base	44,44%	11,11%	3,70%	40,74%
	handle	0,00%	0,00%	66,67%	33,33%
	rim	9,88%	2,47%	0,00%	87,65%
Overall classification accuracy: 65,99%					

Table II: Experimental results for 'Production technique' criterion

Confusion matrix			
	Associated class		
Actual class		hand	wheel
	hand	80,00%	20,00%
	wheel	2,78%	97,22%
Overall classification accuracy: 93,96%			

Table III: Experimental results for ‘Chronology’ criterion (CH: Classical-Hellenistic, H: Hellenistic, HR: Hellenistic-Roman, R: Roman)

Confusion matrix					
	Associated class				
Actual class		CH	H	HR	R
	CH	50,00%	0,00%	0,00%	50,00%
	H	3,57%	64,29%	3,57%	28,57%
	HR	0,00%	75,00%	16,67%	8,33%
	R	0,00%	38,89%	1,85%	59,26%
Overall classification accuracy: 55,65%					

6. Discussion and Future Work

The aim of this study was to assist in the manual classification of sherds that is carried out by pottery specialists, by providing an automatized, computer-based classification approach that allows users to obtain consistent results. This paper presents the preliminary results of the two complementary approaches we have developed that take into consideration different characteristics of the sherds, namely, their profile and texture. The approaches were tested using a challenging dataset, i.e. sherds that were collected during surface survey and have therefore deteriorated due to surface exposure and ploughing.

The profile-based approach proposes to tackle the automated classification of pottery sherds by adapting Kovács’ method on shape representation (1998) to the shape matching problem. Our implementation of Kovács’ method uses as ground truth the images of classified pottery shapes from the books of Rotroff (1997, 2006) and Hayes (1972) that are used by the pottery specialists as reference to create the classification of imported pottery wares at Koroneia. For the moment only a small selection of sherds from Koroneia was used to develop the method. In the future a larger set of sherds will be considered and the results will be used to provide a quantified reference of pottery variations by overlapping the profiles of the sherds that were found at Koroneia with published material, thus offering further insight into Koroneia’s pottery assemblages.

The visual features extraction approach exploits a new ‘Bag-of-Words’ technique that overcomes the limitations of traditional BoW methods. This approach exhibits promising classification results for all supported criteria and most defined classes. However, there are also classes that present low recognition rates, i.e. ‘base’ and ‘Hellenistic-Roman’. The ‘base’ class was confused with ‘body’ and ‘rim’, which can be related to the use of pictures as input data for the classification. In fact, in the case of pottery fragments, where only a small fraction of the original complete shape is preserved, the algorithm may not be able to detect the features from the pictures that are relevant to distinguish between e.g., a fragment of a base and a rim. For this reason, attention must be paid to how the pictures are taken; also, the combination of the visual features approach with the profile-based approach that we described could increase the accuracy of the classification. Regarding the ‘Hellenistic-Roman’ class, this was mainly confused with the ‘Hellenistic’ class. The sherds for which the computer-based and the manual classification have returned different classes are currently under scrutiny by the pottery specialists to establish whether this mismatch could provide further insight into the classification of Hellenistic and Roman sherds. We envision that the visual

extraction approach could be used to highlight the sherds for which conflicting classifications exist, and that therefore need to be reconsidered by the pottery specialists.
As future work, the two approaches will be combined in order to fully exploit their complementarity in the classification of sherds. Also, we propose to include pictures of clay fabric to be classified using the visual features extraction approach. In this way, the automatic classification of the sherds will be integrated with crucial information for the identification of local production of pottery and trade exchanges (see Moody et al. 2003). In addition, we will investigate the potential of unsupervised machine learning techniques to cluster the sherds in order to compare the groupings that the algorithm creates with those created by manual classification. The unsupervised machine learning could help in highlighting new groups or in combining previously separated categories. In this way, a more objective classification could be obtained, by limiting the bias inherently connected with manual classifications (e.g. the so-called "Wallcott's shoehorn" as defined by Gould 1989).

7. Acknowledgements

The work presented in this paper was supported by the European Commission under contract FP7-258749 CEEDs and by the CORES network of the Belgian Programme on Interuniversity Poles of Attraction (http://iap-cores.be/).

8. References

Aha D. W., D. Kibler and M. K. Albert. "Instance-based Learning Algorithms." *Machine Learning* 6 (1991): 37-66.

Bintliff, J. "Coronea (Antiquity)." *Encyclopedia of the Hellenic World, Boeotia.* 2011, http://www.ehw.gr/l.aspx?id=12891

Bintliff, J., B. Slapšak, B. Noordervliet, J. Van Zwienen, I. Uytterhoeven, K. Sarri, M. Van der Enden, R. Shield, C. Piccoli. "The Leiden-Ljubljana Ancient Cities of Boeotia Project 2009 season." *Pharos* 17, no. 2 (2011): 1-63.

Burges, C. J. C. "A Tutorial on Support Vector Machines for Pattern Recognition." *Data Mining and Knowledge Discovery* 2 (1998): 121–167.

Chimlek, S., K. Kesorn, P. Piamsa-Nga and S. Poslad. "Semantically Similar Visual Words Discovery to Facilitate Visual Invariance." In *Proceedings of the 2010 IEEE Int. Conf. on Multimedia and Expo*, 1242-1247. 2010.

Csurka, G., C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual Categorization With Bags of Keypoints." In *ECCV International Workshop on Statistical Learning in Computer Vision* (2004).

Durham, P., P. H. Lewis, and S. J. Shennan. "Artefact Matching and Retrieval Using the Generalised Hough Transform." In *Computer Applications and Quantitative Methods in Archaeology 1993*, edited by J. Wilcock and K. Lockyear, 25-30. Oxford: Tempus Reparatum, 1995.

Gilboa A., A. Karasik, I. Sharon, and U. Smilansky. "Towards Computerized Typology and Classification of Ceramics." *Journal of Archaeological Science* 31, no. 6 (2004): 681-694.

Gould, S. J. *Wonderful Life.* New York: W. W. Norton & Co., 1989.

Guo, Z., L. Zhang and D. Zhang. "Rotation Invariant Texture Classification Using LBP Variance (LBPV) with Global Matching." *Pattern Recognition* 43, no. 3 (2010): 706–719.

Guyon, I., J. Weston, S. Barnhill and V. Vapnik. "Gene Selection for Cancer Classification using Support Vector Machines." *Machine Learning* 46 (2002): 389-422.

Hall, M. A. "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning." In *Proceedings of the Seventeenth Int. Conf. On Machine Learning*, 359–366. 2000.

Hayes, J. W. *Late Roman Pottery*. London: The British School at Rome, 1972.

Hotta, K. "Scene Classification Based on Local Autocorrelation of Similarities with Subspaces." In *Proceedings of the 2009 IEEE Int. Conf. on Image Processing*, 2053-2056. 2009.

Hristov V., and G. Agre. "A Software System for Classification of Archaeological Artefacts Represented by 2D Plans." *Cybernetics and Information Technologies* 13, no. 2 (2013): 82-96.

John G. H., and P. Langley. "Estimating Continuous Distributions in Bayesian Classifiers." In *Proceedings of the 11th Conf. on Uncertainty in Artificial Intelligence*, 338-345. 1995.

Jolliffe, I.T. *Principal Component Analysis*. New York: Springer, 1986.

Jonhson, N.L., S. Kotz and N. Balakrishnan. *Continuous Univariate Distributions.* New York: John Willey, 1994.

Kampel M., and R. Sablatnig. "Rule Based System for Archaeological Pottery Classification." *Pattern Recognition Letters* 28, no. 6 (2007): 740-747.

Kampel M., and R. Sablatnig. "Color Classification of Archaeological Fragments." In *Proceedings of the 15th IEEE Int. Conf. on Pattern Recognition*, 771-774. 2000.

Kandasamy K., and R. Rodrigo. "Use of a Visual Word Dictionary for Topic Discovery in Images." In *Proceedings of the Int. Conf. on Information and Automation for Sustainability*, 510-515. 2010.

Karasik A. and U. Smilansky. "Computerized Morphological Classification of Ceramics." *Journal of Archaeological Science* 38, no. 10 (2011): 2644-2657.

Karasik, A. "A Complete, Automatic Procedure for Pottery Documentation and Analysis." In *Proceedings of the Computer Vision and Pattern Recognition Workshops*, 29-34. 2010.

Kirsch, R. "Computer Determination of the Constituent Structure of Biological Images." *Computers and Biomedical Research* 4 (1971): 315–328.

Kovács, I., Á. Fehér, and B. Julesz. "Medial-point Description of Shape: A Representation for Action Coding and its Psychophysical Correlates." *Vision Research* 38, no. 15 (1998): 2323-2333.

Li-Ying Qi and Wang Ke-Gang. "Kernel Fuzzy Clustering Based Classification of Ancient-Ceramic Fragments," In *Proceedings of the 2nd IEEE Int. Conf. on Information Management and Engineering (ICIME)*, 348-350. 2010.

Liu H. and R. Setiono. "A Probabilistic Approach to Feature Selection - a Filter Solution." In *Proceedings of Int. Conf. on Machine Learning*, 319–327. 1996.

Lowe, D. G. "Object Recognition from Local Scale-invariant Features." In *Proceedings of the 7th IEEE International Conference on Computer Vision*, vol. 2, 1150-1157. 1999.

Lowe, D. G. "Distinctive Image Features from Scale-invariant Keypoints." *International Journal of Computer Vision* 60, no. 2 (2004): 91-110.

Maiza, C., and V. Gaildrat. "Automatic Classification of Archaeological Potsherds." In *Proceedings of the 8th International Conference on Computer Graphics and Artificial Intelligence*, 11-12. 2005.

Makridis M., and P. Daras, "Automatic Classification of Archaeological Pottery Sherds." *ACM Journal on Computing and Cultural Heritage* 5, no. 4 (2012): 1-21.

Mara, H., M. Kampel, and R. Sablatnig. "Preprocessing of 3D-Data for Classification of Archaelogical Fragments in an Automated System." In *Proceedings of the 26th Workshop of the Austrian Association for Pattern Recognition "Vision with Non-Traditional Sensors" (ÖAGM)*, 257-264. 2002.

Mara, H., and M. Kampel. "Profile Extraction of Archaeological Fragments for Classification." In *Proceedings of 8th Computer Vision Winter Workshop*, edited by O. Drbohlav, 62-72. 2003.

Martinez-Carrillo, A. L. "Computer Applications in Archaeological Pottery: a Review and New Perspectives." In *On the Road to Reconstructing the Past. Proceedings of the 36th*

Computer Applications and Quantitative Methods in Archaeology Conference. Budapest, April 2-6, 2008, edited by E. Jerem, F. Redő and V. Szeverényi, 397-401. Budapest: Archaeolingua, 2011

Michelson, A. *Studies in Optics*. Chicago: University of Chicago Press, 1927.

Mom, V. “SECANTO–The Section Analysis Tool.” In *The World is in your Eyes. Proceedings of the 33rd Computer Applications and Quantitative Methods in Archaeology Conference. Tomar, March 2005*, edited by A. Figueiredo and G. Leite Velho, 95-102. 2005.

Mom, V., and J. J. Paijmans. “SECANTO: A Retrieval System and Classification Tool for Simple Artifacts.” In *Layers of Perception: Proceedings of the 35th Computer Applications and Quantitative Methods in Archaeology Conference. Berlin, Germany, April 2–6, 2007,* edited by A. Posluschny, K. Lambers, and I. Herzog, 164-170. Bonn: Dr. Rudolf Habelt GmbH., 2008.

Moody, J., H. Lewis Robinson, J. Francis, L. Nixon, L. Wilson. “Ceramic Fabric Analysis and Survey Archaeology. The Sphakia Survey.” *ABSA* 98 (2003): 37-105.

Ojala, T., M. Pietikainen and T. Maenpaa. “Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, no. 7 (2002): 971-987.

Otsu, N. “A Threshold Selection Method from Gray-level Histograms.” *IEEE Transactions on Systems, Man, and Cybernetics* 9, no. 1 (1979): 62–66

Platt, J. “Fast Training of Support Vector Machines using Sequential Minimal Optimization.” In *Advances in Kernel Methods - Support Vector Learning*, edited by B. Schölkopf, C. Burges and A. Smola, 185-208. Cambridge, MA: MIT Press, 1999.

Poblome, J., P. Bes and R. Willet. “Thoughts on the Archaeological Residue of Networks. A View from the East.” In *Rome, Portus and the Mediterranean, Archaeological Monographs of the British School at Rome,* edited by S. Keay, 393-401. London: The British School at Rome, 2012.

Read, D. W., *Artifact Classification. A Conceptual and Methodological Approach*. Walnut Creek, California: Left Coast Press, 2007.

Reddi, S. S., S. F. Rudin, and H. R. Keshavan. “An Optimal Multiple Threshold Scheme for Image Segmentation.” *IEEE Transactions on Systems, Man, and Cybernetics* 14, no. 4 (1984): 661–665.

Reynolds, P. “Trade Networks of the East, 3rd to 7th Centuries: The View from Beirut (Lebanon) and Butrint (Albania) (Fine Wares, Amphorae and Kitchen Wares).” In *LRCW3 Late Roman Coarse Wares, Cooking Wares and Amphorae in the Mediterranean,* edited by S. Menchelli, S. Santoro, M. Pasquinucci and G. Guiducci, 89-114. Oxford: Archaeopress, 2010.

Robnik-Sikonja M. and I. Kononenko. “An Adaptation of Relief for Attribute Estimation in Regression.” In *Proceedings of the 14th Int. Conf. on Machine Learning,* 296-304.1997.

Rotroff, S. I. *The Athenian Agora Volume XXIX. Hellenistic Pottery, Athenian and Imported Wheelmade Table Ware and Related Material*. Princeton: ASCSA, 1997.

Rotroff, S. I. *The Athenian Agora Volume XXXIII. Hellenistic Pottery: The Plain Wares*. Vol. 33. Princeton: ASCSA, 2006.

Rutter. J.B. “Some Thoughts on the Analysis of Ceramics Data Generated by Site Survey.” In *Archaeological Survey in the Mediterranean Area. BAR International Series,* edited by D. R. Keller and D. W. Rupp, 137-142, Oxford: 1983.

Sheng, X., F. Tao, Li Deren and W. Shiwei. “Object Classification of Aerial Images With Bag-of-Visual Words.” *IEEE Geoscience and Remote Sensing Letters* 7, no. 2 (2010): 366-370.

Schon. R. “Vox Clamantis in Campo: Further Thoughts on Ceramics and Site Survey.” in *Our Cups are Full: Pottery and Society in the Aegean Bronze Age,* edited by W. Gauss, M. Lindblom, R. A. K. Smith and J. C. Wright, 231-241, Oxford: ArchaeoPress, 2011.

Smith, P., D. Bespalov, A. Shokoufandeh and P. Jeppson. "Classification of Archaeological Ceramic Fragments Using Texture and Color Descriptors," In *Proceedings of the Computer Vision and Pattern Recognition Workshops*, 49-54. 2010.

Sumner, M., E. Frank and M. Hall. "Speeding up Logistic Model Tree Induction." In *Proceedings of the 9th European Conf. on Principles and Practice of Knowledge Discovery in Databases*, 675-683. 2005.

Vincent, L. "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms." In *IEEE Transactions on Image Processing* 2, no. 2 (1993): 176-201.