Automatic Classification of Archaeological Pottery Sherds

MICHAEL MAKRIDIS and PETROS DARAS, Information Technologies Institute

This article presents a novel technique for automatic archaeological sherd classification. Sherds that are found in the field usually have little to no visible textual information such as symbols, graphs, or marks on them. This makes manual classification an extremely difficult and time-consuming task for conservators and archaeologists. For a bunch of sherds found in the field, an expert identifies different classes and indicates at least one representative sherd for each class (training sample). The proposed technique uses the representative sherds in order to correctly classify the remaining sherds. For each sherd, local features based on color and texture information are extracted and are then transformed into a global vector that describes the whole sherd image, using a new bag of words technique. Finally, a feature selection algorithm is applied that locates features with high discriminative power. Extensive experiments were performed in order to verify the effectiveness of the proposed technique and show very promising results.

Categories and Subject Descriptors: I.4.1 [**Image Processing and Computer Vision**]: Digitization and Image Capture; I.4.7 [**Image Processing and Computer Vision**]: Feature measurement; I.4.9 [**Image Processing and Computer Vision**]: Applications; I.4.10 [**Image Processing and Computer Vision**]: Image Representation—*Statistical*

General Terms: Algorithms, Experimentation

Additional Key Words and Phrases: Bag of words, feature selection, K-nearest neighbor, local binary patterns, archaeological sherds, Classification

ACM Reference Format:

Makridis, M. and Daras, P. 2012. Automatic classification of archaeological pottery sherds. ACM J. Comput. Cult. Herit. 5, 4, Article 15 (October 2012), 21 pages.
DOI = 10.1145/2399180.2399183 http://doi.acm.org/10.1145/2399180.2399183

1. INTRODUCTION

Computer aided restoration of cultural relics is a two-stage process involving classification and reconstruction. Classification is currently performed manually on large quantities of archaeological sherds in order to roughly classify them into groups from the same or similar relics. Reconstruction involves pairwise comparison of sherds using geometrical, color and texture features, so as to obtain the entire object (relic).

Although numerous techniques have been introduced in the relevant literature for object reconstruction, using torn documents [Biswas et al. 2005; de Smet 2008; Justino et al. 2006], jigsaw puzzles [Bunke and Kaufmann 1993; Goldberg et al. 2004; Yao and Shao 2003; Makridis and Papamarkos 2010] and archaeology fragments [Leitao and Stolfi 2002; Papaodysseus et al. 2002; Kampel and

This work was supported by the EC-funded project CEEDS under contract 258749.
Author's address: M. Makridis; email: makridism@gmail.com.
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.
© 2012 ACM 1556-4673/2012/10-ART15 $15.00
DOI 10.1145/2399180.2399183 http://doi.acm.org/10.1145/2399180.2399183

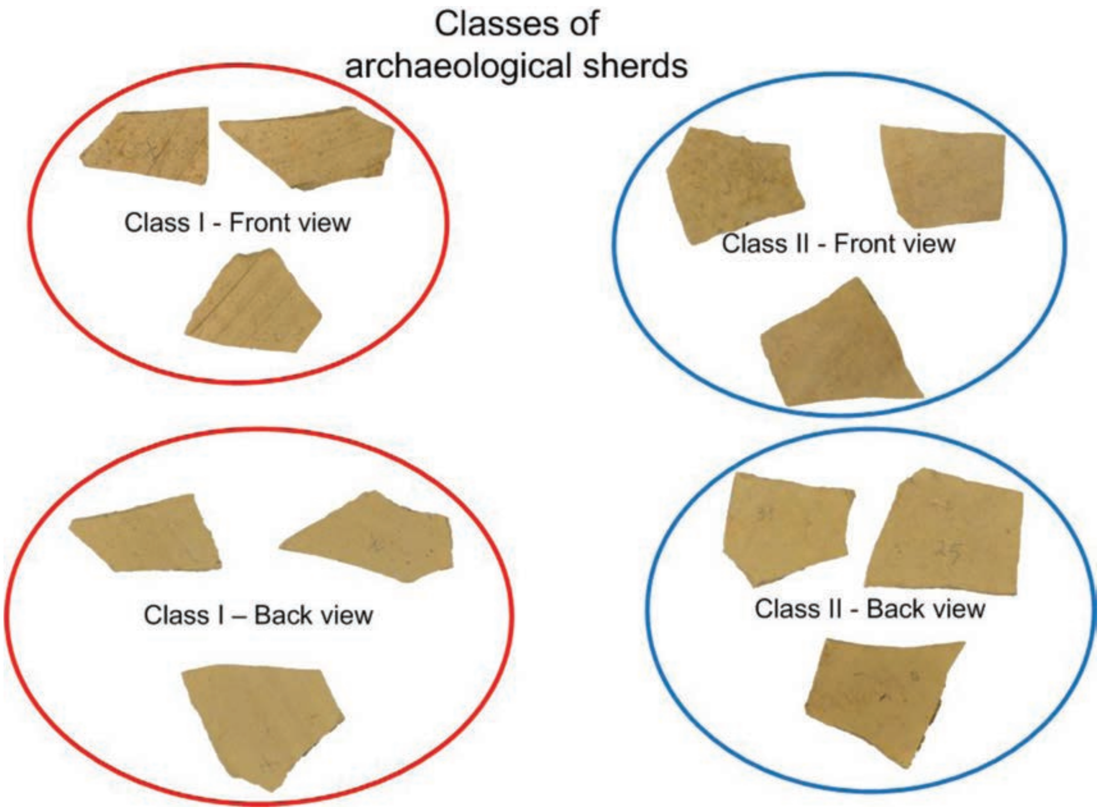

Fig. 1. Images taken from the pottery sherd database. Front and back views of sherd images from two classes are depicted.

Sablatnig 2004], only few techniques addressing the sherd classification problem have been proposed so far. During an archaeological dig, hundreds or even thousands of sherds, constituting fragments of numerous different potteries, can be discovered. These sherds are carefully cleaned and then roughly classified into similar sherd-groups. Each sherd-group contains sherds with similar semantics, for example, pottery of the Byzantine era. Sherd classification is achieved by taking into account mainly color, hue, saturation, chrominance, and material characteristics. Finally, conservators detect matching pairs of sherds in order to reconstruct a relic. This is obviously a highly time consuming task, which may require months of daily effort in order to identify the sherds of a certain object. As a result, millions of sherds are stored into large storage areas and remain unexploited for years.

This article aims at automatic and accurate classification of sherds; a very challenging process, since sherds usually have limited, if any textual information. Most of them have neither graphs, nor figures, nor painted shapes. Their contours are chipped and degraded due to their time exposure, soil substances, erosion and usage. Figure 1 illustrates two sherd classes of the database used in this article. Both front and back views of each sherd are depicted. The visual similarity among sherds belonging to different classes justifies the challenge of automatic and accurate sherd classification. It is hard to distinguish sherds of different classes, as their color is very similar and the differences that appear on the sherds' hue, chrominance, rills (edges), saturation, and texture are only at a local level. Sherds of the first class in Figure 1 are slightly more textured; their front views demonstrate faint rills, while both views appear to have higher saturation than the sherds of the second class. The proposed technique

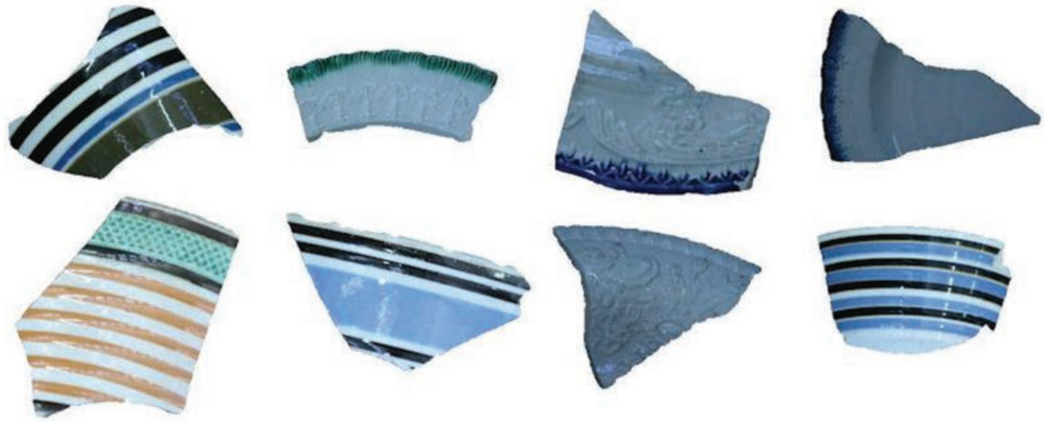

Fig. 2. Indicative images from each class of the ceramic sherd database [Ceramic sherd database 2010], with permission of Drexel Computer Science and NEC Labs.

uses a combination of simple low-level features focusing mainly on the color properties of the sherds. Furthermore, the sherd classification problem is treated statistically in order to achieve accurate classification estimation and simultaneously to avoid classification errors due to local degradation of a sherd. In a nutshell, the proposed technique aims to automate the sherd classification process so as to reduce archaeologists' workloads.

1.1 Background and Related Work

As already mentioned, few works dealing with sherd classification have been proposed so far in the literature. Moreover, each of these methods handles different classification subproblems depending on the data set, for example, potteries, ceramic, textured, nontextured, or mixed sherds, and so on. Furthermore, no existing method exploits both front and back view characteristics of the sherds. Although the front view can describe a sherd more accurately than its back view, the latter can significantly enhance the classification accuracy.

Among the first researchers involved in sherd classification were Kampel and Sablatnik [2000], who classified archaeological sherds based on estimating sherd color so as to automate the classification procedure. Their technique relied on the assumption that the spectral reflectance of materials varies slowly in the visible spectrum. In order to test their color estimation method, they used Tungsten Halogen Floodlamps TL-light, as light resources along with various video cameras. This technique required color calibration with known illuminants and therefore, color estimation was sensitive to lightning variations.

Smith et al. [2010] proposed a method for ceramic sherd classification based on color and texture characteristics. This is the most recent and accurate sherd classification technique and it was applied to a database that consists of ceramic sherds, most of which are highly textured [Ceramic sherd database 2010]. As shown in Figure 2, the test database contains classes with quite different visual patterns. Color similarity between ceramic sherds was based on estimating the joint probability distribution of the color channels [Smith and Chang 1996], and a color histogram in the 3-dimensional RGB space was constructed (Novak and Shafer 1992). Moreover, texture similarity was estimated using a new texture descriptor motivated by the geometric total variation energy (TVG) concept proposed by Burchard [2002]. Finally, a sherd descriptor vector was produced as a combination of TVG and color histograms, which achieved satisfactory results in the textured sherd database [Ceramic sherd

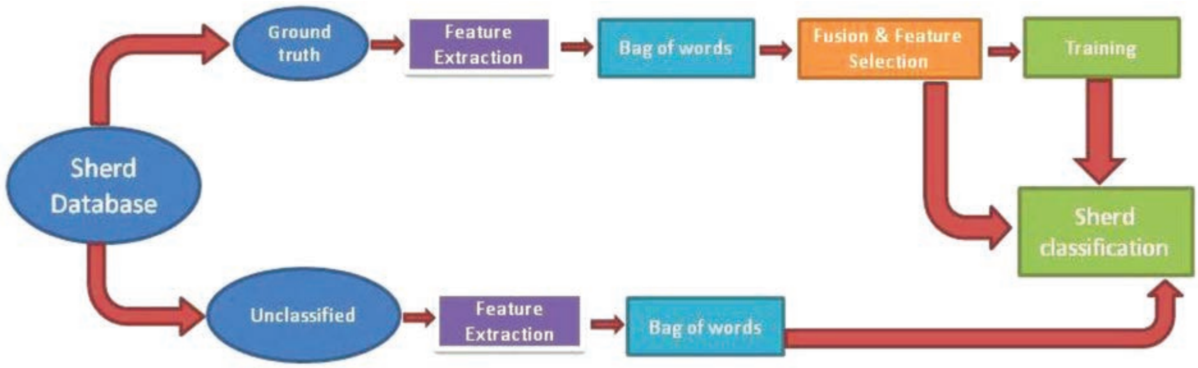

Fig. 3. Block diagram of the proposed technique.

database 2010]. On the other hand, as the authors noticed, their algorithm resulted in generally poor classification rates in low textured sherds with small intensity and color variations.

Li-Ying and Ke-Gang [2010] classified ancient sherds using solely texture-based features, which were extracted by applying Gabor wavelet transformations [Lee 1996]. These features were then used to classify sherds using an unsupervised kernel fuzzy clustering algorithm that is an extension of the fuzzy C-means algorithm [Bezdek et al. 1984]. Ke Gang et al. [2008] proposed another texture-based method for classification of porcelain sherds, where texture features, extracted using the Gabor transformation, were used to distinguish different kinds of porcelain pictures.

Another interesting work on classification of sherds was proposed by Karasik and Smilansky [2011]. This technique performs a morphological analysis and classification of ceramic assemblages based on their profile morphology. However, there are major differences with the proposed technique in this article, due to certain assumptions. Moreover, it is assumed that the assemblages were produced on a wheel and possess axial symmetry. In other words, as the authors state, their shape is entirely characterized by their profiles. It is worth mentioning that Karasik and Smilansky have presented interesting work about the typology, morphology, classification, and 3D scanning of ceramic or pottery assemblages [Adan-Bayewitz et al. 2009; Karasik and Smilansky 2008, 2006; Karasik 2008; Saragusti et al. 2005; Karasik et al. 2005; Gilboa et al. 2004].

In this article, we propose a novel approach for sherd classification using local color-based and texture-based features. It is an attempt to reduce the expert's workload, by automating the steps of the classification procedure. The proposed technique has the following advantages: (a) it performs very well on both high- and low-textured databases; (b) it significantly reduces the computational complexity and enhances the accuracy, by using bag of words and feature selection techniques; and (c) it exploits both front and back views of sherd images, as archaeologists do in their everyday activities.

The rest of the article is organized as follows. An overview of the proposed technique is presented in Section 2. Section 3 presents color and texture features used in the proposed framework. Section 4 introduces the proposed bag of words method used to transform the extracted local features of each sherd into global feature vectors. Feature selection and classification stages are described in Section 5. In Section 6, multiple experiments are given, so as to prove the efficiency of the proposed method. Finally, conclusions are drawn in Section 7.

2. METHOD OVERVIEW

The proposed technique can be summarized in the following steps (Figure 3): (a) local feature extraction from the front and the back views of the archaeological sherds, (b) bag of words creation and feature fusion, in order to form a global sherd descriptor vector that will decrease systems' computational complexity, (c) feature selection that will further increase systems' performance and reduce vectors' dimensionality by keeping only those descriptors that have the highest discriminative power; and (d) classification using at least one ground truth sherd descriptor vector from each class.

More specifically, feature extraction is performed locally for each pixel of each view of a sherd image. Initially, many local features are calculated in order to decide which of them are more appropriate for the specific sherd classification task and lead to the best classification performance. Therefore, midlevel (such as edge map and local binary patterns), low-level (components of color models, standard deviation and contrast), texture-based, and color-based features are extracted.

In general, our aim was to create a global descriptor vector from local features extracted for each sherd image. Local feature extraction on a pixel-by-pixel basis increases the overall computation cost to prohibitive levels. On the other hand, global feature extraction may lead to significant loss of information.

As it is previously mentioned, pottery sherds have certain characteristics that differentiate them from generic images, such as degradation, chipped contour, and low-texture information. Therefore, it is essential for a pottery sherd classification technique to focus mainly on dominant color, chrominance, and texture information, as archaeologists do in their narrative processes. Among the simplest, yet most effective ways of exploiting these characteristics is the use of histograms of local features. Furthermore, histograms of local features have the advantage of being robust to local noise.

In order to deal with the computational complexity problem, we propose a new bag of words (BoW) method. More specifically, words are created using a multithresholding technique over the 256-sized vector of each local feature's histogram. The state-of-the-art multithresholding technique used in this work 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. Finally, the extracted feature vectors are concatenated so as to form a global sherd descriptor vector.

Since global sherd descriptor vectors are formed for all sherd images (ground truth and unclassified), a feature selection technique is applied in order to find features that have the most discriminative power, increase the classification performance and further reduce descriptors' dimensionality. Finally, unclassified descriptor vectors are classified to one of the ground truth classes using a K-nearest neighbor (KNN) classifier [Aha and Kibler 1991]. The KNN classifier was selected due to its ability to perform well in generic image classification problems [Boiman et al. 2008]. This is also proved experimentally in this article (Section 6).

3. FEATURE EXTRACTION

Numerous color- and texture-based local features, both low level and medium level were implemented and presented in this section along with the rationale for the selection.

3.1 Color Model's Components

Each component of a color model has different utility, some of which can be proved efficient and computationally inexpensive when dealing with a specific problem [Agrawal et al. 2011]. Color components are appropriate for sherd classification, since the expected result is a rough classification, based mainly on color, without requiring excessive computational cost. The color model components that were selected as more suitable for describing color, chromaticity, and chrominance of a sherd image are: RGB, HSV, and YIQ.

Transformations from an RGB color model to YIQ and HSV are used to extract the corresponding components and are given in the following equation.

(a) (b)

Fig. 4. Sherds that justify the use of texture features.

$$
\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -0.522591 & 0.311135 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}
$$

$$
H = \begin{cases} 60(b-g), & \text{if } R = M \\ 60(2+r-b), & \text{if } G = M \\ 60(4+g-r), & \text{if } B = M \\ 180, & \text{if } M = 0 \end{cases}
\quad (1)
$$

$$
S = \begin{cases} 0, & \text{if } M = 0 \\ (M-m)/M, & \text{if } M <> 0, \end{cases}
$$

$$
V = M
$$

where $M = \text{maximum}(R, G, B)$, $m = \text{minimum}(R, G, B)$, $r = (M-R)/(M-m)$, $g = (M-G)/(M-m)$, $b = (M-B)/(M-m)$.

3.2 Low Level Features

Standard deviation is a widely used feature in statistical image processing. Although it is susceptible to noise, its histogram [Zhenhua et al. 2010a] can be proved to be a very effective tool in rough classification problems, such as sherd classification. The standard deviation is defined by the following equation:

$$
\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})}, \quad (2)
$$

where n is the total number of pixels neighboring to x_i and $\bar{x}$ is the average pixel value of the pixels in the neighborhood.

Contrast is more appropriate when distinctive objects exist on an image. Although pottery sherd images do not have graphics or symbols on them, degradation due to subsoil substances and time creates faint objects in some sherd classes. Therefore, Michelson contrast [Michelson 1927], which has been adopted by the proposed technique, has proved to be an effective tool for exploiting such information, especially for sherds such as the ones illustrated in Figure 4

As depicted in Figure 4, the skin of the sherd is degraded over time, creating visual color variations on its surface. On the other hand, contrast can lead to misclassification in classes where some sherds

(a) (b)

Fig. 5. Kirsch edge map for a single sherd view.

are degraded while others are intact. Michelson contrast is calculated for the eight neighbors of a pixel according to the following equation.

$$\mathrm{c} = (\mathrm{I}_{\mathrm{max}} - \mathrm{I}_{\mathrm{min}})/(\mathrm{I}_{\mathrm{max}} + \mathrm{I}_{\mathrm{min}}), \tag{3}$$

where I_{max} is the maximum intensity (in gray scale) in the neighborhood of the center pixel, while I_{min} is the minimum intensity (in gray scale) in the neighborhood of the center pixel. If both $I_{max}, I_{min} = 0$ then the contrast is considered as zero, since there is no intensity variation in the neighborhood.

3.3 Medium Level Features

Kirsch edge map [Kirsch 1971] is a first order method [Maini and Aggarwal 2009], generally used for edge detection. In the proposed technique, it is applied on the intensity image of each color sherd image. The intensity image is convolved with filter masks and an edge strength value is assigned to each pixel. Statistical analysis of pixels' edge strength on a sherd image is used to reveal the degradation degree or rills on the image, if present. However, similarly to the contrast feature, the kirsch edge map is not effective in classes, where some sherds are highly degraded while others are not.

Kirsch edge detection is based on image filtering using eight masks corresponding to different directions. Kirsch masks, which detect edges in horizontal and vertical directions at 45, 135, 225, and 315 degree angles, are presented in following.

$$\begin{bmatrix} -3 & -3 & 3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}, \begin{bmatrix} -3 & -3 & 3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}, \begin{bmatrix} 5 & -3 & 3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}, \begin{bmatrix} 5 & 5 & 3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix},$$
$$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}, \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}, \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}, \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \tag{4}$$

For each pixel of the sherd image the largest absolute value is considered as the pixel's edge value. An example of Kirsch edge map is given in Figure 5.

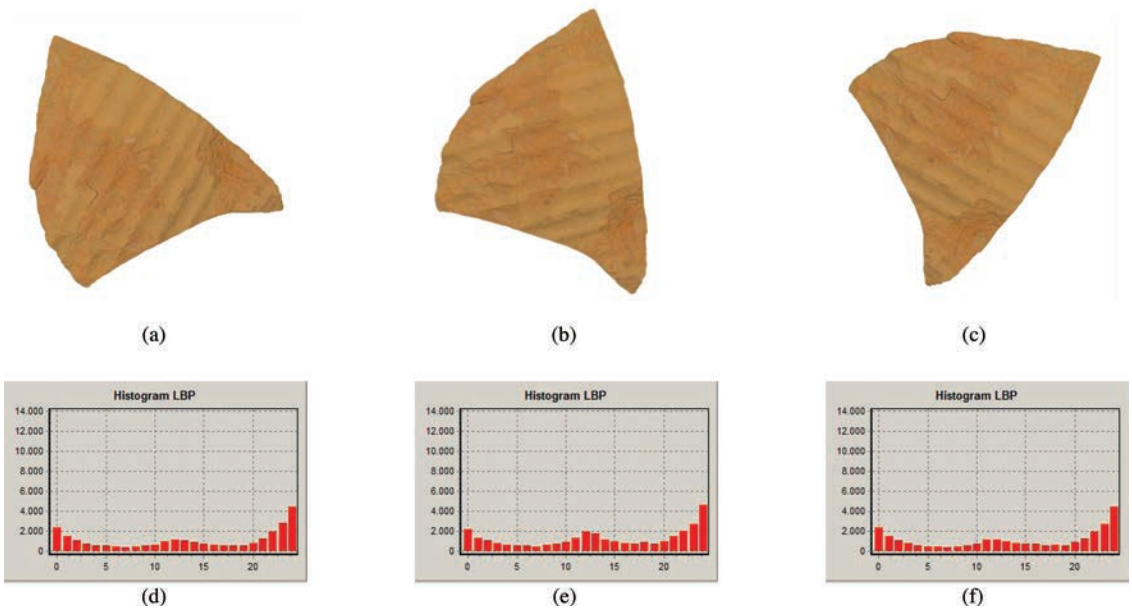

Fig. 6. LBP descriptor vector (histogram) rotation invariance demonstration.

Local binary patterns (LBP), which were first introduced by Ojala et al. [1994], are simple but efficient descriptors designed for texture classification. Since then, various papers have exploited LBP [Khellah 2011; Zhenhua et al. 2010b; Ojala et al. 2002] due to its simplicity and rotation-invariant nature. A fixed-range circular neighborhood is defined around each pixel. Neighboring pixels' intensities are considered as 1 or 0 values, depending upon whether they are higher or lower than the intensity of the center pixel. The local LBP value of the center pixel is calculated according to the number of transitions between ones and zeros around it. Rotation invariance is achieved due to the circular neighborhood. The LBP global descriptor vector is constructed through histogram construction of local LBP values and its length depends on the range of the LBP neighborhood and the distance of the adjacent neighbor pixels. LBP is more efficient in sherds that are highly textured, such as those included in the ceramic sherd database [Ceramic sherd database 2010] and thus is has been included in the proposed technique.

Rotation invariance with LBP can be achieved by shifting bitwise the values or by using uniform patterns. We decided to use the latter since it was proven to be more accurate [Ojala et al. 2002]. Rotation invariance of the LBP descriptor is demonstrated in Figure 6. The original sherd image illustrated in Figure 6(a) was manually rotated in two random angles (Figure 6(b) and (c)). The corresponding LBP histograms are depicted in Figure 6(d)–(f).

4. BAG OF WORDS WITH MULTITHRESHOLDING

Initially, bag of words (BoW) models were applied on words in documents and were related to the frequency of appearance of each word, without preserving the order of appearance in a sentence [Wallach 2006]. Similarly to these models, new bag of words models were soon applied to image features.

BoW models reduce the dimensionality of high-dimensional feature vectors by clustering them into a small fixed vocabulary of visual words, and then using histograms of these visual words to represent the images. In most recent techniques, K-means [Sheng et al. 2010; Kandasamy and Rodrigo 2010;

Hotta 2009; Chimlek et al. 2010] is preferred for clustering data because it is unsupervised and easily implemented. However, K-means has obvious sensitivity in initialization and local searching abilities. In cases where the class centers are not efficiently defined, K-means can get trapped in local minima.

In this article, a new technique for creating bag of words is proposed using the Reddi multithresholding concept [Reddi et al. 1984]. The latter is based on the maximization of the intraclass variances between different classes, which can be seen as valley detection on a histogram. Once maximization is achieved, clusters' range can be easily calculated as the range between neighbor thresholds.

More specifically, let us assume an image I of dimensions K,L, and each pixel values range is [0–255], where the 0 corresponds to black and 255 to white colors (grayscale values). The proposed BoW model can be described by the following steps.

First, histogram extraction for each local feature f_s takes place according to the following equation.

$$h_f(x) = \sum_{i=0}^{K} \sum_{j=0}^{L} d(f_s(i, j)), \quad x = 0, 1, 2, \ldots, 255, \tag{5}$$

where $f(i, j)$ is the feature value in coordinates $f(i, j)$ and d is the delta function.

$$d(f_s(i, j)) = \begin{cases} 1, & \text{if } f_s(i, j) = x \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Then, the accumulative histogram AH_{f_s}, is created for each feature and for all ground truth sherds according to the following equation:

$$AH_{f_s} = \sum_{i=0}^{i=N} h_{i,f_s}(x), \tag{7}$$

where N is the total number of ground truth sherd images, which is equal to the total number of sherd classes (in the case of one ground truth per class).

Finally, Reddi et al. [1984] multithreshording is applied on each feature's accumulative histogram AH. Words are created according to features' values and the extracted thresholds, as described in the preceding.

Using this transformation, the dimensionality of the final global feature vector is reliably decreased from 256 (all histogram bins) to the final number of thresholds. Since we have N ground truth sherd images we define $N - 1$ thresholds leading to N words in each feature's histogram.

After BoW realization, all local features are concatenated forming a global descriptor vector that describes the whole sherd image. A graphical presentation of the proposed BoW technique is depicted in Figure 7.

4.1 Reddi Multithresholding

In this paragraph, we briefly describe the multithresholding technique used in the Bag of Words creation. In the literature, there are several histogram-based multithresholding techniques that 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 powerful methods for global thresholding, to multithresholding. The criterion used is the selection of thresholds so that the interclass variance between dark and bright regions is maximized.

The Reddi multithresholding technique,which is applied to all features used in this article except for LPB the(LBP histogram has only 24 bins), can be summarized in the following steps.

—Define N number of thresholds.

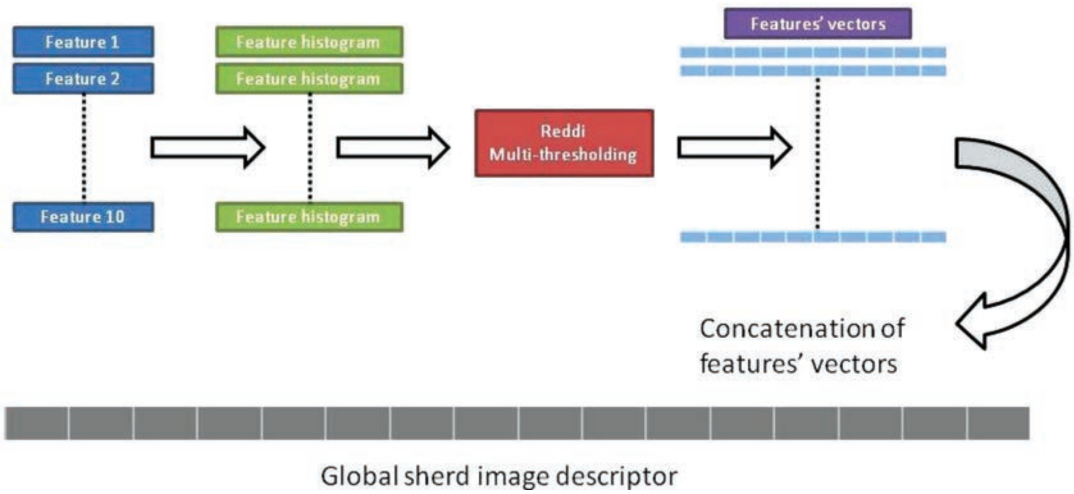

Fig. 7. Graphical presentation of the proposed bag of words procedure.

—Initialize threshold values according to the following equation.

$$k_i = \frac{256}{N+1}, \tag{8}$$

where 256 is the range of possible pixel values in a gray scale image. When this algorithm is applied to a normalized feature, the range of possible values changes from [0,255] to [0,1].

—Calculate the following error values for each threshold k_i.

$$\begin{aligned} e_1(0, k_1) &= \frac{[m(0, k_1) + m(k_1, k_2)]}{2} - k_1 \\ e_2(k_1, k_2) &= \frac{[m(k_1, k_2) + m(k_2, k_3)]}{2} - k_2 \\ &\dots \\ e_n(k_{n-1}, k_n) &= \frac{[m(k_{n-1}, k_n) + m(k_n, 255)]}{2} - k_n, \end{aligned} \tag{9}$$

where $m(k_i, k_j) = \frac{\sum_{k_i}^{k_j} x p_x}{\sum_{k_i}^{k_j} p_x}$, k_i, k_j are neighbor thresholds, x is the position on the histogram, and p_x is the value in this position.

—Calculate new threshold values according to the following equation.

$$k_i = k_i + round(e_i), i = 1, 2, \dots, n. \tag{10}$$

—Repeat until maximum($|e_1|, |e_2|, \dots, |e_n|$) ≤ 0.5.

5. FEATURE SELECTION AND SHERD CLASSIFICATION

Feature selection techniques have been widely used in 2D image classification problems [Chimlek et al. 2010; Kerroum et al. 2009]. In our case, a plethora of features has to be integrated in order to deal with both textured and nontextured databases. On the other hand, only a number of features is usually required, as the selection of the appropriate features depends on the morphology of the sherds.

Table I. Pottery Sherd Database

Classes	Number of sherd images (front and back views)
Class A	46
Class B	22
Class C	16
Class D	8
Class E	8
Class F	10

Therefore, an effective feature selection technique that is able to keep only the most discriminative ones is of paramount importance. This results in increased classification rates, reduced complexity, and increased adaptability of the method. Summarizing, the bag of words technique is the first step in reducing the overall computation cost, while feature selection acts in a complementary fastion to further reduce the complexity and improve the classification performance.

Although feature selection is popular in image classification problems, no reference of feature selection in archaeological sherd classification has been reported so far. In order to choose the most appropriate features, several techniques have been tested. Eventually, the most appropriate one for our problem was found to be the principal component method (PCA) [Jolliffe 1986]. A comparative graph is illustrated in Figure 11 in the experimental results section.

After the feature selection procedure, a feature vector is produced for each sherd view. Final sherd classification is performed using the global sherd feature vector produced by concatenating each sherd's view feature vectors. Although concatenation increases the computational complexity, a sherd's representation by a larger feature vector leads to more accurate classification.

Classification. At the final stage, archaeological sherds are classified using a K-Nearest Neighborhood classifier (KNN) [Aha and Kibler 1991]. The problem addressed by the proposed technique involves only a small set of representative samples for each class. In this case, KNN has low computational complexity. Thus, the rationale behind the choice of KNN is due to its simplicity, low computational complexity, and accuracy [Boiman et al. 2008].

Before concluding that KNN, had the best performance for this problem, the proposed system was tested using various classifiers besides KNN, such as SVM [Chang and Lin 2001], Naïve Bayes [John and Langley 1995], Sequential Minimal Optimization (SMO) [Platt 1998; Keerthi et al. 2001; Hastie and Tibshirani 1997] and Simple Logic [Sumner et al. 2005]. The experimental results that were performed in two sherd databases (Figure 10), pottery and ceramic, justify the selection of KNN.

6. EXPERIMENTAL RESULTS

Several experiments were performed in different areas of the problem so as to provide an objective evaluation of the proposed method.

More specifically, two databases with different characteristics were used. The first is a pottery sherd database, which consists of 110 pottery sherd image views (back and front views), classified in 6 different classes by archaeologists and conservators.[1] Ground truth images (one from each class, front and back views) were also indicated by specialists. Table I summarizes the pottery sherd database indicating the corresponding classes, while Figure 8 depicts the ground truth images from all classes.

Sherd database creation was achieved in cooperation with archaeologists from the 9th Ephorate of Byzantine Antiquities, in Thessaloniki, Greece. All sherds were selected by experts and photographed by the authors. More specifically, a Sony DSC-HX1 SLR-like camera with a resolution of 10 megapixels

[1]Ftp://ftp.iti.gr/pub/incoming/Sherds%20classification%20code.zip.

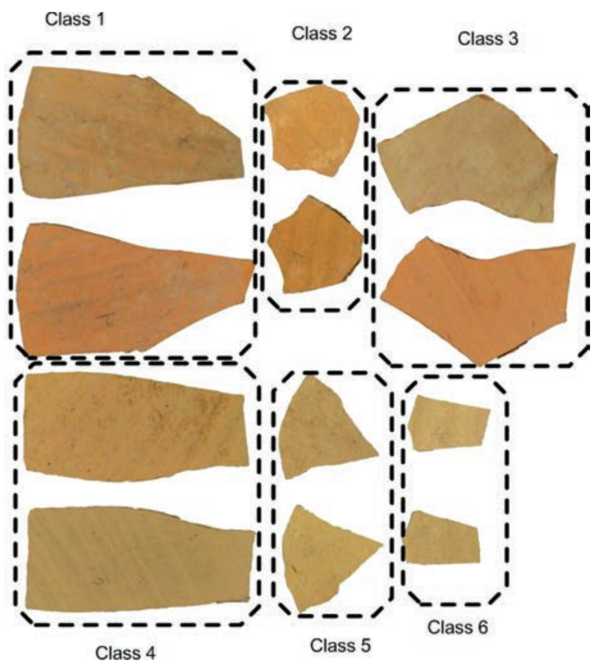

Fig. 8. Ground truth images from the pottery sherd image database.

Table II. Ceramic Sherd Database

Classes	Number of sherds
Class A	9
Class B	9
Class C	9
Class D	16
Class E	2
Class F	10
Class G	7
Class H	18

was used. During the photograph session there were no special parameters except for a camera base in order to keep the camera on a steady distance (10 cm) from the ground.

It should be clearly stated that, the aim of the proposed work is not to replace the experts but to facilitate their work. As it can be assumed, automated classification of archaeological sherds cannot be more successful than manual classification. However, a rough classification with over 50% success rate would be very useful, in terms of speeding up the manual classification procedure. This is exactly the aim of the article.

The ceramic sherd database [2010] a used by Smith et al. [2010], is ideal for evaluation reasons, because it has significant differences from the pottery sherd database. In this dataset, most of the sherds are highly textured, with smooth surface and high contrast. The dataset consists of 80 ceramic sherds classified in 8 classes (Table II). Sample sherds are presented in the following figure.

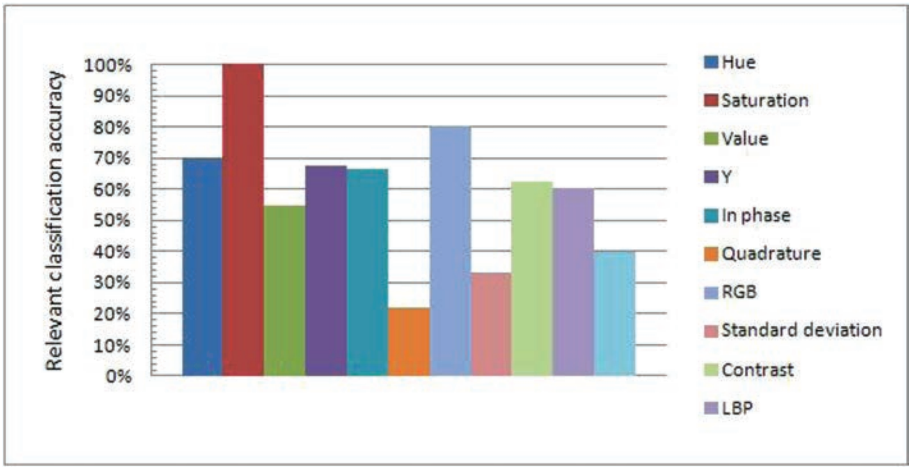

Fig. 9. Comparative mean performance of all features used in the proposed technique. Each feature's percentage is set according to saturation, which had the best performance in the experiments.

Table III. Ceramic Sherd Database

	Images	Resolution (dpi)	Mean (width X height)	Total execution time (sec)
Pottery sherd db	110	200	1206 X 928	227,938
Ceramic sherd db	80	72	357 X 346	28,766

6.1 Advantages of the Proposed Sherd Classification Technique

Relevant performance of features used in the proposed technique. Since several features have been exploited in the proposed framework, it is essential to present comparative performance of each feature separately. Comparative results are presented in Figure 9. It should be noted that some features act in a complementary fashion, while others act competitively. Therefore, a feature selection technique is used to select features with better discriminative power.

Additional texture features were studied (such as those mentioned in Lowe [2004]; Muwei et al. [2009]; Pengyu et al. [2007]; Chamorro-Martinez and Martinez-Jimenez [2009]; and Khellah [2011]), but most of them failed, which justifies why sherd classification depends more on the color rather than the texture.

Results related to execution times of the feature extraction process, for all images, are presented in Table III. It should be noted that in the proposed framework, no code optimization has been applied and the following results are relevant to the implementation and the characteristics of the computer. Experiments have been performed using a Pentium I-5 @ 3.3 Ghz (4GB RAM) and the system was implemented in Pascal Borland Delphi.

Computational complexity reduction using the proposed BoW method: The BoW method takes feature histograms as input and performs clustering on each of them. The number of clusters is defined by the number of final classes. According to this, the vector of each feature is reduced (in percentage) by $100\%(1 - \frac{N}{256})$, where N is the total number of classes, while 256 is the quantity of bins in the histogram. Thus, in the sherd database, the dimensionality of the global feature vector for each ground truth sherd is decreased from $256 * 6 = 1536$ to 36.

Relevant performance of feature selection techniques. Since various tests with different feature selection techniques have been performed, it is worth discussing the comparative performance of the major

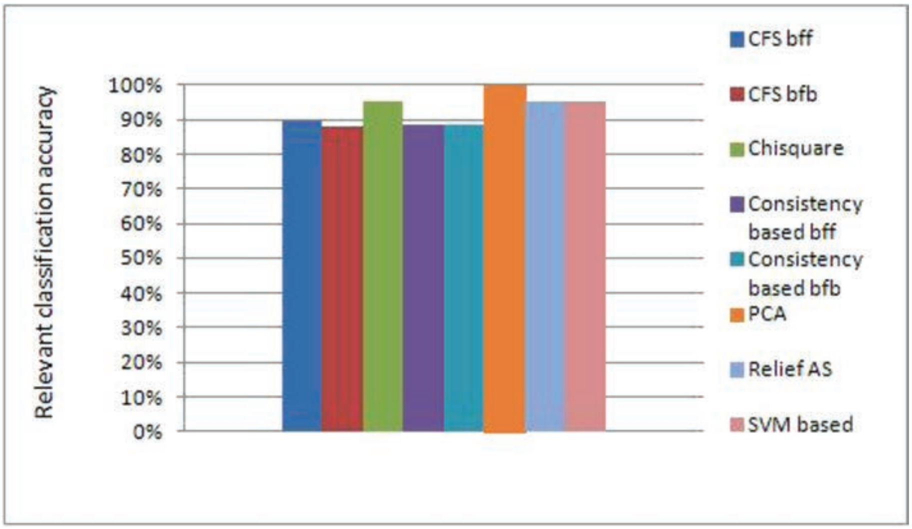

Fig. 10. Comparative mean performance of feature selection algorithms used in the proposed technique. Each algorithm's percentage is according to PCA, which had the best performance in experiments.

techniques used in the proposed article. More specifically, we mention the feature selection techniques tested, as well as the different search strategies used.

—Correlation-based feature selection (CFS bff) [Hall 2000] using best first forward search strategy;
—Correlation-based feature selection (CFS bfb) [Hall 2000] using best first backward search strategy;
—Chi-Square attribute selection (Chisquare) [Jonhson et al. 1994] using ranker search strategy;
—Consistency-based (Consistency based bff) [Liu and Setiono 1996] using best first forward search strategy;
—Consistency-based (Consistency based bfb) [Liu and Setiono 1996] using best first backward search strategy;
—Principal components (PCA) [Jolliffe 1986] using ranker search strategy;
—Relief attribute selection (Relief AS) [Robnik-Sikonja and Kononenko 1997] using ranker search strategy;
—Support vector machines (SVM)-based [Guyon et al. 2002] using ranker search strategy.

In Figure 10, a comparative graph of all feature selection techniques is illustrated. Search strategies were used in order to avoid exhaustive enumeration of all possible feature subsets. The principal components analysis technique (PCA) has been proven to have the best results, in mean terms, for both datasets. More specifically, without using PCA, the overall classification accuracy was 67.74% and 75% in the pottery and the ceramic database respectively. Using PCA, the classification accuracy was increased to 70.96% and 78.26% in the corresponding databases.

Computational complexity reduction due to feature selection. The contribution of feature selection techniques to the reduction of computational cost depends on the number of selected features. Figure 11 depicts the dimensionality reduction of each feature selection technique. In particular, each percentage is the mean performance of each technique in the pottery and the ceramic database. It is obvious that the consistency-based techniques (bff and bfb) achieve the highest dimensionality reduction, while chisquare, relief attribute selection, and SVM-based, very contribute to the overall computational

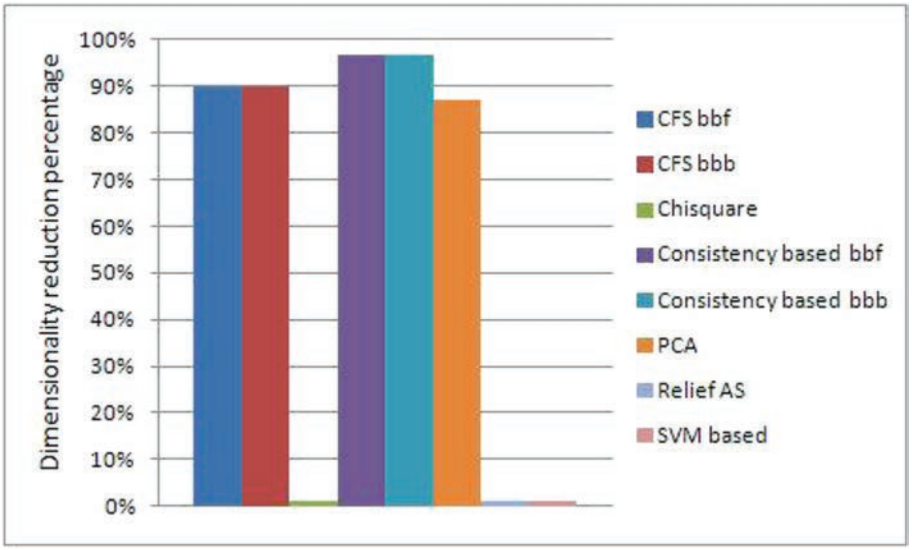

Fig. 11. Descriptor vector dimensionality reduction (in percentage) for each feature selection technique tested.

Table IV. Global Descriptor Dimensionality Reduction after Applying Feature Selection

	Pottery database	Ceramic database
Global descriptor vector length	73	137
Length after applying CFS bbf	6	16
Length after applying CFS bbb	6	16
Length after applying Chisquare	72	136
Length after applying Consistency based bbf	2	5
Length after applying Consistency based bbb	2	5
Length after applying Principal Components	12	13
Length after applying Relief Attribute Selection	72	136
Length after applying SVM	72	136

complexity. The Principal Component analysis method is selected as the most balanced between accuracy improvement (the main purpose of feature selection) and dimensionality reduction. Analytical performance of each feature selection technique for each database is shown in Table IV. All feature selection techniques are executed almost in real time and therefore time-execution results are not presented.

Comparative performance of classification techniques. Classification of the extracted descriptor vectors is performed using KNN. However, four other classifiers were also tested. More precisely, we compared KNN with SVM [Chang and Lin 2001], Naïve Bayes [John and Langley 1995], Sequential Minimal Optimization (SMO) [Platt 1998; Keerthi et al. 2001] and Simple Logic [Hastie and Tibshirani 1997]. Comparative results, which are illustrated in Figure 12, prove that KNN is the most suitable choice for the proposed method.

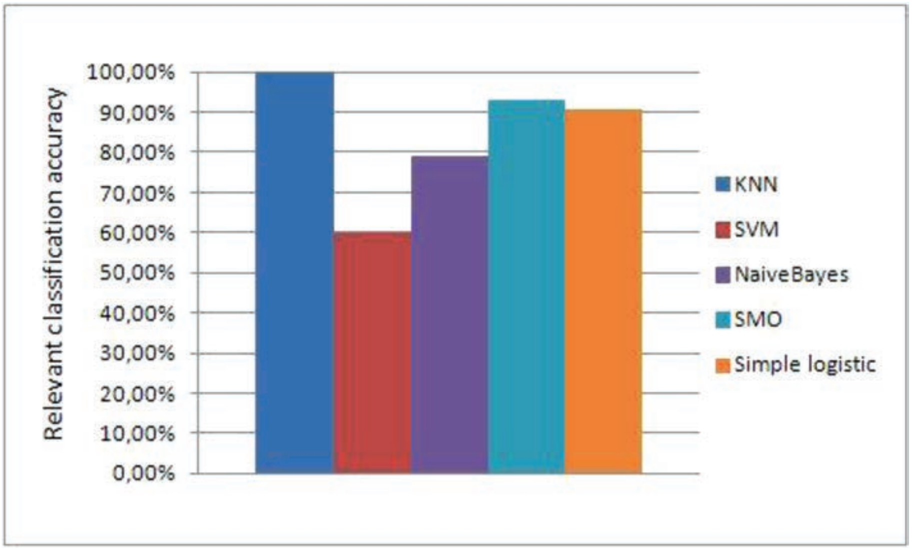

Fig. 12. Comparative mean performance of all classifiers tested on ceramic and pottery datasets, in comparison with KNN.

6.2 Evaluation of the Proposed Technique on the Pottery Sherd Database

The proposed framework has been initially evaluated on the pottery sherd database (Table I). As illustrated in Figure 8, ground truth sherd images from different classes have many similarities. Some differences are detected on sherds' hue, chromaticity, saturation, and small local texture variance. Therefore, any pottery sherd classification technique should focus on these differences in order to discriminate sherd classes. This justifies the use of color related features by the proposed technique. More specifically, concerning the pottery sherd database, the selected features were: in phase (Eq. (1)), contrast (Eq. (3)), Kirsch edge map (Eq. (4)), saturation (Eq. (1)), quadrature (Eq. (1)), and standard deviation (Eq. 2).

Table V presents the confusion matrix, which is also known as the contingency table. The diagonal of the confusion matrix depicts the number of correctly classified sherds. The sum of the diagonals represents the total number of correctly classified sherds. There are 68 correctly classified sherd views, in the proposed experiment. The sum of each row represents the total number of sherd images for each class. Finally, each column represents the number of sherd images that are classified into a certain class, whether they actually belong to this class or not. In this experiment, 12 ground truth sherd views were used for training (one for every class) while 98 sherd views were used for testing. In total, 69.39 percent of the sherds were correctly classified.

In order to evaluate the classification outcome per class, the following measures have been used.

—The *True Positive* (TP rate) rate is the percentage of correctly classified sherds within a class over the entire collection of sherds that belong to that class. True positive is the same as recall.

—The *False Positive* (FP rate) rate is the proportion of sherds that were classified in a certain class, but belong to a different class, among all sherds that are not of this class.

—The *Precision* is the proportion of the sherds that truly belong to a class among all those that were classified in that class

—The *F-Measure* is 2*Precision*Recall / (Precision + Recall), a combined measure for precision and recall.

Table V. Confusion Matrix of the Proposed Technique in the Pottery Sherd Database (Front and Back Views)

	Class A	Class B	Class C	Class D	Class E	Class F	Sum
Class A	26	4	4	2	6	2	44
Class B	2	18	0	0	0	0	20
Class C	0	0	14	0	0	0	14
Class D	0	0	0	2	4	0	6
Class E	2	0	0	0	4	0	6
Class F	0	0	0	4	0	4	8
Correctly Classified	26	18	14	2	4	4	68
Total tested	98	**Success Rate**	69,39%				

Table VI. Evaluation per Class of the Proposed Technique in Pottery Sherd Database (Front and Back Views)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.591	0.074	0.867	0.591	0.703	A
0.9	0.051	0.818	0.9	0.857	B
1	0.048	0.778	1	0.875	C
0.333	0.065	0.25	0.333	0.286	D
0.667	0.109	0.286	0.667	0.4	E
0.5	0.022	0.667	0.5	0.571	F

Table VII. Confusion Matrix of the Proposed Technique in Pottery Sherd Database using 40% as Training Data (Front and Back Views)

	Class A	Class B	Class C	Class D	Class E	Class F	Sum
Class A	18	2	2	0	0	0	22
Class B	4	10	0	0	0	0	14
Class C	0	2	8	0	0	0	10
Class D	0	0	0	2	0	0	2
Class E	2	0	0	0	4	0	6
Class F	2	0	0	0	4	2	8
Correctly Classified	18	10	8	2	4	2	44
Total tested	62	Success Rate	70,97%				

Experimental results' evaluation according to these metrics and classification accuracy are given in Table VI.

In this article, we demonstrate a technique that deals with the sherd classification problem using at least one ground truth sherd image. However, splitting the dataset into train and test samples is used in the majority of the techniques found in the literature. In order to demonstrate the accuracy of the proposed technique in a way that is comparable with the existing literature, we also tested the proposed technique by splitting our dataset into training and test sets. The training percentage usually [Zhenhua et al. 2010b; Lazebnik et al. 2006] varies between 30% and 50%. Therefore, in the proposed technique we perform all experiments by splitting the datasets (ceramic and pottery databases) into 40% training samples (mean value) and 60% test samples. The confusion matrix of results in the pottery sherd database is shown in Table VII.

Table VIII. Confusion Matrix of the Proposed Technique in Ceramic Sherd Database

Classes	A	B	C	D	E	F	G	H	Sum
Class A	5	0	0	0	0	0	0	0	5
Class B	0	5	0	0	0	0	0	0	5
Class C	1	0	6	0	0	0	0	0	7
Class D	0	0	0	4	0	1	3	0	8
Class E	0	0	0	1	0	0	0	0	1
Class F	0	1	0	0	0	2	0	0	3
Class G	0	0	0	0	2	1	1	0	4
Class H	0	0	0	0	0	0	0	13	13
Correctly Classified	5	5	6	4	0	2	1	13	36
Total tested	46	**Rate**	78,26%						

Table IX. Comparative Results with Five Well Known Feature Extraction Techniques

	Ceramic sherd db	Pottery sherd db front views	Pottery sherd db back views
Geomentric blur	43,86%	25,00%	15,63%
PHOW color	42,11%	15,63%	18,75%
PHOW gray	40,35%	25,00%	15,63%
Self similarity	33,33%	18,75%	25,00%
Weighted	43,86%	21,88%	15,63%
Proposed	78,26%	61,29%	70,96%

6.3 Comparison with the Method in Smith et al. [2010] on the Ceramic Sherd Database

The proposed technique was also tested in the ceramic sherd database. Despite the great difference between these databases, the proposed technique resulted in satisfactory success rates in ceramic sherds as well.

Another important reason for using this database was to test the proposed BoW technique (see Figure 7), which appears to work efficiently, when it is fed with the appropriate features. For the ceramic database experiment, the selected features were hue (Eq. (1)), in phase (Eq. (1)), R-G-B components (Eq. (1)), standard deviation (Eq. (2)), saturation (Eq. (1)) and local binary patterns.

Again, the splitting percentage is set to 40% training and 60% testing samples. The same percentage is also used in the following section, where we compare the proposed technique with four other state-of-the-art methods. Table VIII presents the confusion matrix of our method in the ceramic sherd database, which has a success rate of 78.26%. In the same database, the technique of Smith et al. [2010] achieves a 75% success rate using the SIFT texture descriptor and 74% using the TVG descriptor.

6.4 Comparison with Various State of the Art Descriptors in Both Databases

In this section, we present our experiments with four other state-of-the-art techniques: self similarity [Shechtman and Irani 2007], pyramid histogram of words (color/gray) [Lazebnik et al. 2006; Bosch et al. 2007], geometric blur [Zhang et al. 2006; Berg et al. 2005] and a weighted combination of them [Vedaldi et al. 2009]. All four techniques have demonstrated significant performance in generic image databases. However, applying those methods to sherd classification reveals the difficulties of such a problem. Since these techniques are not designed to combine information from both front and back views, the pottery sherd database was considered as two separate databases, front sherd views and back sherd views. The proposed system was also tested in front and back sherd views separately.

Table IX presents the comparative classification rates for this experiment, where it is obvious that the proposed framework outperforms the existing state of the art.

7. CONCLUSION

In this article, we proposed an automatic technique for archaeological pottery sherd classification. The main contributions of the proposed work include a novel BoW model based on multithresholding, the exploitation of the information derived from both front and back views of sherd images, and a new approach for the creation of feature vectors, which can be used to classify sherds of artifacts into a finite number of classes. The proposed technique uses simple local features and a new bag of words model in order to create a global descriptor for each sherd image and to reduce the computational complexity. Furthermore, a feature selection technique was applied in order to select features with higher discriminative power, increase classification performance and further reduce the computational cost. In the proposed technique, at least one ground truth sherd image of each class was used in order to classify the rest of the sherds. Extensive experiments have been compared to other state of the art techniques. The results demonstrate the efficiency of the proposed method.

REFERENCES

ADAN-BAYEWITZ, D., KARASIK, A., SMILANSKY, U., ASARO, F., GIAUQUE, R. D., AND LAVIDOR, R. 2009. Differentiation of ceramic chemical element composition and vessel morphology at a pottery production center in Roman Galilee. *J. Archaeol. Sci. 36*, 2517–2530.

AGRAWAL, S., VERMA, N. K., TAMRAKAR, P., AND SIRCAR, P. 2011. Content based color image classification using SVM. In *Proceedings of the International Conference on Information Technology: New Generations.* 1090–1094.

AHA, D. AND KIBLER, D. 1991. Instance-based learning algorithms. *Springer, Mach. Learn. 6*, 37–66

BERG, A. C., BERG, T. L., AND MALIK, J. 2005. Shape matching and object recognition using low distortion correspondences. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recogniti. Vol. 1*, 26–33.

BEZDEK, J. C., EHRLICH, R., AND FULL, W. 1984. FCM: Fuzzy c-means algorithm. *Comput. Geosci. 10*, 191–203.

BISWAS, A., BHOWMICK, P., AND BHATTACHARYA, B. B. 2005. Reconstruction of torn documents using contour maps. In *Proceedings of the IEEE International Conference on Image Processing.* 517–520.

BOIMAN, O., SHECHTMAN, E., AND IRANI, M. 2008. In defense of nearest-neighbor based image classification. *Comput. Vis. Pattern Recog.* 1–8.

BOSCH, A., ZISSERMAN, A., AND MUNOZ, X. 2007. Representing shape with a spatial pyramid kernel. In *Proceedings of the International Conference on Image and Video Retrieval.* 401–408.

BUNKE, H. AND KAUFMANN, G. 1993. Jigsaw puzzle solving using approximate string matching and best-first search. In *Lecture Notes in Computer Science, vol. 719*, 299–308, Spriger.

BURCHARD. P. 2002. Total variation geometry i: Concepts and motivation. UCLA CAM rep. vol. 2(01).

CERAMIC SHERD DATABASE 2010. mst.cs.drexel.edu/datasets/Main/ACVA2010, with permission of Drexel Computer Science and NEC Labs.

CHAMORRO-MARTINEZ, J. AND MARTINEZ-JIMENEZ, P. 2009. A comparative study of texture coarseness measures. In *Proceedings of the International Conference on Image Processing.* 1337–1340.

CHANG, C.-C. AND LIN, C.-J. 2001. LIBSVM - A library for support vector machines. http://www.csie.ntu.edu.tw/∼cjlin/libsvm/.

CHIMLEK, S., KESORN, K., PIAMSA-NGA P., AND POSLAD, S. 2010. Semantically similar visual words discovery to facilitate visual invariance. In *Proceedings of the IEEE International Conference on Multimedia and Expo.* 1242–1247.

DE SMET, P. 2008. Reconstruction of ripped-up documents using fragment stack analysis procedures. *Elsevier Forensic Sci. Intern. 176*, 124–136.

GILBOA, A., KARASIK, A., SHARON, I., AND SMILANSKY, U. Towards computerized typology and classification of ceramics. *J. Archaeol. Sci. 31,* 681–694.

GOLDBERG, D., MALON, C., AND BERN, M. 2004. A global approach to automatic solution of jigsaw puzzles. *Elsevier Comput. Geom. 28,* 165–174.

GUYON, I., WESTON, J., BARNHILL, S., AND VAPNIK, V. 2002. Gene selection for cancer classification using support vector machines. *Mach. Learn. 46,* 389–422.

HALL, M. A. 2000. Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the 17th International Conference on Machine Learning.* 359–366.

HASTIE, T. AND TIBSHIRANI, R. 1997 Classification by pairwise coupling. In *Proceedings of Advances in Neural Information Processing Systems Conference (NIPS).*

HOTTA, K. 2009. Scene classification based on local autocorrelation of similarities with subspaces. In *Proceedings of the IEEE International Conference on Image Processing*. 2053–2056.

JOHN, G. H. AND LANGLEY, P. 1995. Estimating continuous distributions in Bayesian classifiers. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence.* 338–345.

JOLLIFFE, I. T. 1986. *Principal Component Analysis.* Springer-Verlag, NY.

JONHSON, N. L., KOTZ, S., AND BALAKRISHNAN, N. 1994. *Continuous Univariate Distributions* 2nd Ed. Vol. 1, Chapter 18, John Wiley and Sons.

JUSTINO, E., OLIVEIRA, L. S., AND FREITAS, C. 2006. Reconstructing shredded documents through feature matching. *Elsevier Forensic Sci. Intern. 160*, 140–147.

KAMPEL, M. AND SABLATNIG, R. 2000. Color classification of archaeological fragments. In *Proceedings of the IEEE International Conference on Pattern Recognition.* 771–774.

KAMPEL, M. AND SABLATNIG, R. 2004. On 3D mosaicing of rotationally symmetric ceramic fragments. In *Proceedings of the IEEE International Conference on Pattern Recognition.* 265–268.

KANDASAMY, K. AND RODRIGO, R. 2010. Use of a visual word dictionary for topic discovery in images. In *Proceedings of the International Conference on Information and Automation for Sustainability.* 510–515.

KARASIK, A. AND SMILANSKY, U. 2011. Computerized morphological classification of ceramics. *J. Archaeol. Sci. 38*, 2644–2657.

KARASIK, A. AND SMILANSKY, U. 2008. 3D Scanning technology as a standard archaeological tool for pottery analysis: Practice and theory. *J. Archaeol. Sci. 35,* 1148–1168.

KARASIK, A. 2008. Applications of 3D technology as a research tool in archaeological ceramic analysis. In *Beyond Illustration: 2D and 3D Digital Technology as Tools for Discovery in Archaeology,* B. Frischer and A. Dakouri-Hild, Eds., British Archaeological Reports International Series 1805, 103–116.

KARASIK, A. AND SMILANSKY, U. 2006. Computation of the capacity of pottery vessels based on drawn profiles. In *Excavations at Tel Beth Shean 1989–1996,* vol. I, Chapter 12, Appendix 1A, A. Mazar, Ed.

KARASIK, A., SMILANSKY, U., AND BEIT-ARIEH, I. 2005. New typological analyses of holemouth jars from the early Bronze Age from Tel Arad and Southern Sinai. *J. Instit. Archaeol. Tel Aviv Univ. 32*, 1.

KEERTHI, S. S., SHEVADE, S. K., BHATTACHARYYA, C., AND MURTHY, K. R. K. 2001. Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Comput. 13,* 3, 637–649.

KE-GANG, W., GUO-HUA, G., AND MING-QUAN, Z. 2008. Texture-based ancient porcelain shards classifications using Gabor transform. In *Proceedings of the 3rd International Conference on Intelligent Systems and Knowledge Engineering.* 699–702.

KERROUM, M. A., HAMMOUCH, A., AND ABOUTAJDINE, D. 2009. Textural feature selection by mutual information for multispectral image classification. In *Proceedings of the International Conference on Multimedia Computing and Systems.* 422–427.

KHELLAH, F. M. 2011. Texture classification using dominant neighborhood structure. *IEEE Trans. Image Process.*

KIRSCH, R. 1971. Computer determination of the constituent structure of biological images. *Comput. Biomed. Res. 4,* 315–328.

LAZEBNIK, S., SCHMID, C., AND PONCE, J. 2006. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2169–2178.

LEE, T. S. 1996. Image representation using 2D Gabor wavelets. *IEEE Trans. Pattern Anal. Mach. Intell*. 18, 959–971.

LEITAO, H. AND STOLFI, J. 2002. A multi-scale method for the reassembly of two-dimensional fragmented objects. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 1239–1251.

LIU, H. AND SETIONO, R. 1996. A probabilistic approach to feature selection—a filter solution. In *Proceedings of the International Conference on Machine Learning*. 319–327.

QI, L.-Y. AND KE-GANG, W. 2010. Kernel fuzzy clustering based classification of Ancient-Ceramic fragments. In *Proceedings of the IEEE Conference on Information Management and Engineering.* 348–350.

LOWE, D. G. 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision 60*, 2.

MAINI, R. AND AGGARWAL, H. 2009. Study and comparison of various image edge detection techniques. *Int. J. Image Proc.* 3, 1–11,

MAKRIDIS, M. AND PAPAMARKOS, N. 2010. A new technique for solving puzzles. *IEEE Trans. Syst., Man, Cybern., Syst. B. Cybern.* 40, 789–797.

MICHELSON, A. 1927. *Studies in Optics*. http://books.google.com.

MUWEI, J., LEI, L., AND FENG, G. 2009. Texture image classification using perceptual texture features and Gabor wavelet features. In *Proceedings of the Asia-Pacific Conference on Information Processing (APCIP)*.

NOVAK, C. AND SHAFER, S. 1992. Anatomy of a color histogram. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 599–605.

OJALA, T., PIETIKÄINEN, M., AND HARWOOD D. 1994. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR*). 582–585.

OJALA, T., PIETIKÄINEN, M., AND MAENPAA, T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell. 24*, 7, 971–987,

OTSU, N. 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man. Cybern. 9*, 1, 62–69.

PAPAODYSSEUS, C. PANAGOPOULOS, T., EXARHOS, M., TRIANTAFILLOU, C., FRAGOULIS,, D., AND DOUMAS, C. 2002. Contour-shape based reconstruction of fragmented, 1600 BC wallpaintings. *IEEE Trans. Signal Process. 50*, 1277–1288,

PENGYU, L., KEBIN, J., AND ZHUOZHENG, W. 2007. An effective image retrieval method based on color and texture combined features. In *Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing.*

PLATT, J. 1998. Fast training of support vector machines using sequential minimal optimization. In *Advances in Kernel Methods - Support Vector Learning,* B. Schoelkopf, C. Burges, and A. Smola, Eds.

REDDI, S. S., RUDIN, S. F., AND KESHAVAN, H. R. 1984. An optimal multiple threshold scheme for image segmentation. *IEEE Trans. Syst. Man. Cybern. 14*, 4, 661–665.

ROBNIK-SIKONJA, M. AND KONONENKO, I. 1997. An adaptation of Relief for attribute estimation in regression. In *Proceedings of 14th International Conference on Machine Learning.* 296–304.

SANGHO, Y. AND. GRAY, R. M. 2005. Feature selection based on maximizing separability in Gauss mixture model and its application to image classification. In *Proceedings of the IEEE International Conference on Image Processing.* 1198–1201.

SARAGUSTI, I., KARASIK, A., SHARON, I., AND SMILANSKY, U. 2005. Quantitative analysis of shape attributes based on contours and section profiles in archaeological research. *J. Archaeol. Sci. 32,* 841–853.

SHECHTMAN, E. AND IRANI, M. 2007. Matching local self-similarities across images and videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 1–8.

SHENG, X., TAO, F., LI, D., AND SHIWEI, W. 2010. Object classification of aerial images with bag-of-visual words. *IEEE Geosci. Remote Sens. Lett. 7*, 2, 366–370.

SMITH, J. R. AND CHANG, S.-F. 1996. Tools and techniques for color image retrieval. In *Proceedings of the Symposium on Electronic Imaging: Science and Technology - Storage and Retrieval for Image and Video Databases IV.* Vol. 2670, 426–437.

SMITH, P., BESPALOV, D., SHOKOUFANDEH, A., AND JEPPSON, P. 2010. Classification of archaeological ceramic fragments using texture and color descriptors. In *Proceedings of the IEEE Computer Vision and Pattern Recognition Workshop.* 49–54.

SUMNER, M., FRANK, E., AND HALL, M. 2005. Speeding up logistic model tree induction. In *Proceedings of 9th the European Conference on Principles and Practice of Knowledge Discovery in Databases.* 675–683.

VEDALDI, A., GULSHAN, V., VARMA, M., AND ZISSERMAN, A. 2009. Multiple kernels for object detection. In *Proceedings of the International Conference on Computer Vision.*

WALLAH, H. M. 2006. Topic modeling: beyond bag-of-words. In *Proceedings of the 23rd International Conference on Machine Learning.*

YAO, F.-H. AND SHAO, G.-F. 2003. A shape and image merging technique to solve jigsaw puzzles. *Elsevier Pattern Recog. Lett. 24*, 1819–1835.

ZHANG, H., BERG, A.C., MAIRE, M., AND MALIK, J. 2006. SVM-KNN: Discriminative nearest neighbor classification for visual category recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2126–2136.

ZHENHUA, G., LEI, Z., ZHANG, D., AND SU, Z. 2010a. Rotation invariant texture classification using adaptive LBP with directional statistical features. In *Proceedings of the IEEE International Conference on Image Processing.* 285–288.

ZHENHUA, G., ZHANG, L., AND ZHANG, D. 2010b. Rotation invariant texture classification using LBP variance (LBPV) with global matching. *Pattern Recog. 43,* 3.

Received December 2011; revised April 2012, July 2012; accepted July 2012