Research Journal of Applied Sciences, Engineering and Technology 10(12): 1396-1403, 2015
DOI: 10.19026/rjaset.10.1840
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: March 28, 2015 Accepted: April 22, 2015 Published: August 25, 2015

Research Article

Using Both HSV Color and Texture Features to Classify Archaeological Fragments

Nada A. Rasheed and Md Jan Nordin
Center for Artificial Intelligence Technology, Faculty of Information Science and Technology,
Universiti Kebangsaan Malaysia, Malaysia

Abstract: Normally, the artifacts are found in a fractured state and mixed randomly and the process of manual classification may requires a great deal of time and tedious work. Therefore, classifying these fragments is a challenging task, especially if the archaeological object consists of thousands of fragments. Hence, it is important to come up with a solution for the classification of the archaeological fragments accurately into groups and reassembling each group to original form by using computer techniques. In this study we interested to find the solve to this problem depending on color and texture features, to accomplish that the algorithm begins by partition the image into six sub-blocks. Furthermore, extract HSV color space feature from each block, then this feature represent into a cumulative histogram, as a result we obtain six vectors for each image. Regard to extract the texture feature for each sub-block it will be used the Gray Level Co-occurrence Matrix (GLCM) that include Energy, Contrast, Correlation and Homogeneity. At the final stage, based on k-Nearest Neighbors algorithm (KNN) classifies the color and texture features, this method able to classify the fragments with a high accuracy. The algorithm was tested on several images of pottery fragments and yield results with accuracy as high as 86.51% of original grouped cases correctly classified.

Keywords: Classification, feature extraction, GLCM, HSV color, texture

INTRODUCTION

Archaeology is the scientific study of the last remnants of humanitarian civilization (Son *et al*., 2013) and the reconstruction of fractures of ancient artifacts is important, as it helps archaeologists access to make inferences about past civilizations (Oxholm and Nishino, 2013). Therefore, over the last decade, there has been a trend toward the reconstruction of cultural heritage (Karianakis and Maragos, 2013; Youguang *et al*., 2013), which is considered among the difficult and unsolved problems in the field of computer vision (Papaodysseus *et al*., 2012; Rasheed and Nordin, 2014).

Usually, pottery fragments are found in archaeological excavation sites, randomly mixed with each other. Consequently, classifying them manually is difficult and time consuming, because they commonly exceed thousands of fragments, (Belenguer and Vidal, 2012; Papaodysseus *et al*., 2012; Son *et al*., 2013). Therefore, numerous researches proposed various methods have good achievement to classify the fragments of archaeological pottery object by depending on two-dimensional image. For example, Ying and Gang (2010) focused on surface texture feature, (Smith *et al*., 2010; Zhou *et al*., 2011; Makridis and Daras, 2012) proposed approaches for the classification of ancient ceramic fragments by relying on color and texture. This study suggested a new approach rely on color and texture features that are extracted from each image after divided it into six parts, this way assists to obtain the most important features to achieve more accurate and higher results than the previous work.
Thus, the main aims of this study achieve to:

- Propose a novel algorithm for classification of 2D ancient pottery fragments into groups with the assistance of computers.
- Improve the accuracy of the results.

To highlight the most important contributions of this study, the proposed method assists the archaeologists to:

- Reduce the time and manual effort required
- Reduce the human resources

This is a preparatory stage for the next phase, which is the reconstruction of the archaeological

Corresponding Author: Nada A. Rasheed, Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia

fragments with high accuracy. Therefore, in this study, we propose an algorithm to find a solution that classifies the archaeological fragments into groups, depending on their HSV color space and texture. In order to obtain the color features of an artifact, we consider Hue, Saturation and Value colors between its fragments. Also for the purpose of extracting the texture feature, we adopt the GLCM (Haralick *et al*., 1973) and calculate the Entropy, Contrast, Correlation and Homogeneity. Then we classify the fragments by using the KNN.

METHODOLOGY

System overview: This section is devoted to presenting the proposed method, which is divided into a set of steps, each one responsible for a specific job. As shown in Fig. 1, the proposed method can be represented by a series of steps, which has been programmed using MATLAB R2014 and performed on a standard laptop computer (Intel Core i5 2.5 GHz with 8 GB RAM).

Image acquisition and pre-processing: This procedure is responsible for loading the JPG format image file into memory. This format is used in this study, because this image format is one of the most widely used nowadays, due to its simple format, easy usage and provides high image quality. The dimensions of each image used in this study are (300×210) pixels. Sometimes noisy objects with no meaning appear with different sizes and shapes, because the image is converted from RGB color to the HSV color. Consequently, we applied 2D median filtering (Huang and Yang, 1979) to remove noisy objects.

Feature extraction: Feature extraction is an important component in pattern recognition, where the feature vector is a list of descriptions that include sufficient information to identify the pattern. In this study, the features were extracted from fragments depending on their color and texture, in order to select important features for recognition.

HSV color feature extraction: Color value varies according to three main factors, these factors determine the color that humans are able to see. The color discrimination is based on three elements, namely, Hue, Saturation, Value (HSV) (Levkowitz, 1997). As shown in Fig. 2, the angle above the circle from the axis describe the Hue (H), which has the value a range (0-360°). Saturation (S) is the distance from the axis, which is representing the vibrancy of the color, its value a range between 0 and 100°. Finally, the distance along the axis represents the Value (V), which is the brightness or intensity of the color. Its range is from 0-100°.

In order to extract HSV as a cumulative histogram vector for each sub-block, we apply a procedure that is similar to the one used in Kavitha *et al*. (2011), who explain the proposed algorithm thoroughly. Their algorithm includes the following steps:

Step 1: Divided each of Hue into eight parts, Saturation into three parts and intensity into three parts, depending on the human eye's ability to distinguish.

Step 2: On the basis of this division, it has been converted the values of Hue (H), Saturation (S) and Value (V) according to the following equations:

$$H = \begin{cases} 0 \; if \; h \in [316, 20] \\ 1 \; if \; h \in [21, 40] \\ 2 \; if \; h \in [41, 75] \\ 3 \; if \; h \in [76, 155] \\ 4 \; if \; h \in [156, 190] \\ 5 \; if \; h \in [191, 270] \\ 6 \; if \; h \in [271, 295] \\ 7 \; if \; h \in [296, 31] \end{cases} \tag{1}$$

$$S = \begin{cases} 0 \; if \; s \in [0, 0.2] \\ 1 \; if \; s \in [0.2, 0.7] \\ 2 \; if \; s \in [0.7, 1] \end{cases} \tag{2}$$

$$V = \begin{cases} 0 \; if \; v \in [0, 0.2] \\ 1 \; if \; v \in [0.2, 0.7] \\ 2 \; if \; v \in [0.7, 1] \end{cases} \tag{3}$$

Step 3: Apply the following equation:

$$G = 9H + 3S + V \tag{4}$$

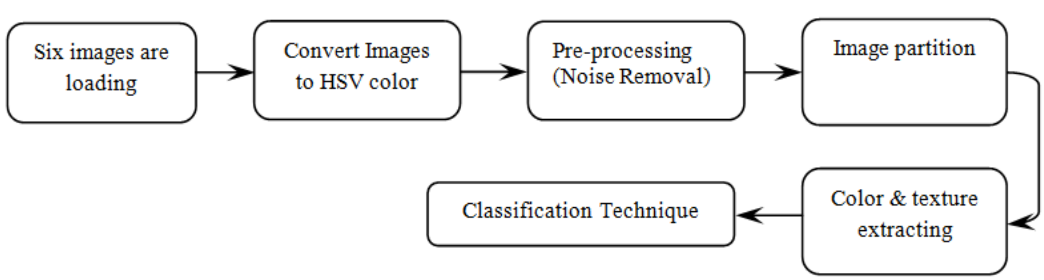


Fig. 1: Structure of the proposed algorithm

Fig. 2: Showing HSV color space

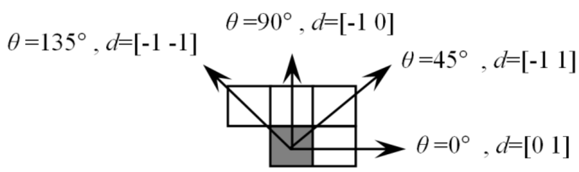


Fig. 3: Measurement of GLCM for four distances d and four angles θ

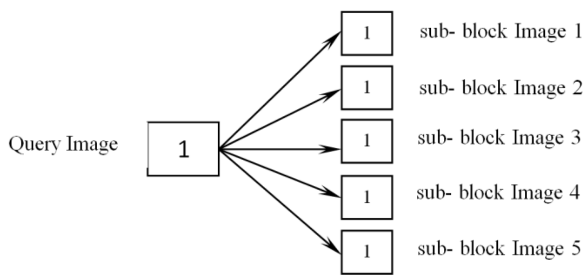


Fig. 4: The classification between each sub-block to the query image with five sub-blocks for other images

As a result, we obtain six histogram vectors for each image, which is normalized into the same range.

Texture feature extraction: One of the most powerful statistical texture analysis algorithms is the GLCM method, which is a two dimensional matrix that represents the number of co-occurrences to the pair of pixels, or two sets of pixels which have a grayscale value and are related by specific relationships. The algorithm consists of the following steps:

Step 1: The co-occurrence matrix method is performed on gray scale intensity images, where each one is divided into six parts.

Step 2: For each partition, a new matrix is created, composed of the probability value defined by P (i, $j|d$, θ), which expresses the probability of the couple pixels at θ direction and d distance that specifies the relationship between the intersected pixel and its neighbor.

In our adopted model, we depend on four different directions: θ = 0°, 45°, 90°, 135°. For each angle, we calculate four features, while d is represented by (0 1), (-1 1), (-1 0), (-1 -1), which corresponds to the 0°, 45°, 90° and 135° respectively. As shown in Fig. 3.

Step 3: On the basis of the matrix obtained, the Texture features are extracted from the archaeological fragments by applying four equations that include Contrast, Correlation, Energy and Homogeneity and can be represented through Eq. (5) to (8) respectively:

$$Contrast = \sum_{i,j}^{k} P_{i,j}(i-j)^2 \tag{5}$$

where,

- i and j = Pixels
- K = Represented (row or column) dimension of a square matrix
- $P_{i,j}$ = The probability of pixel pairs

$$Correlation = \sum_{i,j}^{k} \frac{(i-\mu_i)(j-\mu_j)P_{i,j}}{\sigma_i \sigma_j} \tag{6}$$

$$Energy = \sum_{i,j} P_{i,j}^2 \tag{7}$$

$$Homogeneity = \sum_{i,j} \frac{P_{i,j}}{1+|(i-j)|} \tag{8}$$

In our present study, the co-occurrence matrix method is performed on gray scale intensity images, where each one is divided into six parts.

Classification technique: Classifying images into similar groups manually is difficult for humans when there are many images. Using computer guidance for classifying is more convenient, especially when it is programmed with a series of important steps to classify the images. After converting each image to six vectors with a fixed length, we apply the most common distance function for KNN (Kim *et al*., 2012). The classification process was completed for all images (per sub-block for the query image to per sub-block for the target images depending on the corresponding locations) (Hiremath and Pujari, 2007). As shown in Fig. 4 the classification is between each sub -block to the query image with five sub-blocks for other images. After that compared the results for all sub-blocks to the query image and apply procedure to select the most similar image.

Experimental setup: The aim of this experiment is to use HSV color and texture features to classify the ancient fragments. So the following steps are explaining the experiment in details.

Image acquisition and pre-processing: The algorithm starts by loading six image files as shown in Fig. 5 and

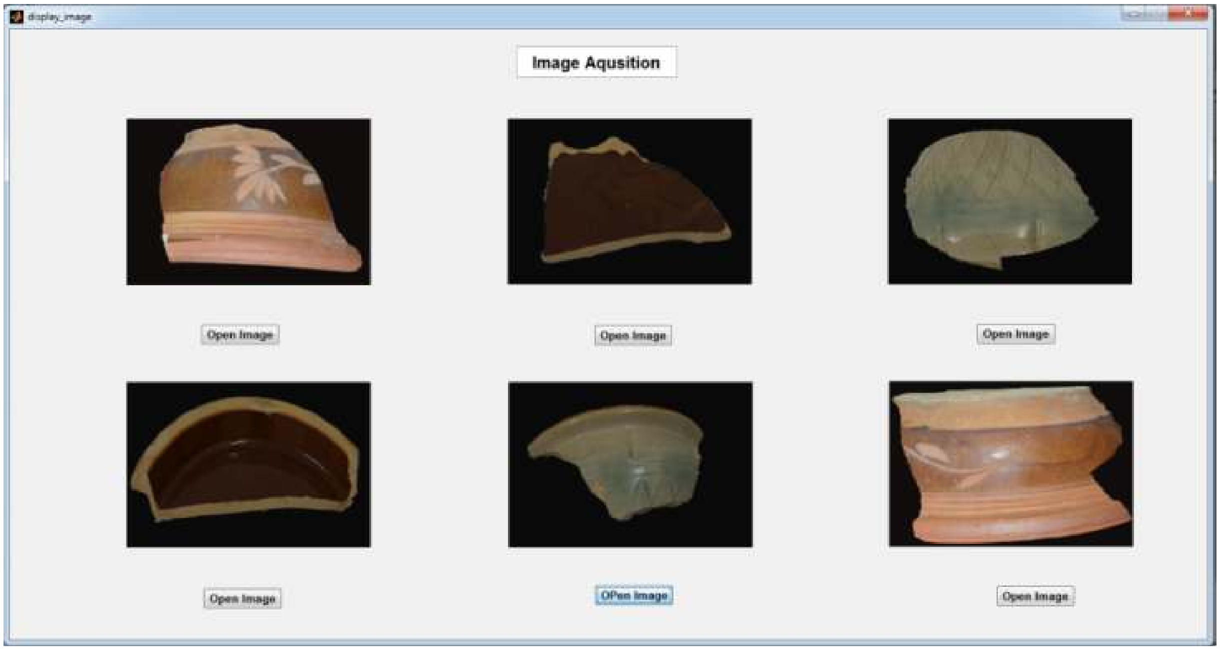


Fig. 5: Upload six images to the memory

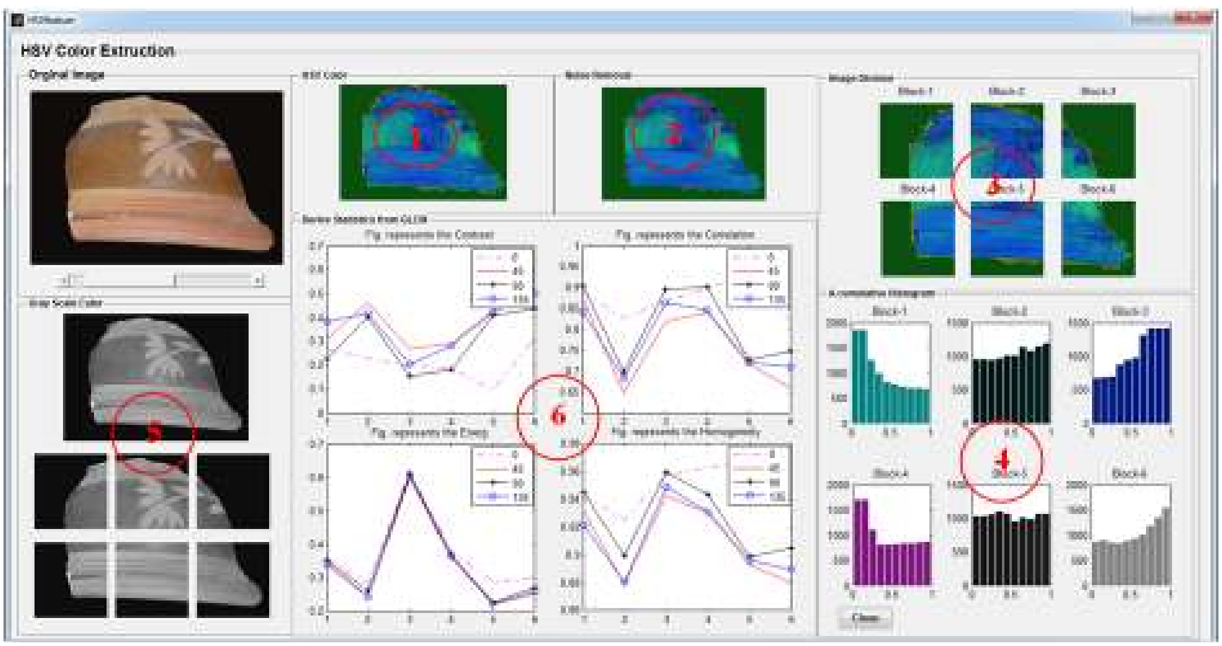


Fig. 6: Feature extraction depending on color and texture

each one is converted from RGB color to HSV. Then apply two procedures one for noise removal and the second procedure for dividing each image into six equal sized and non-overlapping sub-blocks, as shown in Fig. 3 to 6, each sub-block is (50×35) pixels in dimension.

Feature extraction: After computing the HSV from each sub-block, we obtain six histogram vectors for each image, as shown in parts (4) and (6) of Fig. 6, we apply the four equations numbered (5-8) on each sub-block to derive the statistics texture features, which are computed for all the parts of an image. Table 1 to 4 demonstrate the texture feature (GLCM), that results according to all angles (0°, 45°, 90° and 135°, respectively) for the first image. Note that the values of the features for each block according to the three angles (45°, 90° and 135°) almost converge, except for the zero angle, which is slightly different from the remaining angles, in terms of Contrast, Correlation

Table 1: Demonstrates the texture feature (GLCM) results according to the 0° for image#1

Texture feature	Block#1	Block#2	Block#3	Block#4	Block#5	Block#6
Correlation (R)	0.8905	0.8262	0.8686	0.8982	0.9360	0.8209
Contrast (C)	0.2671	0.2286	0.1964	0.1883	0.0962	0.3133
Energy (E)	0.3557	0.2841	0.6130	0.3798	0.2862	0.2957
Homogeneity (H)	0.9429	0.9258	0.9573	0.9629	0.9688	0.9333

Table 2: Demonstrates the texture feature (GLCM) results according to the 45° for image#1

Texture feature	Block#1	Block#2	Block#3	Block#4	Block#5	Block#6
Correlation (R)	0.8742	0.6470	0.8188	0.8414	0.7156	0.6611
Contrast (C)	0.3066	0.4645	0.2676	0.2922	0.4270	0.5910
Energy (E)	0.3465	0.2407	0.5974	0.3599	0.2186	0.2497
Homogeneity (H)	0.9316	0.8785	0.9430	0.9304	0.8928	0.8804

Table 3: Demonstrates the texture feature (GLCM) results according to the 90° for image#1

Texture feature	Block#1	Block#2	Block#3	Block#4	Block#5	Block#6
Correlation (R)	0.9066	0.6968	0.8967	0.9019	0.7283	0.7476
Contrast (C)	0.2271	0.4020	0.1532	0.1805	0.4083	0.4390
Energy (E)	0.3531	0.2592	0.6098	0.3697	0.2241	0.2676
Homogeneity (H)	0.9449	0.8996	0.9596	0.9436	0.8989	0.9051

Table 4: Demonstrates the texture feature (GLCM) results according to the 135° for image#1

Texture feature	Block#1	Block#2	Block#3	Block#4	Block#5	Block#6
Correlation (R)	0.8432	0.6829	0.8630	0.8468	0.7197	0.7111
Contrast (C)	0.3821	0.4172	0.2023	0.2822	0.4208	0.5039
Energy (E)	0.3376	0.2409	0.6048	0.3607	0.2219	0.2566
Homogeneity (H)	0.9218	0.8805	0.9492	0.9313	0.8962	0.8899

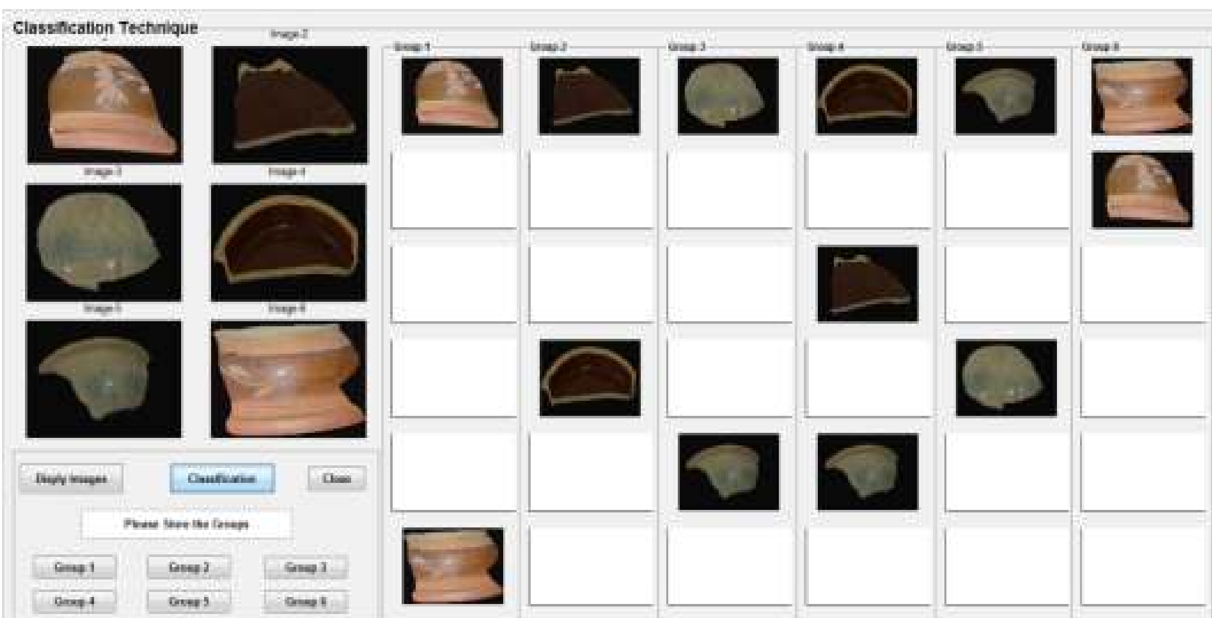


Fig. 7: Classification technique depending on color and texture

and Homogeneity. While in the Energy feature, the values in all angles almost identical.

As a result, for each image, we obtain six vectors of color features and sixteen vectors of texture features. The vectors are unified (color and texture) by combining together features in one vector for each sub-block. This way, we obtain six vectors per image, which will be suitable for classifying the images.

Classification technique: This section describes the classification of the fragments on the basis of color and texture features of the fragment. After computing all of the vectors to all the available sub-blocks, then the fragments are clustered by using KNN. The output of the classification technique is sorted and displayed, as shown in Fig. 7 and 8, Fig. 7 contains several icons in the far left of the model that display the six input images, while the classification technique was applied to the right of the model. Therefore, some of the fragments are classified successfully, such as Groups one, two and three, but failed with the fourth group, where fragment five appears in group four. Moreover,

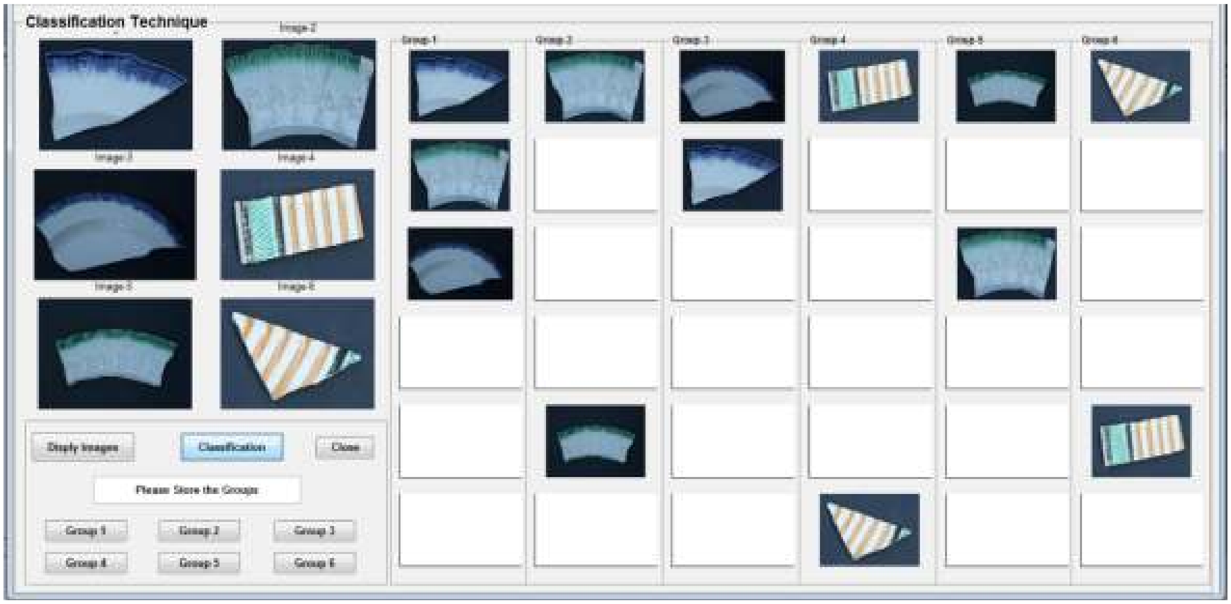


Fig. 8: Classify the Ceramic Sherd Database (2010)

as shown in Fig. 8, the algorithm has achieved success for all groups except the first group, which failed when the second fragment appeared in the wrong group. The user can choose the correct group by pressing on the one icon on the far left in order to save the results and use these as inputs for the next phase that reconstructs archaeological objects.

RESULTS AND DISCUSSION

In order to evaluate the performance of the proposed algorithm, several tests were conducted on the image dataset. In this experiment, nine ceramic fragments were obtained by a Nikon camera with a resolution of 24 MP and eighty fragments were imported from a database available on website (Ceramic Sherd Database, 2010). These artifacts date back to the late eighteenth century and the beginning of the nineteenth century and were discovered during the construction of the National Constitution Center between 2002 and 2003 (Smith *et al*., 2010). Most of this fragments database is highly textured, with smooth surfaces and high contrast, also it classified in 8 classes (Makridis and Daras, 2012).

In order to compare the performance between the proposed system and the other studies is very simple, because all the studies used the same datasets those on the site (Ceramic Sherd Database, 2010). Table 5 demonstrates the performance of the proposed system (P) is 86.51%, the measurement of the algorithm performance was conducted by using the following equation:

$$P = \frac{\sum CF}{\sum F}\% \tag{9}$$

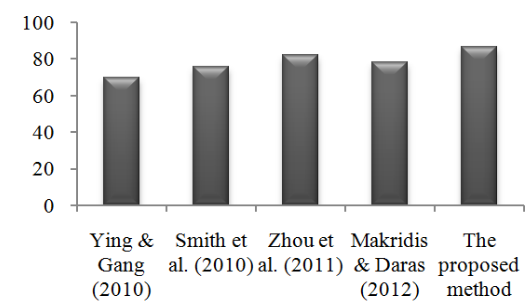


Fig. 9: Compare the performance of different studies

$$P = 86.51\% \tag{10}$$

Figure 9 shown the comparison between the performance of the proposed system and the other systems of classifying archaeological fragments. As demonstrated in Table 6, Zhou *et al*. (2011) study have applied the same features and classifier, but they obtained results less than the proposed method. Because the proposed method relied on sub-blocks of the image, when extract the features from each block and applied the same classifier, gave result a higher than the results of previous studies.

CONCLUSION

Archaeological objects are usually found broken and mixed with each other. Therefore, there is a need to classify these fragments. Currently, the classification of archaeological fragments into similar groups from a large quantity of fragments is performed manually. Hence, classifying them manually is difficult and time consuming, because they commonly exceed thousands

Table 5: Demonstrates the achievement the proposed method for classifying the fragments to each class in the database

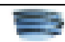
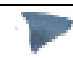
	Group										
	A	B	C	1	2	3	4	5	6	7	8
Image for each class											
No. (F)	3	3	3	9	9	9	16	2	10	7	18
(CF)	3	3	3	7	9	7	10	2	10	7	16
(%)	100	100	100	77.7	100	77.7	62.5	100	100	100	88.8

Table 6: Compare the performance of different studies

Previous works	Feature extraction	Classification technique	System performance (%)
Ying and Gang (2010)	Surface texture	Non-supervision kernel-based fuzzy clustering algorithm	70.00
Smith *et al*. (2010)	Color and texture	Fast vector quantization algorithm	76.00
Zhou *et al*. (2011)	Color and texture	K-Nearest Neighbor classifier (KNN)	82.14
Makridis and Daras (2012)	Chromaticity and chrominance (color), low level features	K-Nearest Neighborhood classifier (KNN)	78.26
The proposed method	Color and texture	K-Nearest Neighbor classifier (KNN)	86.51

of fragments. This study aims to solve the problem with depending on several tasks. In the first step, six images are acquired into memory and converted into HSV color space. Then, the pre-processing procedure is applied to remove noise caused by the conversion process. Moreover, each image is divided into six sub-blocks and the HSV color feature extracted from each sub-block. As a result, we obtain six vectors for each image. Regarding the extraction of the texture feature for each sub-block, the Gray Level Co-occurrence Matrix (GLCM) is used, which includes Energy, Contrast, Correlation and Homogeneity. At the final stage, on the basis of the results we obtained, the fragments are classified by using KNN method. The algorithm was tested on several images of 2D pottery fragments and was obtained accurate results. We conclude that the proposed classification technique is effective and achieves promising results. In Future work of this research, integrate RGB color space with HSV color space to extract the features and classifying the images based on KNN.

ACKNOWLEDGMENT

The researchers wish to thank Universiti Kebangsaan Malaysia (UKM) and Ministry of Education, Malaysia for supporting this work by research grants ERGS/1/2012/STG07/UKM/02/9 and FRGS/1/2014/ICT07/UKM/02/2.

REFERENCES

Belenguer, C.S. and E.V. Vidal, 2012. Archaeological fragment characterization and 3D reconstruction based on projective GPU depth maps. Proceeding of the IEEE 18th International Conference on Virtual Systems and Multimedia (VSMM), pp: 275- 282.

Ceramic Sherd Database, 2010. With Permission of Drexel Computer Science and NEC Labs. Retrieved from: mst.cs.drexel.edu/datasets/Main /ACVA2010.

Haralick, R.M., K. Shanmugan and I. Dinstein, 1973. Texture features for image classification. IEEE T. Syst. Man Cyb., 3(6): 610-621.

Hiremath, P. and J. Pujari, 2007. Content based image retrieval using color, texture and shape features. Proceeding of the International Conference on Advanced Computing and Communications, 1: 780-784.

Huang, T.S. and G.J. Yang, 1979. A fast two-dimensional median filtering algorithm. IEEE T. Acoust. Speech, 27(1): 13-18.

Karianakis, N. and P. Maragos, 2013. An integrated system for digital restoration of prehistoric theran wall paintings. Proceeding of the 18th IEEE International Conference on Digital Signal Processing (DSP, 2013), pp: 1-6.

Kavitha, C., B.P. Rao and A. Govardhan, 2011. Image retrieval based on color and texture features of the image sub-blocks. Int. J. Comput. Appl., 15(7): 33-37.

Kim, J., B.S. Kim and S. Savarese, 2012. Comparing image classification methods: K-nearest-neighbor and support-vector-machines. Proceeding of the 6th WSEAS International Conference on Computer Engineering and Applications and American Conference on Applied Mathematics, pp: 133-138.

Levkowitz, H., 1997. Color Theory and Modeling for Computer Graphics, Visualization and multimedia Applications. Kluwer Academic Publishers, USA, pp: 1.

Makridis, M. and P. Daras, 2012. Automatic classification of archaeological pottery sherds. J. Comput. Cult. Herit., 5(4): 1-21.

Oxholm, G. and K. Nishino, 2013. A flexible approach to reassembling thin artifacts of unknown geometry. J. Cult. Herit., 14(1): 51-61.

Papaodysseus, C., S. Skembris and E. Koukoutsis, 2012. 2D fragmented object reconstruction with the use of the chromatic and thematic content. Pattern Anal. Appl., 15(2): 133-146.

Rasheed, N.A. and M.J. Nordin, 2014. A polynomial function in the automatic reconstruction of fragmented objects. J. Comput. Sci., 10(11): 2339-2348.

Smith, P., D. Bespalov, A. Shokoufandeh and P. Jeppson, 2010. Classification of archaeological ceramic fragments using texture and color descriptors. Proceeding of the IEEE Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW, 2010), pp: 49-54.

Son, K., E.B. Almeida and D.B. Cooper, 2013. Axially symmetric 3D pots configuration system using axis of symmetry and break curve. Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR, 2013), pp: 257-264.

Ying, L. and W. Gang, 2010. Kernel fuzzy clustering based classification of ancient-ceramic fragments. Proceeding of the IEEE Conference on Information Management and Engineering, pp: 348-350.

Youguang, W., X. Li and M. Li, 2013. Color and contour based reconstruction of fragmented image. Proceeding of the 8th IEEE International Conference on Computer Science and Education (ICCSE, 2013). Colombo, Sri Lanka, pp: 999-1003.

Zhou, P., K. Wang and W. Shui, 2011. Ancient porcelain shards classifications based on color features. Proceeding of the 6th IEEE International Conference on Image and Graphics, pp: 566-569.