International Journal of Soft Computing 13 (2): 51-60, 2018
ISSN: 1816-9503

Trace Transform Based Features for Offline Handwritten Jawi Word Recognition

[1]Mohammad F. Nasrudin, [2]Elmi Mahzum, [1]Khairuddin Omar,
[3]Choong-Yeun Liong and [1]Mohamad Shanudin Zakaria
[1]Centre for Artificial Intelligence Technology (CAIT), Universiti Kebangsaan Malaysia (UKM),
43600 Bangi, Selangor, Malaysia
[2]Department of Physics, Faculty of Training and Education, Syiah Kuala University,
Banda Aceh, Indonesia
[3]Pusat Pemodelan dan Analisis Data (DELTA), Pusat Pengajian Sains Matematik,
Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia (UKM),
43600 Bangi, Selangor, Malaysia

Abstract: This study discusses offline handwritten Jawi recognition using the trace transform. We use two attainable kinds of features from the trace transform which are the “object signature” and “triple feature”. They are invariant to affine distortion, generated by the trace transform to discriminate between offline handwritten Jawi sub words. In trace transform, features construction of an image consists of tracing an image with straight lines, along which certain functional of the image function are calculated in all possible orientations. An object signature or a function of the orientation of the parallel lines is produced when a second functional is applied over all values computed along parallel lines. The computed object signature is in a form of string of numbers. A single number, triple feature is produced when a third functional over the string of numbers is applied. If the functional used have some predefined properties, the feature can be used to characterise the handwriting in an affine way. In this study, we also demonstrate the way of determining useful signatures and selecting cross-correlation methods for the signature classification. The results of the recognition experiments show that the object signature is a better feature than the triple feature in recognition of offline Jawi words.

Key words: Trace transform, offline handwritten word recognition, Jawi, feature extraction, signature, offline

INTRODUCTION

The field of offline handwritten Jawi recognition has made huge advancement during the past 10 years as shown by Nasrudin *et al.* (2008 a, b). The need for an automated reading in various offline applications such as form processing and ancient Jawi manuscript transliteration incite many developments of recognition methods. Many of these studies were targeting on the “offline” rather than “online” recognition problem simply because the latter is usually easier than the latter. Besides that, some recent works in Jawi handwritten recognition by Redika *et al.* (2008) and Heryanto *et al.* (2008) were working in word level because a perfect segmentation algorithm that would split Jawi’s cursive word image into complete characters with high confidence is not available.

Jawi is a cursive script that was derived from the Arabic alphabet and was adopted for the use of Malay language writing. The Jawi alphabet contains 36 basic characters of which 28 are similar to the Arabic alphabet. Although, they are using almost the same characters, Jawi words have more tendencies to be in a form of two or more sub-words compared to the Arabic script. According to Nasrudin (2010), it is because Jawi uses vowel characters in most of its syllables to accommodate the necessities in Malay language. Two out of three of its vowel characters are existed only in isolated form. It means that the characters can’t connect directly to the character which immediately follows to give the written text an overall cursive appearance. These differences make Arabic recognition methods such as those that uses word-level recognition can’t easily be employed to recognize Jawi words.

Offline handwritten word recognition requires the use of image features that capture the characteristics of the word shown, so that, they are invariant to the way the word is presented in the image. All well-established Jawi

Corresponding Author: Mohammad F. Nasrudin, Centre for Artificial Intelligence Technology (CAIT), Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia

image features that had been reported by Nasrudin *et al.* (2008b) are defined in a way that our human vision system would recognize in the image. However, features which are formulated away from the human perception may prove to be more appropriate for the characterization of handwritten Jawi word we would like to identify in an image.

A method that allows us to construct thousands of features that do not have physical or other meaning is the Trace transform. It is an alternative image representation and from which we can construct the, so, called object signature and triple feature. They have been developed by Kadyrov and Petrou (2001b) and have shown their effectiveness in various applications such as Korean character recognition (Kadyrov *et al*., 2001a), face recognition (Srisuk *et al*., 2005) and insect footprint recognition (Shin *et al*., 2008). In this study, we explore the use of the both trace transform based features to the problem of handwritten Jawi recognition.

Literature review: The area of offline cursive word recognition has been studied for various scripts. Many methods have been developed form these works. Most of these methods while presenting a large spectrum of perspectives on the problem, share a common vital issue that the recognition method should relate to is the segmentation problem. Ideally a perfect segmentation algorithm would split a cursive word image into complete characters. The common methods will try to find potential segmentation points while others had ventured on various unusually algorithms such as the Voronoi diagram like recently by Ramdan *et al.* Due to ambiguity in cursive words that was best expressed by Sayre's paradox, segmenting word image into perfect characters seems impossible. Based on this problem two approaches were developed, segmentation-based approach and segmentation-free approach.

In segmentation-based approach, a sequence of primitive segments derived from a word image and a group of possible words, one can seek for a possible segmentation that maximizes the average matching score between respective characters and segments. These primitive segments are created by some segmentation algorithm which might be imperfect and therefore cause over or under-segmentation. This approach can be observed by Heryanto *et al.* (2008) who over-segmented a Jawi word image using fixed length segments calculated based on the core zone height of the word. A neural network was used to classify each segment while dynamic programming was used to find the best complete match between blocks of primitive segments and a word's characters. However, this approach is failed to produce a promising result because of its imperfect segmentation algorithm and the existing of non-standard ligatures that is common in Jawi script.

The other approach uses a segmentation-free method where no attempt to split the word image into segments that relate to characters. It tries to recognize the whole word as one entity, by searching for a word with the most similar complete description to the one obtained from the whole input image. Still, it is possible that the image would be split into pieces in order to produce a sequence of observations like symbols. This can be observed by Redika *et al.* (2008) who spitted a Jawi word image to 5 frames of observation and employed Hidden Markov Model (HMM) as the classifier. The model produced decent recognition result using a data set that contains non-standard ligatures and overlaps. The success of this attempt, however was limited to constrained experiments and non-robust features.

Based on the last observations, we conclude that word or sub-word recognition without any segmentation is the way to go forward. However, the features extracted from word images should robust enough to classify the word. Based on the last survey on offline handwritten Jawi recognition done by Nasrudin *et al.* (2008b), they acknowledged the need for a better Jawi image features compared to the existing ones. They also highlighted that the well-investigated structural and statistical Jawi image features have been too human-centric. Some of the features for Jawi images that were proposed by Omar (2000) and Manaf (2002) such as number of isolated dots intersection points, end points and total black pixels, center of gravity orientation, eccentricity and spread of a binary object. Limiting the features among those that make sense to the human vision system somehow restricts the number of features that one can use.

Nasrudin *et al.* (2008a) have demonstrated image representation with no physical meaning can be generated by using the Trace transform. By using only five image features, they successfully recognized almost all printed Jawi characters that were affinely deformed. They concluded that the trace transform allows us to escape from the homocentric anthropocentric form of object description and permits us to use an almost unlimited number of features.

In a previous study, Nasrudin and Petrou (2011), we accomplished several preliminary experiments on the handwritten Jawi words feature construction from the Trace transform. In the study, we generated the Trace transform feature using some preselected functionals and selected them using a ranking measure called normalized discount cumulative gain. We then did the matching using a preselected correlation function. As the

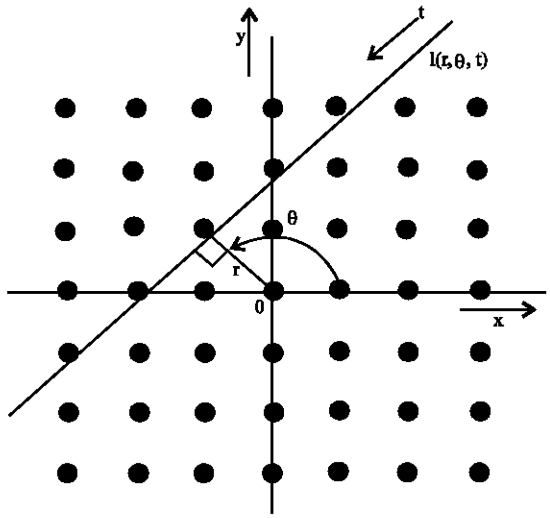


Fig. 1: Definition of the parameters of an imagef (r, y)and tracing line l (r, θ, t)

benchmark, we generated features from the infamous affine moment invariant and angular radial transform, the shape descriptor for the MPEG-7 standard. We compared our results with those obtained from the classed benchmark descriptors using several machine learning algorithms. From the study we found that for this problem, the trace transform was way better than the benchmark descriptors. However, it was worth to explore other improvements to our method. For these reasons in this study, we investigate other matching and feature selection techniques for the trace transform and compare recognition results generated from its both features.

Trace transform: The trace transform can be understood as a generalization of the well-known Radon transform (Radon 1917). The Radon transform of a real image function f(x, y) is a function p(r, θ) defined by computing the integral of f(x, y) along all lines L(r, θ):

$$p(r,\theta)=\iint_D f(x,y)\delta(r\text{-}x\cos\theta\text{-}y\sin\theta)\,dxdy \tag{1}$$

Where:

$r = x\cos\theta + y\sin\theta$	= The normal parameterization of a line l(r, u, t)
r	= The length of the normal from the origin of the axes to the line
θ	= The angle between the normal and the positive x-axis
t	= The parameter along the line
D	= The area of support of f(x, y)
δ	= The Dirac function (Fig. 1)

The trace transform is similar to the Radon transform in the sense that it also calculates a functional of the image function along lines but the functional is not necessarily the integral. The Radon (1970) transform, therefore is a specific case of the trace transform. Consider criss-crossing image f(x, y) with lines l(r, θ, t) in all directions. Denote by L(r, θ) the set of all lines. The trace transform is a function g(T, f, r, θ) defined on L(r, θ) with the help of trace functional T(some functional of the image function f(x, y) when it is considered along line l(r, θ) as a function of parameter t) then:

$$g(T,f,r,\theta)=T\left[f(r,\theta,t)\right] \tag{2}$$

A triple feature, a number which can characterise image f(r, y) is generated with the help of two additional functionals called diametric and circus functionals, designated by P and Φ, respectively (Kadyrov and Petrou 1998). The triple feature Π is defined as:

$$\Pi(f)=\Phi\left[P\left[T\left[f(r,\theta,t)\right]\right]\right] \tag{3}$$

Where:

Π	= The extracted triple feature of image and the three batch processes are
T	= The functional of parameter t
P	= The functional that is applied to parameter distance r, after the previous operations have been performed
Φ	= The functional operating on the orientation variable θ, after the two previous operations have been performed

The extracted triple feature is influenced strongly by the properties of the chosen functionals T, P and Φ. For practical applications, these functionals may be chosen so that the triple feature has one of the following properties (Kadyrov and Petrou, 2001b):

- Invariance to rotation, translation and scaling
- Sensitivity to rotation, translation and scaling so that these parameters can be recovered
- Correlates well with some desired properties which we want to identify in a sequence of images

Using an appropriate combination of functionals T, P and Φ, thousands of triple features can be generated, although most of them will not be useful nevertheless, one can investigate them and make the appropriate choice for the specific task with the help of experimentation.

Figure 2 shows the step of generating triple features from images of two Jawi characters “Syin” and “Pa”. Below each character image is its trace transform. Below the trace transform, we show the circus function of the image which is obtained by applying functional P to the columns of the trace transform. Finally, we apply functional Φ to it to obtain a single number, namely a triple feature which characterises the character.

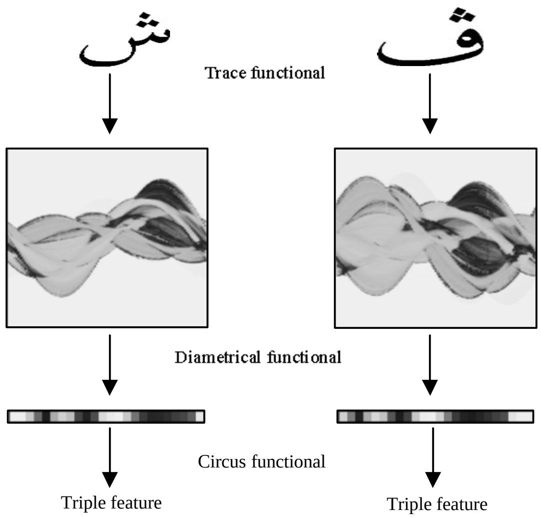


Fig. 2: The procedure of froducing a triple feature of an image

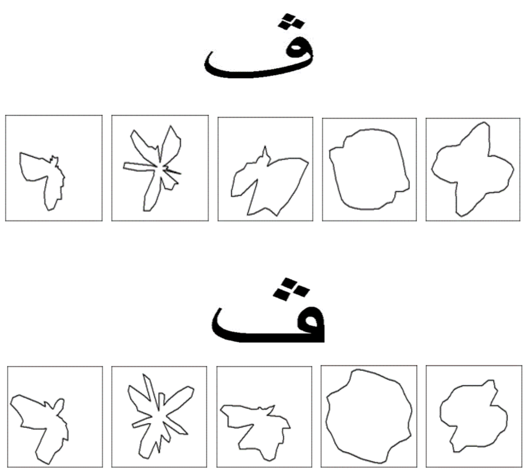


Fig. 3: The signatures (normalised associated circus functions) of Jawi character "Pa" in traditional and simple font types

Several functionals had been proposed by Kadyrov and Petrou (2001b) that one should use to produce features that are invariant to rotation, translation and scaling. Researchers by Petrou and Kadyrov (2004) presented some functionals that produced features invariant to affine distortions. Besides that, Kadyrov and Petrou (2001a) proposed a method that can be used to characterise an object, not by a single number produced by the cascaded application of the three functionals T, P and Φ but by using instead only the first two functionals. Using functionals T and P successively allows one to characterise an object by a string of numbers which is like a signature of the object.

The object signature is a function, called the associated circus, ha(ϕ) defined in terms of the function h(ϕ) which is produced by applying functionals T and P:

$$ha(\phi) = |h(\phi)|^{\frac{1}{\lambda_P K_T - K_P}} \quad (4)$$

parameters λ_P, K_T and K_P are some real valued numbers which characterise functionals T and T (Kadyrov and Petrou 2001a). If $\lambda_P K_T$-K_P = 0 the associated circus is defined as:

$$ha(\phi) = \sqrt{\left|\frac{dh(\phi)}{d(\phi)}\right|} \quad (5)$$

If we plot in polar coordinates the associated circus function of the original image, $h_{a1}(\phi)$ and the associated circus function of the affinely distorted image, $h_{a2}(\phi)$, they will produce two closed shapes which are connected by a linear transformation. In order to be able to compare these two shapes, they have to be normalised so that, their principal axes are coincidental. This can be done by a linear transformation applied to each shape separately as described in detail by Kadyrov and Petrou (2001a). The normalised shapes $h_{a1}(\phi)$ and $h_{a2}(\phi)$ are the signatures of the two images. For practical applications, the task of identifying an object is just comparing two strings of numbers, $h_{a1}(\phi)$ and $h_{a2}(\phi)$ that are circularly shifted and possibly scaled versions of each other. Figure 3 shows two signatures of a Jawi character in two different font types that are very similar in shape and differ mainly by rotation and scaling.

Experiments: The big picture of our experiments is shown in Fig. 4. Each of the steps is briefly discussed in the following subsections.

Data preparation: Generally, the input image of our recognition system is handwritten Jawi word grayscale images. We use the image original pixel values throughout the whole processes without went through any pre-processing filters. We collected 9 sets of scanned articles written by 9 different writers. Each article was designed such that all possible combinations of 36 Jawi characters exist at least once in the text. This was to make sure that all kinds of character combinations were tested, since, cursiveness changes character shape depending on where the character appears within its word or sub word. Each article contains 213 words. Randomly, we selected 6 articles as the training data (reference dictionary) and the rest, 3 articles as the testing data.

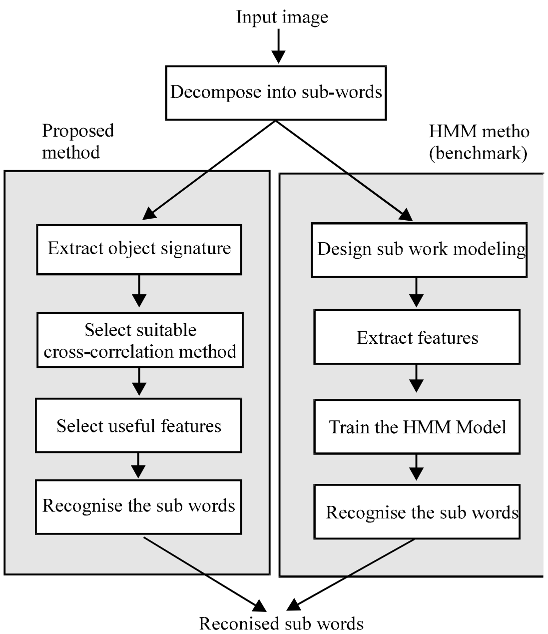


Fig. 4: Basic idea of the proposed method and benchmark method

Decompose into sub-words: All the scanned word images were decomposed into a set of sub-word images using connected component labeling segmentation algorithm provided by Pitas (2000). This step is possible because there are seven Jawi characters that can't be connected to their following character. So, a Jawi word is likely to be divided into several units. Theoretically, the smallest unit can be one character and the largest is the word itself. Of course in the worst case, all characters in a word including the seven characters, may be touching and even overlapping. Therefore, obviously, connected component analysis is not a decomposition method. Either the image of the worse-case scenario has to be in the reference dictionary or a more sophisticated segmentation procedure has to be developed.

The decomposition process generated 540 sub word images for each article. Thus in total there were 3240 images in the training data (reference dictionary) set and 1620 images in the testing dataset. These sub words can be grouped into 216 classes.

MATERIALS AND METHODS

Proposed method; Extraction of object signature: We calculated the object signature, ha by applying the functionals T and P. We tested seven different T functionals and eleven P functional taken from Kadyrov and Petrou (2001a) and Petrou and Kadyrov (2004). Their combinations however depend on two types of functional invariant and sensitive which are discussed in detail by Petrou and Kadyrov (2004). Basically, the allowed combinations are:

- Invariant T functional and invariant P functional
- Invariant T functional and sensitive P functional

The invariant T functional were:

- T_1: Integral of f(t) where f(t) is the value of the image function along the tracing line
- T_2: Max of $|f(t)|$
- T_3: Integral of $|f'(t)|$
- T_4: Integral of $|f''(t)|$
- T5: Lp quasi-norm (p = 0.5) = where q = integral of $\sqrt{|f(t)|}$
- T_6: Median R+: f(t-c) where c is the median abscissa
- T_7: Weighted Median R+: f (t-c) where c is the median abscissa and the weights are

The first seven P functionals are the same as T_1-T_7, called P_1-P_7, respectively. In addition, the following four sensitive P functionals were used:

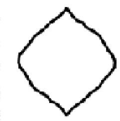
- P_8: t-median index dividing the integral of $|f(t)|$
- P_9: Average of t-index max of $|f(t)|$
- P_{10}: t-gravity center of $|f(t)|$
- P_{11}: t median index dividing integral of $\sqrt{|f(t)|}$

Following the above combination rules, we then generated all permitted pairs using these seven T functionals and eleven P functionals. In total, there were $7 \times 11 = 77$ circus functions or signatures to characterise an image.

Each image was traced by lines one pixel apart, i.e., the value of parameter p for two successive parallel lines used differed by 1. Each line was sampled with points one pixel apart that is to say parameter t took discrete values with steps equal to 1 inter-pixel distance. For each value of p, 48 different orientations were used. This means that the orientations of the lines with the same p differed by 7.5°. Thus that makes each object signature is 48 numbers long. Table 1 shows samples of sub words images along with their signature and combination of trace and diametric functionals.

Proposed method; Selection of the cross-correlation: To match two signature values, one from the test image and one from the reference, we computed their correlation coefficient for all possible shifts as suggested by Kadyrov and Petrou (2001a). We used the normalised circular cross-correlation of the two signatures which is defined as:

Table 1: Samples of sub words images along with their trace functional, diametric functional and signature both in real values and polar coordinates

Image	T	P	Object signature	Polar
	T_7	P_4	8.050e-004 ,8.050e-004, 8.126e-004, 8.278e-004, 8.509e-004,8.821e-004, 9.211e-004, 9.676e-004, 1.020e-003, 1.115e-003, 5.873e-004, 5.844e-004, 5.961e-004, 5.961e-004,5.844e-004, 5.873e-004, 1.115e-003, 1.020e-003, 9.676e-004, 9.211e-004, 8.821e-004, 8.509e-004, 8.278e-004,8.126e-004, 8.050e-004, 8.050e-004, 8.126e-004, 8.278e-004, 8.509e-004, 8.821e-004, 9.211e-004, 9.676e-004,1.020e-003, 1.115e-003, 5.873e-004, 5.844e-004, 5.961e-004, 5.961e-004, 5.844e-004, 5.873e-004, 1.115e-003,1.020e-003, 9.676e-004, 9.211e-004, 8.821e-004, 8.509e-004, 8.278e-004, 8.126e-004	
	T_1	P_3	8.981e-010, 1.007e-009, 9.338e-010, 7.387e-010, 6.569e-010,7.466e-010, 1.028e-009, 9.927e-010, 9.247e-010, 7.684e-010, 1.011e-009, 9.135e-010, 7.950e-010, 8.118e-010,7.835e-010, 7.889e-010, 8.583e-010, 1.014e-009, 8.858e-010, 1.042e-009, 8.645e-010, 6.478e-010, 8.493e-010,8.611e-010, 8.981e-010, 1.007e-009, 9.338e-010, 7.387e-010, 6.569e-010, 7.466e-010, 1.028e-009, 9.927e-010,9.247e-010, 7.684e-010, 1.011e-009, 9.135e-010, 7.950e-010, 8.118e-010, 7.835e-010, 7.889e-010, 8.583e-010,1.014e-009, 8.858e-010, 1.042e-009, 8.645e-010 6.478e-010, 8.493e-010, 8.611e-010	
	T_3	P_5	8.981e-010, 1.007e-009, 9.338e-010, 7.387e-010, 6.569e-010,7.466e-010, 1.028e-009, 9.927e-010, 9.247e-010, 7.684e-010, 1.011e-009, 9.135e-010, 7.950e-010, 8.118e-010,7.835e-010, 7.889e-010, 8.583e-010, 1.014e-009, 8.858e-010, 1.042e-009, 8.645e-010, 6.478e-010, 8.493e-010,8.611e-010, 8.981e-010, 1.007e-009, 9.338e-010, 7.387e-010, 6.569e-010, 7.466e-010, 1.028e-009, 9.927e-010,9.247e-010, 7.684e-010, 1.011e-009, 9.135e-010, 7.950e-010, 8.118e-010, 7.835e-010, 7.889e-010, 8.583e-010,1.014e-009, 8.858e-010, 1.042e-009, 8.645e-010, 6.478e-010, 8.493e-010, 8.611e-010	
	T_3	P_2	4.633e-004, 5.053e-004, 1.061e-003, 7.493e-004, 5.189e-004,5.648e-004, 5.029e-004, 4.841e-004, 6.460e-004, 5.800e-004, 8.256e-004, 6.510e-004, 4.582e-004, 5.091e-004,6.670e-004, 7.999e-004, 7.762e-004, 7.675e-004, 7.630e-004, 6.343e-004, 4.912e-004, 4.170e-004, 4.066e-004,4.373e-004, 4.633e-004, 5.053e-004, 1.061e-003, 7.493e-004, 5.189e-004, 5.648e-004, 5.029e-004, 4.841e-004,6.460e-004, 5.800e-004, 8.256e-004, 6.510e-004, 4.582e-004, 5.091e-004, 6.670e-004, 7.999e-004, 7.762e-004,7.675e-004, 7.630e-004, 6.343e-004, 4.912e-004, 4.170e-004, 4.066e-004, 4.373e-004	

$$CX(d)=\frac{\sum_{t=1}^{N}\left[\left(h_{at}(i)\right)\left(h_{al}(i-d)\right)\right]}{\sqrt{\sum_{t=1}^{N}\left(h_{at}(i)\right)^2}\times\sqrt{\sum_{t=1}^{N}\left(h_{al}(i)\right)^2}} \tag{6}$$

Where:

h_{at} and h_{al} = The signature values of the test image and reference image, respectively

N = The length of the signatures

d = The shift

Two signatures are most similar when their correlation is maximum. The correlation coefficient may also be produced via the Fast Fourier Transform as suggested by Kadyrov and Petrou (2003)

Circular cross-correlation algorithms are used in different variations. Since, circular cross-correlation is the only way to match two signatures, it is worthwhile to investigate the performance of other circular cross-correlation variations than what is proposed by Kadyrov and Petrou (2001a). Thus, we calculated five other circular cross-correlation methods which are defined as:

$$CC_1(d)=\sum_{i=1}^{N}|h_{at}(i)-h_{al}(i-d)|$$

where two signatures are most similar when their CC_1 is minimum:

$$CC_2(d)=\frac{1}{N}\sum_{i=1}^{N}\left(h_{at}(i)-h_{al}(i)\right)^2$$

mean squared error where two signatures are most similar when their $CC_2$is minimum:

$$CC_3(d)=\sum_{i=1}^{N}\left[\left(h_{at}(i)-h_{al}(i-d)\right)^2\right]$$

where two signatures are most similar when their CC_3 is minimum:

$$CC_4(d)=\frac{\sum_{t=1}^{N}\left[\left(h_{at}(i)\right)-\left(h_{al}(i-d)\right)\right]}{\sqrt{\sum_{t=1}^{N}\left(h_{at}(i)\right)^2}\times\sqrt{\sum_{t=1}^{N}\left(h_{al}(i)\right)^2}}$$

where wo signatures are most similar when their CC_4 is minimum:

$$CC_5(d)=\frac{\sum_{t=1}^{N}\left[\left(h_{at}(i)-\bar{h}_{at}\right)-\left(h_{al}(i-d)-\bar{h}_{al}\right)\right]}{\sqrt{\sum_{t=1}^{N}\left(h_{at}(i)-\bar{h}_{at}\right)^2}\times\sqrt{\sum_{t=1}^{N}\left(h_{al}(i)-\bar{h}_{al}\right)^2}}$$

where $\bar{h}_{at}$ and $\bar{h}_{al}$ are the average signature values of the test image and reference image, respectively. Two signatures are most similar when their CC_5 is maximum.

Proposed method; Selection of useful features: The first step in selecting useful features is to eliminate the useless ones. Signatures that have all zeros or all the same constant numbers are discarded. Then, functional combinations that generated object signature which correctly ranks the reference data should be selected. However, to our knowledge, there have been not any methods analogous to feature selection dedicated to object signature ranking. The nearest work is done by Kadyrov *et al.* (2002) which is to select relevant triple features for texture classification. In that study, the resaerchers selected a feature as relevant if its value were stable when the instantiation of the texture changed. They defined the stability measure to be the variance of the values of the feature over the whole set of its values. From the variance, they formulated weights for the features to indicate their relevance.

Obviously, existing feature selection methods for classification such as by Yang and Pedersen (1997) and Forman (2003) are not suitable for ranking. In ranking, a number of ordered categories are used, representing the ranking, relationship between instances while in classification the categories are flat. Besides that in ranking usually precision is more important than recall while in classification both precision and recall are important. In addition in ranking correctly ranking the top-n instances is more critical whereas in classification making a correct classification decision is of equal significance for all instances. In this study, we adopted two widely-used measures in evaluation of ranking methods for information retrieval which are Mean Average Precision (MAP) (Yates and Neto, 1999) and Normalized Discount Cumulative Gain (NDCG) (Jarvelin and Kekalainen, 2002).

MAP is a measure on precision of ranking while assuming that there are two types of item (cross-correlation results) which are positive (relevant) and negative (irrelevant). Precision at n measures the accuracy of the top n sorted items and defined as:

$$P(n)=\frac{top_{pos}\text{-}n}{n} \tag{7}$$

where top_{pos} is number of positive items within the top. Average precision of a query, AP is defined as:

$$AP=\sum_{n=1}^{N}\frac{P(n)\times pos(n)}{all_{pos}} \tag{8}$$

Where:

n = Position
N = Number of results
pos(n) = The binary function indicating whether the item at position
n = Positive
all_{pos} = The number of positive items

MAP is defined as AP averaged over all sorted items. NDCG is designed for measuring ranking accuracies when there are multiple levels of relevance judgment. NDCG at position n in sorted items is defined as:

$$N(N)=Z_n\sum_{j=1}^{N}\frac{2^{R(j)}-1}{\log(1+j)} \tag{9}$$

Where:

n = Position
R (j) = Score for ran k j
Z_n = The normalization factor to guarantee that a perfect ranking's NDCG at position n equals 1

In evaluation, NDCG is further averaged over all queries. In selecting useful features using the MAP and NDCG, the objective is to search the maximum sum of the importance scores of individual object signatures. Testing all possible combinations, however will lead to a problem of optimization. The worst approach is to perform exhaustive search. However, the time complexity of it is too high to make it applicable in real applications. Instead, we used the hill climbing approach to find the optimal combination. It starts with a random object signature, iteratively adds another object signature to the combination and calculates the value of MAP and NDCG. If the combination produces lower result, then, the recently added object signature is discarded. When all object signatures are tested, it terminates.

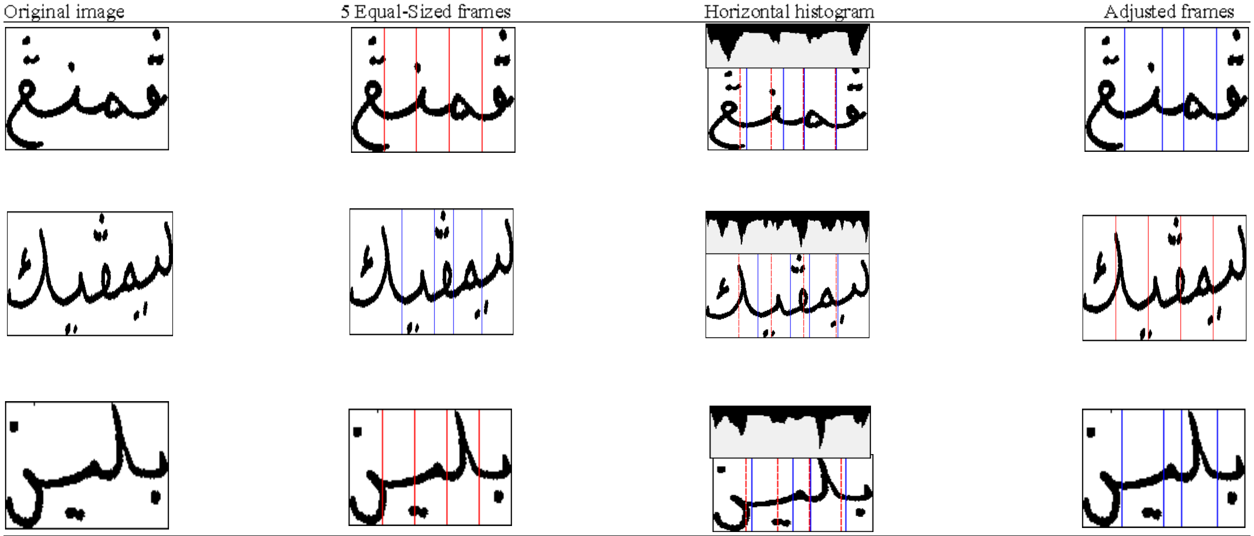
Proposed method; Recognising of the sub-words: The proposed method of recognition is to compare the two circular cross-correlation values, one from the test image and one from the reference. We computed their correlation coefficient over 48 shifts (equal to 48 different orientations used). We used as measure of similarity the sum of these numbers. We ranked the numbers. The smallest or the largest number, depending on the cross-correlation technique used indicates the two most similar signatures.

Benchmark method; Designing of sub word models: The first task in the recognition model is to transform word images into a suitable model of HMM as discussed by Redika *et al*. (2008). In order to adapt the approach, we converted the word image into sequential data taken from fragments of the word image which were called 'frames'. We divided each word image into five equal-sized width vertical frames. We then adjusted the width of each frame according to the nearest smallest horizontal histogram projection. As a result, different-sized frames were obtained as in Table 2. Each observation frame is represented by a sequence of features.

Benchmark method; Features extraction: We collected the shape primitive information of a sub word image using geometrical and low-level pixel-based features. For each frame, we extracted:

Table 2: Samples of sub words images along with their 5 equal-sized frames, horizontal histogram and adjusted frames

Original image	5 Equal-Sized frames	Horizontal histogram	Adjusted frames



* The 32 profile features. We calculated eight profiles features from four directions: left, right, top and bottom
* The 16 ink-crossing features. We calculated eight ink crossing features vertically and horizontally
* The 17 pixel density features. We calculated the pixel density features from eight columns (zones) vertically and horizontally. One pixel density feature was taken from the whole image

Thus, each frame was represented by a 65 dimensional features vector. In order to reduce the sensitivity of the features vectors, each feature value was normalized by the width or height of the frame, depending on whether it was calculated vertically or horizontally.

Benchmark method; Training of the HMM Model: Training the HMM Model means the maximization process of the likelihood of the training data. The training is done by using the segmental k-means or the Baum-Welch algorithm from the HMMPak Software that can iteratively and automatically find the maximum likelihood. From the training, we got a trained HMM Model. All identical Jawi sub-words are grouped into a class and have a trained HMM Model with parameters estimated from various instances of the sub word class. We trained 3240 images of 216 sub word image lexicons. The training processes are shown in Fig. 5.

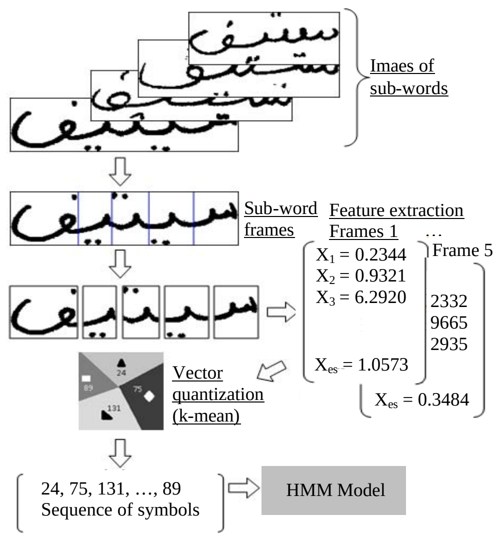


Fig. 5: Training of the HMM Model

Benchmark method; Recognition of the sub-words: In this step, we applied the same procedure to testing sub word image as in the training step which was converting the image into five observation frames, extracting features and encoding the features. Then, we employed the Viterbi algorithm to find the maximum likelihood of testing data over the trained HMM Models. We ranked all model probabilities. HMM Models that were in the top-5 rank were the candidates with high probabilities.

RESULTS AND DISCUSSION

The results of all our experiments are presented in Table 1-3 and divided into three parts: the best cross-correlation method, a set of useful features and comparison of the proposed and the benchmark methods.

Table 3: Percentage of correct sub words recognition based on several cross-correlation methods selection

Methods	Percentage of correct recognition 1-5	6-0	11-15	≥16
CX	36.62	6.57	10.33	46.48
CC_1	26.29	8.45	10.80	54.46
CC_2	24.41	6.10	10.33	59.15
CC_3	19.25	8.92	6.10	65.73
CC_4	46.48	10.80	5.63	37.09
CC_5	38.03	8.45	10.80	42.72

Table 4: Percentage of correct sub-words recognition and its number of features based on MAP and NDCG measurements

Methods	Percentage of correct recognition 1-5	6-10	11-15	≥16	No. of feature
MAP	63.85	7.04	4.69	24.41	20
NDCG	64.79	8.92	5.63	20.66	24

The best cross-correlation method: The result of our experiments on the performance of normalised circular cross-correlation and its variations is presented in Table 3. In order to scale down the magnitude of the experiment, we randomly selected only one set of reference data and one set of test data. Since, the useful features are still unknown, we used five randomly selected signatures in the experiment. In the table, we put in the second column the percentage of correct recognition of the test sub words in the top-five choices. In the third and fourth columns we put the percentages it recognised in the sixth to tenth position and the 11-15th positions, respectively. In the fifth column we put the percentages of recognition in the sixteenth position and beyond.

The result shows that the CC4 cross-correlation method produces a better result compared with other methods. It is obvious that the overall percentage of recognition is very low, since, only five features are used in this experiment. From this result, we decided that the CC4 will be used for the rest of the experiments.

A set of useful features: The result of our experiments on the selection of a set of useful features based on MAP and NDCG ranking evaluation methods is presented in Table 4. As in the previous experiment, here, we also used scaled-down data sets which consist of one set of reference data and one set of test data. In the table, we put in the second column the percentage of correct recognition in the top-five choices. In the third and fourth columns, we put the percentages it recognised in the sixth to tenth position and the eleventh to fifteenth positions, respectively. In the fifth column, we put the percentages of recognition in the sixteenth position and beyond. In the last column, we put the number of selected useful features based on both the MAP and NDCG methods.

Table 5: Comparison of correct sub-words recognition percentages of the proposed and benchmark methods in positions 1-5, 6-10, 11-15, and the 16th and beyond

Methods	Percentage of correct recognition 1-5	6-10	11-5	≥16
Proposed	72.02	6.89	3.71	17.29
Bench-mark	54.91	18.10	8.65	18.34

The result shows that based on the NDCG measurement, 24 out of 77 signatures are useful. The MAP measurement, on the other hand, selected only 20 features and registered a lower recognition rate. Based on the higher overall correct recognition percentage, we decided to use those 24 features for the benchmarking experiment.

Comparison of the proposed and the benchmark methods: In both our proposed and benchmark experiments, we use all 9 sets of data which are 6 sets (3240 images) as reference data and 3 sets (1620 images) as testing data. The comparison of the sub-words recognition percentages of the proposed and benchmark methods are presented in Table 5.

CONCLUSION

We can see that normalised circular cross correlation which is based on the difference rather than the multiplication, produces a better result compared with other cross-correlation variations. It is much better than the result produced by CX, the method proposed by Kadyrov and Petrou (2001). However, at this stage, the best recognition result is still low which is 46.48%.

The features selected based on the MAP and NDCG ranking measurements. Both measurements produced slightly different results. NDCG gave better overall recognition percentages with 4 extra features. This demonstrates that NDCG is a better ranking measurement to calculate the importance scores for individual object signatures. It improves the best recognition rate to 64.79%.

Results of the experiments on comparing the benchmark methods and the proposed method are studied here. Overall, the object signatures clearly show better results compared with the benchmark method. The top-5 recognition rate of the object signature is 72.02% which is better than 54.91% of the top-5 recognition rate of the benchmark method. It shows that by selecting the suitable cross-correlation method and the right combination of object signatures, trace transform is the best alternative to renowned methods such as the Hidden Markov Model in recognizing handwritten Jawi.

ACKNOWLEDGEMENTS

Researchers also would like to thank the University Kebangsaan Malaysia and Ministry of Science, Technology and Innovation (MOSTI), Malaysia for the Research Grant No. FRGS/1/2016/ICT02/UKM/02/1.

REFERENCES

Forman, G., 2003. An extensive empirical study of feature selection metrics for text classification. J. Machine Learn. Res., 3: 1289-1305.

Heryanto, A., M.F. Nasrudin and K. Omar, 2008. Offline Jawi handwritten recognizer using hybrid artificial neural networks and dynamic programming. Proceedings of the 2008 International Symposium on Information Technology (ITSim'08) Vol. 2, August 26-28, 2008, IEEE, Kuala Lumpur, Malaysia, ISBN: 978-1-4244-2327-9, pp: 1-6.

Jarvelin, K. and J. Kekalainen, 2002. Cumulated gain-based evaluation of IR techniques. ACM. Trans. Inf. Syst., 20: 422-446.

Kadyrov, A. and M. Petrou, 2001. The trace transform and its applications. IEEE. Trans. PatternAnal. Mach. Intell., 23: 811-828.

Kadyrov, A. and M. Petrou, 2001. Object descriptors invariant to affine distortions. Proceedings of the Conference on British Machine Vision (BMVC'01), September 10-13, 2001, University of Manchester, Manchester, UK., ISBN: 1-901725-16-2, pp: 391-400.

Kadyrov, A. and M. Petrou, 2001. The trace transform and its applications. IEEE. Trans. Pattern Anal. Mach. Intell., 23: 811-828.

Manaf, M., 2002. Jawi handwritten text recognition using recurrent bama neural networks. Ph.D. Thesis, Universiti Kebangsaan Malaysia.

Nasrudin, M.F. and M. Petrou, 2011. Offline handwritten JAWI recognition using the trace transform. Proceedings of the 2011 International Conference on Pattern Analysis and Intelligent Robotics (ICPAIR'11) Vol. 1, June 28-29, 2011, IEEE, Putrajaya, Malaysia, ISBN: 978-1-61284-407-7, pp: 87-91.

Nasrudin, M.F., 2010. Offline JAWI handwritten recognition using the trace transform (In Malay). Ph.D Thesis, National University of Malaysia, Bangi, Malaysia.

Nasrudin, M.F., C.Y. Liong, K. Omar and M.S. Zakaria, 2008. Trace transform for Jawi character recognition. Proceedings of the International Symposium on Geometric Function Theory and its Applications (GFTA'08), November 10-13, 2008, National University of Malaysia, Bangi, Malaysia, pp: 441-448.

Nasrudin, M.F., K. Omar, M.S. Zakaria and L.C. Yeun, 2008. Handwritten cursive JAWI character recognition: A survey. Proceedings of the 5th International Conference on Computer Graphics, Imaging and Visualisation, August 26-28, 2008, Penang, pp: 247-256.

Omar, K., 2000. Jawi handwritten text recognition using multi-level classifier (in Malay). Ph.D Thesis, Universiti Putra Malaysia, Seri Kembangan, Malaysia.

Petrou, M. and A. Kadyrov, 2001. Affine invariant features from the trace transform. IEEE Trans. Patt. Anal. Mach. Intell., 26: 30-44.

Pitas, I., 2000. Digital Image Processing Algorithms and Applications. John Wiley & Sons, New York, USA., ISBN:0-471-37739-2, Pages: 403.

Radon, J., 1917. [On the determination of functions by their integrating values of the weaving manifolds (In German)]. Rep. Saechsishe Acad. Science. Math. Phys. Klass, 69: 262-277.

Redika, R., K. Omar and M.F. Nasrudin, 2008. Handwritten Jawi words recognition using hidden Markov models. Proceedings of the International Symposium on Information Technology (ITSim'08) Vol. 2, August 26-28, 2008, IEEE, Kuala Lumpur, Malaysia, ISBN:978-1-4244-2327-9, pp: 1-5.

Shin, B.S., E.Y. Cha, K.W. Cho, R. Klette and Y.W. Woo, 2008. Effective feature extraction by trace transform for insect footprint recognition. MSc Thesis, University of Auckland, Auckland, New Zealand.

Srisuk, S., M. Petrou, W. Kurutach and A. Kadyrov, 2005. face authentication system using the trace transform. Pattern Anal. Appl., 8: 50-61.

Yang, Y.M. and J. Pedersen, 1997. A comparative study on feature selection in text categorization. Proceedings of the 14th International Conference on Machine Learning, July 8-12, 1997, Nashville, TN., USA., pp: 412-420.

Yates, R.B. and B.R. Neto, 1999. Modern Information Retrieval. Pearson, London, England, UK., ISBN: 978-81-317-0977-1, Pages: 317.