Assessing the Use of Similarity Distance Measurement in Shape Recognition

Siti Salwa Salleh, Noor Aznimah Abdul Aziz, Daud Mohamad, and Megawati Omar
Faculty of Computer and Mathematical Science,
Research Management Institute
University Technology MARA, 40450 Shah Alam, Selangor, Malaysia
Email: ssalwa@tmsk.uitm.edu.my

Abstract—Distance measure is one of the techniques widely used to measure the similarity between two feature matrices of objects. The objective of this paper is to explore researches on applied distance measures in shape-based recognition. In distance measures computation, patterns that are similar will have a small distance while uncorrelated pattern in the feature space will have a far a part distance. The search for effective distance measures of shape recognition is always active as each measure suffers certain drawbacks and it must be selected appropriately to handle chosen shape features of the objects. Thus in this paper, the Chord, Cosine, Euclidean, Mahalanobis, Trigonometric and Jaccard distance were reviewed and discussed in terms of their contributions, measures strengths and weaknesses. It was found that Jaccard and Mahalanobis have their strengths that they were selected to guide in justifying and identifying appropriate distance measures of our future work on two dimensional sketching images. The new distance measure is expected to perform better and capable to obtain significance degree of accuracy and recognition rate for real time recognition for automatic classifier.

Index Terms—Distance Measures, shape recognition, recognition classifier, similarity measures, shape context.

I. INTRODUCTION

Over the years, object recognition has been employed in several approaches. One of the approaches is a shape descriptor which covers shape space [15], chord context [9], shape context [13], Fourier descriptors [21] and more of its kind. Shape descriptor has been deployed to find the corresponding points or features between shapes [13, 1]. Previous shape recognition researchers widely use neural network [17, 7], hidden markov model (HMM) [19], simple matching techniques or similarity measures [3, 8, 9, 12, 14, 16]. However, recognition using Neural Network and HMM requires a large set of training database [18, 3], while simple matching technique does not return a promising recognition performance. Moreover, in certain application, large dataset are not available, neither nor it practical. Shape is a dominant feature of an object as it consist lines, contours, curves, and vertices and it is normally presented by discrete set of points or set of pixels value sampled from region or internal and external contour on the object [1, 5]. Generally, there are few recognition techniques that based on object properties [5,6,10] which are shape, texture, color, and brightness. But compared to other features, shape is the most unique where it is able to recognize objects more efficiently. Practically, shape recognition recognizes reference shape from other shape, and test whether the reference shape is applicable or otherwise [20, 10]. This recognition technique is being used in industries such as fashion, architectural, journalism, advertisement, education, and entertainment.

Among recognition techniques, similarity measure using distance measure techniques is one of the popular methods. It works in the manner that small distances corresponding to large similarities and large distances corresponding to the small. This approach is widely used for measuring the similarity between objects as reported in [3, 8, 9, 12, 14, 16]. Generally, distance measures is extensively employed in content-based image retrieval for shape-based trademark retrieval, shape-based image retrieval, planar object recognition such as handwritten character recognition and other. Likewise, various similarity measures techniques have also been used to define appropriate distance function that provides the reasonable results for the images comparison. They are Euclidean Distance [14], Mahalanobis Distance [3], Chord Distance [9] Cosine Distance [16], Trigonometric Distance [8] Jaccard Distance [12] and others. But among those, not all distance measures work well on a sketch-based input and each distance measure have its own strength and also suffer from certain weakness.

II. PROBLEM STATEMENT AND OBJECTIVE

A central problem in object recognition is to determine proper distance measure for a particular object. Extensive studies on shape recognition prove some significance progresses but current use of the recognition and shape classification technique has yet to prove satisfactory. Since each distance measure has its own strength and weakness, researchers must be careful to choose the different ways of measuring the distance, which is the chosen approaches must conform to their needs and applications. It is important to note that, there is no general distance measure that works best for every kind of shape feature. Hence, the best way in selecting distance or similarity measure is by identifying and considering the most minimum vector space values by the use of simple function [8].

Another problem we may bear in mind is that however appropriate the distance function is, it still suffers from some kind of shortcomings. A survey by Wang et. al [14] mentioned that most of existing distance function encompasses a complexity of the computation measure. It also brings some difficulty to combine or embed the metric proposed with the
powerful classifier. Therefore, this study will identify a distance measure that fulfills the following criteria (listed based on priority): i) A measure that works well on strokes which work well on sketching; Invariant to object transformation; ii) A measure that can be modified and added with guiding elements; iii) Low computation suitable for real time (automatic) checking purpose.

### III. Analysis of Recent Research

The analysis was conducted by way of examining the strengths and shortcomings of each measure. To iterate, works that applied the distance functions are generally used in the area of image retrieval, shape retrieval and handwritten character recognition. They commonly focus on overcoming certain problems such as computation time, invariant of image transforms, computational complexity. Secondly, the focus is to increase its retrieval accuracy. In this paper, the most recent studies that applied five different distance functions were studied. The studies are conducted by (i) Wang et. al [14] who presented new Euclidean distance for images, called IMED which image metric can be embedded in the existing image recognition methods for 2D image. ii) Nemmour and Chibani [12] who proposed new kernel for the support vector machine for handwritten digit recognition based on the Jaccard negative distance. The computation of Jaccard distance in their study determines the correlated or uncorrelated pattern based on pixel-based description where it takes into account the number of pixels in foreground and background for both patterns that leads to similar or dissimilar patterns. (iii) Chen et. al [3] comes out with the Progressive Mahalanobis distance for financial hand-written Chinese character recognition. Mahalanobis distance derived by a probability density function of multivariate normal distribution or the key is the calculation of covariance matrix. (iv) Zou and Umugwaneza [16] that proposed cosine distance function for shape-based trademark retrieval for 2D images. (vi) Mingjiang et. al [9] who proposed shape descriptor based on chord context use to measuring similarities between 2D two shapes And finally (vii) Li et. al [8] who proposed new similarity measurement method, an algorithm for object shape analysis based on Trigonometric distance.

### IV. Discussion

The performance of recognition effectiveness of different measures investigated theoretically based on the work reported by researchers. We studied works conducted by previous researchers who did some improvement on the following distance function: Chord Distance, Cosine Distance, Euclidean Distance, Mahalanobis Distance, Trigonometric distance and Jaccard Distance. The objects that these researchers work with were in many forms such as photography, hand sketches, drawing and others. The research analyzed were chosen due to their closest input type with our intended research, having a grayscale format as the grayscale format can be computed in minimal time. Moreover, most of the input used contains strokes and tiny lines.

First of all, Euclidean distance is the most commonly used in similarity measurement for its simplicity [8, 14] where it computes the difference of points in magnitude and widely applied in several applications such as character recognition, face recognition, image retrieval and others. On the other hand, the images in Euclidean distance are not necessarily similar in all features. To overcome the shortcoming in Euclidean, Zou and Umugwaneza [16] implemented a distance metrics Cosine distance to improve the formulation by multiplying the first and second coordinates and add the results from Euclidean distance dimensional profile data matrix.

Another measure that contains a simple distance function and requires minimum computation time is Jaccard Distance. In this sense, Jaccard Distance is better that it outdoes Euclidean Distance in shape recognition. Jaccard Distance is basically employed to compare two objects in a binary format and it also does not require a large set of data for training and testing purposes. It works by measuring the asymmetric information on binary variables, the comparison between two vector components. However, despite its strength and being used widely, Jaccard Distance’s disadvantage is its variant to image transformations. This distance function is commonly applied on binary data where the similarity computation computes the values of 0 and 1. The input object must be in binary form and therefore, Jaccard’s method cannot be used if any transformations are employed to the object features. Another requirement is that the size of both binary object and shapes must be of similar size. But Jaccard’s limitation can be improved by adding a pre-processing tasks prior to the computation.

At present, the Mahalanobis distance is one of the commonly distance measures used in CBIR. However, its weakness is that Mahalanobis distance’s computation needs sufficient training samples and long computation time in recognition. Computation time is one of the important distributions in character recognition and retrieval applications, therefore consuming long computation time will affect the recognition performance. Consequently, Mahalanobis distance function will be more time-consuming on large database. Eventhough Chen et al. [3] conducted a research to reduce the computation time of original Mahalanobis Distance by creating a Progressive Mahalanobis Distance to reduce computational loads at satisfactory level, its computation time still consumes more time than the Euclidean Distance. But interestingly, Mahalanobis [3] performs reliable result on financial hand-written Chinese characters that involve variations of size and density; stroke translation; stroke length and width; broken and connected stroke and others. These applicable methods contribute to the accurate automatic retrieving system in trademark images [16] handwritten digit recognition [12] automatic bank cheque processing [3], and texture (Chen and Chu, 2005). Another distance is a Cosine Distance of which whose strength is its consideration of the correlation of the features vector. This measure normalizes all feature vectors to unit length, comparing the angle between two shape features vectors. In [16] worked on trademark retrieval using
Cosine Distance function, where it shows a weakness where this measure is invariant with the scaling of the image content and treating similar images in different ways. However, Cosine Distance produces more accurate results compared to the Euclidean Distance in retrieval effectiveness.

Chord Distance also shows a high degree of retrieval accuracy than other measures on the MPEG-7 CE-1 database and the Kimia silhouettes datasets retrieval test. The strength here is that it is robust to noise and occultation. Another potency is that its capability to describe a frequency distribution of chord lengths with different orientations. Shape descriptor using Chord context [9] for shape description in CBIR showed that the measure is invariant in image translation, rotation, and scaling. However, the propose descriptor perform no special operations to resist affine transformation. Furthermore, in most researches on chord distance, it also provides poor recognition result if the input consists of strokes and lines. Likewise, Trigonometric Distance applied in the object shapes analysis for retrieval application produced better recognition rate on similarity images in different angles. Li et al. [8] uses Trigonometric distances that normalize the distance of two points in image similarity. Trigonometric’s forte is that it reduces the influence of noise and produces better recognition rate than Euclidean distance in image retrieval. However, it requires a well cleaned input whereby any noise can easily affect the recognition performance. Defining and filtering noises in real time application require efficient noise removal algorithm and involve a careful choice of feature extractions steps and which may be involve another difficult study.

Hence in a nutshell, it was found that each distance function posses its own strength weaknesses in the context of strokes and line input as well as for a real time application. In short, sigh sensitivity through small deformation may result in a large Euclidean Distance. The individual distance needs normalized distance to decrease the percentage of missed images in the results for effective image retrieval [16]. While the Mahalanobis Distance is time consuming when tested on large database, and also, limited samples cause a bad influence. Other distance functions do not resist affine transform, contributing to a major weakness in accuracy retrieved, for example, Chord Distance [9]. Considerable efforts have been made to obtain intuitively reasonable. Although there is a remaining setback, some methods have been evolved for better achievement in recognition retrieval. The proposed similarity measure using distance function contribute to strengths in flexibility against handling various type of queries and robust to noise and minor occultation.

Therefore, the Chord Distance, Cosine, Jaccard, Mahalanobis and Trigonometry are found to be better than the Euclidean distance in all courses. Unfortunately, Euclidean Distance’s hitches are on scale properties where it produces far distances; no similarity although the objects are in the normalized scale images and computed base magnitude. Therefore, Euclidean suffers from the most important drawback that makes it offer the lowest degree of recognition. On the other hand, Cosine Distance works well on the objects without rotation, translation and scale.

However, in application that receives direct input using a pen input device faces some disturbances in recognizing the object, where the pen device will show differences from the pen pressures. Different pressure between individuals [4, 11] produces significance differences in the result of recognition. The summary of previous work is shown below (Table 1).

<table>
<thead>
<tr>
<th>Distance Function</th>
<th>Strength(s)</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euclidean</td>
<td>Simple formulation</td>
<td>Low accuracy, tendency of the largest-scaled feature to dominate the others and treating all features in the same way it effects image retrieval</td>
</tr>
<tr>
<td>Jaccard</td>
<td>Minimum computation time and efficient on a binary image</td>
<td>Variant object transform</td>
</tr>
<tr>
<td>Mahalanobis</td>
<td>Accurate in classifying two groups of points, invariant to object transformation. Able to recognize strokes and lines.</td>
<td>High computation time</td>
</tr>
<tr>
<td>Cosine</td>
<td>Invariant to object transformation.</td>
<td>Requires image normalization which involves much pre processing work</td>
</tr>
<tr>
<td>Chord</td>
<td>No special operation to resist affine transformation. Works well for object in different orientations and robust to noise and occultation</td>
<td>Does not work well on thin strokes or lines</td>
</tr>
<tr>
<td>Trigonometry</td>
<td>Able to measure invariant object orientation</td>
<td>Requires accurate pre processing work on image for noise removal</td>
</tr>
</tbody>
</table>

**Conclusions**

Comparison was made on the distance measures and basically, each distance performs differently against common essential properties such as noise resistance, affine invariance, occultation invariance, statistical independence, reliability, translation, rotation, scale invariance and others. Most of the distance functions mentioned above require basic pre processing tasks on the noise removal, object scale and rotation which are important in recognizing shapes. In conclusion, we chose Jaccard and Mahalanobis to be the main references for our future work. Jaccard outshines the others by its simplicity but able to maintain higher accuracy. On the other hand, Mahalanobis is able to deal with strokes and lines which provide us likelihood to produce a more promising equation that combines both distance equations. Again, in spite of its simplicity that allows minimum computation time, Jaccard works efficiently on a binary image and Mahalanobis’s invariant to object transform works best for real time application which will, in turn, perform automatic checking on the object. Hence, our future work will focus on two dimensional sketching inputs from pen device. This may improve one of the selected distance functions and may combine it with other distance functions to improve similarity.
measures performances. It deems that new distance function will perform better on sketching images where it will be able to obtain better significance degree of accuracy and recognition rate for real time recognition for automatic classifier.

ACKNOWLEDGMENT

This work was supported by the FRGS Grant, Ministry of Higher Education, Malaysia (600-RMI/ST/FRGS 5/3 Fst (98/2010)

REFERENCES


