TAR based shape features in unconstrained handwritten digit recognition

P. AHAMED AND YOUSEF AL-OHALI
Department of Computer Science
King Saud University
P.O.B. 51178, Riyadh 11543
SAUDI ARABIA
shamapervez@gmail.com , yousef@ccis.edu.sa

Abstract: - In this research, the recognition accuracy of triangle-area representation (TAR) based shape feature is measured in recognizing the totally unconstrained handwritten digits. The TAR features for different triangles of variable side lengths that are formed by taking the combinations of different contour points were computed. The set of contour points that yielded the best features was experimentally discovered. For classification a curve matching technique is used.

Several experiments were conducted on real-life sample data that was collected from postal zip codes written by mail writers. The highest recognition result of 98.5 % was achieved on the training data set and 98.3% on the test data set.

Key-Words: - Triangle area representation, Shape descriptors, Digit recognition, Contour points, and zip codes

1 Introduction

One of the advantages of shape descriptor based features is that they provide structural description of the underlying pattern for recognition. This has attracted researcher's attention [1]- [14]. As a result, a large number of shape description techniques have evolved. These techniques are broadly classified into contour-based and region-based techniques. If a technique extracts shape features from the contour (boundary) only then it is classified as a contour-based technique, and if a technique extracts the property of a region like the area covered by a region then it is classified as the region based technique. These techniques can extract structural and global features. A contour based technique can extract structural features like chain code, polygon, B-spline and invariants while a region-based technique can extract structural features like convex hull and media axis and the like. Similarly, global features like perimeter, compactness, eccentricity, shape signature, Fourier descriptors can be extracted from the contour and global features like area and moments can be extracted from the regions [1].

In this paper, the performance of a contour-based technique referred to as triangle-area representation (TAR) signature has been measured in recognizing the totally unconstrained handwritten postal zip code digits [15] and [16].

A TAR signature reflects contour characteristics like concavity or convexity at contour points at which it is computed [2][3] [4] and [5]. In addition, it is very effective in capturing the local as well as global shape characteristics by simply varying the triangle side-lengths. Its' main advantage is that it can be used to define translation, rotation, and scale invariant shape features (also referred to as shape descriptors in this paper).

There are many variations of TAR signatures. In their paper Roh and Kweon [2] have defined TAR based shape features using five equally spaced contour points $P_1(t)$, $P_2(t)$, $P_3(t)$, $P_4(t)$ and $P_5(t)$ form a list of N contour points where the point $P_i(t)$ is the co-ordinate $(x_i(t), y_i(t))$ of the i^{th} contour point, and t denotes the set of contour points selected at time t. For each selection t=1,2,3,…,N they defined the shape invariant $I(t)$ as: $I(t) = \dfrac{\|P_5(t)P_1(t)P_4(t)\| \times \|P_5(t)P_2(t)P_3(t)\|}{\|P_5(t)P_1(t)P_3(t)\| \times \|P_5(t)P_2(t)P_4(t)\|}$,

where, $\|P_5(t)P_1(t)P_4(t)\|$, $\|P_5(t)P_1(t)P_4(t)\|$, $\|P_5(t)P_1(t)P_4(t)\|$ and $\|P_5(t)P_1(t)P_4(t)\|$ are the areas of triangles formed by joining the points: {(P5(t), P1(t), P4(t)), (P5(t) P2(t)P3(t)), (P5(t), P1(t), P3(t)) and (P5(t) P2(t) P4(t)}, and the five contour points were determined as:

$$P_1(t)=(x(t), y(t)),$$
$$P_2(t)=(x(N/5+t), y(N/5+t)),$$
$$P_3(t)=(x(2N/5+t), y(2N/5+t)),$$
$$P_4(t)=(x(3N/5+t), y(3N/5+t)),$$
$$P_5(t)=(x(4N/5+t), y(4N/5+t)).$$

They obtained the shape signature of a pattern by plotting the value $I(t)$ versus t for different values of t=1,2,3,…, N, and tested the performance this descriptor in image retrieval. Through a series of experiments, they have demonstrated that these features are invariant under rotation, translation and scale, and also stable and robust.

In a similar approach, El Rube *et. al* [5] have tested the robustness of a multi-scale triangle-area representation for 2-D shapes and discovered that the representation is least affected by noise. In further experiments, Alajlana *et. al* [3] and [4] have discovered that the TAR based shape descriptors are adaptive as they can capture the varying levels of details simply by controlling the selection of the number of contour points while generating the signature. These properties of the TAR signature based shape descriptors are the main motivation to test their performance in terms of recognition accuracy on real-life samples of unconstrained handwritten digits. For this purpose, the test data set that is described in Ahmed and Suen [4] and Abuhaiba and Ahmed and [5] is used. This data set reflects natural variations and distortions in shape, size and orientation. The data was obtained from the totally unconstrained handwritten postal zip codes that were collected by the US postal service department from the dead letter envelopes at offices across the country and they were digitized on 16 gray levels.

In the experiments, reported in here, samples possessing almost all the natural variations except breaks in shapes were used. Fig. 1 below shows representative samples from the data set.

In search of a better recognition result, stage-wise classification schemes were developed in which the first stage is the prediction stage that is referred to as prediction module hereafter. This module was designed to predict the probable classes of an unknown sample. The second stage is referred to as the recognition module. This module accepts the list of the predicted classes and determines the actual identity of the unknown sample within these classes.

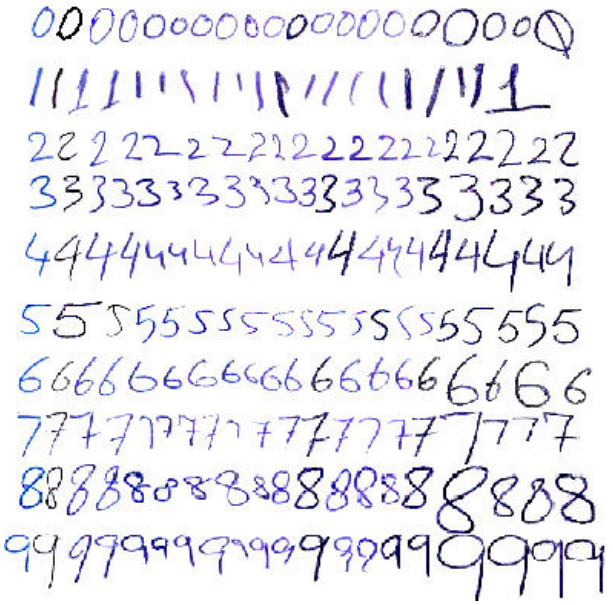

Fig.1 Data Set Representative Samples

The prediction module uses a zone-based prediction method. In this method, characteristics of the joint distribution of two TAR feature elements were studied to determine a viable zone size. The prediction helped in narrowing down the identity of the unknown patterns.

As mentioned before, the recognition stage recognizes the pattern as one of the member of the predicted pattern classes. In this stage a curve matching technique is used. The curve matching is an active area of research and its applications are being explored in searching shapes where shapes are represented as parameterized curve in two-dimensions [17]-[24]. This representation reduces the shape matching problem to the problem of matching two curves in a two-dimensional Euclidean space, and curve matching techniques, like dynamic space warping described in [3], and Fr´echet distance [18]-20],[22] and [24] which is a natural measure for matching curves, are being devised and tested.

We have developed a new curve matching technique that uses the cumulative feature curve obtained from the unknown pattern and matches it against a set of prototypical curves representing the average cumulative feature curves of the known pattern classes. The prototypical curves are feature growth curves that are estimated from the training set samples.

The stage-wise classification scheme has yielded a very promising correct recognition result of 98.5 % and 98.3% on the training and test set data respectively.

The TAR technique is presented in Section 2. The design of the prediction stage is described in Section 3 and the recognition stage in Section 4. The experimental results are represented in Section 5 and a discussion on the performance in Section 6.

2 Triangular-Area Representation (TAR)

The Triangular Area Representation (TAR) is computed from a circular list of contour points. The list is obtained from the pattern image **I**[M,N] having M rows and N columns. The point **I**[0,0] is the top leftmost point and **I**[M-1, N-1] is the bottom rightmost point. Let the list of contour points be $P_0(x_0, y_0)$, $P_1(x_1, y_1),\ldots, P_{n-1}(x_{n-1}, y_{n-1})$, where n is the total number of contour points. The point P_i (x_i, y_i), denotes the x and y co-ordinate of the ith contour point P_i . The point $P_0(x_0, y_0)$ is the fist contour point that is detected by scanning the image row-by-row starting from the point **I**[0,0]. Once the first boundary point is detected, starting from that point the rest of the contour points were detected and recorded by following the contour clockwise. The contour following process stops at the last contour point $P_{n-1}(x_{n-1}, y_{n-1})$, which is in the contour point in the 8-neighborhood of the start point $P_0(x_0, y_0)$.

To compute a TAR value at the point P_i (x_i, y_i), three contour points from the list, say P_k (x_k, y_k), P_i (x_i, y_i) and P_r (x_r, y_r), where k < i < r were selected and the area of the triangle formed by these points was computed by evaluating the determinant value A_i as shown in equation 1.

It can be easily verified that the value of A_i is 0 whenever (x_k, y_k), (x_i, y_i) and (x_r, y_r) are co-linear, and $A_i < 0$ for convex and $A_i > 0$ for concave regions (Alajlana *et. al,* 2007). The signed value of A_i provides an estimate of the distribution of the three contour points (x_k, y_k), (x_i, y_i) and (x_r, y_r).

$$A_i = \frac{1}{2}\begin{vmatrix} x_k & y_k & 1 \\ x_i & y_i & 1 \\ x_r & y_r & 1 \end{vmatrix} \qquad (1)$$

A TAR signature is a plot of triangle number versus the area of the triangle. For example, if P_k (x_k, y_k), P_i (x_i, y_i) and P_r (x_r, y_r), where k < i < r, and i=1,2,…, m, are a set of m triangles, then the plot of (i, A_i) for i=1,2,…, m is the TAR signature.

The TAR signature of the image of a handwritten digit zero is shown in Fig.2. This signature was obtained by triangles formed by three consecutive contour points (x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}) for i= *1*, …, *n*-1, where ***n*** =65 is the total number of contour points. The signature is normalized to one by dividing each TAR value A_i by M×N (M=31 and N=19).

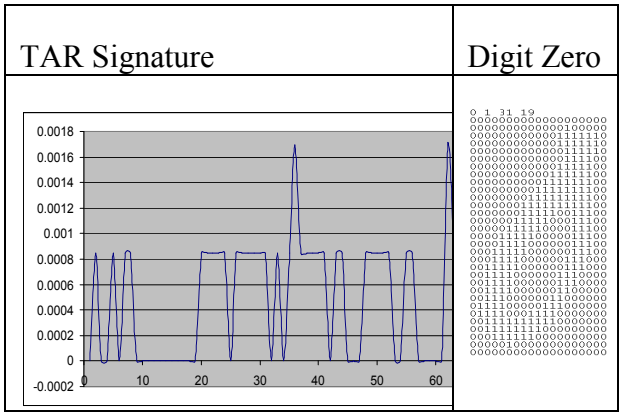

Fig.2 TAR signature of the Digit Zero Image

3 Prediction Module Design

The two best TAR features (the features that yielded the maximum correct recognition) were experimentally determined for prediction module design. Let these features be denoted as f_1 and f_2 respectively. In implementation, these features are the TAR values of triangles $P_0P_1P_2$ and $P_2P_3P_0$ as shown in Fig 3 (a). The points P_0 and P_2 are the start and midpoint in the contour and their coordinates are (x_0,y_0) and $(x_{n/2},y_{n/2})$ respectively, where 0 and n/2 are their indices in the contour point list. Similarly, the points P_1 and P_3 are at the ¼rth and ¾ rth positions of the contour point list and their coordinates are $(x_{n/4},y_{n/4})$ and $(x_{3n/4},y_{3n/4})$ respectively, where n/4 and 3n/4 are the indices of these points in the contour point list.

To predict the probable classes of an unknown pattern, the joint distribution of the feature values of f_1 and f_2 in two dimensional feature space was estimated from the training set samples, and the prediction zones were estimated. Here a zone is defined as a rectangular region. The joint distribution of f_1 and f_2 for 100 training set samples of each pattern class 0 to 9 is shown below in Fig. 3 (b). In the figure, the values of f_1 are used for horizontal axis and the values of f_2 for vertical axis. Each feature pair (f_1 , f_2) in the plot represents a pattern class.

The zones were formed by partitioning the feature axes f_1 and f_2 into n equal parts. If these parts are denoted as $f_1^1, f_1^2, f_1^3, ..., f_1^n$ and $f_2^1, f_2^2, f_3^3, ..., f_2^n$ respectively, then zones are defined as rectangles $\{(f_1^i, f_2^j)_{bl}, (f_1^i, f_2^j)_{tr}\}$, for i,j =1,2,3, …, n, where, n is the number of horizontal and vertical partitions, and $(f_1^i, f_2^j)_{bl}$ and $(f_1^i, f_2^j)_{tr}$ are the bottom-left and top-right points of the rectangle. The zones are recorded as a tuple $Z_k=\{(f_1^k, f_2^k)_{bl}, (f_1^k, f_2^k)_{tr}, L_k\}$, where the first two fields are the bottom-left and top-right corner points that define the zone and L_k is a list that contains labels of those sample patterns whose feature values f_1 and f_2 are lying in zone Z_k

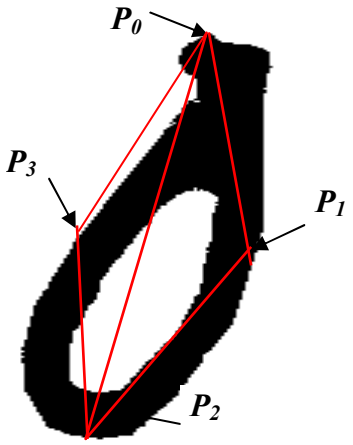

Fig. 3 (a) Selected contour points

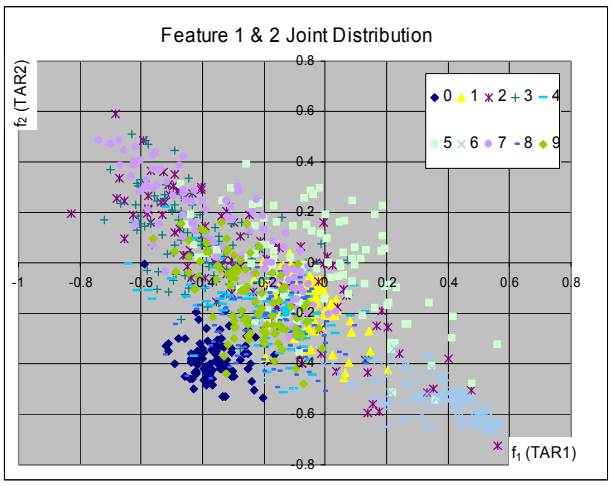

Fig.3 (b): Joint distribution of features f_1 and f_2 in digits 0, 1,2, …, 9

The zone width and height were estimated from the training set samples by estimating an optimal partition size that yielded the best zone formation i.e., no zone contained more than a pre-specified number of distinct pattern classes. After estimation, all the non-empty zones were recorded as a tuple in an array Z of tuples [Z_1 Z_2 Z_3 ,…, Z_m], where m is the total number of non-empty zones in $n \times n$ rectangles. To predict the classes for a given feature pair (f_1,f_2) the array Z is searched to locate the zone in which the feature pair lies. If it lies, say in zone Z_k, then the predicted classes are the labels list L_k that belongs to Z_k.

4 Recognition Module Design

This module uses TAR features but these features are different from those two features that are used in the prediction module. These features are extracted from the TAR signatures that are defined to capture more boundary details which are obtained by taking more than two triangles of varying side lengths at different contour locations. The extraction of these features is described in Section 4.1. For classification, the prototypes that represented the classes were generated for the TAR signatures of each class samples. The prototype generation process is described in Section 4.2. To obtain a better recognition score a curve matching based classification procedure was developed which is given in Section 4.3.

4.1 Feature Extraction

As mentioned before, triangles having sides of different lengths and are located at different contour positions were used for TAR signature generation. One of the advantages of this approach is that it gives flexibility in devising exploratory feature extraction techniques. Consequently, it helps in determining an optimal discriminatory TAR feature set from the combinations of triangles having sides of different lengths and they are located at different contour positions. In search of such a feature set for unconstrained digit recognition several experiments were conducted and the definition and extraction of the feature set that yielded better recognition result is described below.

In this case, the feature set was obtained from the contour part P_0 to $P_{n/3}$ in n point contour list P_0, P_1, …, P_{n-1}. This contour part was partitioned into $\aleph$ equal segments. By choosing different values of $\aleph$, different feature sets can be formed. Assume that the endpoints of each of the $\aleph$ segments are: ($P_{(k-1)\times \tau}$, $P_{k\times \tau}$), where k=1,2,3,…, $\aleph$, are segment numbers and $\tau = n/(3\times\aleph)$ is the segment length. For the segment number 1 ($k=1$) the segment endpoints are P_0 and $P\tau$, where P_0 is the contour start point and $P\tau$ is the point after τ points from the start point.

Similarly, for $k=\aleph$ the segment endpoints are $P_{n/3-}\tau$ and $P_{n/3}$. For each k=1,2,3,…, $\aleph$ a triangle with three contour points ($P_{(k-1)\times}\tau$, $P_{n/3+(k-1)\times}\tau$, $P_{2n/3+(k-1)\times}\tau$) can be formed and its TAR value is considered as the feature f_k. The feature values are normalized to 1.

For clarity consider Fig. 4 and assume that it contains 45 contour points (n=45). In this figure, the contour part P_0 P_{15} is divided into 3 ($\aleph$=3) segments: P_0 P_5 , P_5 P_{10} and P_{10} P_{15} where the segment length is τ =5 points. In this case, three features f_1, f_2 and f_3 can be generated by computing the areas of the triangles $\Delta(P_0P_{15}P_{30})$, $\Delta(P_5P_{20}P_{35})$ and $\Delta(P_{10}P_{25}P_{45})$ respectively.

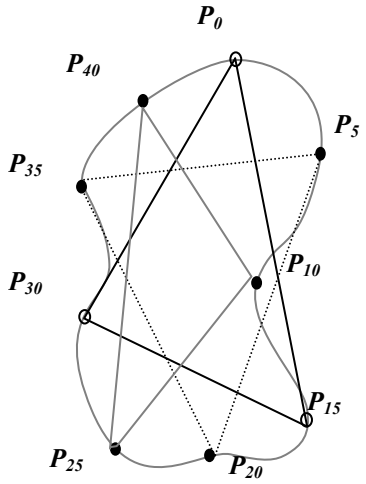

Fig. 4 TAR Signature Computation

Fig.5 shows a plot of the average of each feature value of 16 ($\aleph$=16) features observed in 1000 sample with 100 samples per digit. In the plot, digit classes are represented by different colors (see the side legend).

To study the contribution of individual feature in the accumulative growth of feature values, features f_k , k=1,2,…, $\aleph$, were transformed into a growth function form: $s(k)=\sum_{j=1}^{k} f_j$, where $s(k)$ is an increasing function as $s(k) \le s(k+1)$ for k =1,2,…, $\aleph$. The advantage of this representation is that it gives simple representation (see Fig. 6) as compared to the original signature (Fig.5).

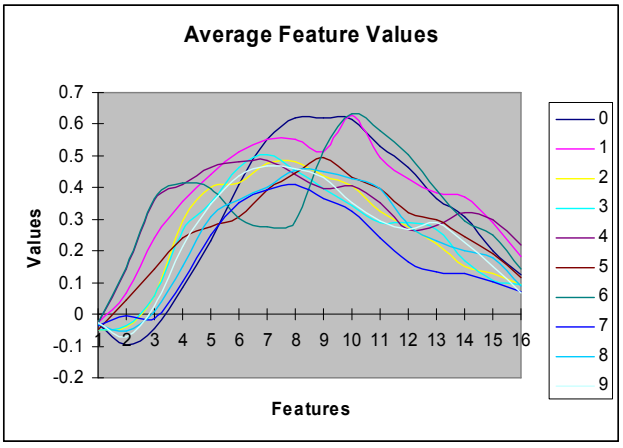

Fig.5: Plot of Average Feature Values

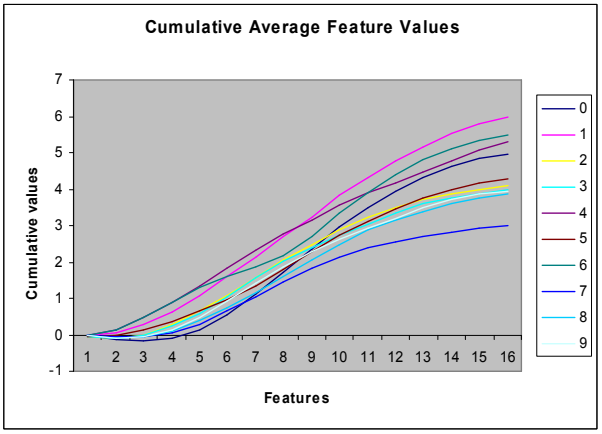

Fig.6: The Cumulative Average Feature Values

4.2 Prototype Creation

The prototypes are class representatives. In this research prototypes (one for each class) were created from the average cumulative feature values $\bar{s}^c(k)$ for k=1,2,…, $\aleph$ and for class c = 1, 2, …, m, where m is the number of classes and $\aleph$ is the number of features. The values of $\bar{s}^c(k)$ are computed as $\bar{s}^c(k)=\sum_{i=1}^{N_c} s_i^c(k)/N_c$, where $s_i^c(k)$:is the cumulative value of the k^{th} feature of the i^{th} sample in class c, and N_c is the total number of samples in class c.

As mentioned before, in this research the average feature values $\bar{s}^c(k)$ are used as prototype. One of the advantages of using the cumulative feature values is that the prototypes can be modeled using known growth functions. In that case, instead of storing the average values, only the model parameters of the growth function need to be estimated, and the prototypes can be generated in real-time and prototype curve shapes can be varied

in real-time to accommodate the variations; otherwise to capture the variations several prototypes may be required for each class.

4.3 Classification

The classification module accepts the cumulative feature curve $s(k)$ of an unknown pattern, where k=1,2,…, $\aleph$ and $\aleph$ is the total number of features. The classifier classifies $s(k)$ by comparing it against the prototypes $\bar{s}^c(k)$, k=1,2,…, $\aleph$ of each class c = 1,2,…, m. The classification criteria is to classify $s(k)$ into class c, if the value of the function Ψ ($s(k)$, $\bar{s}^c(k)$) is maximum for c = 1,2,…, m, where function Ψ ($s(k)$, $\bar{s}^c(k)$) compares the two curves $s(k)$ and $\bar{s}^c(k)$ and yields a similarity index. The step-wise curve comparing process is as follows.

1. Create a feature-wise difference table $\delta_{k,c}$ between the two curves $s(k)$ and $\bar{s}^c(k)$ by computing the value $\delta_{k,c} = \left|s(k) - \bar{s}^c(k)\right|$, for k= 1,2,…, $\aleph$ and c=1,2,…, m.
2. Create a rank table $\gamma_{k,c}$ by assigning a rank to the elements of each row of $\delta_{k,c}$ by the following rules:
   - $\gamma_{k,c}$ = 1, if $\delta_{k,c}$ is the smallest value,
   - $\gamma_{k,c}$ = 2, if $\delta_{k,c}$ is the next higher value and so on.
   - …
   - ….
   - Finally, $\gamma_{k,c} = m$ if $\delta_{k,c}$ is the highest value.

<u>Note</u>: In the ranking process a tie may occur. In which case, assign average rank to all the elements that are in tie. For example, if δ_{k,c_1}, δ_{k,c_2} and δ_{k,c_3} have equal values for the classes c_1, c_2 and c_3, then the ranks for $\gamma_{k,c_1} = \gamma_{k,c_2} = \gamma_{k,c_3} = (c_1 + c_2 + c_3,)/3$.

C. Compute $d_c = \sum_{k=1}^{\aleph} \gamma_{k,c}$

The d_c value can also be used to classify an unknown pattern into class c if d_c is the minimum of d_i, i=1, 2, 3, …, m and $c \neq i$. Experiments were conducted with this classification procedure but the result shows poor recognition performance (see Section 5). To improve the recognition performance a weighted similarity measure, described below, was developed and used.

A. Compute the rank frequency table $\Re^i$ for each class i =1, 2,3,..., m using the training set samples. In this table, each entry $\Re^i = (\gamma^i_{k,c})$ is the frequency with which the k^{th} feature ranks the training sample of some i^{th} class as the member of class c.
B. For an unknown pattern obtain $\gamma_{k,c}$ as described in steps 1 and 2 before.
C. For each class i=1,2,3,…, m compute $\phi^i_{k,c} = \frac{\gamma_{k,c}}{\gamma^i_{k,c}}$ for all $\boldsymbol{k}$ = 1,2,3,…, $\aleph$ and c = 1,2,3,…, m.
D. Compute the function $\psi(s(k), \bar{s}^c(k)) = \sum_{k=1}^{\aleph} \phi^i_{k,c}$.
E. Classify an unknown pattern having cumulative feature values $s(k)$ into class c if $\Psi(s(k), \bar{s}^c(k))$ is the maximum for the class c.

This classification procedure improved the recognition accuracy considerably (Section 5)

5 Experiments

As mentioned before, several experiments were conducted in search of the best possible recognition score. This section describes these experiments along with their respective feature extraction and classification methods, and recognition performance. Experiment 5 is the main experiment. The other experiments are described just to illustrate the performance of the techniques that provided insight that has lead to the development of improved techniques.

5.1 Experiment 1

The objective of the first experiment was to assess the applicability of TAR feature in unconstrained handwritten digit recognition. In this experiment the TAR feature of a single triangle of side length $n/3$, where n is the total number of contour points was used. The triangle was formed by selecting the contour points ($P_{i-n/3}$, P_i , $P_{i+n/3}$) for a given $i \in [n/3$, $2n/3]$ and the TAR feature was computed (as described in Fig 4 Section 4.1).

During the training phase, for every selected triangle the average feature value $\bar{f}_c$ and its variance σ_c^2 for each class c were estimated. Using these values an unknown pattern was classified as coming from the class c if its TAR value belongs to the interval $\bar{f}_c \pm k\sigma_c^2$ where k is a chosen constant. In

case, the TAR value does not belong to any class interval then the unknown pattern is declared unrecognized. Several experiments were conducted using different triangle locations and the recognition performance for different TAR values was estimated. The best correct recognition performance of 22.2% was observed on 1000 training set samples for k=0.2.

5.2 Experiment 2

In this experiment the joint distribution of two TAR feature values that were obtained from the two triangles was studied. Like experiment 1 the side length of both the triangles were taken as *n/3* and they were selected from two different locations in the range as described in experiment 1. From the training set samples, the class feature means for each $\overline{f_1^c}$ and $\overline{f_2^c}$, and class feature variances $\sigma_{1,c}^2$ and $\sigma_{2,c}^2$ were estimated and the intervals of the form $\left(\overline{f}_1^c \pm k_1\sigma_{1,c}^2, \overline{f}_2^c \pm k_2\sigma_{2,c}^2\right)$ were formed for each class *c*, where k_1 and k_2 are the chosen constants. Their values were estimated from the training set samples.

In this case, an unknown pattern having feature values f_1 and f_2 was recognized as coming from the class *c* if the feature values f_1 and f_2 fall in the class interval $\left(\overline{f}_1^c \pm k_1\sigma_{1,c}^2, \overline{f}_2^c \pm k_2\sigma_{2,c}^2\right)$.

Experiments were conducted for several triangle combinations and for each combination k_1 and k_2 values were chosen. In these experiments the triangles formed by the points P_0, P_1, P_2 and P_3 shown in Fig.3 (a) yielded the best recognition result. The horizontal triangle combinations ($\triangle P_3P_0P_1$ and $\triangle P_1P_2P_3$) yielded 42.2% and the vertical triangle combinations ($\triangle P_0P_1P_2$ and $\triangle P_2P_3P_0$) yielded 43.7% correct recognition on the 1000 training samples that were used in experiment 1.

5.3 Experiment 3

In this experiment, features obtained from the horizontal and vertical sets of triangles, described in experiment 2, were used but in two different classification schemes. Both the schemes use prediction and recognition modules. To classify a pattern, first its features were presented to the prediction module. The output of this module is a set of probable classes which is sent to the recognition module along with the feature elements. The recognition module recognizes the unknown pattern as one of the members of the predicted classes.

In the first scheme, the TAR features obtained from the vertical set of triangles were used to create the zone-based prediction module and hence to predict the probable classes. The TAR features obtained from the horizontal set of triangles were used to recognize an unknown pattern as one of the members of the predicted classes using the interval method described in experiment 2. This method improved the correct recognition percentage to 49.8%.

The second scheme was similar to the first scheme but in this case the horizontal set of triangles was used to design the prediction module and vertical set of triangles to test the recognition performance. In this case, the recognition performance was 52.3% which is slightly better. The experiment 3 was conducted on the same 1000 training set samples that were used in earlier experiments.

5.4 Experiment 4

In this experiment, several combinations of two TAR feature values were studied to improve the prediction performance, and multiple TAR features representing 4, 8, 12 and 16 features were studied to improve the recognition performance.

Like experiment 3, prediction zones were estimated and it was observed that the two vertical triangles that are shown Fig.3 (a) yielded the best prediction zones in which no zone contains more than three prediction classes. Using 1000 samples we observed that the zone height=0.05 and width = 0.05 yielded the best prediction zones.

To recognize an unknown pattern, a classifier using $d_c = \sum_{k=1}^{K} \gamma_{k,c}$ as described in Section 4.3 was used.

It was observed that 16 TAR feature set produced the best correct recognition percentage of 63.4% on the same 1000 training samples that were used in experiments 1 to 3.

5.5 Experiment 5

This experiment was designed using the prediction module and the TAR feature set of experiment 4, and the weighted rank classifier described in Section 4.3. This classifier classifies an unknown pattern having feature values $s(k)$ as the member of one of the predicated classes *c*: if $\Psi\ (s(k),\ \bar{s}^c(k))$ is maximum for that class.

This classification procedure improved the recognition accuracy significantly to 98.5% that was achieved for 16 features on 1000 training samples. The confusion matrix of this experiment is shown in Table I, which shows that the highest confusion occurs between digits 4 and 8 and the next highest confusion between 9 and 8. It looks like structural variations in digits 4 and 8, and 9 and 8 may have yielded almost similar shapes.

The recognition accuracy was tested on a larger test sample of size 2500 digits, and it was observed that the accuracy dropped to 85.12%. The confusion table of this test is shown in Table II.

Table: I. Confusion Table Training Set (1000 Samples)

Unknown digit \ Recognized As	0	1	2	3	4	5	6	7	8	9
0	100	0	0	0	0	0	0	0	0	0
1	0	98	0	0	0	0	0	0	2	0
2	0	0	100	0	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	95	0	0	0	5	0
5	0	0	0	0	0	100	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0
7	0	0	0	0	0	0	0	100	0	0
8	0	0	0	0	4	0	0	0	95	1
9	0	0	0	0	1	0	0	0	2	97

Table-II: Confusion Table Test Set (2500 Samples)

Unknown digit \ Recognized As	0	1	2	3	4	5	6	7	8	9
0	250	0	0	0	0	0	0	0	0	0
1	0	135	3	0	30	17	0	33	32	0
2	0	0	248	0	0	0	0	2	0	0
3	0	0	9	227	0	0	0	10	0	4
4	0	3	1	0	193	0	0	1	25	27
5	0	0	0	0	0	250	0	0	0	0
6	0	0	0	0	0	0	250	0	0	0
7	0	0	36	18	0	0	0	195	0	1
8	0	0	0	0	16	0	0	0	192	42
9	0	0	2	0	32	0	0	3	25	188

Table II shows that the digit 1 has the lowest recognition percentage of 54%, and most of the errors occur between digits 1 and 4, 1 and 7 and 1 and 8. To investigate the reasons, samples of digit one were examined, and it was discovered that the size and rotation variations are likely reasons because the rotated digit 1 and digits 4, 7 and 8 have majority of contour points that are lying at the right side boundary are common, while other contour points in digit one may have not contributed significant discriminatory features. Training set size might have contributed to the error. So, to measure the effect of the larger training set size, the 2500 test samples were used as training set, and 1000 sample that were used in training set were used as the test set. The confusion of table of the training set and test set are shown in Table-III & IV.

The recognition percentage on the training set is 96.06% and test set is 98.3%. The training set percentage is low because of the size and shape variations in digit one.

Table-III: Confusion Table Training Set (2500 Samples)

Unknown digit \ Recognized As	0	1	2	3	4	5	6	7	8	9
0	250	0	0	0	0	0	0	0	0	0
1	0	185	2	0	13	9	0	19	22	0
2	0	0	250	0	0	0	0	0	0	0
3	0	0	0	250	0	0	0	0	0	0
4	0	1	1	0	243	0	0	1	2	2
5	0	0	0	0	0	250	0	0	0	0
6	0	0	0	0	0	0	250	0	0	0
7	0	0	2	2	0	0	0	245	0	1
8	0	0	0	0	3	0	0	0	241	6
9	0	0	1	0	3	0	0	2	6	238

Table-IV: Confusion Table Test Set (1000 Test samples)

Unknown digit \ Recognized As	0	1	2	3	4	5	6	7	8	9
0	100	0	0	0	0	0	0	0	0	0
1	0	100	0	0	0	0	0	0	0	0
2	0	0	100	0	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	98	0	0	0	2	0
5	0	0	0	0	0	100	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0
7	0	0	1	0	0	0	0	99	0	0
8	0	2	0	0	1	0	0	0	91	6
9	0	0	0	0	0	0	0	0	5	95

6 Conclusion

This research started with a simple experiment that was designed to measure the effectiveness of the TAR features in recognizing the totally unconstrained handwritten digits. Initially, a little success was achieved but it was observed that TAR features can be defined in many ways. Thus, it can

be used as an exploratory technique to explore the best possible combination of TAR features.
In the process, a series of experiments were conducted and every experiment yielded an encouraging result that reached close to 98.5% correct recognition score which is an excellent score in the light of the data quality of the test and training sets.

The data was obtained from the totally unconstrained handwritten postal zip codes. These zip codes were collected by the US postal service department from the dead letter envelops at offices across the United Sates of America. The data reflects expected variations in writing styles, stroke thickness, writing material, ink other similar attributes. The zip-codes were digitized on the 16 gray levels using very rudimentary scanning technology that produced some structurally damaged images.

The TAR features are structural features. Extraction of these features from the broken images may not reflect the true situation. Therefore their performance on the structurally damaged samples is expected to deteriorate. However, the experimental results show that on the structurally undamaged samples having almost all the natural variations except the breaks in shapes, the technique yielded excellent results.

References:

[1] D. Zhang and G. Lu, Review of shape representation and description techniques, *Pattern Recognition*, Vol. 37, 2004, pp. 1 – 19.

[2] K. Roh and I. Kweon, 2-d object recognition using invariant contour descriptor and projective refinement, *Pattern Recognition* Vol. 31 No. 4, 1998, pp. 441–445.

[3] N. Alajlana *et al*., Shape retrieval using triangle-area representation and dynamic space warping, *Pattern Recognition* Vol. 40, 2007, pp. 1911 – 1920.

[4] Naif Alajlan, Mohamed S. Kamel and George H. Freeman, Geometry-Based Image Retrieval in Binary Image Databases, *IEEE Transactions on Pattern analysis and Machine Intelligence*, Vol. 30, No. 6, 2008, pp. 1003-1012.

[5] El Rube, N. Alajlan, M. Kamel, M. Ahmed, G. Freeman, Robust multiscale triangle-area representation for 2d shapes, in: IEEE, International Conference on Image Processing (ICIP), Genoa, Italy, September 2005, pp. 545–548.

[6] S. Belongie, J. Malik, and J. Puzicha, Shape matching and object recognition using shape contexts., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 24, 2002, pp.509.522.

[7] G. Chauang and C. Kuo, Wavelet descriptor of planar curves: Theory and applications, *IEEE Transaction on Image Processing,* Vol. 5, No. 1, 1996, pp. 56-70.

[8] E. Klassen, A. Srivastava,W. Mio, and S. Joshi, Analysis of planar shapes using geodesic paths on shape spaces., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26**,** No. 3, 2004, pp. 372-383.

[9] F. Mokhtarian and A. Mackworth, Scale-based description and recognition of planar curves and two dimensional shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8 , No. 1, 1986, pp. 34-43.

[10] F. Mokhtarian and A. Mackworth, A theory of multi-scale, curvature-based shape representation for planar curves, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14 No. 8, 1992,789-805.

[11] E.G.M. Petrakis, A. Diplaros, and E. Milios, Matching and retrieval of distorted and occluded shapes using dynamic programming, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No. 11, 2002 pp. 1501-1516.

[12] T. Sebastian, P. Klein, and B. Kimia, Recognition of shapes by editing their shock graphs., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26**,** No. 5, 2004, 550-571.

[13] Chunjing Xu, Jianzhuang Liu, and Xiaoou Tang, 2D Shape Matching by Contour Flexibility, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 1, 2009, pp. 180-186.

[14] T. Adamek and N.E. O'Connor, "A Multiscale Representation Method for Nonrigid Shapes with a Single Closed Contour," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 14, No. 5, 2004, pp. 742-753.

[15] P. Ahmed and C. Y. Suen, Computer Recognition of Totally Unconstrained Handwritten Zip Codes, *Int. J. of Pattern Recognition and Artificial Intelligence*, Vol. 1 No.1, 1987, pp. 1-15.

[16] I. S. I. Abuhaiba and P. Ahmed, "A fuzzy Graph Theoretic Approach to Recognize the Totally Unconstrained Handwritten Numerals," *Pattern Recognition*, Vol. 26, No. 9, 1993, pp. 1335-

1350.

[17] P. F. Felzenszwalb and J. Schwartz, Hierarchical Matching of Deformable Shapes, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol. 1, 2007, pp. 1-8.

[18] Kevin Buchin, Maike Buchin, Carola Wenk, Computing the Fr´echet Distance Between Simple Polygons, *22nd European Workshop on Computational Geometry*, 2006, pp. 103-106

[19] H. Alt and M. Godau, Computing the Fr´echet distance between two polygonal curves, *Internat. J. Comput. Geom. Appl.*, Vol. 5, 1995, pp. 75–91.

[20] Adrian Dumitrescu and G¨unter Rote, On the Fr´echet distance of a set of curves, *16th Canadian Conference on Computational Geometry ,Montreal, Quebec,* 2004, pp. 163-165.

[21] Xiao-Diao Chen, Linqiang Chen, Yigang Wanga, Gang Xua, Jun-Hai Yong and Jean-Claude Paul, Computing the minimum distance between two Bézier curves, *Journal of Computational and Applied Mathematics*, Vol. 229, 2009, pp. 294-301.

[22] H. Alt, C. Knauer, and C. Wenk, *Matching polygonal curves with respect to the Frechet distance*, In: Proc. 18th Int. Symp. Theoret. Aspects Comput. Sci., STACS 2001 (A. Ferreira and H. Reichel, eds.), Lect. Notes Comp. Sci., vol. 2010, Springer-Verlag, 2001, pp. 63-74.

[23] M. Cui, J. Femiani, J. Hu, P. Wonka and A. Razdan, Curve matching for open 2D curves, *Pattern Recognition Letters*, Vol. 30, 2009, pp. 1–10

[24] R. Sriraghavendra, Karthik K. and Chiranjib Bhattacharyya, Fr´echet Distance based Approach for Searching Online Handwritten Documents, *Ninth International Conference on Document Analysis and Recognition* (ICDAR*)*, Vol. 1, 2007 pp. 461 - 465.