RESEARCH ARTICLE OPEN ACCESS

Offline Signiture and Numeral Recognition in Context of Cheque

Bhagyashri Renukadas Bakshi[1], Prof. Zameer Farooqui[2]
[1]M.E. Student, Electronics and telecommunication Department,Aditya College of Engineering , Beed, India.
[2]Associate Professor & Head, Electronics and telecommunication Department,Aditya College of Engineering , Beed, India

Abstract
Signature is considered as one of the biometrics. Signature Verification System is required in almost all places where it is compulsory to authenticate a person or his/her credentials to proceed further transaction especially when it comes to bank cheques. For this purpose signature verification system must be powerful and accurate. Till date various methods have been used to make signature verification system powerful and accurate. Research here is related to offline signature verification. Shape Contexts have been used to verify whether 2 shapes are similar or not. It has been used for various applications such as digit recognition, 3D Object recognition, trademark retrieval etc. In this paper we present a modified version of shape context for signature verification on bank cheques using K-Nearest Neighbor classifier.
Keyword: Centroid, KNN Classifier, Offline Signature Verification, Shape Context.

I. INTRODUCTION

Handwritten signature of each person is unique hence it is used as one of the biometrics to authenticate that person. Today various governments and financial institutions accept signature as a legal means of verifying identity. Any offline signature verification system consists of five phases as Data Acquisition, Preprocessing, Feature Extraction, Classification and Verification. Offline signature is a handwritten signature written on a paper which is obtained by scanning through optical scanner. This scanned digital image of signature needs preprocessing to clear distortion or any noise occurred while scanning image. Preprocessing includes image normalization, binarization, thinning etc. This step is important to extract features from signature. In feature extraction a feature vector is calculated considering different features from signature. These features may include features from local parts of signature features whole signature. During classification input signature is categorized. This input signature is then tested with set of reference signatures that are trained by classifier into a database. If it matches above certain threshold then it is considered as genuine otherwise it is forged. This basic working of signature verification system is as in Fig. 1

II. BASIC CONCEPTS

Shape context is a feature descriptor used in object recognition.

A. Shape matching approaches

There are 2 basic approaches for shape matching, namely Feature based approach [1],[3],[7] which include extraction of feature points, usually edge or corner points, from the image and reduce the problem to point set matching. Brightness based approach [8], [9] make more direct use of pixel brightness and use the image itself as a feature descriptor. Brightness based approach is more robust but it takes huge computation time than feature based approach.

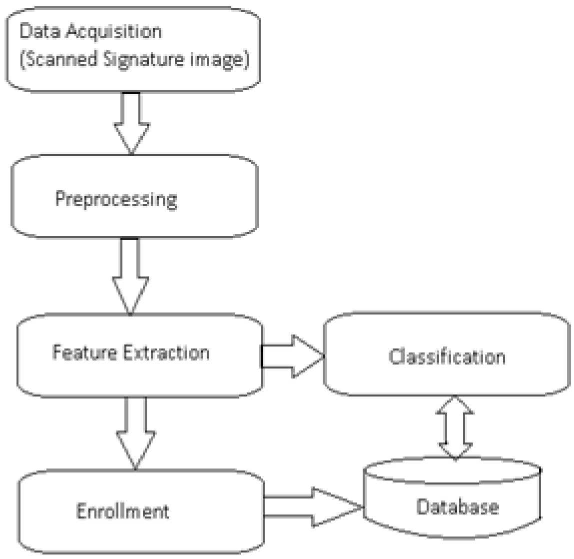

Fig. 1 Basic Signature Verification System

B. Shape Context

Shape context have been mainly used for matching similarities between shapes. In [2]. Introduced the shape contexts descriptor, which describes the coarse distribution the rest of the shape with respect to the given point on the shape. Where shape descriptor of a given point on the shape is the histogram of polar coordinates relative to all other points. For each point pi on first shape, we want to find the "best" matching point qj on the second shape. Shape context descriptor as introduced in [2] by Belongie et al. play such a role in shape matching. The basic idea is to pick n points on the contours of a shape. For each point pi on the shape, consider the n − 1 vectors obtained by connecting pi to all other points.

For the point pi, the coarse histogram of the relative coordinates of the remaining n − 1 points, hi (k) = # q != pi : (q − pi) ∈ bin(k) is defined to be the shape context of pi. The bins are normally taken to be uniform in log-polar space. The fact that the shape context is a rich and discriminative descriptor can be seen in the Fig. 2 in which the shape contexts of two different versions of the letter "A" are shown.

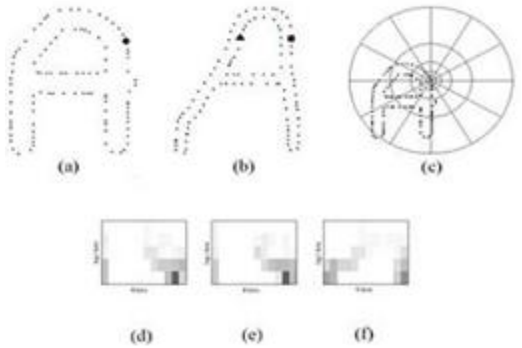

Fig. 2 (a) and (b) The Sample Edge Points of the Two Shapes, (c)Diagram of Log-Polar Bins Used to Compute The Shape Context.(d) Shape Context for the Circle (e) Shape Context for the diamond, and (f) Shape Context for the triangle.

As in Fig. 2 (d) and (e) are the shape contexts for two closely related points, they are quite similar, while the shape context in (f) is very different. Rotation invariance can be achieved in shape contexts. One way is to measure angles at each point relative to the direction of the tangent at that point (since the points are chosen on edges). This results in a completely rotationally invariant descriptor. Scale invariance will be obtained by normalizing all radial distances by the mean distance α between all the point pairs in the shape. Translational invariance comes naturally to shape context.

In [2] shape contexts matching approach consist of 3 stages:

1. solving the correspondence problem between 2 shapes

   By shape contexts descriptor and bipartite matching method.
2. Applying the correspondences to estimate an aligning transform use thin plate spline(TPS) modeling transformation and
3. Computing the distance between the two shapes as a sum of matching errors. The distance will be estimated as weighted sum of three terms: shape context distance, image appearance distance, and bending energy. It is defined as: $D=1.6Dac+Dsc+0.3Dbe$

Where shape context distance is the symmetric sum of shape context matching costs over best matching points. Appearance cost is the sum of squared brightness differences in Gaussian windows around corresponding image points. And bending energy is a measure of how much transformation is required to bring two shapes into alignment.

III. PREPROCESSING

The preprocessing step is applied both in training and testing phases. Signatures are scanned in gray. The purpose in this phase is to make signatures standard and ready for feature extraction. The preprocessing stage includes four steps: Background elimination, noise reduction, width normalization and skeletonization. The preprocessing steps of an example signature are shown in Fig. 3.

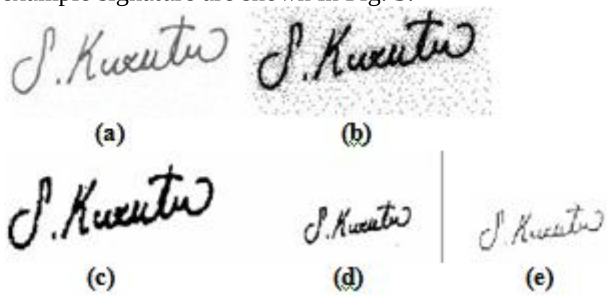

Fig 3. Preprocessing steps: (a) scanning, (b) background elimination, (c) noise reduction, (d) width normalization,(e) thinning applied signatures.

3.1 Background Elimination

Data area cropping must be done for extracting features. P-tile thresholding was chosen to capture signature from the background. After the thresholding the pixels of the signature would be "1" and the other pixels which belong to the back-ground would be "0".

3.2 Noise Reduction

A noise reduction filter is applied to the binary image for eliminating single black pixels on white background. 8-neighbors of a chosen pixel are examined. If the number of black pixels is greater than number of white pixels, the chosen pixel will be black otherwise it will be white.

3.3 Width Normalization

Signature dimensions may have intrapersonal and interper-sonal differences. So the image width is adjusted to a default value and the height will change without any change on height-to-width ratio. At the end of width normalization width dimension is adjusted to 100.

3.4 Thinning

The goal of thinning is to eliminate the thickness differences of pen by making the image one pixel thick. In this system Hilditch's Algorithm is used.

IV. FEATURE EXTRACTION

Extracted features in this phase are the inputs of training phase. The features in this system are global features, mask features and grid features. Global features provide information about specific cases of the signature shape. Mask features provide information about directions of the lines of the signatures. Grid features provide overall signature appearance information.

4.1 Global Features

Signature area is the number of pixels which belong to the signature. This feature provides information about the signature density.

Signature height-to-width ratio is obtained by dividing signature height to signature width. Signature height and width can change. Height-to-width ratios of one person's signatures are approximately equal.

Maximum horizontal histogram and maximum vertical histogram: The horizontal histograms are calculated for each row and the row which has the highest value is taken as maximum horizontal histogram. The vertical histograms are calculated for each column and the column which has the highest value is taken as maximum vertical histogram.

Horizontal and vertical center of the signature are calculated using the formulas in Eq. 1 [10].

4.2 Mask Features

Mask features provide information about directions of the lines of the signatures. The angles of the signatures have interpersonal differences. In this system 8 different 3x3 mask features are used. Each mask is taken all around the signatures and the number of 3x3 parts of the signature, which are same with the mask, is calculated.

4.3 Grid Features

Grid features are used for finding densities of signature parts [10]. In this system 60 grid features are used. Signature is divided into 60 equal parts and the image area in each di-vided part is calculated.

V. PROPOSED WORK

In this paper we propose a feature-based method using shape context descriptor. We calculate shape context with respect to centroid relative to n (possibly 100) sample points on the contour of the test signature considering centroid as a reference point of signature and then compute shape context distance between test signature and template signatures from the database after calculating the shape contexts of their centroid relative to n sample points on the contour of the signatures. This method will reduce the number of matching candidates and processing time.

VI. EXISTING SYSTEM AND ITS LIMITATIONS

Existing work include shape context which computes the radial and angular distances with respect to all points on the shape considering one sample point randomly as a reference point from the set of n sample points. The basic shape context for signature verification requires extra work for shape alignment which is time consuming task.the shape contexts of their centroid relative to n sample points on the contour of the signatures. This method will reduce the number of matching candidates and processing time.

VII. ADVANTAGES OF PROPOSED SYSTEM

Proposed system uses shape contexts for feature extraction which is more robust on complex backgrounds such as cheques. Aim of this paper is to build accurate signature verification system for bank cheques using shape contexts with modification of reference point as a centroid. ThereTransformation cost is reduced and system gives accurate results in less time.

VIII. PROPOSED SYSTEM ALGORITHM

1. Load scanned cheque image and crop the signature part.
2. Apply normalization to the cropped signature as Thinning, Binarization, rotation.
3. Apply feature extraction. That is calculate centroid of the signature as:

Given signature shape viewed as a binary function

$$f(x,y)=\begin{cases} 1 & if\ (x,y)\in D \\ 0 & Otherwise \end{cases}$$

Where D is the domain of this binary shape. Then Centroid C (gx,gy) of this shape given by:-

$$g_x= \frac{1}{6A}\sum_{i=0}^{N-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$$

$$g_x= \frac{1}{6A}\sum_{i=0}^{N-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \quad (1)$$

Where A is the area of the given shape,

$$A= \frac{1}{2}\sum_{i=0}^{N-1}(x_i y_{i+1} - x_{i+1} y_i)$$

4. Select train/template signature database.
5. Apply steps 1-3 on each signature in the train/template database.
6. For a test signature and template signatures centroid is considered as a reference point.

Compute shape context for each point from n sample points (possibly 100) on the contour of test signature and template signatures, considering centroid as a reference point.

A. Calculation of shape context

For each point pi on the first shape (shape of test signature), we want to find the "best" matching point qj on the second shape (shape of template signature). Shape context here is a set of vectors originating from a centroid of signature to all the sample points on the signature's boundary. These vectors express the configuration of the entire shape relative to the centroid. We compute a coarse histogram hi of the relative coordinates of around 100 boundary points with respect to centroid as,

$$hi(k) = \#\, q \neq pi : (q-pi) \in bin(k) \qquad (2)$$

This histogram is defined to be the shape context of pi. We use bins that are uniform in log polar space, making descriptor more sensitive to positions of nearby sample points than to those of points farther away.

7. Apply same procedure as in 6 to calculate shape contexts of template signatures from database.
8. Calculate the shape distance between each pair of points on two shapes (test signature and template signature) as a weighted sum of shape context distance and image appearance distance.

B. Calculations of Shape distance

In basic shape context implementation, shape distance is a weighted sum of shape context distance Dsc, image appearance distance Dac and the amount of transformation necessary to align the shapes i.e. bending energy Dbe. In our Proposed system shape distance between two signatures is calculated using only shape context distance and image appearance distance since we are calculating shape context distance using centroid as a reference point, no transformation is required for shape alignment. Therefore no need to calculate bending energy. Shape distance will be calculated as:

$$D = D_{sc} + 1.6 D_{ac} \qquad (3)$$

Where,

$$D_{sc}(P,Q) = \frac{1}{n}\sum_{p\in P} \arg\min_{q\in Q} C(p,q) + \frac{1}{m}\sum_{q\in Q} \arg\min_{p\in P} C(p,q) \qquad (4)$$

$$D_{ac}(P,Q) = \frac{1}{n}\sum_{i=1}^{n}\sum_{\Delta\in Z^2} G(\Delta)\left[I_P(p_i+\Delta) - I_Q(q_{\pi(i)}+\Delta)\right]^2 \qquad (5)$$

Where Ip and IQ are the grey-level images corresponding to P and Q, respectively. Δ denotes some deferential vector offset and G is a windowing function typically chosen to be a Gaussian, thus putting emphasis to pixels nearby.

C.K-NNClassifier

KNearestNeighborclassifierisbasedonnon-parametricmethodusedforclassification.Itisoneofthesimplestmachinelearningalgorithms.Itclassifiesanobjectbyamajorityvoteofitsneighbors,withtheobjectbeingassignedtotheclassmostcommonamongitsknearestneighbors(kisapositiveinteger,typicallysmall).Ifk=1,thentheobjectissimplyassignedtotheclassofthatsinglenearestneighbor.Thetrainingphaseofthealgorithmconsistsonlyofstoringthefeaturevectorsandclasslabelsofthetrainingsamples.Intheclassificationphase, *k*isauser-definedconstant,andanunlabeledvector(aqueryortestpoint)isclassifiedbyassigningthelabelwhichismostfrequentamongthe *k* training samples nearest to that querypoint.

10. Ifthe shape distancedifference <= threshold, calculate the cost matrix for computing shape contexts and find out minimumcost as asimilarity measure.

D.CalculationofShapecontextcost

Considerapoint$p_i$onthefirstshapeandthepoint$q_j$onthe secondshapethenC_{ij}willdenotethecostofmatchingthese two pointsgivenby,

$$C_{ij} = C(p_i,q_j) = \frac{1}{2}\sum_{k}\frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)} \qquad (6)$$

Where$h_i(k)$and$h_j(k)$denotetheK-binnormalizedhistogram at $p_i$and q_j, respectively.

11**.**Theoutputwillbeeithertestsignatureisauthenticatedandacceptedortestsignatureisnotauthenticatedandaccepted.Inthiswayifthecroppedsignaturefromthebankchequeisauthenticatedchequeundertestwillalso be authenticated.

InputTestSignatureNormalization
(LoadScannedCheque

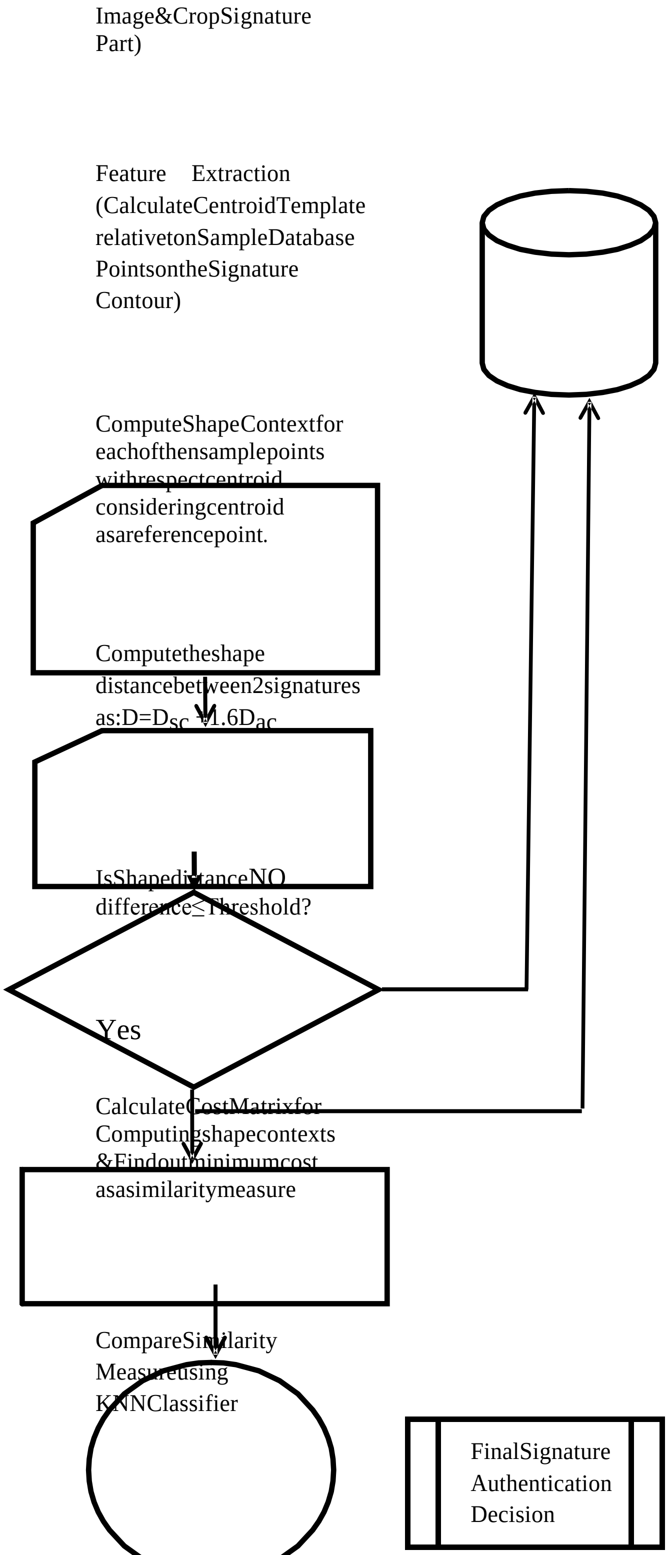

Fig. 4 SystemFlowchart

IX. CONCLUSION

In this paper we have proposed the modified shape context for offline signature recognition. The shape context feature proposed by Belongie et al. [2] computes the radial and angular distances with respect to all points on the shape. Here we have modified this concept by calculating the distance with respect to centroid of the shape. Shape context for n sample boundary points is calculated considering centroid of the signature as a reference point. K nearest neighbor classifier is used to compare Shape distance of test signature with template signatures. Since there is no alignment work needed total computation time is reduced, hence proposed work is much promising. Proposed system will help to prevent cheque frauds in banking sector.

REFERENCES

[1] Miguel A. Ferrer, J. Francisco Vargas, Aythami Morales, and Aaron Ordonez, ”*Robustness of Offline Signature Verification Based on Gray Level Features*”, IEEE Transactions on Information Forensics and Security, Vol 7,No. 3, June 2012, PP 966-977

[2] Serge Belongie, Jitendra Malik, Jan Puzicha, “*Shape Matching and Object Recognition Using Shape Contexts*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 24, No 24, April 2002, PP 509-522

[3] Greg Mori, Serge Belongie, Jitendra Malik,”*Efficient Shape Matching Using Shape Contexts*”,Vol 27,No. 11,November 2005.

[4] Marcin Adamski, Khalid Saeed, ”*Offline Signature Verification Based on Shape Contexts Using Shared and User-Specific Thresholds*”, Journal of Medical Informatics & Technologies, Vol. 22, Issn 1642-6037,Year 2013, PP 195-201

[5] ADAMSKI M., SAEED K., “*Offline Signature Identification and Verification using Noniterative Shape Context Algorithm*”, Journal of Medical Informatics and Technologies, Vol. 13, pp. 47-58, 2009.

[6] Chien-Chou Lin,Chun-Ting Chang,”*A Fast Shape Context Matching Using Indexing*”, Fifth International Conference on Genetic and Evolutionary Computing IEEE,pp.17-20,2011

[7] Persoon, E. and K.S. Fu,”*Shape Discrimination Using Fourier Descriptors*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(3):p.388-397

[8] Oren, M. et al., “*Pedestrian Detection using Wavelet Templates*”, Proceedings of IEEE computer society Conference on computer vision and pattern recognition, 1997.

[9] Moghaddam, B., T. Jebara, and A. Pentland, “*Bayesian face recognition*”, Pattern Recognition, 2000, 33(11): p.1771-1782.