

# Signature Verification for Automated Cheque Authentication System Based on Shape Contexts

Sangeeta Girish Narkhede<sup>1</sup>, Prof. Dinesh D. Patil<sup>2</sup>

<sup>1</sup>M.E. Student, Computer Science & Engineering Department,  
Shri Sant Gadge Baba College of Engineering & Technology, Bhusawal, India

<sup>2</sup>Associate Professor & Head, Computer Science & Engineering Department,  
Shri Sant Gadge Baba College of Engineering & Technology, Bhusawal, 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.

**Keywords:** 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 or features from 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

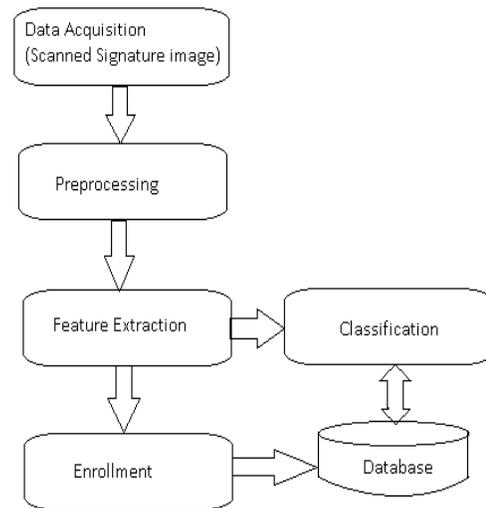


Fig. 1 Basic Signature Verification System

## II. BASIC CONCEPTS

Shape context is a feature descriptor used in object recognition. Serge Belongie and Jitendra Malik were first to propose the term in their paper "Matching with Shape Contexts" in 2000

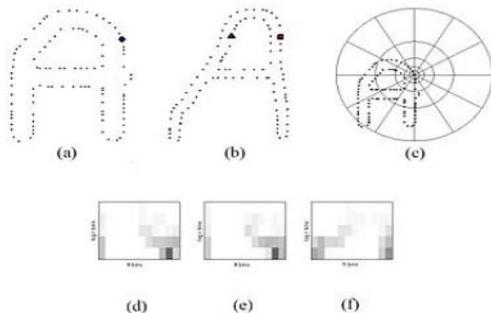
### A. Shape matching approaches

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.

### B. Shape Context

Shape context have been mainly used for matching similarities between shapes. In [2] Belongie et al. 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  $p_i$  on first shape, we want to find the “best” matching point  $q_j$  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  $p_i$  on the shape, consider the  $n - 1$  vectors obtained by connecting  $p_i$  to all other points. For the point  $p_i$ , the coarse histogram of the relative coordinates of the remaining  $n - 1$  points,  $h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\}$  is defined to be the shape context of  $p_i$ . 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  $\alpha$  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:  $D=1.6D_{ac}+D_{sc}+0.3D_{be}$

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. RELATED WORKS

Two algorithms for rapid shape retrieval presented as in [3] by Greg Mori et al. as: representative shape contexts, performing comparisons based on a small number of shape contexts, and shapemes, using vector quantization in the space of shape contexts to obtain prototypical shape pieces. A system for offline signature verification based on shape context descriptors using shared and user specific thresholds as in [4] presented by Marcin Adamski and Khalid Saeed. The feature vector is built from Shape Context Descriptors computed for selected points on skeletonized signature line. The verification process is based on the distance measure that uses Shape Context Descriptors. The presented system is evaluated using random and skilled forgeries with shared and user-specific thresholds.

Offline signature identification and verification using noniterative shape context algorithm had been presented in [5] by Adamski M., Saeed K. The system had an aim to recognize offline handwritten signatures using only one reference sample per person.

A method for fast shape context matching using indexing proposed in [6] by Chien-Chou Lin and Chun-Ting Chang. Algorithm they proposed uses the mean distances and standard deviations of shape contexts as the index of shapes. The best-fit ellipse modeling is adopted as the preprocessing for normalizing.

### IV. 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.

### 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. 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. Thereby

transformation cost is reduced and system gives accurate results in less time.

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

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

$$g_x = \frac{\sum_{x=0}^A \sum_{y=0}^A f(x,y)(x_1 + x_{1+1})(x_1 x_{1+1} - x_{1+1} x_1)}{A}$$

$$g_y = \frac{\sum_{x=0}^A \sum_{y=0}^A f(x,y)(y_1 + y_{1+1})(x_1 x_{1+1} - x_{1+1} x_1)}{A} \quad (1)$$

Where A is the area of the given shape,

$$A = \sum_{x=0}^A \sum_{y=0}^A f(x,y)(x_1 x_{1+1} - x_{1+1} x_1)$$

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  $p_i$  on the first shape (shape of test signature), we want to find the “best” matching point  $q_j$  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  $h_i$  of the relative coordinates of around 100 boundary points with respect to centroid as,

$$h_i(k) = \# \{ q \neq p_i : (q-p_i) \in \text{bin}(k) \} \quad (2)$$

This histogram is defined to be the shape context of  $p_i$ . 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  $D_{sc}$ , image appearance distance  $D_{ac}$  and the amount of transformation necessary to align the shapes i.e. bending energy  $D_{be}$ . 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:

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

Where,

$$D_{sc}(P,Q) = \frac{1}{2} \sum_{p_i \in P} \text{avg} \min_{q_j \in Q} C(p_i, q_j) + \frac{1}{2} \sum_{q_j \in Q} \text{avg} \min_{p_i \in P} C(p_i, q_j) \quad (4)$$

$$D_{ac}(P,Q) = \frac{1}{2} \sum_{p_i \in P} \sum_{q_j \in Q} G(\Delta) [I_p(p_i + \Delta) - I_q(q_j + \Delta)]^2 \quad (5)$$

Where  $I_p$  and  $I_q$  are the grey-level images corresponding to P and Q, respectively.  $\Delta$  denotes some differential vector offset and G is a windowing function typically chosen to be a Gaussian, thus putting emphasis to pixels nearby.

9. To identify the test signature, use K-Nearest-Neighbor classifier to compare its shape distance to shape distances of template signatures.

**C. K-NN Classifier**

K Nearest Neighbor classifier is based on non-parametric method used for classification. It is one of the simplest machine learning algorithms. It classifies an object by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If  $k = 1$ , then the object is simply assigned to the class of that single nearest neighbor. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the classification phase,  $k$  is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the  $k$  training samples nearest to that query point.

10. If the shape distance difference  $\leq$  threshold, calculate the cost matrix for computing shape contexts and find out minimum cost as a similarity measure.

**D. Calculation of Shape context cost**

Consider a point  $p_i$  on the first shape and the point  $q_j$  on the second shape then  $C_{ij}$  will denote the cost of matching these two points given by,

$$C_{ij} \equiv C(p_i, q_j) = \frac{|h_i(k) - h_j(k)|}{h_i(k) + h_j(k)} \quad (6)$$

Where  $h_i(k)$  and  $h_j(k)$  denote the K-bin normalized histogram at  $p_i$  and  $q_j$ , respectively.

11. The output will be either test signature is authenticated and accepted or test signature is not authenticated and accepted. In this way if the cropped signature from the bank cheque is authenticated cheque under test will also be authenticated.

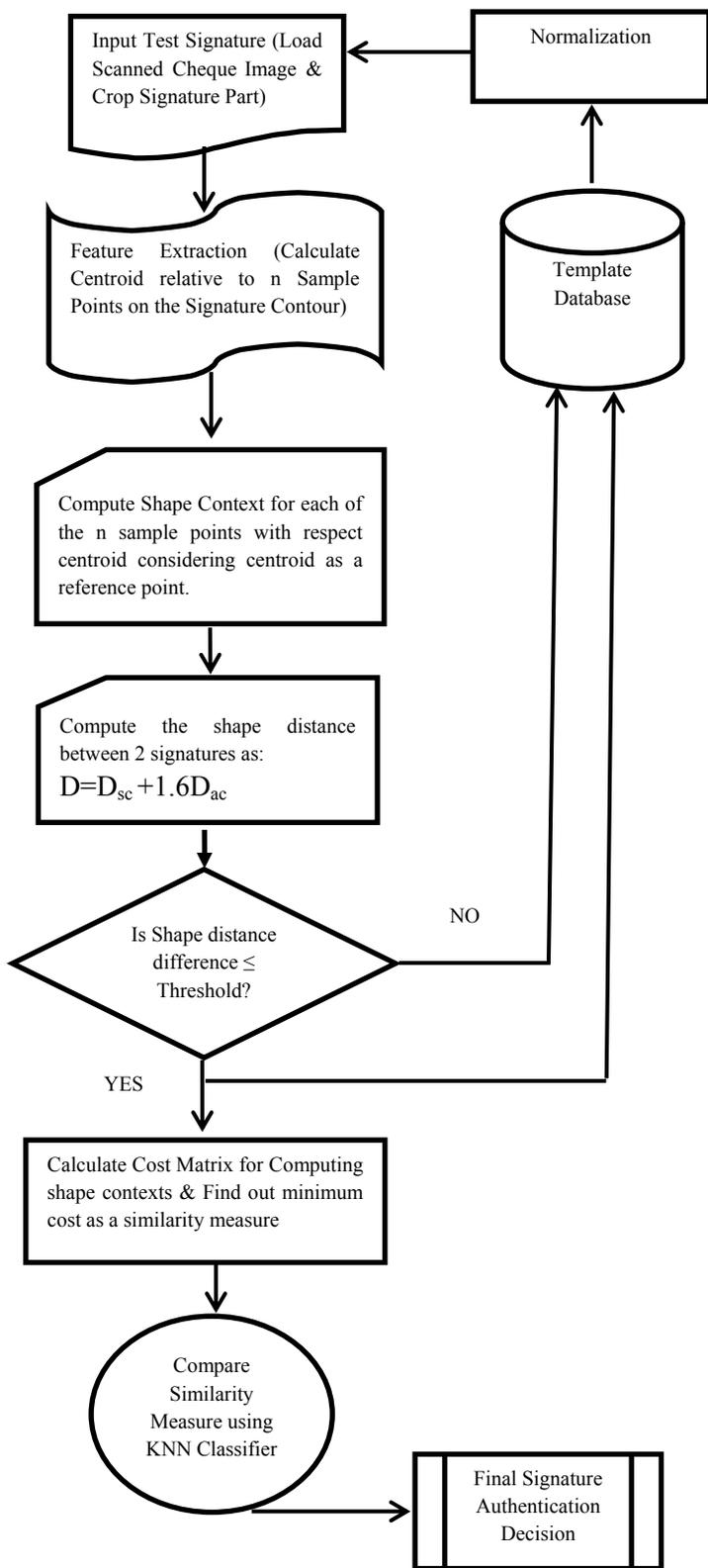


Fig. 3 System Flowchart

### VIII. 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.