International Journal of Scientific Research in Computer Science, Engineering and Information Technology
 | Volume 3 | Issue 3 | ISSN : 2456-3307

Principal Component Analysis and Support Vector Machine approach for Gujarati Handwritten Numeral Recognition

Dr. Mrs. Mamta Jagdish Baheti

Department of Computer Science, Hislop College, Nagpur, Maharashtra, India

ABSTRACT

In this paper we have proposed an algorithm for recognition of handwritten Gujarati Numerals. While reviewing the reported work, it was found that Gujarati is used all across globe including India. The proposed algorithm is applied to noisy numerals. In the algorithm we have used invariant moments as feature extraction technique and PCA and SVM as classifiers. We compared the results for both the classifiers and found that for our database SVM gave 90.55% which were better results as compared to by PCA 80.6% for Invariant moments as feature extraction technique. These results can be improved over good quality images.

Keywords : Principal Component Analysis, Support Vector Machine, Gujarati, Handwritten, Numeral Recognition

I. INTRODUCTION

Handwriting consists of simulated graphical inscription on a plane and its purpose is to exchange a few impressive words and also to increase the person reminiscence. It is a talent intimately connected to the individuality, which explain its immense changeability. In spite of all this changeability, by the instance they are five years old, the majority of the kids know how to recognize digit and writing. Tiny typescript and big typescript or rotated all are predictable without difficulty by the young ones. The typescript may be written on a jumbled background, on wrinkled document or may even be to some extent occluded. Subsequent to 50 years of study, blueprint of a wide-ranging writing identification device residue a hard to pin down object [1]. In the company of the increasing insist on to create a paperless globe, many OCR algorithms have been built-up larger than the existence [2-7]. On the other hand, a good number of OCR systems are script-dependent i.e. they can read typescript written in one exacting script only. Nowadays, there is a renewed interest in the character recognition field, which involves the recognition of both printed and handwritten characters. It is also supported now by the more compact and powerful computers and more accurate electronic equipment such as scanner, camera, electronic tablet etc, which tries to utilize advanced methodologies.

In this paper we have used invariant moments treated with SVM and PCA as a classifier. Section I describes the introduction of the paper. Section II gives the detailed literature survey about the topic. Section III describes the data collection and acquisition done for the system. Section IV elaborates the feature extraction method used in the paper along with classifiers SVM and PCA. Section V describes the algorithms proposed for comparison of classifiers for invariant moments as feature extraction technique. The results are elaborated in Section VI followed by conclusion described in Section VII.

II. LITERATURE SURVEY

Initially efforts for printed Gujarati character recognition were made by authors [8] using K-Nearest Neighbor (KNN) and minimum hamming distance classifiers with regular and invariant moments. The accuracy reported as 67% and 48% with KNN and minimum hamming distance classifiers respectively. The authors [9] implemented a template matching system for printed Gujarati character recognition using Fringe distance and achieved the overall accuracy of 72.3%. The authors have explored eight neighbours and prepared feature extraction vector and recognized digit on the basis of total. This method was applied on over 100 samples and the accuracy obtained to recognize printed Gujarati digit was approximately 95%.[10] In another reported work, [11] Neural Network is used for recognising over 3900 handwritten Gujarati digits on three different image size group i.e. 7X5, 14X10, 16X16 and the recognition rate obtained was 87.29%, 88.52%, 88.76% respectively.

In the paper [12] efforts are taken to recognize Gujarati handwritten numeral using low level stroke features like end points, curve segments, line segments, junction points along with K-nearest neighbour classifier, support vector machine and radial basis functions as classifiers resulting the accuracy of 98.46%. In another reported work, [13] zoning based feature extraction method along with Naïve Bayes classifier and multilayer feed forward neural network classifiers. The authors [14] have used orientation of strokes as feature extraction method to prepare feature set and linear support vector machine as classifier.

A reported work [15] stated in their work that for Gujarati numerals they have used three methods feature extraction; global direction with freeman chain code, Fourier descriptor and discrete cosine transform coefficients with the classifiers Knearest neighbour, Support vector machine and back propagation neural network. They reported to have highest recognition rate for DFT. In the paper [16] used Supervised Locality Preserving Projection (ESLPP) coefficients for feature extraction and MLP with BPNN as a classifier and achieved accuracy of 96%. Structural feature extraction based method for printed Gujarati characters was reported by the author [17]. They identified almost 30 strokes which can formulate almost all printed Gujarati character set. They achieved 95% of recognition rate.

A work [18], have used radial histogram and Euclidean distance classifier and achieved accuracy of 26.86%. The authors [19] used structural features like connected and disconnected components, number of end point, and number of close loop for recognition of five isolated handwritten Gujarati characters along with decision tree classifier was and achieved 88.78 % accuracy.

III. DATA COLLECTION AND ACQUISITION

As there was no standardized database available for Gujarati handwritten numerals we prepared the database of Gujarati handwritten Numerals. A specially designed sheet was prepared and on that data was collected from people who knew to write Gujarati, irrespective of their profession or gender. Ten samples of each digit from 100 persons were collected. While creating database there was no restriction of pen and color of ink. Data acquisition was done manually and the data sheet was then scanned with resolution 200 dpi using a HP 2000 flatbed scanner.

IV. FEATURE EXTRACTION AND CLASSIFIERS

Invariant Moments

Feature-based recognition of printed and handwritten characters dependent on their position, size, orientation, slant and other variations has been the goal of ongoing research. Finding efficient invariant features is the key to solving this problem. Taking into consideration the independencies of basic transformation, we use Hu's [20] moment invariants technique for feature extraction. A set of seven invariant moments can be derived from (1) out of which six moments are absolute orthogonal invariants (and one skew orthogonal invariants) [21-23].

$$
\left.\begin{aligned}
&\phi_1 = \eta_{20} - \eta_{02} \\
&\phi_2 = (\eta_{20} - \eta_{02})^2 - 4\cdot\eta_{11}^2 \\
&\phi_3 = (\eta_{30} - 3\cdot\eta_{12})^2 + (3\cdot\eta_{21} - \eta_{03})^2 \\
&\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
&\phi_5 = (\eta_{30} - 3\cdot\eta_{12})\cdot(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3\cdot(\eta_{21} + \eta_{03})^2] + \\
&(3\cdot\eta_{21} - \eta_{03})\cdot(\eta_{21} + \eta_{03})[3\cdot(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\phi_6 = (\eta_{20} - \eta_{02})\cdot(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\
&4\cdot\eta_{11}\cdot(\eta_{30} + \eta_{12})\cdot(\eta_{21} + \eta_{03}) \\
&\phi_7 = (3\cdot\eta_{21} - \eta_{03})\cdot(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3\cdot(\eta_{21} + \eta_{03})^2] + \\
&(3\cdot\eta_{12} - \eta_{30})\cdot(\eta_{21} + \eta_{03})[3\cdot(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}\right\} \quad (1)
$$

This set of moments is invariant to translation, rotation and scale change. We computed these features on the character image and its changed versions that were prepared by applying various parameters in preprocessing.

Principal Component Analysis

Principal Components Analysis (PCA) is a multivariate procedure, which rotates the data such that maximum variability's are projected onto the axes [24]. The main use of PCA is to reduce the dimensionality of a data set while retaining as much information as is possible. It computes a compact and optimal description of the data set. Data points are vectors in a multidimensional space. PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on [25-27]. PCA is theoretically the optimum transform for a given data in least square terms. The Principal Component Analysis module in proposal system generates a set of data, which can be used as features in building feature vector section. Figure 1shows a co-ordinate system (X_1, X_2). Choose a basis vector such that these vector points in the direction of max variance of the data, say (Y_1, Y_2), and can be expressed as

$$
\left.\begin{aligned}
Y_1 &= X_1 \cos\theta - X_2 \sin\theta \\
Y_2 &= X_1 \sin\theta + X_2 \cos\theta
\end{aligned}\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)
$$

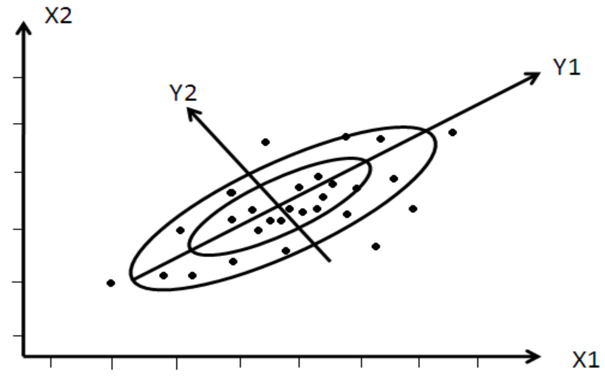

Figure 1. Ellipse Distribution with PCA

Principal Component Analysis can be used for dimensionality reduction in a data set by retaining those characteristics of the data set that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the "most important" aspects of the data. Feature vector X_k is derived from affine invariant moments where k=1…r. r is total number of images. For that prepared feature vector we can calculate mean M, difference vector R and covariance matrix Σ.

$$
M_k = \frac{1}{r}\sum_{k=1}^{r} X_k \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)
$$

$$\left.\begin{aligned} R_k &= X_k - M_k \\ \Sigma &= \frac{1}{r}\sum_{k=1}^{r} R_k R_k^t \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Where:

$$\left. X = \begin{bmatrix} X_1 \\ X_2 \\ . \\ X_r \end{bmatrix}, M = \begin{bmatrix} M_1 \\ M_2 \\ . \\ M_r \end{bmatrix}, R = \begin{bmatrix} R_1 \\ R_2 \\ . \\ R_r \end{bmatrix} \right\} \ldots\ldots\ldots\ldots\ldots(5)$$

Principal components are calculating from the eigenvectors Φ and eigenvalues λ of the covariance matrix Σ. The eigenvectors Φ are normalized, sorted in order eigenvalue, highest to lowest and transposed, to obtain transformation matrix W, where K is the number of dimensions in the dimensionally reduced subspace calculated by:

$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{l} \lambda_i} \geq p \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

where: p is assumed as threshold [30].The matrix W is given by:

$$W = \begin{bmatrix} \Phi_1^1 \ldots \Phi_1^K \\ \ldots\ldots\ldots\ldots \\ \Phi_l^1 \ldots \Phi_1^K \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

After features projection into eigenvectors space we do not use all eigenvectors, but only those with maximum eigenvalues, this gives the components in order of significance. The eigenvector associated with the largest eigenvalue is one that reflects the greatest variance in the input data. That is, the smallest eigenvalue is associated with the eigenvector that finds the least variance. They decrease in exponential fashion, means approximtely 90% of the total variance is contained in the first 5% to 10% of the dimensions [28]. The projection of X into eigenvectors space is given by:

$$\left.\begin{aligned} Y &= W(X - M) \\ Y &= \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ Y_r \end{bmatrix} \end{aligned}\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

The final data set will have less dimensions than the original [25], after all we have r column-vector for each input feature vector derived from affine invariant moments with K values:

$$Y_k = \begin{bmatrix} y_1 \\ y_2 \\ . \\ y_r \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

Support Vector Machine Classifier

A support vector machine (SVM) is a concept in computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes the input is a member of, which makes the SVM a non-probabilistic binary linear classifier. Support vector machine [29-31] is new classifier that is extensively used in many pattern recognition applications. On pattern classification problem, SVM demonstrate excellent generalization performance in practical applications. Support Vector Machines are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships. Construction of a hyper plane in a feature space requires transformation of the *n*-dimensional input vector into an *N*-dimensional feature vector, i.e.

$$\Phi : \Re^n \rightarrow \Re^N \ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

An *N*-dimensional linear separator w and a bias b are then constructed for the set of transformed vectors. Classification of an unknown vector x is done by first transforming the vector to the feature space and then computing $\text{sgn}(w \cdot \phi(x) + b)$ The vector w can be written as a linear combination of a small set of vectors in the feature space. This can be mathematically expressed as

$$w = \sum \alpha_i \cdot y_i \cdot \phi(x_i), \quad \ldots\ldots\ldots\ldots\ldots\ldots(11)$$

where the summation is over all vectors in the training set whose corresponding α's are non-zero. These vectors are called support vectors [32].

$$\text{sgn}\left(\sum_{SVs} y_i \cdot \alpha_i \phi(x) \cdot \phi(x_i) + b \right) \ldots\ldots\ldots\ldots\ldots\ldots(12)$$

The above equation is prohibitively expensive to implement directly in the feature space, since it would involve a dot product computation in a very high dimensional space. However if we could define a function in the input space which equals the dot product in the feature space, the overall complexity of the process can be drastically reduced since we significantly lower the dimensionality. The existence of such a function is guaranteed by Mercer's conditions [29, 33-34]. Such functions are referred to as a kernel in the SVM approach. . A kernel is utilized to map the input data to a higher dimensional feature space so that the problem becomes linearly separable. The kernel plays a very important role. To construct an optimal hyper plane, SVM employs an iterative training algorithm that is used to minimize an error function. It can be written as

$$\text{sgn}\left(\sum_{SVs} y_i \cdot \alpha_i K(x, x_i) + b \right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(13)$$

where K is the kernel.

There are number of kernels that can be used in Support Vector Machines models. These include linear, polynomial, radial basis function (RBF) and sigmoid:

Linear kernel function is given by

$$K(x_i, x_j) = x_i^T x_j \ldots\ldots\ldots\ldots\ldots(14)$$

Polynomial kernel function is given by

$$K(x_i, x_j) = (\sigma \cdot x_i^T x_j + coeff)^{degree} \ldots\ldots\ldots\ldots\ldots(15)$$

Gaussian Radial basis function (kernel function) is given by-

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots(16)$$

Sigmoid kernel function

$$K(x_i, x_j) = \tanh(\sigma \cdot x_i^T x_j + coeff) \ldots\ldots\ldots\ldots\ldots\ldots(17)$$

The RBF is by far the most popular choice of kernel types used in Support Vector Machines. This is mainly because of their localized and finite responses across the entire range of the real x-axis.

V. PROPOSED ALGORITHM

Algorithm based on Invariant Moments using Support Vector Machine

1. Take the input image from database
2. Resize it to 40x40
3. Complement the image
4. Binarize the image with thresholds 0.1, 0.2 0.3 and 0.4 separately
5. Dilate the binarized image with structuring elements 'line' and 'diamond'
6. Thin the image with various iterations (1,2,3,4,5,Inf)
7. Apply Invariant Moments Approach
8. Apply Image Slicing Approach (Matrix 4x4, 5x5 and 8x8)
9. Use Support Vector Machine as classifier
10. Compute the recognition rate on the basis of misclassified and classified numerals

Algorithm based on Invariant Moments using Principal Component Analysis

1. Take the input image from database
2. Resize it to 40x40
3. Complement the image
4. Binarize the image with thresholds 0.1, 0.2 0.3 and 0.4 separately
5. Dilate the binarized image with structuring elements 'line' and 'diamond'
6. Thin the image with various iterations (1,2,3,4,5,Inf)
7. Apply Invariant Moments Approach
8. Apply Image Slicing Approach (Matrix 4x4, 5x5 and 8x8)
9. Use Principal Component Analysis for dimensionality reduction
10. Use Euclidean similarity measure to find classified and misclassified numerals
11. Compute the recognition rate on the basis of misclassified and classified numerals

VI. RESULTS AND DISCUSSION

Invariant Moments Approach using Support Vector Machine as classifier

Table 1. Confusion matrix for invariant moments using SVM Classifier

	0	1	2	3	4	5	6	7	8	9
0	**77**	0	0	0	0	1	0	1	0	1
1	1	**71**	4	2	0	0	1	2	0	0
2	0	2	**72**	0	3	1	1	1	0	0
3	0	1	1	**71**	0	5	1	0	1	0
4	0	0	3	0	**73**	2	0	0	1	1
5	0	1	0	0	2	**71**	6	0	0	0
6	0	2	2	0	3	1	**72**	0	0	0
7	1	2	1	0	2	2	1	**70**	0	1
8	1	1	0	0	0	0	0	0	**78**	0
9	1	3	0	0	0	0	3	3	1	**69**

Numeral 0 is misclassified by numerals 1, 7 and 9 whereas numeral 1 shows misrecognition with 0, 2, 3, 6 and 7. For numeral 9, it is misclassified with numerals 0, 1, 6, 7 and 8. Numeral 8 have shown a recognition rate of 96.75%. Numerals 0, 4 and 8 have shown good results for SVM classifier. The overall recognition rate is found to be **90.55%.**

Invariant Moments Approach using Principal Component Analysis as dimensionality reduction and Euclidean similarity measure

Table 2. Confusion matrix for invariant moments using PCA Classifier

	0	1	2	3	4	5	6	7	8	9
0	**69**	1	3	0	0	1	2	3	0	1
1	1	**59**	4	2	0	1	1	11	1	0
2	0	3	**61**	4	3	1	1	4	2	1
3	0	1	13	**53**	0	5	5	0	3	0
4	0	0	3	0	**73**	2	0	0	1	1
5	0	2	3	2	2	**62**	6	2	1	0
6	0	1	1	2	3	8	**62**	2	0	1
7	7	2	1	0	2	2	1	**64**	0	1
8	2	1	2	0	1	0	0	0	**72**	2
9	2	1	2	0	2	0	1	3	1	**68**

For numeral 0 we have reported the recognition rate as 87%, whereas for numeral 1 is 73%. For numerals 2, 5 and 6 recognition rates is found to be 76%, 78% and 78% respectively. Numerals 3 reported to have less recognition rate as 67%. Numerals 7 and 9 have shown some better results as compared to 2, 5 and 6 as 80% and 86% respectively. Numeral 4 is reported to have maximum recognition rate 91% among all these recognition rates. At next rank in score is numeral 8 with recognition rate 90%. The overall recognition rate is **80.6%**. It has shown good results for numerals 4 and 8.

Numerals 1, 2, 3, 5, 6 and 7 have shown less recognition rate in case of PCA as compared to others. Numerals 8, 4 and 0 show optimum results for the PCA. When we see the recognition rates of SVM from table 3, we find that SVM have shown good results for all numerals. As compared to PCA, SVM have shown improved results of 10% for numeral zero. Numeral one is being recognized 15% more by SVM than that by PCA. Numeral two shows the difference of 13.25% in positive side for SVM as that for PCA. There is just a difference of approximately 0.25% for numeral four when compared by both the classifiers. Likewise numerals five and six have shown approximate 11% enhancement in results through SVM classifier. Among all recognition rates numeral 0, 1 and 2 had highest recognition rate of 97%, 95% and 90% respectively for SVM.

Table 3. Comparison of Recognition rate in % for invariant moments for SVM and PCA classifiers

Numerals	SVM	PCA
Zero	97	87
One	89	73
Two	89.25	76
Three	88.75	67
Four	90.75	91
Five	89	78
Six	89.25	78
Seven	88	80
Eight	96.75	90
Nine	87.75	86
Average	**90.55**	**80.6**

VII.CONCLUSION

As compared to overall recognition rate from Table 3, SVM have shown recognition rate of 90.55% where as PCA shows 80.6%. SVM have proved to be better as compared to PCA as per our database. The results were compared with [8, 35-37] and were found to be better because we have applied our algorithm on noisy numerals. In future we will try to improve the recognition by improvising the pre-processing methods.

VIII. REFERENCES

[1]. Luiz Eduardo Soares Oliveira (2003) "Automatic Recognition of Handwritten Numerical Strings" Ph. D Theses

[2]. G. Nagy.(2000) "Twenty years of document image analysis in PAMI", IEEE Trans. On Pattern Analysis and Machine Intelligence, 22(1):38–62.

[3]. C.Y. Suen, M. Berthod, and S. Mori (2003) "Automatic Recognition of Hand printed Characters – The State of The Art," Proc. IEEE, vol. 68, no. 4, pp. 469-487.

[4]. J. Mantas (1986) "An Overview of Character Recognition Methodologies," Pattern Recognition, vol. 19, no. 6, pp. 425-430.

[5]. V.K. Govindan and A.P. Shivaprasad (1990) "Character Recognition – A Review," Pattern Recognition, vol. 23, no. 7, pp. 671-683.

[6]. S. Mori, C.Y. Suen, and K. Yamamoto, "Historical Review of OCR Research and Development," Proc. IEEE, vol. 80, no. 7, pp. 1029-1058, Jul. 1992.

[7]. H. Bunke and P.S.P. Wang (eds.), Handbook of Character Recognition and Document Image Analysis, World Scientific Publishing, Singapore, 1997.

[8]. S. Antani, L. Agnihotri, "Gujarati character recognition", in: Document Analysis and Recognition, 1999. ICDAR'99. Proceedings of the Fifth International Conference on, IEEE, pp. 418 – 421.

[9]. Prof S K Shah, A Sharma, "Design and implementation of optical character recognition system to recognize gujarati script using template matching",IE(I) Journal ET, Vol 86 , (January 2006).

[10]. Hetal R. Thaker, Vivek J. Vyas, and Vaishali G. Sanghvi (2016) "Classification and Recognition of Gujarati Numeral using Nearest Neighbour Approach", International Journal of Scientific

Research in Science, Engineering and Technology, vol. 2(3), pp 49-52.

[11]. A. R. Vasant, “Performance Evaluation of Different Image Sizes for Recognizing Offline Handwritten Gujarati Digits using Neural Network Approach,” in 2012 International Conference on Communication Systems and Network Technologies, 2012, pp. 271– 274.

[12]. Mukesh M. Goswami, Suman K. Mitra, “Offline handwritten Gujarati numeral recognition using low- level strokes”, International Journal of Applied Pattern Recognition , volume 2 issue 4, DOI: http://dx.doi.org/10.1504/IJAPR.2015.075955

[13]. A. K. Sharma, D. M. Adhyaru, T. H. Zaveri and P. B. Thakkar, "Comparative analysis of zoning based methods for Gujarati handwritten numeral recognition," 2015 5th Nirma University International Conference on Engineering (NUiCONE), Ahmedabad, 2015, pp. 1-5. doi: 10.1109/NUICONE.2015.7449632

[14]. R. Nagar and S. K. Mitra, "Feature extraction based on stroke orientation estimation technique for handwritten numeral," Advances in Pattern Recognition (ICAPR), 2015 Eighth International Conference on, Kolkata, 2015, pp. 1-6. doi: 10.1109/ICAPR.2015.7050654

[15]. A. N. Vyas and M. M. Goswami, "Classification of handwritten Gujarati numerals," Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on, Kochi, 2015, pp. 1231-1237. doi: 10.1109/ICACCI.2015.7275781

[16]. Mandar Chaudhary,Gitam Shikkenawis, Suman K. Mitra Mukesh Goswami, “Similar looking gujarati printed character recognition using locality preserving projection and artificial neural networks",Third International Conference on Emerging Applications of Information Technology (EAIT),IEEE,(2012), 978-1-4673-1827-3/12.

[17]. Mukesh Goswami and Suman Mitra, “Structural feature based classification of printed gujarati characters",Springer-Verlag Berlin Heidelberg,(2013)LNCS 8251, P. Maji et al. (Eds.): PReMI 2013, LNCS 8251, pp. 82–87, 2013.

[18]. Lipi Shah, Ripal Patel, Shreyal Patel, Jay Maniar, “Handwritten Character Recognition using Radial Histogram",International Journal of Research in Advent Technology,(2014) Vol.2, No.4, E-ISSN: 2321-9637.

[19]. Hetal R. Thaker , C. K. Kumbharana, “Structural feature extraction to recognize some of the offine isolated handwritten Gujarati characters using decision tree classifier", International Journal of Computer Applications(0975 – 8887),(August 2014) Volume 99 – No.15

[20]. M. K. Hu, “Visual Pattern Recognition by Moment Invariants”, IRE Trans. On Information Theory, Vol. IT-8, pp. 179-187, 1961.

[21]. Trier O D, Jain A K, Taxt T(1996) Feature extraction methods for character recognition- a survey. Pattern Recognition 29:641-662.

[22]. Dhandra BV, Hangarge M, Hegadi R, Malemath V S (2006) Handwritten Script Identification Using Fuzzy K-Nearest Neighbor. Proc. of IEEE 1st ICSIP II: 587-591.

[23]. Andrews M J (1999) Moment Representation of blobs in 2-D Intensity Images, http//wwwgweepnet/~rocko/Moment/ Accessed 15 June 2007

[24]. Ramteke R.J., Borkar P.D. and Mehrotra; “Recognition of isolated Marathi Handwritten Numerals: An Invariant Moments Approach” Proceedings of the International Conference on Cognition and Recognition 2005, pp 482-489

[25]. Smith L. I.. A Tutorial on Principal Components Analyzing – Discussion and singular value decomposition, http://www.cs.otago.ac.nz/cosc453/

student_tutorials/principal_components.pdf (2002).

[26]. Shlens J. A Tutorial on Principal Component Analysis, available at: http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition _jp.pdf (2005).

[27]. Jolliffe I. T. Principal Component Analysis, Springer Series in Statistics, 2nd ed., Springer. (2002)..

[28]. Kim K. Face Recognition using Principle Component Analysis, DCS, University of Maryland. (2003).

[29]. V. Vapnik, "The Nature of Statistical Learning Theory" Springer Verlag, 1994.

[30]. C. Burges, "A Tutorial on support Vector machines for pattern recognition" Data mining and knowledge discovery, vol.2,1998,pp.1-43.

[31]. V.N. Vapnik, Statistical Learning Theory. John Wiley and sons, 1998..

[32]. http://sybil.ece.ucsb.edu/pages/poorman3dreg/index.html

[33]. Nello Cristianini and John Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods", Cambridge University Press, 2000.

[34]. B. Schölkopf, Support Vector Learning., Ph.D. Thesis, R. Oldenbourg Verlag Publications, Munich, Germany, 1997.

[35]. Prasad, J.R. ; Kulkarni, U.V. ; Prasad, R.S. (2009) Template Matching Algorithm for Gujarati Character Recognition In Proc. Of 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET), pp 263 - 268

[36]. Desai A A (2010) Gujarati handwritten numeral optical character reorganization through neural network. Pattern Recognition 43: 2582-2589.

[37]. Desai A. A. (2010) Handwritten Gujarati Numeral Optical Character Recognition using Hybrid Feature Extraction Technique. IPCV pp 733-739

Cite this article as :

Dr. Mrs. Mamta Jagdish Baheti, "Principal Component Analysis and Support Vector Machine approach for Gujarati Handwritten Numeral Recognition", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 3, Issue 3, pp.2139-2147, March-April-2018.

Journal URL : http://ijsrcseit.com/CSEIT1952295