

# An Efficient Method for Similarity Measure in Independent PCA based Classification

Dr. H. B. Kekre

Professor, Computer Engineering
Mukesh Patel School of Technology Management and Engineering,
NMIMS University, Vileparle(w) Mumbai 400–056, India
hbkekre@vahoo.com

Dr. Tanuja K. Sarode
Associate Professor, Computer Engineering,
Thadomal Shahani Engineering College, Bandra(W),
Mumbai 400-050, India
tanuja 0123@yahoo.com

Jagruti K. Save
Ph.D. Scholar, MPSTME, NMIMS University,
Associate Professor, Computer Engineering
Fr. C. Rodrigues College of Engineering, Bandra(W), Mumbai 400-050, India
jagrutik\_save@yahoo.com

## **ABSTRACT**

The paper presents a new approach of finding nearest neighbor in image classification algorithm by proposing efficient method for similarity measure. Generally in supervised classification, after finding the feature vectors of training images and testing images, nearest neighbor classifier does the classification job. This classifier uses different distance measures such as Euclidean distance, Manhattan distance etc. to find the nearest training feature vector. This paper proposes to use Mean Squared Error (MSE) to find the nearness between two images. Initially Independent Principal Component Analysis (PCA),which we discussed in our earlier work, is applied to images of each class to generate Eigen coordinate system for that class. Then for the given test image, a set of feature vectors is generated. New images are reconstructed using each Eigen coordinate system and the corresponding test feature vector. Lowest MSE between the given test image and new reconstructed image indicates the corresponding class for that image. The experiments are conducted on COIL-100 database. The performance is also compared with distance based nearest neighbor classifier. Results show that the proposed method achieves high accuracy even for small size of training set.

## Indexing terms/Keywords

Feature Vector, Nearest neighbor classifier, Principal Component Analysis, Similarity Measures, Supervised Classification.

# Council for Innovative Research

Peer Review Research Publishing System

Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 9, No 2 editor@cirworld.com www.cirworld.com, member.cirworld.com

**1099** | Page July 20, 2013



#### I. INTRODUCTION

With the explosion of image data on the Internet and the availability of large-scale image databases automatically classifying these large collections of images is becoming an important challenge. The recognition or classification of images can provide valuable information for image retrieval, analysis[1] and image annotation. Although the classification of images has been researched for long, the involved technologies such as feature extraction and matching still require substantial advancements. In image classification, an image is classified according to its visual content. Classification algorithms employ two phases of processing: training and testing. In the initial training phase, characteristic properties of typical image features are found out and, based on these, a unique description of each classification category, i.e. training class, is created. In the testing phase, these feature-space partitions are used to classify image features. The features should have two properties: 1. Discriminatory (image features in different training class should have significantly different descriptions) and 2. Reliable (all image features within a training class should share the common definitive descriptions of that group). Generally image features are represented with feature vector  $(v_1, v_2, ..., v_n)$ . This representation allows us to consider each image feature vector as occupying a point in n dimensional training class sub-space. Then the classification problem is that of determining to which sub-space class each test feature vector belongs. Feature extraction is most critical because this process directly influence the efficacy of the classification task. Lot of research has been done on the topic of feature extraction. There are different types of features such as color, shape, texture[2][3][4] ,low level, high level[5][6][7], local and global features[8][9]. Feature extraction can be done in many ways [10][11][12]. Feature extraction follows by various classification methods like nearest neighbor classifier[13], artificial neural network[14], support vector machine[15], genetic algorithm[16] and so on. In nearest neighbor classification, some distance measure such as Euclidean distance is used for comparing feature vectors. However, this may not give good results. A number of recent approaches in computer vision compare images using measures of similarity that are not Euclidean, and in fact not even metric. Much research in psychology suggests that human similarity judgments are also not metric [17][18]. The contribution of this paper is in proposing a new and effective method of classification with non Euclidean similarity measures. The paper uses independent PCA[19] for generation of new Eigen subspaces and feature vectors. The paper is organized as follows: section II describes Principal Component Analysis (PCA), section III gives proposed algorithm. Results are given in section IV and conclusions in section V followed by references in section VI.

## II. PRINCIPAL COMPONENT ANALYSIS (PCA)

Given a set of multivariate measurements, the purpose of PCA is to find a smaller set of variables with less redundancy, that would give as good a representation as possible. It reduces the dimensionality of the description by projecting the points onto the principal axes, where orthonormal set of points are in the direction of maximum covariance of the data. PCA was invented in 1901 by Karl Pearson[20], as an analogue of the principal axes theorem in mechanics; it was later independently developed (and named) by Harold Hotelling [21] in the 1930s. PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Hence PCA has a lot of applications in image processing. Examples of its many applications include data compression, image processing, visualization, exploratory data analysis, pattern recognition etc. A complete discussion of PCA can be found in book [22]. The procedure to get the principal components of the data set by PCA method is as follows:

Step 1: Subtract the mean: Subtract the mean from each of the data dimensions to obtain a zero-mean data set. Step 2: Calculate the covariance matrix from zero mean data: It is a square, symmetric matrix with diagonals are the variances, off-diagonals are the covariances.

Step 3: Calculate the Eigen vectors and Eigen values of the covariance matrix.

Step 4:Once Eigenvectors are found from the covariance matrix, the next step is to order them by Eigen value, highest to lowest. This gives the components in order of significance. The eigenvector with the highest Eigen value is the principle component of the data set.

### III. PROPOSED ALGORITHM

The proposed method can be divided into 3 sections. Section A gives the step wise procedure to generate Eigen images in turn Eigen coordinate system for each class. Section B explains the calculation of feature vectors of test image. Finally section C gives the method of classification of test image.

## A. Generation of Eigen images for each training class

Let us consider that there are 'P' classes of training images, each class having 'n' images. All images are converted into gray scale. The size of each image is 128 x 128. The following procedure is to be applied for each training class.

Step1 : Find the average image  $'I_{avg}'$  .

Step2 : Find zero mean training images( $Z_1...Z_n$ ) as shown in Fig.1



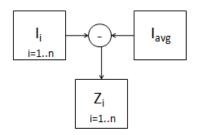


Fig.1 Finding zero mean images

Step 3: Zero mean images are converted into one dimensional vectors and organized in a matrix form (refer Fig.2 ).

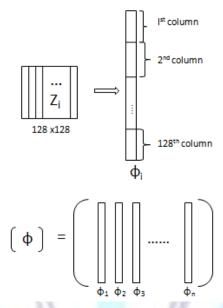


Fig. 2 Matrix of zero mean images

Step 4: Calculate the covariance matrix A by formula shown in Fig.3

Fig. 3 Calculation of Covariance matrix

Step 5: Solve the equation 1 to get Eigen values and Eigen vectors. Arrange the Eigen vectors  $(V_1..V_n)$  in decreasing order of corresponding Eigen values. For COIL-100 database, it has been observed that the lowest Eigen value is very low compared to other Eigen values. So it is not considered in further calculation.

$$[A - \lambda I]V = 0 \tag{1}$$

Step 6: Construct Eigen images  $(F_1..F_{n-1})$ . The steps are shown in Fig. 4 and Fig. 5. These images are orthogonal and hence they are used as coordinate system.



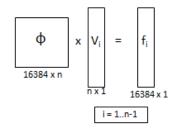


Fig. 4 Construction of Eigen Images: step 1

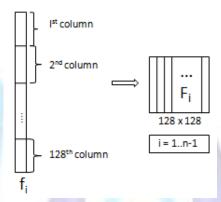


Fig. 5 Construction of Eigen Images: step 2

## **B.**Feature vectors generation for test image

Test image is converted into gray scale. Find zero mean test image for each class by subtracting average image of that class from test image as shown in Fig. 6.

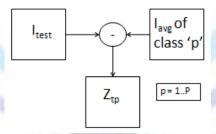


Fig.6 Finding zero mean test images

Calculate the feature vector of test image for each class 'p' using the equation 2.

$$W_{testp} = \begin{bmatrix} w_{1p} \\ w_{2p} \\ \vdots \\ w_{(n-1)p} \end{bmatrix} \quad for \ p = 1..P$$
 (2)

where each coefficient  $w_{jp}$  is given in equation 3.



$$w_{jp} = \frac{1}{\mu_{jp}} \sum_{x,y} Z_{tp}(x,y) F_{j}(x,y)$$

$$where \mu_{jp} = Cumulative \ energy \ of \ the$$

$$eigen \ image \ F_{j} \ of \ class \ 'p'$$

$$Z_{tp} = Zero \ mean \ test \ image \ for \ class \ 'p'$$

$$for \quad p = 1..P$$

$$j = 1..n - 1$$
(3)

## C.Classification of test image

Procedure in section A gives average image and Eigen images( $F_1..F_{n-1}$ ) for each class. Procedure in section B gives 'P' feature vectors, each of size (n-1), for test image. The following procedure explains the classification of test image to an appropriate training class.

Step 1. The new test image is reconstructed by multiplying Eigen images with feature vector coefficient and adding them with average image as shown in Fig.7 for each class 'p' where p=1 to 'P'.

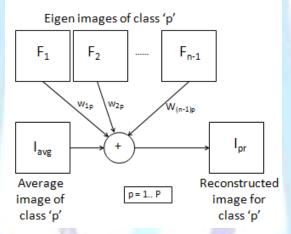


Fig.7 Reconstruction of new image

Step 2. Fig. 8 shows that, MSE(mean squared error) between test image and reconstructed image for each class is to be calculated .

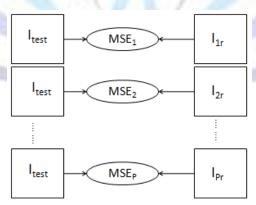


Fig.8 MSEs of 'P' classes

Step3. Lowest MSE indicate the class for test image. The test image is assigned to that class.(refer Fig.9). The proposed algorithm is applied to all test images.



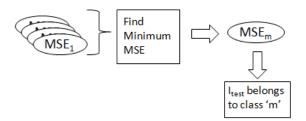


Fig.9 Classification of test image

## IV. RESULTS

The implementation of the proposed method is done in MATLAB 7.0 using a computer with Intel Core i5, CPU (2.50GHz and 6 GB RAM). The proposed technique is tested on the COIL-100[23] image database. Columbia Object Image Library (COIL-100) is a database of color images of 100 objects. The objects have a wide variety of complex geometric and reflectance characteristics. The objects were placed on a motorized turntable against a black background. The turntable was rotated through 360° to vary object pose with respect to a fixed color camera. Images of the objects were taken at pose intervals of 5 degrees. This corresponds to 72 poses per object. Thus this database has total 7200 images(100 object classes each of 72 images). Fig. 10 shows the sample images of coil database objects. Fig.11 shows the sample of different images in a particular class.

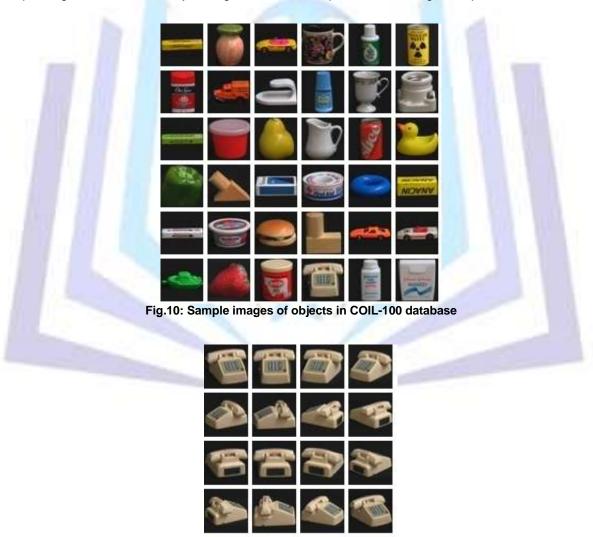


Fig.11: Sample images in one class of COIL-100 database

Training is done using 10 images per class. Thus 1000 images are used for training purpose and remaining 6200 images are used for testing purpose. The training images are orderly selected from database at an angle of 35°,70°,105°,140°,....so on. The performance is shown in table I. Fig.12 shows the number of classes in different accuracy ranges. The accuracy for each class is calculated as given in equation 4.



$$\% Accuracy = \frac{(Number of correctly classified images per class \times 100)}{Number of testing images per class}$$
(4)

Table I: Number of correctly classified images per class

Training Set: 10 images per class

Coil database Object number	Number of Correctly classified images (out of 62 per class)						
1	50	26	62	51	59	76	48
2	<mark>62</mark>	27	44	52	<mark>62</mark>	77	60
3	62	28	62	53	<mark>62</mark>	78	62
4	<mark>62</mark>	29	<mark>62</mark>	54	56	79	56
5	<mark>62</mark>	30	62	55	<mark>62</mark>	80	57
6	55	31	51	56	<mark>62</mark>	81	<mark>62</mark>
7	<mark>62</mark>	32	62	57	<mark>62</mark>	82	60
8	43	33	62	58	<mark>62</mark>	83	<mark>62</mark>
9	<mark>62</mark>	34	62	59	<mark>62</mark>	84	33
10	<mark>62</mark>	35	62	60	<mark>62</mark>	85	<mark>62</mark>
11	<mark>62</mark>	36	62	61	57	86	<mark>62</mark>
12	59	37	58	62	59	87	<mark>62</mark>
13	61	38	57	63	61	88	<mark>62</mark>
14	<mark>62</mark>	39	62	64	<mark>62</mark>	89	<mark>62</mark>
15	51	40	<mark>62</mark>	65	55	90	<mark>62</mark>
16	<mark>62</mark>	41	<mark>62</mark>	66	62	91	55
17	<mark>62</mark>	42	62	67	32	92	61
18	<mark>62</mark>	43	62	68	54	93	<mark>62</mark>
19	53	44	50	69	44	94	<mark>62</mark>
20	<mark>62</mark>	45	62	70	<mark>62</mark>	95	<mark>62</mark>
21	49	46	54	71	<mark>62</mark>	96	56
22	<mark>62</mark>	47	62	72	<mark>62</mark>	97	<mark>62</mark>
23	56	48	62	73	<mark>62</mark>	98	43
24	<mark>62</mark>	49	62	74	<mark>62</mark>	99	<mark>62</mark>
25	<mark>62</mark>	50	<mark>62</mark>	75	62	100	54

Note: Numbers in green indicate 100% accuracy.

Observations: Out of 100 object classes, 65 object classes give 100% accuracy

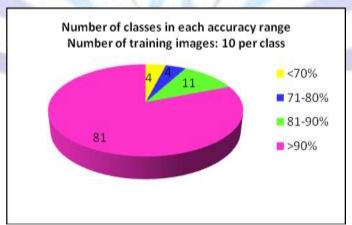


Fig. 12 Accuracy performance of 100 classes when 1000 training images are used.

Then training set size is increased from 10 to 18 per class. These images are orderly selected from database at an angle 0°,20°,40°,60°,....so on. Remaining 54 images per class are used for testing purpose. So now training set includes 1800 images and testing set includes 5400 images. Table II shows the individual class performance. Fig.13 shows the number of classes in different accuracy ranges.



Table II: Number of Correctly classified images per class
Training set: 18 images per class

Coil database Object number	Number of Correctly classified images (out of 54 per class)	Coil database Object number	Number of Correctly classified images (out of 54 per class)	Coil databas e Object number	Number of Correctly classified images (out of 54 per class)	Coil database Object number	Number of Correctly classified images (out of 54 per class)
1	49	26	<mark>54</mark>	51	<mark>54</mark>	76	48
2	<mark>54</mark>	27	47	52	<mark>54</mark>	77	<mark>54</mark>
3	<mark>54</mark>	28	<mark>54</mark>	53	<mark>54</mark>	78	<mark>54</mark>
4	<mark>54</mark>	29	<mark>54</mark>	54	<mark>54</mark>	79	53
5	<mark>54</mark>	30	<mark>54</mark>	55	<mark>54</mark>	80	<mark>54</mark>
6	52	31	48	56	<del>54</del>	81	<mark>54</mark>
7	<mark>54</mark>	32	<mark>54</mark>	57	<mark>54</mark>	82	<mark>54</mark>
8	50	33	<mark>54</mark>	58	<mark>54</mark>	83	<mark>54</mark>
9	<mark>54</mark>	34	<mark>54</mark>	59	<mark>54</mark>	84	37
10	<mark>54</mark>	35	<mark>54</mark>	60	<mark>54</mark>	85	<mark>54</mark>
11	<mark>54</mark>	36	<mark>54</mark>	61	53	86	<mark>54</mark>
12	<mark>54</mark>	37	<mark>54</mark>	62	<mark>54</mark>	87	<mark>54</mark>
13	<mark>54</mark>	38	53	63	<mark>54</mark>	88	<mark>54</mark>
14	<mark>54</mark>	39	<mark>54</mark>	64	. <mark>54</mark>	89	<mark>54</mark>
15	48	40	<mark>54</mark>	65	49	90	<mark>54</mark>
16	<mark>54</mark>	41	<mark>54</mark>	66	<del>54</del>	91	50
17	<mark>54</mark>	42	<mark>54</mark>	67	33	92	<mark>54</mark>
18	<mark>54</mark>	43	<mark>54</mark>	68	52	93	<mark>54</mark>
19	53	44	53	69	49	94	<mark>54</mark>
20	<mark>54</mark>	45	<mark>54</mark>	70	<mark>54</mark>	95	<mark>54</mark>
21	49	46	51	71	<mark>54</mark>	96	51
22	<mark>54</mark>	47	<mark>54</mark>	72	<mark>54</mark>	97	<mark>54</mark>
23	53	48	<mark>54</mark>	73	<del>54</del>	98	40
24	<mark>54</mark>	49	<mark>54</mark>	74	<del>54</del>	99	<mark>54</mark>
25	<mark>54</mark>	50	<mark>54</mark>	75	<mark>54</mark>	100	<mark>54</mark>

Note: Numbers in green indicate 100% accuracy.

Observations: Out of 100 object classes, 77 object classes give 100% accuracy

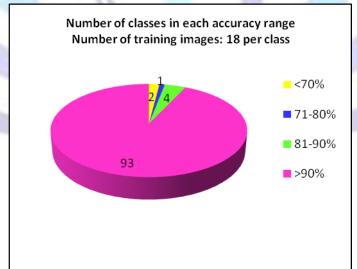


Fig. 13 Accuracy performance of 100 classes when 1800 training images are used.

Then the number of training images increased from 18 to 24. These images are orderly selected from database at an angle 0°,15°,30°,45°,....so on. So 2400 images are used for training and the rest are used for testing. Table III shows results in the case and Fig. 14 shows the corresponding accuracy range performance.



Table III: Number of Correctly classified images per class
Training set: 24 images per class

Coil database Object number	Number of Correctly classified images (out of 48 per class)	Coil databas e Object number	Number of Correctly classified images (out of 48 per class)	Coil database Object number	Number of Correctly classified images (out of 48 per class)	Coil database Object number	Number of Correctly classified images (out of 48 per class)
1	<mark>48</mark>	26	<mark>48</mark>	51	<mark>48</mark>	76	46
2	<mark>48</mark>	27	46	52	<mark>48</mark>	77	<mark>48</mark>
3	<mark>48</mark>	28	<mark>48</mark>	53	<mark>48</mark>	78	<mark>48</mark>
4	<mark>48</mark>	29	<mark>48</mark>	54	<mark>48</mark>	79	<mark>48</mark>
5	<mark>48</mark>	30	<mark>48</mark>	55	<mark>48</mark>	80	<mark>48</mark>
6	<mark>48</mark>	31	43	56	<mark>48</mark>	81	<mark>48</mark>
7	<mark>48</mark>	32	<mark>48</mark>	57	<mark>48</mark>	82	<mark>48</mark>
8	<mark>48</mark>	33	<mark>48</mark>	58	<mark>48</mark>	83	<mark>48</mark>
9	<mark>48</mark>	34	<mark>48</mark>	59	<mark>48</mark>	84	39
10	48	35	48	60	<mark>48</mark>	85	<mark>48</mark>
11	48	36	<mark>48</mark>	61	<mark>48</mark>	86	<mark>48</mark>
12	48	37	<mark>48</mark>	62	<mark>48</mark>	87	<mark>48</mark>
13	<mark>48</mark>	38	<mark>48</mark>	63	<mark>48</mark>	88	<mark>48</mark>
14	48	39	48	64	<mark>48</mark>	89	<mark>48</mark>
15	45	40	<mark>48</mark>	65	46	90	<mark>48</mark>
16	<mark>48</mark>	41	<mark>48</mark>	66	<mark>48</mark>	91	<mark>48</mark>
17	<mark>48</mark>	42	<mark>48</mark>	67	36	92	<mark>48</mark>
18	<mark>48</mark>	43	<mark>48</mark>	68	<mark>48</mark>	93	<mark>48</mark>
19	<mark>48</mark>	44	<mark>48</mark>	69	46	94	<mark>48</mark>
20	<mark>48</mark>	45	<mark>48</mark>	70	<mark>48</mark>	95	<mark>48</mark>
21	47	46	<mark>48</mark>	71	<mark>48</mark>	96	<mark>48</mark>
22	<mark>48</mark>	47	48	72	<mark>48</mark>	97	<mark>48</mark>
23	<mark>48</mark>	48	48	73	<mark>48</mark>	98	38
24	<mark>48</mark>	49	48	74	<mark>48</mark>	99	<mark>48</mark>
25	<mark>48</mark>	50	48	75	<mark>48</mark>	100	<mark>48</mark>

Note: Numbers in green indicate 100% accuracy.

Observations: Out of 100 object classes, 90 object classes give 100% accuracy

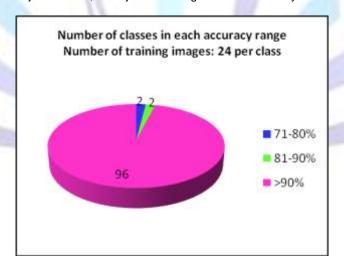


Fig.14 Accuracy performance of 100 classes when 2400 training images are used.

The overall performance for our new approach consisting MSE for matching instead of distance criteria is shown and compared with our earlier technique of classification using independent PCA ( with two distance measures for matching) in table IV.



Table IV: Number of correctly classified images.

Methods	Training images : 10 per class Total Training images: 1000 Testing images: 6200		Training images : 18 per class Total Training images: 1800 Testing images: 5400		Training images : 24 per class Total Training images: 2400 Testing images: 4800	
wethous	Number of Correctly classified images (out of 6200)	Accuracy	Number of Correctly classified images (out of 5400)	Accuracy	Number of Correctly classified images (out of 4800)	Accuracy
Classification using Independent PCA with Euclidean Distance as similarity measure	2700	43.55%	3602	66.70%	3695	76.98%
Classification using Independent PCA with <b>Manhattan Distance</b> as similarity measure	2799	45.15%	3699	67.94%	3640	75.83%
Proposed Method	5881	94.85%	<mark>5279</mark>	97.76%	<mark>4752</mark>	99%

**Observations**: Analyzing the results, we can observe that the proposed method gives very high accuracy even when the number of training images are only 10 i.e only around 14% of data is used for training the accuracy achieved above 90%. When 25% of data is used for training the accuracy goes above 95%. With around 33% of data is used as a training set, the performance is best giving 99% of accuracy.

#### V. CONCLUSIONS

The paper presents an efficient approach of classification that achieves very high accuracy. In our earlier work, after applying independent PCA, nearest neighbor classification is done with Euclidean and Manhattan distance. The overall accuracy was not satisfactory. After doing lot of experimental study, it has been noticed that image reconstructed using Eigen coordinate system is visually similar to input test image even when the distance between feature vectors seems large. So finding the mean squared error (MSE) between input image and the reconstructed image proves better than finding the difference between feature vectors. The proposed technique is applied on COIL-100 database. This database is very large consisting of total 7200 images (72 images per class, 100 classes). With such a large database, to achieve accuracy above 90% is highly complex and difficult task. With training set sizes as 10 images per class, 18 images per class and 24 images per class, the overall accuracy achieved is 95%, 97% and 99% respectively. In view of individual class accuracy, out of 100 classes 65,77, and 90 classes show the 100% accuracy in respective training set sizes. Most of the classes give accuracy more than 90%. In this paper the technique is applied on grayscale image. It can be extended to all three planes of color image.

### VI. REFERENCES

- [1] Shih-Ming Huang and Jar-Ferr Yang, "Unitary Regression Classification With Total Minimum Projection Error for Face Recognition," IEEE signal processing letters, Vol. 20, No. 5, May 2013, pp. 443-446
- [2] N. Jain, S. Sharma and R. Sairam, "Content Base Image Retrieval using Combination of Color, Shape and Texture Features," International Journal of Advanced Computer Research, Vol. 3 No. 1, Issue-8 March 2013, pp. 70-77
- [3] H.B.Kekre, S.D.Thepade, T.K.Sarode and V.Suryawanshi, "Image Retrieval using Texture Features extracted from GLCM, LBG and KPE," International Journal of Computer Theory and Engineering, Vol. 2, No. 5, October- 2010, pp. 695-700
- [4] H.B.Kekre, V.A. Bharadi, A.R. Sawant and B.Markarkandy, "Color &Texture Based Image Retrieval using Fusion of Modified Block Truncation Coding (MBTC) and Kekre Transform Patterns," (IJACSA) Special Issue on Selected Papers from International Conference & Workshop On Emerging Trends In Technology 2012, pp. 12-17, www.ijacsa.thesai.org
- [5] D.Stan and I.K.Sethi, "Mapping low-level image features to semantic concepts," Proc. SPIE 4315, Storage and Retrieval for Media Databases 2001, 172 (January 1, 2001); http://dx.doi.org/10.1117/12.410925
- [6] S.A.Chatzichristofis and Y.S.Boutalis, "FCTH: Fuzzy Color And Texture Histogram A Low Level Feature For Accurate Image Retrieval," in Proceedings of Ninth IEEE International Workshop on Image Analysis for Multimedia Interactive Services, 2008, pp. 191- 196
- [7] A.L.Koerich, A.S.Britto Jr., L.E.S.deOliveira and R.Sabourin, Fusing high- and low-level features for handwritten word recognition, in Proceedings of the International Workshop on Frontiers in Handwriting Recognition, France 2006, pp. 151–156



- [8] D. A. Lisin, M. A. Mattar, M. B. Blaschko, E. G. Learned-Miller and M. C. Benfield, "Combining Local and Global Image Features for Object Class Recognition," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) Workshops - Volume 03, Washington, DC, USA 2005, pp.47-55
- [9] X. Meng, Z. Wang and L. Wu, "Building Global Image Features For Scene Recognition," Journal on Patteren Recognition, Vol. 45, Issue 1, Elsevier Science Inc. New York, NY, USA, Jan-2012,pp. 373-380
- [10] R. S. Chora's, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems," International Journal Of Biology And Biomedical Engineering, Vol. 1, Issue 1,2007, pp.6-16
- [11] H.B.Kekre, D.Mishra, "Discrete Sine Transform Sectorization for Feature Vector Generation in CBIR," Universal Journal of Computer Science and Engineering Technology, ISSN 2219-2158, Vol. 1, Issue. 1, Oct. 2010, pp. 6-15
- [12] H.B.Kekre, T.K.Sarode and J.K.Save, "Image Classification in Transform Domain," International Journal of Computer Science and Information Security (IJCSIS), Vol.10,No.3, Mar 2012, pp. 91-97
- [13] O. Boiman, E. Shechtman, and M. Irani, "In Defense of Nearest-Neighbor Based Image Classification," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2008
- [14] S.B.Park, J. Won Lee, and S. K. Kim, "Content-based image classification using a neural network," Pattern Recognition Letters, Vol. 25, Issue 3, February 2004, Pages 287–300, EISEVIER publication
- [15] M. A. Oskoei, and H. Hu, "Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb," IEEE Transactions On Biomedical Engineering, Vol. 55, No. 8, August 2008, pp.1956-1965
- [16] N. Suguna and K. Thanushkodi, "An Improved k-Nearest Neighbor Classification Using Genetic Algorithm," International Journal of Computer Science Issues (IJCSI), Vol.7, Issue 4, No 2, July 2010, pp. 18-21
- [17] Jacobs, D.W., Weinshall, D., and Gdalyahu, Y., "Classification with nonmetric distances: image retrieval and class representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, Issue 6, 2000, pp. 583-600
- [18] F.Wang, S.Jiang, L.Herranz and Q. Huang, "Improving Image Distance Metric Learning by Embedding Semantic Relations," Advances in Multimedia Information Processing-PCM 2012, Lecture notes in Computer Science, Vol. 7674,2012, pp. 424-434
- [19] H.B.Kekre, T.K.Sarode and J.K.Save, "Classification of Image Database Using Independent Principal Component Analysis," International Journal of Advanced Computer Science and Applications (IJACSA), accepted for publication, Jun 2013
- [20] K.Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," Philosophical Magazine 2 (11),1901,pp. 559–572.
- [21] H.Hotelling, "Analysis of a complex of statistical variables into principal components," Journal of Educational Psychology, Vol 24, No. 6, Sep 1933, 417-441.
- [22] I.T.Jolliffe, "Principal Component Analysis," Springer series in statistic, 2nd edition, New York: Springer; 2002.
- [23] Nene S. A., Nayar S. K. and Murase H.: Columbia Object Image Library (COIL-100), Technical Report CUCS-006-96 (1996)

### Author' biography with Photo



**Dr. H. B. Kekre** has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engineering and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering. College, Mumbai. Now he is Senior Professor at MPSTME, SVKM's NMIMS University. He has guided 17 Ph.Ds., more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National

/International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE Recently twelve students working under his guidance have received best paper awards and six research scholars have beenconferred Ph. D. Degree by NMIMS University. Currently 7 research scholars are pursuing Ph.D. program under his guidance.



Tanuja K. Sarode has Received Bsc.(Mathematics) from Mumbai University in 1996, Bsc.Tech.(Computer Technology) from Mumbai University in 1999, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She has more than 10 years of experience in teaching. Currently working as Associate Professor in Dept. of Computer Engineering at Thadomal Shahani Engineering College, Mumbai. She is life member of IETE, ISTE, member of International Association of Engineers (IAENG) and International Association of Computer Science and Information Technology (IACSIT), Singapore. Her areas of interest are Image Processing, Signal Processing and

Computer Graphics. She has more than 100 papers in National /International Conferences/journal to her credit.



Jagruti K. Save has received B.E. (Computer Engg.) from Mumbai University in 1996, M.E. (Computer Engineering) from Mumbai University in 2004, currently Pursuing Ph.D. from Mukesh Patel School of Technology, Management and Engineering, SVKM's NMIMS University, Vile-Parle (W), Mumbai, INDIA. She is currently working as Associate Professor in Dept. of Computer Engineering at Fr. Conceicao Rodrigues College of Engg., Bandra, Mumbai. Her areas of interest are Image Processing, Neural Networks, Fuzzy systems, Data base management and Computer Vision. She has 8 papers in National /International Conferences/journal to her credit.