A comparative study between Ridgelet PCA and PCA using different distance measure technique for 2D shape recognition and retrieval

Muzameel Ahmed
Dept. of Computer Science and Engineering
Jain University
Bengaluru, INDIA
muzchk@yahoo.com

V.N. Manjunath Aradhya
Dept. of MCA
Sri Jayachamarajarendra College of Engineering
Mysore, INDIA
aradhya1980@yahoo.co.in

Abstract

In this paper, we have proposed a novel method for two-dimensional shape object recognition and retrieval. The proposed method is based on Ridgelet Principal Component Analysis (Ridgelet PCA). In our proposed approach we first use the ridgelet transform to extract line singularity features and point singularity features by applying the radon and wavelet transform respectively and then applying PCA to extract the effective features. For recognition and retrieval we have conducted a study by using seventeen different distance measure techniques. The training and testing process is conducted using leave-one-out strategy. The retrieval process is carried out by considering standard test 'bullseye' score. The proposed method is tested on the collected standard dataset MPEG-7. Experimental results of Ridgelet PCA are compared with the existing PCA method, which show that our approach results are favorable compared to the reference methods, in terms of recognition and retrieval rate.

keywords : 2D Object Recognition, Retrieval, Principal Component Analysis, Ridgelet Transform, Distance Measure Techniques.

1. Introduction

Artificial intelligence is a field in which machines are trained and equipped with all the senses of human being, like smell sense, touch sense, hearing sense, taste sense and vision sense. Machine vision is a field which depicts the effect of human eye in the case of machine. Computer vision is the ability to view and recognize object in a scene. In the recent year computer vision have made enormous progress in this field to achieve high quality visual perception and object recognition. To recognition an object, their are several properties that can be used for the purpose of recognition and classification, like object shape, object color, object texture and object brightness. Of all the properties shapes is the most intrinsic feature used for recognition of objects. Shape representation is done using two major approaches, one the boundary based approach which uses contour information and the second approach needs a holistic representation, requiring general information about the shape [15].

Bribiesca and Wilson [7], presented a approach for 2D shape object dissimilarity. The shape of the different object to be compared are mapped to a representation invariant under translation, rotation and Scaling. Bandera et al [1], proposed a algorithm, were contour are represented by their curvature function, decomposed in the Fourier domain as linear combination of a set of representative object and object are identified by multilevel clustering. Kumar and Rockett [19], proposed a method representing scaling, translation and rotation based on the invariance properties of angle of the triangle which are used to construct signature histogram of local shape. Guerra [3], presented a approach using reconfigurable mesh architecture with horizontal and vertical broadcasting. The object models are described in terms of a convex/concave multiscale boundary decomposition that is represented by a tree structure. Khalil and Bayoumi [11], proposed a method to recognize 2D object under translation, rotation and scale transformation, using the technique based on the continuous wavelet transform and neural networks. Mcneill and Vijaykumar [8], present a corresponding-based technique for efficient shape classification and retrieval. Shapes are represented by a large number of points on the boundary, the points lie at fixed intervals on the boundary or radial angle. Which gives a robust description of shapes. Belongie et al [2], present a method to measure similarity between shapes, and exploit it for recognition. In this framework it solves for correspondences between points on the two shapes, by using the correspondence to estimate an aligning transforms.

Nagabhushan et al [16], propose a technique which is based on 2-Directional 2-Dimensional Fisher's Linear Discriminate analysis for object/face representation and recognition. Sun et al [20], propose a method that employs the eigen-values of covariance matrix, re-sampling, and autocorrelation transformation to extract unique features, and then use minimum euclidean distance method and backpropagation neural networks for classification. Nam et al [17], presented a scheme for similarity-based leaf images retrieval. The method compares the effective measurement of leaf similarity, by considering shape and venation features. Arodz [21], proposed a method using the radon transform properties to drive the invariant transformation involving translation compensation, angle representation and 1-D Fourier transform. Daliri and Torre [15], proposed a algorithm based on dynamic programming to find the object match.

From the literature survey carried out, it is clearly evident that method device so far on shape based object recognition have their own merits and demerits. So a detail observation was made and conclusion was drawn that, new methods can be tried to over come the short coming in the existing methods.

Bandera et al [1], method work with reduced number of input pattern and its required error rate to be reduced in clustering process. Bribiesca et al [7], in this method when shapes are irregular, the extraction of feature or primitive is difficult. Daliri and Torre [15], method is more complex and it is found that it is not the fastest approach in recognition of shapes in the objects. Du et al [6], The MMC classifier method can work with only limited features and with 20 species of plant leaves. Mai et al [13],method limitation is that the curves for registration are to be closed, and it has problems in dealing with occlusion cases. Mcneill et al [8], method ML classifier is not robust to the increased probability of the algorithm, and decreases performance as well. Ruberto et al [5], the method has short coming in effective and efficient discrimination between shapes and in handling noise and occlusion.

The present paper is organized into following section. In section 2 we will discuss the proposed method for recognition and retrieval. Section 3 presents the experimental results. section 4 we discuss the analysis of result and in section 5 we present the discussion and conclusion.

2. Proposed Method

In this section, we explain the idea of proposed method for object recognition and retrieval. The proposed method consists of two stages. The first stage consists of applying Ridgelet Transform, second stage consists of applying PCA.

2.1. Ridgelet Transform

The Ridgelet Transform is the combination of Radon Transform and One Dimensional Wavelet Transform. The Radon transform have received wide range of application in recent years. This transform is able to transform two dimensional images with lines into a domain of possible line parameters, where each line in the image will give a peak positioned at the corresponding line parameters. This have lead to many line detection applications within image processing. [14], One dimensional wavelet transform are efficient at representing zero-dimensional or point singularity. Therefore the ridgelet transform is used to map line singularity into point singularity precisely.The continuous ridgelet transform, defined by [10], produced from the Radon transform, instrumental in its implementaion [9].

Given an integrable bivarite function $f(x_1, x_2)$ its Radon transform (RDN) is defined by

$$RDN_f(\theta, t) = \int_{R^2} (f(x_1, x_2)\delta(x_1 cos\theta + x_2 sin\theta - t))dx_1 dx_2 \tag{1}$$

The radon transform operator maps the spatial domain into the projection domain (θ, t), in which each point corresponds to straight line in the spatial domain. Conversely, each point in the spatial domain becomes a sine curve in the projection domain.

The continuous Ridgelet transform (CRT) is simply the application of a mono-dimensional wavelet $(\psi_{ab}(t)) = a^{-1/2}\psi((t-b)/a)$ to the slice of the radon transform:

CRT(a,b,θ) $= \int_R \psi_{a,b}(t) RDN_f(\theta, t)dt. =$

$= \int_{R^2} (\psi_{a,b,\theta}(x_1, x_2) f(x_1, x_2) dx_1.dx_2.$

where the $\psi_{a,b,\theta}(\overline{x})$ in 2-D are defined from a wavelet-type function $\psi(t)$ as:

$$\psi_{a,b,\theta}(x_1, x_2) = a^{-1/2}\psi((x_1 cos\theta + x_2 sin\theta - b)/a). \tag{2}$$

This show that the ridgelet function is constant along the lines where $x_1 cos\theta + x_2 sin\theta = const.$

Wavelets are very effective in representing objects with isolated point singularities, while ridgelets are very effective in representing objects with singularities along straight lines.

Discrete transform is needed, to apply ridgelet to a digital images. For this reason Do and Vetterli[16] have proposed Finite Ridgelet Transform (FRIT). FRIT is based on the Finite Radon Transform (FRAT), which is defined as summation of image pixels over a certain set of lines. Those lines are defined in a finite geometry in a similar way as the lines for the continuous Radon transform in the Euclidean geometry. Denote $Z_p = 0, 1, ..., p-1$, where p is a prime number and Z_p is finite field with modulo p operations.

The FRAT of real discrete function f on the finite grid Z_p^2 is defined as:

$$FRAT_f(k, l) = \frac{1}{\sqrt{p}} \sum_{(i,j) \in L_{k,l}} f(i, j). \tag{3}$$

Here $L_{k,l}$ denotes the set of points that make up a line on the lattice Z_p^2, i.e.

$$L_{k,l} = \begin{cases} (i,j) : j = (ki + l)(modp), i \in Z_p & \text{if } 0 \le k \le p \\ (l,j) : j \in Z_p & \text{if } k = p \end{cases}$$

Most of the energy information can be found in the low-pass of ridgelet image decomposition. Normally, feature vectors are typically several thousands elements wide [4].

2.2. Principal Component Analysis

To reduce the large dimension feature vector of Ridgelet Transform, we apply PCA [12].The PCA method uses the Karhunen-Loeve transform to produce the most effective subspace for image representation and recognition. In our study we use PCA method for effective feature extraction from the out come of ridgelet transform. The PCA technique is explained as fallows:

Let M be the number of vectors of size $N(L \times L)$, p_i's be the pixel values and $i = 1, ..., M$.

$$x_i = [p_1...p_n]^T \tag{4}$$

The images are mean centered by subtracting the mean image from each image vector. Let m represent the mean image which is N $\times$ 1.

$$m = \frac{1}{M}\sum_{i=1}^{M} x_i \tag{5}$$

Next the covariance matrix C is calculated using:

$$C = \frac{1}{M}\sum_{i=1}^{M} w_i w_i^T \tag{6}$$

Next the eigenvectors e_i and the corresponding eigenvalues λ_i are calculated. From the above M eigenvectors, only k should be chosen corresponding to largest eigenvalues. The eigenvectors of the highest eigenvalues describe more characteristic features of an image. Using the k eigenvectors e_i and $i = 1, \ldots, M$ feature extraction computed by PCA is as follows:

$$F_i = e_k^T(x_i - m) \tag{7}$$

When we have feature vector F_i of each image, identification of a image can be performed. After projecting a new unknown image into the eigenspace we get its feature vector F_{inew}.

Then after calculate the distances between unknown image and each known image using different distance measure classification techniques for classification purpose.

2.3. Distance Measure Classification Techniques

Clustering is the process of grouping together object or instance of similar type, so their should some means to classify the object based on their similarity or dissimilarity. Distance measure and similarity measure techniques are used for the purpose of object classification. The distance between two instance x_i and y_j is denoted as $d(x_i, y_j)$. Distance measure is also called metric distance measure if it satisfy the properties. 1). Triangle inequality $d(x_i, y_k) \leq d(x_i, y_j) + d(x_j, y_k) \; \forall \; x_i, x_j, x_k \in S$. 2). $d(x_i, y_j) = 0 \Rightarrow x_i = x_j \;\; \forall \; x_i, x_j \in S$. A good distance measure should be symmetric and obtain minimum value(usually zero) in case of identical vectors.

We have explored seventeen different distance measure techniques for classification. The distances between feature vector of trained and test images are calculated using the distance measure techniques. The techniques used are Euclidean, Manhattan, Mahalanobis, Minkowski, Modified Manhattan, Modified Squared Euclidean, Squared Euclidean, Weight Angle, Weight Manhattan, Weight Modified Manhattan, Canberra, Modified Normed Distance, Mean Squared Error, Weight Modified SSE, Weight SSE, Angle and Corr. co-efficient [18].

3. Experimental Results and Comparative Study

In this section we will present and compare the performance of our method. The Ridgelet PCA method is tested upon the standard MPEG-7 dataset. The MPEG-7 dataset is a collection of both natural and artificial objects, which has 70 different object, each in a class of 20 samples, for a total of 1400 samples. The dataset is a challenging one as it has samples which are very dissimilar to the other samples in the same class and samples which are very similar to samples in the other class. By large MPEG-7 database has a good collection of samples compared to the other database.

A notebook computer with CORE i5 processor, 2GB RAM memory and Matlab 10.0 platform were used to compute results. The recognition test is carried out using the leave-one-out strategy, where one sample is left out in a class and remaining nineteen are trained. The sample left out of training will be tested. The sample is considered as recognized if it test matches in the same class.

The retrieval process is carried out by considering the standard test 'bullseye' score, where each sample is tested. Retrieval is considered to be correct if the test sample belongs to the same class as being tested. The number of correct matches in the top 40 result are counted, including the self match. Retrieval rate for each method is reported as percentage of the maximum possible number of correct retrievals. That is 28,000 (1400 shapes * 20 correct retrievals) [15]. The experimental results were extracted using seventeen different distance measure techniques. And also the experimental results were extracted by varying the projection vector value between 10 to 50, by increment value in terms of 10 for recognition and retrieval.

Table 1 and 2 show the result of PCA and Ridgelet PCA, recognition rate respectively. The recognition accuracy of PCA is 87.50 (using mod.manhattan) shown in table 1, and Ridgelet PCA is 89.50 (using weight sse) shown in table 2. Retrieval rate of PCA and Ridgelet PCA, is shown in table

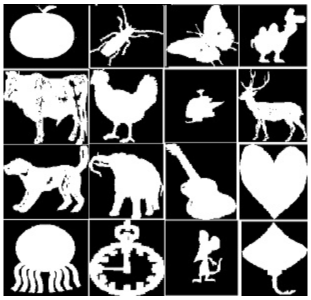

Figure 1. MPEG-7 Dataset Images

3 and 4 respectively. Retrieval rate for PCA is 64.38 (using weight.angle) shown in table 3, Ridgelet PCA is 65.97 (using weight.angle) shown in table 4.

4. Analysis of Results

In this section we analysis the results of PCA and Ridgelet PCA for recognition and retrieval accuracy rate. Ridgelet PCA recognition accuracy result with Weighted SSE (weighted sum of squared errors) distance measure technique outperforms all the other distance measure techniques mentioned. And also PCA recognition results using seventeen distance measure techniques mentioned. In Weighted SSE distance measure the quality of clustering, which is also know as scatter. It calculates the error of each data point that is, its euclidean distance to the closest centroid and then computes the total sum of the squared error. So therefore in recognition of binary image object their is a lesser squared error rate and its has closest centroid in the cluster. Therefore Weighted SSE gives higher recognition rate compared to other distance measure technique.

For the retrieval rate, Ridgelet PCA with Weighted Angle distance measure technique outperform all other distance measure techniques mentioned,and also PCA retrieval result using different distance measure techniques. Weighted Angle distance measure technique, deals with the issue of length normalization. Because long data would be more similar to each other by virtue of length. However data can be implicitly normalized by taking angle instead of similarity between vectors. For the data d_i and query q can be computed as vector product. For binary vector the inner product is a number of matched query terms in the data. For weighted term vector it is a sum of product of the weight of matched term. Inner product favors long data with a large numbers of unique terms, and measures the matched terms. Distance between vectors d_1 and d_2 is captured by the cosine of the angle x between them. So the cosine measure is used for the purpose of retrieval.

5. Conclusion

Recognition and retrieval plays a important role in the field of computer vision applications. Developing an efficient and accurate system for recognition and retrieval is a real challenge as extracting shape based features in comparison with, complex and extraordinary human vision perception is not a easy job. In this paper, we have made a study and comparison of shape based feature extraction and representation methods. The Ridgelet PCA, method is tested on MPEG-7 database, which has a collection of 70 different type of object, categorized as 70 class with 20 samples in each of them (70 x 20 = 1400). The recognition process is conducted using leave-one-out strategy. And the retrieval rate are computed by counting the top 40 matches in the test query. Seventeen different distance measure techniques for categorization are used. And experimental results are obtain by varying the projection vector value between 10 to 50, in increments of 10. The comparative study of the methods result is done, and it is found that the Ridgelet PCA has better and encouraging result, to that of PCA in terms of recognition and retrieval. In the Ridgelet PCA, Ridgelet transform map the line singularity into point singularity in the radon transform. Then the wavelet transform is used to effectively handle the point singularity in the radon domain. Which allows features extraction in a highly efficient manner.

References

[1] F. Arrebola Bandera. A, C. Urdiales and F. Sandoval. 2d object recognition based on curvature functions obtained from local histograms of the contour chain code. *Pattern Recognition Letters*, 20:49–55, 1999.

[2] Malik. J Belongie. S and Puzicha. J. Shape matching and object recognition using shapes contexts. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24(24):509–522, 2002.

[3] Guerra. C. 2d object recognition on a reconfigurable mesh. *Pattern Recognition*, 31(1):83–88, 1998.

[4] Naveena. C and Manjunath Aradhya. V. N. An effect of ridgelet transform on various distance measure technique in bi-lingual handwritten character recognition. *In proceedings of WICT*, pages 618–621, 2012.

[5] Ruberto. C . D. and Luigi Cinque. Decomposition of two-dimensional shapes for efficient retrieval. *Image and Vision Computing*, 27:1097–1107, 2009.

[6] Wang. X Du. J and Zhang. G. Leaf shape based plant species recognition. *Applied Mathematics and Computing*, 185:883–893, 2007.

[7] Bribiesca. E and R. G. Wilson. A measure of 2d shape-of-object dissimilarity. *Application Mathematics Letters*, 10(6):107–115, 1997.

Classifier Methods	**10**	**20**	**30**	**40**	**50**
Manhattan	78.57	84.36	85.50	86.14	86.50
Euclidean	79.36	83.71	85.71	85.71	85.93
Mahalanobis	09.00	20.43	31.07	39.57	47.57
M.D between normed distance	53.07	78.86	83.86	85.57	86.79
Minkowski	76.93	81.00	83.00	83.29	83.64
Modified Manhattan	79.00	85.43	86.50	86.93	**87.50**
Modified Squared Euclidean	79.29	84.43	86.57	86.29	86.71
Mean Squared Error	79.36	83.71	85.71	85.71	85.93
Squared Euclidean	79.36	83.71	85.71	85.71	85.93
Weighted Angle	53.07	78.86	83.86	85.57	86.79
Weighted Manhattan	78.36	84.00	84.64	85.21	85.79
Weighted Modified Manhattan	77.86	84.86	85.86	86.64	87.21
Weighted Modified SSE	79.14	84.64	85.86	86.14	87.00
Weighted SSE	79.07	84.14	85.71	85.79	86.07
Canberra	79.00	85.00	85.86	86.64	87.29
Angle	78.57	83.86	86.21	85.86	86.50
Correlation co-efficient	77.14	84.14	85.93	85.50	86.50

Table 1. Recognition Rates using PCA with different Classifier Methods and by varying Projection values between 10 to 50

Classifier Methods	**10**	**20**	**30**	**40**	**50**
Manhattan	82.86	87.79	88.79	88.71	88.86
Euclidean	83.14	86.36	87.79	87.79	87.86
Mahalanobis	01.43	01.43	01.64	01.50	01.50
M.D between normed distance	44.29	71.43	78.64	81.79	83.71
Minkowski	83.14	86.36	87.79	87.79	87.86
Modified Manhattan	83.64	87.71	88.07	88.14	88.43
Modified Squared Euclidean	83.00	86.29	87.07	86.79	87.14
Mean Squared Error	83.14	86.36	87.79	87.79	87.86
Squared Euclidean	83.14	86.36	87.79	87.79	87.86
Weighted Angle	44.29	71.43	78.64	81.79	83.71
Weighted Manhattan	84.57	88.79	89.07	88.79	88.29
Weighted Modified Manhattan	83.00	86.93	87.79	87.00	87.07
Weighted Modified SSE	83.64	86.50	86.64	86.57	86.14
Weighted SSE	85.14	88.36	89.36	89.36	**89.50**
Canberra	83.29	88.07	88.71	88.86	88.79
Angle	83.79	86.93	87.57	87.57	87.57
Correlation co-efficient	82.71	86.71	87.43	87.71	87.50

Table 2. Recognition Rates using Ridgelet PCA with different Classifier Methods and by varying Projection values between 10 to 50

[8] McNeill. G and Vijayakumar. S. 2d shape classification and retrieval. *Proceeding of International Joint Conference on Artificial Intelligence*, pages 1483–1488, 2005.

[9] Moschetti. F Grania. L and Vandergheynst. P. Ridgelet transform applied to motion comprensated images. *In proceedings of ICASSP2003*, pages 381–384, 2003.

[10] Candes. E. J and Donoho. D. L. Ridgelet: a key to higher dimensional intermittency. *Philosophical Transaction of the royal society*, pages 2495–2509, 1999.

[11] Khalil. I. M and Mohamed. M. Bayoumi. Invariant 2d object recognition using the wavelet modulus maxima. *Patten Recognition Letters*, 21:863–872, 2000.

[12] Turk. M and Pentland. A. Eigenfaces for recognition. *Journal of Cognition and Neuroscience*, 3:71–86, 1991.

[13] C .Q .Chang Mai. F and Y. S. Hung. A subspace approach for matching 2d shapes under affine distortions. *Patten Recognition*, 44:210–221, 2011.

[14] Miciak Miroslaw. Radon transformation and principal component analysis method applied in postal address recognition task. *International Journal of Computer Science and Application*, 7(3):33–44, 2010.

[15] Daliri. R. Mohammad and Vincent Torre. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 41:1782–1798, 2008.

[16] D. S. Guru Nagabhushan. P and B. H. Sheker. (2d)2 fld: An efficient approach for appearance based object recognition. *Neurocomputing*, 69:934–940, 2006.

[17] Hwang. E Nam. Y and Kim. D. A similarity-based leaf image retrieval scheme : Joining shape and venation features. *Computer Vision and Image Understanding*, 110:245–259, 2008.

[18] Vytautas Perlibakar. Distance measure for pca-based face recognition. *Patten Recognition Letters*, 25:711–724, 2004.

[19] Kumar Rajeev and Rockett Peter. Triplet geometric representation : a novel scale, translation and rotation invariant feature

Classifier Methods	**10**	**20**	**30**	**40**	**50**
Manhattan	55.93	59.96	61.18	61.53	61.13
Euclidean	55.96	59.35	60.67	61.38	61.38
Mahalanobis	17.41	28.35	35.94	39.83	42.37
M.D between normed distance	52.56	60.93	63.28	64.31	64.38
Minkowski	53.08	54.97	55.67	55.83	55.80
Modified Manhattan	56.64	61.23	62.93	63.44	63.54
Modified Squared Euclidean	56.43	60.18	61.75	62.67	62.81
Mean Squared Error	55.96	59.35	60.67	61.38	61.38
Squared Euclidean	55.96	59.35	60.67	61.38	61.38
Weighted Angle	52.56	60.93	63.28	64.31	**64.38**
Weighted Manhattan	55.93	59.68	59.88	59.00	57.57
Weighted Modified Manhattan	56.95	61.73	63.14	62.81	61.97
Weighted Modified SSE	56.72	61.08	62.61	63.16	63.11
Weighted SSE	56.11	59.78	60.81	61.28	60.90
Canberra	56.57	61.24	63.04	63.63	63.78
Angle	56.02	59.82	61.23	62.01	62.03
Correlation co-efficient	54.59	59.52	61.10	61.84	61.89

Table 3. Retrieval Rates using PCA with different Classifier Methods and by varying Projection values between 10 to 50

Classifier Methods	**10**	**20**	**30**	**40**	**50**
Manhattan	60.98	63.76	64.89	65.21	65.04
Euclidean	58.36	60.46	61.07	61.19	61.23
Mahalanobis	06.47	08.66	09.91	10.73	11.26
M.D between normed distance	56.56	63.51	65.26	65.66	65.97
Minkowski	58.36	60.46	61.07	61.19	61.23
Modified Manhattan	60.98	63.66	64.77	64.97	65.03
Modified Squared Euclidean	57.84	59.53	60.09	60.12	60.10
Mean Squared Error	58.36	60.46	61.07	61.19	61.23
Squared Euclidean	58.36	60.46	61.07	61.19	61.23
Weighted Angle	56.56	63.51	65.26	65.66	**65.97**
Weighted Manhattan	62.88	64.83	65.75	64.90	64.17
Weighted Modified Manhattan	60.60	62.29	62.28	61.20	60.29
Weighted Modified SSE	59.34	60.13	60.18	59.23	58.48
Weighted SSE	62.22	64.46	65.36	65.51	65.44
Canberra	61.43	64.30	65.53	65.78	65.72
Angle	61.40	62.84	63.13	63.04	62.85
Correlation co-efficient	59.75	62.74	63.16	63.06	62.94

Table 4. Retrieval Rates using Ridgelet PCA with different Classifier Methods and by varying Projection values between 10 to 50

representation based on geometric constraints for recognition of 2d object features. *Image and Vision Computing*, 15:235–249, 1997.

[20] Chi-Shuan Liu Sun Te-Hsiu and Fang-Chin Tien. Invariant 2d object recognition using eigenvalues of covariance matrices, re-sampling and autocorrelation. 35:1966–1977, 2008.

[21] Arodz Tomasz. Invariant object recognition using radon-based transform. *Computing and Information*, 24:183–199, 2005.