INTERNATIONAL JOURNAL FOR TRENDS IN ENGINEERING & TECHNOLOGY
VOLUME 20 ISSUE 1 – FEBRUARY 2017 - ISSN: 2349 – 9303

Image Classification and Recognition Using PCA

G.K.Abishek Prasanna[1]
[1]Dr.Mahalingam College of Engineering and Technology,
Department of Information Technology,
abishekprasanna6468@gmail.com

E.Arunasaithambi[3]
[3]Dr.Mahalingam College of Engineering and Technology,
Department of Information Technology,
arunasaithambiea@gmail.com

R.Karthick[2]
[2]Dr.Mahalingam College of Engineering and Technology,
Department of Information Technology,
karthickkd126@gmail.com

J.Dhyaneswaran[4]
[4]Dr.Mahalingam College of Engineering and Technology,
Department of Information Technology,
dhyan123@drmcet.ac.in

Abstract— This paper demonstrates the classification of images using principal component analysis. PCA is an image classification algorithm which presents accuracy of classification. First, the image should be pre-processed that is the images are in different sizes, some may be noisy or larger in size, so all the images are resized to a specific size, then it is easy to classify the images. Second, the images are colorful, so it is converted into grey scale image color images are 3-dimensional it is hard to process an 3d image so it is converted into grey scale which is 2d in nature. PCA is an image classification technique which extracts the major features of an image. First, the system is trained using some images, in order to produce an image feature. After it is trained, when a new image is passed it compares with the image feature and displays the output image based on the input. The development of image classification has been improved due to growth in volume of images, as well as the widespread application in multiple fields. Final output is to retrieve the related images based on the input.

Index Terms— Classification, Eigen value, Eigen Vector, Image, Principal Component.

1 INTRODUCTION

Recently the image processing has grown rapidly. So, image plays a vital role in image classification. My work in this work is based on the content of the image which is a major part in the image classification. The growth in the technologies and web, there has been a large complexity to handle the data, which is stored, analyzed and transmitted. In order to make use of this vast data, many effective techniques have been used to retrieve information depending on its content need to be developed. Image classification is challenging & emerging field of image processing. This application importance in image analysis and design field.

2 LITERATURE SURVEY

In the existing system, they have used particle swarm optimization. This algorithm applies in noisy feature of images. This algorithm is swarm intelligence based. The observation of swarming habits of birds or fish. This algorithm maintains multiple potential solutions at one time. During each of the iterations, the solution is evaluated by an objective function. In this they used satellite images to process it.

3 PROPOSED SYSTEM

In proposed system, we have used Principal Component Analysis technique which is a face recognition algorithm which provides more accuracy than the existing system. In this system, the images are classified by the principal component.

4 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a statistical procedure which uses orthogonal transformation to convert a set of observed values of possibly correlated variables in to a set of linearly uncorrelated variables called the principal component. The principal component number is less than or equal to the number of original variables. This type of transformation is defined in a way such that the first principal component has the highest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the largest variance possible under the constraint that it is Orthogonal to the preceding component. The resulting vectors are an interrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.

5 MODULES

5.1 DATASET

The dataset consists of 100 images of different images like cat, flower, lion and car. Each of the image are equally distributed in the dataset. These images are in different sizes. The images are stored in a single dataset. After this, the image has to be pre-processed so that only we can classify the images.

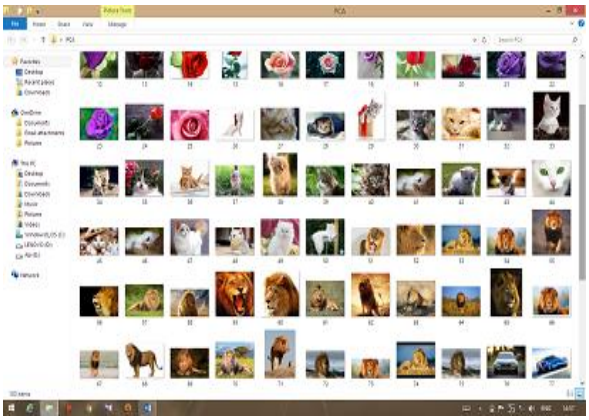

Fig. 1 . Represents the image database with different images like flower, cat, lion and car.

5.2 IMAGE PRE-PROCESSING

In this each image are in different sizes, so the images are converted to a unique size 250x250. These images are in standard size. Then the images are colorful, so it is converted in to gray scale which is more efficient for classification.

5.3 MATRIX FORMATION

We are using MATLAB to process the image, in MATLAB the images are stored in a array matrix (m x n, m-rows-columns). Then the images are converted into mn x 1 format, so that it is easy to process. For example, if there are 100 images, these images are stored as mn x 100.

5.4 COVARIANCE MATRIX

Covariance is such a measure. Covariance is measured *between only* 2 dimensions. If we calculate the covariance for 1 dimension and *itself*, you get the variance. So, if you had a 3-dimensional data set (x,y,z), then you could measure the covariance between x-y, y-z, x-z respectively. Measuring the covariance between same variables it would give you the variance of the x, y and z dimensions respectively.
here is, the formula for covariance:

COV (X, Y) =E([X-E(X)] [Y-E(Y)])

where,
X is the original image matrix,
Y is the transposed matrix of X,
E(X) is the mean subtracted matrix of X,
E(Y) is the mean subtracted matrix of Y.

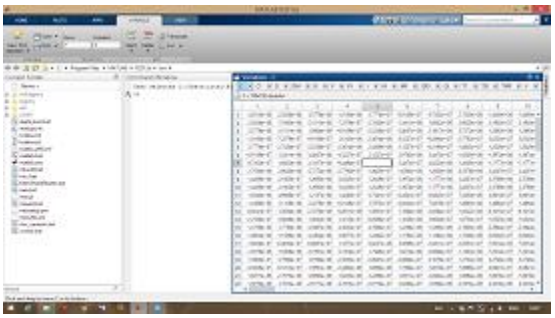

Fig. 2 . Represents the covariance value of the image database in the matrix format.

5.5 CALCULATION OF EIGEN VALUES

You should first know that eigen values are applicable only to a square matrix, so that the square matrix is formed with m x n matrix. Another property of Eigen vectors is that even if I scale the vector by some amount of value before I multiply it, I still get the same multiple of it as a result. This is because, if you measure a vector by an amount, you are doing it longer, but not changing its direction. Lastly, all the eigen vectors of a matrix are p*erpendicular*, i.e. At right angles to each other. Another name for word perpendicular is orthogonal.

Another important thing to know is that when the mathematicians find eigen vectors, they would also like to find the eigen vectors whose length is equal to one. So, in order to keep eigen vectors standard, we find an eigen vector and scale it to make it have a length of 1, so that all eigen vectors have the same length. The eigen value for the covariance matrix is found out using the formula, with the eigen value eigen vector is calculated, the maximum eigen value is the Principal Component for the image, so the image is classified with the help of the principal component.

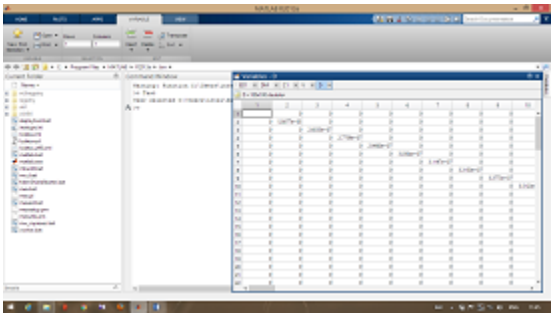

Fig. 3. The Eigen values of the covariance matrix which is diagonally represented.

5.6 CHOOSING THE PRINCIPAL COMPONENT

After, the calculation of the eigen vectors there are some values in it, which is a square matrix. Now, order the eigen vector from higher to lower order. Among these values the value with highest significance is the principal component of the image. You can also neglect the lesser significance value, but you may lose some information in it, but not much. So, first p eigen vectors are only taken out from the

feature, so that you have only p component.

Feature vector = (eig1 eig2 eig3 ……. eign)

Feature vector is nothing but the matrix of vectors. This is constructed from the list eigen vectors, that we have chosen.

5.7 DERIVING A NEW DATASET

This is the last step in PCA, after choosing the component formed from a feature vector, we take a transpose of a feature matrix and multiply to the original vector.

Final Data = Row Feature Vector x Row Data Adjust

Where Row Feature Vector is the column matrix of the eigen vector transposed, Row Data Adjust is the original data.

5.8 GETTING THE DATA BACK

After all, process have been done, the last step is retrieval of images. This is done by finding the principal component of the new image and comparing the component with the final data that we have produced. If there is any match between the values retrieve the images or else notify there are no images present in the database.

6 CONCLUSION

Image Classification in terms of a image processing is a vast field. This image classification will be useful in many fields like detection of a person, security, object recognition, and in terms of a traffic signals. This paper provides the most accuracy of the other methods in terms of classification.

7 REFERENCES

[1] P. Kamavisdar, S. Saluja, and S. Agrawal, “A Survey on Image Classification Approaches and Techniques,” vol. 2, no. 1, pp. 1005–1009, 2013.
[2] D. L. & Q. Weng, “A survey of image classification methods and techniques for improving classification performance,” Int. J. Remote Sens., 2007.
[3] Mokhairi Makhtar, Engku Fadzli, and Shazwani Kamarudin, “The Contribution of Feature Selection and Morphological Operation For On-Line Business System ’ s Image Classification,” World Appl. Sci. J., 2014.
[4] J. F. Nunes, P. M. Moreira, and J. M. R. S. Tavares, “Shape Based Image Retrieval and Classification,” 2010, pp. 433–438.
[5] Fei-Fei, Fergus, and Perona, “Caltech 101 dataset,” 2004.
[6] A. C. Berg, T. L. Berg, J. Malik, and U. C. Berkeley, “Shape Matching and ObjectRecognition using Low Distortion Correspondences,” 2005.
[7] L. Fei-Fei, R. Fergus, and P. Perona, “Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object
Categories,” Conf. Comput. Vis. Pattern Recognit. Work., pp. 178–178, 2004.
[8] S. Banerji, A. Sinha, and C. Liu, “New image descriptors based on color, texture, shape, and wavelets for object and scene image classification,” Neurocomputing, vol. 117, pp. 173–185, Oct. 2013.
[9] P. A. B. Miss Hetal J. Vala, “A Review on Otsu Image Segmentation Algorithm,” Int. J. Adv. Res. Comput. Eng. Technol., vol. Volume 2, no. Issue 2, 2013.