International Journal of Scientific & Engineering Research, Volume 7, Issue 2, February-2016
ISSN 2229-5518

Face Recognition using Principal Component Analysis

Prof. Pritika V. Mamankar, Prof. Sonika A. Chorey, Prof. Rachana S. Sawade

Abstract— Image processing is process of pictures victimization mathematical operations by victimization any kind of signal process that the input is a image; the output of image process could also be either a picture. Image process sometimes refers to digital image process, however optical and analog image process are also potential. In camera work, pictures are manually made up of physical models of objects, environments, and lighting, rather than being non inheritable from natural scenes, as in most animated movies. Laptop vision, on the opposite hand, is commonly thought of high-level image process out of that a machine/computer/software intends to decipher the physical contents of a image or a sequence of pictures.

Index Terms— Image Face recognition, Principal Component Analysis

1 Introduction

A FACE recognition system is a computer vision and it automatically searches a human face from database images. The face recognition problem is difficult task as it needs to account for all possible appearance variation caused by change in illumination, facial features, occlusions, etc. This gives a

Neural and PCA based algorithm for efficient and robust face recognition. Holistic approach, feature-based way and hybrid way are some of the ways for face recognition. Here, a holistic approach is used in which the whole face part is taken into concern as input data. This is based on principal component-analysis technique, which is used to solve a dataset into lower dimension while retaining the characteristics of dataset. Pre-processing, PCA and Back Propagation Neural Algorithm are the major implementations of this paper. Pre-processing is done for two purposes:

(i) To decreases noise proportion and possible convolute effects of interfering system,
(ii) To convert the image into a variable space where classification may prove
easier by exploitation of certain features.

PCA is a common statistical ways for identifying the patterns in high dimensional data's feature extraction, also called Dimensionality Reduction, is done by PCA for a three main effects like
i)To decrease the dimension of the data to more tractable limits
ii) To identify salient class-specific features of the data
iii) To remove redundancy.

Here face recognition is processed by both PCA and Back propagation Neural Networks. BPNN mathematically models the behavior of the feature vectors by particular descriptions and then exploits the statistical behavior of the feature vectors to define decision regions related to different classes. Any new pattern can be classified depending on which decision region it would be falling in.

The Algorithm for Face recognition using neural classifier is as follows:
a) Pre-processing stage –Images are made up of zero-mean and unit-variance.
b) Dimensionality Reduction stage: PCA - Input data is decreases to a last dimension to facilitate classification.
c) Classification stage - The decreased vectors from PCA are applied to train BPNN classifier to
identify the recognized image.

PCA is employed during this work as a preliminary statistical procedure technique. Statistical procedure thinks about with information sets that have quite one response variable for every empiric unit. The information sets may be summarized by data matrices X with n rows and p columns, the rows representing the observations, and therefore the columns the variables. The most division in variable ways is between people who assume a given structure, as an example, dividing the cases into teams, and people that ask for to find the structure from the proof of the information matrix alone, additionally known as data processing. In pattern recognition language the excellencies between supervised and unattended ways. The stress of this paper is Khaled Labib and V. Rao Vemuri totally on mistreatment PCA as AN unattended technique with belief of no a priori data of the structure of the input file.

In PCA, the extractions of PC can be developing using either original multivariate data set or by using covariance matrix if the initial data set is unavailable. In deriving PC, the correlation matrix may be used, instead of the covariance matrix, when different variables in the data set are calculated using different measurement or if different variables have differ-

- Prof. Pritika V. Mamankar is working in Department of IT, PRMIT&R,Badnera, Amravati, India. Email: pritikamamankar11@gmail.com
- Prof. Sonika A. Chorey is working in Department of IT, PRMIT&R,Badnera, Amravati, India. Email: sonikachorey@gmail.com
- Prof. Rachana S. Sawade is working in Department of IT, PRMIT&R,Badnera, Amravati, India. Email: rachanas56@gmail.com

ent variances. Using correlation matrix is similar to standardizing the variables to zero mean and unit standard deviation.

The PCA model can be represented by:
mx1 mxd dx1 u =W x
where u, m-dimensional vector, is a projection of x - the original d-dimensional data vector (m << d).

It can be shown that the m projection vectors that maximize the variance of u, called the principal axes, are shown by the eigen vectors e1, e2, …, em of the data set's covariance matrix S, corresponding to the m largest non-zero eigen values λ1, λ2, … λm.

The data set's covariance matrix S can be find from: ( )( )T n i x x n S − μ − μ − = ∑1 =1 1 where μ is the mean vector of x. The eigen vectors ei can be found by solving the set of equations: ( ) S − λiI ei = 0 i =1,2,...,d where λi are the eigen values of S. After calculating the eigen vectors, they are sorted by the magnitude of the corresponding eigen values. Then the m vector with the largest eigen values are chosen.

1.1 Principal Component Analysis

Principal element Analysis(PCA) could be a method that uses associate orthogonal transformation to convert a collection of observations of probably correlative variables into a collection of values of linearly unrelated variables known as principal parts. The amount of principal parts is a smaller amount than or adequate the amount of original variables. This transformation find is in such a way that the primary principal element has the biggest attainable variance (that is, accounts for the maximum amount of the variability within the knowledge as possible), and every succeeding element successively has the very best variance attainable below the constraint that it's orthogonal to the preceding parts. The ensuing vectors area unit associate unrelated orthogonal basis set. The principal parts area unit orthogonal as a output of they are the Eigen vectors of the variance matrix, that is bilateral. PCA is sensitive to the relative scaling of the initial variables. [1]

1.2 PCA Algorithm

The algorithm is used for principal component analysis is as follows:
(i) Identify an initial set of M face images (the training set) & calculate the eigen-faces
from the training set, keeping M' eigen faces that correspond to the highest
eigen value.
(ii) Calculate the corresponding distribution in M'-dimensional weight space for each
known individual, and calculate a set of weights based on the input image.
(iii) Identify the weight pattern as either a known person or as unknown, according to its distance to the closest weight vector of a known person.

PCA is that the simplest of truth eigenvector-based variable analyses. Often, its operation may be thought of as revealing the inner structure of information. |The information during a means that best explains the variance within the data. If a variable information set is visualized as a group of coordinates during a high-dimensional data house (1 axis per variable), PCA will offer the user with a lower-dimensional image, a projection or "shadow" of this object once viewed from its (in some sense; see below) most informative viewpoint. This is often done by victimization solely the primary few principal elements in order that the spatial property of the reworked information is reduced.

PCA is closely associated with co relational analysis. Corelational analysis usually incorporates a lot of domain specific assumptions regarding the underlying structure and solves eigenvectors of a rather completely different matrix.

PCA is additionally associated with canonical correlation analysis (CCA). CCA defines reference systems that optimally describes the cross-covariance between two datasets whereas PCA defines a brand new orthogonal coordinate system that optimally describes variance in an exceedingly single dataset.

A. PCA And Qualitative Variables

In PCA, it's normal that we might wish to introduce qualitative variables as supplementary elements. For example, many quantitative variables square measure measured on plants. For these plants, some qualitative variables unit offered as, for example, the species thereto the plant belongs. These information were subjected to PCA for quantitative variables. Once analyzing the results, it's natural to connect the principal elements to the qualitative variable species. For this, the following results unit created. Identification, on the factorial planes, of the various species e.g. exploitation totally different colors. Representation, on the factorial planes, of the centers of gravity of plants belonging to the same species.

For each center of gravity and each axis, p-value to judge the significance of the difference between the center of gravity and origin [2] [3].

B. Relation Between PCA And K- Means Clustering

It was declared in this the relaxed answer of k-means bunch, such by the cluster indicators, is given by the PCA (principal component analysis) principal parts, and so the PCA set spanned by the principal directions is also a twin of the cluster centre of mass set. However, that PCA may be a helpful relaxation of k-means bunch wasn't a innovative result, and it is easy to uncover counter examples to the statement that the cluster centre of mass set is spanned by the principal directions [4] [5].

C. Relation Between PCA And Factor Analysis

Principal component analysis creates variables that unit of measurement linear combos of the primary variables. The new variables have the property that the variables unit of measurement all orthogonal. The principal elements are accustomed notice clusters in associate passing set of data. PCA might be a variance-focused approach seeking to breed the general variable variance, inside that elements replicates every common and distinctive variance of the variable. PCA is usually hottest for functions of data reduction but not once the goal is to search out the latent construct/factors.[6]

Corelational analysis is like principal element analysis, throughout this statistical procedure besides involves linear combos of variables. Totally completely totally different from PCA, statistical procedure could also be a correlation-focused approach seeking to breed the inter-correlations among variables, within that the factors "represent the common variance of variables, excluding distinctive variance. In terms of the matrix, this corresponds with that think about explaining the off-diagonal terms (i.e. shared co-variance), whereas PCA focuses on explaining the terms that sit on the diagonal. However, as a facet result, once making an effort to breed the on-diagonal terms, PCA besides tends to suit comparatively well the off-diagonal correlations. Results given by PCA and statistical procedure square measure terribly similar in most things, however generally |this will be} typically not invariably the case, and there square measure some issues wherever the results square measure considerably totally completely totally different. Statistical procedure is sometimes used once the analysis purpose is detection organization (i.e., latent constructs or factors) or causative modeling [7].

REFERENCES

[1] M.A.Turk and A.P.Petland, “Eigen faces for Recognition,” *Journal of Cognitive Neuroscience.* vol. 3, pp.71-86, 1991.

[2] H. Zha, C. Ding, M. Gu, X. He and H.D. Simon, "Spectral Relaxation for K-means Clustering" (PDF). *Neural Information Processing Systems,* vol.14 (NIPS 2001)(Vancouver, Canada): 1057–1064., Dec 2001

[3] Chris Ding and Xiaofeng He. "K-means Clustering via Principal Component Analysis" (PDF). *Proc. of Int'l Conf. Machine Learning* (ICML): 225–232, July 2004.

[4] Drineas, P.; A. Frieze; R. Kannan; S. Vempala; V. Vinay,. "Clustering large graphs via the singular value decomposition" (PDF). *Machine learning* 56: 9–33.doi:10.1023/b:mach.0000033113.59016.96. Retrieved 2012-08-02., 2004.

[5] Cohen, M.; S. Elder; C. Musco; C. Musco; M. Persu,. "Dimensionality reduction for k-means clustering and low rank approximation (Appendix B)". ArXiv. Retrieved 2014-11-29.

[6] Timothy A. Brown. Confirmatory Factor Analysis for Applied Research Methodology in the social sciences. Guilford Press, 2006

[7] I.T. Jolliffe. Principal Component Analysis, Second Edition. Chapter 7, 2002