IJCSE International Journal of Computer Sciences and Engineering **Open Access**

Research Paper **Vol-6, Special Issue-4, May 2018** **E-ISSN: 2347-2693**

Role of Feature Extraction Techniques : PCA and LDA for Appearance Based Gait Recognition

K. Annbuselvi[1*], N. Santhi[2], S. Sivakumar[3]

[1*] Department of Computer Science, V.V.V. College for Women, Virudhunagar, India
[2] Department of Computer Science, V.V.V. College for Women, Virudhunagar, India
[3] Department of Computer Science, CPA College, Bodinayakanur, India

Available online at: www.ijcseonline.org

***Abstract* —** Feature extraction is one of the most important step in image pattern recognition. Some sources of difficulty are the presence of irrelevant information and the relativity of a feature set to a particular application. Feature extraction and description are essential components of various computer vision applications. The concept of feature extraction and description refers to the process of identifying points in an image (interested points) that can be used to describe the image's contents. The One major goal of feature extraction is to increase the accuracy of learned models by compactly extracting prominent features from the input data, while also possibly removing noise and redundancy from the input. Additional objectives include low-dimensional representations for data imagining and compression for the purpose of reducing data storage requirements as well as increasing training and implication speed. The aim of this paper is to report an descriptive study of most popular feature extraction methods PCA and LDA which are generally used in pattern recognition and the role of PCA and LDA in gait feature extraction.

***Keywords* —** Feature extraction, PCA, LDA, Gait Feature Extraction

I. Introduction

Feature extraction is playing a vital role in image pattern recognition applications. The aim of feature extraction is to find the most pertinent data from the original data and represent that data in low dimensional space [**1**]. When the input data to the algorithm is very large and redundant then the input data to be transformed into reduced set of features. The process of transforming the input data into a reduced set of features is called feature extraction. Feature extraction has been given as "extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the between class pattern variability". Thus, selection of a suitable feature extraction technique according to the input to be applied need to be done with utmost care. Taking into consideration all these factors, it becomes essential to look at the various available techniques for feature extraction in a given domain, covering vast possibilities of cases. The efficiency of feature extraction method enhances the further processing of an image to a great extent. These features can be used in image matching, pattern recognition and retrieval.

II. FEATUE EXTRACTION TECHNIQUES

A. PRINCIPAL COMPONENT ANALYSIS (PCA)

The main aim of a principal component analysis [**2**] [**3**] [**4**] [**5**] is finding the directions of maximum variance in high-dimensional (n) dataset and project it onto a smaller dimensional subspace while retaining most of the information. In other words, it is a procedure that uses an orthogonal transformation to convert a set of possibly M correlated variables into a set of K uncorrelated variables are called as principal components / eigen vectors (K < M). The flowchart of Principal Component Analysis is given in figure 1.

Principal Component Analysis Algorithm

Inputs : a set of images visualized as a set of coordinates in a high-dimensional(n) dataset.

Outputs : produces a lower dimensional picture a shadow (eigen images) of given m images when viewed from its most informative points.

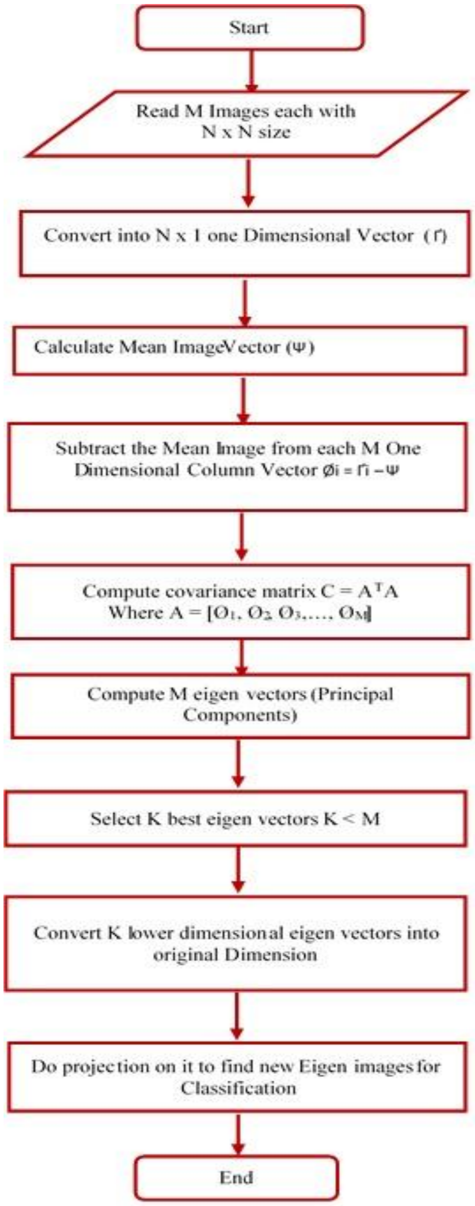

Figure 1. Flowchart of Principal Component Analysis

Steps :

1. Convert each m input images (I1,I2,…,Im) of size N X N into N X 1 one dimensional column vectors (Γ1, Γ2…, Γm)
2. Normalize N X 1 column vector. It means that the all the common features from each M images are removed so that each image has only unique features.
   i. Calculate an average/mean image vector (Ψ)
   ii. Subtract the average/mean image vector (Ψ) from each M one dimensional column vectors Γ
   (i.e.) Normalized Image Vectors Øi = Γi – Ψ
3. Calculate K significant Eigen Vectors (principal components/axes) and eigen values (variance) from a covariance with reduced dimensionality

   Eigen vectors => Determines the direction of the new feature space / axes

   Eigen Values => Determines the magnitude / variance of the data along the new feature space / axes

   i) Calculate the Covariance matrix $C = A^T A$ where $A = [Ø_1, Ø_2, Ø_3, \ldots, Ø_M]$

   Dimension of $A = N^2 M$ & $A^T = M N^2$

   Hence $C = (MN^2)\ (N^2 M) \Rightarrow (M\ M)$

   ii) Generate M eigen vectors (Principal Components / Axes)
4. Select K best Eigen Vectors K<M
5. Convert lower dimensional K eigen vectors into original Dimensionality and do projection along these eigen images to find new features for classification.

B. LINEAR DISCRIMINANT ANALYSIS (LDA)

The goal of the Linear discriminant analysis technique is to project the original data matrix onto a lower dimensional space. The Linear discriminant analysis technique [**5**] is developed to transform the features into a lower dimensional space, which maximizes the ratio of the between-class variance to the within-class variance, thereby guaranteeing maximum class separability. To achieve this goal, three steps to be performed. The first step is to calculate the separability between different classes (i.e. the distance between the means of different classes), which is called the between-class variance or between-class matrix. The second step is to calculate the distance between the mean and the samples of each class, which is called the within-class variance or within-class matrix. The third step is to construct the lower dimensional space which maximizes the between-class variance and minimizes the within class variance. The flowchart of Linear Discriminant Analysis is given in figure 2.

Linear Discriminant Analysis Algorithm

Input : Given a data matrix N X M. Where N denotes number of samples X = (x1, x2, x3, …….xN). Each sample xi is represented as a vector with a length of M.

Output : Generates a lower dimensional picture a shadow (fisher images) of given N images.

Steps:

1. Read N sample images. Each image is represented as a row vector with a length of M features and form a data matrix X with a size of N X M.
2. Partition the data matrix into c classes (ie x1, x2, … xc). Each class xj has M features. Find the mean of each class μj (1 × M) using the equation

$$\mu_j = \frac{1}{n_j}\sum_{i=1}^{M} x_i$$

3. Find the total mean of all data μ(1×M) using the equation

$$\mu = \frac{1}{c}\sum_{i=1}^{c} \mu_i$$

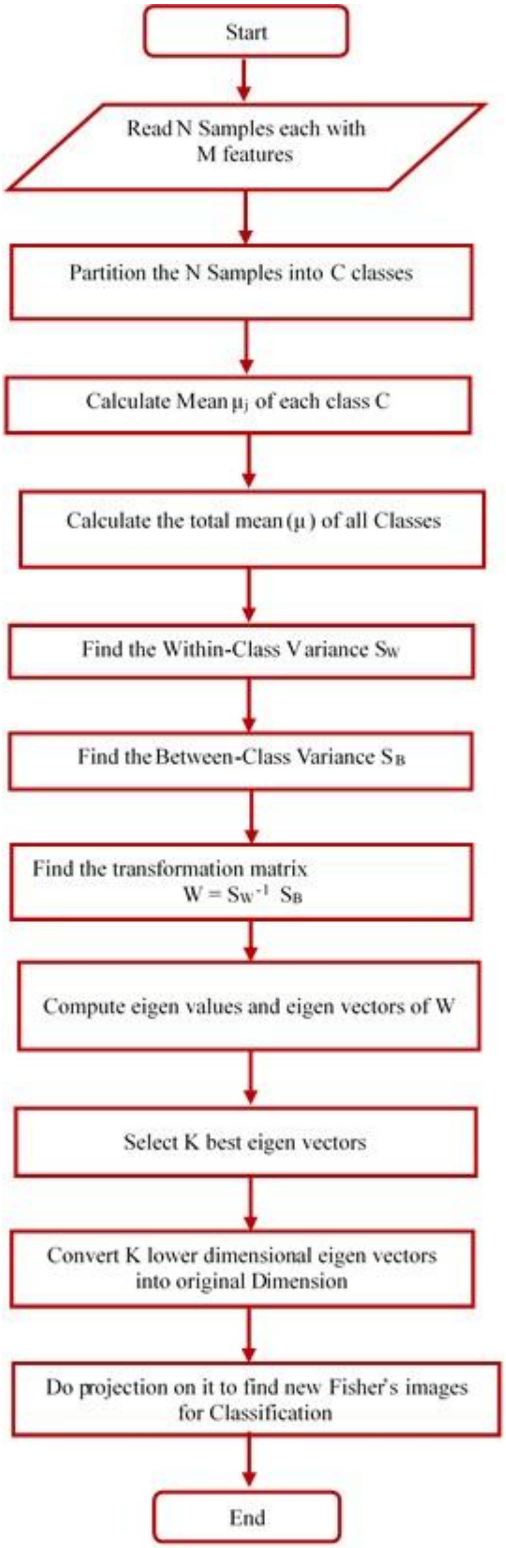

Figure 2. Flowchart of Linear Discriminant Analysis

4. Find the within-class variance S_W. It represents the difference between the mean and the samples of that class using the equation
$$S_W = \sum_{j=1}^{c}\sum_{i=1}^{n_j}(x_{ij}-\mu_j)(x_{ij}-\mu_j)^T$$
5. Find the between-class variance (S_B). Represents the difference between the mean and the samples of that class using the equation
$$S_B = \sum_{i=1}^{c} n_i(\mu_i-\mu)(\mu_i-\mu)^T$$
6. Find the transformation matrix (W)
$$W = S_W^{-1}S_B$$
7. The eigenvalues (λ) and eigenvectors (V) of W are then calculated.
$$S_W W = \lambda S_B W$$
Where the eigenvalues are
$$\lambda = \{\lambda_1, \lambda_2, \lambda_3 \ldots\ldots \lambda_M\}$$
and eigenvectors are
$$V = \{v_1, v_2, v_3 \ldots\ldots v_M\}$$
8. Sorting the eigenvectors in descending order according to their corresponding eigenvalues. The first k eigenvectors are then used as a lower dimensional space (V_k).
9. Project all original samples (X) onto the lower dimensional space of LDA using the equation.
$$Y = XV_k$$

Where each sample (X_i) which was represented as a point a M-dimensional space will be represented in a k-dimensional space by projecting it onto the lower dimensional space (V_k) as follows, $Y_i = X_iV_k$.

III THE ROLE OF PCA AND LDA FOR GAIT FEATURE EXTRACTION

Gait is an important biometric feature which defines the way of motion. Human Gait is used as major identifying feature to generate the unique gait sequence for each individual. Gait is one of the biometric which can be identified at a distance or at low resolution. Compared with other kinds of biometrics (such as face, iris, and fingerprint), gait has the merit of non-contact, unobtrusive and can be used for human recognition at a distance when other biometrics are obscured. Gait includes both the body appearance and the dynamics of human walking motion. Intuitively, recognizing people by gait depends greatly on how the silhouette shape of an individual changes over time in an image sequence. So, we may consider gait motion to be composed of a sequence of static body poses and expect that some distinguishable signatures with respect to those static body poses can be extracted and used for recognition by considering temporal variations of those observations.

The gait pattern is described by the periodic sequence width images, which contain the static and dynamic gait feature [6]. Gait feature of movement is a high dimension eigenvector, which not only need large storage room, but also increases the space and time complexity of system. Therefore, to reduce the dimensionality of gait feature, some process is necessary carry out before recognition [7].

The generation of binary silhouette frames of walking subjects is the initial step. Some distinguishable gait features, viz., centroid, aspect ratio, orientation, height and width are extracted from the silhouette frames to acquire feature vectors [8] [9].

The generation of binary silhouette frames is shown in figure 3.

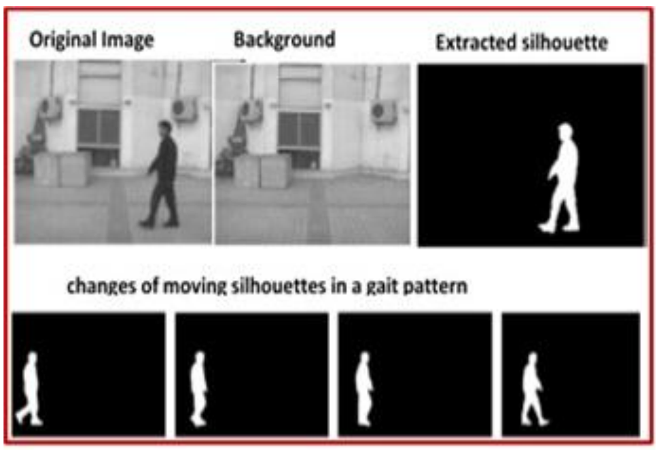

Figure 3. Generation of binary silhouette frames

C. Gait Based PCA Algorithm

1. First the binary silhouettes of each frame were extracted and centralized.
2. Represent all binary silhouettes in a complete gait cycle.
   $Gc = \{ Gc_1, Gc_2, \dots ,Gc_N \}$
   Where N is the number of binary silhouettes in a gait cycle.
3. Construct a set of averaged silhouettes,
   Avs = { Avs(1), Avs(2),…, Avs(i) } for this compute average of silhouettes in a gait cycle is shown in figure 4.

$$Avs(i) = \frac{1}{N}\sum_{l=1}^{N} Gc(l)$$

4. Eigen space transformation which is based on Principal Component Analysis (PCA) is applied to the averaged silhouettes to reduce the dimensionality of the feature space. Let $Avs_1, Avs_2, \dots, Avs_m$ be a set of averaged silhouettes. The largest eigenvectors of the matrix

$$C = \sum_{i=1}^{m} Avs_i \, Avs_i^T$$

   make a subspace that can rebuild the averaged silhouette with less dimensions. We applied a threshold to ignore small eigenvalues and their eigen vectors .
5. Do projection along these eigen images to find new features for classification.

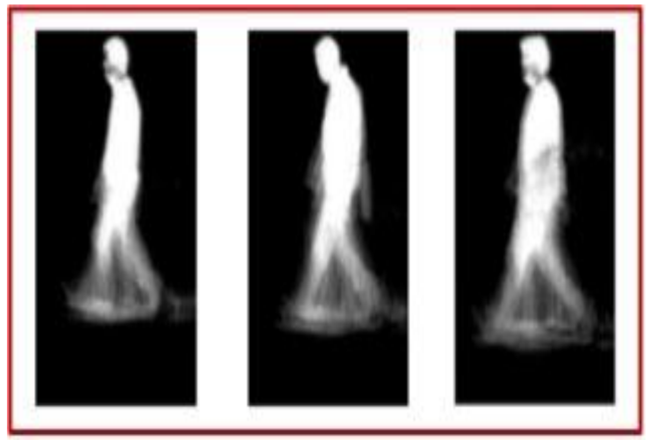

Figure 4. Average of silhouettes

The flowchart of Gait Based PCA is shown in figure 5.

D. Gait Based LDA Algorithm

1. Using background subtraction segment contours of human body then a gait cycle is computed with the change of body silhouette.
2. Using the images in gait cycle sequences compute the average of binary silhouette images for each class μ_j and overall class μ.
3. Find within-class variance S_W and between class variance S_B.
4. Find transformation matrix W.
5. The eigen values (λ) and eigenvectors (V) of W are then calculated.
6. Sort the eigen vectors in decreasing order and select the first K low dimensional eigen vectors.
7. Project all original samples onto the lower dimensional space of LDA.

The flowchart of Gait Based LDA is shown in Figure 6.

IV RELATED WORKS

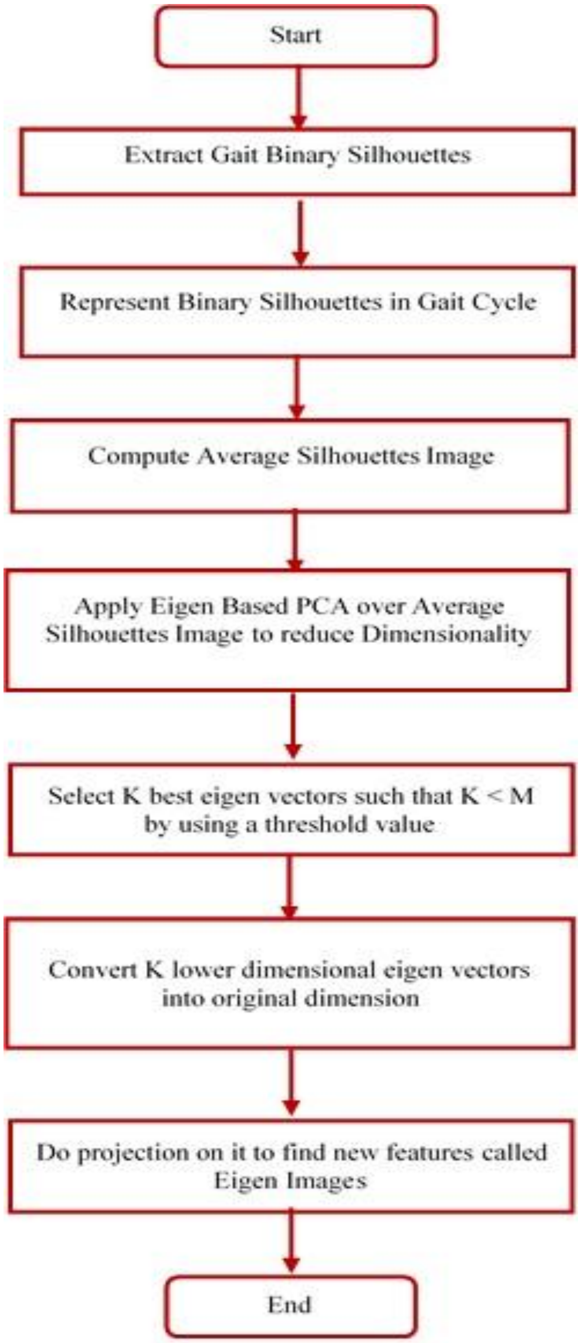

Figure 5. Flowchart of Gait Based PCA

Human identification at a distance has recently gained growing interest from computer vision researchers. Gait recognition aims essentially to address this problem by identifying people based on the way they walk. The eigen space transformation based on Principal Component Analysis (PCA) is applied [10] to time-varying distance signals derived from a sequence of gait silhouette images to reduce the dimensionality of the input feature space. Both PCA and LDA together applied [11] to improve the topological structure and reduce the dimensionality of the feature space. Combined PCA and Multiple Discriminant Analysis (MDA) [12] applied to process Gait Energy Image (GEI). PCA-based methods only preserve those features which contribute most to variance, which may be not optimal for classification. A general tensor discriminant analysis (GTDA) [13] proposed to preserve discriminative information of Gabor features and used LDA for classification. More recently, discriminative locality alignment (DLA) [14] is utilized to reduce dimensionality of biologically inspired features, while Hu et al. [15] apply a two-stage PCA+DLA to get Periodicity Feature Vector (PFV) and shape features. A two-dimensional locality preserving projections (2DLPP) is used by Zhang etal. [16] to improve the discriminative power of features extracted based on active energy image (AEI). While most of the aforementioned approaches focus on feature dimensionality reduction, Guo and Nixon [17] [18] select gait feature subset by maximizing the mutual information of gait features. Average Gait Differential image (AGDI) is proposed [19] to generate the accumulation of the silhouettes difference between adjacent frames, the advantage of this method lies in that as a feature image it can preserve both the kinetic and static information of walking. Two-dimensional Principal Component Analysis (2DPCA) [20] is based on 2D image matrices rather than 1D vectors so the image matrix does not need to be transformed into a vector prior to feature extraction. Instead, an image covariance matrix is constructed directly using the original image matrices, and its eigenvectors are derived for image feature extraction. A new multi-view gait recognition approach [21] creates a so called View Transformation Model (VTM) based on spatial-domain Gait Energy Image (GEI) by adopting Singular Value Decomposition (SVD) technique, to further improve the performance of the proposed VTM, Linear Discriminant Analysis (LDA) is used to optimize the obtained GEI feature vectors.

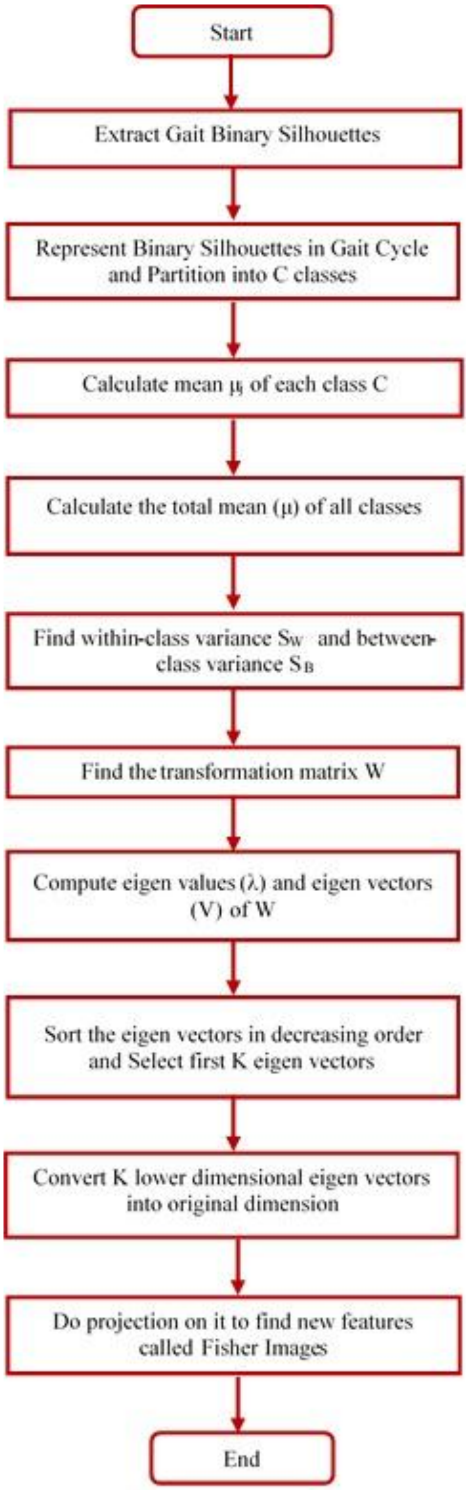

Figure 6. Flowchart of Gait Based LDA

V COMPARATIVE STUDY

Appearance-based methods are widely used in object recognition systems. Within this hypothesis, PCA and LDA can be used for extracting features from images in many applications particularly in gait and face recognition. Both Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) are linear transformation techniques that are usually used for dimensionality reduction. PCA can be defined as an "unsupervised" algorithm, since it "ignores" class labels and its goal is to find the directions (the so-called principal components) that maximize the variance in a dataset. In contrast to PCA, LDA is "supervised" and calculates the directions ("linear discriminants") that will represent the axes that maximize the separation between multiple classes. Specifically the complexity of PCA is $O(Xk^3+ k^2)$and LDA has $O(Xk^2 + k^3)$ time complexity, where k is the number of gait features.[22][23].

Although it might sound intuitive that LDA is greater to PCA for a multi-class classification task where the class labels are known, this might not always the case. When a small (or non representative) training data set is used, there is no guarantee that LDA will outperform PCA.

VI CONCLUSION

The aim of feature extraction is to find the most pertinent data from the original data and represent that data in low dimensional space. Principal Component Analysis is a simple and fast method to identify the principal components of the input image there by reducing the dimensionality of the input image called eigen images. But Eigen image Approach might not be applicable to the real system because it needs to be more robust and to have other discriminant features. While Fisher image Approach (LDA) is a simple yet popular method for handling high dimensional data as class labels are available. LDA is often superior to PCA in feature extraction for classification but does not always perform better. It does seem to be superior to PCA when the training data set is large.

REFERENCES

[1] Gaurav Kumar, Pradeep Kumar Bhatia, A Detailed Review of Feature Extraction in Image Processing Systems, 2014 Fourth International Conference on Advanced Computing & Communication Technologies.

[2] Turk M, Pentland A. Eigenfaces for recognition. J Cognitive Neurosci. 1991; 3(1):71–86.

[3] Lindsay I Smith, A tutorial on Principal Components Analysis, 2002.

[4] M. Murali, Principal Component Analysis based Feature Vector Extraction, Indian Journal of Science and Technology, Vol 8(35), 2015.

[5] Alaa Tharwat, Tarek Gaber, Abdelhameed Ibrahim, Aboul Ella Hassanien, Linear discriminant analysis: A detailed Tutorial, AI Communications, 2017.

[6] Han Su, Zhi-Wu Liao, Guo-Yue Chen, A gait recognition method using L1-PCA and LDA, IEEE, International Conference on Machine Learning and Cybernetics, 2009.

[7] Pranjit Das, Sarat Saharia, Human Gait Recognition Based on Principal Component Analysis, International Journal of

Computer Sciences and Engineering, Volume-4, Special Issue-7, Dec 2016.

[8] Qiong Cheng, Bo Fu, and Hui Chen, Gait Recognition Based on PCA and LDA, Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09), 2009.

[9] Gait recognition based on gait energy image and linear discriminant analysis, IEEE International conference on signal processing, communications and computing, 2015.

[10] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu, Silhouette Analysis-Based Gait Recognition for Human Identification, IEEE Trans. on Pattern analysis and machine intelligence, vol. 25, no. 12, december 2003.

[11] T. Daoliang, H. Kaiqi, Y. Shiqi, and T. Tieniu, "Orthogonal Diagonal Projections for Gait Recognition," in 2007 IEEE International Conference on Image Processing, 2007, pp. I - 337-I - 340.

[12] P.N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711-720, 1997.

[13] T. Dacheng, L. Xuelong, W. Xindong, and S. J. Maybank, "General Tensor Discriminant Analysis and Gabor Features for Gait Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 1700-1715, 2007.

[14] Y. Mu and D. Tao, "Biologically inspired feature manifold for gait recognition," Neurocomputing, vol. 73, pp. 895-902,2010.

[15] R. Hu, W. Shen, and H. Wang, "Recursive spatio temporal subspace learning for gait recognition," Neuro computing,vol. 73, pp. 1892-1899, 2010.

[16] E. Zhang, Y. Zhao, and W. Xiong, "Active energy imageplus 2DLPP for gait recognition," Signal Processing, vol.90, pp. 2295-2302, 2010.

[17] G. Baofeng and M. S. Nixon, "Gait Feature Subset Selectionby Mutual Information," IEEE Transactions on Systems,Man and Cybernetics, Part A: Systems and Humans vol. 39, pp. 36-46, 2009.

[18] G. Baofeng and M. S. Nixon, "Gait Feature Subset Selection by Mutual Information," in First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007, pp. 1-6.

[19] Jinyan Chen Jiansheng Li, Average Gait Differential Image Based Human Recognition, Scientific World Journal. 2014. 27.

[20] Yang J, Zhang D, Frangi AF, Yang J-Y. Two-dimensional PCA: a new approach to appearance based face representation and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004;26(1):131–137.

[21] Worapan Kusakunniran, Qiang Wu, Hongdong Li, Jian Zhang, Multiple views gait recognition using View Transformation Model based on optimized Gait Energy Image, IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), 2009

[22] Deng Cai, Xiao fei He, Jia wei Han."Training Linear Discriminant Analysis in Linear Time," IEEE International Conference on Data Engineering, 2008(24):pp. 209-217.

[23] Fradkin, Dmitriy1,Madigan, David2."Experiments with random projections for machine learning," IEEE 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'03. 2003(8):pp 517-522.

