International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 07 | July-2016 www.irjet.net p-ISSN: 2395-0072

Various classifiers based on their accuracy for age estimation through facial features

Anchal Tomar

M.Tech Student, Computer Science and Engineering, M.I.E.T college (Meerut) APJAKTU University,(lucknow), Uttar Pradesh ,India,anchal.tomar.2@gmail.com

---

Abstract - *Facial recognition technique is one of array of technologies being developed to address a fundamental concern of modern societies: the problem of "disembodied identities", or the existence of geometrical and textual representation of individuals that give independent information of their physical appearance. But as we know recognition problem is not an easy task as it has great variability in moving body and tilt, lighting intensity and angle, facial expression, aging, etc. Hence it is important to focus on technology which provides best results his paper concerns with simply comparing some existing techniques like fisher face or linear discriminate analysis (LDA), Eigen face or principal component analysis (PCA) , support vector machine(SVM),(k nearest neighbor) KNN to find which one attain good accuracy to estimate age group using face features. Age estimation process involves three stages: pre-processing, feature extraction and classification.*

***Key Words*: Fisher face or linear discriminate analysis (LDA), Eigen face or principal component analysis (PCA, support vector machine (SVM)), (k nearest neighbor) KNN**

1. INTRODUCTION

An issue of personal verification and identification is an effectively developing region of research due to which biometric face recognition technology has gotten noteworthy consideration in the previous quite a long while because of its potential for a wide assortment of utilizations in different fields for the most part like law implementation and non-law requirement[2][5].
Face, voice, fingerprint, iris, ear, retina are normally utilized authentication methods but face recognition has distinct preferences due to its non-contact process. Research in those areas has been conducted for more than 35 years [3][4]. Facial recognition systems are commonly used for security purposes yet are progressively being utilized as a variety of other real world applications such as land enrollment, passports, driver's licenses, and acknowledgment of a human in a security territory[5].I observe during my research that numerous scientist incorporate me first attempt to discover best algorithm for research and afterward begin reason along these lines, in this paper for age classification using feature selection used many classification techniques (**KNN,LDA,PCA ,SVM…**) which analyzed and finish up which of the algorithm is best in age grouping at what circumstance. To address this task, we study various classification techniques, namely, support vector machine (SVM), linear discriminant analysis (LDA) and K nearest neighbor (KNN) [6]. We observe that despite KNN is as effective as SVM but KNN prohibits its usage due to high response time when data is in high dimensions [16]. To speed up KNN retrieval, some researcher's propose a feature reduction technique using principle component analysis (PCA) to facilitate near real time face recognition along with better accuracy[15][16]. Hence, this paper explains various classification approaches, namely, SVM, KNN, PCA, LDA and survey purposes behind the adequacy of KNN over SVM and LDA, PCA.

2. RELATED WORK

Age Estimation approach fall with two standards. As per the main stream the issue is dealt with as a standard classification problem, illuminated utilizing standard classifiers where age estimation is performed by assigning a set of facial features to an age group[1][3]. Inside this setting facial components utilized might be connected with the general appearance of a face or might be related to age-related elements (e.g. wrinkles)[4]. As an option age estimation approaches that rely on the demonstrating of the aging process have been developed. In this segment typical approaches described in the literature are quickly introduced. The aim of this review is not to present an exhaustive literature review of the topic but instead to highlight the development of the subject [5]. A more point by point presentation of the related literature is exhibited by Ramanathan (Ramanathan 2009) and Fu (Fu 2010)[2][4].
One of the main attempts to create facial age estimation algorithms was accounted for by Kwon and Lobo (Kwon 1999)[1]. Kwon and Lobo use two principle sorts of features: Geometrical ratios calculated based on the distance and the size of certain facial characteristics and an estimation of the amount of wrinkles detected by deformable shapes (snakes) in facial areas where wrinkles are typically experienced[1][2]. In view of these elements Kwon and Lobo (Kwon 1999) group faces into babies, adults and seniors based on a computational theory for visual age classification from facial pictures. To begin with, primary features of the face, namely the eyes, nose, mouth, chin, and virtual top of the head, are found.

According to "Facial Age Classification using Sub pattern-based Approaches" by Fatemeh Mirzaei and Önsen Toygar , sub pattern-based approaches are utilized to take care of the age classification issue on facial images[17]. Sub pattern-based approaches named Local Binary Patterns (LBP), (spPCA) sub pattern-based Principal Component Analysis and (mPCA) modular Principal Component Analysis are inspected to exhibit the age classification performance on female and male facial images of human beings with a few parameter settings.Classification of age intervals are directed independently on female and male facial images subsequent to the maturing procedure for female and male is distinctive for individuals in real life[17]. Sub pattern-based LBP, spPCA and mPCA are utilized for feature extraction on various datasets chose from FG-NET and MORPH databases. Exploratory results exhibit the prevalence of sub pattern-based LBP over spPCA and mPCA methods. Age classification performance using these three sub pattern-based techniques with different parameter settings on the selected datasets is likewise exhibited[17][4].
All presume that research in age-estimation started in early 1990s[5] and up to now, many approaches have been proposed. Here are some algorithmn explains used in research.

2.1. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) , also known as "Karhunen-Loeve expansion", is a general technique of analysis for correlated multi-variable dataset widely used in the areas of pattern recognition and computer vision such as face recognition [8][11]. It is a good tool to reduce multidimensional data to some lower dimensions while retaining important information of data. It also covers standard deviation, covariance, and eigenvectors [9].

2.1.1. Advantage and disadvantage of PCA

The advantages of PCA are listed below:

1) Lack of data redundancy [10][11].
2) Smaller database representation since only the trainee images are stored [8].
3) Low noise sensitivity as the maximum variation basis is chosen and so the small variations in the background are ignored automatically [9].

The disadvantages of PCA are:

1) PCA is an unsupervised technique hence only suitable for databases that contain images with no class labels.[8]
2) Low accuracy as covariance matrix is difficult to evaluate.[9]
3) PCA is not as powerful as other methods to distribute classes in small datasets.[10]
4) PCA is failed to capture even the simplest invariance unless the training data explicitly provides this information.[8][11].

2.2. Linear Discriminate Analysis (LDA)

Linear discriminant analysis (**LDA**) is a speculation of **Fisher's linear discriminant**, a technique utilized as a part of measurements, pattern recognition and machine learning to find a linear combination of features that characterizes or isolates two or more classes of objects or events[11][12]. The subsequent mix might be utilized as a linear classifier or, more regularly, for dimensionality reduction before later grouping.

2.2.1. Advantage and disadvantage of LDA

Advantages

1) Reduce dimensionality without loss of information and speed up the classifier for training [12].
2) LDA based algorithms beat PCA based ones, since the previous enhances the low dimensional representation of the objects with focus on the most discriminant feature extraction while the last accomplishes simply object reconstruction [13].

Disadvantages

1) An intrinsic restriction of traditional LDA is the so called singularity issue, that is, it fizzles when all scatter matrices are singular [16][10].
2) However, a critical problem using LDA, especially in face recognition area, is the Small Sample Size (SSS) Problem. This issue is experienced practically speaking following there are regularly an expansive number of pixels available, but the yet the aggregate number of training samples is less than the dimension of the feature space [11]. This infers all scatter matrices are singular and thus the traditional LDA algorithm fails to use.

2.3 SVMs (Support Vector Machines)

SVMs (Support Vector Machines) are a valuable strategy for data classification and are still under serious research [1][12]. SVM is fundamentally a classier technique that performs classification tasks by building hyper planes in a multidimensional space that isolates instances of various class labels. SVM bolsters both regression and classification tasks and can deal with multiple continuous and categorical variables [16].
In machine learning, support vector machine**s** are supervised learning models with related learning algorithms that dissect data utilized for classification and regression analysis [2]. Given a set of training illustrations (Figure1), each marked for belonging to one of two categories, a SVM training algorithm assembles a model that doles out new case into one classification or the other, making it a non-probabilistic binary linear classifier[16][17]. A SVM model is a representation of the case as points in space, mapped so that the examples of the different classifications are partitioned by a clear gap that is as wide as could reasonably be expected. New illustrations are then mapped into that same space and anticipated to

have a place with a class in view of which side of the gap they fall on.

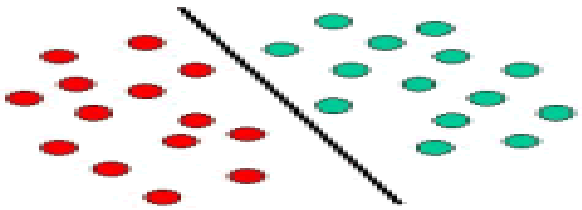

Figure1: - classic example of a linear classifier

2.3.1 Disadvantages of SVM

1) Although SVMs have great speculation execution, they can be wretchedly moderate in test stage, an issue tended to in (Burges, 1996; Osuna and Girosi, 1998)[12].
2) SVMs is the high algorithmic complexity and broad memory necessities requirements of the required quadratic programming in vast -scale tasks.

2.4 k-Nearest Neighbor (K-NN)

K-NN is a sort of instance-based learning, or lazy learning, where the function is just approximated locally and all calculation is conceded until classification [3]. An object is characterized by the ―distance from its neighbors, with the object being allotted to the class most common among its k distance-nearest neighbors Figure2. In the event that k = 1, the calculation basically turns out to be nearest neighbor algorithm and the object is classified to the class of its nearest neighbor [5][6].

2.4.1 KNN Algorithm

The fundamental steps of the k-NN algorithm are;

1) To compute the distances between the new sample and all previous samples, have as of now been classified into clusters [14].
2) To sort the distances in expanding arrange and select the *k* samples with the smallest distance values [2][14].
3) To apply the voting standard. A new sample will be included (classified) to the largest cluster out of *k* selected samples

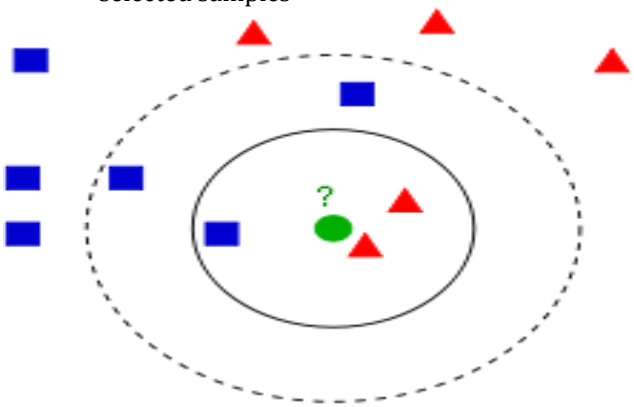

Figure2: - Example of KNN classification

2.4.2 Advantages:

1) K-NN has some solid consistency results and simple to actualize as contrast with other.
2) Analytically tractable and straightforward usage [3][14].
3) Uses nearby data, which can yield exceedingly versatile conduct.
4) Lends itself effortlessly to parallel executions.

3. IMPLEMENTATION

The classifier involve following process step:

Preprocessing:

- Face image is processed to get a transformed face image to increase the quality of the face image, holding the vital attributes [1][3].
- Data pre-processing includes cleaning, normalization, transformation, featur e extraction and selection, and so on. The result of data pre-processing is the final training set [7].

Feature extraction

- It begins from an underlying arrangement of measured data and assembles inferred values planned to be informative and non-redundant, encouraging the subsequent learning and generalization steps, and sometimes prompting better human elucidations[5][17].
- When the input data to an algorithm is too extensive to be processed and it is suspected to be redundant, then it can be changed into a reduced set of features (also named a "features vector")[8]. This process is called *feature extraction*. It involves 2 processes for feature extraction: geometric features and wrinkle features. Below given diagram shows complete process of the system [13].

Classification

1) Age ranges are classified progressively relying upon on number of groups using K-Means clustering algorithm effectively [2].
2) Experimental results demonstrate that only wrinkle features is more critical for age group classification [4].

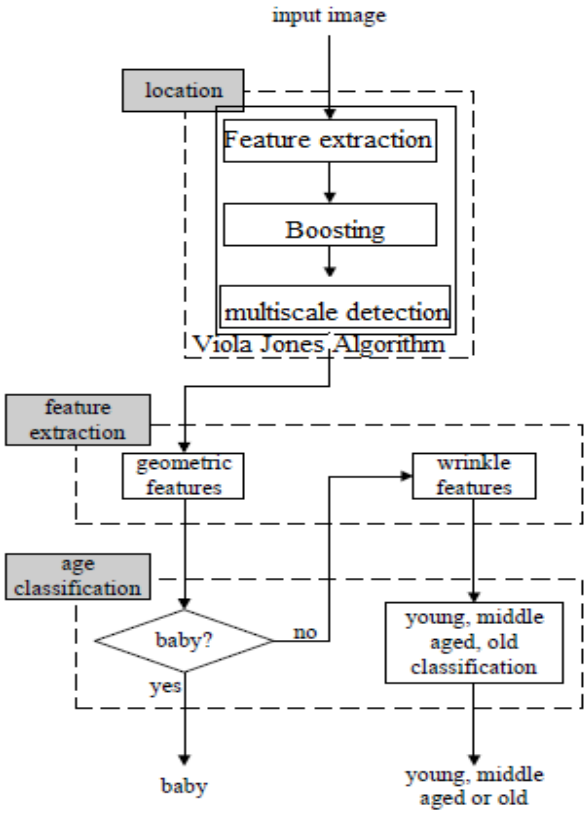

Figure3:-Process of the system

4. Conclusion

Face recognition comprise a wide range of methods to perceive a face [1]. We have experienced PCA and LDA, KNN, SVM, among all PCA is the most straightforward, efficient and oldest one yet learning is time consuming so we prefer LDA [17]. In LDA the singularity problem comes anyway it best tries to sort the illumination issue. PCA is neglected to catch even the most straightforward invariance unless the training data explicitly provides this information. In any case, a critical issue using LDA, especially in face recognition area, is the Small Sample Size (SSS) Problem [4][12]. This problem is encountered in practice since there are frequently an expansive number of pixels accessible, yet the aggregate number of training samples is less than the dimension of the feature space[4][2]. This implies that all scatter matrices are singular and thus the traditional LDA algorithm fails to use. Despite the fact that SVMs have good generalization performance, they can be moderate in test stage, a problem addressed in (Burges, 1996; Osuna and Girosi, 1998).Hence we discover KNN all the more simple excessively learn, effective and consistent in result among every one of them[3][4].

References

[1]Y. Kwon and N. Lobo, "Age classification from facial images," *Compute. Vis. Image Understand.*, vol. 74, no. 1, pp. 1–21, 1999.

[2] N.Ramanathan and R. Chellappa, "Face verification across age progression", in Proc. IEEE Conf. Computer Vision and Pattern Recognition, San Diego, CA, 2005, vol.2, pp.462-469.

[3] A.Lanitis, C.Draganova, and C.Christodoulou, "Comparing different classifiers for automatic age estimation", IEEE Trans.Syst.Man, Cybern.B, Cybern, vol34, no.1, pp.621-628, Feb.2004.

[4] N.Ramanathan and R. Chellappa, "Modeling Age Progression in young faces", in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), vol.1, pp.387-394, 2006.

[5] H. H. Barret. Foundations of Image Science. John Wiley & Sons, New Jersey, U.K., third edition, 2004.

[6] R. Jana, H. Pal, A. R. Chowdhury, "Age Group Estimation Using Face Angle", IOSR Journal of Computer Engineering (IOSRJCE), Volume 7, Issue 5, PP 35-39, Nov-Dec. 2012.

[7] R. C. Gonzales and R. E. Woods. Digital Image Processing. Prentice Hall, second edition, 2002. 795 pages, ISBN 0-201-18075-8.

[8] W. Miziolek and D. Sawicki, "Face Recognition: PCA or ICA," PrzegladElektrotechniczny, Vol. 88, 2012, pp. 286-288.

[9] C. Li, Y. Diao, H. Ma and Y. Li, "A Statistical PCA Method for Face Recognition," in Intelligent Information Technology Application, 2008, pp. 376-380.

[10] H. Duan, R. Yan and K. Lin, "Research on Face Recognition Based on PCA," in Future Information Technology and Management Engineering, 2008, pp. 29-32.

[11] 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, no. 7,pp. 711– 720, 1997.

[12] Anissa Bouzalmat, Jamal Kharroubi and Arsalane Zarghili, "Face Recognition Using SVM Based on LDA", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 4, No 1, July 2013

[13] W. Zhao, R. Chellappa, and P. Philips. Subspace Linear Discriminant Analysis for Face Recognition. Technical Report CAR-TR-914. Center for Automation Research, University of Maryland, 1999.

[14] A.Lanitis and C.J.Taylor, "Towards Automatic Face Identification Robust to Ageing Variation", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.24, no.24, p.442-455, 2002.

[15] A.Lanitis, C.Draganova, and C.Christodoulou, "Comparing different classifiers for automatic age estimation", IEEE Trans.Syst.Man, Cybern.B, Cybern, vol34, no.1, pp.621-628, Feb.2004.

[16] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morph able model", IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(9):1063 –1074, September 2003.

[17] Fatemeh Mirzaei and Önsen Toygar,"Facial Age Classification using Sub pattern-based Approaches", *Proc. SPIE* 8009, Third International Conference on Digital Image Processing (ICDIP 2011), 80091F (July 08, 2011); doi:10.1117/12.896285