ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

# Comparison of various face detection and face recognition algorithms

[1] A. Mallikarjuna Reddy, [2] K Naresh, [3] S.Shiva Kumar, [4] Madar Bandu
[1][2][3][4] Assistant Professor, Dept. of CSE, Anurag Group of Institutions, Hyderabad, T.S., India.

**Abstract: The biometric is a study of human behavior and features. Face detection and Face recognition is a technique of biometric. Various approaches are used for it. A lot of Face detection and Face recognition algorithms have been developed during the past decades. Face detection and Face recognition is emerging branch of biometric for security as no faces can be defeated as a security approach. The present paper presents several famous Face detection and Face recognition algorithms, such as Holistic-based, Feature based, Template matching and part based , knowledge-based, feature invariant and appearance-based methods.**

**Keywords: Face detection; Face Recognition; Pattern Recognition; Image Processing.**

## I. INTRODUCTION

Face detection and Face recognition are the most significant applications of image analysis. Face detection and Face recognition are the most interesting and significant research fields in the past two decades. The reasons come from the need of automatic recognitions and surveillance systems, the interest in human visual system on face recognition, and the design of human-computer interface, etc. These researches involve knowledge and researchers from disciplines such as neuroscience, psychology, computer vision, pattern recognition, image processing, and machine learning, etc. A lot of papers have been published to overcome dissimilarity factors and achieve better recognition rate, while there is still no robust technique against uncontrolled practical cases which may involve kinds of factors concurrently. Structures of recognition, significant issues and factors of human faces, critical techniques and algorithms, and finally give a comparison and conclusion. Readers who are interested in face recognition could also refer to published surveys [1-3] and website about face recognition [4].

pattern recognition[5] can be classified into four categories:(i)Template matching, (ii)statistical approaches, (iii)syntactic approach, and (iv)neural networks. The template matching category builds several templates for each label class and compares these templates with the test pattern to achieve a suitable decision. The statistical approaches is the main category that will be discussed in this paper, which extracts knowledge from training data and uses different kinds of machine learning tools for dimension reduction and recognition.

The syntactic approach is often called the rule based pattern recognition, which is built on human knowledge or some physical rules, for example, the word classification and word correction requires the help of grammars. The term, knowledge, is referred to the rule that the recognition system uses to perform certain actions. Finally, the well-known neural network is a framework based on the recognition unit called perceptron. With different numbers of perceptrons, layers, and optimization criteria, the neural networks could have several variations and be applied to wide recognition cases.

There're two main categories of (i) dimension reduction techniques: (ii) domain knowledge approaches and data-driven approaches. The domain-knowledge approaches perform dimension reduction based on knowledge of the specific pattern recognition case. For example, in image processing and audio signal processing, the discrete Fourier transform (DFT) discrete cosine transform (DCT) and discrete wavelet transform are frequently used because of the nature that human visual and auditory perception have higher response at low frequencies than high frequencies. Another significant example is the use

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

of language model in text retrieval which includes the contextual environment of languages.The holistic-based viewpoint claims that human recognize faces by the global appearances, while the feature-based viewpoint believes that significant features such as eyes, noses, and mouths play dominant roles in identifying and remembering a person.face detection algorithms [7] are classified as follows: (i)knowledge-based, (ii)feature invariant, (iii)template matching, and (iv)appearance-based method. The present paper is organized as follows. The section (II) describes the related work; section (III) presents the conclusions.

## II. RELATED WORK

G. Yang et.al [7] proposed a method is composed of the multi-resolution hierarchy of images and specific rules defined at each image level. The hierarchy is built by image sub-sampling. The face detection procedure starts from the highest layer in the hierarchy (with the lowest resolution) and extracts possible face candidates based on the general look of faces. Then the middle and bottom layers carry rule of more details such as the alignment of facial features and verify each face candidate. This method suffers from many factors described in Section 3 especially the RST variation and doesn't achieve high detection rate , while the coarse-to-fine strategy does reduces the required computation and is widely adopted by later algorithms.

C. Kotropoulos et.al [8] proposed a method uses the fairly simple image processing technique, the horizontal and vertical projection . Based on the observations that human eyes and mouths have lower intensity than other parts of faces, these two projections are performed on the test image and local minimums are detected as facial feature candidates which together constitute a face candidate. Finally, each face candidate is validated by further detection rules such as eyebrow and nostrils.This me-thod is sensitive to complicated backgrounds and can't be used on images with mul-tiple faces.

Hsu et al. [9] proposed to combine several features for face detec-tion. They used color information for skin-color detection to extract candidate face regions. In order to deal with different illumination conditions, they extracted the 5% brightest pixels and used their mean color for lighting compensation. After skin-color detection and skin-region segmentation, they proposed to detect invariant facial features for region verification. Human eyes and mouths are selected as the most significant features of faces and two detection schemes are designed based on chrominance contrast and morphological operations, which are called "eyes map" and "mouth map". Finally, form the triangle between two eyes and a mouth and verify it based on (i) luminance variations and average gradient orientations of eye and mouth blobs, (ii) geometry and orientation of the triangle, and (iii) the presence of a face boundary around the triangle. The regions pass the verification are denoted as faces and the Hough transform are performed to extract the best-fitting ellipse to extract each face.

Leung et al. [10] proposed a probabilistic method to locate a face in a cluttered scene based on local feature detectors and random graph matching. Their motivation is to formulate the face localization problem as a search problem in which the goal is to find the arrangement of certain features that is most likely to be a face pattern. In the initial step, a set of local feature detectors is applied to the image to identify candidate locations for facial features, such as eyes, nose, and nostrils, since the feature detectors are not perfectly reliable, the spatial arrangement of the features must also be used for localize the face.

M. Kass et.al [12] proposed the deformation constraints are determined based on user-defined rules such as first- or second-order derivative properties. These constraints are seeking for the smooth nature or some prior knowledge, while not all the patterns. Furthermore, the traditional techniques are mainly used for shape or boundary matching, not for texture matching. Kass et al. [12] proposed a ASM (active shape model) exploits information from training data to generate the deformable constraints. They applied the principal component analysis (PCA) [14] [15] to learn the possible variation of object shapes. Table 1 shows the summary of face detection techniques.

| S. NO | Category | Method | Characteristics |
|---|---|---|---|
| 1 | Knowledge -based | Hierarchical knowledge-based [7] | Coarse-to-fine procedure |
| | | Horizontal projection [8] | Coarse-to-fine procedure |
| 2 | Feature-based | Face Detection Using Color Information [9] | Combining skin-color detection, face shape verification, and facial feature configuration for detection |
| | | Face detection based on random labeled graph matching [10] | Combining simple features with statistical learning and esti-mation |
| 3 | Template matching | Active appearance model [15] | Learning facial shape and appearance variation by data |
| 4 | Appearance-based | Example-based learning [16] | Learning the face and non-face distribution by mixture of |
| | | Haar features with Adaboost [37] | Adaboost for speed-up |
| 5 | Part-based | Generative models [20] | Unsupervisedly extracting significant facial features, and learning the relation among parts and discrimination be- tween face and non-face by the graphical model structure. |
| | | Component-based with SVM [21] | Learning global and local SVM for detection |

***Table1: The summary of face detection techniques***

Although the principal component analysis can't exactly capture the nonlinear shape variation such as bending, this model presents a significant way of thinking: learning the deformation constraints directly from the possible variation.

The appearance-based methods consider not the facial feature points but all regions of the face. Given a window size, the appearance-based method scans through the image and analyze each covered region.

Sung et al. [16], proposed the window size of 19x19 is selected for training and each extracted patch can be represented by a 381-dimensional vector. A face mask is used to disregard pixels near the boundaries of the window which may contain background pixels, and reduce the vector into 283 dimensions. In order to better capture the distribution of the face samples, the Gaussian mixture model [17] is used. Given samples of face patches and non-face patches, two six-component Gaussian mixture models are trained based on the modified K-means algorithm [17]. The non-face patches need to be carefully chosen in order to include non-face samples as many as possible, especially some naturally non-face patterns in the real world that look like faces when viewed in a selected window. To classify a test patch, the distances between the patch and the 12 trained components are extracted as the patch feature, and a multilayer neural network [18][19] is trained to capture the relationship between these patch features and the corresponding labels.

R. Fergus et.al [20] proposed to learn and recognize the object models from unlabeled and unsegmented cluttered scenes in a scale invariant manner. Objects are modeled as flexible constellations of parts. The object model is generated by the probabilistic representation and each object is denoted by the parts detected by the entropy-based feature detector. Aspects including appearances, scales, shapes, and occlusions of each part and the object are considered and modeled by the probabilistic representation to deal with possible object variances.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

Bernd et.al [21] proposed the face detection algorithm consisting of a two-level hierarchy of support vector machine (SVM) classifiers [15][17]. On the first level, component classifiers independently detect components of a face. On the second level, a single classifier checks if the geometrical configuration of the detected components in the image matches a geometrical model of a face.

Belhumeur et.al [23] proposed to use the linear discriminative analysis (LDA) [14] for bases finding. The objective of applying the LDA is to look for dimension reduction based on discrimination purpose as well as to find bases for projection that minimize the intra-class variation but preserve the inter-class variation.

Bartlett et al. [25], they derived the ICA bases from the principle of optimal information transfer through sigmoidal neurons. In addition, they proposed to architectures for dimension-reduction decomposition, one treats the image as random variables and the pixels as outcomes, and the other one treats the pixels as random variables and the image as outcomes.

The Laplacian faces proposed by He et al. [26] used the locality preserving projections (LPP) [29] to find an embedding that preserves local information, and obtains a face subspace that best detects the essential face manifold structure.

Wright et al. [30] proposed to use the sparse signal representation for face recognition. They used the over-complete database as the projection basis, and applied the L1-minimization algorithm to find the representation vector for a human face. They claimed that if sparsity in the recognition problem is properly harnessed, the choice of features is no longer critical. What is crucial, however, is that whether the number of features is sufficiently large and whether the sparse representation is correctly computed.

Lades et.al [31] proposed the elastic graph matching framework is used for finding feature points, building the face model and performing distance measurement, while the Gabor wavelets are used to extract local features at these feature points, and a set of complex Gabor wavelet coefficients for each point is called a jet.

Wiskott et.al [32] proposed an improved elastic graph matching framework to deal with the computational-expensive problem above and enhance the performance.
They employed object-adaptive graph to model faces in the database, which means the vertices of a graph refer to special facial landmarks and enhance the distortion-tolerant ability

Ahonen et.al [33] proposed to extract the local binary pattern (LBP) histo-grams with spatial information as the face feature and use a nearest neighbor classifier based on Chi square metric as the dissimilarity measure. The idea behind using the LBP features is that the face images can be seen as composition of micro-patterns which are invariant with respect to monotonic gray scale transformations. Combining these micro-patterns, a global description of the face image is obtained.

The original LBP operator, introduced by Ojala et.al [34], is a powerful means of texture description. The operator labels the pixels of an image by thresholding the -neighborhood of each pixel with the center value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. Later the operator was extended to use neighborhoods of different sizes based on circular neighborhoods and bilinear inter-polation of the pixel values [54]. The notation (P,R), where P means the number of sampling points on a circle of radius R, is adopted.

Another extension to the original operator uses so called uniform patterns [36]. A local binary pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. Ojala et al. no-ticed that in their experiments with texture images, uniform patterns account for a bit less than 90 % of all patterns when using the (8,1) neighborhood and for around 70% in the (16,2) neighborhood.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

Traditional template-matching is pretty much like using distance metric [11]for face recognition, which means selecting a set of symbolic templates for each class (person), the similarity measurement is computed between a test im-age and each class, and the class with the highest similarity score is the selected as the correct match. Recently, deformable template techniques are proposed.

Edwards et.al [56] proposed to use the Mahalonobis distance measure for each class and generate a class-dependent metric to encounter the intra-class variation. To better exploit the inter-class variation against the intra-class variation, they also used the linear discri-minant analysis (LDA) for dimension reduction and classification task.

Heisele et.al [38] compared the performance of the component-based face recognition against the global approaches. In their work, they generated three different face recognition structures based on the SVM classifier: a component-based algorithm based on the output of the component-based face detection algorithm, a global algorithm directly fed by the detected face appearance, and finally a global approach which takes the view variation into account. Table 2 shows the summary of face recognition techniques.

| S. NO | Category | Method | Characteristics |
|---|---|---|---|
| 1 | Holistic-based | PCA [22] | PCA for learning eigenfaces, unsupervised |
| | | LDA [23] | LDA for learning fisherfaces, supervised |
| | | 2D-PCA [24] | 2D-PCA for better statistical properties |
| | | ICA [25] | ICA for catch facial independent components, two archi- tectures are proposed |
| | | Laplacian faces [26] | Nonlinear dimension reduction for finding bases, LPP |
| | | Evolutionary pursuit [27] | Using the genetic algorithm for finding the best projection bases based on generalization error |
| | | Kernel PCA And Kernel LDA [28] | Mapping the image into higher-dimensional space by the kernel function, and exploit the PCA and LDA bases there |
| | | Sparse representation [30] | Using L1 minimization and over-complete dictionary for finding sparse representation |
| 2 | Feature-based | Gabor and dynamic link architecture [31] | Gabor features extracted at facial feature locations, while performing one-by-one matching |
| | | Gabor and elastic bunch graph matching [32] | Gabor features extracted at facial feature locations, and obtaining the robust representation by the FBG matching. |
| | | LBP [33] | Local binary patterns are introduced |
| | | LTP [35] | Binary into ternary |
| 3 | Template matching | AAM [37] | AAM parameters for classification learning |
| 4 | Part-based | Component-base [38] | Comparing global and component representation, while a SVM classifier for each person is not suitable in practice. |

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

| | | | |
|---|---|---|---|
| | | SIFT [40] | Using SIFT feature with spatial constraints to compare faces |

*Table 2: The summary of face recognition techniques*

The scale-invariant feature transform (SIFT) proposed by Lowe et.al [39] has been widely and successfully applied to object detection and recognition. In the works of Luo et al. [40], they proposed to use the person-specific SIFT features and a simple non-statistical matching strategy combined with local and global similarity on key-point clusters to solve face recognition problems.

## III. CONCLUSION

In this paper, we will give summaries different types of face detection and face recognition techniques during the past two decades. PCA and LDA are two principle algorithms used for decreasing data dimension. PCA processes data in a way to minimize the noise and redundancy and is thus widely used in data compression, while LDA aims to maximize the distance between different classes and thus obtain good reputation in pattern classification. We have identified two key problems for any face recognition systems one is illumination problem and second is pose problem. Finally, there is still no robust face detection and recognition technique for unconstraint real-world applications.

## REFERENCES

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, “Human and machine recognition of faces: a survey,” Proc. IEEE, vol. 83, no. 5, pp. 705-740, 1995.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face recognition: a literature survey,” Tech-nical Report CAR-TR-948, Center for Automation Research, University of Maryland (2002).

[3] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, “2D and 3D face recognition: a survey,” Pattern Recognition Letter, vol. 28, no. 14, pp. 1885–1906, 2007.

[4] M. Grgic, and K, Delac, “Face recognition homepage.” [online] Available: http://www.face-rec.org/general-info. [Accessed May. 27, 2010].

[5] A. K. Jain, R. P. W. Duin, and J. C. Mao, “Statistical pattern recognition: a review,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4–37, 2000.

[6]M. H. Yang, D. J. Kriegman, and N. Ahuja, “Detecting face in images: a survey,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, pp. 34–58, 2002.

[7] G. Yang and T. S. Huang, “Human face detection in complex background,” Pattern Recognition Let-ter, vol. 27, no. 1, pp. 53-63, 1994.

[8] C. Kotropoulos and I. Pitas, “Rule-based face detection in frontal views,” Proc. Int’l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.

[9] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, ”Face detection in color images,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, 696–706, 2002.

[10] T. K. Leung, M. C. Burl, and P. Perona, “Finding faces in cluttered scenes using random labeled graph matching,” Proc. Fifth IEEE Int’l Conf. Computer Vision, pp. 637-644, 1995.

[11] S. Theodoridis and K. Koutroumbas, Pattern recognition, 4th ed., Academic Press, 2009.

[12] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” Int. Journal of Computer Vision, vol. 1, pp. 321–331, 1988.

[13] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, “Active shape models - their training and application,” Computer Vision and Image Understanding, vol. 61, no. 1, pp. 38-59, 1995.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

[14] R. O. Duda, P. E. Hart, D. G. Stoke, Pattern classification, 2nd ed., John Wiley & Sons, 2001.

[15] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed., Springer, 2005.

[16] K. K. Sung and T. Poggio, “”Example-based learning for view-based human face detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no.1, 39–51, 1998.

[17] E. Alpaydin, Introduction to machine learning, 2nd ed., The MIT Press, 2010.

[18] C. M. Bishop, Neural networks for pattern recognition, Oxford University Press, 1995.

[19] C. M. Bishop, Pattern recognition and machine learning, Springer, 2006.

[20] R. Fergus, P. Perona, and A. Zisserman, “Object class recognition by unsupervised scale-invariant learning,” IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 264–271, 2003.

[21] B. Heisele, T. Serre, M. Pontil, and T. Poggio, “Component-based face detection,” IEEE Conf. Computer Vision and Pattern Recognition, pp. 657–662, 2001.

[22] M. Turk and A. Pentland, “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, vol. 3, no.1, pp. 72-86, 1991.

[23] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, 711–720, 1997.

[24] J. Yang D. Zhang, A. F. Frangi, and J. Y. Yang, “Two-dimensional PCA: A new approach to appear-ance-based face representation and recognition,” IEEE Trans. Pattern Analysis and Machine Intelli-gence, vol. 26, no. 1, 131–137, 2004.

[25] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face recognition by independent component analysis,” IEEE Trans. Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.

[26] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, “Face recognition using Laplacianfaces,” IEEE Trans.Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 328-340, 2005.

[27] M. H. Yang, ““Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using kernel methods,” AFGR, pp. 205–211, 2002.

[28] C. Liu and H. Wechsler, “Evolutionary pursuit and its application to face recognition,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 6, pp. 570–582, 2000.

[29] X. He and P. Niyogi, “Locality Preserving Projections,” Proc. Conf. Advances in Neural Information Processing Systems, 2003.

[30] J. Wright, A. Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 2, 210-227, 2009.

[31] M. Lades, J. C. Vorbriiggen, J. Buhmann, J. Lange,C. von der Malsburg, R. P. Wiirtz, and W. Konen, “Distortion invariant object recognition in the dynamic link architecture,” IEEE Trans. on Computers, vol. 42, no. 3, pp. 300-311, 1993.

[32] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg, “Face recognition by elastic bunch graph matching,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no.7, pp. 775-779, 1997.

[33]T. Ahonen, A. Hadid, and M.Pietik¨ainen, “Face recognition with local binary patterns,” Proc. Eighth European Conf. Computer Vision, pp. 469-481, 2004.

[34] T. Ojala , M. Pietikainen, D. Harwood, "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, vol. 29, pp. 51–59, 1996.

[35] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult light-ing conditions," In: Zhou, S.K., et al. (eds.) AMFG 2007. LNCS, vol. 4778, pp. 168-182. Springer, Hei-delberg, 2007.

[36] T. Ojala , M. Pietikainen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, pp. 971–987, 2002.

[37] G.J. Edwards, T.F. Cootes, and C.J. Taylor, "Face recognition using active appearance models," Proc. European Conf. Computer Vision, vol. 2, pp. 581-695, 1998.

[38] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face Recognition: Component-based versus global ap-proaches," Computer Vision and Image Understanding, vol. 91, nos. 1-2, pp. 6-21, 2003.

[39] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int'l Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[40] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B. L. Lu, "Person-Specific SIFT Features for Face Recognition," Int'l Conference on Acoustic, Speech, and Signal Processing, 2007.

 