

Detecting and Recognising Faces in Images

Lei Liu, 9669373

May 18, 2016

1 Introduction

A human can identify individuals' face based on his/her own knowledge and memory once been taught. Even a person can be recognised in a few second despite the great change in appearance due to the variant of surroundings, a view of parts of a face, ageing, facial expression and hair style. In the field of artificial intelligence and machine vision, therefore, the goal of a face recognition system is to give a machine the similar capability that identify a face from a scene in a robust and rapid way. More specifically, given an input image, the system is able to locate the regions that contain a face, and then determine the identification of each extracted face. This derives the two main tasks for researching the face information in an image, which are face detection that find the position of a face form a digital image, and face recognition that recognise a specific individual by comparing the face image with per-defined database.

A face recognition system could be widely employed on practical applications, such as criminal identification, security system and human-interaction application especially the video games. However, it is a challenge for the computer to perform such a complicated task since the complexity of facial structure and its variance as well [TP91]. Besides, the difference of conditions for capturing photos would also affect the intensity of image dramatically, thus, affect the performance of face detection and recognition. In the following section, I will introduce several works that handled either the task of face detection or face recognition.

2 literature review

The basic framework for face detection and recognition is that: for any input image, the detector or recognizer would extract critical features and project

them on the pre-trained feature space (could be feature coordinates or pre-defined template), then assign the most similar label. Here, I will discuss six papers related to the face detection and recognition problems, which are based on machine learning algorithm and information theory. Generally, four of them could be categorized into eigenvalue-based approaches, while the other are non-eigenvalue approaches. The eigenvalue-based approaches basically applied the eigenvalue decomposition either for principal component analysis (PCA), or combined with other theory such as probability estimate, kernel function and multilinear analysis.

2.1 Eigenface

Firstly, the pure principal component analysis for face recognition was proposed in 1991 [TP91]. This method defined a face space consisted of M eigenvectors that have the largest eigenvalues which is computed from a set of training images. Then, for each known face, its coordinates in face space can be trained. When testing on a new image, the distances between its coordinates in the same face space and both the trained face coordinates and face space could be produced. The final result (non-face, unknown face or face label) depends on the manually selected thresholds. This method is sufficient in grouping similar faces as a unsupervised learning algorithm, and perform well on constant environment once the eigenface space has been defined. Although this method is sensitive to the lighting variation, varying in scale and head orientation as well as different face views, it provides a fast and simple method for extract the critical feature from high-dimensional image space.

2.2 Probability Estimation

As we know, the PCA is widely used for finding the essential subspace that can be used to encode the facial feature. Moghaddam et al. [MP97] introduced a probabilistic formulation that using density estimation function to estimate the likelihoods of a particular object represented in eigenspace. Basically, it is a template matching algorithm that for each object class, estimate the class samples (data have been represented by selected principal components) and generate a likelihood function for that data. Therefore, for every testing input, a maximum likelihood detection can be applied to find the best match. In the application of face detection and recognition, it performed well on locating the facial features including the left and right eyes, the nose and the centre of mouth [MP97] despite the variant of position and scale, but still the performance were affected by lighting condition, and head

position. Compared with eigenface, this method performed better on solving the problem of varying position and scale, but decrease computational efficiency as distance calculation and template matching are needed.

2.3 Kernel PCA

Another way to understand PCA is that PCA performs a linear transformation of coordinates of the input space. It assumes that the orthogonal coordinates are linear combined. While, a method proposed by Schölkopf et al. [SSM98] is able to handle the nonlinear component analysis problem by introducing a kernel function. Specifically, when mapping the input space into higher dimensional space (since nonlinear problem could become a linear analysis problem after projecting into high-dimensional space), it is extremely difficult to calculate the nonlinear mapping function. While, this mapping function could be approximated by a kernel function, so as to be able to perform the linear PCA in high-dimensional space. This method is an extension of the standard PCA, thus, can be applied to any situation that using a standard PCA, and performs better than the linear PCA in recognition problem [SSM98].

2.4 TensorFace

A conventional PCA could perform well on the dataset with single factor (including lighting, pose, facial view and expression) variations. In practical, all these factors could be varied at same time which place a challenge for face recognition system. Vasilescu et al. [VT02] proposed the TensorFaces which is a multifactor structure that considering all the above factors simultaneously. The TensorFaces is constructed by multilinear analysis which combines numerous conventional PCA matrix with respect to different factors. In other words, the principal components for TensorFaces are computed from higher-order PCA and SVD matrix. The advantage of this approach is the consideration of several factors at the same time, thus, could be robust under several kinds of variations. However, it would be strict to build a suitable face dataset as each person should be captured under different conditions and the computational cost would increase compared with conventional PCA.

2.5 FisherFace

Different from eigenface, the FisherFace for face recognition, described by Belhumeur et al [BHK97], is a method that employed both the within class

and between class information and produce a projection coordinate based on linear discriminant analysis. This method is similar to eigenface. However, rather than maximizing the covariance of data, this tries to maximize the ratio of the between class scatter by the within class scatter [BHK97]. Consequently, the projection has the property that distances between different classes are maximum, while that within a same class is minimum. However, one of the problem is that it is difficult to find an optimum projection which is able to separate multiple face classes at the same time. When compared with eigenface in face recognition, FisherFace performed better under variation of lighting and expression conditions.

2.6 Rapid Object Detection

Viola et al. [VJ01] introduced a novel framework for object detection, and built a frontal face detection system with high accuracy and real-time performance. Basically, the framework consists of three main components. Firstly, a new image representation is extracted based on the Haar-like features which are related to the ad-hoc domain knowledge and are more meaningful than a single pixel. Secondly, from these extracted features, an AdaBoost algorithm is applied to select an important small set of features so as to build a built a feature classifier with strong boundaries and reduce the searching space of features. The last component is a cascade classifier that combines numerous complex classifiers sequentially. The former classifiers are simple and with weak boundary that eliminate a major of negative sub-windows, while the following classifiers are complex and have high accuracy that are able to focus on much more critical area and output results with high confidence. The most important advantage of this framework is finding the trade-off between minimum computation time and high accuracy. However, this approach is limited in frontal face detection as it employs the Haar-like features which is affected by the variations of lightness, pose and scale.

3 Comparison

Here a discussion of the above six approaches will be given. The eigenface is the simplest approach for face detection and recognition as PCA is an unsupervised learning algorithm that solves the only eigenvalue problem and can perform in real-time application. However, its performance is limited under varying conditions of lightness, facial expression, head pose and facial view. While, the extension of PCA like, Probability Estimation (section 2.2) Kernel PCA (section 2.3), TensorFace (section 2.4), and FisherFace (section 2.5)

all performed better with respect to different variations, though increased the computational time and searching space due to the extra computation when introducing other theories. The Rapid Object Detection (section 2.6) is a different approach that employed the Haar-like features and AdaBoost algorithm to reduce the computation and searching space dramatically, but depended heavily on the features which are sensitive to the above variations. With respect to lighting variation, the eigenface could failed since the first principal component could be mainly related to lighting condition. While the FisherFace would successful, as well as the TensorFaces. With variation in both lighting and facial expression, the Fisherface could perform better than the others. The rapid object detection could fail under variation of head pose and facial viewing as the feature would change dramatically in this condition. The kernel PCA would work successful on those nonlinear separable dataset.

4 Conclusion

In conclusion, these methods have their own strengths under different conditions. Thus, the choice of algorithm depends on the requirement of application. It is difficult to find a best solution that could handle all kinds of variations and ensure computational efficiency as well. However, for task of face detection, I would choose the Rapid Object Detection algorithm as it works in real-time and have high confidence for decision. While, for face recognition, the TensorFace could be the best among these methods as this one could perform well under variations of several factors at the same time.

References

- [BHK97] 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*, 19(7):711–720, Jul 1997.
- [MP97] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696–710, Jul 1997.
- [SSM98] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998.

- [TP91] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.
- [VJ01] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.
- [VT02] M Alex O Vasilescu and Demetri Terzopoulos. Multilinear analysis of image ensembles: Tensorfaces. In *Computer Vision ECCV 2002*, pages 447–460. Springer, 2002.