INSTIGATION OF FACE RECOGNITION

How face recognition evolved through the years

Vipin Benny

IMCA

SCMS School of Technology and Management

Kochi, India

vipinbenny31@gmail.com

ABSTRACT: This paper talks about how face recognition has changed over the years to be what it today, face recognition today is part of all secure system. All the current mobile devices today come with this feature. The main context of this paper is to evaluate the evolution of this biometric security over the years and understand the changes in the system, how it may change and the possibility of its further evolution in the next decade.

INTRODUTION

The idea of ‘Face recognition’ dates back to 1960’s where the most prominent development was the us e of IC chips. IBM used IC chips in its “computers”. The first implementation of facial recognition was in 1964, pioneers of this project include Woody Bledsoe, Helen Chan Wolf and Charles Bission.Its during the years 1964 and 1965 the three worked on using computers to recognize human faces. This project later paved the way for thedevelopment of the feature. In 1966 the project held the disadvantage of large nee. The initial approach of the three pioneers was much more physical than technical, Bledsoe marked various landmarks on the face such as the eye centers, mouth, and these were mathematically rotated by computer to compensate for pose variation. The distance between the landmark was also computed and compared to determine the identity.

Further details of this project were held back due it being the private funded, with the references available the main problem faced by them was the amount of data needed. A variation in the slightest would make the recognition much more, in the words of Bledsoe:

“*This recognition problem is made difficult by the great variability in head rotation and tilt, lighting intensity and angle, facial expression, aging, etc. Some other attempts at face recognition by machine have allowed for little or no variability in these quantities. Yet the method of correlation (or pattern matching) of unprocessed optical data, which is often used by some researchers, is certain to fail in cases where the variability is great. In particular, the correlation is*

*very low between two pictures of the same person with two different head rotations*”- Woody Bledsoe, 1966

In about 1996 the work was continued by Peter Hart who enthusiastically recalled the project with the exclamation “it really worked”.

Looking through the evolution of face recognition is denser as its the evolution of more than software, algorithms, face recognition can also be considered to change many fields,3D technology is one of the numerous fields that has an influence in the technology.

Up till the 1980’s face recognition needed more manual work than a complete program, it was decades of technological advancement which makes facial recognition what it is today. The major advancement in the 1970 was by Goldstein, Harmon and Lesk who were able to increase the accuracy of the manual recognition system. As with Bledsoe’s system, the actual biometrics had to still be manually computed.

In the late 1980’s the question regarding where complete automation of face recognition stated and as a result in 1988 Sirovich and Kiriby began applying linear algebra to the problem of facial recognition. Sirovich and Kriby were able to show that feature analysis on a collection of facial images could form a set of basic features. They were also able to show that less than one hundred values were required in order to accurately code a normalized face image. This is a major breakthrough which helped form the Eigenfaces, which can be considered as one of the five different face recognition algorithms.

In 1987 the very first algorithm of face recognition was introduced called EIGENFACES. It was a revolutionary change in facial recognition. It can be considered as the first accepted face recognition algorithm in the history.

I.EIGEN FACES

Eigen faces is an appearance-based approach to face recognition. It’s a recognition method developed by Sirovich and Kirby in 1987 and used by Matthew Turk and

Alex Pentland in face classification. The name comes the set of eigenvectors used in the system. An eigenvector is a nonzero vector that changes at most by a scalar factor when that linear transformation is applied to it. Eigen faces can be considered as the first working face recognition by many and served as basis for one of the top commercial face recognition technology products. The motivation of Eigen faces can be concluded into two main conclusions:

* Extract the relevant facial information, which may or may not be directly related to human intuition of face features such as the eyes, nose, and lips. One way to do so is to capture the statistical variation between face images.
* Represent face images efficiently. To reduce the computation and space complexity, each face image can be represented using a small number of parameters.

Computing of Eigenfaces is done by using Principle Component Analysis (PCA), but before proceeding to generate Eigenfaces, face images are normalized to line up eyes and mouth, then all resampled at the same pixel resolution. Extraction of Eigenface by PCA is as follow:

1.Given M face images with the size of h×w, each image is transformed into a vector of size D(=hw) and placed into the set

{Γ1,Γ2,⋯,ΓM}

The face images should be appropriately scaled and aligned, and the backgrounds (and possibly non-face areas such as hair and neck) should be constant or removed

2.Each face differs from the average by the vector Φi=Γi−Ψ, where the average face is defined by

Ψ=

3. The covariance matrix **C** ∈ is defined as

where **A**={ Φ 1,Φ2,⋯,Φ*M*}∈

4. Determining the eigenvectors of **C** is an intractable task for typical image sizes when *D*≫*M*. However, to efficiently compute the eigenvectors of C, one may first compute the eigenvectors of the much-smaller M×M matrix **A**. The eigenvector and eigenvalue matrices of **A** are defined as

**V**= {v1, v2,⋯vr}

and Λ=*diag* {λ1,λ2,⋯λr} , λ1≥λ2≥⋯λr>0 , where is the rank of **A** . Note that eigenvectors corresponding to eigenvalues of

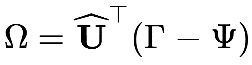
zero have been discarded.

5.The eigenvalue and eigenvector matrices of **C** are Λ and **U**=**AV**Λ−1/2, where **U**={**u***i*} is the collection of eigenfaces.

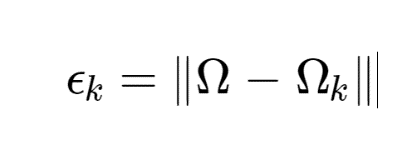
There are mainly two steps in the processing of face in the system, face detection and face recognition:

Face detection is the first step in the process, it ca n be considered as detecting image patches that lie close to the face space. The distance between the face image and its projection can be computed as

Where **I** is the identity matrix.

On the other hand, in face recognition a face Γ is projected into the face space by 

where Uˆ is the set of significant eigenvectors. Note that the weight vector Ω is the representation of the new face in face space. One simple way to determine which face class Γ belongs to is minimizing the Euclidean distance



Where is weight vector representing the kth fac. The face Γ is considered as belonging to class k if the minimum is smaller than some predefined threshold otherwise, it is classified as unknown.

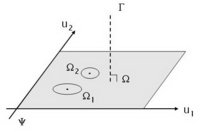


Fig: Visualization of a 2D face space, with axes representing two Eigenface.

The Eigenface was an important step in facial recognition and overtime this was modified to face challenges in the reality, that is even when eigen face can identify and recognize faces with the technology of the time it could never produce accurate reading if there is change in the facial feature. Eigen face was very sensitive to lighting, scale and translation and required highly controlled environment. Any change in expression can also cause problem in the recognition. It was at the end of the 90’s that the next notable face recognition system can to light.

II.LBPH

LBPH stands for Local Binary Histograms, it was introduced at 1996, the entire concept of LBPH revolves around Local Binary Patterns (LBP). LBP is a type of visual descriptor used for the classification in computer vision, an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. LBP was first described in 1994, it has since been found to be a powerful feature for texture calculation. Using LBP combined with histogram face image can be represented with simple data vector.The following is the method by which LBPH identify and recognize faces

1.Checking parameters, parameters of LBPH include

* Radius: the radius is used to build the circular local binary pattern and represents the radius around the central pixel. It is usually set to 1.
* Neighbors: the number of sample points to build the circular local binary pattern. Keep in mind: the more sample points you include, the higher the computational cost. It is usually set to 8.
* Grid X: the number of cells in the horizontal direction. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8.
* Grid Y: the number of cells in the vertical direction. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector. It is usually set to 8.

2.Training the algorithm: using data set with the facial image of the person that is to be recognized, also an ID for each image so that the algorithm will use the information to recognize an input image and gives a output. It is to be noted to put the images of the same person under a single ID.

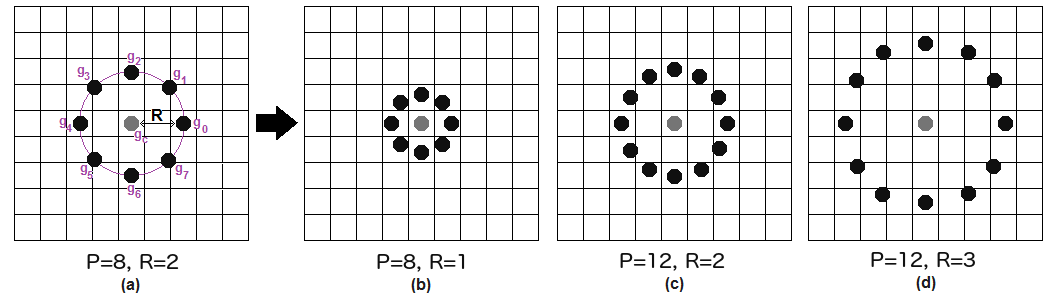
3. Applying the LBP operation: The first computational step of the LBPH is to create an intermediate image that describes the original image in a better way, by highlighting the facial characteristics. To do so, the algorithm uses a concept of a sliding window, based on the parameters radius and neighbors. The figure below provide the example of this procedure



Based on the image above, let’s break it into several small steps so we can understand it easily:

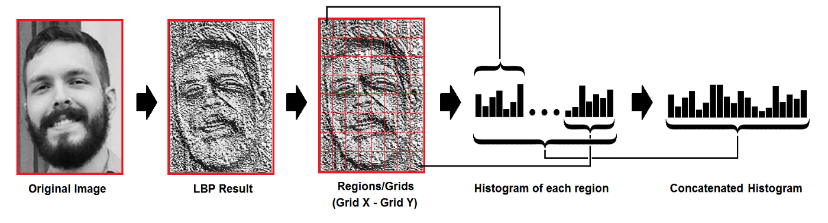
* Suppose we have a facial image in grayscale.
* We can get part of this image as a window of 3x3 pixels.
* It can also be represented as a 3x3 matrix containing the intensity of each pixel (0~255).
* Then, we need to take the central value of the matrix to be used as the threshold.
* This value will be used to define the new values from the 8 neighbors.
* For each neighbor of the central value (threshold), we set a new binary value. We set 1 for values equal or higher than the threshold and 0 for values lower than the threshold.
* Now, the matrix will contain only binary values (ignoring the central value). We need to concatenate each binary value from each position from the matrix line by line into a new binary value (e.g. 10001101). Note: some authors use other approaches to concatenate the binary values (e.g. clockwise direction), but the final result will be the same.
* Then, we convert this binary value to a decimal value and set it to the central value of the matrix, which is actually a pixel from the original image.
* At the end of this procedure (LBP procedure), we have a new image which represents better the characteristics of the original image.

Note: The LBP procedure was expanded to use a different number of radius and neighbors, it is called Circular LBP.



It can be done by using bilinear interpolation. If some data point is between the pixels, it uses the values from the 4 nearest pixels (2x2) to estimate the value of the new data point.

4. Extracting the Histograms: Now, using the image generated in the last step, we can use the Grid X and Grid Y parameters to divide the image into multiple grids, as can be seen in the following image:



Based on the image above, we can extract the histogram of each region as follows:

* As we have an image in grayscale, each histogram (from each grid) will contain only 256 positions (0~255) representing the occurrences of each pixel intensity.
* Then, we need to concatenate each histogram to create a new and bigger histogram. Supposing we have 8x8 grids, we will have 8x8x256=16.384 positions in the final histogram. The final histogram represents the characteristics of the image original image.

5. Performing the face recognition: In this step, the algorithm is already trained. Each histogram created is used to represent each image from the training dataset. So, given an input image, we perform the steps again for this new image and creates a histogram which represents the image.

* So, to find the image that matches the input image we just need to compare two histograms and return the image with the closest histogram.
* We can use various approaches to compare the histograms (calculate the distance between two histograms), for example: Euclidean distance, chi-square, absolute value, etc. In this example, we can use the Euclidean distance (which is quite known) based on the following formula:
* So the algorithm output is the ID from the image with the closest histogram. The algorithm should also return the calculated distance, which can be used as a ‘confidence’ measurement. Note: don’t be fooled about the ‘confidence’ name, as lower confidences are better because it means the distance between the two histograms is closer.
* We can then use a threshold and the ‘confidence’ to automatically estimate if the algorithm has correctly recognized the image. We can assume that the algorithm has successfully recognized if the confidence is lower than the threshold defined.

LBPH is hence one of the easiest face recognition algorithms, it can represent local features in the image itself. After the LBPH the next face recognition system to be used is Fisherfaces which was introduced in 1997.

III.FISHERFACES

Fisherface can be considered to be an improvement of the

Eigenface algorithm and it also uses the principal of PCA, Principal Component Analysis and LDA, Liner Discriminant Analysis. The general steps involved in this face recognition is are

* Capturing
* Feature Extraction
* Comparison
* Match/Non-match

We know that the major flaw in Eigen face is that it considers illumination as one of the important features of the face, but it actually isn’t. Fisherface algorithm extracts principal components that separates one individual from another. Image recognition using this algorithm is based on reduction of face space dimensions using PCA method and then applying LDA method also known as Fisher Linear Discriminant (FDL) method to obtain characteristic features of image.

LDA is used to find a linear combination of features that separates two or more classes or objects. It can be used for dimension reduction before further classification. It attempts to model the difference between classes of data .

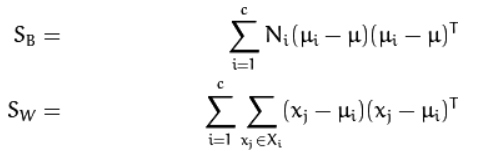
* Data is assumed to be uniformly distributed in each class.
* Aim is to maximize the ratio of between-class scatter matrix and the within-class scatter matrix.
* It can produce good results even in varying illumination.

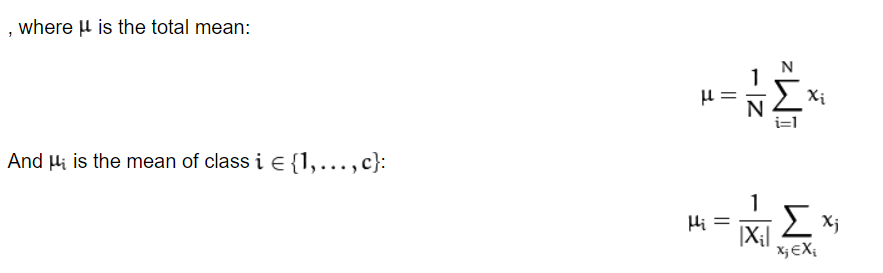
The algorithm of Fisherface is given below

Let X be a random vector with samples drawn from c classes:

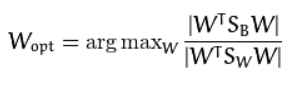


The scatter matrices S\_{B} and S\_{W} are calculated as:





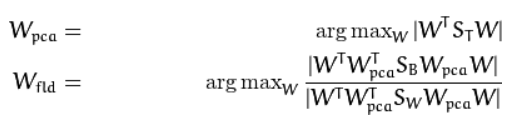
Fisher’s classic algorithm now looks for a projection W, that maximizes the class separability criterion:



A solution for this optimization problem is given by solving the General Eigenvalue Problem:



The optimization problem can then be rewritten as:



The transformation matrix W, that projects a sample into the (c-1)-dimensional space is then given by:

tm

The fisherface is immune to noise-induced images and blurring effects on the image, it’s the features of the image that dominate the other persons feature in Fiserface

* By using Fisherfaces we can prevent features of an individual from being dominant but it still considers illumination an important feature. But we know that illumination is not an important feature as it's not even a part of face.
* PCA dimension reduction process can cause some loss of discriminant information useful in the LDA process which is a disadvantage of this process.
* The problem of computation in face recognition using fisherface method is because the computation process is very complicated and complex.

IV.SIFT

Scale Invariant Feature Transform (SIFT) is an image descriptor for image-based matching and recognition developed by David Lowe (1999,2004). This descriptor as well as related image descriptors are used for a large number of purposes in computer vision related to point matching between different views of a 3-D scene and view-based object recognition. The SIFT descriptor is invariant to translations, rotations and scaling transformations in the image domain and robust to moderate perspective transformations and illumination variations. Experimentally, the SIFT descriptor has been proven to be very useful in practice for image matching and object recognition under real-world conditions.

In its original formulation, the SIFT descriptor comprised a method for detecting interest points from a grey-level image at which statistics of local gradient directions of image intensities were accumulated to give a summarizing description of the local image structures in a local neighborhood around each interest point, with the intention that this descriptor should be used for matching corresponding interest points between different images. Later, the SIFT descriptor has also been applied at dense grids (dense SIFT) which have been shown to lead to better performance for tasks such as object categorization, texture classification, image alignment and biometrics. The SIFT descriptor has also been extended from grey-level to colour images and from 2-D spatial images to 2+1-D spatio-temporal video.

The original SIFT descriptor (Lowe 1999, 2004) was computed from the image intensities around interesting locations in the image domain which can be referred to as interest points, alternatively key points. These interest points are obtained from scale-space extrema of differences-of-Gaussians (DoG) within a difference-of-Gaussians pyramid. The concept of difference-of-Gaussian bandpass pyramids was originally proposed by Burt and Adelson (1983) and by Crowley and Stern (1984).

A Gaussian pyramid is constructed from the input image by repeated smoothing and subsampling, and a difference-of-Gaussians pyramid is computed from the differences between the adjacent levels in the Gaussian pyramid. Then, interest points are obtained from the points at which the difference-of-Gaussians values assume extrema with respect to both the spatial coordinates in the image domain and the scale level in the pyramid.

V.SURF

In computer vision, speeded up robust features (SURF) is a patented local feature detector and descriptor. It can be used for tasks such as object recognition, image registration, classification, or 3D reconstruction. It is partly inspired by the scale-invariant feature transform (SIFT) descriptor. The standard version of SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT.

SURF descriptors have been used to locate and recognize objects, people or faces, to reconstruct 3D scenes, to track objects and to extract points of interest. SURF was first published by Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, and presented at the 2006 European Conference on Computer Vision. An application of the algorithm is patented in the United States. An "upright" version of SURF (called U-SURF) is not invariant to image rotation and therefore faster to compute and better suited for application where the camera remains more or less horizontal.

SURF has three main parts in its algorithm they are interest point detection, local neighborhood description, and matching.

Detection: SURF uses square-shaped filters as an approximation of Gaussian smoothing. (The SIFT approach uses cascaded filters to detect scale-invariant characteristic points, where the difference of Gaussians (DoG) is calculated on rescaled images progressively.) Filtering the image with a square is much faster if the integral image is used

Descriptor: The goal of a descriptor is to provide a unique and robust description of an image feature, e.g., by describing the intensity distribution of the pixels within the neighborhood of the point of interest. Most descriptors are thus computed in a local manner; hence a description is obtained for every point of interest identified previously.

The dimensionality of the descriptor has direct impact on both its computational complexity and point-matching robustness/accuracy. A short descriptor may be more robust against appearance variations, but may not offer sufficient discrimination and thus give too many false positives.

The first step consists of fixing a reproducible orientation based on information from a circular region around the interest point. Then we construct a square region aligned to the selected orientation, and extract the SURF descriptor from it.

Matching: By comparing the descriptors obtained from different images, matching pairs can be found.

VI.SOCIAL MEDIA

In 2010, Facebook began implementing facial recognition functionality that helped identify people whose faces may be featured in the photos that Facebook users update daily. While the feature was instantly controversial with the news media, sparking a slew of privacy-related articles, Facebook users at large did not seem to mind. Having no apparent negative impact on the website’s usage or popularity, more than 350 million photos are uploaded and tagged using face recognition each day. Although there are error in recognition some time its pretty much became better as time went in especially after 2017, which changed the way face recognition was looked as a biometric security.

It was after this introduction that in 2011, face recognition was introduce in airport. This was done by the program Facefirst, shortly after implementation, the system resulted in the apprehension of multiple Interpol suspects. Another major use the facial recognition is the confirmation of identity using facial recognition in fact facial recognition was used to help confirm the identity of Osama bin Laden after he was killed in a U.S. raid.

VII. DOT-PROJECTION

Dot projection can be considered as the latest improvement in the face recognition, dot projection has high similarity between the first discovery of face recognition, we can see that just like when the first face recognition was introduced that latest both relay on a machine. Dot-projector is introduced by apply in its product iPhone x, in 2017 calling it faceID. The interesting fact in this is that using dot projector the faceID can recognize features of the face even when there is change in face like growing a beard, trimming, putting on glasses. The Dot projector in the iPhone X is designed using high intensity LEDs that throws a dotted light pattern of 173 X 173 dots on the face. So, the dot projector is basically a structured light source that has one or two flash component which emits a predetermined light pattern on the face, and all these happen in few milliseconds. The best part of dot projector is that it is a lot different than power-hungry laser beams which is traditionally used for 3D imagery. It not only consumes a low amount of power but also offers better 3D depth than traditional systems.

Every face comes with a unique facial structure, so it becomes easy for the dot projector to create an exclusive 3D facial map. So, every time you use the Face ID, the infrared camera captures the dot structure, and the neural engines match it with the original map.

The facial 3D map stays within the phone, and they are not uploaded in the iCloud as there are chances of a security breach. So, the use of dot projector in the facial recognition system of iPhone X is simple but highly effective.

The figure above shows dots projected toward the face of the user. Using this accurate information FaceID is able to recognize the user with near total accuracy, making it one of the best technological advancement in facial recognition.

CONCLUSION

Hence, we can see that over time the invention of more accurate facial recognition and different approach to the facial recognition. Its to be noted that with the path its following we can say that in about a decade we can recognize individuals without any equipment and also at high accuracy like the FaceID, the face recognition which was a idea of 1960 has influenced a lot of change in the recent past. FaceID is used as a secure means of even bank transactions, this was a huge change considering that face recognition was not considered as a secure biometric method, it was after 2011, using FaceID that face recognition played an important role int the security. Nowadays face recognition has become an important security method in the smartphone industry and high recognition method are used to decrease any fault in the system.

REFERENCES

[wikipedia.org](http://www.wikipedia.org)

[facefirst.com](http://www.facefirst.com)

[scholarpedia.org](http://www.scholarpedia.org)

[towardsdatascience.com](http://www.towardsdatascience.com)

[openbio.sourceforge.net](http://openbio.sourceforge.net/)