FACE RECOGNITION BASED AUTHENTICATION SYSTEM FOR DATA CENTER

Submitted By,

Amit H. Vallecha (202004104610016),

Akhilesh K. Solanki (202004104610017),

Mrugesh R. Modi (202004104610018)

Guided By,

Mr. Jitendra B. Upadhyay,

Dr. Kalpesh B. Lad

for partial fulfillment of the requirements

for the Degree of Master of Computer Applications

Shrimad Rajchandra Institute of Management and Computer Application

Uka Tarsadia University.

May, 2021.

Introduction

Maintaining security in an organization specifically in the workplaces like examination hall in the education sector, lockers room in banking places, server room or control room of an organization, is the dominant thing and keeping the environment free of wrongdoing and also one of them is the data-center.

A data center is basically a physical facility that organizations use to house their critical applications and data. It holds proprietary and sensitive information such as intellectual property, customer data, and financial records, and therefore needs to be both physically and digitally secure. Traditional ways are failing to tackle modern challenges of efficiency, security. There is an immense need for reform and leading technology to protect several areas from unauthorized users.

Authentication may be defined as “providing the right person with the right privileges the right access at the right time.” In general, there are three approaches to authentication. In order of least secure and least convenient to most secure and most convenient, they are Something you have - card, token, key. Something you know- PIN, password. Something you are- a biometric.

The most common person authentication methods are passwords, PINs, smart keys, smart cards, but have limitations and vulnerabilities. For instance, passwords and PINs can be easily forgotten, hard to remember, or stolen. Smart keys can be easily lost or duplicated. Smart cards with magnetic strips can be forged. But a person’s biometrics or biological traits cannot be stolen, forgotten, or misplaced. As a result, they provide a much more secure and reliable way to authenticate an individual when compared to the traditional methods.

Biometrics is the measurement and statistical analysis of people's unique physical and behavioral characteristics. The term *biometrics* is derived from the Greek word *bio*, meaning *life*, and *metric*, meaning *to measure*. Every human being possesses certain unique features in terms of both physical and behavioral characteristics that are different from everybody else on the planet. Based on that several types of biometric exist like Fingerprint, voice recognition, facial recognition, iris recognition, retina scan, etc. Some confusion still exists about which types of biometrics organizations should use. Based on the Organization's needs we can use one of them.

The proposed solution implies resolving complications by facial recognition. Facial recognition is an easy task for humans. Experiments have shown [1], that even one to three-day-old babies can distinguish between known faces but the same process for the computer to do the same task is difficult. How computer read the images? Basically, the computer interpreter the image as a combination of pixels. Each pixel contains a different number of channels. Pixels are organized in an ordered rectangular array. The size of an image is determined by the dimensions of this pixel array called Digital Image. When Processing the images there may be several problems that exist with images like blurred images, illumination, etc. Here we need Digital image processing techniques to process the images.

Digital image processing focuses on two major tasks–Improvement of pictorial information and Processing of image data for storage, transmission and representation for autonomous machine perception [2].

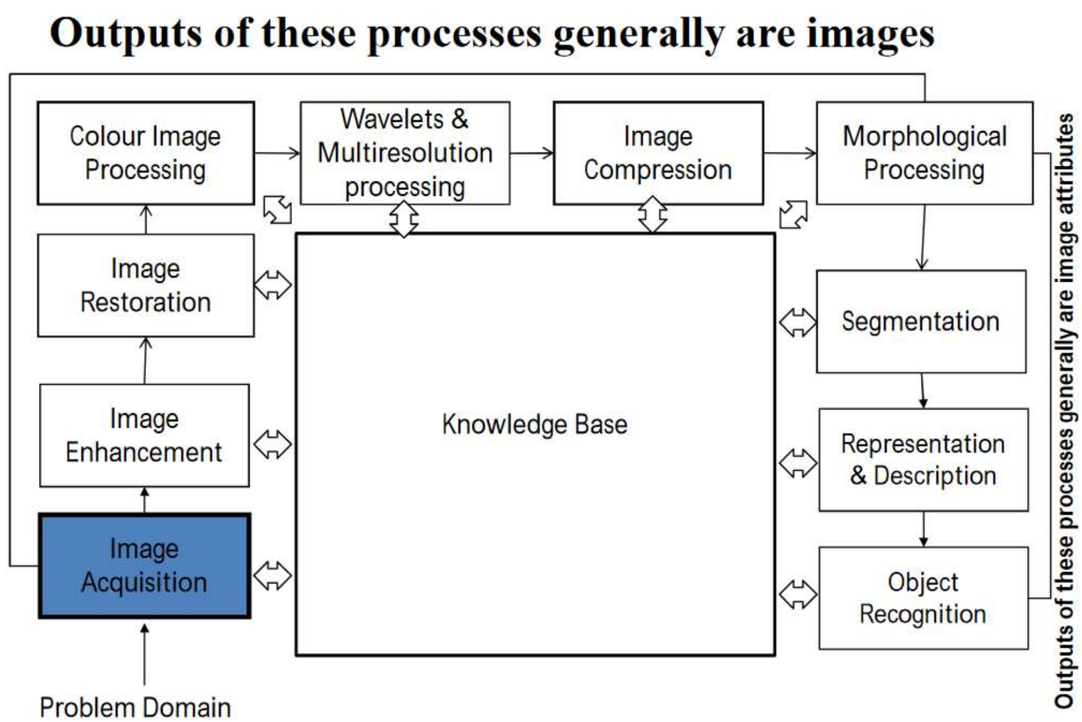


Figure -1: Fundamental steps in digital image processing [2]

Consider Figure 1 that shows the fundamental steps in digital image processing. It starts with Image Acquisition, in that the image is captured by a sensor (e.g., Camera), and digitized if the output of the camera or sensor is not already in digital form, using an analog-to-digital convertor. Image Enhancement is the process of manipulating an image so that the result is more suitable than the original. The idea behind enhancement techniques is to bring out hidden details, or simply to highlight certain features of interest in an image. Image Restoration, for Improving the appearance of an image. Color Image Processing uses the color of the image to extract features of interest in an image. Wavelets Are the foundation of representing images in various degrees of resolution. It is used for image data compression. Compression Techniques for reducing the storage required to save an image or the bandwidth required to transmit it. Morphological Processing for extracting image components that are useful in the representation and description of shape. In this step, there would be a transition from processes that output images, to processes, that output image attributes. Image Segmentation procedures partition an image into its constituent parts or objects. The more accurate the segmentation, the more likely recognition is to succeed. Representation Make a decision whether the data should be represented as a boundary or as a complete region. It almost always follows the output of a segmentation stage. The description also called, feature selection, deals with extracting attributes that result in some information of interest. Recognition is the process that assigns a label to an object based on the information provided by its description. Knowledge about a problem domain is coded into an image processing system in the form of a knowledge database [2].

System Flow

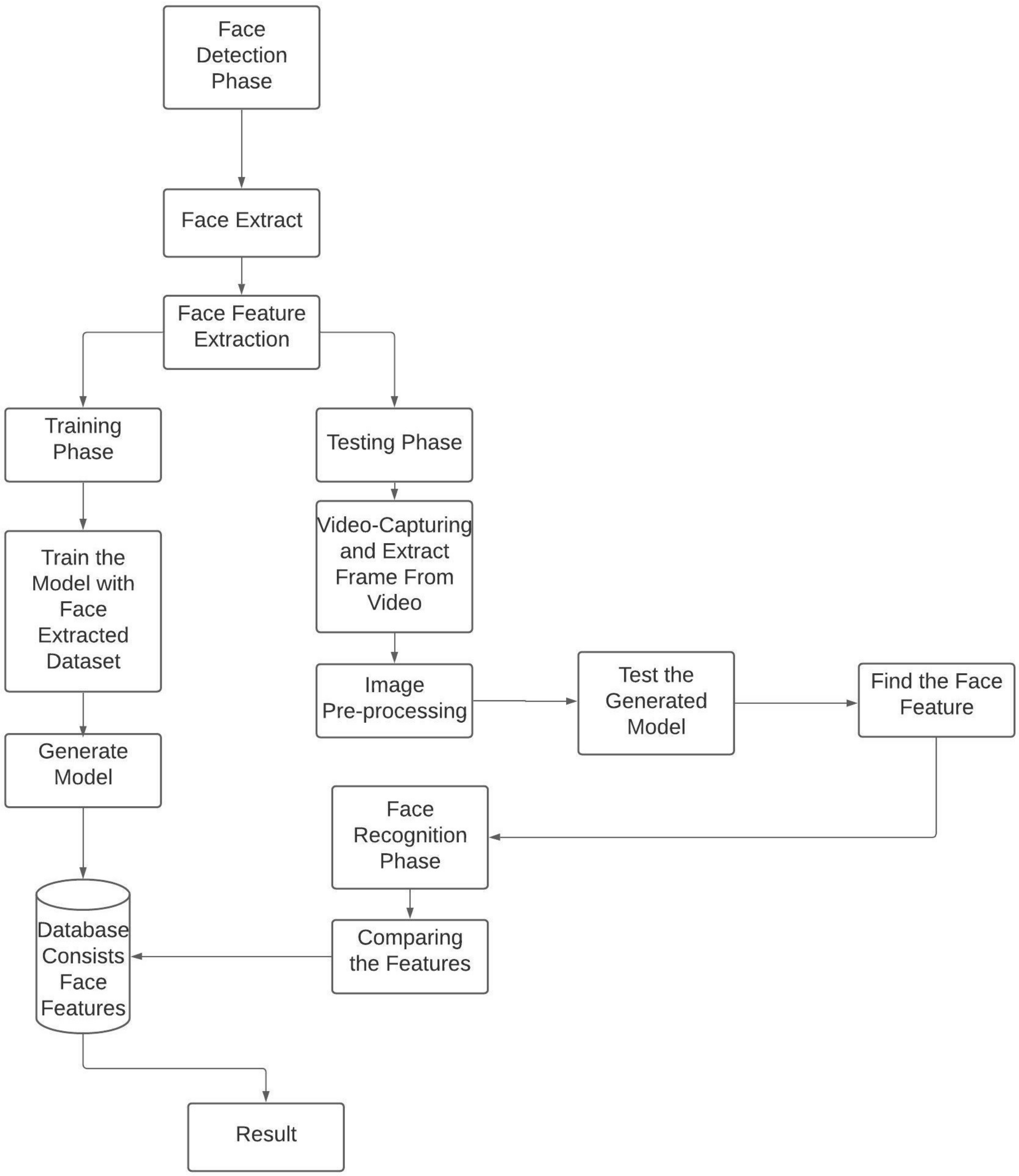


Figure-2: System Work flow of face feature extraction & face recognition

Consider the above diagram wherein the Face detection training phase, the color FERET standard dataset was used for the positive images. Color FERET dataset consists of 1208 different person images, each person having two frontal-face images with the three different image dimensions that respectively 512*768, 256*384, and 128*192. We trained our model with 1000 positive images of 512*768 dimension and 2000 negative images of the same dimension as consider for positive images. To train, the model viola-jones face detection algorithm is used.

Testing consists of video-capturing and Extract frame from the video phase wherein still images are used then in the pre-processing stage, the image converts into the RGB to greyscale, and in the last phase, the still image check within the trained model, and then results are displayed that whether the face is detected or not.

Consider figure-2 wherein face detection phase output was the detection of face. Now the output of that phase is input for the face feature extraction phase. That is further divided into two tasks that are training phase and testing phase.

In training phase, the model is trained against the face dataset that was extracted by the face detection phase. Based on the literature reviews there are two major methods to train the model that is template matching and feature matching. Template matching method [3] that involves the direct comparison of pixel intensity values taken from facial images. For the Feature matching method, there are several methods available. Most popular are Eigenface method and Fisher's Face method. In the Eigenface method [4], that calculated the eigenvectors and eigenvalues of the covariance matrix of facial images and only the principal components were preserved and compared. The Fisher's face method [5] used both principal component analysis and linear discriminant analysis to produce a subspace projection matrix

Literature Review

Face Feature Extraction A Complete Review [6]

In this paper, the researcher focuses on the general feature extraction framework for robust face recognition. They collect about 300 papers regarding face feature extraction. They believe that a general framework for face feature extraction consists of four major components: filtering, encoding, spatial pooling, and holistic representation. Then, they provide a brief review of deep learning networks, which can be seen as a hierarchical extension of the traditional framework. They provide a detailed performance comparison of various features on the LFW and FERET face database.

The filtering phase is to generate robust features for a given face image, it is beneficial to convolve the image with a specific filter, either using a pre-defined pattern or using a discriminative filter learned from the training dataset. In the Encoding phase, the output of the encoding phase could be a histogram or a feature vector. Spatial pooling can be seen as a way to further compress the coding vector according to the spatial layout of the image to form a final holistic feature. There are two classical pooling methods: average pooling that preserves the average response, and max pooling that preserve the maximum response. Block division can be seen as a part of feature pooling, as features extracted from different blocks are aggregated. This could be quite beneficial to face recognition, as the frontal structure of a face includes several distinct areas (eyes, nose, mouth, etc.). In Holistic Representation, they discourse about template matching methods and Feature Extraction methods. Eigenface method, Fisher's face method, and Principal Component Analysis methods are used for the Feature Extraction. [7-8]

After that, they discuss various filtering techniques including Gabor, LBP, Difference of Gaussians. Gabor wavelets are widely used in the image processing field in that they capture local structure corresponding to spatial frequency, spatial localization as well as orientation selectivity. In LBP, pixels of an image are labeled by thresholding the neighborhood of the pixel locally compared to the pixel itself to a binary number. Specifically, its feature vector is built by comparing the pixel with each of its neighboring pixels, and it interpolates values bilinearly at non-integer coordinates.

Feature selection using principal component analysis [9]

In this paper, researchers used principal component analysis for feature selection. PCA has been widely used in a variety of fields such as image processing, pattern recognition, data compression, data mining, machine learning, and computer vision. The principal components analysis methodology has been applied to face recognition, image denoising, and machine learning, etc.

The proposed method indeed deals with the feature selection issue from a viewpoint of numerical analysis. It exploits the eigenvector to evaluate the contribution, to the feature extraction result, of each feature component. The feature extraction result of an arbitrary sample is a linear combination of all the feature components of this sample and the entries of the eigenvector are the coefficients. That is, in the linear combination, the coefficient of i^{th} feature component of the sample is indeed the i^{th} entry of the eigenvector. As a result, if the i^{th} entry of the eigenvector has a very small absolute value, the i^{th} feature component of all the samples will statistically have a little effect on the feature extraction result.

For testing, they used 3 standard datasets ORL Dataset, Feret, and AR Datasets. Rates of classification errors of AR is 0.5047, Feret is 0.4583, ORL is 0.05. The rates of classification errors on the original samples obtained using the nearest neighbor classifier. The feature components selected by their method can obtain a lower classification error rate than the original samples. For example, while the classification error rate on the original samples of the AR database is 0.5047, scheme 1 of their method obtains a classification error rate of 0.3192 when it selects 600 feature components from all the feature components of the original samples of the AR database. we also see that scheme 1 of our method outperforms scheme 2 of our method.

Human Face Feature Extraction and Recognition Base on SIFT [10]

In this paper, researchers used SIFT (scale-invariant feature transform) method. SIFT was firstly presented by David Lowe in 2004 and is successful in video image matching. This method mainly includes four steps: extracting feature in multi-scale; locating the key-point; calculating the direction feature of the key-point and generating the descriptor of the key-point. This method has many advantages, such as scale invariance, rotation invariance, affine invariance, etc., and has strong robustness for the occlusion problem. Now, this method has been successfully used in object recognition and panorama stitching, etc. The image features extracted by SIFT have such advantages as scale invariability, rotation invariance, affine invariance, which could reduce mismatching because of occlusion, confusion, noise, and keep a high matching rate.

Firstly, extreme value is detected in the different scale of the image, and it is the key-point; secondly, the position of key-point is located by the filter of the key-points; thirdly, the gradient direction of the key-point is determined by the key-point neighborhood; fourthly, the feature descriptor is calculated by the feature of the key-point.

In this paper, the database of the test faces is the ORAL. This face DB is composed of forth people, and everyone has ten face photos that have different lightness, position, and expression. Those photos of the same people have apparent differences from each other. The size of the photo is 92 ×112. To check the performance of this algorithm, the number of training photos is from 3 to 7. The TN is the training number; the CT is the correct rate of face recognition; the ET is the misclassification rate; the RT is the refusal rate.

TN	*3*	*4*	*5*	*6*	*7*
CT(%)	92.1	94.6	96.3	96.8	97.1
ET(%)	6.3	4.5	2.4	1.8	1.3
RT(%)	0.06	0.9	1.3	1.4	1.6

The Result of Face Recognition

Face Recognition: Features versus Templates [3]

This paper focuses on two traditional classes of techniques applied to the recognition of digital images of frontal views of faces under roughly constant illumination. The first technique is based on the computation of a set of geometrical features from the picture of a face. This was the first approach toward an automated recognition of faces. The second class of techniques is based on template matching. The paper has focused on a comparative analysis of these two different approaches to face recognition.

The two approaches are Geometric, Feature-Based Matching and Template matching Geometric, Feature-Based Matching: A face can be recognized even when the details of the individual features (such as eyes, nose, and mouth) are no longer resolved. The remaining information is, in a sense, purely geometrical and represents what is left at a very coarse resolution. The idea is to extract relative position and other parameters of distinctive features such as eyes, mouth, nose, and chin. Goldstein and Kaya showed that a computer program provided with face features extracted manually could perform recognition with apparently satisfactory performance. Template Matching: In the simplest version of template matching, the image, which is represented as a bidimensional array of intensity values, is compared using a suitable metric (typically the euclidean distance) with a single template representing the whole face. There are, of course, several, more sophisticated ways of performing template matching. For instance, the array of grey levels may be suitably preprocessed before matching. Several full templates per each face may be used to account for the recognition from different viewpoints. Still another important variation is to use, even for a single viewpoint, multiple templates.

The results obtained by using this approach (about 90% correct recognition using geometrical features and perfect recognition using template matching) favors the implementation of this template-matching approach.

Facial feature extraction for face recognition: a review [11]

This paper provides an up-to-date review of major human facia recognition research. The overall process which is conveyed in the paper focuses on face detection, feature extraction, and face recognition. The importance of facial features for face recognition cannot be overstated. Many face recognition systems need facial features in addition to the holistic face, as suggested by studies in psychology. Several Techniques of facial feature extraction are covered in this paper are Color segmentation techniques and Appearance-based approaches.

The Color segmentation technique makes use of skin color to isolate the face. Any non-skin color region within the face is viewed as a candidate for eyes and/or mouth. The performance of such techniques on facial image databases is rather limited, due to the diversity of ethnic backgrounds.

The concept of “feature” in these approaches differs from simple facial features such as eyes and mouth. Any extracted characteristic from the image is referred to as a feature. Methods such as principal component analysis (PCA), independent component analysis, and Gabor-wavelets [45] are used to extract the feature vector. These approaches are commonly used for face recognition rather than personal identification.

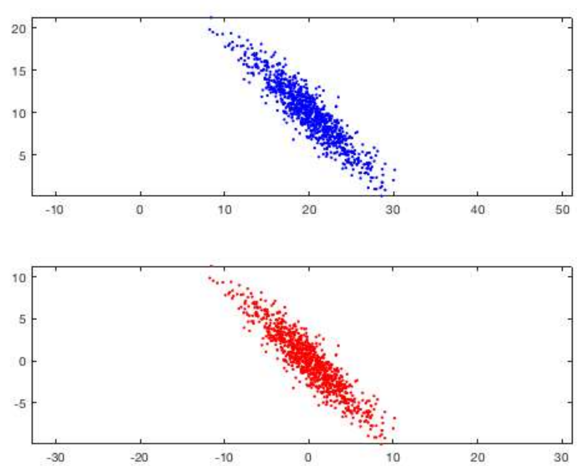
In this paper, they have reviewed the works done in facial feature extraction, concerning resource discovery. They have also focused on and evaluated different techniques and ideas, used to reduce bandwidth consumption during this process.

Face Feature Extraction Methodology

Eigenfaces Method Approach [3-10]

Eigenfaces refer to an appearance-based approach to face recognition that seeks to capture the variation in a collection of face images and use this information to encode and compare images of individual faces in a holistic (as opposed to a parts-based or feature-based) manner. Specifically, the eigenfaces are the principal components of the distribution of faces, or equivalently, the eigenvectors of the covariance matrix of the set of face images, where an image with *N* pixels is considered a point (or vector) in *N*-dimensional space.

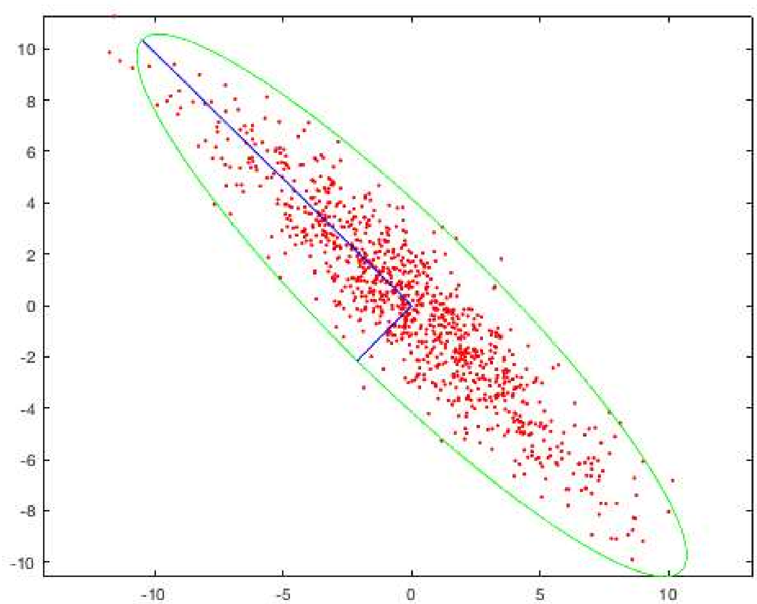
Eigenface is primarily a dimension reduction method, a system can represent many subjects with a relatively small set of data. Dimensionality reduction is a type of unsupervised learning where we want to take higher-dimensional data, like images, and represent them in a lower-dimensional space. following image as an example.



These plots show the same data, except the bottom chart zero-centers it. Notice that our data do not have any labels associated with them because this is *unsupervised learning*! In our simple case, dimensionality reduction will reduce these data from a 2D plane to a 1D line. If we had 3D data, we could reduce it down to a 2D plane or even a 1D line. All dimensionality reduction techniques aim to find some hyperplane, a higher-dimensional line, to *project* the points onto. We can imagine a *projection* as taking a flashlight perpendicular to the hyperplane we project onto and plotting where the shadows fall on that hyperplane. For example, in our above data, if we wanted to project our points onto the x-axis, then we pretend each point is a ball and our flashlight would point directly down or up (perpendicular to the x-axis) and the shadows of the points would fall on the x-axis. This is a *projection.*

Principal Component Analysis

One technique of dimensionality reduction is called principal component analysis (PCA)**.** The idea behind PCA is that we want to select the hyperplane such that when all the points are projected onto it, they are maximally spread out. In other words, we want the *axis of maximal variance.* Let's consider our example plot above. A potential axis is an x-axis or y-axis, but, in both cases, that's not the best axis. However, if we pick a line that cuts through our data diagonally, that is the axis where the data would be most spread.

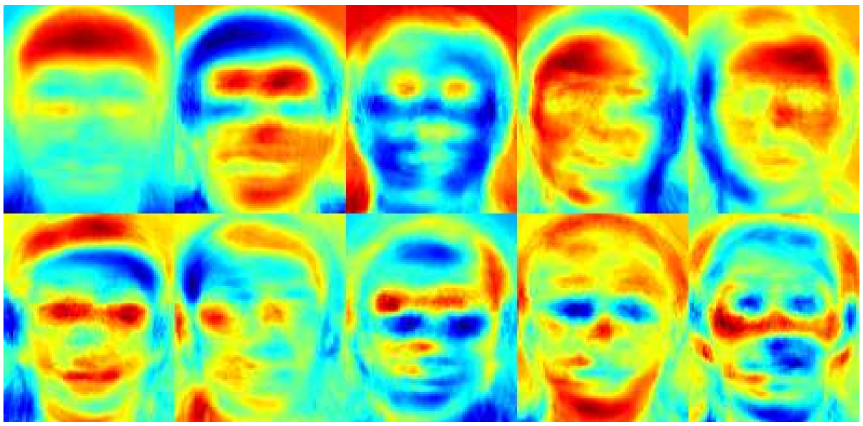


The longer blue axis is the correct axis! If we were to project our points onto this axis, they would be maximally spread. We can borrow a term from linear algebra called eigenvectors**.** This is where eigenfaces get its name. Essentially, we compute the covariance matrix of our data and consider that covariance matrix's largest eigenvectors**.** Those are our *principal axes* and the axes that we project our data onto to reduce dimensions. Using this approach, we can take high-dimensional data and reduce it down to a lower dimension by selecting the largest eigenvectors of the covariance matrix and projecting them onto those eigenvectors.

EigenFaces Face Recognizer

This algorithm considers the fact that not all parts of a face are equally important and equally useful. When you look at someone you recognize him/her by his distinct features like eyes, nose, cheeks, forehead and how they vary concerning each other. so, you are actually focusing on the areas of maximum change (mathematically speaking, this change is variance) of the face. For example, from eyes to nose there is a significant change and the same is the case from nose to mouth. When you look at multiple faces you compare them by looking at these parts of the faces because these parts are the most useful and important components of a face. Important because they catch the maximum change among faces, change the helps you differentiate one face from the other. This is exactly how EigenFaces face recognizer works.

EigenFaces face recognizer looks at all the training images of all the persons as a whole and tries to extract the components which are important and useful (the components that catch the maximum variance/change) and discards the rest of the components. This way it not only extracts the important components from the training data but also saves memory by discarding the less important components. These important components it extracts are called principal components. Below is an image showing the principal components extracted from a list of faces.



You can see that the principal component actually represents faces and these faces are called eigenfaces and hence the name of the algorithm.

So, this is how EigenFaces face recognizer trains itself (by extracting principal components). Remember, it also keeps a record of which principal component belongs to which person. One thing to note in the above image is the Eigenfaces algorithm also considers illumination as an important component.

Algorithm Description [12]

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a random vector with observations $x_i \in R^d$.

1. Compute the mean μ

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

2. Compute the Covariance Matrix *S*

$$S = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T$$

3. Compute the eigenvalues λ_i and eigenvectors v_i of S

$$S v_i = \lambda_i v_i, i = 1, 2, \ldots, n$$

4. Order the eigenvectors descending by their eigenvalue. The k principal components are the eigenvectors corresponding to the k largest eigenvalues.

The k principal components of the observed vector x are then given by:

$$y = W^T (x - \mu)$$

where $W = (v_1, v_2, \ldots, v_k)$.

The reconstruction from the PCA basis is given by:

$$x = Wy + \mu$$

where $W = (v_1, v_2, \ldots, v_k)$.

References

1. Chiara Turati, Viola Macchi Cassia, Francesca Simion, and Irene Leo. Newborns' face recognition: Role of inner and outer facial features. Child development, 77(2):297–311, 2006.
2. Gonzalez, Rafael C, and Richard E. Woods. Digital Image Processing. , 2002. Print.
3. R. Brunelli and T. Poggio, "Face recognition: features versus templates," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052, Oct. 1993, doi: 10.1109/34.254061.
4. L. Sirovich and M. Kirby, ''Low-dimensional procedure for the characterization of human faces,'' J. Opt. Soc. Amer. A, Opt. Image Sci. Vis., vol. 4, no. 3, pp. 519–524, 1987.
5. P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, ''Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997
6. H. Wang, J. Hu and W. Deng, "Face Feature Extraction: A Complete Review," in IEEE Access, vol. 6, pp. 6001-6039, 2018, doi: 10.1109/ACCESS.2017.2784842.
7. L. Best-Rowden, H. Han, C. Otto, B. F. Klare, and A. K. Jain,''Unconstrained face recognition: Identifying a person of interest froma media collection,''IEEE Trans. Inf. Forensics Security, vol. 9, no. 12,pp. 2144–2157, Dec. 2014.
8. Y. Huang, Z. Wu, L. Wang, and T. Tan, ''Feature coding in image classifi-cation: A comprehensive study,''IEEE Trans. Pattern Anal. Mach. Intell.,vol. 36, no. 3, pp. 493–506, Mar. 2014.
9. F. Song, Z. Guo and D. Mei, "Feature Selection Using Principal Component Analysis," 2010 International Conference on System Science, Engineering Design and Manufacturing Informatization, 2010, pp. 27-30, doi: 10.1109/ICSEM.2010.14.
10. H. Yanbin, Y. Jianqin and L. Jinping, "Human Face Feature Extraction and Recognition Base on SIFT," 2008 International Symposium on Computer Science and Computational Technology, 2008, pp. 719-722, doi: 10.1109/ISCSCT.2008.249.
11. E. Bagherian and R. W. O. K. Rahmat, "Facial feature extraction for face recognition: a review," 2008 International Symposium on Information Technology, 2008, pp. 1-9, doi: 10.1109/ITSIM.2008.4631649.
12. https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html