

El Shorouk Academy Higher Institute of Engineering

Communications and Computer Engineering Department

Project Background

**Names:**

**Omar Gamal Saad**

**(01009317165) (122170252@sha.edu.eg)**

**Abdullah Said Mohamed-Shokri Abbas (01126902308) (122170214@sha.edu.eg)**

**Abdullah Gamil Mohamed Ahmed (01098337244) (**[**122170915@sha.edu.eg**](mailto:122170915@sha.edu.eg)**)**

**Abdelrahman Mahmoud Ahmed Taha (01116692155) (122170841@sha.edu.eg)**

**To Dr: Ahmed El-Mahalawy**

PCA

Principal Component Analysis commonly uses the eigenfaces

in which the probe and gallery images must be the same size

as well as normalized to line up the eyes and mouth of the

subjects whining the images. Approach is then used to reduce

the dimension of data by the means of image compression

basics and provides most effective low dimensional structure

of facial pattern. This reduction drops the unuseful

information and decomposes the face structure into

orthogonal (uncorrelated) components known as eigenfaces.

Each face image is represented as weighted sum feature vector

of eigenfaces which are stored in 1-D array. A probe image is

compared against the gallery image by measuring the distance

between their respective feature vectors then matching result

has been disclosed. The main advantage of this technique is

that it can reduce the data needed to identify the individual to

1/1000th of the data presented. [12]

PCA solves the recognition problem within a representation

space of lower dimension than image space. PCA is an

eigenface method which helps in the reduction of the

dimensionality of the original data space. But there is a

disadvantage of PCA which says that recognition rate

decreases under varying pose and illumination.

A face recognition system can be considered as a good system

if we extract with the help of Principal Component Analysis

and for recognition back propagation Neural Network are

used.

LDA

Linear Discriminant Analysis is an appearance based

technique used for dimensionality reduction and recorded a

great performance in face recognition. It provides us with a

small set of features that carry the most relevant information

for classification purposes.

LDA is a statistical approach for classifying samples of

unknown classes based on training samples with known

classes. This technique aims to maximum between-class

(across users) variance and minimum within class (within

user) variance. In these techniques a block represents a class,

and there are a large variations between blocks but little

variations within classes. It searches for those vectors in

underlying space that best discriminate among classes (rather

than those that best describe the data). More formally given a

number of independent features relative to which the data is

described. LDA creates a linear combination of these which

yields the largest mean difference between desire classes.

Table

Description automatically generated

SVM

Support Vector Machine (SVM) is a binary classifier as a

method for learning. SVM is a classification method that

separate two data sets with maximum distance between

them.The concept is to extend the spatial resolution around

the margin by a conformal mapping, such that the divisibility

between classes is increased. SVM cannot be applied directly

when some of the features (face pixels) are occluded. In this

case, values for those dimensions are unknown. SVM cannot

be used when the feature vectors defining our samples have

missing entries.

Support vector machines (SVMs) are formulated to solve a

classical two class pattern recognition problem

SIFT

In 2004 Lowe, invents SIFT descriptor which is invariant to

scale, rotation, affine transformation, noise, occlusions and is

highly distinctive. SIFT features consist of four major stages

in detection and representation; they are (1) finding scale-

space extrema; (2) key point localization and filtering; (3)

orientation assignment; (4) key point descriptor. The first

stage is to construct the key points of images by using

Difference-of-Gaussian (DoG) function. The second stage,

candidate key points are restricted to sub-pixel accuracy and

removed if found to be unreliable. The third stage represents

the dominant orientations for each essential point of the

images. The final stage constructs a descriptor for each key

point location depends upon the image gradients in its local

neighborhood. Then the SIFT descriptor is accepting the 128-

dimensional vector which used to identify the neighborhood

around a pixel. The SIFT extracts the key points (locations

and descriptors) for all the database images. Then given an

altered image SIFT extracts the key point for that image and

compares that point to the dataset.

SURF

The SURF also extracts the key points from both the database

images and the altered images. This method matches the key

points between altered image and each database image. In

2008, H. Bay invents SURF descriptor which is invariant to a

scale and in-plane rotation features. It consists of two stages

such as interest point detector and interest point descriptor. In

the first stage, locate the interest point in the image. Use the

Hessian matrix to find the approximate detection. [2]

SURF is a scale and in-plane rotation invariant detector and

descriptor. [16] SURF detectors are find the interest points in

an image, and descriptors are used to extract the feature

vectors at each interest point just as in SIFT. SURF uses

Hessian-matrix approximation to locate the interest points

instead of difference of Gaussians (DoG) filter used in SIFT.

SURF as a descriptor uses the first-order Haar wavelet

responses in x and y, whereas the gradient is used by SIFT.

SURF usually uses 64 dimensions in SURF to reduce the time

cost for both feature matching and computation. SURF has

three times better performance as compared to SIFT.

Table

Description automatically generated

## CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional neural network (CNN) is one of the breakthroughs of [artificial neural networks](https://recfaces.com/articles/ai-facial-recognition) (ANN) and AI development. It’s one of the most popular algorithms in deep learning, a type of machine learning in which a model learns to perform classification tasks directly on an image, video, text, or sound. The model shows impressive [results](https://ieeexplore.ieee.org/abstract/document/8308186) in several fields: computer vision, natural language processing (NLP), and the largest image classification data set (Image Net). CNN is a normal neural network with new layers — convolutional and pooling. CNN can have dozens and hundreds of these layers, and each of them learns to detect different imaging features.

## EIGENFACES

Eigenfaces is a face detection and [recognition](https://www.researchgate.net/publication/277930278_Face_Recognition_using_Eigenfaces) method that determines face variance in image data sets. It uses these variances to encode and decode faces with machine learning. A set of eigenfaces is a collection of “standardized face ingredients” determined by statistical analysis of a large number of face images. Facial features are assigned mathematical values, as this method doesn’t use digital pictures but rather statistical databases. Any human face is a combination of these values with different percentages.

## FISHERFACES

Fisherfaces is one of the most popular facial recognition algorithms; it’s considered superior to many of its alternatives. As an improvement to the Eeigenfaces algorithm, it’s often compared to Eigenfaces and considered more successful in class distinction in the training process. The key advantage of this algorithm is its ability to interpolate and extrapolate over lighting and facial expression variation. There are [reports](https://iopscience.iop.org/article/10.1088/1742-6596/1028/1/012119/pdf) of 93% accuracy of the Fisherfaces algorithm when combined with the PCA method at the preprocessing stage.

## KERNEL METHODS: PCA AND SVM

**The principal component analysis (PCA)** is a universal statistical [method](https://www.researchgate.net/profile/Liton-Paul/publication/318362885_Face_recognition_using_principal_component_analysis_method/links/5dc70ae34585151435fb427b/Face-recognition-using-principal-component-analysis-method.pdf) with many practical applications. When used in the face recognition process, PCA aims to reduce the source data size while preserving the most relevant information. It generates a set of weighted eigenvectors that, in their turn, build up eigenfaces — extensive sets of different human face images. A linear combination of eigenfaces represents every image in the training set. The PCA is used to receive these eigenvectors from the covariance matrix of a training image set. For each image, its main components are calculated (from 5 to 200). The other components encode minor differences between faces and noise. The recognition process includes comparing the unknown image’s main component to the components of all other images.

**Support vector machine (SVM)** is a machine learning algorithm that uses a two-group classification principle for distinguishing faces from “not-faces.” For each category, an SVM model receives a labeled training data set to categorize new test data. Researchers apply linear and nonlinear SVM training models for face recognition. The recent [results](https://ieeexplore.ieee.org/document/1285674) show that the nonlinear training machine has a larger margin and better recognition and classification results.

## HAAR CASCADES

Haar Cascade is an object detection method used to locate objects on images. The algorithm learns from a large number of positive and negative samples — the former contains an object of interest, and the latter contains anything other than the object you are looking for. After training, the classifier can find an object of interest on new images. The method was [used](https://www.researchgate.net/publication/345896024_22_Face_Recognition_using_Haar_-_Cascade_Classifier_for_Criminal_Identification) in criminal identification in combination with the local binary pattern algorithm to recognize faces. The Haar cascade classifier uses 200 (out of 6000) features, which ensures an 85-95% recognition rate even with varying expressions.

## THREE-DIMENSIONAL RECOGNITION

The underlying idea of 3D face recognition technology is the human skull’s unique structure. Each person’s skull structure is unique and can be described by several dozen parameters. This facial recognition method is based on comparing a 3D facial scan to the database patterns. It has an essential advantage — makeup, facial hair, glasses, and similar factors don’t affect the detection and recognition process. The latest [research](https://www.researchgate.net/publication/6506621_Three-Dimensional_Face_Recognition_in_the_Presence_of_Facial_Expressions_An_Annotated_Deformable_Model_Approach) has used the technology of mapping the 3D geometry information on a regular 2D grid. It allows the combination of 3D data’s descriptiveness with 2D data’s computational efficiency and shows the highest performance reported on the FRGC v2 (Face Recognition Grand Challenge 3D facial database).

## SKIN TEXTURE ANALYSIS

Skin recognition technology has many applications — face detection algorithms, objectionable image filtering, hand gesture analysis, etc. It usually uses high-resolution images. Particular cases of skin texture analysis use different unique parameters like moles, skin color, skin tones, and many others. Recent [research](https://arxiv.org/ftp/arxiv/papers/1311/1311.6049.pdf) based on a combination of texture features and skin color showed interesting results. The researchers used a neural network to develop and test a skin recognition system. The feed-forward neural networks used in the project classified input texture images as “skin” and “non-skin” and showed an impressive performance.

## THERMAL CAMERAS

A thermal camera is a device used for monitoring the temperature distribution of the examined surface. The temperature distribution is displayed as a colored picture with different colors corresponding to temperatures. The technology already has several [practical applications](https://academic.oup.com/jlb/article/7/1/lsaa038/5857112) adapting to global changes — smartphone-based immunity certificates, remote fever detection, and thermal facial recognition. Thermal face recognition models are based on the unique temperature patterns of a human face. Human consistent temperature “signatures” are measured with thermal infrared (IR) imaginary. Using the thermal method in face recognition has an undeniable benefit — makeup, facial hair, hats, and glasses don’t affect its accuracy. Moreover, it distinguishes twin siblings.

## ANFIS

An adaptive neuro-fuzzy interference system (ANFIS) is a type of artificial neural network. This method integrates the principles of neural networks with fuzzy logic principles and combines their advantages in a single structure. ANFIS is used to classify image features extracted from datasets on the preprocessing stage. Data scientists combine this method with a variety of feature extraction algorithms. Thus, some studies [reported](https://www.researchgate.net/publication/282240664_Face_recognition_using_2DPCA_and_ANFIS_classifier) incredible 97.1% ANFIS classification accuracy after feature extraction with 2D principle component analysis.

## LOCAL BINARY PATTERNS HISTOGRAMS (LBPH)

This method uses local binary patterns (LBP), a simple, effective texture operator in computer vision that marks pixels in an image by setting each pixel’s neighborhood threshold and treating the result as a binary number. At the learning stage, the [LBPH](https://thesai.org/Downloads/Volume10No5/Paper_35-LBPH_based_Enhanced_Real_Time_Face_Recognition.pdf) algorithm creates histograms for each image that is labeled and classified. Each histogram represents each image from the training set. This way, the actual recognition process implies comparing histograms of any two images.

## FACENET

The face recognition system FaceNet, developed by Google researchers in 2015, is based on face recognition benchmark datasets. Available pre-trained models and various open-source third-party implementations make this system quite wide-spread. FaceNet shows excellent [results](https://ieeexplore.ieee.org/document/8985786) in research-conducting surveys, testing performance, and accuracy compared to other algorithms developed earlier. FaceNet accurately extracts face embeddings, high-quality features used for training face identification systems at later stages.

## NEC

The solution developed by the Japanese technology company NEC allows highly accurate identification of people while recognizing age changes. The solution uses Adaptive Region Mixed Matching, a model that focuses on highly similar segments for mapping. The NEC technology divides input and registered images into small segments and focuses only on greater similarity segments. It allows the system to show higher identification accuracy, even in the case of the face wearing a mask or glasses. As its underlying [algorithm](http://www.mva-org.jp/Proceedings/2011CD/papers/13-02.pdf), the NEC solution uses generalized learning vector quantization (GLVQ).

## MEGVII (FACE++)

Chinese technology company Megvii has become known worldwide after introducing its face recognition application platform. It’s multi-purpose software. The Megvii algorithm is based on [graph detection](https://www.sciencedirect.com/science/article/abs/pii/S0957417414002000#:~:text=Another%20approach%20called%20Local%20Graph,each%20pixel%20of%20an%20image.) and fuzzy image search technology. The tech solution uses the company’s proprietary deep learning framework MegEngine, built on big data. The company’s technology successfully performs facial information extraction and includes several key features: face and human detection and tracking, face recognition and clustering, key-point detection, face attribute estimation, and [face search engine](https://recfaces.com/articles/face-search).

## FACIAL RECOGNITION: COMBINING DIFFERENT TECHNIQUES

Each facial recognition technology has its effective features. However, recent research has proved that the best results are achieved with combinations of different algorithms and methods. These combinations aim to solve the facial recognition process’s many routine problems — differences in facial expressions, posing, lighting conditions, image noise, etc. The latest [experiments](https://www.sciencedirect.com/science/article/pii/S2590005619300141) combine the LBP algorithm with advanced image processing techniques: bilateral filter, histogram equalization, contrast adjustment, and image blending. Such a technique shows impressive improvement to the LBP code and looks very promising for further research.

Graphical user interface, application

Description automatically generated