Intelligent Global Face Recognition

Adnan Khashman
Near East University
Northern Cyprus

1. Introduction

Face recognition by humans is a natural process that we perform on daily basis. A quick glance at a face and we are able to recognize the face and, most of the time, name the person. Such a process occurs so quickly that we never think of what exactly we looked at in that face. Some of us may take a longer time while trying to name the person, however, the recognition of the familiar face is usually instantaneous.

The complexity of a human face arises from the continuous changes in the facial features that take place over time. Despite these changes, we humans are still able to recognize faces and identify the persons. Of course, our natural recognition ability extends beyond face recognition, where we are equally able to quickly recognize patterns, sounds and smells. Unfortunately, this natural ability does not exist in machines, thus the need for artificially simulating recognition in our attempts to create intelligent autonomous machines.

Face recognition by machines can be invaluable and has various important applications in real life, such as, electronic and physical access control, national defense and international security. Simulating our face recognition natural ability in machines is a difficult task, but not impossible. Throughout our life time, many faces are seen and stored naturally in our memories forming a kind of database. Machine recognition of faces requires also a database which is usually built using facial images, where sometimes different face images of a one person are included to account for variations in facial features.

Current face recognition methods rely on: detecting local facial features and using them for face recognition or on globally analyzing a face as a whole. The first approach (local face recognition systems) uses facial features within the face such as (eyes, nose and mouth) to associate the face with a person. The second approach (global face recognition systems) uses the whole face for identifying the person.

This chapter reviews some known existing face recognition methods and presents one case study of a recently developed intelligent face recognition system that uses global pattern averaging for facial data encoding prior to training a neural network using the averaged patterns.

The development of intelligent systems that use neural networks is fascinating and has lately attracted more researchers into exploring the potential applications of such systems. The idea of simulating the human perceptions and modeling our senses using machines is great and may help humankind in medical advancement, space exploration, finding alternative energy resources or providing national and international security and peace. Intelligent systems are being increasingly developed aiming to simulate our perception of

Face Recognition

various inputs (patterns) such as images, sounds... etc. Biometrics is an example of popular applications for artificial intelligent systems. The development of an intelligent face recognition system requires providing sufficient information and meaningful data during machine learning of a face.

This chapter presents a brief review of known face recognition methods such as Principal Component Analysis (PCA) (Turk & Pentland, 1991), Linear Discriminant Analysis (LDA) (Belhumeur et al., 1997) and Locality Preserving Projections (LPP) (He et al., 2005), in addition to intelligent face recognition systems that use neural networks (Khashman, 2006).

There are many works emerging every year suggesting different methods for face recognition, however, these methods are appearance-based or feature-based methods that search for certain global or local representation of a face.

The chapter will also provide one detailed case study on intelligent global face recognition system. In this case a neural network is used to identify a person upon presenting his/her face image. Global pattern averaging is used for face image preprocessing prior to training or testing the neural network. Averaging is a simple but efficient method that creates "fuzzy" patterns as compared to multiple "crisp" patterns, which provides the neural network with meaningful learning while reducing computational expense.

Intelligent global face recognition considers a person’s face and its background and suggests that a quick human “glance” can be simulated in machines using image pre-processing and global pattern averaging, whereas, the perception of a “familiar” face can also be achieved by exposing a neural network to the face via training (Khashman, 2006).

The chapter is organized as follows: section 1 contains an introduction to the chapter. Section 2 presents a review on problems and difficulties in face recognition. Section 3 describes known conventional face recognition methods and a selection of intelligent face recognition techniques. Section 4 presents in details our case study on intelligent global face recognition. Section 5 presents analysis and discussion of the results of implementing the work that is described in section 4. Finally, section 6 concludes this chapter and provides a discussion on the efficiency of intelligent face recognition by machines.

2. Problems with Face Recognition

The databases used in developing face recognition systems rely on images of human faces captured and processed in preparation for implementing the recognition system. The variety of information in these face images makes face detection difficult. For example, some of the conditions that should be accounted for, when detecting faces are (Yang et al., 2002):

- Occlusion: faces may be partially occluded by other objects
- Presence or absence of structural components: beards, mustaches and glasses
- Facial expression: face appearance is directly affected by a person’s facial expression
- Pose (Out-of Plane Rotation): frontal, 45 degree, profile, upside down
- Orientation (In Plane Rotation): face appearance directly varies for different rotations about the camera's optical axis
- Imaging conditions: lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, gain control, lenses), resolution

Face Recognition follows Face detection. Face recognition related problems include (Li & Jain, 2005):

- Face localization
Intelligent Global Face Recognition

- Aim to determine the image position of a single face
- A simplified detection problem with the assumption that an input image contains only one face
- Facial feature extraction (for local face recognition)
  - To detect the presence and location of features such as eyes, nose, nostrils, eyebrow, mouth, lips, ears, etc
  - Usually assume that there is only one face in an image
- Face recognition (identification)
- Facial expression recognition
- Human pose estimation and tracking

The above obstacles to face recognition have to be considered when developing face recognition systems. The following section reviews briefly some known face recognition methods.

3. Face Recognition Methods

With the increase in computational power and speed, many face recognition techniques have been developed over the past few decades. These techniques use different methods such as the appearance-based method (Murase & Nayar, 1995); where an image of a certain size is represented by a vector in a dimensional space of size similar to the image. However, these dimensional spaces are too large to allow fast and robust face recognition. To encounter this problem other methods were developed that use dimensionality reduction techniques (Belhumeur et al., 1997); (Levin & Shashua, 2002); (Li et al., 2001); (Martinez & Kak, 2001). Examples of these techniques are the Principal Component Analysis (PCA) (Turk & Pentland, 1991) and the Linear Discriminant Analysis (LDA) (Belhumeur et al., 1997).

PCA is an eigenvector method designed to model linear variation in high-dimensional data. Its aim is to find a set of mutually orthogonal basis functions that capture the directions of maximum variance in the data and for which the coefficients are pairwise decorrelated. For linearly embedded manifolds, PCA is guaranteed to discover the dimensionality of the manifold and produces a compact representation. PCA was used to describe face images in terms of a set of basis functions, or “eigenfaces”.

LDA is a supervised learning algorithm. LDA searches for the projection axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other. Unlike PCA which encodes information in an orthogonal linear space, LDA encodes discriminating information in a linearly separable space using bases that are not necessarily orthogonal. It is generally believed that algorithms based on LDA are superior to those based on PCA. However, other work (Martinez & Kak, 2001) showed that, when the training data set is small, PCA can outperform LDA, and also that PCA is less sensitive to different training data sets.

Another linear method for face analysis is Locality Preserving Projections (LPP) (He & Niyogi, 2003) where a face subspace is obtained and the local structure of the manifold is found. LPP is a general method for manifold learning. It is obtained by finding the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the manifold. Therefore, though it is still a linear technique, it seems to recover important aspects of the intrinsic nonlinear manifold structure by preserving local structure. This led
to a recently developed method for face recognition; namely the Laplacianface approach, which is an appearance-based face recognition method (He et al., 2005). The main difference between PCA, LDA, and LPP is that PCA and LDA focus on the global structure of the Euclidean space, while LPP focuses on local structure of the manifold, but they are all considered as linear subspace learning algorithms. Some nonlinear techniques have also been suggested to find the nonlinear structure of the manifold, such as Locally Linear Embedding (LLE) (Roweis & Saul, 2000). LLE is a method of nonlinear dimensionality reduction that recovers global nonlinear structure from locally linear fits. LLE shares some similar properties to LPP, such as a locality preserving character. However, their objective functions are totally different. LPP is obtained by finding the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the manifold. LPP is linear, while LLE is nonlinear. LLE has also been implemented with a Support Vector Machine (SVM) classifier for face authentication (Pang et al., 2005).

Approaches that use the Eigenfaces method, the Fisherfaces method and the Laplacianfaces method have shown successful results in face recognition. However, these methods are appearance-based or feature-based methods that search for certain global or local representation of a face. None so far has considered modeling the way we humans recognize faces.

One of the simplest methods for modeling our way of recognizing faces is neural network arbitration. This has been explored with the aim of developing face recognition systems that incorporate artificial intelligence using neural networks in order to provide an intelligent system for face recognition.

The use of neural networks for face recognition has also been addressed by (Lu X. et al., 2003); (Zhang et al., 2004); (Pang et al., 2005); (Fan & Verma, 2005). More recently, Li et al. (Li G. et al., 2006) suggested the use of a non-convergent chaotic neural network to recognize human faces. Lu et al. (Lu K. et al., 2006) suggested a semi-supervised learning method that uses support vector machines for face recognition. Zhou et al. (Zhou et al., 2006) suggested using a radial basis function neural network that is integrated with a non-negative matrix factorization to recognize faces. Huang and Shimizu (Huang & Shimizu, 2006) proposed using two neural networks whose outputs are combined to make a final decision on classifying a face. Park et al. (Park et al., 2006) used a momentum back propagation neural network for face and speech verification.

Many more face recognition methods that use artificial intelligence are emerging continually; however, one particular method; namely Intelligent Global Face Recognition, will be studied in this chapter, and is therefore presented in the following section.

4. Intelligent Face Recognition Using Global Pattern Averaging

One of our commonly referred five senses is “Seeing”. We see and perceive objects in different ways depending on our individuality. However, we share the ability to recognize objects or patterns quickly even though our experience of these objects is minimal. A quick “glance” onto a “familiar” face and recognition occurs. The following section presents our hypothesis where we aim to simulate our way of recognizing faces in machines using a neural network model.
4.1 Hypothesis of Simulating Glance and Familiarity

This case study presents an intelligent face recognition system that uses global pattern averaging of a face and its background and aims at simulating the way we see and recognize faces. This is based on the suggestion that a human “glance” of a face can be approximated in machines using pattern averaging, whereas, the “familiarity” of a face can be simulated by a trained neural network (Khashman, 2006). A real-life application will be presented using global averaging and a trained back propagation neural network to recognize the faces of 30 persons from our databases.

4.2 Databases and Method

One common problem with processing images is the large amount of data that is needed for meaningful results. Although neural networks have the advantage of parallel processing, there is still a need to pre-process images to reduce the amount of data while retaining meaningful information on the images. This is an important requirement for an efficient system that has low time and computational expense.

There are 30 persons of various gender, ethnicity and age whose faces were to be recognized and thus their face images would be used as the database for the work presented within this case study. Each face has three different projections, which were captured while looking: Left (LL), Straight (LS) and Right (LR) as shown in Figure 1 resulting in 90 images that are used for implementing the intelligent system. Figures 2, 3 and 4 show these 90 images representing 30 persons looking straight (LS), right (LR) and left (LS) respectively.

All original images are gray and of size (512x512) pixels. The images were compressed and their size reduced to 128x128 pixels. A window of size 100x100 pixels; that contains the face and its background, is then extracted and the data within this relatively smaller size image is used for training and eventually testing the neural network.

Figure 1. Person 21 looking: a- right (LR) b- straight (LS) c- left (LL)
Figure 2. Own face database of 30 persons looking straight (LS)
Figure 3. Own face database of 30 persons looking right (LR)
Figure 4. Own face database of 30 persons looking left (LL)
4.3 Glance Simulation (Global Pattern Averaging)

The method used for presenting the images to the neural network is based on global pattern averaging, which provides the glance approximation. A face image of size 100x100 pixels is segmented and the values of the pixels within each segment are averaged. The result average values are then used as input data for the neural network.

The averaging of the segments within an image reduces the amount of data required for neural network implementation thus providing a faster recognition system. This also provides flexible mathematical inputs for neural networks that simulate the quick glance of a human which is sufficient for pattern recognition. Global pattern averaging can be defined as follows:

\[
\text{PatAv}_i = \frac{1}{S_k S_l} \sum_{k=1}^{s_k} \sum_{l=1}^{s_l} p_i(k,l),
\]

(1)

where \(k\) and \(l\) are segment coordinates in the \(x\) and \(y\) directions respectively, \(i\) is the segment number, \(S_k\) and \(S_l\) are segment width and height respectively, \(P_i(k,l)\) is pixel value at coordinates \(k\) and \(l\) in segment \(i\), \(\text{PatAv}_i\) is the average value of pattern in segment \(i\), that is presented to neural network input layer neuron \(i\). The number of segments in each window (of size \(X\times Y\) pixels) containing a face, as well as the number of neurons in the input layer is \(i\) where

\[
i = \{-0, 1, 2, \ldots, n\},
\]

(2)

and

\[
n = \left( \frac{X}{S_k} \right) \left( \frac{Y}{S_l} \right).
\]

(3)

Segment size of 10x10 pixels \((S_k = S_l = 10)\) has been used and average values representing the image were obtained, thus resulting in 100 average values in total \((n = 100)\) that were used as the input to the neural network for both training and testing.

Figure 5 shows an example of this pre-processing phase. The original 512x512 pixel image is reduced to 256x256 pixels and then to 128x128 pixels. This is followed by extracting a region of size 100x100 pixels that contains the face. The extracted region is then segmented (tiled) and averaged yielding a 10x10 pixel pattern that represents the original image.

Figure 5. Image pre-processing before neural network training or testing
4.4 Familiarity Simulation (Neural Network Implementation)

The multilayer perceptron neural network, which was developed as part of this global face recognition system, is based on the back propagation learning algorithm, with a total number of three layers, comprising, input layer, hidden layer and output layer. The input layer has 100 neurons, each receiving an averaged value of the face image segments. The hidden layer consists of 99 neurons, whereas the output layer has 30 neurons according to the number of persons. Figure 6 shows the topology of this neural network and image data presentation to the input layer.

![Figure 6. Global pattern averaging and neural network design](image)

The approach within this case study is based on simulating the “glance” and “familiarity” of faces in humans. The glance effect is approximated via image pre-processing and global pattern averaging as described in (section 4.3), whereas, familiarity of a face is simulated by training the neural network using face images with different orientations.

The implementation of a neural network consists of training and testing. In this work a total of 90 face images (corresponding to 30 persons) were used. For training the neural network 60 face images (looking left LL and looking right LR) were used. The 30 remaining face images (looking straight LS) were used for testing purposes where the system is expected to recognize the person looking straight at the camera by training it on face images looking left and right. This simulates the familiarity of a face in machines, even though the test images (looking straight) present a neural network with different pixel values as a result of the difference in the orientation of the face.

A recognition system “sensitivity” feature was also developed as part of the neural network classification of input face images. Three levels of tolerance, namely Low (minimum 80% face resemblance), Medium (minimum 65% face resemblance) or High (minimum 50% face resemblance) can be used depending on the required level of accuracy. The results that are presented in the next section were obtained using Low tolerance (i.e. minimum 80% face resemblance).
5. Results and Discussion

The back propagation neural network, within the intelligent system, learnt and converged after 4314 iterations and within 390 seconds, whereas the running time for the trained neural network after training and using one forward pass was 0.21 seconds. These time cost results were obtained using the following system configuration: 2.4 GHz PC with 256 MB of RAM using Windows XP operating system, C-language source code and Borland C++ compiler. Table 1 lists the final parameters of the successfully trained neural network.

All training images (60 face images - looking left and right) were correctly recognized when used for testing the trained neural network yielding 100% recognition rate as would be expected. The recognition of the testing face images (30 face images – looking straight) indicates the success and robustness of this intelligent system, as these face images had not been presented to the neural network before. Additionally, the “look straight” face images have different orientation and, thus, different pixel values in comparison to the training face images “look left” and “look right” at similar coordinates. Testing the neural network using these different test images yielded a successful 96.67% recognition rate where 29 out of 30 faces were correctly identified.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Layer Nodes</td>
<td>100</td>
</tr>
<tr>
<td>Hidden Layer Nodes</td>
<td>99</td>
</tr>
<tr>
<td>Output Layer Nodes</td>
<td>30</td>
</tr>
<tr>
<td>Bias Neurons Value</td>
<td>1</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>0.008</td>
</tr>
<tr>
<td>Momentum Rate</td>
<td>0.32</td>
</tr>
<tr>
<td>Minimum Error</td>
<td>0.002</td>
</tr>
<tr>
<td>Initial Random Weights Range</td>
<td>-0.3 to +0.3</td>
</tr>
<tr>
<td>Iterations</td>
<td>4314</td>
</tr>
<tr>
<td>Training Time (seconds)</td>
<td>390</td>
</tr>
<tr>
<td>Generalization/Run Time (seconds)</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Table 1. Trained neural network final parameters using global face data

Figure 7. Incorrect identification of Person 23 as person 21
The only incorrect result, out of the testing image set, was person 23 being identified as person 21. Both persons have close face resemblance, where a quick “glance” may not be able to distinguish. This incorrect identification occurred only when presenting the neural network with the face image looking straight (LS). Figure 7 shows both persons. Table 2 shows the recognition rates where a total recognition rate of 98.89% has been achieved.

<table>
<thead>
<tr>
<th>Image Set</th>
<th>Recognition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Set (60 images)</td>
<td>(60/60) %100</td>
</tr>
<tr>
<td>Testing Set (30 images)</td>
<td>(29/30) %96.67</td>
</tr>
<tr>
<td>Total (90 images)</td>
<td>(89/90) % 98.89</td>
</tr>
</tbody>
</table>

Table 2. Intelligent global face recognition results for 30 persons

In summary, the recognition process has two phases. First, simulating the quick look (glance) via image pre-processing which involves face image size reduction, cropping, segmentation and global pattern averaging. This phase yields segment pattern average values that are global representations of the face and consequently form the input to a neural network. The second phase (simulating familiarity) is training the neural network using the output of the first phase. Once the network converges or learns, classification and face recognition is achieved.

Further tests were carried out to investigate the effect of the presence or absence of structural components such as beards, mustaches or glasses on the recognition results. The effect depends on the differences in pixel values due to the structural component. A large difference in pixel values would marginally change the averaged pattern value, whereas a small difference would cause a minimal change in averaged pattern values. This problem can be solved by updating the intelligent global recognition system with any changes to a face due to a structural component; in other words familiarizing the intelligent system with any changes to a face.

This problem was investigated by testing the trained neural network using face images of “person 3” wearing a dark hat, thus resulting in minimal changes to the averaged pattern value. The system was able to correctly recognize “person 3” with and without the hat (see figure 8).

Figure 8. Further recognition system tests: Person 3 with and without a hat: Recognized
On the other hand, two extra face images of “person 2” clean shaven and also with full beard were used for further tests. The intelligent system yielded “unknown person” result, thus requiring updating the recognition system with the new look of the person, after which “person 2” was correctly recognized (see figure 9).

Figure 9. Further recognition system tests: Person 2 with different looks: Unrecognized

Another interesting result of further tests was the correct recognition of person 20 and person 28, who happen to be identical twins (see figure 10). The intelligent systems recognized both persons without the need for further training. This demonstrates the flexibility of the developed system where face image database can be updated as required.

Figure 10. Further recognition system tests: Persons 20 and 28 are identical twins: Recognized

6. Conclusion

The recognition of a face that has been seen before is a natural and easy task that we humans perform everyday. What information we pick from a face during a glance may be mysterious but the result is usually correct recognition. Do we only look at features such as eyes or nose (local feature detection) or do we ignore these features and look at a face as a whole (global face recognition)? Many research works on face recognition attempt to answer these questions, however, one common concept that is shared by most methods is that the detection of a face requires facial information, which can be obtained locally (using local facial features such as eyes) or globally (using a whole face).
The diversity of the different methods and approaches is more evident when investigating
the development of artificially intelligent face recognition systems. These intelligent systems
aim to simulate the way we humans recognize faces, and the methods that are developed to
achieve such an aim are as diverse as our natural individual perception of faces.

This chapter presented a review of related works on face recognition in general and on
intelligent global face recognition in particular. The presented case study is based on using
global (complete face and background) data averaging and a neural network in order to
simulate the human “glance” and face “familiarity”.

The glance effect is approximated via image pre-processing and global pattern averaging.
When we humans have a quick look (glance) at faces, we do not observe the detailed
features but rather a general global impression of a face. This can be approximated by
averaging the face image instead of searching for features within the face. The averaged
patterns are representation of a face regardless of its expression or orientation. The quick
glance is followed by familiarity with a face, which is simulated by training a neural
network using face images with different orientations.

The neural network within the intelligent system learnt to classify the faces within 390
seconds, whereas the running time for the trained neural network was 0.21 seconds. These
time costs can be further reduced by using faster machines, which will inevitably occur in
the near future.

The implementation of the intelligent global face recognition system used 90 face images of
30 persons of different gender, age and ethnicity. A total recognition rate of 98.89% was
obtained using 90 face images (combining training and testing images) of the 30 persons in
different orientations. Only one person’s face image (looking straight) was mistaken for
another person (looking straight too) as shown in Figure 7. The robustness and success of
this face recognition system was further demonstrated by its quick run time (one neural
network forward pass) of 0.21 seconds. Time cost was kept minimal through image-pre-
processing and reduction of input/hidden layer neurons in the topology of the neural
network.

Further tests of the trained neural network within the intelligent system investigated the
effects of the presence or absence of structural components such as beards or hats on the
recognition results. The outcome of these tests suggests that some of the “familiar” faces
may not be recognized if there is a drastic change on the face, this is due to the large
difference in pixel values which would marginally change the averaged global pattern
values. On the other hand, a small difference would cause a minimal change in averaged
global pattern values, and thus would not affect the recognition results. This problem can be
solved by updating the intelligent global recognition system with any changes to a face due
to a structural component; in other words familiarizing the intelligent system with the new
look of a person.

Additionally, three levels of tolerance can be used when implementing the system that was
presented in the case study. The choice of the tolerance level depends on the required level
of accuracy: low tolerance (80% face resemblance), medium tolerance (65% face
resemblance) or high tolerance (50% face resemblance). All results that were shown in the
case study on intelligent global face recognition were obtained using the low tolerance
classification, where a minimum of 80% face resemblance was required. This is believed to
be a good resemblance ratio considering the neural network is trained using globally
averaged patterns of faces and backgrounds.
Further work on intelligent global face recognition will investigate the acquisition of facial data using only left and right portfolios of a face. The developed neural network for this task would use different parameters obtained using other methods (e.g. edge detection) in order to associate the face with the person.

Finally, despite successful implementations of artificial intelligent face recognition systems such as our case study, there are questions that are yet to be answered before we can completely trust a machine whose intelligence “evolves” in minutes in comparison with our natural intelligence that took thousands of years to evolve. There is no doubt that the advancement in technology provides us with the means to develop efficient artificially intelligent systems, however, the question remains: how intelligent are they really are?

7. References


This book will serve as a handbook for students, researchers and practitioners in the area of automatic (computer) face recognition and inspire some future research ideas by identifying potential research directions. The book consists of 28 chapters, each focusing on a certain aspect of the problem. Within every chapter the reader will be given an overview of background information on the subject at hand and in many cases a description of the authors’ original proposed solution. The chapters in this book are sorted alphabetically, according to the first author's surname. They should give the reader a general idea where the current research efforts are heading, both within the face recognition area itself and in interdisciplinary approaches.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: