

# **A survey of Pattern Recognition algorithms and the link with facial expression detection**

---

**Business Mathematics and Informatics Paper**

Hassnae Belkasim

**Business Mathematics and Informatics Paper**  
VU University Amsterdam  
Faculty of Sciences  
De Boelelaan 1081  
1081HV Amsterdam  
The Netherlands

## **A survey of Pattern Recognition algorithms and the link with facial expression detection**

---

Author: Hassnae Belkasim

Supervisor: dr. Zoltan Szlavik

## Abstract

In the last decades, the concept and applications of face recognition have become a popular item in different branches. A considerable amount of research has been conducted within this area especially because of the rise of biometry techniques use for security aims and credit cards for instance and even in some social networks. The purpose of this paper is to mainly cover the mathematical aspect and idea behind two algorithms designed for Face Recognition, and to make some comparison on the efficiency of those algorithms in terms of accuracy. The Face Reader model will be also discussed in detail. The scope of this paper also includes the relationship between face- and expression detection and the feasibility of those algorithms in the context of Automatic Expression Analysis.

**Keywords:** Face recognition, expression detection, feature extraction, PCA, LDA, AAM, FaceReader

## Preface

Writing a research paper is a compulsory part of the Business Mathematics and Informatics master at the VU University. The essential purpose of this paper is providing a survey of the algorithms used to detect patterns and facial expressions.

I would like to take this opportunity to thank my supervisor for his patience and support.

Octobre, 2012  
Hassnae Belkasim

## Summary

Pattern recognition has become one of the popular research issues in the data mining field. This is due to its different applications. This paper tries to treat the following questions: How do recognition algorithms work? How are face detection techniques linked to the automatic expression analysis field? And what are the success factors as well as the limitations that can impact the accuracy of the recognition process?

There are several studies done on how to increase the recognition rate under different conditions of lighting, shading, profiling, and orientation. Many of them show that preprocessing by extracting features the original data images has a strong impact on the obtained accuracy. Principal Component Analysis (PCA) is an efficient method to accomplish this task. Linear Discriminant Analysis (LDA) is a handy technique that is based on linearity; however it depends strongly on the first stage related to the trained data. It is suggested therefore by (Weng et al) to combine both methods in order to reach higher accuracy. PCA is also suitable for preprocessing images data for facial expression detection. This is successfully used in the Face Reader model for detecting common universal face expressions, with which the overall accuracy was 89%.

## Contents

Abstract .....	3
Preface.....	4
Summary .....	5
1. Pattern Recognition .....	7
2.1 Background.....	7
2.2 Face detection .....	7
2.3 Expression detection .....	8
2. Approaches and algorithms facial recognition .....	9
2.4 Principal Component Analysis .....	9
2.4.1 Procedure and mathematical aspect.....	9
2.5 Linear Discriminant Analysis .....	11
2.5.1 Procedure and mathematical aspects .....	11
2.5 Comparison.....	12
3. Face Reader Model for expression detection .....	13
3.1 Face defining .....	13
3.2 Face modeling.....	13
3.3 Face classification.....	14
4. Conclusions .....	15
References .....	16

# 1. Pattern Recognition

## 2.1 Background

A considerable amount of research has been carried out around recognition systems for human faces. However the two areas of faces recognizing processes and facial expressions detecting processes have been in nearly all recent studies studied separately.

## 2.2 Face detection

While developing or even classifying techniques for recognizing faces, we can look at them from different perspectives. One option is that we can classify them into these two categories according to [8]:

- Geometric: This perspective is structure-based and focuses on characterizing features included in the images data
- Photometric: This is a statistical approach of converting images into values and trying to find matches with standard templates. This approach is more effective compared to the geometric one in terms of accuracy

Besides of that, there is a couple of misleading factors that can have a serious impact on the output accuracy of the recognizing process. Those factors can be summarized under the following categories:

- Orientation and Profiling: head poses. Frontal and profile images could mislead during the process and give different classification results.
- Thatcherization: Thatcher illusion or inverted faces explained in [10]. It often goes unnoticed - that's where the face is upside-down but the eyes and the mouth are changed to remain the right way up.
- Shading, pigmentation and lighting conditions
- Facial Expression: (un)smiling faces, etc..
- Noisy data: too much data and high dimensionality that is mostly caused by the high variation of the features

To limit the negative impact of these factors, the following steps are required:

- a. The first step in each pattern recognition process is to preprocess the data and to extract features that are needed later for the classification task. Feature extraction is a form of reducing dimensionality by transforming some given input data of images to a representation set of features, mostly under the form of vectors. This is required to reduce the commutation time and the high variation of information in the data. This can help in case of large data size and when suspicion of noise and redundancy is high. Required in the features set is that it should include relevant information to fulfill the task using it instead of the whole data set. This can be achieved using PCA explained in chapter 2.
- b. Original pattern should be also transformed into some representation that can be easily manipulated programmatically. This step depends on the structure of the data set
- c. Classification: matching data based on measurements similarities with other patterns, mostly via an artificial neural network

### 2.3 Expression detection

Facial expressions are the most important way of non-verbal communication. Human faces are multi-dimension senders and receivers of signals essential for interaction, and detecting those signals can have great advantages in medical, psychological and social contexts.

One of the most famous and successful works in this area is the study of Ekman Paul that resulted in the so called Facial Action Coding System (FACS). This sophisticated system is based on studying different combinations of face muscles producing various expressions. The FACS measurements are defined by nearly 46 Action Units (AUs) which consist about one or more muscles that are necessary for a given facial action. Besides that, Ekman makes a clear distinction between emotion and expression and explains also the relationship between the two notions in [5, 7]. The FACS study is a generalization of the studies based on common universal expressions like the FaceReader model. This FaceReader model will be explained in chapter 3.



## 2. Approaches and algorithms facial recognition

### 2.4 Principal Component Analysis

This method has been developed in 1991 and was based upon Karhunen-Loeve's transformation. The effectiveness of PCA has been studied extensively in recent years by many researchers. The authors Zhang and Zhou have conducted a very interesting research about two-dimensional PCA and obtained a recognition rate of nearly 90%. The results are represented and explained in [12]. In this paper, the focus will be on the one-dimensional PCA as higher-dimensional PCA's are generalizations of the same basic idea.

The key idea of PCA is to transform the face images into a set of characteristics feature images called eigenfaces. The eigenfaces are called the principal components of the initial training set of the face images. PCA tries to maximize the total scatter of all classes [2].

#### 2.4.1 Procedure and mathematical aspect

In the study of Thakur & Sing [2], a training set of images, consisting about a multi-dimensional vector representation of some human face is considered. PCA is used in order to reduce the dimensionality by finding some multi-dimensional subspace where the basic vectors are corresponding to the maximum variance direction in the original data set. Those basic vectors are defined as eigenvectors of the scatter matrix and the image elements are random variables.

A training set consisting of  $R$  face images  $X = (X_{i1}, X_{i2}, \dots, X_{iD})$  is assumed, where each image  $X_i$  of size  $m \times n$  can be represented as a vector of  $D$  pixels, where  $D = m \times n$ . The covariance matrix is given by:

$$Cov = \frac{1}{R} \sum_{i=1}^R (X_i - X_{av})(X_i - X_{av})^T$$

Where:  $Cov$  is a square matrix and  $X_{av} = \frac{1}{R} \sum_{i=1}^R X_i$  is the mean image and the eigenvalues and the eigenvectors can be calculated from the covariance matrix. If we consider a set of vectors  $Q$  being the eigenvectors that correspond to  $r$  eigenvalues, where  $0 < r < R$ , then each of the  $r$  eigenvectors is an eigenface.

Thus the projection of the original face images into the eigenface space gives the eigenface-based features:

$Z_i = Q^T Y_i / i = 1, 2, \dots, R$  and  $Y_i$  is the mean-subtracted image of  $X_i$ . After applying PCA for feature extraction, the classification task has been carried out by using the eigenfaces that were resulted in the previous step as inputs for the three-layer RBF neural network. The RBF (Radial Basic Function) neural network is a feed-forward network and forms an appropriate classifier tool in this case as it possesses a simple structure and good learning abilities. The detailed features of an RBF network are explained among others in this study [11].

The used RBF network is illustrated in the following figure:

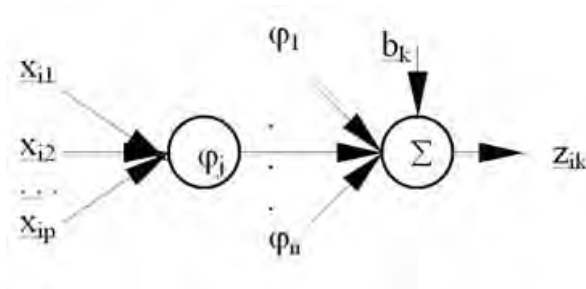


Figure1: Structure of a hidden and output layer of a neuron of the RBF-NN [2]

This study points out that the performance of this RBF network strongly depends on the training procedure. It has also shown that the highest performance is achieved when trying 60 principal components and 120 hidden layer neurons. The recognition rate was measured by looking to the ratio of the number of correct recognitions by the classifier to the number of the test set images for each run, thus:

$$R_{avg} = \frac{\sum_{i=1}^q n_{corr}^i}{q * n_{tot}}$$

Where:  $q$  is the number of runs

$n_{corr}^i$  is the number of correct recognition in the  $i^{th}$  run

$N_{tot}$  is the total number of faces under test in each run

The performance of this method has been tested under different conditions of rotation and scaling. One of the better results was achieved when trying this method on a database of 400 images using 400 runs, very high success rates were achieved; 97.75% and 99.94%

## 2.5 Linear Discriminant Analysis

### 2.5.1 Procedure and mathematical aspects

This analysis is mainly based on the linear discriminant functions under the form:

$$g(x) = w^T x + w_0$$

For a two-dimensional case, we classify  $x$  in class 1 if the function is positive and in class 2 otherwise.

So the following decision rule holds:

$$\begin{cases} w_1, & \text{if } g(x) > 0 \\ w_2, & \text{if } g(x) < 0 \end{cases}$$

The special case  $f(x)=0$  represents geometrically the decision line/surface that determines the two sub-spaces  $R_1$  and  $R_2$  as illustrated in this figure:

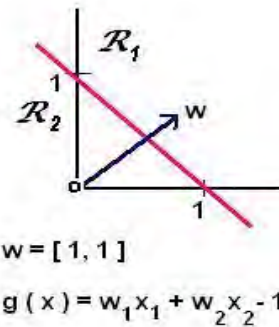


Figure2: The linear boundary separation for a 2-dimensional case [7]

As a generalization of the function above, we can formulate  $g(x)$  as:

$$g(x) = w_0 + \sum_{i=1}^d w_i x_i + \sum_{i=1}^d \sum_{j=i}^d w_{ij} x_i x_j$$

Where the third term represents additional term. The value of the LD function depends on the signed distance from the  $x$ .

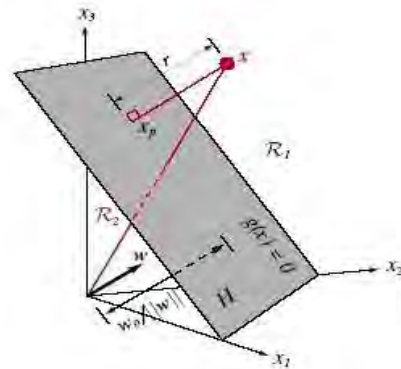


Figure3: The boundary separation in a multi-dimensional case [7]

## 2.5 Comparison

Both PCA and LDA try to automate recognition process. While PCA focuses on transforming the input faces images into reduced matrixes, LDA provides linear classification of trained input. It has been shown in [2] that it leads to a better recognition rate. However it is hard to deny that, during this transformation, some risk of losing information does exist as a direct consequence of reducing the within-class variance.

There is even a suggestion to combine the two techniques in one process in a study carried out by (Weng et al). He made such a suggestion because he has noticed that LDA does not perform well in the following cases:

- Human faces that are not included in the training set
- Different samples of trained classes
- Samples with different background

### 3. Face Reader Model for expression detection

“The core problem of face analysis is how to simultaneously account for the three major source of variance in face images: pose/orientation, expression and lighting.”[1]

Several methods have been developed in this context, basing on images or video fragments. One of the simple, but also efficient, ones is the image-based FaceReader model that consists of three main steps:

#### 3.1 Face defining

In this stage, a basic and neutral facial position should be found. This can be done using the Active Template Method (ATP) [6]

#### 3.2 Face modeling

In this step, a method called The Active Appearance Model (AAM) is used. The main advantage of AAM is that it is able not only to treat frontal images but also face images under different conditions like variations in orientation, lighting and emotions. This model uses, according to [6] a set of images to find out the sources of variance and tries to produce an artificial template, considered as mean faces, that points:

- The location of the facial key components
$$\mathbf{S} = ((x_1, y_1), (x_2, y_2), \dots, (x_M, y_M))^T$$
- The texture of the face in a low dimensionality using the PCA (treated in chapter 4) to reduce its dimensionality

Normal distribution is assumed and new input of faces can be considered as a deviation of the mean face which is predetermined by a vector called “appearance vector”. The image in the middle with value 0 is the neutral position and the images within  $[-3\lambda, 3\lambda]$  are the deviated positions. See figure 4:

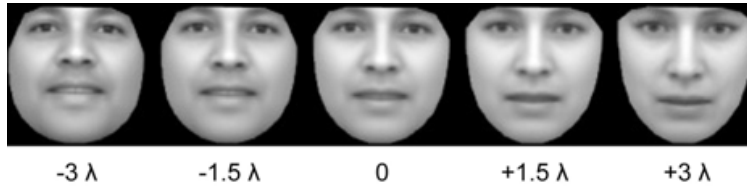


Figure 4: Mean face and deviated faces from the appearance vector [6]

### 3.3 Face classification

This is the final stage within the process according to the FaceReader model. With a considerable amount of data we can train the network to learn any facial expression that is included in the synthesized images. The input that should be trained is the constructed AAM appearance vector using the images of which the variation of features has been reduced. The characteristics that should be distinguished are the neutral position, being the mean face, and six basic emotional expressions, which are also known by “universal common expressions”. The following matrix represents the results obtained in the study of den Uyl & van Kuilenburg:

	<i>Happy</i>	<i>Angry</i>	<i>Sad</i>	<i>Surprised</i>	<i>Scared</i>	<i>Disgust</i>	<i>Neutral</i>	<i>Recall</i>
Happy	<b>138</b>	0	1	0	0	0	1	0.99
Angry	1	<b>116</b>	2	1	3	11	0	0.87
Sad	3	4	<b>109</b>	19	2	1	1	0.78
Surprised	0	1	6	<b>128</b>	0	0	0	0.95
Scared	0	8	5	2	<b>115</b>	5	3	0.83
Disgust	1	5	3	0	3	<b>125</b>	0	0.91
Neutral	0	11	2	1	1	0	<b>125</b>	0.89
Precision	0.97	0.8	0.85	0.85	0.93	0.88	0.96	<b>0.89</b>

Table1: Classification results for universal facial expressions [6]

The rows represent the real expressions and the columns the predicted expressions. The values on the matrix diagonal are the correct classified expressions and the overall recognition rate is 89% [6].

This study is carried out basing on the so called universal face expressions. However the work of Paul Ekman, in which he has studied 15 different facial expressions, is much more sophisticated and detailed. His article about the widely used FACS, which is mentioned at [3], can help us in understanding his study.

## 4. Conclusions

- No one best recognition algorithm
- Recognition rate depends on different factors: extraction method, number of hidden layer of the used network, the classification tool and also on how the success rate is measured
- PCA has a significant positive effect on the recognition rate for both face- and expression detection
- LDA is not suitable in all conditions and requires input data that is very similar to the trained data set
- Preprocessing techniques are as important as classifier methods and this is due to the different variation in the images input in real life cases
- Face recognition and expression detection are independent research areas. However there are many common features especially in the pre-processing stage

## References

- [1] Proceedings of Measuring Behavior (Wageningen, 30 August – 2 September 2005) by Eds. L.P.J.J. Noldus, F. Grieco, L.W.S. Loijens and P.H. Zimmerman
- [2] Face Recognition using Principal Component Analysis and RBF Neural Networks by S. Thakur, J. K. Sing, D. K. Basu<sup>2</sup>, M. Nasipuri, M. Kundu
- [3] The Facial Action Coding System by Ekman, P.; Friesen, W.V; Hager, J.C (2002). Weidenfeld & Nicolson, London
- [4] The Inner And Outer Meanings Of Facial Expressions (1983) by Joseph C. Hager and Paul Ekman
- [5] The Nature of Emotion - Fundamental Questions (Oxford University Press, New York, 1994) by P.Ekman and R.J. Davidson
- [6] The FaceReader: Online Facial Expression Recognition (2005) by M.J.den Uyl and H.van Kuilenburg
- [7] Universal Facial Expressions of Emotion (California Mental Health Research Digest) by P.Ekman
- [8] Pattern Recognition using Neural networks (December 2010) by Santaji Ghorpade, Jayshree and Shamla Mantri
- [9] Neural Networks and Pattern Recognition (2004) by Anna Bartkowiak
- [10] Margaret Thatcher: a new illusion (Perception, 1980, volume 9) by Peter Thompson
- [11] An Efficient MDL-based construction of RBF networks (Neural Networks 11, 1998, pages 963-973) by Ales Leonardis and Horst Bischof
- [12] Two-directional two-dimensional PCA for efficient face representation and recognition by Daoqiang Zhang and Zhi-Hua Zhou