Facial Expression Recognition using Local Binary Patterns

with classification based on Support Vector Machines

AALBORG UNIVERSITY
Department of Electronic Systems
Vision, Graphics and Interactive Systems
9^{th} Semester project

Autumn 2012
Maxime Coupez
Kim-Adeline Miguel
Julia Alexandra Vigo

AALBORG UNIVERSITY

Department of Electronic Systems
Fredrik Bajers Vej 7
DK-9220 Aalborg Ø
http://es.aau.dk

Title:
Project Title

Theme:
Interactive Systems

Project Period:
Fall Semester 2012

Project Group:
12gr942

Participant(s):
Maxime Coupez
Kim-Adeline Miguel
Julia Alexandra Vigo

Supervisor(s):
Zheng-Hua Tan

Copies: 1

Page Numbers: 27

Date of Completion:
November 9, 2012

Abstract:

Since the last decade, a lot of researches have been carried out about emotion recognition. The number of projects conducted in this field demonstrates the interest and the importance of systems which can recognize human mood.
In this project, an emotion recognition system is developed, using a Microsoft Kinect. This recognition is achieved in 3 steps: Face detection, extraction and classification of facial features, this structure being the usual modus operandi in emotion recognition research.
Face detection is performed using Viola-Jones' algorithm, then Local Binary Patterns (LBP) are used to extract facial features. Finally, Support Vector Machines (SVM) classify these features into six predefined emotions.
The system is implemented to run on a computer using a Kinect and works for one person in front of it. The classifier is trained with the Cohn-Kanade database, which includes enough different faces to obtain a satisfying result.

Preface

This report documents the semester project entitled *Facial expression recognition using Local Binary Patterns*. The project was carried out during the 9th semester of specialization *Vision, Graphics, and Interactive Systems* under the Department of Electronic Systems at Aalborg University in Autumn 2012.

The report is divided into four parts plus appendices: *Introduction*, *Feature Detection*, *Feature Classification*, *Implementation* and *Evaluation*. The first part review the general structure of a facial expression recognition system and its main issues, and concludes with a state of the art of existing systems. Analysis of possible solutions and design of our system are contained in the following two parts, and the fourth part describes our implementation. The last part evaluates the performance and accuracy of our system and concludes on the project as a whole.

References to secondary literature sources are made using the syntax [number]. The number refers to the alphabetically sorted bibliography found at the end of the report, just before the appendices.

We would like to thank our supervisor at Aalborg University Zheng-Hua Tan for supporting us in this challenging project.

A CD is attached to this report which includes:

- Source code of the developed program.
- PDF file of this report.

Aalborg University, November 9, 2012

Maxime Coupez
<mcoupe12@es.aau.dk>

Kim-Adeline Miguel
<kmigue12@es.aau.dk>

Julia Alexandra Vigo
<jvigo12@es.aau.dk>

Contents

Part I

Introduction

Contents

The main motive of this project is to understand real-time facial expression recognition systems and their applications. Then a review of the architecture of such systems will be done, along with a state of the art of already existing algorithms for each step. After this study, one algorithm for each step will be chosen, considering their difficulty and their efficiency, in order to improve it. In the last part, the problem will be formulated.

Chapter 1

Motivations

A facial expression is a visible manifestation of the effective state, cognitive activity, intent, personality, and psychopathology of a person [6]; facial expressions play a significant role in human dialogue and interaction. Indeed, facial expressions carry more informations than mere speech, informations on which humans can relay for interaction. Facial expressions have a considerable effect on a listening interlocutor; a speaker facial expressions accounts for about 55 percent of the effect, 38 percent of the latter is conveyed by voice intonation and 7 percent by the spoken words [12].

Since Antiquity, researchers have been interested in emotion and more particularly in emotion recognition. But one of the important studies on facial expression analysis impacting on the modern day science of automatic facial expression recognition was the work carried out by Charles Darwin [3]. In 1872, Darwin wrote a treatise that established general expression principles and expression means for both humans and animals [5]. He also classified various kinds of expressions. This can be considered as the beginning of facial expression recognition.

Now, with the emergence of new technologies and computers, research is now focused on computer-based automatic facial expression recognition. Because facial expressions are major factors in human interaction, this research field will broaden the domain of Human-Machine Interaction. Indeed, emotion recognition will enable computers to be more responsive to users' emotions, and allow interactions to become more and more realistic.

Another domain where facial expression recognition is an important issue is robotics. With the advances made in robotics, robots nowadays tend to mimic human emotion and react as as human-like as possible, especially for humanoid robots. However, since robots are being more and more present in our daily lives, they need to understand and recognize human emotions.

A lot of real time applications in the robotics field have already been created. For example, Bartlett et al. have successfully used their face expression recognition system to develop an animated character that mirrors the expressions of the user (called CU Animate) [2]. They have also been successful in deploying the recognition system on Sony's Aibo Robot and ATR's RoboVie [2]. Another interesting application has been demonstrated by Anderson and McOwen, called "EmotiChat" [1]. It is a regular chatroom, except the fact that their facial expression recognition system is connected to the chat and convert the users' facial expressions into emoticons. Because facial expression recognition systems' robustness and reliability are constantly increasing, lots of innovative applications will appear.

There are also various other domains where emotion recognition can be used: Telecommunications, behavioural science, video games, animations, psychiatry, automobile safety, affect-sensitive music jukeboxes and televisions, educational software, etc [3].

Our project focuses on real-time facial expression recognition from a video stream. Indeed, facial expression recognition can be performed *statically* on input images, or *dynamically* on video sequences. Systems can also be *obtrusive*, or *non-obtrusive*, the former based on a device mounted on the user's head or body, therefore following each of his movements and perform facial expression recognition without much losses, while the latter can encounter difficulties if the user is not properly situated. However, non-obtrusive systems allow more natural user interactions. We chose our system to be non-obtrusive, and will detail further its setup in the next section.

1.1 Environment Setup

Our system will use the camera embedded into a Microsoft Kinect to record the user's video input, and we will consider a casual use of the camera, with the user sitting in front of the computer, the camera being next to it, as seen in **Insert figure & ref to figure**. This camera provides a 640×480 pixels frame resolution, while recording at 30 FPS.

For development and training purposes we will use some pre-existing emotion datasets, in order to validate the efficiency of the system before testing it in real conditions.

1.2 Emotion Datasets

Databases are very important for facial expression recognition system.

Using the same databases for studies that aims to improve existing systems is very useful. It allows to compare he results and to see if the new system is indeed better than the existing one. A lot of research studies work is based on the same databases than previous studies in order to compare the efficiency of their algorithm.

But databases are hard to construct. It has to fill all the requirements and that is why most of the work on facial expression recognition is based on existing databases.

The hardest requirement to fill is to have a standardized database. Most of the actual databases use posed expressions and not spontaneous expression, and both are very different. New versions of the databases are coming out with spontaneous expressions in order to be more complete. Even with this transition from posed expressions to spontaneous expressions, there are other requirements that should be respected to have a database standardized for training and testing. It should contain images and video sequences and both should be of different resolutions. It should also contain people displaying expressions under different conditions: it could be change in the lighting, occlusions or rotations of the head [12].

Following are the databases that will be used to test this facial expression recognition system. These are part of the databases that are popular, freely available and mostly used in the past few years.

1.2.1 Japanese Female Facial Expression Database (JAFFE)

The database contains 213 images of 7 facial expressions (6 basic facial expressions: happy, angry, afraid, disgusted, sad, surprised + 1 neutral) posed by 10 Japanese female models. Each expression has been photographed three or four times. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. The database was planned and assembled by Miyuki Kamachi, Michael Lyons, and Jiro Gyoba [11].

This DB contains only posed expressions. The photos have been taken under strict controlled conditions of similar lighting and with the hair tied away from the face [12].

Following are an example of images contained in the database. Here the subject is a woman and she displays 7 different emotional expressions (neutral, happy, angry, afraid, disgusted, sad, surprised):

1.2.2 Karolinska Directed Emotional Faces Database (KDEF)

The Karolinska Directed Emotional Faces (KDEF) is a set of totally 4900 pictures of human facial expressions of emotion. The material was developed in 1998 by Daniel Lundqvist, Anders Flykt and Professor Arne Ohman at Karolinska Institutet, Department of Clinical Neuroscience, Section of Psychology, Stockholm, Sweden [9].

The material was originally developed to be used for psychological and medical research purposes. More specifically material was made to be particularly suitable for perception, attention, emotion, memory and backward masking experiments. Hence, particular attention was for instance paid to create a soft, even light, shooting expressions in multiple angles, use of uniform T-shirt colors, and use of a grid to center participants face during shooting, and positioning of eyes and mouths in fixed image coordinates during scanning [9].

The set contains 70 individuals (35 males and 35 females), ranging from 20 to 30 years, each displaying 7 different emotional expressions (neutral, happy, angry, afraid, disgusted, sad, surprised), each expression being photographed (twice) from 5 different angles (-90, -45, 0, +45, +90 degrees: i.e. full left profile, half left profile, straight, half right profile, full right profile) [9].

Following are an example of images contained in the database. Here the subject is a woman photographed from a straight angle and she displays 7 different emotional expressions (neutral, happy, angry, afraid, disgusted, sad, surprised):

1.2.3 Montreal Set of Facial Displays of Emotion Database (MSFDE)

The database consists of emotional facial expressions by men and women of European, Asian, and African descent. Each expression was created using a directed facial action task and all expressions were FCAS coded to assure identical expressions across actors [16].

The set contains expressions of happiness, sadness, anger, fear, disgust, and embarrassment as well as a neutral expression for each actor. All expressions have been morphed into 5 different levels of intensity [16].

Following are an example of images contained in the database. Here the subject is an african woman and she displays 7 different emotional expressions (neutral, happy, angry, afraid, disgusted, sad, ashamed):


SAD
ANGER
DISGUST
NEUTRAL
SHAME
HAPPY
FEAR

Chapter 2

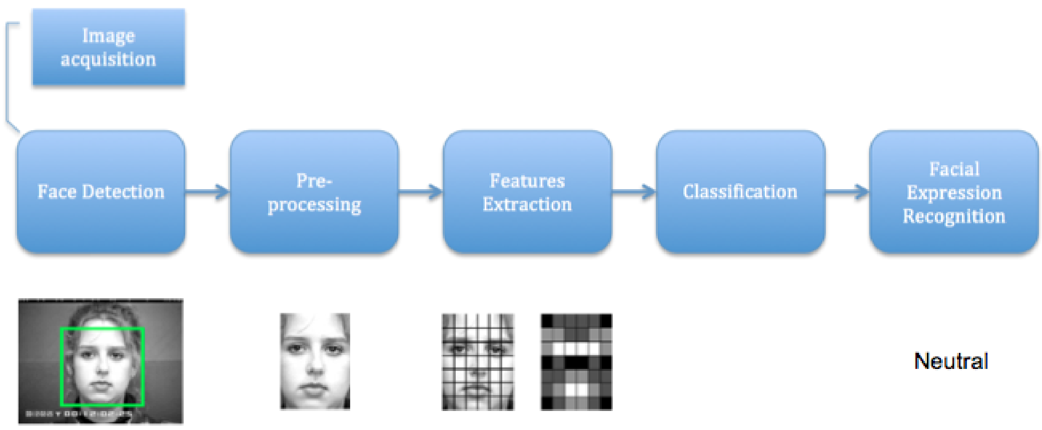
Facial expression recognition

Introductory paragraph of the chapter

2.1 General structure

Facial expression recognition is a system enabling an automatic recognition of emotions displayed by a human face. Facial expression recognition can be image or video-based; it can also be computed real-time. Most of the time, researchers try to recognize emotions out of images of human faces. This can also be achieved real-time on video streams : While the person displays his/her emotions, the facial expression recognition system analyses the video, and detect in real-time the displayed emotion.

In both cases, facial expression recognition process is structured as follows:



2.1.1 Image Acquisition

First step of the process is "Image Acquisition". Images used for facial expression recognition can be static images or image sequences. Image sequences give more informations about the facial expression, as the steps in muscles movement. About static images, facial expression recognition systems usually need 2D greyscale images as inputs. We can however expect future systems to use colour images; first because of the increasing affordability of technologies and devices capable of capturing images or image sequences; then because colours can give more information on emotions, i.e blushing [4].

2.1.2 Face Detection

Second step is "Face Detection". Indeed, in a static image and even more in an images sequence, this is an obvious need. Once the face has been detected, all other non-relevant information can be deleted, since only the face is needed. It could hence be included in the next step, which is "Pre-processing", but because of its importance it represents a step in itself. In a real-time facial expression recognition system working with image sequences, the face has to be detected, but also tracked. One of the most used and famous detection and tracking algorithm is the Viola-Jones Algorithm, which we will explain in detail later in this report. This algorithm can be trained to detect all kind of objects, but is mostly used for face detection.

2.1.3 Pre-processing

Third step is "Pre-processing", which is about applying image processing algorithms to the image, in order to prepare it for the next step. Pre-processing is usually about noise removal, normalization against the variation of pixel position or brightness, segmentation, location, or tracking of parts of the face. Emotion recognition is also sensitive to transformation, scaling and rotation of the head in the image or image sequence. In order to solve this problem, the image can be geometrically standardized. References used for this standardization are usually the eyes [4].

2.1.4 Features Extraction

Once the image has gone through the "Pre-processing" step, the next one is "Features Extraction". In this step, data is converted into a higher representation of shape, motion, colour, texture, and spatial configuration of the face or its components. One of the main goals of this step is to reduce the dimensionality of the input data. The reduction procedure should retain essential information possessing high discrimination power and high stability [4]. There are a lot of features extraction methods. The most famous are : Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Problem Based Learning (PBL), Hidden Markov Models (HMM), Eigenfaces, Gabor Wavelets. The extracted data is then used in the "Classification" step.

2.1.5 Classification

INSERT PARAGRAPH ON CLASSIFICATION - explanation
- existing algorithms : PCA et LDA, mais aussi neural networks, svm

2.2 Existing systems

Before developing a facial expression recognition project, it is important to know what already exist; the state of the art of facial expression recognition system. In this chapter, an overview will be given of the existing systems before to decide on a system for the project. **Introductory paragraph on which methods are we going to focus on, appearance-based & geometry-based**

2.2.1 Principal Component Analysis (PCA)

This is a statistical method; one of the most used in linear algebra. PCA is mainly used to reduce high dimensionality of data and to obtain the most important information from this data. Because Facial Expression Recognition needs to reduce the dimensionality of data during features extraction, PCA is commonly used. It helps transforming high dimensionality of data to a new coordinate system of lower dimensions while still preserving the most important information. PCA computes a covariance matrix and a set of values called the eigenvalues and eigenvectors from the original data [7].

2.2.2 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis is also a statistical method, used to classify a set of objects into groups. It is done by observing a set of features that describe the objects. LDA as PCA are used to establish a linear relationship between the dimensions of the data. The main difference is that LDA uses the linear relationship to model the differences into classes of objects and PCA does not take any differences into account in the linear relationship. The idea is to perform a linear transformation on the data to obtain a lower dimensional set of features [7].

2.2.3 Local Binary Patterns (LBP)

This is an appearance-based method. It can be used to describe texture and shape. LBP extracts some informations from the neighbourhood of a central pixel. It compares the intensity values of the neighbourhood pixels with the intensity value of the central pixel [7]. This method is the one that will be used for this Facial Expression Recognition system.

2.2.4 Hidden Markov Models (HMM)

These models are a set of statistical models used to characterize the statistical properties of a signal [14]. It is developed to recognize expressions based on the maximum likelihood decision criterion [10].

2.2.5 Eigenfaces

Eigenfaces are a set of eigenvectors which are derived from the covariance matrix of a set of face images in a high-dimensional vector space. The eigenvectors are ordered and each one represents the different amount of the variation among the face images. It all together characterizes the variation between face images [15].

2.2.6 Gabor Filters

Gabor filters are applied in order to extract a set of Gabor wavelet coefficients. When convolving these Gabor filters with a simple face image, filter responses are obtained. These representations display desirable locality and orientation performance [8]. However, it has a limitation which is the processing time of Gabor feature extraction. It is very long and its dimension is prohibitively large [13].

2.3 Issues

bla bla bla

2.4 Requirements

bla bla bla

Part II

Feature detection

Contents

Bla bla bla

Part III

Feature classification

Contents

Bla bla bla

Part IV

Implementation

Contents

Bla bla bla

Part V

Evaluation

Contents

Bla bla bla

Conclusion

In case you have questions, comments, suggestions or have found a bug, please do not hesitate to contact me. You can find my contact details below.

Jesper Kjær Nielsen
jkn@es.aau.dk
http://kom.aau.dk/~jkn
Niels Jernes Vej 12, A6-302
9220 Aalborg Ø

Bibliography

[1] Keith Anderson and Peter W. McOwan. A real-time automated system for recognition of human facial expressions. *IEEE Trans. Systems, Man, and Cybernetics Part B*, 36(1):96–105, 2006.

[2] Marian Stewart Bartlett, Gwen Littlewort, Ian Fasel, and Javier R. Movellan. Real time face detection and facial expression recognition: Development and application to human computer interaction. *Proc. CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction*, 5:53, 2003.

[3] Vinay Bettadapura. Face expression recognition and analysis: The state of the art. *Tech Report*, 2012.

[4] Claude C. Chibelushi and Fabrice Bourel. Facial expression recognition: A brief tutorial overview, 2003.

[5] Charles Darwin. *The Expression of the Emotions in Man and Animals.* John Murray, 2. ed. edition, 1904.

[6] Gianluca Donato, Marian Stewart Bartlett, Joseh C. Hager, Paul Ekman, and Terrence J. Sejnowski. Classifying facial actions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(10):974–989, 1999.

[7] Abhiram Ganesh. Evaluation of appearance based methods for facial expression recognition, 2008.

[8] Yousra Ben Jemaa and Sana Khanfir. Automatic local gabor features extraction for face recognition. *International Journal of Computer Science and Information Security*, 3(1), 2009.

[9] Emotion Lab. Karolinska directed emotional faces (kdef). `http://www.emotionlab.se/resources/kdef`.

[10] Jenn-Jier James Lien. Automatic recognition of facial expressions using hidden markov models and estimation of expression intensity, 1998.

[11] Michael Lyons. The japanese female facial expression (jaffe) database.

[12] Maja Pantic and Leon J.M. Rothkrantzi. Automatic analysis of facial expressions: the state of the art. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1424–1445, 2000.

[13] Lekshmi V Praseeda and M Sasikumar. Facial expression recognition from global and a combination of local features. *IETE Tech Rev*, 26(1):41–46, 2009.

[14] Lawrence R. Rabiner and Biing-Hwang Juang. Fundamentals of speech recognition, 1993.

[15] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[16] UQUAM. The msfde. `http://www.er.uqam.ca/nobel/r24700/Labo/Labo/MSEFE.html`.

Appendix A

Appendix A name

Here is the first appendix