Entropy Based Feature Selection for 3D Facial Expression Recognition

Kamil Yurtkan

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy
in
Electrical and Electronic Engineering

Eastern Mediterranean University
September 2014
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Elvan Yılmaz
Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Doctor of Philosophy in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Hasan Demirel
Chair, Department of Electrical and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Philosophy of Doctorate in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Hasan Demirel
Supervisor

Examining Committee

1. Prof. Dr. Adnan Yazıcı
2. Prof. Dr. Fikret S. Gürgen
3. Prof. Dr. Hüseyin Özkaramanlı
4. Assoc. Prof. Dr. Hasan Demirel
5. Asst. Prof. Dr. Önsen Toygar

ABSTRACT

Human face is the most informative part of the human body that carries information about the feelings of the human. Recent improvements in computer graphics and image processing fields of computer science make facial analysis and synthesis algorithms applicable with the current digital Central Processing Units (CPUs). The information embedded to the human face can be analyzed with facial movements and mimics. The extracted parameterized data can be used in defining the facial expressions. Improvements in Human-Computer Interaction (HCI) systems have placed face processing research studies into a crucial stage in order to develop algorithms and applications. Therefore, facial expression recognition is an essential part of face processing algorithms.

The thesis presents novel entropy based feature selection procedures for person independent 3D facial expression recognition. The coarse-to-fine classification model and the expression distinctive classification model which are both based on Support Vector Machine (SVM) are used for the proposed feature selection procedures. Information content of the facial features is analyzed in order to select the most discriminative features which maximize expression recognition performance. Entropy and variance have been employed as information content metrics.

The input features are 3D facial feature points provided in MPEG-4 standard. A face is represented with 3D positions of geometric facial feature points. The feature selection algorithm selects the best feature points using novel entropy based method and represents the face with the selected points. Selections are done depending on Fisher's criterion. High entropy facial feature points maximizing Fisher's criterion are selected. The main contributions of the thesis are entropy based feature selections based on two different classifier models. The first one is a two-level coarse-to-fine classifier model and the second one is expression distinctive classifier model. For each model, entropy based feature selection is applied. Feature selection in two-level classifier model is accomplished in two levels. First, the best features are selected that classify the unknown input face into the one of the big expression classes, which are Class 1 and Class 2. Class 1 includes anger, disgust and fear expressions, where Class 2 includes happiness, sadness and surprise expressions. In the second level, the best features for each class are selected that classifies an expression into one of the three expressions presented in the selected class. As a result, three different feature models are proposed for the two-level coarse-to-fine classifier model. One feature model in order to classify into Class 1 and Class 2, and the two other feature models for each class's inner class classification processes. The second classifier model is the expression distinctive model in which entropy based feature selection method is applied to each expression specifically. Thus, the feature selection algorithm proposes six different feature models that maximize Fisher's criterion for each expression.

The proposed algorithms are tested in BU-3DFE and Bosphorus databases and the experimental results provide significant improvements on recognition rates. Proposed methods achieve comparable recognition rates for all of the six basic expressions which overcome the problem of having very high recognition rates for some of the expressions and unacceptable rates for some others, resulting in good average rates.

Keywords: Facial expression recognition, feature selection, face biometrics, entropy, information content.

ÖZ

İnsan yüzü, insan bedeninde kişinin duygu durumu hakkında bilgi taşıyan en önemli kısımdır. İçinde bulunduğumuz dönemdeki bilgisayar bilimi araştırmalarının sunduğu yenilikler ve hatırı sayılır gelişme gösteren bilgisayar grafikleri ile imge işleme algoritmaları, günümüz sayısal işlemcilerinin insan yüzünü işleyebilmelerine olanak tanımaktadırlar. Parametrelendirilmiş yüz hareketleri, yüz ifadelerinin analizi ve tanınması için kullanılabilmektedir. Temelinde İnsan Makine Etkileşimi (İME) olan uygulamaların gelişimi, insan yüzünün sayısal işlemciler tarafından işlenmesi gereksinimini çok kritik bir safhaya taşımıştır. İnsan yüzünde gömülü olan bilginin çıkarımı yüzdeki hareketlerin ve mimiklerin tespiti ile mümkündür. Bu nedenle, yüz ifadeleri analizi, yüz işlemeyi kullanan algoritmalar için vazgeçilmez bir kısım konumundadır.

Tezde, kişiden bağımsız yüz ifadeleri tanınmasına yönelik öznitelik seçimi için geliştirilen özgün yöntemler sunulmaktadır. Destek Vektör Makinesi (DVM) tabanlı iki farklı sınıflandırma modeli sunulmuştur. Bunlar kabadan inceye doğru sınıflandırıcı ve yüz ifadesine özel sınıflandırıcı modelleri olarak ikiye ayrılır. Yüz özniteliklerinin seçimi için önerilen yöntem her iki modele de uygulanmıştır. En ayrıştırıcı özniteliklerin seçimi için yüz ifadelerinin oluşumu esnasında öznitelikerin bilgi içeriği incelenmiştir. Bilgi içeriğinin ölçümünde entropi ve varyans ölçüm metrikleri olarak kullanılmıştır. Yüz ifadelerinin oluşumu esnasında en çok bilgiyi taşıyan ve tanınma başarısını geliştiren yüz öznitelikleri seçilmektedir.

Sistemin girdi öznitelikleri MPEG-4 standardında tanımlanan yüz öznitelik noktalarıdır. Yüz, bu geometrik noktaların 3 boyutlu konum bilgisinden temsil

edilmektedir. Özgün öznitelik seçim yöntemi, söz konusu öznitelik noktalarından entropiye göre seçim yapmakta ve seçilen özniteliklerle yüzü temsil etmektedir. Öznitelik seçimleri Fisher kriteri göz önünde bulundurularak yapılmıştır. Fisher kriterinin en büyük olduğu yüksek entropiye sahip noktalar seçilmektedir. Tezin iki ana katkısı öznitelik seçimlerinin iki farklı sınıflandırma modeline yönelik yapılması ve sonucunda farklı öznitelik modellerinin önerilmesidir. Birinci model iki seviyeli kabadan inceye doğru sınıflandırma modeli, ikinci model ise yüz ifadesine özel sınıflandırıcı modelidir. Öznitelik seçim yöntemi her iki modele farklı şekilde uygulanmıştır. İki seviyeli modelde öznitelik seçimi önce birinci seviye olan ve bilinmeyen yüz vektörünün iki büyük sınıfa ayrıldığı seviyede yapılmıştır. Bunlar Sınıf 1 ve Sınıf 2 olarak isimlendirildiğinde, Sınıf 1 içerisinde öfke, ,iğrenti ve korku ifadeleri, Sınıf 2 içerisinde ise mutluluk, üzüntü ve sürpriz ifadeleri bulunmaktadır. İkinci seviye için ise mevcut üç ifade arasında en ayrıştırıcı öznitelikleri bulmak için öznitelik seçimi yapılmıştır. Bu seviyede seçilen öznitelikler her bir sınıfın sınıf içi sınıflandırma başarısını artıracak şekilde yapılmıştır. Buna göre ilk seviye için bir ve ikinci seviyedeki herbir sınıf için öznitelik seçimi yapılmış, toplamda üç farklı öznitelik modeli önerilmiştir. Yüz ifadesine özel sınıflandırıcı modelinde ise entropiye dayalı öznitelik seçimi her bir temel yüz ifadesi için ayrı ayrı yapılmış ve sonuç olarak bu model için Fisher kriterinin en büyük olduğu altı farklı öznitelik modeli önerilmiştir.

Önerilen yöntemler BU-3DFE ve Bosphorus veritabanları üzerinde test edilmiş ve ümit verici sonuçlar elde edilmiştir. Önerilen yöntemlerde tüm temel yüz ifadelerinin yakın ve yüksek oranlarda tanınma başarısı gösterdiği gözlemlenmiştir. Mevcut bazı sistemlerde görülen bir problem olan belli yüz ifadelerinin çok yüksek, diğer yüz ifadelerinin ise başarıyla tanınması fakat ortalama tanınma başarısının yüksek olması,

önerilen metodlarda tüm yüz ifadelerinde yakın tanınma oranları elde edilerek aşılmıştır.

Anahtar Kelimeler: Yüz ifadeleri tanıması, öznitelik seçimi, entropi, yüz biyometrisi, bilgi içeriği.

ACKNOWLEDGMENT

I would like to express my gratitude to my thesis supervisor Assoc. Prof. Dr. Hasan Demirel for his continuous support and supervision throughout the whole thesis process. I would like to thank him for his significant contributions to the publications. Without his continuous encouragement and support, this work could have never been accomplished at all.

I also would like to thank to my family for their continuous support in all phases of my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / LIST OF ABBREVIATIONS

$\Delta\ i$ (x, y, z)	Displacement function of a vertex
W_i	Weight of a vertex i
W'	FAP influence weight
V_i	Vertex
V_{i_x}	X coordinate of a vertex
V_{i_y}	Y coordinate of a vertex
V_{i_z}	Z coordinate of a vertex
FV_j	J^{th} Face vector
FM	Face matrix
q	Number of vertices in BU-3DFE database
n	Number of face vectors in BU-3DFE database
r	Number of vertices in Bosphorus database
p	Number of face vectors in Bosphorus database
$w_k(x)$	Degree of belonging to a cluster
k	Number of clusters
C_k	k^{th} Cluster
s	Number of samples
μ	Mean
X	Random variable
$E[X]$	Expected value of random variable X
Var	Variance
$H(x)$	Entropy of event X
$P(x_i)$	Probability of even X_i
m	Mean

s^2	Variance
$\|S_B\|$	Determinant of between class scatter matrix
$\|S_W\|$	Determinant of within class scatter matrix
J	Fisher's criterion function
φ	Ratio of determinant of the scatter matrices (criterion)
δ_i	Magnitude of vertex i
$FV(\delta)$	Face vector of vertex magnitudes
$FM(\delta)$	Face matrix of *FV(D)s*
A	Set of bins used in probability $P(x_i)$
λ	Number of training samples used in probability $P(x_i)$
HCI	Human Computer Interaction
2D	2 Dimensional
3D	3 Dimensional
MPEG	Moving Pictures Expert Group
FA	Facial Animation
FDP	Facial Definition Parameter
FAP	Facial Animation Parameter
AU	Action Unit
FACS	Facial Action Coding System
SVM	Support Vector Machine
FCM	Fuzzy C-Means
PCA	Principal Component Analysis
ICA	Independent Component Analysis
GW	Gabor wavelet
LDA	Linear Discriminant Analysis
LBP	Local Binary Patterns

BU-3DFE	Binghamton University 3D Facial Expression Database
HMM	Hidden Markov Model
ANN	Artificial Neural Network
MLP	Multi-layer Perceptron
PNN	Probabilistic Neural Network
CK	Cohn - Kanade
FRGC	Face Recognition Grand Challenge
NSGA-II	Non Dominated Sort Genetic Algorithm II
PSO	Particle Swarm Optimization
FS	Feature Selection

Chapter 1

INTRODUCTION

Developments of Human-Computer Interaction (HCI) systems together with recent advances in computer graphics and image processing make it possible to develop algorithms for face processing. Face processing studies can be categorized into two main categories which are facial analysis and facial synthesis. The analysis studies are mainly focused on the processing of facial images for face detection, face recognition, face tracking and facial expression recognition. Two of the key elements in facial analysis is the step of facial feature extraction and selection processes which affect the overall performance of the system. Whereas, the synthesis studies are concentrated on face modeling, facial animation and facial expression synthesis. Figure 1 shows research directions on face processing.

Since early 1990s, there have been a lot of researches in face processing research fields with the continuous improvements in image processing capabilities of digital computers. Facial synthesis studies gained acceleration with the improvements in computer graphics. High speed graphics processing capabilities make it possible to synthesize a face in high resolution. One of the milestones in facial synthesis studies is the face model Candide, which is proposed by Stromberg [1]. Candide parameterized face model is still popular in many research labs.

While facial synthesis studies are based on accurate face modeling, facial analysis studies are focused on face detection, face tracking and facial feature extraction and selection. One of the most important parts of facial analysis studies is the accurate face detection. A popular algorithm is the face detector introduced by Viola & Jones [2]. Research directions in face processing are shown in Figure 1.1.

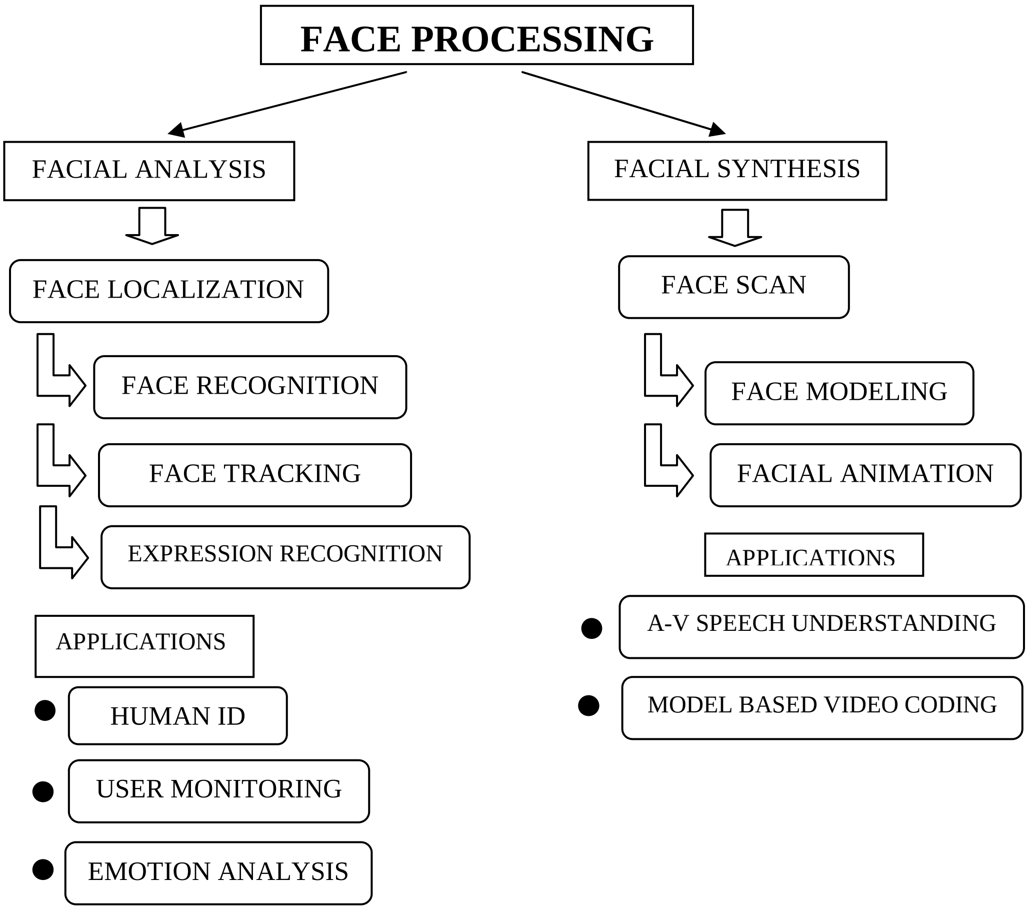


Figure 1.1: Face processing research directions.

Human face contains most of the information about the feelings of a person and human-computer interaction highly depends on accurate facial analysis. This information can be expressed as facial movements and facial expressions. Therefore, these aspects of the face have been parameterized by earlier systems in the literature. The two popular systems that parameterize facial movements are the Facial Action Coding System (FACS) and Action Units (AUs) [3, 4] and Moving Pictures Expert Group’s MPEG-4 Facial Animation Parameters (FAPs) [5] in the literature. An important milestone in 1970s in facial expression research was the studies of Paul Ekman and his colleagues [3, 4]. Ekman's research findings were about the classification of facial expressions into seven basic categories namely anger, disgust, fear, happiness, neutral, sadness and surprise. Later, Ekman and Friesen developed Facial Action Coding System (FACS) to code facial expressions and facial movements which they called action units [4]. Most of these parameters can be utilized in determining facial expressions of the faces. The thesis study is built on Ekman's findings about the classification of facial expressions and employs 3D geometrical facial features defined in MPEG-4.

1.1 Problem Definition

Since the early 1990s, many studies related to facial expression recognition have been published. The approaches differ according to the feature extraction method used, person dependency and classifier design. On the other hand, facial representation is also an important part of a facial expression recognition system. The face can be represented using texture information, 2D or 3D geometry, or the fusion of both. Besides facial representation, the feature extraction and feature selection processes are also vital for a successful expression recognizer [6].

Recently, facial expression recognition studies have been growing. Current research activity in facial expression recognition is focused on automatic facial expression recognition and has achieved acceptable recognition performances under controlled conditions. There are still several challenges in facial expression recognition studies on which research is focused. One of the most important challenges is to select the most discriminative features and represent the face accordingly for expression classification. The facial representations used in the expression recognition studies are general representations for a face. However, not all the facial features carry information when the face is deformed for an expression. For example, the centre of the forehead is almost static in all facial expressions and animations. Thus, using all the features presented in facial descriptors may confuse the classifier with similar feature spaces over different expressions. Using redundant features may mislead the classifier. The problem here can be celarly stated as a feature selection problem which is the method of finding the most discriminative facial features. It should be done specifically under facial expressions in order to provide facial features that are the best features discriminating facial expressions. Also, considering the spontanous behaviour of facial expressions, the real-time feature extraction process is an important challenge for the current systems. Extraction and tracking of facial features in order to analyze facial movements should be done in real-time. Considering the computational complexity of a real-time facial feature tracker, the extraction and tracking of unnecessary features will decrease the system performance. A successful feature selection process relaxes the feature extraction part so that not all the feature are tracked in real-time, only the most discriminative features are extracted and tracked. Redundant facial features are omitted in the feature exraction part. Other challenges can be listed as follows:

- Facial expression categories can be extended further from the six basic expressions and those expressions also should be recognized.
- Head rotations and different views of the face (angles) affect automatic recognition process.
- Recognition of spontaneous expressions.
- Real-time facial feature detection for expression recognition. Most expressions are recognized with a very high rate of accuracy and a few with

low rates of accuracy so the overall recognition rate is fairly high, not all expressions are being accurately recognized [7].

In the thesis study, the problem of feature selection for improved facial expression recognition has been focused. Information content of the facial features has been analyzed in order to find the most discriminative features. Variance and entropy have been used as the metrics of information content. The representations of faces are done by using 3D facial geometry including 3D positions of and 3D distances between the facial feature points which are described in Chapter 3 in details.

1.2Thesis Objectives

The thesis concentrates on the problem of feature selection for 3D facial expression recognition. The first objective is to use a good face representation in order to be a base for feature selection. There are several representations used in the literature for face description [8]. Mainly, the human face can be defined with appearance based or 3D geometry based features [9]. For facial expression recognition, the changes on the facial information should be extracted. Skin based features represent the facial characteristics well, however, these features are processed in 2D space. When the face is deformed with one of the facial expressions, the most of the changes are in 3D space. Therefore, the face representation should include 3D information about the facial characteristics. On the other hand, the face representation should be robust under limiting conditions such as illumination changes. As a result, the 3D geometric facial feature points are selected as they represent the changes on the face. The details about the face representation are provided in Chapter 3.

The second objective of the thesis is to find a good classifier to be used for facial expression recognition system. By representing a face with 3D facial feature points, the next objective is to classify 3D facial feature points as one of the six basic facial expressions. Using geometric feature points yield representation of a face as a row vector. Then, the facial expression recognition problem becomes a vector classification problem. Support Vector Machine (SVM) is employed as the classifier as it is a well known, successful classifier for the vector classification problems [10]. SVM is used in two different ways, multi-class and group of 2-class classifiers. The details about the classifier are presented in Chapter 4.

The third and the main objective of the thesis is to further find the most efficient facial features based on 3D facial geometry which will discriminate facial expressions in the best way. Under this objective, the aim is to find which facial features carry the most of the information during expression deformations of the face. Also, this objective includes finding the features that provide the best discrimination as well as the least confusion between facial expressions. Currently, one of the most important challenges in facial expression recognition systems is that the most of the expressions are recognized with a very high rate of accuracy and a few with low rates of accuracy so that the overall recognition rate remains high. In the thesis study, this issue has been focused and the feature selection algorithms are proposed to recognize each expression with closely high rates. This objective is achieved by analyzing the information content of 3D facial feature points during facial expression deformations. Information content is analyzed using the strong metrics which is entropy. Novel feature selection procedures which are the main objectives of the thesis are presented in Chapter 5.

1.3 Thesis Contributions

The thesis is focused on facial expression recognition problem using 3D facial feature points. The thesis mainly contributes on feature selection procedure. The contributions of the thesis are listed below.

1.3.1 Variance Based Feature Selection for Expression Classification

Feature selection is an important part of a facial expression recognizer. Variance based feature selection is one of the contributions of the thesis. The 3D positions of facial feature points vary under different facial expressions. The points with high variance carry information during expression. The thesis proposes a novel variance based feature selection algorithm for 3D facial expression recognition, explained in Chapter 5, in Section 5.1 [11, 12].

1.3.2 Entropy Based Feature Selection for Expression Classification

The main contribution of the thesis is the novel entropy based feature selection algorithm for person independent 3D facial expression recognition. The variance based feature selection algorithm is improved with the next information content metric, which is entropy. During facial expression deformations, some facial feature points have high entropy. Low entropy means a feature point is not affected from facial expression, even some of them are stable points. Whereas, high entropy means a feature point is very dynamic during expression deformations and it is rich in terms of information content. These high entropy feature points contribute to facial expression recognition more than the low entropy points. Thus, a feature point with significant entropy means information rich point to select. Therefore, the feature selection algorithm selects high entropy points and shows that they improve the recognition performance significantly [13].

In preliminary experiments, it is observed that some expressions are confused like anger and sadness. Most of the anger expressions that are wrongly classified are classified as sadness. These confusions motivate the study in order to solve these confusions. The attempt to solve these confusions is to find out which expression couples are confused. A study has been done about the clustering of facial expressions using Fuzzy C-Means (FCM) clustering algorithm. According to the clustering algorithm results, anger, disgust and fear expressions are grouped in group 1. The other expressions, happiness, sadness and surprise are also grouped in group 2. Therefore, a classifier model which classifies an unknown expression in two levels is used in the thesis. First, the unknown expression is classified as in Class 1 or Class 2 where Class 1 includes anger, disgust, fear, and Class 2 includes happiness, sadness, surprise expressions. The entropy based feature selection algorithm selects the most discriminating feature points for the first level of classification: Class 1 or Class 2. Then, the entropy based feature selection continues for each class, selecting different features for classification under each class in the first level. The details about the feature selection based on two-level classification model are provided in Chapter 5 in Section 5.2 [13].

1.3.3 Expression Distinctive Classifier Model

The next contribution of the thesis is the feature selections for expression specific classifier model for 3D facial expression recognition. Each facial expression is studied individually with the neutral expression and the feature points which are the most discriminative for that specific expression are selected. While some feature points are discriminative, some others may confuse the classifier. As a result, six different facial feature combinations are proposed that are the most discriminative features for each one. The classifier is modified as set of accept-reject classifiers. An unknown face vector undergoes to accept-reject classifiers for the six basic expressions. There are six accept-reject classifiers each of dedicated to a specific expression. It is expected that, for example, an anger face is accepted by the anger's accept-reject classifier, and rejected by the others. If so, it is recognized as anger. In case of multiple accepts within the accept-reject classifiers, a decision module is running to make the final recognition decision for the unknown face vector. Feature selection procedures are based on entropy. High entropy features which are analyzed between neutral and a specific expression are selected [7]. The detailed information is given in Chapter 5 in Section 5.3.

1.4 Overview of Thesis

The thesis expounds contributions which are entropy based feature selections for enhanced facial expression recognition. The thesis is important for enhancing facial expression recognition with novel feature selection procedures. The proposed feature selection procedures are expressed with two different classification models. Chapter 2 includes review about facial expression studies in miscellaneous research directions. Overview covers research findings related to face representation, feature selection and automatic facial expression recognition. In Chapter 3, the details of face representations used in the thesis study are presented as well as the databases used in simulations. Classification models used for novel feature selection procedures are explained in Chapter 4. Furthermore, the motivations behind the classification models used and their classifier architectures are expressed. The proposed entropy based feature selection procedures are provided in Chapter 5. Chapter 6 is the last chapter of the thesis in which overall conclusions and future directions are discussed.

Chapter 2

FACIAL EXPRESSION RECOGNITION

2.1 Introduction

Human facial expression studies have their origins in early 1600s where first categorization of expressions takes place [14]. Since then, facial expressions are in the scope of psychological studies. Latest developments in digital computing in 19th century allow today’s digital processors to be able to detect and analyze human face in digital images. Thus, facial expressions are studied by computer scientists in order to develop systems for automatic facial recognition.

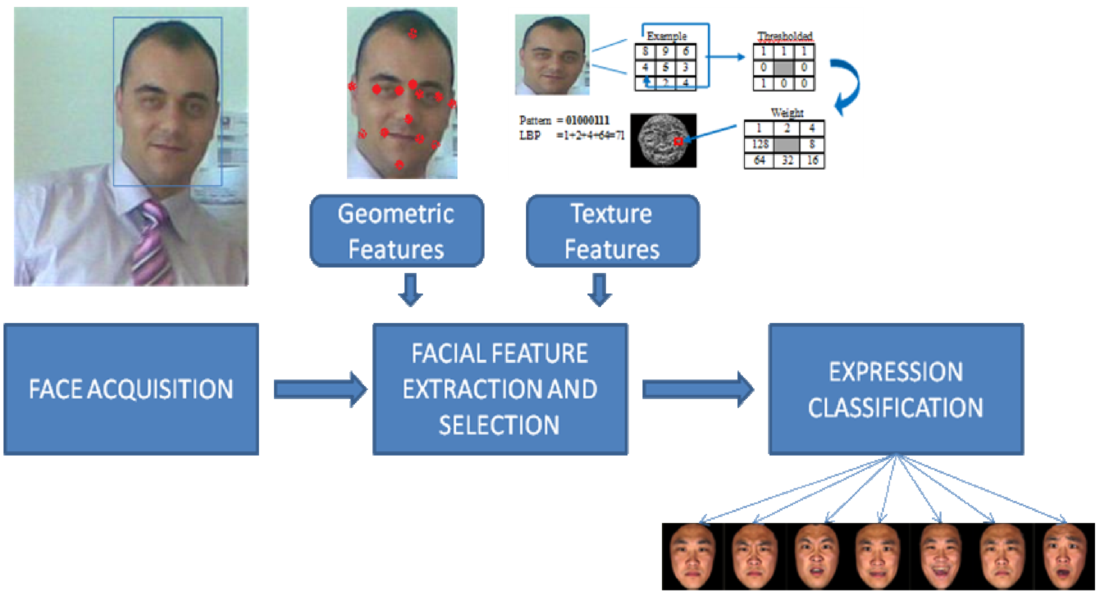


Figure 2.1: A typical facial expression recognition system.

Facial expression recognition methodologies have been developed mostly on still images which is called static facial expression recognition. Furthermore, facial expression recognition methodologies are applied to dynamic facial images in video. In addition, dynamic recognition systems may employ a temporal modeling of the expression as a further step. In the thesis, the proposed methods are applied to static facial expression recognition.

A typical facial expression recognition system can be investigated in three main parts which are face acquisition, facial feature extraction and selection, and classification. Face acquisition is the first step that locates the face in the image. The second step after locating the face is to define the face with facial features. There are several methods for facial feature extraction till now, mainly categorized in two broad categories which are facial geometry based features and appearance based features. After feature extraction and selection, classification step takes place to classify input features as one of the facial expressions. A typical facial expression recognizer is illustrated by the steps given in Figure 2.1.

2.2 Historical Background

Earlier studies of facial expressions in 19th century were pioneered by Charles Darwin. In 1872, Darwin established the principles of expression and the descriptions of expressions. He did expression analysis for humans and animals. Darwin's expression analysis also categorized facial expressions in various categories. Darwin's findings about facial expression categories were anxiety, joy, anger, surprise, disgust, sulkiness and shyness. Each of them includes sub categories. He also defined facial deformations related to each expression category [15].

Charles Darwin's categorization of facial expressions also includes the following categorization where several kinds of expressions are grouped into similar categories [8].

- low spirits, anxiety, grief, dejection, despair
- joy, high spirits, love, tender feelings, devotion
- reflection, meditation, ill-temper, sulkiness, determination
- hatred, anger
- disdain, contempt, disgust, guilt, pride
- surprise, astonishment, fear, horror
- self-attention, shame, shyness, modesty

An important milestone in 1970s in facial expression research was the studies of Paul Ekman and his colleagues [3, 4]. Ekman's research findings were about the classification of facial expressions into seven basic categories which are anger, disgust, fear, happiness, neutral, sadness and surprise. Later on, Ekman and Friesen developed Facial Action Coding System (FACS) to code facial expressions and facial movements which they called action units (AUs) [4]. Facial Action Coding is related to muscles. It includes the facial muscles that make changes in the face. These activities of the face are called Action Units (AUs). Their work is important for the literature in the way that many researchers followed for the developments of current recognition systems.

Another milestone in facial analysis studies is the model proposed by Moving Pictures Expert Group (MPEG) in 1999. They standardize face and facial animation parameters in MPEG-4 standard. The model introduced a face model with 83 feature points defined (Figure 2.4), Facial Definition Parameters (FDPs), describing a face in its neutral state. MPEG-4 standard also defined 68 Facial Animation Parameters (FAPs) which are used to animate the face by the movements of the feature points. FAPs can be used to animate the faces and to synthesize basic facial expressions [5]. Besides, FAPs can be used for facial expression representation on a generic face model. MPEG-4 FAPs are widely used in most of the research labs for facial expression synthesis and analysis studies [16, 17, 18].

Table 2.1: Expression related emotion categories [8].

Basic Expression	**Related Emotions**
Anger	Rage, outrage, fury, wrath, hostility, ferocity, bitterness, hate, loathing, scorn, spite, vengefulness, dislike, resentment.
Disgust	Revulsion, contempt.
Fear	Alarm, shock, fright, horror, terror, panic, hysteria, mortification.
Happiness	Amusement, bliss, cheerfulness, gaiety, glee, jolliness, joviality, joy, delight, enjoyment, gladness, jubilation, elation, satisfaction, ecstasy, euphoria.
Sadness	Depression, despair, hopelessness, gloom, glumness, unhappiness, grief, sorrow, woe, misery, melancholy.
Surprise	Amazement, astonishment.

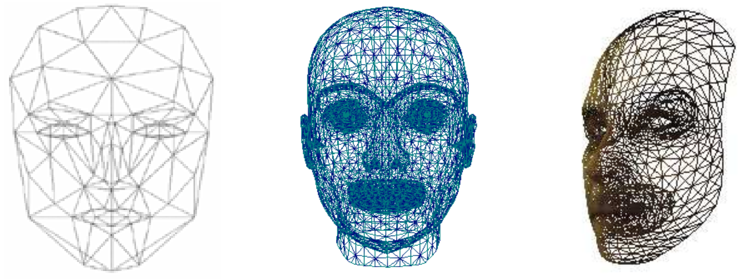
Human facial expressions may differ between different nations or cultures. Ekman's findings about six basic facial expressions have been also stated that these are the most distinguishable expression classes among all the cultures in the World [3, 4]. Besides, other than six basic expressions, human face is capable of expressing other various emotions. Table 2.1 summarizes related emotions belonging to six basic prototypic expressions proposed by Ekman [8].

2.3 Face Acquisition

Face acquisition is the first step for an expression recognizer. This step includes face detection and localization and it is a key step in all face and facial expression recognition systems. The main purpose is to localize and extract the face region and facial information from the image, by extracting the face region from the background.

Several techniques have been developed for detection of the faces in still images. The classification of the methods differs according to the criteria used. Modeling based approaches can be classified into two broad categories which are local feature based methods and global methods. Local features based method first localizes the critical regions on the face such as eyes and mouth. Then the face vectors are constructed using these localized features. In the global methods, the entire facial image is coded and considered as a point in high-dimensional space [19].

Face acquisition methods also involve reconstruction of a 3D face from 2D face images. One of the promising methodologies for 3D face reconstruction from 2D face images is the 3D morphable model method. In this method, 3D laser scans of the faces are used to generate the model and present statistical representation of the facial geometry and texture. Another well-known methodology is the model fitting. Here, the 2D images are analyzed and the facial features are extracted, then the 3D model is synthesized using the adapted 3D morphable model and 2D texture information. A publicly available 3D morphable model is presented in [20, 21]. One of the most popular 3D face model is the Candide, which is also widely used in facial animation [1]. In [16] and [18], 3D face reconstruction from two orthogonal images and single frontal face image are presented. The face models are shown in Figure 2.2.



(a) **(b)** **(c)**

Figure 2.2: 3D face models (a) Candide version 3, (b) 3D face model used in [16], (c) 3D face model used in [18].

Locating 2D faces in the images which is the face detection stage is an essential part of many face related solutions including facial expression recognition. The most widely used algorithm is the one proposed by Viola and Jones. In this popular face detector, Haar-like features and AdaBoost algorithm have been employed [2].

In the thesis study, BU-3DFE database has been employed. In BU-3DFE database, facial shape models, frontal view textures and 83 3D geometrical feature point positions are included for each subject. The details about the database are provided in Chapter 3, in section 3.3.1.

2.4 Facial Feature Extraction and Selection

Representing a face from the face image is of utmost importance for accurate facial expression recognition. Besides, the selection of the most discriminative facial features is a critical step for a successful classification process. The human face can be parameterized with 2D or 3D geometry and texture. There are appearance based or facial geometry based representations in the literature. A number of techniques were successfully developed using 2D static images [10]. In these studies, the face is considered as a 2D pattern with certain textures that expression variations can be observed. Some methods analyze facial appearance changes on the face or on critical regions of the face. These approaches include the application of Principal Component Analysis (PCA), Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA) and Gabor wavelets (GW) to facial images [9]. In the literature, considerable studies have been performed based on appearance. Wang et al. [22] employed LDA based classifier system and achieved 83.6% overall recognition rate on the BU-3DFE database. Lyons et al. [23] achieved 80% average recognition rate using 2D appearance feature based Gabor-wavelet approach. Jiangang Yu and Bhanu improved evolutionary feature synthesis for facial expression recognition [24].

One of the most successful texture representations used in facial analysis is the Local Binary Patterns. Local Binary Patterns (LBP) is investigated by Ojala et al [25] in 1996 for texture classification. Basic LBP operator is shown in Figure 2.3. It is applied to majority of texture classification problems including face recognition. In 2008, Shan et al. evaluated LBP features for person independent facial expression recognition and concluded that LBP features are effective and efficient for facial expression recognition by supplying accurate facial representation that describes appearance changes well [26].

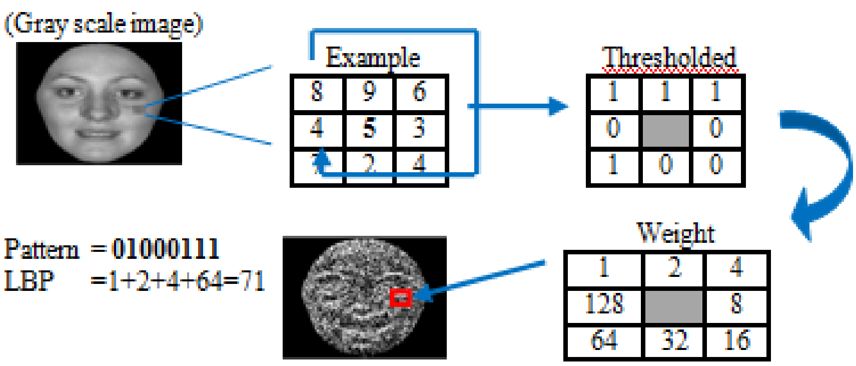


Figure 2.3: Sample LBP applied to a face image from BU-3DFE database [27].

The other efficient facial representation is the one based on facial geometry. There are two main representations that researchers are following. The one is the Facial Action Coding System proposed by Ekman and Friesen [4].

Table 2.2: Action Unit main codes [28].

Action Unit No.	Action Unit Name	Action Unit No.	Action Unit Name
0	face	**24**	Lip Pressor
1	Inner Brow Raiser	**25**	Lips Part
2	Outer Brow Raiser	**26**	Jaw Drop
4	Brow Lowerer	**27**	Mouth Stretch
5	Upper Lid Raiser	**28**	Lip Suck
6	Cheek Raiser	**29**	Jaw Thrust
7	Lid Tightener	**30**	Jaw Sideways
8	Lips Toward Each Other	**31**	Jaw Clencher
9	Nose Wrinkler	**32**	[Lip] Bite
10	Upper Lip Raiser	**33**	[Cheek] Blow
11	Nasolabial Deepener	**34**	[Cheek] Puff
12	Lip Corner Puller	**35**	[Cheek] Suck
13	Sharp Lip Puller	**36**	[Tongue] Bulge
14	Dimpler	**37**	Lip Wipe
15	Lip Corner Depressor	**38**	Nostril Dilator
16	Lower Lip Depressor	**39**	Nostril Compressor
17	Chin Raiser	**41**	Glabella Lowerer
18	Lip Pucker	**42**	Inner Eyebrow Lowerer
19	Tongue Show	**43**	Eyes Closed
20	Lip Stretcher	**44**	Eyebrow Gatherer
21	Neck Tightener	**45**	Blink
22	Lip Funneler	**46**	Wink
23	Lip Tightener		

FACS introduces main face codes, head movement codes, eye movement codes, visibility codes and gross behavior codes. Table 2.2 shows main codes of FACS. Different expressions contain different combination of AUs. For six basic expressions, the related AUs are provided in Table 2.3 [28].

Table 2.3: Expression related facial action units [28].

Expression	Action Units
Anger	4,5,7,23
Disgust	9,15,16
Fear	1,2,4,5,20,26
Happiness	6,12
Sadness	1,4,15
Surprise	1,2,5B,26

The second efficient representation is the one proposed by Moving Pictures Experts Group (MPEG). In order to provide a standardized facial control parameterization, the MPEG defined the Facial Animation (FA) list of conditions in the MPEG-4 standard. The first release of the MPEG-4 standard became the international standard in 1999. Facial expression recognition studies have used the MPEG-4 to define facial expressions.

The neutral face model which has 83 geometric feature points was introduced in the standard as shown in Figure 2.4. MPEG-4 standard also defined 68 Facial Animation Parameters (FAPs) used to animate the face by the movements of the feature points. FAPs can be used to synthesize basic facial expressions and to animate the faces [16, 29]. In addition, FAPs can be used for facial expression representation. MPEG-4 FAPs are popular in most of the research labs today for facial expression synthesis and analysis studies [18, 29].

FAPs represent a complete set of basic facial actions including head motion, tongue, eyes and mouth control. Studies about facial expression modeling and synthesis [16, 17, 18] showed that the action of the FAPs on the whole 3D generic face model is obtained by displacement functions for each vertex. The displacement functions Δi (x, y, z) of a vertex are computed according to the position of the vertex in the influence area, to the intensity of the FAP in the related feature point(s) and an additional weight for design issues as shown in Equation 2.1 [5].

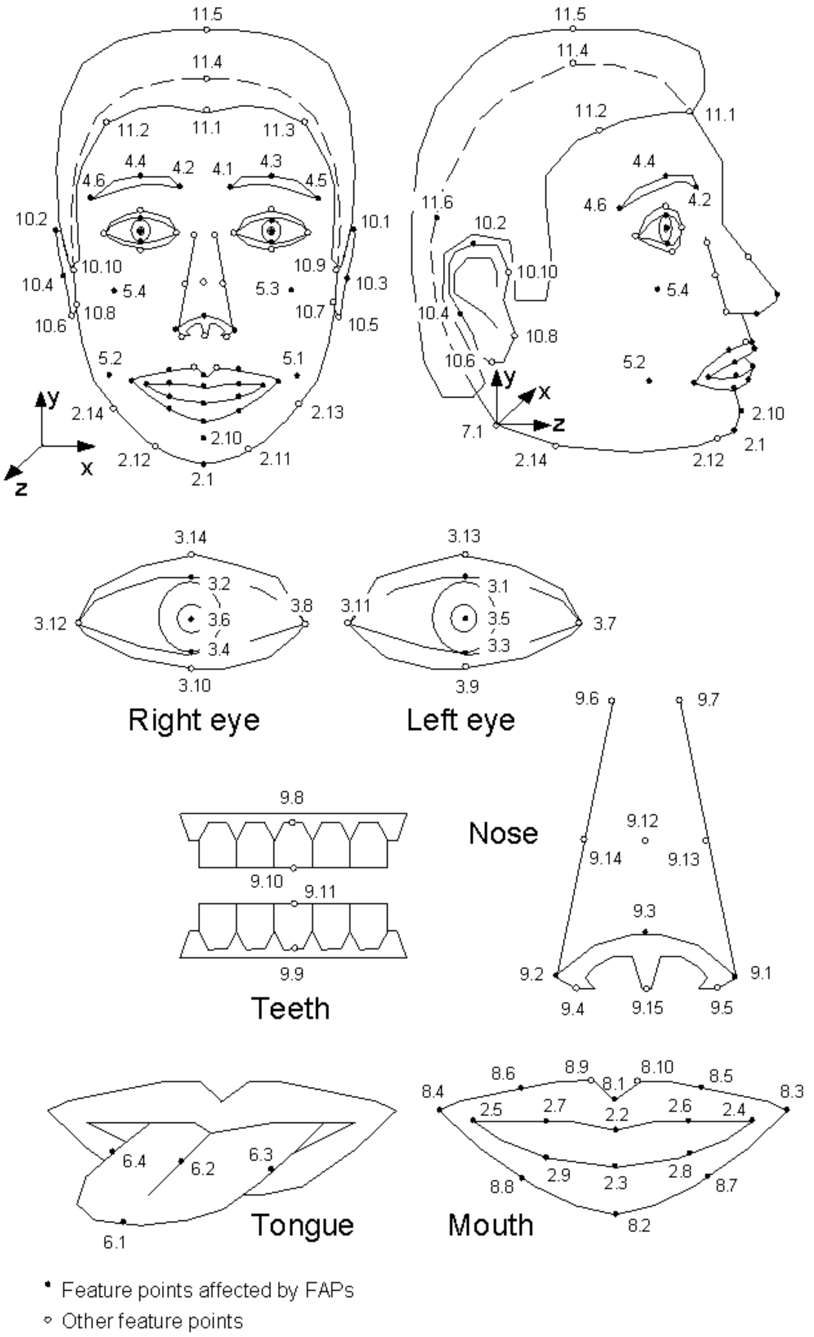


Figure 2.4: MPEG-4 geometric feature points [5].

The weight W_i in Equation 2.1 is based on the distance of the vertex from the feature point and the weight spreads decreasingly from the centre of the influence area. If a vertex is not to be affected by the FAP, $W'_j = 0$ may be chosen.

$$\begin{bmatrix} \Delta i_x \\ \Delta i_y \\ \Delta i_z \end{bmatrix} = \begin{bmatrix} Wi_x \\ Wi_y \\ Wi_z \end{bmatrix} * W' j * FAP_{x,y,z} \tag{2.1}$$

2D representations in both texture and geometry models are effective, however, facial features that affect changes on the face are mostly in 3D space rather than 2D surface. Also, many expressions include skin wrinkles, for example, forehead deformations. Due to the limitations in describing facial surface deformations in 2D, there is a need for 3D space features in order to represent 3D motions of the face successfully. In this context, 3D geometrical feature point data are employed in the thesis study from BU-3DFE database.

BU-3DFE database consists of 100 individuals with 6 basic prototypic expressions and the neutral expression. All expressions contain 4 different intensities, 1 being the lowest and 4 being the highest intensity for the corresponding expression. The aim is to model spontaneous facial expressions. Database includes facial shape models, frontal view textures and 83 3D geometrical feature point positions for each subject. The details about the BU-3DFE database are provided in Chapter 3, section 3.3.1.

Feature selection is one of the most important stages after feature extraction process. The extracted features, geometric or appearance based, always include redundant information. On the other hand, using all the extracted features may confuse the classifier where some features play similar roles among the expressions [13]. Thus, an efficient feature selection algorithm can select the best features that best discriminate the facial expressions. There are research studies in the literature about the feature selection procedures employed for facial features specific to expression recognition. Entropy based feature selection algorithm for facial expression recognition using 3D geometric features is one of the contributions of the thesis. Also AdaBoost and GentleBoost are the two algorithms used in feature selection [30, 31]. Fisher criterion and Kullback-Leibler divergence are employed to find the most discriminative features [32, 33, 34].

There are considerable research studies published on BU-3DFE database. Soyel et al. [35] proposed NSGA-II based feature selection algorithm and achieved %88.3 overall recognition rate using 3D feature distances on BU-3DFE database. Later they proposed a facial expression recognition method based on localized discriminative scale invariant feature transform and reached 90.5% average recognition rate [36].

2.5 Facial Expression Classification

The final stage in a typical expression recognition system is the classification stage where input face vectors are classified as one of the prototypic expressions. After defining a face with extracted and selected facial features, a face can be expressed as a row vector. Thus, facial expression recognition problem is then considered as a vector classification problem. Considering facial expression recognition as a vector classification problem needs a strong classifier.

There are several classifiers used in the literature in order to classify vectors defining facial information. Principal Component Analysis (PCA) has been further used in facial expression recognition systems in order to reduce the dimensionality [33, 37]. The most widely used classifiers in 3D facial expression recognition are Linear Discriminant Analysis (LDA), Bayesian classifiers, Support Vector Machines (SVMs), Hidden Markov Models (HMMs) or Neural Networks.

LDA is a well known linear classifier that can be applied to facial expression recognition. Wang et al. [22] used LDA based classifier system and reached 83.6% overall recognition rate on the BU-3DFE database.

Bayesian classifiers are probabilistic classifiers which depend on Bayes' theorem. They are also employed in facial expression studies. Sebe et al. [38] evaluated Bayesian classifier for authentic facial expression analysis.

SVMs are other common classifiers which are used in facial analysis. SVM is a supervised learning model that can classify the new patterns according to the input known patterns. Sebe et al. [38] also evaluated SVM for authentic facial expression analysis. Kostia and Pitas [22 state of the art] employed multi-class SVM on Candide's geometric data for facial expression recognition. In the thesis study, SVM has been applied as a classifier. The details about the SVM classifier used are given in Chapter 4.

HMMs are also employed in facial expression analysis studies. HMM is a statistical Markov model in which unobserved system states present. Pardas and Bonafonte [39] employed HMMs in their facial expression recognition system and achieved 84% overall recognition performance on Cohn-Kanade database, which is described in Section 2.6.

Several neural network variations are applied to expression classification problem. Artificial Neural Networks (ANN), Multi-Layer Perceptrons (MLP), the Hopfield model, and Probabilistic Neural Networks (PNN) are some examples. Tian et al. [40] employed ANNs, Kotsia et al. [41] employed multiclass SVM and MLP in their studies. In 1999, the Hopfield model has been applied in [69]. In 2008, PNN has been used as a classifier in [42] and achieved 87.8% recognition rate on BU-3DFE database.

2.6 Facial Expression Databases

Starting from the early facial expression analysis up to today's recognition systems, several challenges have been faced. In order to attack these challenges, scientists developed different databases for facial expression recognition. Although most of the databases available for facial analysis are developed for face recognition, there are public face databases available which are dedicated to facial expression recognition.

Facial expressions are the results of spontaneous movements and animations in human face. Thus, a natural facial expression database should be created with subjects' uninformed pure expressions which are the snapshots from their real life expressions. Sebe et al. [38] analyzed the main difficulties of capturing real expressions of humans when a database is to be created. According to their observations, the following conclusions have been reached:

- Emotions are observed in different intensities among subjects. Each subject has different intensity level to express an emotion.
- When subjects are informed before capturing the expressions, their facial expressions become unrealistic.
- Because of laboratory conditions, a subject's facial expression may not reflect his/her natural expression.

There are many databases available today that enable scientists to analyze facial expressions. Currently a number of 3D databases are publicly available facial expression modeling and recognition. The first 3D expression database appears in the study of Y.Chang et al. [43] in 2005 where six basic expressions of the subjects are recorded in a real-time 3D video by using a camera – projector scanning system. This database is not publicly available. Later on, more systematic databases which are publicly available have been developed.

The BU-3DFE database is one of the most widely used public databases available which is introduced in 2006 by L.Yin et al. [45] They aim to foster the research studies on 3D facial expression recognition by offering this first publicly available 3D facial expression database. It includes facial shape models, frontal view textures and 83 3D geometrical feature point positions for 100 adults including 56 female and 44 male subjects. Also, 2D facial textures of face models are included. The six basic expressions for each subject that are anger, disgust, fear, happiness, sadness and surprise are provided in 4 different intensities, 1 being the lowest and 4 being the highest intensity. The neutral faces are also provided for each subject with 2D texture models and geometric feature points. There are 2500 samples in total including 25 for each subject.

The Bosphorus Database represents a 3D face database enriched with six basic expressions in different poses. The expressions covered are anger, disgust, fear, happiness, sadness and surprise. In addition to six basic expressions, the database contains action units of FACS. Facial occlusions such as hand, hair or eyeglasses are also included for the subjects. There are 105 subjects in total, 60 men and 45 women. Besides action units, each face scan has been manually labeled with 24 3D facial feature points. These points are selected from the regions such as nose, eye corners. There are 4652 samples in total including 31 to 54 samples for each subject [46].

The proposed algorithms of the thesis are tested in two well known facial expression databases which are BU-3DFE and Bosphorus databases. The face representations related to these two databases are explained in the sections 3.2 and 3.3.

The Cohn–Kanade facial expression database is another widely used database for 3D facial expression recognition [8]. It is also known as CMU-Pittsburg database. The database includes 486 sequences from 97 subjects in total. Action units, FACS, and their combinations are included for each subject. Each subject was directed to 23 different facial displays. Six basic expressions of the face are included. Also, the database includes the posed expressions. In 2010, the extended version of this database has been released, CK+, which includes 593 sequences from 123 subjects. Image sequences vary in duration from 10 to 60 frames. In the extended version 7 expressions which are anger, disgust, fear, happiness, sadness, surprise and contempt are included together with the neutral. Also, 30 AUs are presented [44].

Another popular expression database is the Face Recognition Grand Challenge (FRGC) database version 2 (v2). It includes samples mainly for face recognition. In addition to this, there are 1 to 22 samples for 466 subjects including expressions. In total, there are 4007 samples including 466 subjects' different expressions which are anger, disgust, happiness, sadness, surprise and puffy [47].

In the following table, the most widely used publicly available 3D facial expression databases which are mentioned above are listed comparatively.

Table 2.4: 3D facial expression databases.

Database	No of Subjects	No of Samples per Subject	Total Samples	Included Expressions
BU-3DFE [45]	100	25	2500	Anger, disgust, fear, happiness, sadness, surprise and neutral
Bosphorus [46]	105	31-54	4652	Anger, disgust, fear, happiness, sadness, surprise and neutral + Action Units
Extended Cohn-Kanade [44]	123	593 image sequences	5930 - 35580	Anger, disgust, fear, happiness, sadness, surprise, contempt and neutral + Action Units
FRGC v2 [47]	466	1-22	4007	Anger, disgust, happiness, sadness, surprise and puffy.

2.7 State of the Art

Since the 1990s, research on facial expression recognition has been growing rapidly. Current research activity in facial expression recognition is focused on automatic facial expression recognition and has achieved acceptable recognition performances under controlled conditions. There remain several challenges in facial expression recognition studies on which research is focused. One of the most important challenges is that not all expressions are being accurately recognized. Most expressions are recognized with a very high rate of accuracy and a few with low rates of accuracy so the overall recognition rate is fairly high. In our study, we have focused on this issue and propose different independent face definitions to recognize each expression with high rates. Other challenges can be listed as follows.

- Recognition of other expressions than the six basic expressions.
- The automatic recognition of expressions and AUs from different head angles and rotations.

- Recognition of spontaneous expressions.
- Real-time facial feature detection for expression recognition.

Since the early 1990s, many studies into facial expression recognition have been published. The approaches differ according to the feature extraction method used, person dependency and classifier design. On the other hand, facial representation is also an important part of a facial expression recognition system. The face can be represented using texture information, 2D or 3D geometry, or the fusion of both. This is also related to the feature extraction and selection processes [48]. Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA) and Gabor wavelets are widely used methods in facial expression analysis [49]. In 2001, Tian et al., reported 96.4% recognition rate of upper face AUs and 96.7% recognition rate of lower face AUs on their own database [50]. Bourel et al., achieved 83.31% recognition rate using Tree Augmented Naive Bayesian classifier on their own database [51] , Cohen et al., proposed the use of Hidden Markov Model to automatically partition a video into different expression partitions in order to develop a real-time system and reported 83.62% recognition rate with Stochastic Structure Search classifier in 2003 [52]. Pantic and Patras generated mid-level parameters by tracking 15 facial feature points and achieved 93.66% recognition rate with Multi-Stream Hidden Markov Model in 2006 [53]. Wang and Yin proposed a topographic modeling approach that treats the grey-scale image as a 3D surface and studied the robustness against the disturbances of face regions and facial expression intensity levels [54]. Kotsia et al., employed Gabor features with multiclass SVM and Multi-Layer Perceptron (MLP) and reported 91.4% recognition rate [55] . Lyons et al. [56]

reported 80% average recognition rate using 2D appearance feature based Gabor-wavelet (GW) approach.

On BU-3DFE database, Wang et al. [22] used LDA based classifier system and reached 83.6% overall recognition rate. Tang and Huang [57] reached 87.1% average recognition rate on BU-3DFE database using line segments connecting 3D facial feature points. Later in 2008, Mpiperis et al. [58] explored the use of bilinear models for facial expression recognition and achieved 90.5% average recognition rate. Soyel et al. [35] proposed Non Dominated Sort Genetic Algorithm II (NSGA-II) based feature selection algorithm and achieved 88.3% overall recognition rate using 3D feature distances on BU-3DFE database. Later they proposed a facial expression recognition method based on localized discriminative scale invariant feature transform and reached 90.5% average recognition rate [36]. Facial expression recognition is still under research for next level robust recognition systems.

Chapter 3

FACE REPRESENTATIONS FOR FACIAL EXPRESSION RECOGNITION

3.1 Introduction

Creation of facial representation from the face image is of utmost importance for accurate facial expression recognition. On the other hand, selection of the most discriminative facial features is also a vital step for expression classification. There are appearance based or facial geometry based representations in the literature. The conventional methods for facial expression recognition focuses on the extraction of data needed to describe the changes on the face. A number of techniques were successfully developed using 2D static images [10]. They consider the face as a 2D pattern with certain textures that expression variations can be measured. However, facial features that affect changes on the face are mostly in 3D space rather than 2D surface. Also, many expressions include in-depth skin motion, for example, forehead deformations. Due to limitations in describing facial surface deformation in 2D, there is a need for 3D space features in order to represent 3D motions of the face accurately [45]. Therefore, we employed 3D geometrical feature point data in our research study.

In the literature, there are recognition systems using face representations with geometric data on the face or the texture, appearance based representations. Also the fusion of geometric and appearance based models for facial representation have been used.

One of the important representations used in the literature is the one standardized by MPEG-4 in 1999. MPEG-4 standard defined a face model in its neutral state with 83 geometric feature points called Facial Definition Parameters (FDPs). MPEG-4 standard also defined 68 Facial Animation Parameters (FAPs) which are used to animate the face based on FDPs. FAPs are also used to synthesize basic facial expressions [16, 17, 18]. Also, FAPs are effective for facial expression representation. MPEG-4 standard is still popular in most of the facial expression studies.

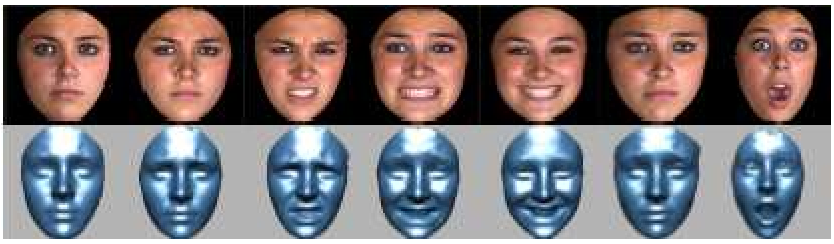
Another important representation is the one proposed by Paul Ekman and Wallace V. Friesen in 1978 [4] based on the model proposed by Carl-Herman Hjortsjö [59]. It is called Facial Action Coding System (FACS) and codes the movements of the face. Movements of individual facial muscles are encoded by FACS from instant variances in facial appearance [3, 4]. FACS became a common model to classify the physical expression of emotions.

In the thesis, 3D geometric feature point data have been used which are included in MPEG-4 FDPs.

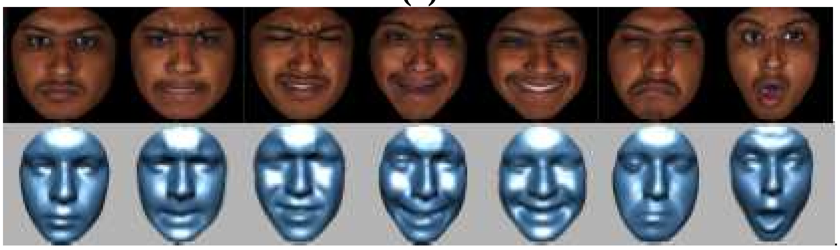
3.2 Face Representation Using BU-3DFE Database

In BU-3DFE, all expressions contain 4 different intensities, 1 being the lowest and 4 being the highest intensity for the corresponding expression. The aim is to model spontaneous facial expressions. Database includes facial shape models, frontal view textures and 83 3D geometrical feature point positions for each subject. Facial shape models are created with 3D face digitizer and facial range data is captured. The system projects a random light pattern to the object and captures the shape with synchronized cameras. Then, a single 3D polygon surface mesh is created for the face by merging all the information coming from six cameras [45]. 83 selected feature points are then picked from the 3D face model, shown in Figure 1 (c). The 3D pose of the face affects this process, so obtained facial models inherently contain varying poses. The pose of the model is calculated by considering three vertices; two from eye corners and one from nose tip, and the model is oriented using a normal vector with respect to the frontal projection plane. The feature detection algorithm used in the creation of BU-3DFE database already incorporates some of the corruptions that can be introduced by possible movements of the head, including rotations. Model projections with respect to the frontal projection plane are open to corruptions on some of the 3D feature point positions, and those corruptions are already embedded in the available data. Therefore, 83 3D feature point positions from BU-3DFE database reflect facial behavior for real life application and can represent a face with high accuracy in 3D. Figure 3.1 illustrates facial shape models, texture models and 83feature points from BU-3DFE database.

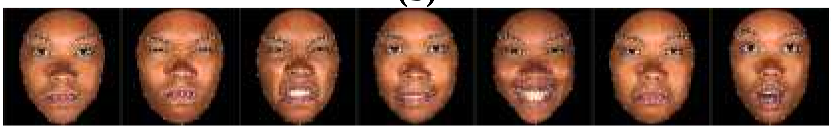
The 3D pose of the face affects geometrical feature extraction process, so obtained facial models inherently contain varying poses. The feature detection algorithm used in the creation of BU-3DFE database already incorporates some of the corruptions that can be introduced by possible movements of the head, including rotations. Model projections with respect to the frontal projection plane are open to corruptions on some of the 3D feature point positions, and those corruptions are already embedded in the available data. Therefore, we employed 83 3D feature point positions from BU-3DFE database that they reflect facial behavior of real life application and can represent a face with high accuracy in 3D.



(a)



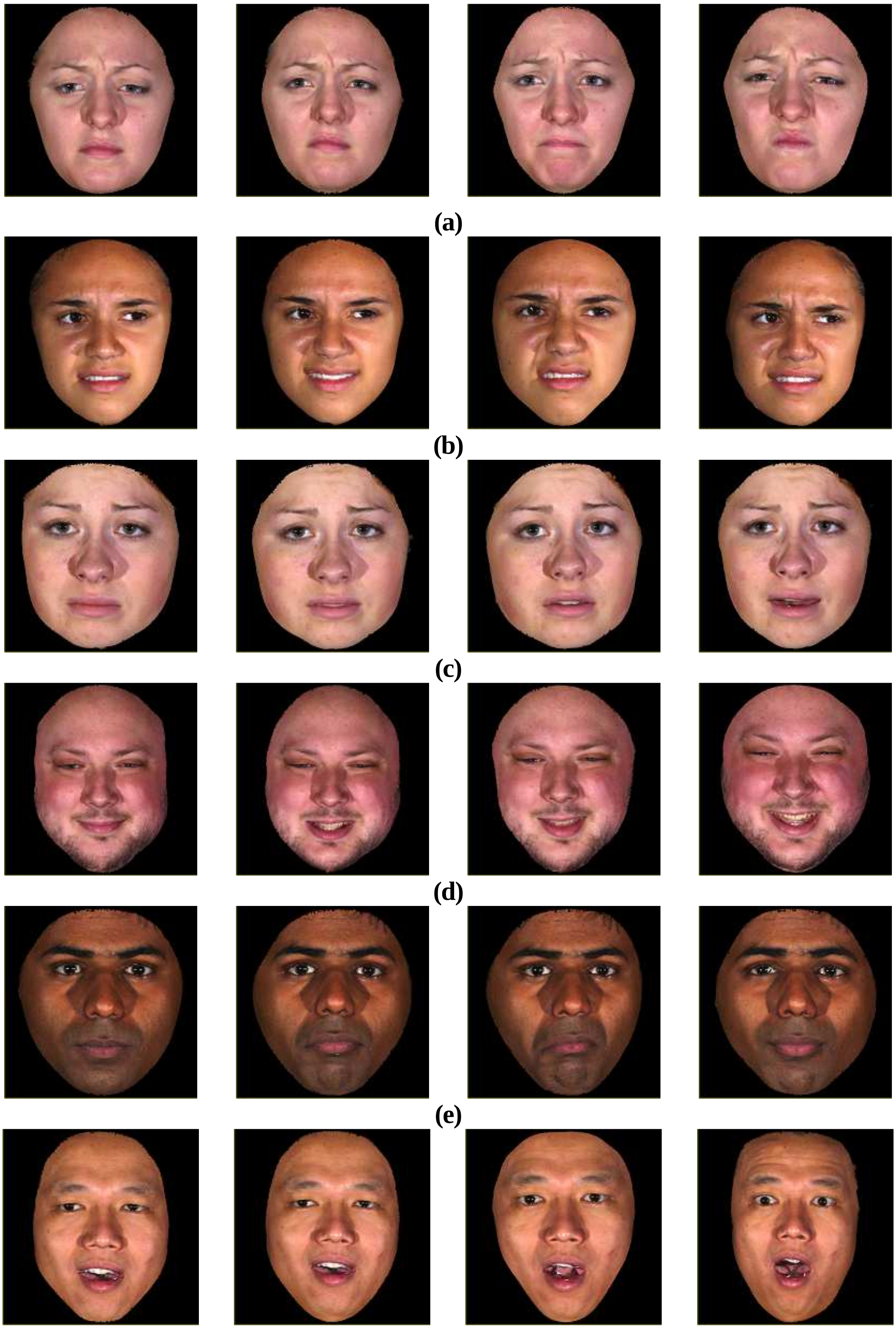
(b)



(c)

Figure 3.1: 2 individuals from BU-3DFE Database, **(a)** a female sample with frontal texture **(first row)** range image based facial shape model **(second row), (b)** a male sample with frontal texture **(first row)** range image based facial shape model **(second row)** and **(c)** 83 facial feature points selected.

The computational complexity of the feature extraction stage depends on the extraction of geometric features provided in BU-3DFE database. The database is created from the facial range data and the shape models obtained from 3D face digitizer. The system projects a random light pattern to the object and captures the shape with synchronized cameras. Then, a single 3D polygon surface mesh is created for the face by merging all the information coming from six cameras. The feature points are then picked from the 3D face model [45]. This process is limited by the speed of the face digitizer used and captures six prototypic expressions with 4 different intensities. In real-time, facial expressions are the spontaneous behaviors of the face. Thus, these 4 levels are designed in order to reflect this spontaneous behavior of the face. Examples including intensity levels are provided in Figure 3.2.



(f)

Figure 3.2: Facial expressions with 4 intensity level, 1 is the lowest (leftmost) and 4 is the highest intensity (rightmost) provided in BU-3DFE Database.
(a) Anger, **(b)** Disgust, **(c)** Fear, **(d)** Happiness, **(e)** Sadness, **(f)** Surprise.

In the thesis, the face representation used is based on the geometric feature points provided in the MPEG-4 FDPs. The Figure 3.3 shows the feature points considered for face representation. The feature selection algorithm uses all the feature points as a basis for further improvements on expression recognition success.

Consider a 3D facial feature point consisting of three vertices as given in Equation 3.1. By using facial feature positions, each face is represented by a face vector, *FV*. This face vector is obtained from the ordered arrangement of 3D feature point vertices (x, y and z for each point) and is created for each expression of the subject. Equation 3.2 shows how a face is represented as a vector of 3D feature positions. Face vectors are then combined into a matrix, face matrix, *FM*, and recognition tests are performed using this matrix, shown in Equation 3.3, where n denotes number of face vectors. Training and test sets for the classifier are derived from the subdivisions of *FM* into two parts.

$$V_i = \begin{bmatrix} V_{i_x} & V_{i_y} & V_{i_z} \end{bmatrix} \tag{3.1}$$

$$FV_j = \begin{bmatrix} V_1 & V_2 & V_3 \; ....V_q \end{bmatrix} \tag{3.2}$$

$$FM = \begin{bmatrix} FV_1 \\ FV_2 \\ .... \\ FV_n \end{bmatrix} \tag{3.3}$$

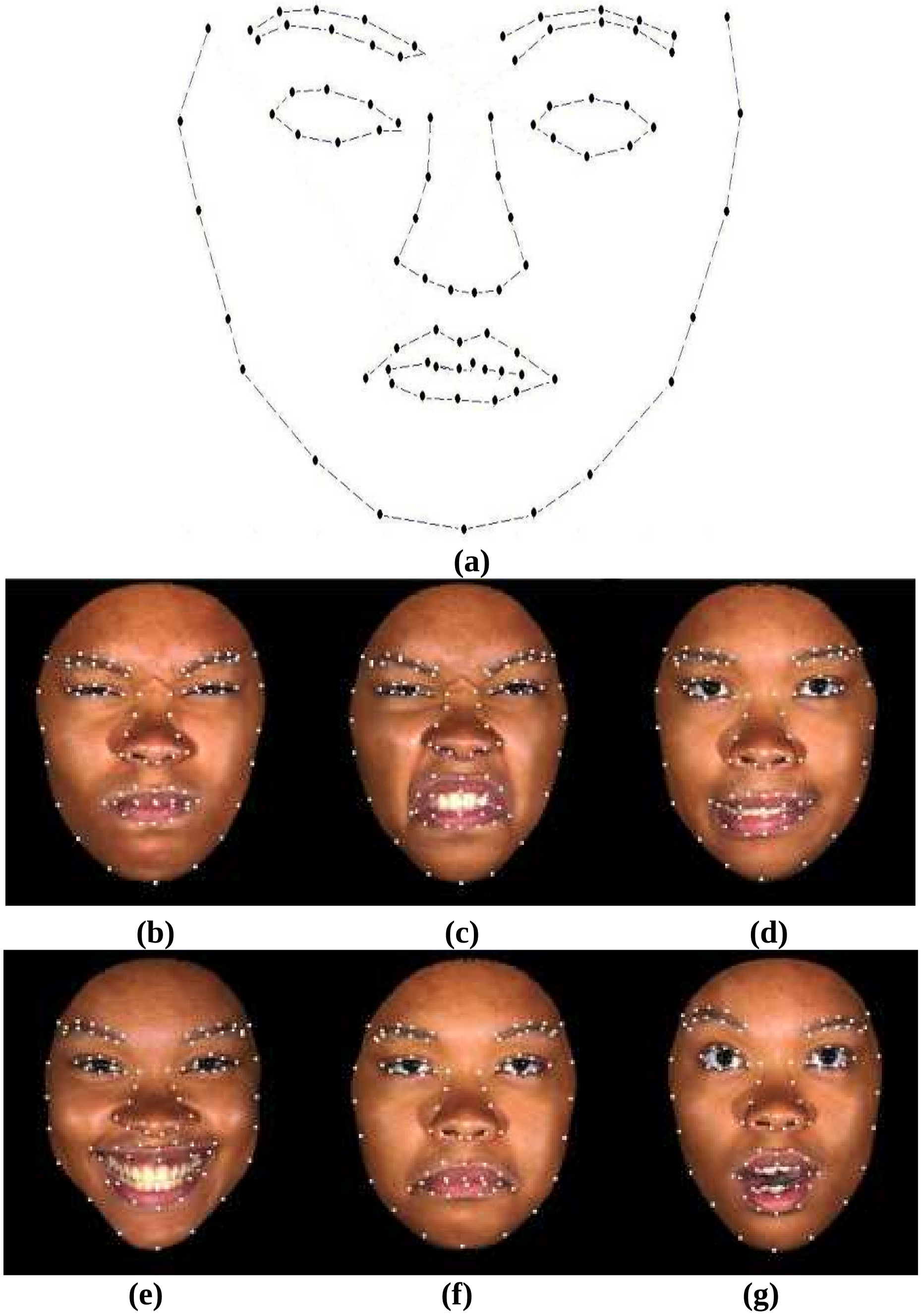


Figure 3.3: 83 Facial feature points used for face representation [45] **(a)** on neutral face, **(b)** Anger, **(c)** Disgust, **(d)** Fear, **(e)** Happiness, **(f)** Sadness, and **(g)** Surprise expressions.

3.3 Face Representation Using Bosphorus Database

Face representation based on 3D geometric facial feature points is also implemented by using the 3D feature point data provided in Bosphorus database. In Bosphorus database, there are 105 subjects, 60 male and 45 female. Occlusions such as eyeglasses or hair are added together with various poses. There exist 54 face scans for each subject, except 34 of them which has 31 face scans. The total number of face scans is 4652.

Each sample includes a color image, a 2D landmark file with the corresponding labels, a 3D landmark file with the corresponding labels and a coordinate file including both 3D and 2D coordinates. Sample facial images of a male and a female from Bosphorus database are illustrated in Figure 3.4. The geometric feature points are labeled manually on each face as shown in Figure 3.5 [46].

(a) **(b)** **(c)** **(d)** **(e)** **(f)**

Figure 3.4: Two individuals, one male and one female, from Bosphorus Database with 6 basic facial expressions: **(a)** anger, **(b)** disgust, **(c)** fear, **(d)** happiness, **(e)** sadness and **(f)** surprise.

Considering a 3D facial feature point consisting of three vertices as given in Equation 3.1, a face vector, FV is obtained from the ordered arrangement of 3D feature point vertices similarly with the implementation based on 83 feature points presented in BU-3DFE database. But in Bosphorus database, because of the number of 3D facial feature points, the face vector FV_j takes the form as in Equation 4 where m represents number of manually labeled feature points. Equation 3.4 shows how a face is represented as a vector of 3D feature positions. Face vectors are then combined into a matrix, face matrix, *FM*, and recognition tests are performed using number of face vectors as *p*, shown in Equation 3.5.

$$FV_j = \begin{bmatrix} V_1 & V_2 & V_3 \; ....V_m \end{bmatrix} \tag{3.4}$$

$$FM = \begin{bmatrix} FV_1 \\ FV_2 \\ .... \\ FVp \end{bmatrix} \tag{3.5}$$

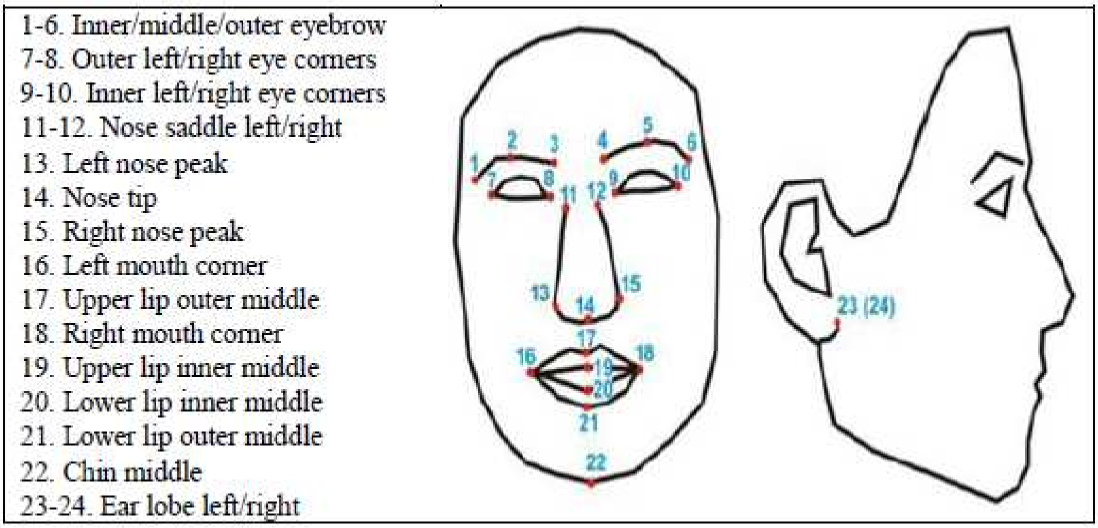


Figure 3.5: 3D Facial feature points contained in Bosphorus database [46].

Chapter 4

CLASSIFICATION MODELS FOR FACIAL EXPRESSION RECOGNITION

4.1 Support Vector Machine

Facial expression recognition can be considered as a vector classification problem after representing a face with facial feature points as a row vector. Thus, this classification problem requires a strong classifier. Support Vector Machine (SVM) is selected as the classifier as it shows high performance for vector classifications [61].

SVMs are supervised learning models that trains known data and recognize patterns. An SVM training method constitutes a classifier that assigns unknown samples into one of the trained classes. An SVM model first trained with a given set of training examples. Training phase results in a linear classifier that separates the training examples into known two classes. New examples are then mapped into that same space and assigned a category which they fall.

SVM constructs a hyper-plane in a high dimensional space that is to be used for classification. A good hyper-plane that separates the training examples into two classes is obtained by finding the largest distance to the nearest training data points of the classes.

The SVM algorithm was proposed by Corinna Cortes and Vladimir N. Vapnik in 1993 [62]. Later, it has been used by many researchers to solve classification problems like facial expression recognition [9, 61].

The facial expression recognition is a multi-class classification problem in which faces are classified as one of the six basic expressions. Thus, a multi-class classifier is needed. SVM can be modeled in order to be used in multi-class classification in two popular ways. The first way is to use one-by-one classifiers between every pair of samples and then applying majority voting strategy. The second way is to distinguish between one of the classes and the rest that is one versus all method. Then, a winner-takes-all strategy is applied where the classifier with the highest output function is selected. The multi-class SVM implementations which are one-versus-one and one-versus-all methods are explained in the sub sections 4.1.1 and 4.1.2 respectively.

4.1.1 Multi-Class SVM Classifier Model Using One-vs-One Approach

In the thesis study, the Ekman’s findings about the classification of facial expressions have been followed. According to Ekman’s classification, facial expressions can be categorized in six basic expressions which are anger, disgust, fear, happiness, sadness and surprise [3, 4]. Thus, these six expressions constitute the classes of the multi-class classification problem.

In two-class SVM classifier model, many SVMs are used in order to implement the multi-class approach. Considering six basic prototypic expressions, the system employs six classifier modules each consisting of fifteen 2-class SVM classifiers. Each SVM classifier uses a linear kernel function, dot product, which maps the training data into kernel space. The classifier module used for each expression is depicted in Figure 4.1. This classifier module uses an unknown facial feature point definition as a face vector and runs fifteen 2-class SVM classifiers. The fifteen 2-class classifiers are the result of all paired combinations of the six basic expressions, that are anger-disgust, anger-fear, anger-happiness, anger-sadness, anger-surprise, disgust-fear, disgust-happiness, disgust-sadness, disgust-surprise, fear-happiness, fear-sadness, fear-surprise, happiness-sadness, happiness-surprise and sadness-surprise classifiers. Then, majority voting is applied to determine the recognized expression. Each SVM classifier module is trained separately.

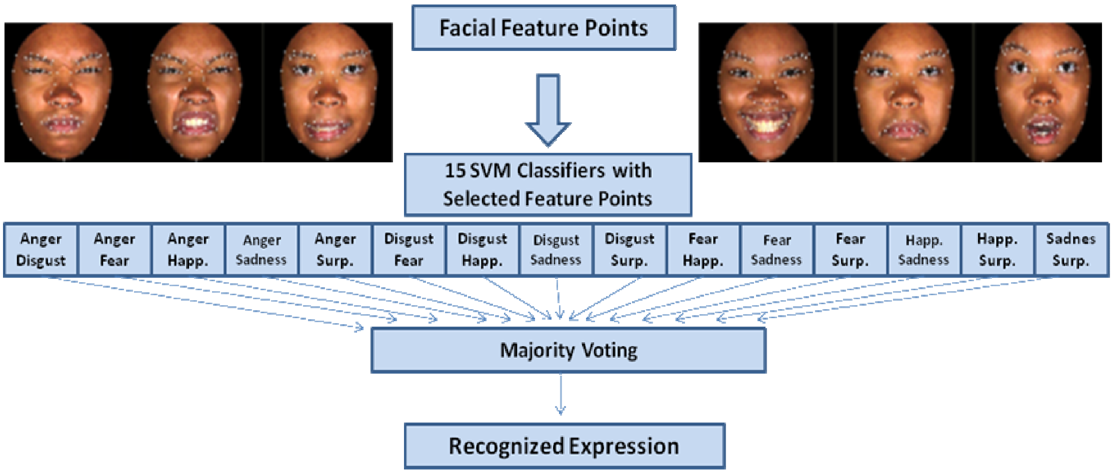


Figure 4.1: Multi-class SVM classifier module including 2-class classifiers for basic facial expressions with one-vs-one approach.

Initial test results are given in Table 4.1 using the multi-class SVM implementation given in Figure 4.2. All 83 3D facial feature point positions are used which are presented in BU-3DFE database. The intensity level considered in our experiments is level 4, which is the highest intensity. In total, face matrix *FM* (Chapter 3, Equation 3) is constructed by using 600 row vectors of 100 sample persons with 6 expressions describing 600 faces. The classifier model employed includes 15 two-class SVM classifiers including all the paired combinations of 6 expressions. These two-class classifiers are trained separately with 90% of the row vectors of *FM* and 10% of row vectors are used in testing. The test results are reported after applying 10-fold cross validation.

Table 4.1: Confusion Matrix for Recognition Rates of SVM Classifier in Figure 4.1 using 83 feature points.

Expression	**Recognition Rate (%)**					
	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**
Anger	**90**	0	0	0	10	
Disgust	0	**80**	10	10	0	0
Fear	10	10	**70**	10	0	0
Happiness	0	0	10	**90**	0	0
Sadness	10	0	0	0	**80**	10
Surprise	0	0	0	0	10	**90**
Overall	**83.33**					

4.1.2 Multi-Class SVM Classifier Model Using One-vs-All Approach

In this classifier model, the six basic expression classes are used in one-versus-all logic. Two-class SVMs are modeled in this manner to implement multi-class version. First, an unknown face vector is classified according to anger expression. The classification is done by classifying as anger or the class including all the rest of the expressions. If the result of the first classification is the second class (other 5 expressions), then the unknown face vector undergoes into second classification process. The second classification process includes disgust or the rest classification. If again the unknown vector is classified as the second class, it continues with the third classification process for fear expression and so on. The nested two-class classifiers implement the multi-class architecture. There are 5 classifiers in total which are anger – rest (disgust, fear, happiness, sadness, and surprise) classifier, disgust – rest (fear, happiness, sadness, and surprise) classifier, fear – rest (happiness, sadness, and surprise) classifier, happiness – rest (sadness, surprise) classifier, and sadness – surprise. The last classifier is a two-class classifier that classifies among sadness and surprise expressions. The resulting multi-class expression classifier is shown in Figure 4.2.

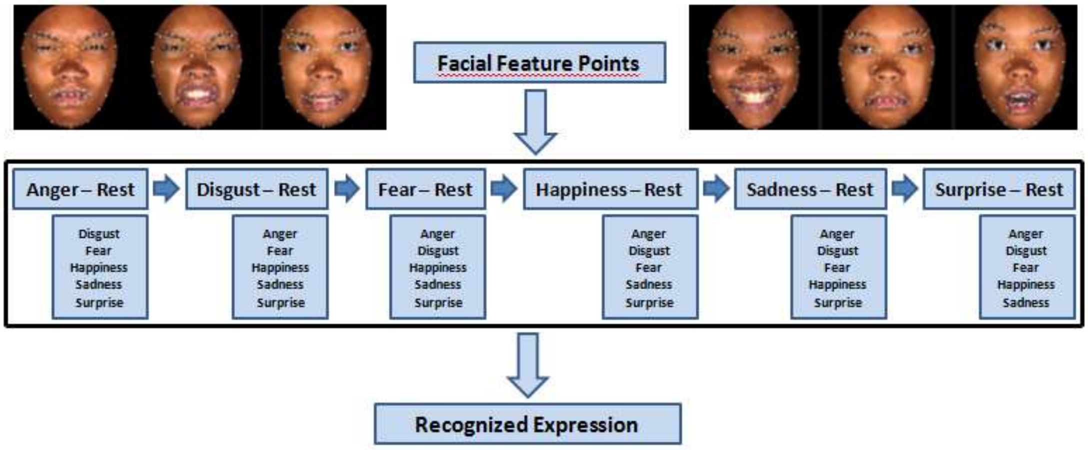


Figure 4.2: Multi-class SVM classifier module including 2-class classifiers for basic facial expressions with one-vs-all approach.

Initial test results are given in Table 4.2 using the multi-class SVM implementation given in Figure 4.3. All 83 3D facial feature point positions are used which are presented in BU-3DFE database with the intensity level 4. In total, face matrix *FM* (Chapter 3, Equation 3.3) is constructed by using 600 row vectors of 100 sample persons with 6 expressions describing 600 faces, same as the previous test. Each of the two-class classifiers is trained separately with 90% of the row vectors of FM and 10% of row vectors are used in testing. The test results are reported after applying 10-fold cross validation.

It is seen from Tables 4.1 and 4.2 that the performance of the multi-class SVM implementation using two-class model which is based on one-vs-one SVM classifiers and majority voting, is better than the other one. Also, the evaluation of multi-class SVM models in [67] states that one-vs-one model performs better. Thus, at this point the first multi-class SVM classifier model is selected and the rest of the study uses this model as the basis classifier. Using this classifier model, a coarse-to-fine classification model and the proposed expression distinctive classification models are developed for facial expression recognition which are explained in sections 4.3 and 4.4.

Table 4.2: Confusion Matrix for Recognition Rates of SVM Classifier in Figure 4.2 using 83 feature points.

Expression	**Recognition Rate (%)**					
	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**
Anger	**76.25**	6.25	5	0	12.5	0
Disgust	7.5	**80.00**	7.5	1.25	1.25	2.5
Fear	2.5	11.25	**68.75**	5	10	2.5
Happiness	1.25	0	12,5	**83.75**	1.25	1.25
Sadness	12.5	5	1.25	1.25	**80.00**	0
Surprise	0	3.75	3.75	1.25	0	**91.25**
Overall	**80.00**					

4.2 Fuzzy C-Means (K-Means) Clustering

Fuzzy C-means (FCM) or K-means clustering is a popular method for unsupervised learning algorithms. It divides the data into classes which is called the clustering process. The main aim here is to collect similar samples into the same class whereas dissimilar samples should appear in different classes. Depending on the nature of the data and the purpose for which clustering is being used; different measures of similarity may be employed. Distance is one of the most widely used examples of similarity measures.

In *FCM* clustering, rather than separating the data into strict classes, every sample has a degree of belonging to clusters. This is the result of fuzzy logic. Thus, samples on the edges may be in the cluster to a lesser degree than points in the center of cluster [66].

FCM clustering provides better results than k-means algorithm for overlapped datasets [68]. Also, a sample may belong to more than one cluster. These advantages of FCM clustering motivate the study to use it for facial expression clustering.

Each sample *x* has group of coefficients indicating the belonging to the k^{th} cluster. The centre sample of a cluster is the mean of all samples, weighted by degree of belonging, $w_k(x)$, shown in Equation 4.1.

$$c_k = \frac{\sum_x w_k(x)^m x}{\sum_x w_k(x)^m} \tag{4.1}$$

1. Begin – initialize number of samples (s), number of clusters (c), mean points and probabilities of belonging to clusters.
2. Normalize probabilities of belonging to clusters
3. Do
4. Classify n samples according to nearest mean
5. Compute mean again
6. Compute the probability (coefficient) of n belonging to clusters
7. While means and coefficients are not changing significantly

Figure 4.3: Fuzzy C-Means (FCM) algorithm.

Samples are classified into clusters according to a similarity metric. The algorithm minimizes with respect to this parameter. The most widely used similarity metric is the distance. The options can be listed as squared Euclidean distance, sum of absolute differences, cosine, in which one minus the cosine of the included angle between points, Hamming distance where it is applied to binary data and correlation, which is one minus the sample correlation between points. The FCM clustering approach is used in the thesis in order to cluster the expressions into two big classes. Correlation has been employed as the metric of distance which measures the similarity between the feature points. The FCM algorithm is given in Figure 4.3.

4.3 Two-Level Coarse-to-Fine Classification Model

Multi-class SVM implementations (Figures 4.1 and 4.2) provide acceptable recognition rates for facial expression recognition (Tables 4.1 and 4.2). However, from Tables 4.1 and 4.2 it is observed that the confusion matrix shows significant confusions between some expressions. For example, in Table 4.1 row 1 shows that 15% of anger expressions are classified as sadness. Similarly, 11.25% of fear expressions that are tested are classified as disgust. These confusions motivate a clustering study among the expressions.

In order to create a discriminative expression space, we adapt SVM classifier into two-level classification process. This proposal is motivated by preliminary test results showing confusions between classes. Moreover, Fuzzy C-Means (FCM) clustering algorithm has been employed among 6 basic expressions to investigate the correlations between the expressions. The algorithm separates the expression data into two clusters. As FCM parameters, exponent for the partition matrix is selected as 2.0 and the minimum amount of improvement as e-5. Correlation is selected as the measure of distance. With the number of iterations being 100, the clustering algorithm separates more than the halves of the expressions anger, disgust and sadness in cluster 1, whereas, more than the halves of the expressions fear, happiness and surprise are in cluster 2. However, with the increased number of iterations, we achieved the grouping of anger, disgust and fear expressions in cluster 1, and happiness, sadness and surprise expressions in cluster 2. The results of the FCM clustering algorithm on BU-3DFE expression data is given in Table 4.3. Values in Table 4.3 indicate the number of occurrences in each cluster, for every expression. There are 600 samples, 100 for each expression and the algorithm separates them into two clusters. Therefore, the grouping of anger, disgust and fear expressions in one class is the result of the FCM clustering algorithm and the confusions between them also supports this conclusion.

Table 4.3: Results Obtained Using Fuzzy C-Means Clustering.

No of Iterations /Expression	**100 Iterations**		**200 Iterations**		**500 Iterations**	
	Cluster 1	**Cluster 2**	**Cluster 1**	**Cluster 2**	**Cluster 1**	**Cluster 2**
Anger	**65**	35	**65**	35	**65**	35
Disgust	**51**	49	**55**	45	**54**	46
Fear	42	**58**	**58**	42	**58**	42
Happiness	45	**55**	45	**55**	44	**56**
Sadness	**52**	48	48	**52**	48	**52**
Surprise	45	**55**	45	**55**	45	**55**

Accordingly, we have organized the classifier system in two levels. In the first level, the unknown input vector is classified into one of the two main classes: Class 1 including anger, disgust and fear, and class 2 including happiness, sadness and surprise expressions. First classification step uses all of the 15 2-class SVM classifiers and employs majority voting among two main classes. Then, the input vector undergoes another classification process, and is classified as one of the six basic expressions in the second level of the classification: Class 1 or Class 2. Hence, an unknown face vector undergoes a two-level classification process and classified using SVM classifiers, as shown in Figure 4.4.

The second level of classification process uses 3 SVM classifiers for each main class including all the combinations of expressions involved. Similar to first level classification, the second level also employs majority voting among 3 expressions. Class 1 of second layer classifier includes anger-disgust, anger-fear and disgust-fear classifiers, whereas Class 2 includes happiness-sadness, happiness-surprise and sadness-surprise classifiers. The expression with the maximum number of classifications is selected as the recognized expression. The two level architecture of the classifier is depicted in Figure 4.4.

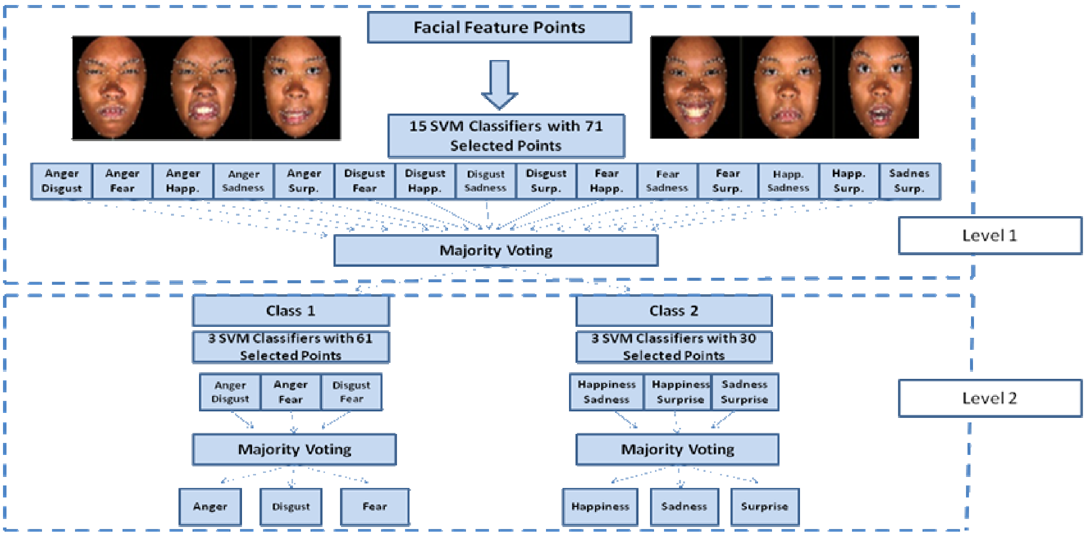


Figure 4.4: 2-Level SVM classifier system used for facial expression recognition.

4.4 Proposed Expression Distinctive Classification

Furthermore, the classification performance of an expression can be maximized by considering the expression distinctive features. Thus, in order to apply expression distinctive feature selection, there is a need for expression specific classifiers.

In proposed expression distinctive classification model, the classifier modules are designed for each of the basic expressions resulting in a total of 6 classifier modules.

Each module uses its own facial feature points which are the outcomes of entropy based feature selection algorithm. Then, each classifier module generates a decision as acceptance or rejection. Each classifier module generates acceptance if the unknown face vector is recognized as the expression name of the classifier module. Otherwise, it is rejected. For example, if the anger classifier module recognizes the unknown expression as anger, then it is accepted by the anger classifier module. If one of the other expressions is recognized in the anger classifier module then it generates a rejection. A similar process is completed for the other classifier modules. The overall classifier design is shown in Figure 4.5.

Obviously, an anger expression is expected to be accepted by the anger classifier module and rejected by the other five classifier modules and vice versa. But some of the test cases show that an unknown expression can be accepted by more than one classifier module. In case of multiple acceptances, the overall classification process goes through a decision module to classify the unknown expression as one of the accepted expressions. In order to achieve this final recognition, the decision module uses the 2-level SVM classifiers including the accepted expressions. If there are more than two accepted expressions, then the decision module employs all combinations of the 2-class SVM classifiers of the accepted expressions. The final recognition is achieved by majority voting. The decision module is depicted in Figure 4.6.

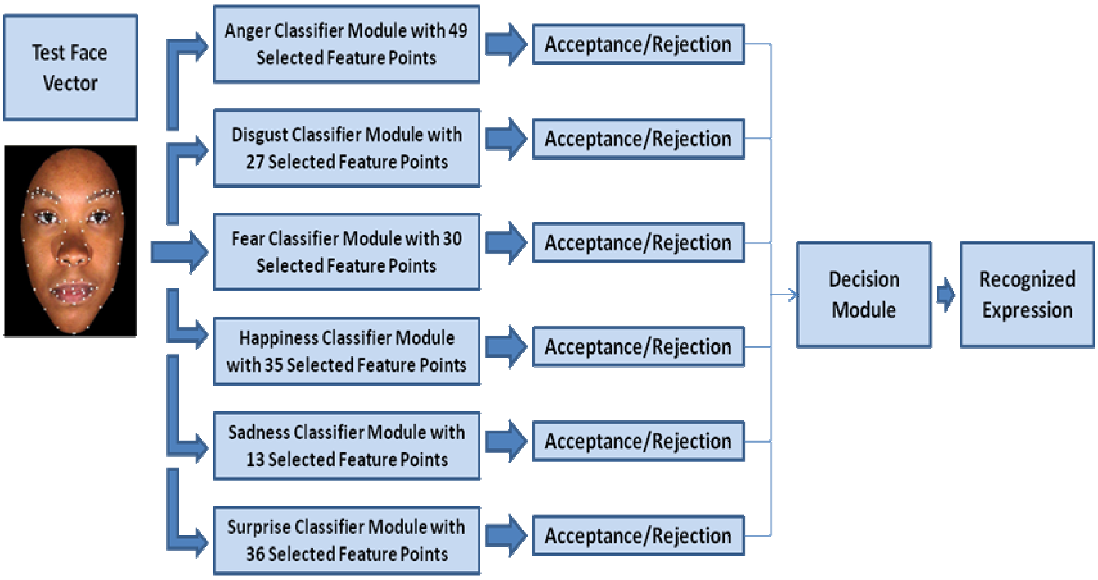


Figure 4.5: Accept-reject classifier system for facial expression recognition.

SVM classifiers use 3D facial feature point positions as the inputs. 3D facial features are arranged as a row vector and used for face definition vector. Each face vector consists of 3D facial feature positions describing a face. All selected facial feature positions are arranged as a row vector for each face. Consider the row vector definition of a facial feature point as shown in Equation 1, where V_i is a 3D vector definition of the i^{th} facial feature point and x, y, z refers to the dimensions of the 3D space respectively.

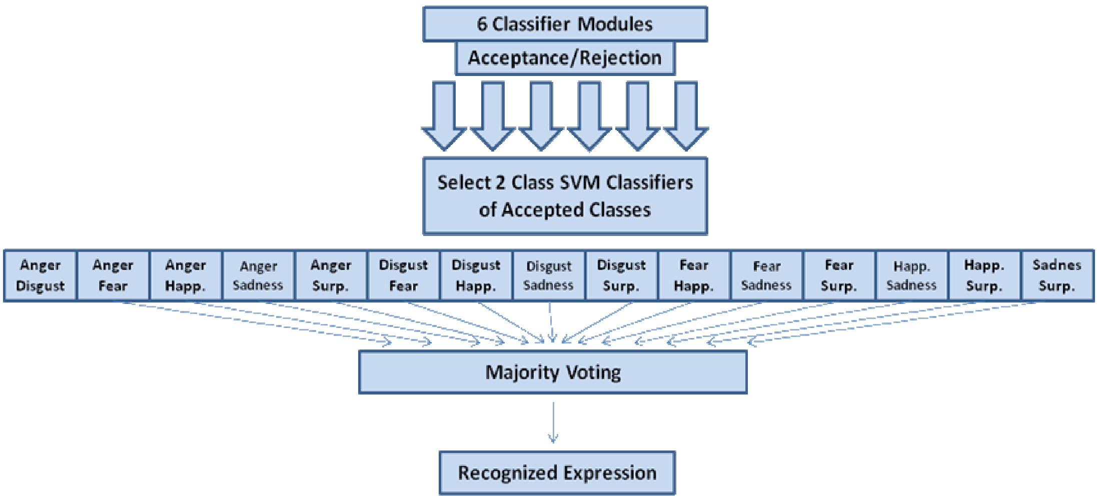


Figure 4.6: Decision module for final recognition of accepted expressions.

Chapter 5

FEATURE SELECTIONS FOR ENHANCED 3D FACIAL EXPRESSION RECOGNITION

5.1 Introduction

The selection of most discriminative features for expression recognition problem is also a critical stage. MPEG-4 defines FDPs and FAPs on a generic face model to represent basic facial movements as well as expressions [3]. Facial expressions are modeled with the deformations on the neutral face by using MPEG-4 FAPs [3]. The motivation of the novel feature selection method proposed in the thesis is therefore the working principle of MPEG-4 FAPs.

FAPs represent a complete set of basic facial actions including head motion, tongue, eye and mouth control. Our previous studies about facial expression modeling and synthesis [1, 2] showed that the action of the FAPs on the whole 3D generic face model is obtained by displacement functions for each vertex. The displacement functions $\Delta\, i\ (x, y, z)$ of a vertex are computed according to the position of the vertex in the influence area, to the intensity of the FAP in the related feature point(s) and an additional weight for design issues as shown in Equation 5.1 [3].

$$\begin{bmatrix} \Delta i_x \\ \Delta i_y \\ \Delta i_z \end{bmatrix} = \begin{bmatrix} Wi_x \\ Wi_y \\ Wi_z \end{bmatrix} * W^{'} j * FAP_{x,y,z} \tag{5.1}$$

The weight Wi is based on the distance of the vertex from the feature point and the weight spreads decreasingly from the centre of the influence area. If we want that vertex not to be affected by the FAP, $W'j = 0$ may be chosen.

It is seen from MPEG-4 FAP implementation that although most of the FDPs are affected among 6 basic expressions, some FDPs are weakly influenced or even not affected by facial expression deformations. This fact justifies the motivation of the feature selection procedure which is applied to extract the most informative feature points that can identify facial expressions.

5.2 Variance and Entropy

Variance and entropy are strong metrics that have been used to measure uncertainty and information content. They show similar verdicts on information content of distributions when the data are informative on the mean with known variance, such as Gaussian distribution [60].

Variance measures how set of numbers is distributed. A variance of zero indicates that all the values are identical. It is a non-negative value. Being a small value means the data is spread around the mean. High variance means that the data is distributed away from the mean and from the others. The variance of a random variable X is the squared deviation from the mean (expected value) μ = $E[X]$ as shown in Equation 5.2.

$$Var(x) = E[(X - \mu)^2] \tag{5.2}$$

When we consider facial expression recognition based on 3D geometric positions of the facial feature points, variance metric can extract the feature points which are active during expression deformations. The variance based feature selection algorithm is explained in Section 5.4.

Entropy is a measure of randomness, in other words, it is a measure of unpredictability of information content. The origin of entropy concept arises from the studies of Ludwig Boltzmann in 1877. Entropy has been given a probabilistic interpretation in information theory by Claude Shannon in 1948. In the 1960s, Henri Theil developed applications of information theory in economics collected in Economics and Information Theory (1967) and Statistical Decomposition Analysis (1972) based on entropy [63].

Entropy conveys the expected information content of a probability distribution. Let X_i stand for an event and p_i for the probability of the event X_i. Let there be n events $X_1 , \ldots, X_n$ with probabilities $p_1 ,\ldots, p_n$ where summation of all p_i gives 1. Since the occurrence of events with smaller probability means more information, a measure of information H is a decreasing function of p_i. In 1948, Shannon proposed a logarithmic function to express information $H(X)$, X with possible values of x_1 to x_n given in equation 5.3. In Equation 5.3, b is the base of the logarithm (common values are 2 and 10) [63].

$$H(X) = -\sum_i P(x_i) \log_b P(x_i) \tag{5.3}$$

The proposed feature selection process employs information content metric, which is entropy that selects the most informative feature points for facial expression recognition. Higher entropy refers to the feature point which is affected significantly from the expression deformation. Hence, high entropy feature points contain more information on the expression. Extracting and tracking only the selected high entropy features help improving the recognition performance. Therefore, features are selected with the assumption that the level of entropy is correlated with the level of information that respective feature point carries.

5.3 Fisher's Criterion

Fisher's Linear Discriminant Analysis (LDA) is a method of projecting data into a line in order to find the optimum separation. After projection, Fisher's LDA performs classification on the projected line. The criterion is to find the maximum distance between the means of the two classes on the projected line. At the same time, within class variance is minimized. This defines the Fisher's criterion which is widely used in pattern recognition applications including facial expression recognition [8]. The criterion function J is shown in Equation 5.4 for all projections of w where m represents the mean, and s^2 represents variance. The subscripts denote the class number. Rather than using variances of the samples, a common method is to define scatter for the projected samples. Then the two matrices which are called within-class scatter matrix and between-class scatter matrix are used. A simple scalar measure of scalar is the determinant of the matrix. Using this measure, the Fisher's criterion function J becomes the ratio of determinant of the scatter matrices, given in the Equation 5.5. The Fisher's criterion used in the thesis can be expressed as φ in Equation 5.5 where $|S_B|$ stands for the determinant of between class scatter matrix and $|S_W|$ stands for the determinant of within class scatter matrix [65].

$$J(w) = \frac{|m_1 - m_2|}{s_1^2 + s_2^2} \tag{5.4}$$

$$\Phi = |S_B| \,/\, |S_W| \tag{5.5}$$

Fisher's criterion given in Equation 5.5 is computed for several classes and maximized for feature selection procedures.

5.4 Variance Based Feature Selection for Expression Recognition

Variance based feature selection method employs variance as the metric of information content. During facial expression deformations, facial feature points that have high variance are selected by assuming that they carry more information about the expressions. In order to achieve this selection of the facial feature points which are highly affected by facial expression deformations, the variances of each feature point in 3D for the 6 basic expressions and a neutral face are measured and analyzed. Equation 5.6 is used to compute the variances where *X* is a 3D feature point position and μ is the mean value for the related feature point among 6 basic expressions and neutral.

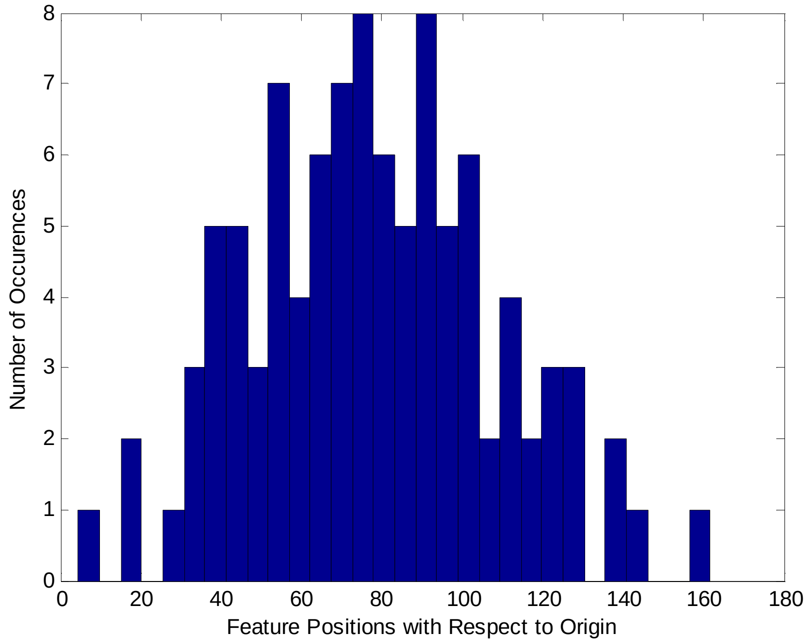
$$Var(x) = E[(X-\mu)^2] \tag{5.6}$$

The feature selection algorithm starts with the distributions of 83 feature points. Feature point positions are used in 3D, shown in Figure 5.3. The 3D data of each feature point is transformed into a magnitude value by using Equation 5.7. Then, the face matrix *FM* takes the form *FM(D)* in Equation 5.9 including *FV(D)* values for each face vector shown in Equation 5.8. For each distribution, considering the minimum and the maximum feature point magnitude value, a histogram is generated.

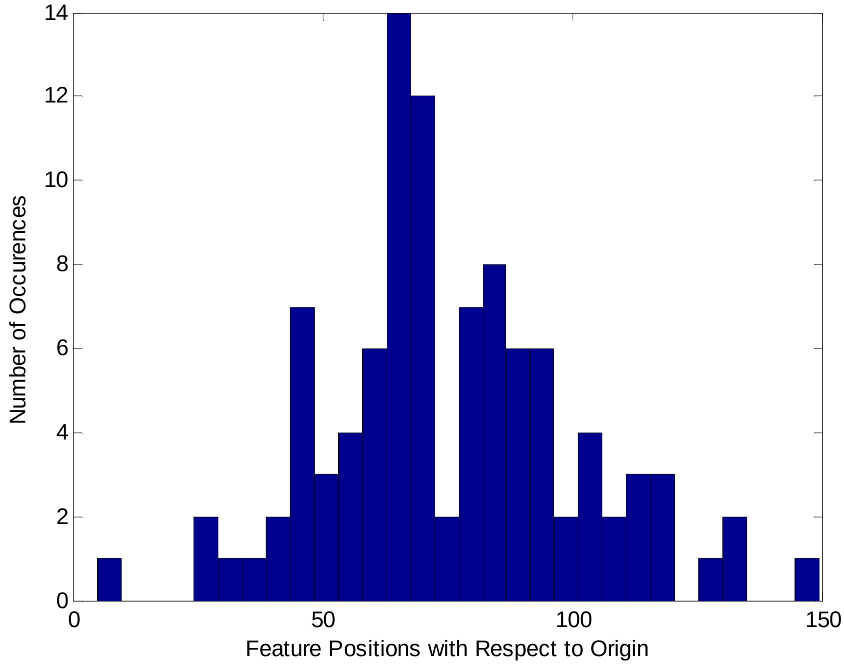
$$V_i(D) = \sqrt{V_{i_x}^{\ 2} + V_{i_y}^{\ 2} + V_{i_z}^{\ 2}} \tag{5.7}$$

$$FV(D)_j = \left[V_1(D) \quad V_2(D) \quad V_3(D) ....V_k(D)\right] \tag{5.8}$$

$$FM(D) = \begin{bmatrix} FV_1(D) \\ FV_2(D) \\ .... \\ FV_n(D) \end{bmatrix} \tag{5.9}$$



(a)



(b)

Figure 5.1: Histograms of (a) the highest variance (b) the lowest variance feature points.

The histograms of feature points with the highest and the lowest variance among 6 basic expressions are shown in Figure 5.1. Training samples from BU-3DFE database are used in variance analysis. The database includes 100 samples with 6 basic expressions of intensity level 4 (highest intensity). 90% of these samples are used for training and 10% are used for testing. The training and testing sets are selected in the same way that they are selected in the state of the art methods which are compared. So, as a result, 90 training samples are employed in the variance analysis. For each feature point, the variance of its positions among neutral and six basic expressions is computed by using Equation 5.6. These variance values are then sorted in descending order. Figure 5.2 shows the decreasing graph of sorted variance values of feature points where *x-axis* denotes sorted index of feature point and *y-axis* represents variance.

It is seen from Figure 5.2 that after sorting the variances in descending order, there are feature points having relatively higher variances with respect to other points when face is deformed from neutral to any one of the expressions. Thus, the algorithm continues and eliminates low variance feature points in face vectors and *Face Matrix* shown in Equation 5.9 is formed accordingly. The elimination of low variance points is done by considering significant decreases in variances. The points in which there is a sudden decrease in variance more than the standard deviation are considered as breaking points. There are 8 breaking points observed in Figure 5.2 in this sense. These breaking points are the 9. , 17. , 27. , 36. , 44. , 48. , 55. and 71. points. The selected breaking points indicate the feature points to be used in forming the face vectors.

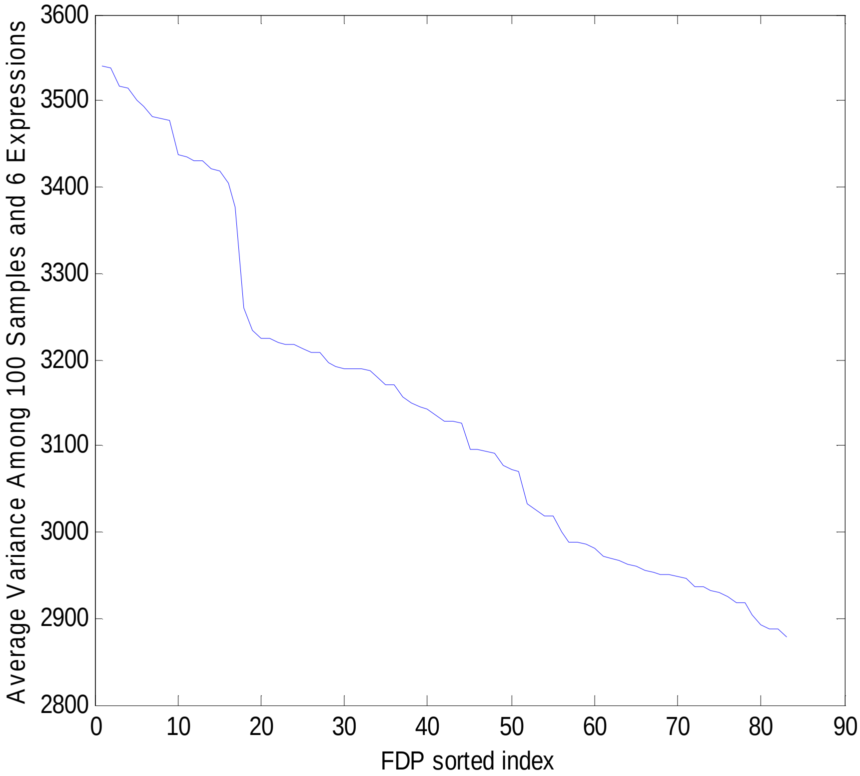
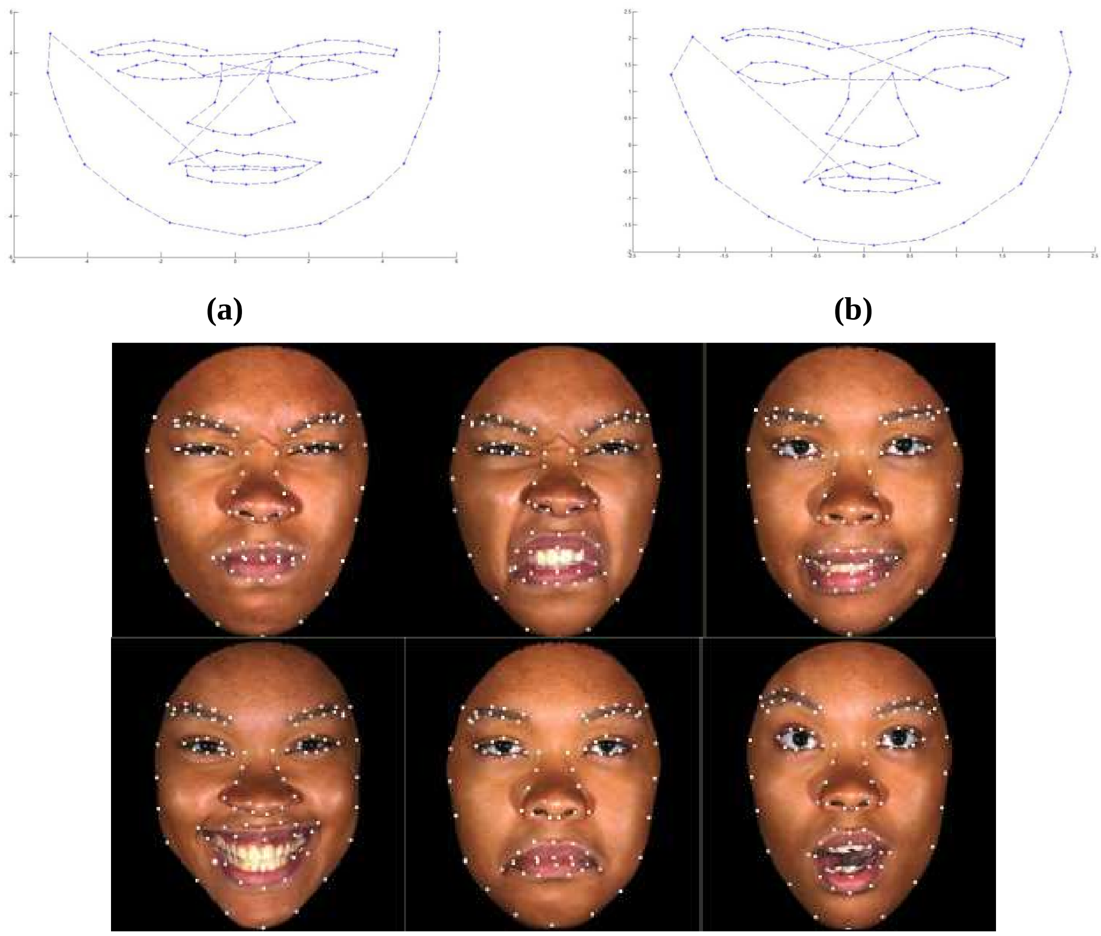


Figure 5.2: Variance analysis of 83 feature points of 85 training samples from BU-3DFE database among neutral and 6 basic expressions.

Then, the feature selection algorithm runs a brute force search by adding next feature points to the selected features and reports acceptance rates for every feature point combination corresponding to breaking points. As a result, the feature selection algorithm picks the first 71 high variance feature points providing the highest rates.



(a) (b)

(c)

Figure 5.3: 83 3D Facial feature points used from BU-3DFE database [45]. (a) and (b) two different sequences given in database files which are pre-processed in the experiments (c) feature points on static image samples.

5.4.1 Performance Analysis on BU-3DFE Database

The performance of variance based feature selection algorithm is measured on BU-3DFE database [6] and recognition rates are reported among six basic facial expressions which are anger, disgust, fear, happiness, sadness and surprise.

Face matrix, *FM*, is constructed as described in section 2 including 100 samples with 6 basic expressions. In total, matrix includes 600 row vectors describing 600 faces. The maximum intensity is selected for expressions from available 4 levels of intensities given in BU-3DFE database.

Table 5.1: Recognition rates for SVM Classifier (Figure 4.1) using all 83 feature points.

Expression	**Recognition Rate (%)**
Anger	93.33
Disgust	86.67
Fear	60.00
Happiness	93.33
Sadness	80.00
Surprise	86.67
Overall	**83.33**

15 2-class SVM classifiers are trained separately with 85% of the row vectors of *FM* and 15% of row vectors are used in testing. As described in section 4.1.1, facial expressions are classified using 15 2-class SVM classifiers and applying majority voting among 6 expression classes.

Table 5.1 presents recognition rates for 6 basic expressions using the classifier explained in section 4.1.1 with 83 feature points. To validate the test results, tests are repeated for two different combinations of training and test sets which are the first 85% for training and the last15% for the testing, and the last 85% for training and first 15% for the testing. The recognition performance is reported is the average of these tests. Also, in [64], the recognition performance of 3D feature points provided in BU-3DFE database by using Linear Discriminant Aanalysis (LDA) classifier is reported as 83%. Table 5.1 verifies the recognition rates reported in [64]. In Table 5.2, improvements in the recognition rates after applying variance based feature selection procedure can be observed. Overall recognition rate is improved from 83.33% to 85.55%.

Table 5.2: Recognition Rates Using 71 Selected Feature Points.

Expression	Recognition Rate (%)
Anger	100
Disgust	86.7
Fear	60
Happiness	93.3
Sadness	80
Surprise	93.3
Overall	**85.55**

It is seen from Table 5.1 and 5.2 that fear expression has an average recognition rate around 60% which is not in the acceptable range. The main reason for this low average recognition rate is that fear and sadness expressions are highly related and their geometric feature positions are close to each other. These kinds of correlations motivate the algorithm to improve a 2-level coarse-to-fine classification process as described in section 4.3 where the highly correlated expressions are grouped.

Table 5.3: Comparison between variance based feature selection algorithm results and other methods tested on BU-3DFE database.

Expression	Recognition Rates (%)				
	Soyel et al.[35]	**Wang et al.[22]**	**Mpiperis et al.[58]**	**Tang et al.[57]**	**Var. Based Feat. Sel.**
Anger	85.9	80	83.6	86.7	100
Disgust	87.4	80.4	100.0	84.2	86.7
Fear	85.3	75.0	97.9	74.2	60
Happiness	93.5	95.0	99.2	95.8	93.3
Sadness	82.9	80.4	62.4	82.5	80
Surprise	94.7	90.8	100.0	99.2	93.3
Overall	**88.3**	**83.6**	**90.5**	**87.1**	**85.55**

The comparison of the variance based feature selection method with the current methods in the literature which are also tested on BU-3DFE database is given in Table 5.3. It is seen from Table 5.3 that variance based feature selection method achieves competitive recognition rates compared to the current systems in the literature.

Although variance based feature selection algorithm improves the average recognition rate of the 6 basic facial expressions, the fear expression is recognized with a moderate rate whereas anger and surprise expressions are recognized with high rates. This result signals a new direction in order to improve the recognition rate of fear expression while keeping others near. On the other hand, variance is a naive method of measuring information content. Instead, a strong information content metric, which is entropy, is employed and further improvements are achieved in the next sections.

5.5 Proposed Entropy Based Feature Selection for Expression Recognition

Entropy conveys the expected information content of a probability distribution. The proposed feature selection algorithm employs entropy as information content metric. Features are selected with the assumption that the level of entropy is correlated with the level of information that respective feature point carries.

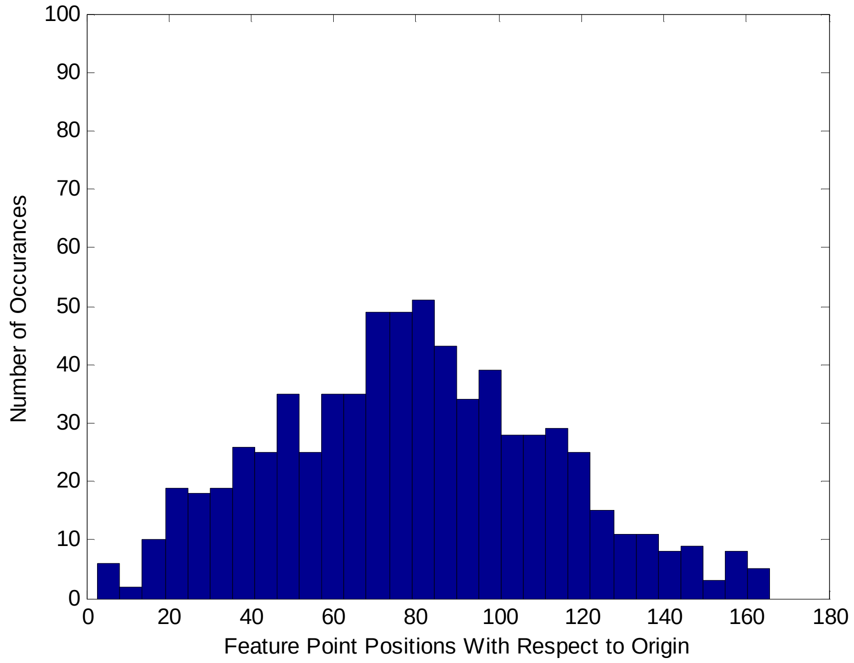
5.5.1 Feature Selection for Coarse-to-fine Classifier on BU-3DFE Database

During facial expressions, facial feature points that have high entropy are selected by assuming that they carry more information about the expressions. Therefore, entropy of each feature point for the 6 basic expressions and a neutral face are analyzed.

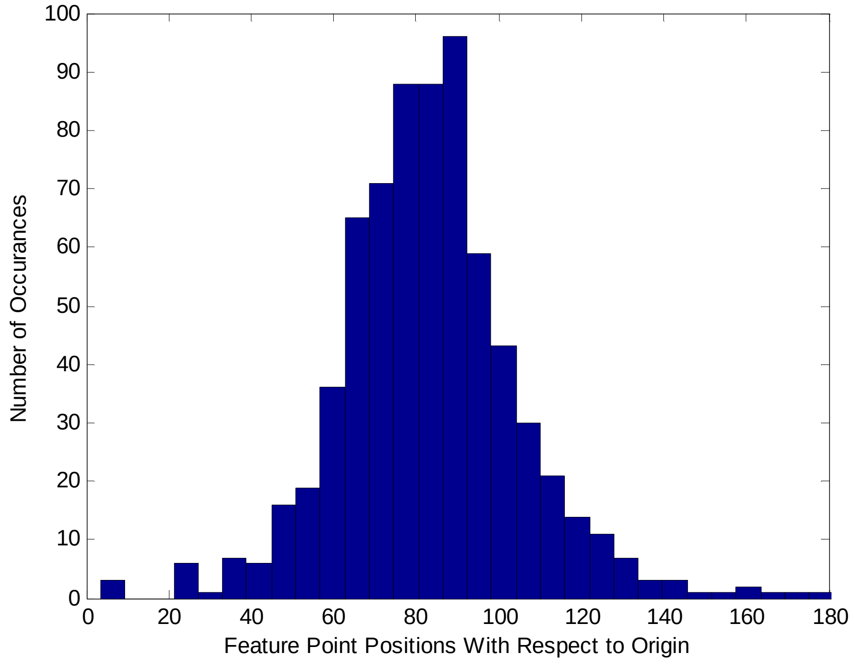
Consider the histograms of highest and lowest entropy feature points for training samples among 6 basic expressions shown in Figure 5.4. Similarly, all feature point distributions under the six basic expressions are analyzed. Then, entropies of each feature point for the 6 basic expressions are computed, and the overall entropy values of the feature points are analyzed by the use of Equation 5.10, where H is non-negative. Entropy analyses are completed on the training samples on BU-3DFE database by computing the entropy of each feature point among six basic facial expressions.

Feature selection algorithm is applied to coarse-to-fine classifier model which is explained in Chapter 4, Section 4.3. Figure 5.5 shows the classifier model used. The first level uses feature points selected from entropy analysis of all of the expressions. The first class of the second layer classifier uses feature points selected from entropy analysis of anger, disgust, fear expressions whereas the second class uses feature points selected from entropy analysis of happiness, sadness and surprise expressions.

The feature selection algorithm starts with the distributions of 83 feature points among six basic expressions. The 3D data of each feature point is transformed into a magnitude value by Equation 5.8. Then, the face matrix *FM* takes the form $FM(\delta)$ in Equation 6 including $FV(\delta)$ values for each face vector shown in Equation 5.11. Each feature point distribution is observed according to the 6 basic expressions. Therefore, employing 90 training subjects for the selected expressions, histograms are calculated for λ training vectors from $FM(\delta)$ matrix shown in Equation 5.12. For each distribution, considering the minimum and the maximum feature point magnitude value, a histogram is generated with 10 equally spaced bins from minimum to maximum. Probabilities for each bin are calculated resulting in 10 different probability values for each feature point magnitude value shown in Equation 5.11 where the value λ is the number of training samples used. This λ value is 540 for the entropy analysis of 6 basic expressions that there are 90 training subjects with 6 expressions resulting in 540.



(a)



(b)

Figure 5.4: Histograms of (a) highest entropy (b) lowest entropy feature points among 6 basic expressions.

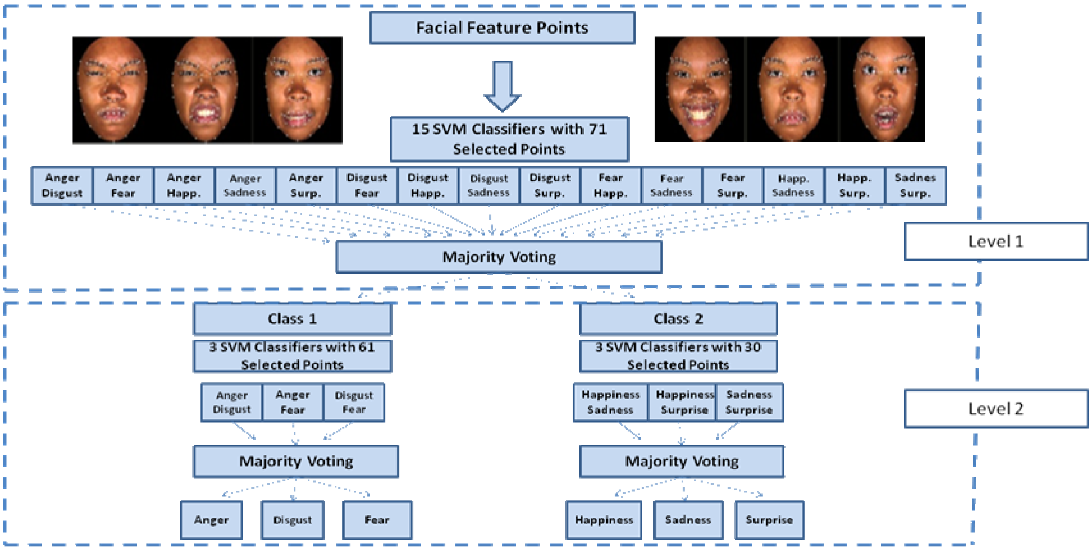


Figure 5.5: 2-Level coarse-to-fine SVM classifier model.

$$\delta_i = \sqrt{V_{i_x}{}^2 + V_{i_y}{}^2 + V_{i_z}{}^2} \tag{5.10}$$

$$FV(\delta)_j = \lfloor \delta_1 \quad \delta_2 \quad \delta_3 .... \delta_k \rfloor \tag{5.11}$$

$$FM(\delta) = \begin{bmatrix} FV_1(\delta) \\ FV_2(\delta) \\ .... \\ FV_n(\delta) \end{bmatrix} \tag{5.12}$$

$$P(a) = \frac{\sum_{i \in a} V_i(\delta)}{\lambda} \tag{5.13}$$

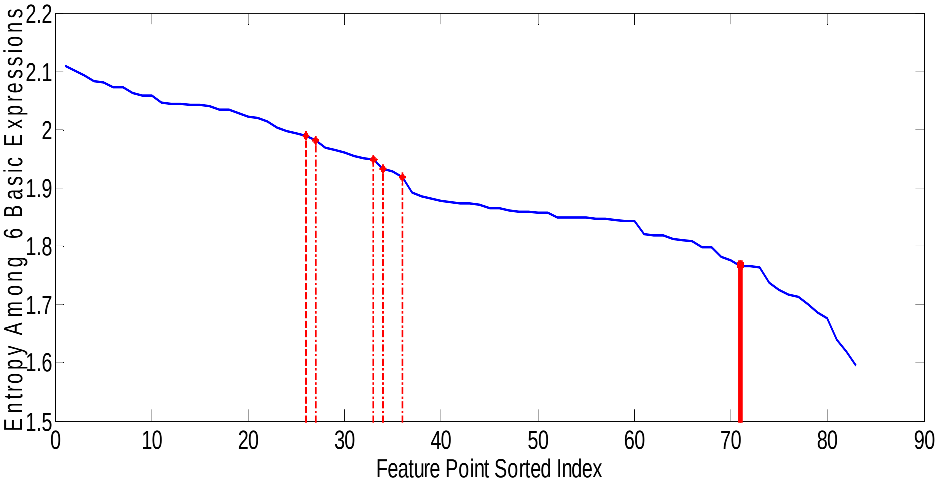
Similarly, λ is chosen as 270 for the second level classifier entropy analysis. Each class in the second level includes 3 expressions. Using 90 subjects with 3 expressions makes 270 face vectors in total.

$$H(a) = -\sum_{a \in A} P(a) \log P(a) \tag{5.14}$$

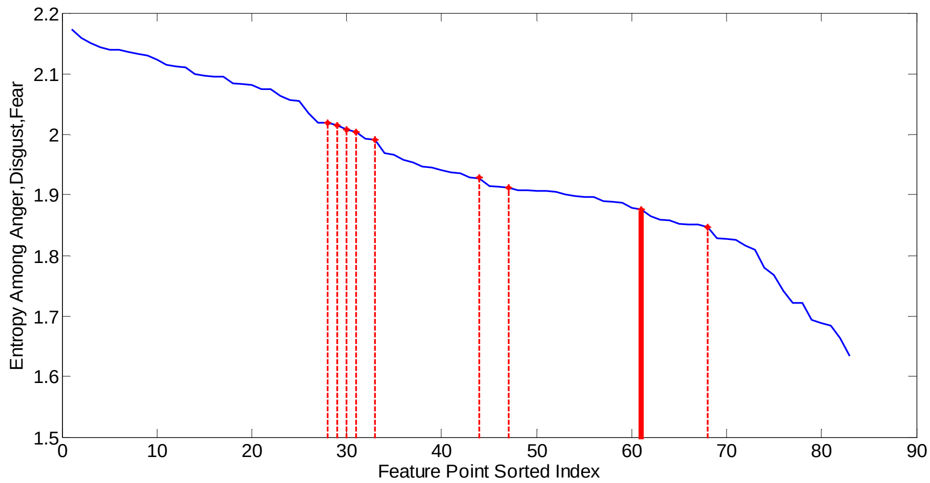
Then, the entropy formula in Equation 5.14 [13] is applied where *a* is the bin index, *A* represents the set of bins and *P(a)* is the probability of a feature point magnitude value for the container *a*, out of 10 different bins. *P(a)* value is calculated by Equation 5.13 where *i* stands for index number of facial feature magnitude values falling in the corresponding bin *a*. Entropy of each feature point is computed distinctively for each phase of the classification. The entropy values are then arranged for each class and sorted in descending order resulting in 3 decreasing functions of feature point index numbers. Figure 5.6 (a), (b) and (c) shows the graphs of sorted entropy values of feature points for each class.

It is seen from Figure 5.6 (a), (b) and (c) that after sorting the entropy values in descending order, there are feature points having significantly higher entropy values than the others when face is deformed to any one of the 6 basic expressions. Thus, we eliminate low entropy feature points in face vectors for our recognition tests and formed *Face Matrix* shown in Equation 5.12 accordingly.

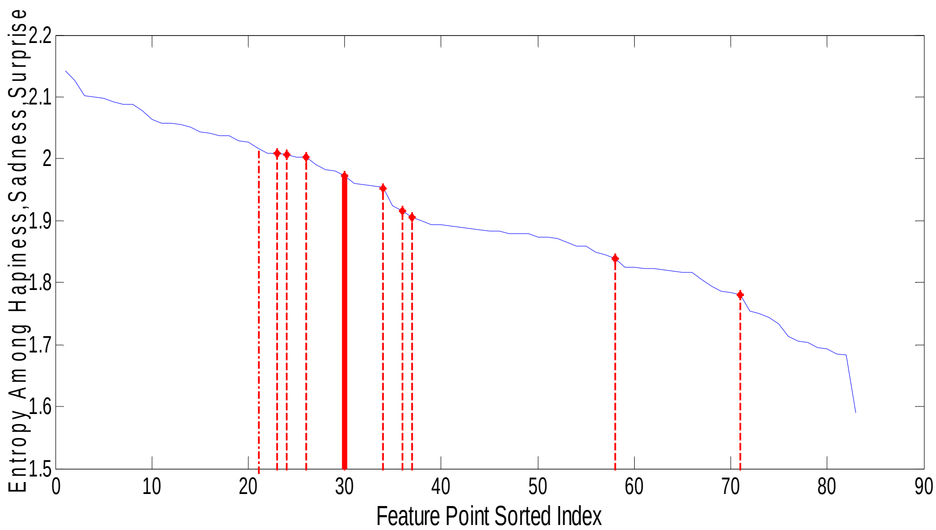
Analyzing the entropy values sorted in descending order, we see that there are breaking points in which the entropy value shows significant decreases which are more than the standard deviation. Because of 2-level classifier implementation, we employ entropy analysis for 6 basic expressions for the first level, and for the 3 basic expressions included in the second level classes separately. This process results in different features selected for each level of classifier. The breaking points among 6 basic expressions, which are for the first level of the classifier, are the first 26, 27, 33, 34, 36 and 71 feature points. Then, the algorithm selects the breaking point which maximizes the Fisher Criterion, φ, given in Equation 5.15 where $|S_B|$ stands for the determinant of between class scatter matrix and $|S_W|$ stands for the determinant of within class scatter matrix. Between and within classes are selected according to groupings shown in Figure 2.



(a)



(b)



(c)

Figure 5.6: Entropy analysis of feature points for 90 training samples **(a)** among 6 basic expressions **(b)** among anger, disgust and fear expressions and **(c)** among happiness, sadness and surprise expressions. Breaking points are marked with dashed lines and the selected breaking point with a solid line.

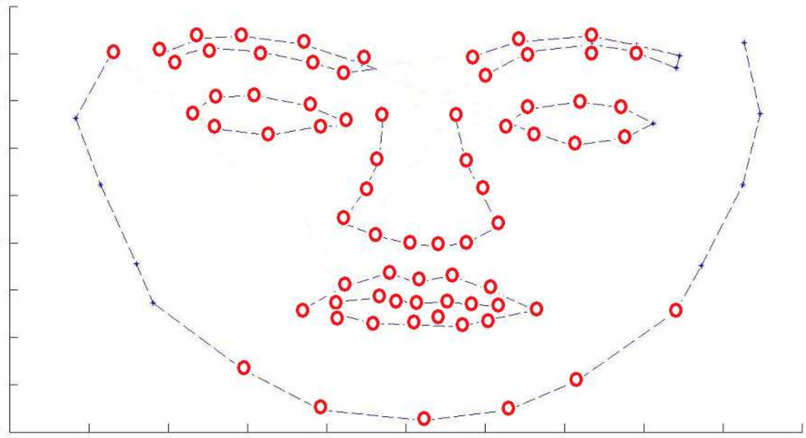
As a result, the feature selection algorithm picks the first 71 high entropy feature points providing the highest Fisher criterion value for the first level classifier.

$$\Phi = |S_B| / |S_W| \tag{5.15}$$

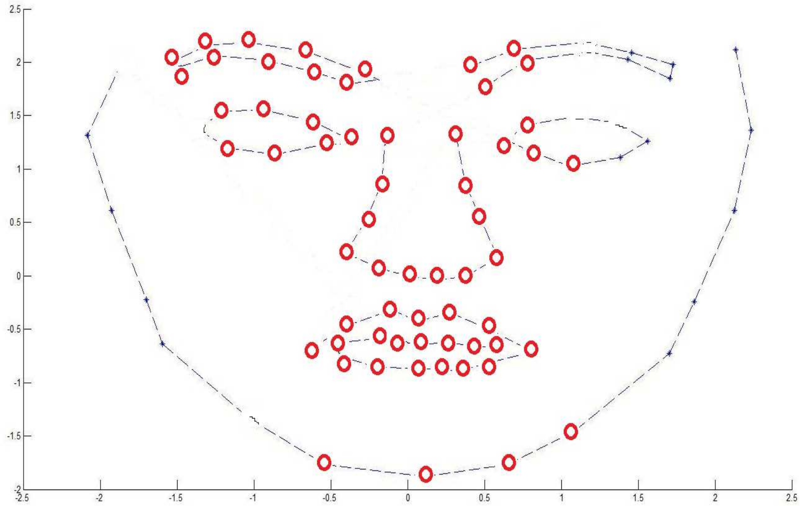
For level 2, we employ similar entropy analysis but now among the expressions included in each class. This results in entropy analysis for anger, disgust and fear expressions and happiness, sadness and surprise expressions. The corresponding entropy graphs are provided in Figure 5.6 (b), (c). Similar to first level feature selection, the entropy values are sorted in descending order and breaking points are extracted according to standard deviation. The breaking points are first 28, 29, 30, 31, 33, 44, 47, 61 and 68 feature points for the Class 1 of second level. Class 2 of second level includes the first 21, 23, 24, 26, 30, 34, 36, 37, 58 and 71 feature points as the breaking points. Then these breaking points which maximize the Fisher Criterion given in Equation 5.15 are selected similar to the level 1 classification.

First 61 feature points for the Class 1, and the first 30 feature points for the Class 2 are selected based on the maximum Fisher Criterion value.

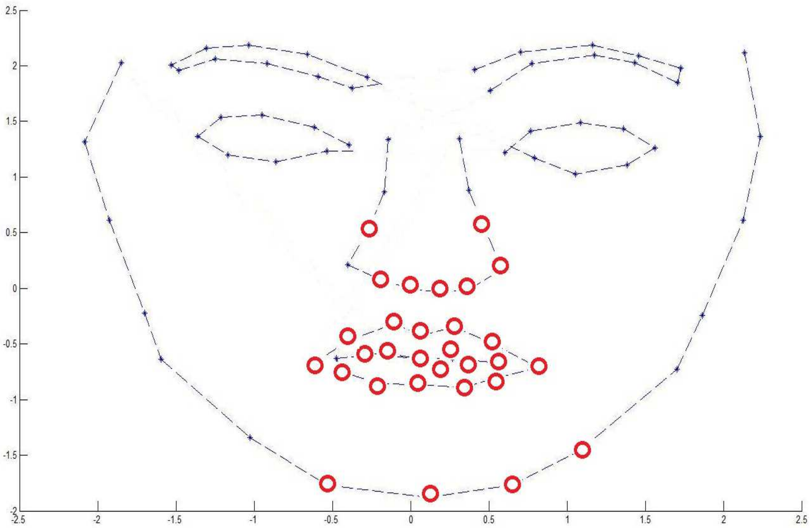
In order to calculate within and between class scatter matrices for the second level, multi-class Fisher Criterion is applied. As a result of feature selection process, the two-level classifier uses 71 points which are the outcomes of the overall entropy analysis in the first level. 61 high entropy points are selected for Class 1 including anger, disgust, fear expressions. 30 high entropy feature points are selected for Class 2 including happiness, sadness and surprise expressions. The selected feature points are illustrated in Figure 5.7 (a), (b) and (c) on the neutral face.



(a)



(b)



(c)

Figure 5.7: Selected feature points: **(a)** 71 feature points for first classifier level **(b)** 61 feature points for Class 1 of the second classifier level **(c)** 30 feature points for Class 2 of the second classifier level.

It is observed from Figure 5.7 (a), (b) and (c) that mostly the mouth points are affected from the expression deformations. The points on the outline of the face are eliminated as a result of feature selection algorithm. The movements of the head cause less distortion to the global feature points of the head (i.e. outliers) than the local feature points of the head (i.e. points around the nose tip). Also, some of the points on the chin are selected as a result of feature selection algorithm. Due to dynamic joints that chin is connected; some of the expressions lower the chin position more than the other expressions. Most of the subjects with fear have wide opened mouth as part of the expression which would affect the chin position. For example, when chin is lower, there is higher probability to have fear.

It is also observed from Figure 5.7 (a), (b) and (c) that the points on the chin are selected for the first level which includes all the expressions, and mostly the Class 1 of the second level classification which includes anger, disgust and fear. This is due to the presence of fear expression samples in these classes. Observing the features selected for the Class 2 including happiness, sadness and surprise, we see that features include few chin points because of the chin movements in the surprise expression samples. The surprise expression is the most discriminative expression according to others. It includes the most significant feature point changes as well as high chin point movements with the open mouth. So, few points on the chin are enough to discriminate the surprise expression from happiness and sadness in the second level classification.

If we further analyze the findings of feature selection algorithm, we can see that the nose points also contribute to the facial expressions. Although most of the points along the nose are expected to be static with negligible entropy, most of the points along the sides of the nose are affected due to the dynamics of disgust expression. We see that only seven nose points are affected in the Class 2 of the second level classification process due to the dynamics of happiness and surprise expressions. The first level and the Class 1 of the second level classification include most of the nose points because of the presence of disgust expression.

5.5.1.1 Performance Analysis

The performance of the proposed feature selection procedure for facial expression recognition system is tested on BU-3DFE database [45]. Recognition rates are reported among six basic facial expressions.

FM matrix is constructed as described in Section 5.5.1 in Equation 5.12 including 100 subjects with 6 basic expressions, containing 600 samples which are presented in BU-3DFE database. BU-3DFE database includes facial expression data of 100 subjects with 4 different levels of intensity. In the training and test phases, expressions with the highest intensity which is 4 are selected. In total, *FM* matrix includes 600 row vectors describing 600 faces. 15 SVM classifiers are trained separately with 90% of the row vectors of *FM* matrix with expression intensity being 4, and 10% of row vectors with intensity being 4 and all of the remaining intensities are used in testing. In the first level classification, the classifiers are trained and tested using 71 selected features. In the second level, the classifiers in Class 1 are trained and tested with the 28 selected feature points, and the classifiers in Class 2 are trained and tested with 21 selected feature points. Training and test sets are selected similar to the other approaches that are compared [22, 35, 57, 58]. A sample which is tested never appears in the training set.

As described in Section 5.5.1 and shown in Figure 5.5, we classify expressions in two classes using 15 2-class SVM classifiers and apply majority voting among 6 expression classes in the first level. Then, in the second level, we employ 3 SVM classifiers for each class to recognize one of the six basic expressions. Class 1 includes anger-disgust, anger-fear and disgust-fear classifiers, class 2 includes happiness-sadness, happiness-surprise and sadness-surprise classifiers. To verify the experimental results, 8-fold cross validation is applied.

In Table 5.4, the classification performance of the first level (depicted in Figure 5.5) classifiers are reported by using the selected feature points. The results in the first level classification rates indicate the classification performance for the first two classes containing anger, sadness, surprise expressions in the first class and containing disgust, fear, happiness expressions in the second class. Second level classification rates indicate the performance of the classifiers in the second level as Class 1 and Class 2. Table 5.5 reports the recognition rates of individual expressions after second level of classification (Figure 5.5). Improvements in the recognition rates after applying our proposed feature selection procedure can be observed from Table 5.6 in which correct classification rates are reported for the overall system. Average recognition rate is improved to 88.28% after feature selection.

The recognition of fear expression is improved from 60% to 76.55% after using the selected features. Fear expression is recognized with lower rate than the others but it is still in the usable range. Surprise and happiness expressions are recognized with the highest rates. This can be explained by the significant movements on the mouth and chin during happiness and surprise expressions. Anger expression is recognized with a high rate around 90.19%. Disgust expression is the second lowest recognized expression which is also in high rate such as 85.17%. Sadness expression is now recognized with a high rate of 88.89%.

Table 5.4: Recognition Rates for the proposed system: First level recognition rates for class 1 and class 2 classifications refer to Figure 5.5.

Expression	**Recognition Rates (%)**	
	Class 1	**Class 2**
Anger	97.50	2.50
Disgust	98.75	1.25
Fear	88.75	11.25
Happiness	3.75	96.25
Sadness	10.00	90.00
Surprise	2.50	97.50
Average Correct Classification	**94.79**	

Table 5.5: Recognition rates for the proposed system: Confusion matrix for second level recognition refers to Figure 5.5 for each individual expression.

Expression	**Recognition Rates (%)**					
	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**
Anger	**92.50**	2.50	5.00	-	-	-
Disgust	1.25	**86.25**	12.50	-	-	-
Fear	6.875	6.875	**86.25**	-	-	-
Happiness	-	-	-	**97.50**	-	2.50
Sadness	-	-	-	1.25	**98.75**	-
Surprise	-	-	-	2.50	-	**97.50**
Average Correct Classification	**93.13**					

Table 5.6: Recognition rates for individual expressions.

Expression	**Recognition Rates (%)**
Anger	**90.19**
Disgust	**85.17**
Fear	**76.55**
Happiness	**93.84**
Sadness	**88.89**
Surprise	**95.06**
Average Recognition Rates	**88.28**

The proposed system is compared with the 4 recent systems as shown in Table 5.7. We see from Table 5.7 that the proposed system achieves comparable recognition rates compared to the current systems in the literature which are also tested on BU-3DFE database. Sadness expression is recognized with the highest rate compared to other methods. Besides, the proposed system obtains high recognition rates for the happiness and surprise expressions which are both around 95%. Disgust expression is recognized with the second highest rate compared to other person independent methods. Mpiperis et al. [58] reported the highest rate, which recognizes disgust expression without confusion, 100%. Fear expression is recognized with 76.55% rate which is still in the acceptable range. But considering the recognition rate of the fear expression after applying the proposed feature selection procedure, we see that it is significantly improved. This result also shows the advantage of the proposed novel feature selection procedure.

Table 5.7: Comparison of the Proposed Recognition System on BU-3DFE Database.

	Recognition Rates (%)				
	Soyel et al.[35]	**Wang et al.[22]**	**Mpiperis et al.[58]**	**Tang and Huang.[57]**	**Proposed**
Database	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**
Training / Test	**90% / 10%**	**90% / 10%**	**90% / 10%**	**90% / 10%**	**90% / 10%**
Person Dependency	**Independent**	**Independent**	**Independent**	**Independent**	**Independent**
Anger	85.9	80	83.6	86.7	90.19
Disgust	87.4	80.4	100.0	84.2	85.17
Fear	85.3	75.0	97.9	74.2	76.55
Happiness	93.5	95.0	99.2	95.8	93.84
Sadness	82.9	80.4	62.4	82.5	88.89
Surprise	94.7	90.8	100.0	99.2	95.06
Average	**88.3**	**83.6**	**90.5**	**87.1**	**88.28**

Furthermore, observing the recognition rates of the other methods listed in Table 5.7, we see that some expressions are recognized with high rates, and some other are reported with low rates, resulting in an acceptable average recognition rate. For the proposed method, it is obviously seen from the Table 5.7 that most of the expressions show high recognition rates above 85% which is a desired result for expression recognition systems. Only the fear expression is recognized with the recognition rate below 85% but still it stays in an acceptable rate.

5.5.2 Feature Selection for Expression Distinctive Classifier on BU-3DFE Database

The proposed expression distinctive feature selection is also based on entropy. Facial expressions are modeled with the deformations on the neutral face. Thus, the facial feature points which are affected from the expression deformations mostly are those being discriminative features. The assumption is same with the coarse-to-fine model that a feature point with low entropy implies a point which is not dynamic during expression deformations. Whereas, a feature point having high entropy means it is dynamic during expression deformations. Considering the entropy of a feature point among neutral and one of the basic expressions, more entropy means more discrimination power for that expression. Therefore, the most discriminative feature points concerning each of the six basic expressions are selected according to entropy values measured among neutral and the expressions distinctively. This method provides different facial feature combinations being discriminative for each expression.

Similar to the previous method, a feature point is modeled with a probabilistic distribution when the face model is deformed with the six basic expressions. However, in this method, each expression is analyzed specifically with neutral. Consider the histogram of the highest and the lowest entropy point during anger expression given in Figure 5.8. Similarly, the distributions of each feature point are analyzed under each specific expression.

The feature selection algorithm starts with the distributions of 83 feature points as in the previous method. Then, the 3D data of each feature point is transformed into a magnitude value by Equation 5.10 similarly with the previous method. Then, the face matrix *FM* takes the form *FM(δ)* including *FV(δ)* values for each face vector as shown in the previous method, Equation 5.11. The neutral expression and the six basic expressions are used and for each expression the distribution is observed with the neutral expression. Therefore, employing 90 training samples for the selected expression and the neutral face, histograms are calculated for 180 training vectors from *FM(δ)* matrix shown in Equation 5.16.

$$FM(\delta) = \begin{bmatrix} FV_1(\delta) \\ FV_2(\delta) \\ \ldots\ldots \\ FV_n(\delta) \end{bmatrix} \tag{5.16}$$

$$P(a) = \frac{\sum_{i \in a} V_i(\delta)}{\lambda} \tag{5.17}$$

$$H(a) = -\sum_{a \in A} P(a) \log P(a) \tag{5.18}$$

For each distribution, considering the minimum and the maximum feature point magnitude value, a histogram is generated with 10 equally spaced bins from minimum to maximum. Probabilities for each bin are calculated resulting in 10 different probability values for each feature point magnitude value shown in Equation 5.17 where the value λ is the number of training samples used which are 180. Then, entropy formula in Equation 5.14 is applied where *a* is the bin index, *A* represents the set of bins and *P(a)* is the probability of a feature point magnitude value for the bin *a*, which are used for 10 different bins. *P(a)* value is calculated by Equation 5.17 where *i* stands for index number of facial feature magnitude values falling in the corresponding bin *a*. Entropy of each feature point is computed distinctively for each basic expression using Equation 5.18. Then, the algorithm provides 6 entropy values for each feature point. These entropy values are then arranged and sorted in descending order resulting in 6 decreasing functions of feature point numbers. Figure 5.9 shows the graph of sorted entropy values.

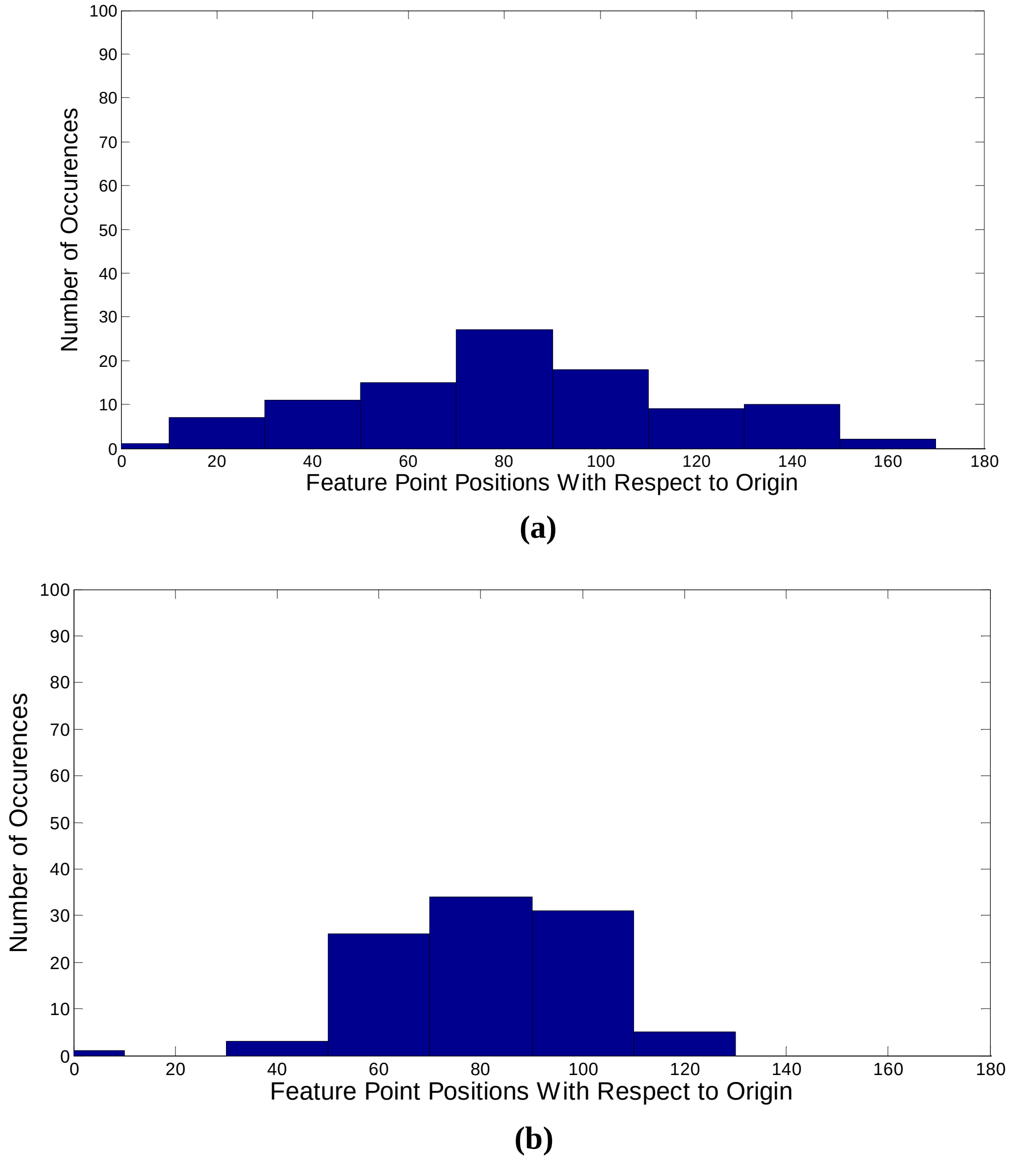
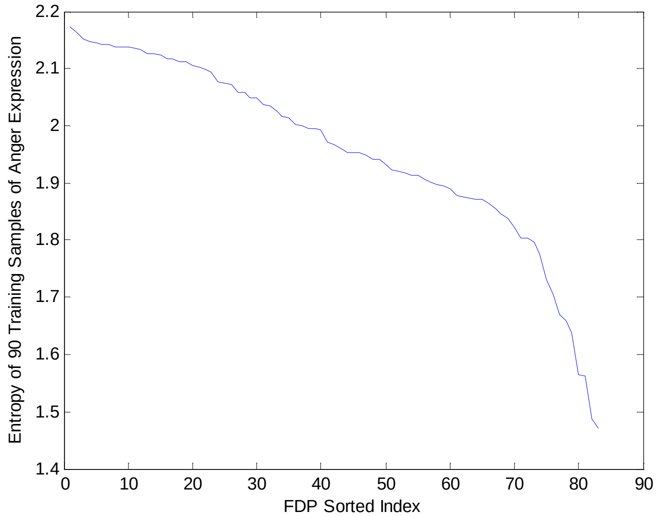
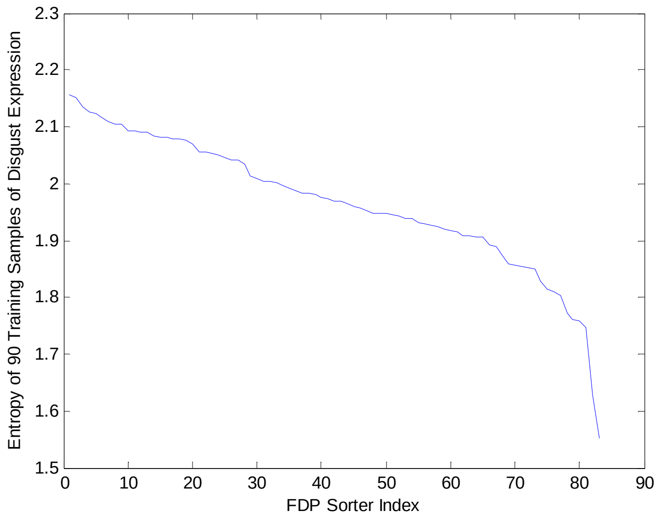


Figure 5.8: Histograms of (a) highest entropy (b) lowest entropy feature points for anger expression.

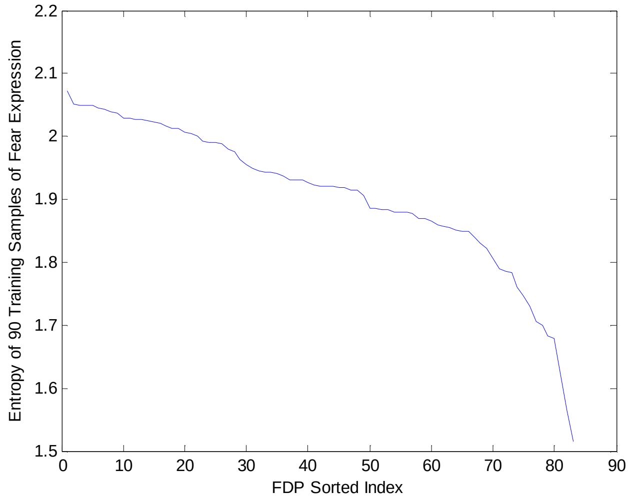
The feature selection algorithm searches for the best feature point sets by analyzing the decreasing functions of entropies. Starting from the highest entropy point, the feature selection algorithm runs a brute force search by adding next feature points to the selected features.



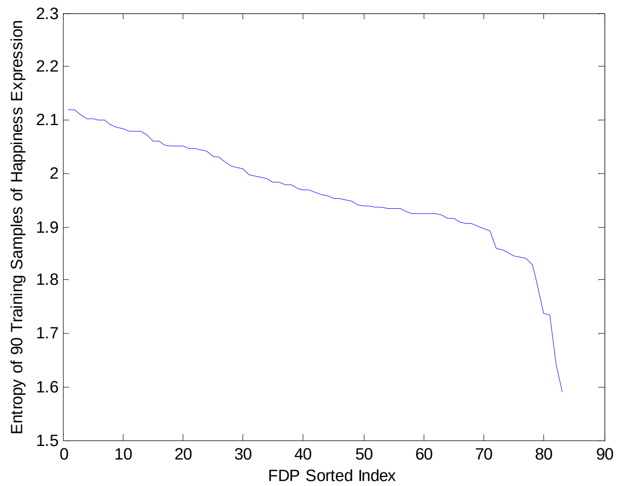
(a)



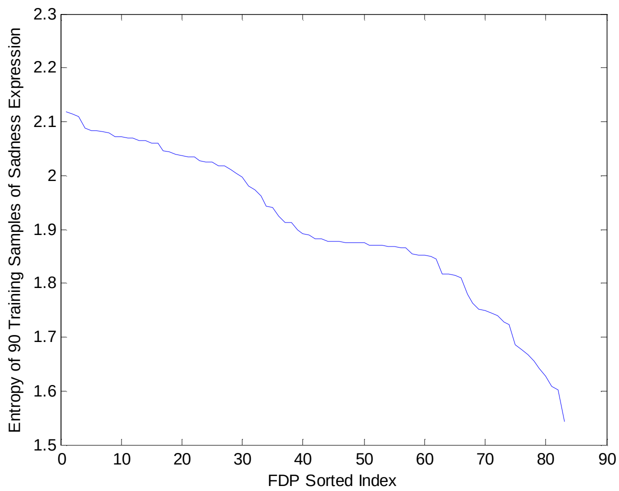
(b)


Entropy of 90 Training Samples of Fear Expression
FDP Sorted Index
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
0
10
20
30
40
50
60
70
80
90

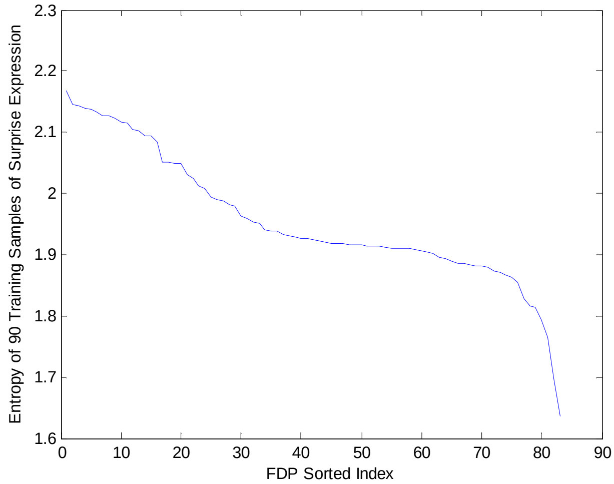
(c)


Entropy of 90 Training Samples of Happiness Expression
FDP Sorted Index
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
0
10
20
30
40
50
60
70
80
90

(d)


Entropy of 90 Training Samples of Sadness Expression
FDP Sorted Index
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
0
10
20
30
40
50
60
70
80
90

(e)

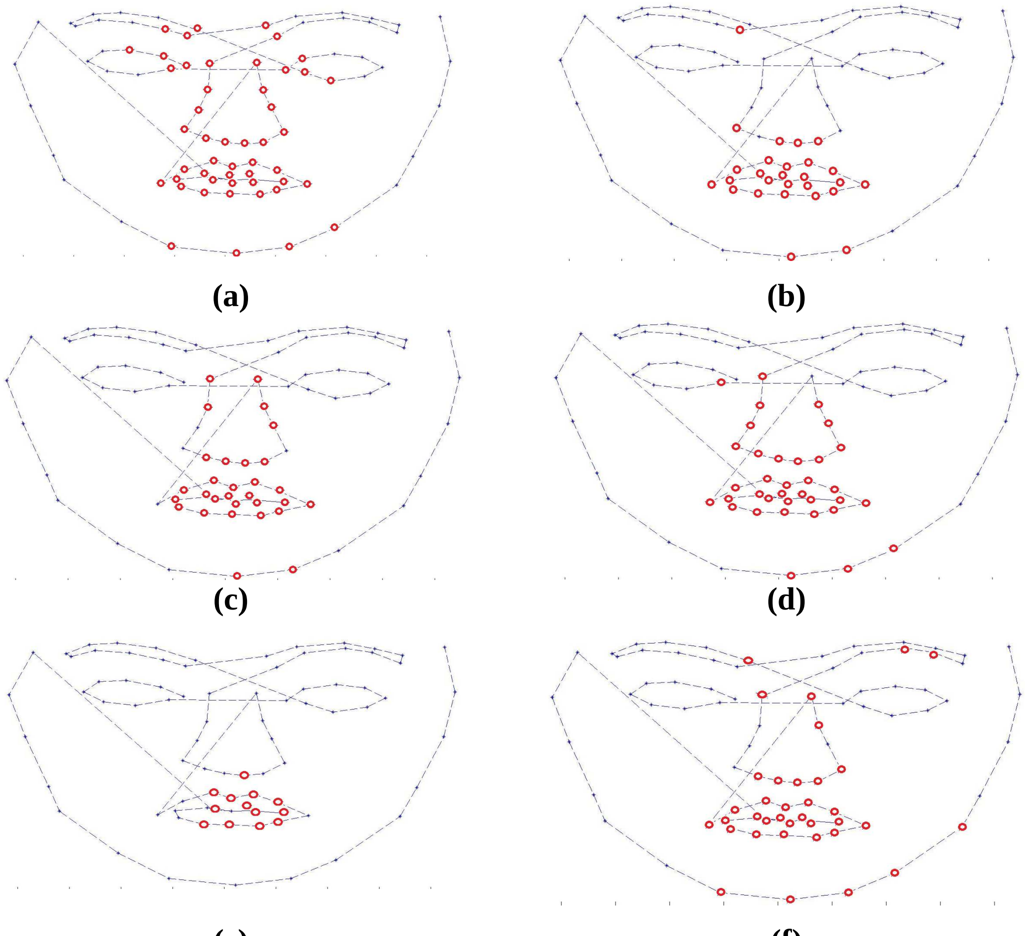


(f)

Figure 5.9: Entropy analysis of 83 feature points for 90 training samples among neutral face and 6 basic expressions sorted in descending order: **(a)** Anger, **(b)** Disgust, **(c)** Fear, **(d)** Happiness, **(e)** Sadness, **(f)** Surprise.

The algorithm reports Fisher criterion value which is computed by Equation 5.19 for each feature point combination. S_B is the between class scatter matrix and S_W is the within class scatter matrix. Scatter matrices are computed in one-versus-all manner, where a specific expression is picked from the training set forming the first class, and the remaining expressions form the second class. Between and within class scatter matrices are computed in this way. The algorithm runs brute force search for all the other expressions to form six different feature point combinations. The proposed algorithm selects 49, 27, 30, 35, 13 and 36 feature points for the six basic expressions, anger, disgust, fear, happiness, sadness and surprise respectively. The selected feature positions are shown in Figure 5.9. Then, the selected feature points for each expression are then employed in the corresponding accept-reject classifier module. Figures 5.11 and 5.12 show the classifier model and the decision module used, which was explained in Chapter 4. Different expressions employ different set of selected feature points. Anger classifier module is using 49 selected feature point positions whereas the sadness classifier module uses 13 selected feature point positions and etc.

$$\Phi = |S_B| / |S_W| \tag{5.19}$$



(a) (b) (c) (d) (e) (f)

Figure 5.10: Outcomes of expression distinctive feature selection algorithm: The most discriminative facial features for **(a)** Anger, **(b)** Disgust, **(c)** Fear, **(d)** Happiness, **(e)** Sadness, **(f)** Surprise expressions selected from 83 feature points.

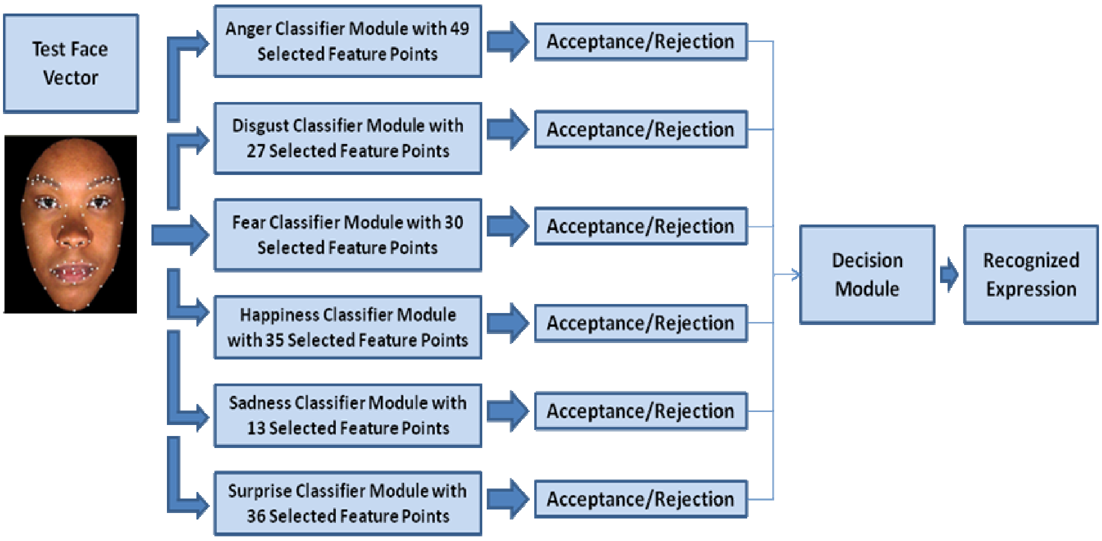


Figure 5.11: The proposed expression distinctive classifier scheme.

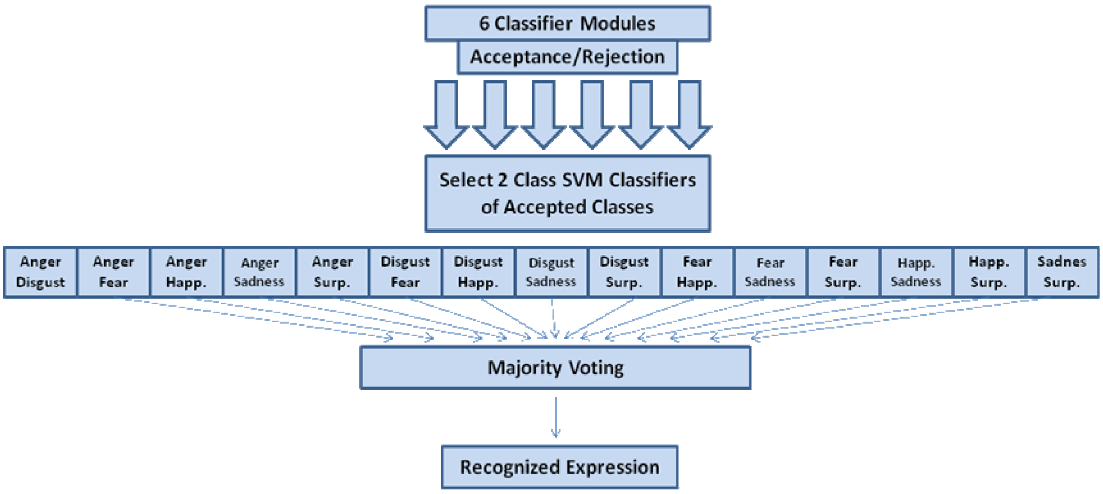


Figure 5.12: Decision module for the final recognition of accepted expressions in expression distinctive classification.

5.5.2.1 Performance Analysis

The performance analysis of the proposed expression distinctive feature selection algorithm is completed on BU-3DFE database and provided promising results. Experimental setup is similar with the previous method, in which 100 samples with 6 different expressions are considered. Intensity level 4 is selected for each expression. In total, face matrix *FM* (Equation 5.16) is constructed from 600 row vectors of 100 sample persons with 6 expressions describing 600 faces. Two-class classifiers of each accept-reject classifier module are trained separately with 90% of the row vectors of *FM* and 10% of row vectors are used in testing. To validate the test results, tests are repeated for two different combinations of training and test sets which are the first 90% for training and the last10% for the testing, and the last 90% for training and first 10% for the testing. The recognition performance is reported as the average of these tests.

Table 5.8: Confusion Matrix after Applying the Proposed Expression Distinctive Feature Selection Method.

Expression	**Recognition Rate (%)**					
	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**
Anger	**90%**	0	0	0	10%	0
Disgust	0	**80%**	10%	10%	0	0
Fear	0	10%	**80%**	10%	0	0
Happiness	0	0	5%	**95%**	0	0
Sadness	0	0	10%	0	**90%**	0
Surprise	0	0	5%	0	0	**95%**
Overall	**88.33%**					

In Table 5.8, the recognition rates of the proposed facial features with SVM classifier system is given as confusion matrix. Improvements in the recognition rates after applying the proposed feature selection procedure can be observed from Table 5.8. The results in the classification rates with entropy driven selected feature point subsets indicate less confusion than using all of the 83 feature points. Considering the recognition rates of each expression separately, there is a significant increase in the performance. Although the recognition performance of anger and disgust expressions remains unchanged, there is obvious increase in the recognition performance of the other expressions. Surprise and sadness expressions are recognized with around and above 90%. The performance increase in sadness reaches to 10%. Similarly, the recognition of fear and happiness expressions shows significant improvements. The overall recognition rate is improved to 88.33% when applying entropy based feature selection.

The performance of the proposed selected features based facial expression recognition method is compared with 4 other methods in the literature. The comparisons are based on the average recognition rates as shown in Table 5.9. The results given in Table 5.9 show that the proposed system achieves higher average recognition rate compared to the current systems in the literature tested using the same database. The performance of the proposed method is considered to provide high recognition rates for all expressions and high overall recognition rate.

Table 5.9: Performance Comparison of the Proposed Expression Distinctive Feature Selection Method.

	Recognition Rates (%)				
	Soyel et al.[35]	**Wang et al.[22]**	**Mpiperis et al.[58]**	**Tang et al.[57]**	**Expression Distinctive Feat. Sel.**
Database	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**
Training / Test	**90% / 10%**	**90% / 10%**	**90% / 10%**	**90% / 10%**	**90% / 10%**
Person Dependency	**Independent**	**Independent**	**Independent**	**Independent**	**Independent**
Anger	85.9	80	83.6	86.7	90.0
Disgust	87.4	80.4	100.0	84.2	80.0
Fear	85.3	75.0	97.9	74.2	80.0
Happiness	93.5	95.0	99.2	95.8	95.0
Sadness	82.9	80.4	62.4	82.5	90.0
Surprise	94.7	90.8	100.0	99.2	95.0
Average	**88.3**	**83.6**	**90.5**	**87.1**	**88.33**

5.5.3 Feature Selection for Coarse-to-fine Classifier on Bosphorus Database

The feature selection algorithms which are variance based and entropy based (coarse-to-fine method) are also tested on Bosphorus database. The algorithms are applied as explained in Sections 5.4 and 5.5 by using feature points provided in the Bosphorus database.

There are 105 individuals in Bosphorus database but because of limitations 40 samples are selected with six basic expressions. The limitations include the absence of several feature point data, varying feature point locations and the absence of samples with all the six basic expressions.

For the selected 40 samples, all feature point distributions under the six basic expressions are analyzed. Then, entropies of each feature point for the 6 basic expressions are computed, and the overall entropy values of the feature points are analyzed similar to Section 5.5.1 by the use of Equation 5.10, where H is non-negative. Entropy analyses are completed on the training samples. Again training and testing set ratio is 90% and 10% respectively.

Variance based feature selection algorithm is applied which is explained in Section 5.4. Variance based algorithm detects the first 9, 11, 12, 14 and 21 high variance feature points. Then it selects the first 14 high variance features and the performance is reported in Table 5.10. An improvement of 1.25% overall recognition rate is achieved (see Table 5.10).

Entropy based feature selection algorithm is applied to coarse-to-fine classifier model which is explained in Chapter 4, Section 4.3. Figure 5.5 shows the classifier model used. The first level uses feature points selected from entropy analysis of all of the expressions. The first class of the second layer classifier uses feature points selected from entropy analysis of anger, disgust, fear expressions whereas the second class uses feature points selected from entropy analysis of happiness, sadness and surprise expressions.

The algorithm runs similarly to the one explained in the Section 5.5.1. The feature selection algorithm starts with the distributions of 83 feature points among six basic expressions. The 3D data of each feature point is transformed into a magnitude value by Equation 5.8. Then, the face matrix *FM* takes the form *FM(δ)* in Equation 6 including *FV(δ)* values for each face vector shown in Equation 5.11. Each feature point distribution is observed according to the 6 basic expressions. Therefore, employing 90% training subjects which makes 36 samples for the selected expressions, histograms are calculated for λ training vectors from *FM(δ)* matrix shown in Equation 5.12. For each distribution, considering the minimum and the maximum feature point magnitude value, a histogram is generated with 10 equally spaced bins from minimum to maximum. Probabilities for each bin are calculated resulting in 10 different probability values for each feature point magnitude value shown in Equation 5.11 where the value λ is the number of training samples used. This λ value is 216 for the entropy analysis of 6 basic expressions that there are 36 training subjects with 6 expressions resulting in 216.

Figure 5.13 shows the entropy graph for the training samples used. The entropy values are sorted in descending order. Similar to Section 5.5.1, a break point extraction process is applied which uses the standard deviation. The extracted break points are the first 7, 15, 16, 17 and 21 points. According to the Fisher's criterion, the first 17 high entropy feature points are selected for the first level classification.

Entropy analyses for the second level classes are shown in Figures 5.14 and 5.15. The second level of classification includes two classes; Class 1 and Class 2 where Class 1 includes anger, disgust and fear expressions whereas Class 2 includes happiness, sadness and surprise expressions. Class 1 entropy analyses are completed on anger, disgust and fear expressions in the training set. Class 2 entropy analyses are completed on happiness, sadness and surprise expressions in the training set. The breaking points extracted for the Class 1 features are the first 6, 7, 9, 14 and 20 high entropy feature points. According to Fisher's criterion, the first 20 high entropy feature points are selected.

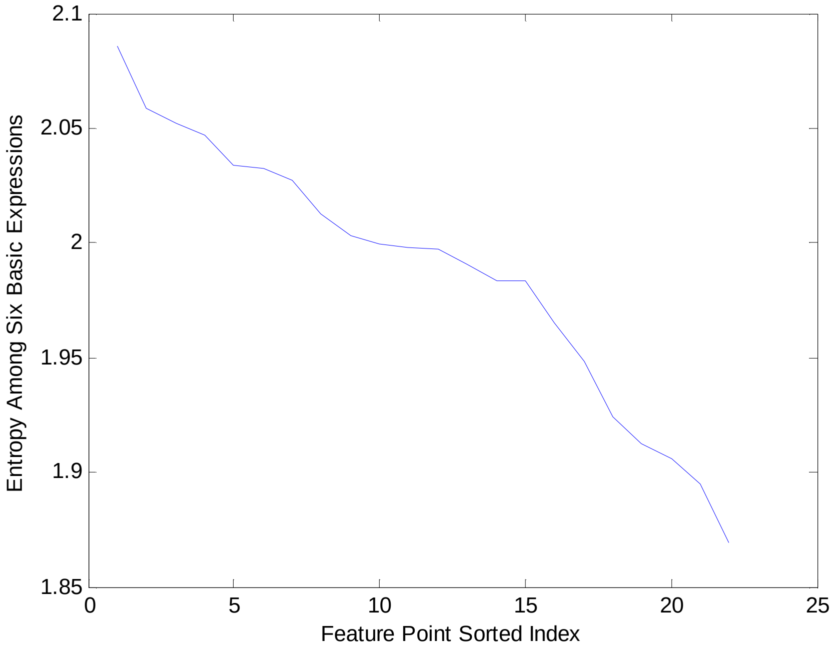


Figure 5.13: Entropy analyses of feature points for 36 training samples among 6 basic expressions.

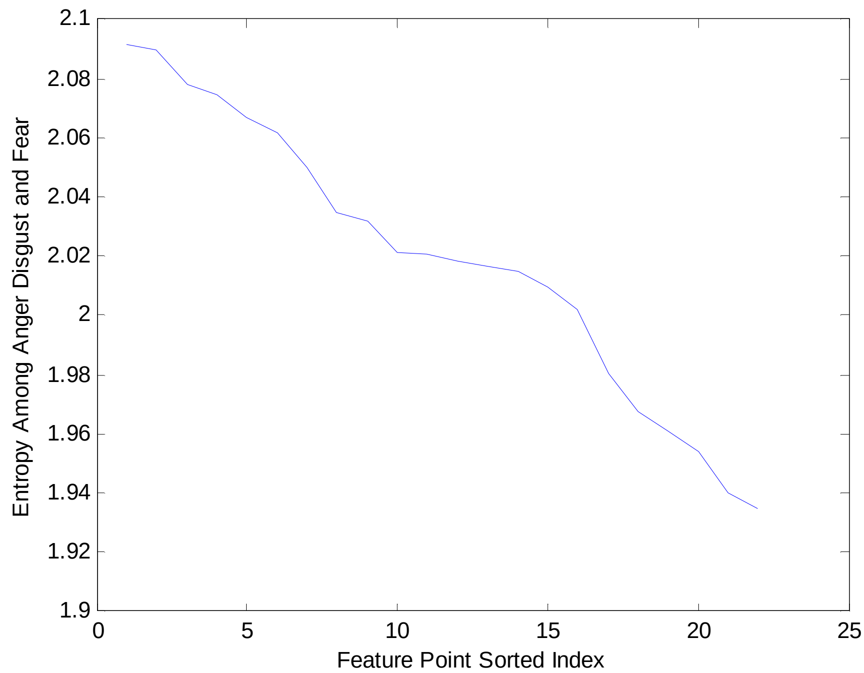


Figure 5.14: Entropy analyses of feature points for 36 training samples among anger, disgust and fear expressions.

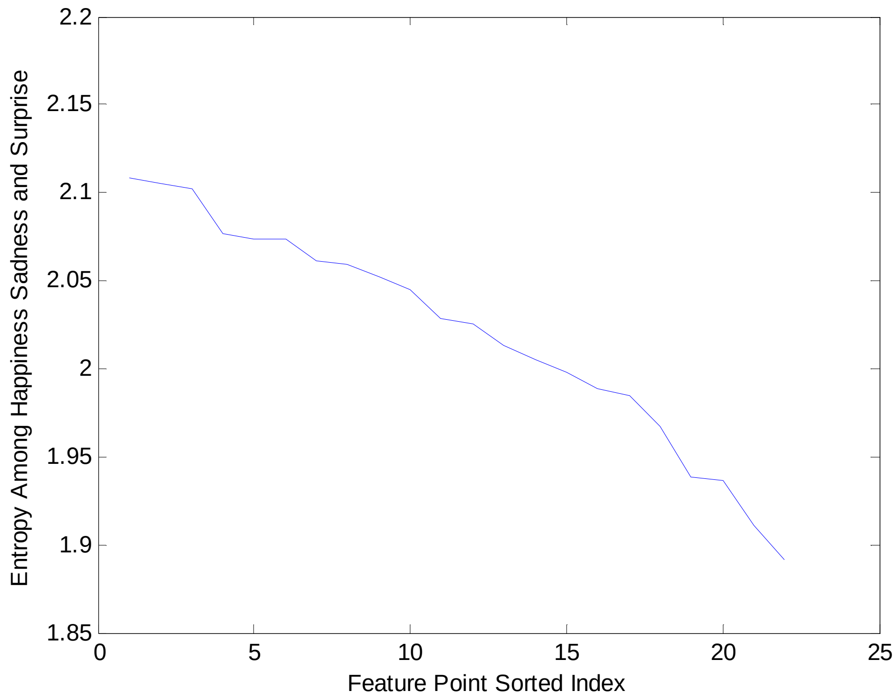


Figure 5.15: Entropy analyses of feature points for 36 training samples among happiness, sadness and surprise expressions.

Bosphorus Database Feature Points Used

1- Outer left eyebrow	12- Nose saddle right
2- Middle left eyebrow	13- Left nose peak
3- Inner left eyebrow	14- Nose tip
4- Inner right eyebrow	15- Right nose peak
5- Middle right eyebrow	16- Left mouth corner
6- Outer right eyebrow	17- Upper lip outer middle
7- Outer left eye corner	18- Right mouth corner
8- Inner left eye corner	19- Upper lip inner middle
9- Inner right eye corner	20- Lower lip inner middle
10- Outer right eye corner	21- Lower lip outer middle
11- Nose saddle left	22- Chin middle

(a)

Sorted Entropy Indices for Level 1:

9 10 6 2 1 8 7 3 5 22 16 18 11 12 4 13 15 20 14 21 19 17

(b)

Selected Features for Level 1:

9 10 6 2 1 8 7 3 5 22 16 18 11 12 4 13 15

(c)

Sorted Entropy Indices for Class 1:

9 12 15 4 10 11 3 8 6 22 1 5 19 21 2 18 13 7 16 20 14 17

(d)

Selected Features for Class 1:

9 12 15 4 10 11 3 8 6 22 1 5 19 21 2 18 13 7 16 20

(e)

Sorted Entropy Indices for Class 2:

8 2 7 6 10 9 1 16 13 3 5 11 12 4 15 22 18 17 19 20 21 14

(f)

Selected Features for Class 2:

8 2 7 6 10 9 1 16 13 3 5 11 12 4 15 22 18 17

(g)

Figure 5.16: Feature points used from Bosphorus database [46]. (a) Feature point descriptions (b) Sorted indices according to entropy in descending order for Level 1 (Figure 5.5) (c) Selected high entropy feature points for Level 1 (d) Sorted indices according to entropy in descending order for Level 2 Class 1 (Figure 5.5) (e) Selected high entropy feature points for Level 2 Class 1 (f) Sorted indices according to entropy in descending order for Level 2 Class 2 (Figure 5.5) (g) Selected high entropy feature points for Level 2 Class 2.

Class 2 entropy analyses outcomes are the breaking points as the first 9, 10, 11, 12, 13, 17 and 18 high entropy feature points. After checking for the Fisher's criterion, the first 18 high entropy feature points are selected.

The high entropy feature points are different for each stage of the selection process. Figure 5.16 shows the feature point descriptions together with the sorted numbers according to their entropy for each class.

5.5.3.1 Performance Analysis

The performance analysis of the feature selection algorithm for coarse-to-fine classifier depicted in Figure 5.5 is performed on Bosphorus database and provided promising results. Experimental setup is similar with the previous methods. 40 samples with 6 different expressions are considered. In total, face matrix *FM* (Equation 5.16) is constructed from 240 row vectors of 40 sample persons with 6 expressions describing 240 faces. Two-class classifiers of each classifier module are trained separately with 90% of the row vectors of *FM* and 10% of row vectors are used in testing. To validate the test results, 10-fold cross validation is applied and the average recognition rates are reported.

Table 5.10: Recognition rates using 22 feature points from Bosphorus database.

Expression	**Recognition Rates (%)**	
	Using 22 Feature Points	Using 14 High Variance Selected Feature Points
Anger	80	67,5
Disgust	65	67,5
Fear	60	65
Happiness	95	95
Sadness	75	70
Surprise	62.5	80
Average Recognition Rates	**72.92**	**74.17**

Table 5.10 shows the recognition rates obtained by using all the 22 feature points provided in Bosphorus database. It is observed from Table 5.10 that the overall recognition rate is moderately high. Individual expressions provide high recognition rates like 95% for the happiness expression and the others are also in the acceptable

range. This result also verifies that the information content of the 3D facial feature points is rich in terms of expression information.

Table 5.11: Recognition Rates after feature selection: First level recognition rates for Class 1 and Class 2 classifications refer to Figure 5.5.

Expression	**Recognition Rates (%)**	
	Class 1	Class 2
Anger-Disgust-Fear	85	15
Happiness-Sadness-Surprise	15.83	84.17
Average Correct Classification	**84.59**	

Table 5.12: Recognition rates after feature selection: Level 2 rates.

Expression	**Recognition Rates (%)**
Anger	95
Disgust	85
Fear	90
Happiness	97.5
Sadness	90
Surprise	90
Average Recognition Rates	**91.25**

Table 5.13: Recognition rates for individual expressions before and after feature selection.

Expression	**Recognition Rates (%)**		
	Before FS	**After FS-Entropy Based**	**After FS-Variance Based**
Anger	80	80.75	67,5
Disgust	65	72.25	67,5
Fear	60	76.5	65
Happiness	95	82.07	95
Sadness	75	75.75	70
Surprise	62.5	75.75	80
Average Recognition Rates	**72.92**	**77.18**	**74.17**

Table 5.11 and 5.12 shows the recognition rates for the Level 1 and 2 of coarse-to-fine classifier. It is observed from these tables that the second level features are very successfully classifying the expressions and improve recognition rates of low

expressions like fear. Although the second level features are very successful, the first level features still have room to improve. The average recognition rate of the first level features is around 85% whereas second level features have around 91% average recognition rate.

Considering the recognition rates before and after the feature selection algorithm, we see that there is a significant increase from 72.92 to 77.18. This result shows that the feature selection algorithm improves the recognition performance database independently. As a result, selecting high entropy feature points will improve the recognition rates while maintaining close rates for each expression.

It is seen from Table 5.13 that the proposed method improves the system performance significantly on a different database as well. This shows that the proposed methods are database independent and improves system performance when 3D geometric points are provided.

5.5.4 Performance Comparison of the Proposed Methods

The proposed methods are compared in Table 5.14. It is observed from Table 5.14 that when using all of the 83 feature points provided in MPEG-4 feature points, the system has an acceptable performance for the classification of six basic expressions. Thus, 3D geometry is a good base for further improvements by applying feature selection. The proposed feature selection procedures, which are based on variance and entropy are compared in Table 5.14. It is obviously seen that entropy based algorithms (columns 4 and 5) perform better than the variance based algorithm. Furthermore, entropy based feature selection methods which are the main contributions of the thesis opens new directions to be followed in feature selections for facial expression recognition.

Although a complete study that analyzes the computational complexity of the proposed methods has not been possessed, there are some experimental observations. According to the experiments, the computation times of the proposed methods can be compared. It is observed that the expression distinctive method is slower than the coarse-to-fine method significantly. This is the result of the number of 2-class SVM classifiers. The testing procedures including more 2-class SVM classifiers require more computation time obviously. Thus, the variance based method has the least number of 2-class SVM classifiers and is expected to be the fastest one. The coarse-to-fine method is slower than the variance based but faster than the expression distinctive method. Expression distinctive method includes all the 2-class classifiers dedicated to each individual expression that makes it slower but more accurate.

Table 5.14: Performance Comparison of the Proposed Methods.

	Recognition Rates (%)			
	Using 83 Feature Points	**Variance Based FS**	**2-Level Coarse-to-Fine FS**	**Expression Distinctive FS**
Database	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**	**BU-3DFE**
Training / Test	**90% / 10%**	**90% / 10%**	**90% / 10%**	**90% / 10%**
Anger	93.33	100	90.19	90.0
Disgust	86.67	86.7	85.17	80.0
Fear	60.00	60	76.55	80.0
Happiness	93.33	93.3	93.84	95.0
Sadness	80.00	80	88.89	90.0
Surprise	86.67	93.3	95.06	95.0
Average	**83.33**	**85.55**	**88.28**	**88.33**

Chapter 6

CONCLUSION

6.1 Conclusion

The thesis proposes novel feature selection procedures for the recognition of six basic facial expressions by utilizing 3-Dimensional (3D) geometrical facial feature positions. The thesis presents systems for classifying expressions into one of the six basic emotional categories which are anger, disgust, fear, happiness, sadness and surprise. Two classification models are presented that are based on SVM. Classification methods and feature selection procedures are both person independent. The thesis contributes on feature selection procedures in two different models. The first model is the two level coarse-to-fine classification model where expressions are classified in two levels. In the first level, the expression is classified as Class 1 or Class 2 where Class 1 includes anger, disgust and fear, and Class 2 includes happiness, sadness and surprise expressions. Then in the second level, expressions are classified in each class. The novel feature selection algorithm proposes three feature models for the first and the second levels of this coarse-to-fine classification model. Secondly, the novel feature selection procedure is applied expression distinctively in order to find six different feature models specific to each individual expression. Therefore, the classification model is also adapted as accept-reject classifier. The outcomes of the feature selection procedure are six different facial feature models which are the most discriminative features among expressions.. Feature selection for each expression is completed independently and achieves high recognition rates with the proposed geometric facial features selected.

The system performances are evaluated on the 3D facial expression databases which are BU-3DFE [45] and Bosphorus [46] facial expression databases. The experimental results on classification performance show that the SVM classifier models with the selected features achieve high overall recognition rates while keeping individual recognition rates closely high. Comparison of the proposed methods with the recent methods reported on the same databases show that the proposed method outperforms other person independent methods in various aspects. The performance results show that the proposed methods obtain encouraging recognition rates and they are open for further improvements.

The proposed system provides better aspects according to the conventional methods in the literature. The system is using limited information which is the face geometry and high recognition rates have been achieved. Although the performance of the proposed system is high, by enriching the information used with the appearance features, the proposed system is ready for further improvements. On the other hand, the proposed system overcomes a common problem of many facial expression recognition systems currently available, which is the balance between the recognition rates of individual expressions. The proposed system reports closely high rates for all the expressions. Another better aspect of the system is that the feature selection process decreases the number of feature points needed to extract, so in a real-time application, the feature extraction process will require less computations.

6.2 Future Directions

Although the recognition of the six basic expressions is enhanced by the proposed novel feature selection methods, there is still room for further improvements. The thesis opens new directions with the proposed feature selection method.

The first direction is to apply the feature selection procedures to a more robust facial representation which is the use of 3D distances between the geometrical facial feature points instead of their positions. This representation is evaluated before in [35] and achieved encouraging results.

The proposed methods employ entropy as the metric of information content. It is used to search for optimal features which maximize Fisher criterion. The second direction of future work is to use entropy as a criterion with a search algorithm like genetic algorithm. Entropy as a criterion should be applied in order to minimize within class entropy and maximize between class entropy. Selection of such features can be achieved by use of evolutionary multi-objective optimization with the possible optimization algorithms like Non Dominated Sort II (NSGA-II) or Particle Swarm Optimization (PSO).

Facial expressions are spontaneous activities on the face. Thus, another direction is to apply the current methods to spontaneous facial expression databases. This spontaneous behavior of the facial expressions in real-time is included in the BU-4DFE version of the updated database. One of the future works is to apply the algorithm to spontaneous behavior of expressions.

The proposed feature selection methods are able to select the most informative features among the expressions. This enables us to study the transitions between the expressions. In real-time, a series of expressions can be observed on the human face. The transitions from one expression to other are also a future direction of research that the proposed methods can be adapted. We believe that the selection of the most discriminative features during the transitions of expressions will improve the performance and decrease the computational complexity during feature extraction process in real-time. Considering transitions among the expressions, instead of 3D geometric data of feature points, the directions of the feature points can also be employed. This will bring a new face representation using feature directions.

Another future direction is to apply the proposed feature selection methods to appearance based features like Local Binary Patterns (LBP). As it is explained in Chapter 3, facial expressions are 3D activities on the face, thus, a face representation should reflect these 3D activities well. Therefore, 3D texture models like 3D LBP can be used in feature selection procedures. Furthermore, the fusion of selected appearance based features and geometric features may represent faces and discriminate facial expressions better.

LIST OF PUBLICATIONS

- Kamil Yurtkan and Hasan Demirel, "Person Independent Facial Expression Recognition Using 3D Facial Feature Positions", 27th International Symposium on Computer and Information Sciences, ISCIS 2012, 3-5 October 2012, Institut Henri Poincare, Paris, France.

- Kamil Yurtkan, Hamit Soyel and Hasan Demirel, "Feature Selection for Enhanced 3D Facial Expression Recognition Based on Varying Feature Point Distances", 28th International Symposium on Computer and Information Sciences, ISCIS 2013, 28-29 October 2013, Institut Henri Poincare, Paris, France.

- Kamil Yurtkan and Hasan Demirel, "Entropy-Based Feature Selection for Improved 3D Facial Expression Recognition", Signal Image and Video Image Processing, Springer, Volume 8, Number 2, February 2014, pp. 267-277, DOI: 10.1007/s11760-013-0543-1, ISSN 1863-1711.

- Kamil Yurtkan and Hasan Demirel, "Feature Selection for Improved 3D Facial Expression Recognition", Pattern Recognition Letters, Elsevier, Volume 38, March 2014, pp. 26-33 DOI: 10.1016/j.patrec.2013.10.026, ISSN 0167-8655.

REFERENCES

[1] Ahlberg, J., “Candide-3: an updated parameterized face”, *Report No. LiTH-ISY-R-2326*, (Linkoping University, Sweden), 2001.

[2] P. Viola and M. Jones, “Robust real-time face detection”, *International Journal of Computer Vision*, vol. 57, pp. 137–154, 2004.

[3] Ekman, P. & Friesen, W., “Pictures of Facial Affect. Palo Alto”, *CA: Consulting, Psychologist*, 1976.

[4] Ekman, P. & Friesen, W., “The Facial Action Coding System: A Technique for the Measurement of Facial Movement”, *Consulting Psychologists Press*, San Francisco, 1978.

[5] G.Abrantes, F.Pereira, “MPEG-4 facial animation technology: survey, implementation and results”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.9, no. 2, pp. 290 – 305, March 1999.

[6] Seyed Mehdi Lajevardi and Zahir M. Hussain, “Automatic facial expression recognition: feature extraction and selection”, *Signal, Image and Video Processing*, Volume 6, Issue 1, pp 159-169, March 2012.

[7] Kamil Yurtkan and Hasan Demirel, “Entropy-Based Feature Selection for Improved 3D Facial Expression Recognition”, *Signal Image and Video Image Processing*, Springer, Volume 8, Issue 2, February 2014, pp. 267-277, 2014.

[8] V. Bettadapura, “Face Expression Recognition and Analysis: The State of the Art”, *Tech Report, arXiv:1203.6722*, April 2012.

[9] G. Rajamanoharan, S. Zafeiriou, M. Pantic, L. Yin., “Static and Dynamic 3D Facial Expression Recognition: A Comprehensive Survey”, *Image and Vision Computing*, 30(10): pp. 683 - 697, 3D Facial Behaviour Analysis and Understanding. 2012.

[10] Caifeng Shan, Shaogang Gong, and Peter W. McOwan, “Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study”, *Image and Vision Computing*, Vol. 27, Issue 6, pp. 803-816, 2009.

[11] Kamil Yurtkan and Hasan Demirel, “Person Independent Facial Expression Recognition Using 3D Facial Feature Positions”, *27th International Symposium on Computer and Information Sciences, ISCIS 2012*, Institut Henri Poincare, Paris, France, 3-5 October 2012.

[12] Kamil Yurtkan, Hamit Soyel and Hasan Demirel, “Feature Selection for Enhanced 3D Facial Expression Recognition Based on Varying Feature Point Distances”, *28th International Symposium on Computer and Information Sciences, ISCIS 2013*, Institut Henri Poincare, Paris, France, 28-29 October 2013.

[13] Kamil Yurtkan and Hasan Demirel, “Feature Selection for Improved 3D Facial Expression Recognition”, *Pattern Recognition Letters*, Elsevier, Volume 38, 1st March 2014, pp. 26-33, 2014.

[14] J. Montagu, “The Expression of the Passions: The Origin and Influence of Charles Le Brun’s ‘Conférence sur l'expression générale et particulière,”, *Yale University Press,* 1994.

[15] C. Darwin, edited by F. Darwin, “The Expression of the Emotions in Man and Animals”, *2nd edition, J. Murray*, London, 1904.

[16] Kamil Yurtkan, Hamit Soyel, Hasan Demirel, Hüseyin Özkaramanlı, Mustafa Uyguroğlu, and Ekrem Varoğlu, “Face Modeling and Adaptive Texture Mapping for Model Based Video Coding”, *Lecture Notes on Computer Science (LNCS) 3691*, pp. 498 – 505, 2005.

[17] Kamil Yurtkan and Hasan Demirel, “Facial Expression Synthesis from Single Frontal Face Image”, *6th Int. Symp. on Electrical and Computer Systems*, Eueropean University of Lefke, Gemikonağı, TRNC, 25-26 Nov. 2010.

[18] Kamil Yurtkan, Turgay Celik and Hasan Demirel, “Automatic Facial Synthesis from Single Frontal Face Image for Model Based Video Coding”, *The 13th IASTED International Conference on Computer Graphics and Imaging, CGIM 2010*, 15-17, Innsburck, Austria, February 2010.

[19] Yongzhong Lu, Jingli Zhou, Shengsheng Yu, “A Survey of Face Detection, Extraction and Recognition”, *Computers and Artificial Intelligence* 22(2): 163-195, 2003.

[20] V. Blanz, T. Vetter, “A morphable model for the synthesis of 3D faces”, *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 187–194, 1999.

[21] V. Blanz, T. Vetter, “Face recognition based on fitting a 3D morphable model”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1063–1074, 2003.

[22] J. Wang, L. Yin, X. Wei, Y. Sun, “3D facial expression recognition based on primitive surface feature distribution”, *Computer Vision Pattern Recognition*, *CVPR 2006*, pp. 1399-1406, 2006.

[23] Lyons, M. Budynek, J. & Akamatsu, S., “Automatic Classification of Single Facial Images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1357–1362, 1999.

[24] Jiangang Yu and Bir Bhanu, “Evolutionary Feature Synthesis for Facial Expression Recognition”, *Pattern Recognition Letters*, Volume 27, Issue 11, pp. 1289-1298, August 2006.

[25] Ojala, T., Pietikäinen, M., Harwood, D., “A comparative study of texture measures with classification based on feature distributions”, *Pattern Recognition*, 29(1), 51–59 (1996).

[26] Caifeng Shan, Shaogang Gong, and Peter W. McOwan, “Robust Facial Expression Recognition Using Local Binary Patterns”, *Proceedings of IEEE International Conference on Image Processing (ICIP'05)*, Genova, Italy, Sept. 2005.

[27] Rawaz Ameen, Hüseyin Öztoprak and Kamil Yurtkan, “Analysis of Local Binary Patterns in Different Color Models for Person Independent Facial Expression Recognition”, *IEEE 22. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SİU 2014,* Karadeniz Technical University, Trabzon, Turkey, 23-25 April 2014.

[28] Wikipedia, http://en.wikipedia.org/wiki/Facial_Action_Coding_System, June 2014.

[29] Narendra Patel and Mukesh Zaveri, “3D facial model reconstruction, expressions synthesis and animation using single frontal face image”, *Signal, Image and Video Processing,* November 2011.

[30] G. Sandbach, S. Zafeiriou, M. Pantic, D. Rueckert, “A dynamic approach to the recognition of 3D facial expressions and their temporal models”, *Automatic Face Gesture Recognition and Workshops (FG 2011),* 2011 IEEE International Conference on, pp. 406–413, 2011.

[31] G. Sandbach, S. Zafeiriou, M. Pantic, D. Rueckert, “Recognition of 3D facial expression dynamics”, *Image Vision Computing*, 30 (10): pp.762-773, *3D Facial Behaviour Analysis and Understanding*, 2012.

[32] H. Tang, T. Huang, “3D facial expression recognition based on automatically selected features”, *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08*, IEEE Computer Society Conference on, pp. 1–8, 2008.

[33] H. Soyel, H. Demirel, “Optimal feature selection for 3D facial expression recognition with geometrically localized facial features”, *Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, 2009. ICSCCW 2009. Fifth International Conference on, pp. 1–4, 2009.

[34] H. Soyel, H. Demirel, “Optimal feature selection for 3D facial expression recognition using coarse-to-fine classification”, *Turkish Journal of Electrical Engineering and Computer Science*, 18 (6), 1031–1040, 2010.

[35] Hamit Soyel, Umut Tekguc, Hasan Demirel, “Application of NSGA-II to Feature Selection for Facial Expression Recognition”, *Computers and Electrical Engineering*, Elsevier, vol. 37, no. 6, pp. 1232-1240, Nov. 2011.

[36] Hamit Soyel and Hasan Demirel, “Localized discriminative scale invariant feature transform based facial expression recognition”, *Computers and Electrical Engineering*, October 2011.

[37] T. Sha, M. Song, et al., “Feature Level Analysis for 3D Facial Expression Recognition”, *Neurocomputing*, 74: 2135-2141, 2011.

[38] N. Sebe, M.S. Lew, Y. Sun, I. Cohen, T. Gevers, and T.S. Huang, “Authentic Facial Expression Analysis”, *Image and Vision Computing*, vol. 25, pp. 1856-1863, 2007.

[39] M. Pardas and A. Bonafonte, “Facial animation parameters extraction and expression recognition using Hidden Markov Models”, *Signal Processing: Image Communication*, vol. 17, pp.675–688, 2002.

[40] Y. Tian, T. Kanade and J. Cohn, “Recognizing Action Units for Facial Expression Analysis”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97–115, 2001.

[41] I. Kotsia, I. Buciu and I. Pitas, “An Analysis of Facial Expression Recognition under Partial Facial Image Occlusion”, *Image and Vision Computing*, vol. 26, no. 7, pp. 1052-1067, July 2008.

[42] H. Soyel, H. Demirel, “3D facial expression recognition with geometrically localized facial features”, *Computer and Information Sciences, 2008. ISCIS'08*, 23rd International Symposium on, pp. 1–4, 2008.

[43] Y. Chang, M. Vieira, M. Turk, L. Velho, “Automatic 3D facial expression analysis in videos”, *AMFG'05 Proceedings of the Second International Conference on Analysis and Modelling of Faces and Gestures*, 293–307, 2005.

[44] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I., “The Extended Cohn-Kanade Dataset (CK+): A complete facial expression dataset for action unit and emotion-specified expression.”, *3rd IEEE Workshop on Computer Vision and Pattern Recognition, CVPR, for Human Communicative Behaviour Analysis*, 2010.

[45] Yin, L. Wei, X. Sun, Y. Wang, J. and Rosato, M., "A 3D Facial Expression Database for Facial Behavior Research", *7th International Conference on Face and Gesture Recognition*, pp. 211-216, United Kingdom, 2006.

[46] Arman Savran, Nese Alyuz, Hamdi Dibeklioğlu, Oya Celiktutan, Berk Gokberk, Bulent Sankur, Lale Akarun, "Bosphorus Database for 3D Face Analysis", *The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008),* Roskilde, Denmark, 7-9 May 2008.

[47] Face Recognition Grand Challenge, http://www.nist.gov/itl/iad/ig/frgc.cfm, June 2014.

[48] Seyed Mehdi Lajevardi and Zahir M. Hussain, "Automatic facial expression recognition: feature extraction and selection", *Signal, Image and Video Processing*, Volume 6, Issue 1, pp 159-169, March 2012.

[49] I. Buciu, C. Kotropoulos and I. Pitas, "Comparison of ICA approaches for facial expression recognition", *Signal, Image and Video Processing*, Volume 3, Issue 4, pp 345-361, December 2009.

[50] Y. Tian, T. Kanade and J. Cohn, "Recognizing Action Units for Facial Expression Analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97–115, 2001.

[51] F. Bourel, C.C. Chibelushi and A. A. Low, “Recognition of Facial Expressions in the Presence of Occlusion”, *Proceedings of the 12th British Machine Vision Conference*, vol. 1, pp. 213–222, 2001.

[52] I. Cohen, N. Sebe, A. Garg, L.S. Chen, and T.S. Huang, “Facial Expression Recognition From Video Sequences: Temporal and Static Modelling”, *Computer Vision and Image Understanding*, vol. 91, pp. 160-187, 2003.

[53] M. Pantic and I. Patras, “Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form Face Profile Image Sequences”, *IEEE Transactions on Systems, Man and Cybernetics Part B*, vol. 36, no. 2, pp.433-449, 2006.

[54] J. Wang and L. Yin, “Static Topographic Modeling for Facial Expression Recognition and Analysis”, *Computer Vision and Image Understanding*, vol.108, pp. 19–34, 2007.

[55] I. Kotsia, I. Buciu and I. Pitas, “An Analysis of Facial Expression Recognition under Partial Facial Image Occlusion”, *Image and Vision Computing*, vol. 26, no. 7, pp. 1052-1067, July 2008.

[56] Lyons, M. Budynek, J. & Akamatsu, S., “Automatic classification of single facial images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21, pp. 1357–1362, 1999.

[57] Tang H, Huang TS., “3D Facial expression recognition based on properties of line segments connecting facial feature points”, *IEEE International Conference on Automatic Face and Gesture Recognition*, 2008.

[58] I. Mpiperis, S. Malassiotis, M.G. Strintzis, “Bilinear models for 3D face and facial expression recognition”, *IEEE Transactions on Information Forensics Security*, pp. 498–511, March 2008.

[59] Hjortsjö, CH (1969), “Man's face and mimic language”, *1969, English, Book edition*, 111 p., Lund, Sweden : Studentlitteratur, c1969.

[60] Ebrahimi and E. Soofi, in D.J. Slottje (ed.), “Measuring Informativeness by Entropy and Variance”, *N. Advances in Econometrics, Income Distribution, and Methodology of Science (Essays in Honor of Camilo Dagum),* Springer-Verlag, 1999.

[61] S.Maji, A.C.Berg, J.Malik, “Classification using Intersection Kernel Support Vector Machines is Efficient”, *IEEE Computer Vision and Pattern Recognition Conference, CVPR 2008,* 2008.

[62] Wikipedia, http://en.wikipedia.org/wiki/Support_vector_machine, June 2014.

[63] Frenken, K., “Entropy statistics and information theory”, in H. Hanusch and A.Pyka (eds.), *The Elgar Companion to Neo - Schumpeterian Economics* (Cheltenham, UK and Northampton MA: Edward Elgar), pp. 544-555, 2007.

[64] Binghamton University BU-3DFE, http://www.cs.binghamton.edu/

~lijun/Research/3DFE/3DFE_Analysis.html, June 2014.

[65] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification", *2nd Edition*, *ISBN: 978-0-471-05669-0*, 2000.

[66] Nock, R. and Nielsen, F., "On Weighting Clustering", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (8), 1–13, 2006.

[67]MIT Courses Media, http://courses.media.mit.edu/2006fall/mas622j/

Projects/aisen-project/, June 2014.

[68] Data Clustering Algorithms, https://sites.google.com/site/

dataclusteringalgorithms/fuzzy-c-means-clustering-algorithm, September 2014.