Khan, Masood Mehmood

Cluster-analytic classification of facial expressions using infrared measurements of facial thermal features

Cluster-analytic classification of facial expressions using infrared measurements of facial thermal features

Masood Mehmood Khan

A thesis presented to
The University of Huddersfield
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

03 March 2008

School of Computing and Engineering
University of Huddersfield

Dedication

To

my **father** Dr. Hafiz Baber Khan,

my **mother** Meher Al-Nissa,

my **wife** Najia Masood,

my **son** Rafay Ahmed Khan, and

my **daughters** Sheema Masood Khan and Samha Ebadat Khan

for supporting my intellectual endeavors

and to

my sisters, brothers and friends for their encouragement and support.

Acknowledgement

I came across a paper that described how affective states could be automatically recognised with the help of human physiological signals, sometimes in the first quarter of 2003. Several discussions on this subject with my thesis supervisor Dr. Robert D. Ward instigated interest in automated recognition of facial expressions. That is how I ended up working on this thesis. I think Dr. Ward desrves all the credit for this work being carried out. I could never complete this thesis without his inspiration, encouragement and support.

I am highly grateful to Dr. Michael Ingleby for his valuable reviews, comments and support. I also owe many thanks to Dr. Janet Finlay for accepting to serve on my doctoral supervisory committee. It was an honour to have Dr. Ingleby and Dr. Finlay on the supervisory committee and work with them.

I must thank the students of the American University of Sharjah, United Arab Emirates, for allowing me to photograph and capture their visible and infrared images. Capturing images and developing a database of visible-spectrum and infrared images was an important and challenging part of this work. Many thanks to my cheerful student volunteers who turned this difficult task into a joyous experience.

Abstract

In previous research, scientists were able to use transient facial thermal features extracted from **T**hermal **I**nfra-**R**ed **I**mages (TIRIs) for making binary distinction between the affective states. For example, thermal asymmetries localised in facial TIRIs have been used to distinguish anxiety and deceit. Since affective human-computer interaction would require machines to distinguish between the subtle facial expressions of affective states, computers' able to make such binary distinctions would not suffice a robust human-computer interaction. This work, for the first time, uses affective-state-specific transient facial thermal features extracted from TIRIs to recognise a much wider range of facial expressions under a much wider range of conditions. Using infrared thermal imaging within the 8-14 μm, a database of 324 discrete, time-sequential, visible-spectrum and thermal facial images was acquired, representing different facial expressions from 23 participants in different situations. A facial thermal feature extraction and pattern classification approach was developed, refined and tested on various Gaussian mixture models constructed using the image database. Attempts were made to classify: neutral and pretended happy and sad faces; multiple positive and negative facial expressions; six (pretended) basic facial expressions; partially covered or occluded faces; and faces with evoked happiness, sadness, disgust and anger.

The cluster-analytic classification in this work began by segmentation and detection of thermal faces in the acquired TIRIs. The affective-state-specific temperature distributions on the facial skin surface were realised through the pixel grey-level analysis. Examining the affective-state-specific temperature variations within the selected regions of interest in the TIRIs led to the discovery of some significant **F**acial **T**hermal **F**eature **P**oints (FTFPs) along the major facial muscles. Following a multivariate analysis of the **T**hermal **I**ntensity **v**alues (TIVs) measured at the FTFPs, the TIRIs were represented along the **P**rincipal **C**omponents (PCs) of a covariance matrix. The resulting PCs were ranked in the order of their effectiveness in the between-cluster separation. Only the most effective PCs were retained to construct an optimised eigenspace. A supervised learning algorithm was invoked for linear subdivision of the optimised eigenspace. The statistical significance levels of the classification results were estimated for validating the discriminant functions.

The main contribution of this research has been to show that: the infrared imaging of facial thermal features within the 8-14 μm bandwidth may be used to observe affective-state-specific thermal variations on the face; the pixel-grey level analysis of TIRIs can help localise FTFPs along the major facial muscles of the face; cluster-analytic classification of transient thermal features may help distinguish between the facial expressions of affective states in an optimized eigenspace of input thermal feature vectors. The Gaussian mixture model with one cluster per affect worked better for some facial expressions than others. This made the influence of the Gaussian mixture model structure on the accuracy of the classification results obvious. However, the linear discrimination and confusion patterns observed in this work were consistent with the ones reported in several earlier studies.

This investigation also unveiled some important dimensions of the future research on use of facial thermal features in affective human-computer interaction.

Contents

List of Figures

List of Tables

Abbreviations

AAR	Automated Affect Recognition
AFEC	Automated Facial Expression Classification
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BOLD	Blood Oxygen Level Dependant Data
CD	Compact Disk
CHI	Computer Human Interaction
DRAM	Dynamic Random Access Memory
ECG	Electroencephalogram
ECG	Electrocardiogram
EMG	Electromyogram
EOG	Electroocculogram
FACS	Facial Action Coding Systems
fMRI	Functional Magnetic Resonance Imaging
FTFP	Facial Thermal Feature Point
FTFPs	Facial Thermal Feature Points
GSR	Galvanic Skin Response
HCI	Human Computer Interaction
HDD	Hard Disk Drive
HMM	Hidden Markov Model

IR	Infrared (Infra Red)
KNN	K-nearest Neighbor Network
LDA	Linear Discriminant Analysis
LR	Logistic Regression
MANOVA	Multivariate Analysis of Variance
MAX	Maximally Discriminative Affect Coding System
MB	Mega Bytes
NVAFEC	Non-vision based Automated Facial Expression Classification
PC	Principal Component
PCA	Principal Component Analysis
PCMCIA	Personal Computer Memory Card International Association
PIXEL	Picture [PIX] Element [EL]
RPNS	Respiratory Patterns
SV	Singular Value
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVs	Singular Values
TI	Thermal Imaging
TIRI	Thermal Infrared Imaging
TIRIs	Thermal Infrared Images
TIV	Thermal Intensity Value
TIVs	Thermal Intensity Values

Chapter 1

INTRODUCTION

1.1 Motivation

Scientific studies confirm that people use a variety of auditory and visual cues such as voice levels, gait information, gestures and facial expressions to understand others' emotions (Bartneck 2001; Brooks 2002; Picard 2000). One important source of visual information is facial expression (Du and Lin 2003). Using these auditory and visual cues appears to be a casual, simple, and effortless task for humans (Redford 2000). What, however, appears to be easy and simple tasks for humans translate into a set of complex computational activities for computers. Despite growing processing power and multiplicity of input-output modalities, computers possess limited abilities to recognise, understand and interpret emotions (Bartneck 2001).

Nevertheless, the potential benefits of computers' able to express and respond to emotions inspired researchers to design and implement socially intelligent systems. This inspiration is evident in the recent scholarly works on automated recognition, interpretation and expression of emotion (Bartneck 2001; Brooks 2002; Busso et al. 2004; Klein et al. 2002). Many recent systems have demonstrated some limited capabilities of recognising, interpreting and expressing emotions (Cohen et al. 2003; Essa and Pentland 1997; Gao et al. 2003; Morishima 2001). A wide range of potential applications of these so called socially intelligent computers have been reported in the literature on human-computer interaction (HCI), robotics, bio-informatics, security and surveillance, and psychotherapy (Cohen et al. 2003; Eveland et al. 2003; Socolinsky et al. 2003; Klein et al. 2002; Lisetti and Schiano 2000; Picard 2000; Reeves and Nass 1996).

A large number of the existing models of socially intelligent systems, as reported in the literature, relies on the visual cues to recognise and classify the facial expressions and interpret emotions (Abidi et al. 2004). Sophisticated algorithms, hardware accessories and tools for implementing the vision-based Automated Facial Expression

Classification (AFEC) systems are being developed, tested and made available. In particular, efforts are being made to enhance the performance of the vision-based AFEC capable systems. Some of the recently developed vision-based AFEC systems claimed over 70 % accuracy in recognising the facial expressions of affective states (Baldwin et al. 1998; Cohen et al. 2003; Gao et al. 2003). Since the vision-based AFEC systems have been around for several years now, their strengths and limitations are well understood. The underlying theories and implementation details of salient vision-based AFEC systems are discussed in (Black and Yacoob 1997; Ekman et al. 1993; Fasel and Luettin 2003; Huang and Huang 1999; Pantic and Rothkrantz 2000).

Despite their claimed success in controlled environments, critics find the vision-based AFEC systems less effective outside the research laboratories (Baldwin et al. 1998; Ekman et al. 1993; Sugimoto et al. 2000). A number of technical limitations are believed to deter the performance of the vision-based AFEC systems in life like situations. Factors such as the deformability and the transient nature of the facial features, and the influence of the ambient light intensity while a face is being observed pose problems in facial feature extraction. Also, dependence of the feature extraction process on a physically-based structural model of the face in the vision-based AFEC systems is considered problematic. The works by Fried (1976) and Friedman (1970) on development of the physically-based structural models of human face provided basis for modeling the face, defining the anatomical components of the face, and representing the interaction between the anatomical components of a face. Some recent scientific studies have raised questions about the theoretical foundations of the works by Fried (1976) and Friedman (1970) and the suitability of using the physically-based structural models in AFEC and AAR systems (Ekman et al. 1993; Morishima 2001). The factors that supposedly deter the performance of the vision-based AFEC and AAR systems are discussed in (Baldwin et al. 1998; Ekman et al. 1993; Fasel and Luettin 2003; Pantic and Rothkrantz 2000; Sugimoto et al. 2000).

Such limitation of the vision-based AFEC and AAR capable systems inspired researchers to explore the possibilities of using non-visual cues for AFEC and AAR. Recent works in the areas of computational intelligence, psychology, physiology, neuropsychology, pattern recognition, machine learning and HCI demonstrate a growing interest in the use of non-visual signals for designing AFEC and AAR capable systems (Abidi et al. 2004; Ang et al. 2004; Mase 1991; Picard 2000).

A number of human bio-physiological signals are considered useful in providing emotion-specific human information. One or more of human bio-physiological cues have been employed in some of the non-vision based AFEC (NVAFEC) and AAR systems. A direct contact with the human body is needed to acquire most of the human bio-physiological signals. Hence, the NVAFEC and AAR capable systems that rely on bio- physiological cues remain intrusive (Prokoski and Iedel 1999). The underlying theories and implementation of non-vision based AFEC and AAR systems were discussed in (Boulic and Thalmann 1998; Christie and Friedman 2004; Critchley et al. 2000; Jones et al. 1988; Kakadiaris 2005a; Kakadiaris 2005b; Naemura et al. 1993; Niemic 2002; Picard 2000; Pollina et al. 2006; Posamentier and Abdi 2003; Prokoski and Iedel 1999; Puri et al. 2005; Schwarz et al. 2002; Yoshitomi et al. 2000).

The major operational difference between the vision-based AFEC systems and the existing NVAFEC systems is that the former systems can perform in a non-invasive and non-contact manner whereas the later systems are primarily intrusive. The intrusive nature of NVAFEC and AAR systems remains a major obstacle in their acceptability and application. However, recent advances in digital thermal infrared imaging have made it possible to acquire a very useful human bio-physiological signal, the body temperature, through non-intrusive and non-contact means (Phillips 2002). Human skin temperature, considered a function of thermo-muscular, hæmodynamic and metabolic factors (Bales 1989), can be measured through the thermal infrared imaging in a non-contact, non-invasive and illumination invariant manner (Jones and Plassmann 2002; Otsuka et al 2002).

Earlier researchers were able to use the facial hæmodynamic variations and thermal features to classify affects and their expressions (Pollina et al. 2006; Puri et al. 2005; Yoshitomi et al. 2000). Some recent studies have demonstrated that pixel grey-levels in the thermal infrared images provide a reliable measure of skin surface radiance and allow measuring the skin temperature distribution patterns (Bales 1998; Jones and Plassmann 2002; Otsuka et al. 2002). Typically, standard image processing methods are invoked to enhance the thermal images and extract the pixel grey-level information for recognising the facial expressions of affective states (Bales 1998; Jones and Plassmann

2002; Otsuka et al. 2002; Pollina et al. 2006; Puri et al. 2005; Sugimoto et al. 2000; Yoshitomi et al. 2000). Investigators were able to recognise the stress levels, deceit and facial expressions of positive and negative affective states using the pixel grey-level information extracted from the thermal images, notably in a dichotomous discrimination manner (Pavlidis 2004; Pavlidis and Levine 2002; Pollina et al. 2006; Puri et al. 2005). Some recent studies confirm that infrared measurement of facial skin temperature can lead to non-invasive, non-contact recognition of common expressions of affective states (Dimberg 1990a; Dimberg 1990b; Khan et al. 2004; Khan et al. 2005; Khan et al. 2006; Pavlidis 2004; Pollina et al. 2006; Puri et al. 2005).

Motivated by the success of previous investigations, this work explores the possibilities of recognising the facial expressions of affective states with the help of facial skin temperature measurements. However, the scope of this work is much broader than that of the previous investigations in that it attempts to recognise the pretended and involuntarily evoked expressions of most common affective states using the temporal facial thermal features. In effect, this work aims to distinguish between the facial thermal features for recognising the expressions of most common affective states. This broader framework of the thesis would require development of an effective temporal facial thermal feature extraction mechanism and design of an effective facial expression classifier.

1.2 Research premises

This work is based on the scientific theories suggesting that (a) the human body metabolism changes with a change in emotive state resulting in emotion-specific bio-physiological variations in the human body; (b) a change in affective state would result in the blood volume flow variations under the facial skin; (c) the facial expression of emotion also causes some musculo-thermal activities under the facial skin; and (d) any change in blood volume flow and associated musculo-thermal changes cause variations in the facial skin temperature.

Based on the scientific evidence available in (Bales 1998; Jones and Plassmann 2002; Pavlidis 2004; Pollina et al. 2006; Puri et al. 2005; Sugimoto et al. 2000; Yoshitomi et al. 2000), this thesis examines if an appropriate analysis of the pixel grey-levels within the time sequential thermal images would allow extracting the facial thermal features to classify the facial expressions of most common affective states. Unlike previous investigations, this work does not analyse the observed thermal symmetries or asymmetries in the infrared images. Instead, this thesis proposes that temperature measurements in the regions of interest within the time-sequential infrared images can be subtracted to discover the affective state-specific variations in the facial thermal features. The thesis further proposes that some principal directions in the affective state-specific facial thermal variations can be discovered. Hence, the thermal images can be represented as uncorrelated vectors along the principal components of a covariance matrix. The thesis takes the position that an appropriate supervised learning method such as discriminant analysis can be used for direct estimation of the posterior probabilities. Thus effective discriminant functions can be generated to allocate an unknown thermal face to a particular cluster of facial expression.

A cluster-analytic approach supported by the statistical classification schema was preferred for classifying the facial thermal features in this work. The employed statistical classification approach would allow representing the facial thermal features for developing a set of optimal discriminant functions by regression. The employed statistical pattern recognition approach would also result in an implicit estimation of the class densities. Hence, it was possible to confidently estimate the *posterior* probabilities of class membership and develop a person-independent classifier. Some popular competing classification approaches such as the syntactical classifiers and neural networks were also considered for this work but were found less relevant to the scope and objectives of this work.

Syntactical classifiers would employ some primitives for feature representation and use a set of grammar rules for developing the classification functions. Thus they impose a rigid a feature representation schema. Such a representation might add complexities in developing a compact and optimised decision space needed for separating a set of complex and overlapped classes.

Neural networks learn complex and non-linear input-output relationships using well-connected sequential training procedures and adapt to the training data. Hence neural networks provide an excellent classification schema. However, they generally employ a non-parametric model as their underlying learning mechanism focuses on *adaptive* error correction. The scope of this work suggested development of an explicit cost function and needed an appropriate parametric model. Since a model-free neural network classifier would not establish the class-conditional probabilities for classification, it would be less appropriate for this thesis.

This work can be considered a logical progression of the few past attempts to achieve the NVAFEC and AAR functionality using the facial thermal features. However, this thesis is distinct in that it uses the facial thermal features, for the first time, to (a) classify the complex facial expressions along the direction of valance; (b) classify common facial expressions of affective states; (c) classify facial expressions on occluded or covered faces; (d) recognise the involuntarily evoked facial expressions of affective states.

1.3 Research focus

A significant number of sophisticated algorithms and advanced computational methods for implementing the vision-based AFEC systems is available in the image processing and pattern recognition literature (Belhumeur et al. 1997; Black and Yacoob1997; Cohn et al. 1999; Dubuisson, et al. 2002; Fasel and Luettin 2003; Huang and Huang 1999; Pantic and Rothkrantz 2000). However, little work has been done on extracting, selecting, representing and classifying the bio-physiological signals for AFEC and AAR. Particularly, automated facial expression classification using the facial skin temperature measurements has not been fully explored yet.

This work focuses on developing effective computational approaches and methods of non-invasive thermal feature extraction, selection and representation, and their classification for developing a robust NVAFEC and AAR capable systems.

The work began by investigating an effective mechanism for extracting the facial thermal features pertaining to the expressions of affective states from the thermal infrared images. An attempt was made to design a realistic model for discovering the most influential and relevant facial thermal features from the thermal infrared images. It was envisaged that the discovery of the most effective facial thermal features would help in direct estimation of the probabilities of facial expression group membership of a thermal image through regression. It was further assumed that estimations of the probabilities of facial expression group membership would help develop a set of efficient discriminant functions for person-independent classification of expressions of affective states. The thesis therefore focuses on developing a robust and non-invasive "emotion detection through facial expression recognition" mechanism that might allow for example, a robot, to gain a higher degree of social intelligence and effectively interact with people.

1.4 Research Contributions

Considered in a broader context, this work contributes in the areas of affective computing, human-computer interaction, thermal infrared imaging application, automated affect recognition and applied perception. The thesis focuses on extracting the facial thermal features form the thermal infrared images, selection of the most useful facial thermal features and their effective representation in a decision space to classify the facial expressions of affective states. The aforementioned scope set for this investigation allowed making following contributions.

First, the emotion-specific human bio-physiological cues and their respective effectiveness in developing a non-invasive AFEC capable system were reviewed. The viability and effectiveness of using infrared measurements of facial skin temperature measurements for automated classification of facial expressions were also explored, analysed and reported.

Second, the scientific foundations of non-invasive, thermal infrared sensing of emotions and affective states were reviewed, analysed and established. Instead of exploiting the thermal symmetries or asymmetries in the infrared images for binary classification of positive and negative affective states, a novel mechanism was developed to extract the facial thermal features from a set of time-sequential thermal infrared images.

Third, an architectural framework was developed to implement the facial skin temperature measurements based classifier network. An optimal pattern representation scheme was developed and for the first time, the facial thermal classification patterns were modeled as stochastically independent and identically distributed clusters. An appropriate supervised learning approach was developed to construct the optimal discriminant rules. The stochastically independent clusters of facial expressions were separated as linear spaces within an optimal decision space allowing person-independent recognition of the facial expressions of affective states.

Fourth, using a parametric estimation of the posterior probabilities, the facial expressions of affective states were first classified along the direction of valence. At a later stage, a complex decision space was constructed to classify the six common facial expressions of affective states.

Fifth, the possibilities of recognising facial expressions using facial thermal features extracted from the partially covered or occluded faces were also investigated.

Sixth, the differences in the patterns of thermal representation of pretended and naturally occurring facial expressions were examined. The differences in a classifier's ability to recognise pretended and evoked facial expressions using the facial thermal features were also studied.

Seventh, an agenda was proposed for future work on the use of bio-physiological cues in designing the AFEC and AAR capable systems.

Finally, new knowledge was added to the existing body of relevant knowledge. The results of this investigation were reported in prestigious and fully refereed publications. Appendix II provides a complete list of publications that emanated from this thesis.

1.5 Thesis overview and organisation

This document comprises of 10 chapters, 2 appendices and a list of references.

Chapter 2 first discusses the existing and potential application of automated facial expression classification (AFEC) systems. The system design approaches used for developing the AFEC and AAR systems are then analysed. Strengths and limitations of existing AFEC systems are also examined. Scientific information about the human physiological information are used for proposing the use of facial skin temperature measurements in AFEC and AAR.

Chapter 3 begins by reporting the recent studies that suggest an association between the emotion and bio-physiological signals. Previous investigations carried out to explore emotion-specific musculo-thermal, physiological and autonomic activities are reviewed. Possibilities of measuring emotion-specific body information are investigated and scientific studies proposing an association between the core body temperature and the emotional states are presented. Important methods and tools used for measuring the skin temperature are also discussed.

Chapter 4 begins by reporting the infrared image acquisition procedure. The ethical considerations for the experiment design and the thermal image processing methods are reported. Approaches employed to extract the transient thermal features from the participant faces in the thermal images are also explained. The thermal data were analysed to examine if the acquired data were suitable for invoking the multivariate analytical methods and pattern recognition algorithms.

Chapter 5 first introduces the generic architecture of an AFEC capable system and then examines the possibilities of adapting a typical AFEC system architecture for this investigation. Some pattern recognition approaches used in the previously developed AFEC capable systems are reviewed. Based on an analysis of the existing AFEC systems, the architecture and functional design of a facial skin temperature based AFEC system is proposed. The computational methods, algorithmic implementation and potential advantages of the proposed approach are also examined and presented.

Chapter 6 reports an attempt to recognise and classify the neutral and pretended happy and sad facial expressions using the facial thermal data. Classification of the two positive (happiness and surprise) and the two negative (angry and disgusted) facial expressions is also reported in this chapter. Detailed analyses of the classification results conclude this chapter.

Chapter 7 is dedicated for reporting the results of classifying the six basic facial expressions (happy, sad, disgust, surprise, angry and fear) using the facial skin temperature measurements. Classification results are analysed and the observed results are discussed.

Chapter 8 begins by discussing the influence of factors like facial hair, glasses, lighting conditions, pose and occlusion on the real life performance of an AFEC system. An argument is then made for bio-physiological signal based classification of affective states when the face is covered or occluded. The facial muscle grouping approach used for representing the covered and occluded faces is presented. Finally, classifier implementation details and classification results are presented and analysed.

Chapter 9 realises that in a life like situation, the AFEC is performed on the naturally occurring, spontaneous and evoked or reactive expressions. The investigations reported in this chapter examine the effectiveness of facial thermal features in classifying the evoked facial expressions. Details of the evoked facial thermal data acquisition approach, data analyses and classification are reported. Observed results are analysed and the classifier performance is compared with that of the previously developed classifiers.

Chapter 10 provides a summary of this work. Observed results are discussed and analysed. The discussion and analyses provide rationale for making several inferences about the viability of developing the bio-physiological signals based AFEC and AAR capable systems. An agenda for the future work is proposed using the discussion and analyses of the employed computational approach.

Appendix I reports the human protection practices observed during the design of experiments and data acquisition.

Appendix II presents a list of accepted, published and submitted publications that emanated from this work.

References section is appended at the end of this document. It provides a list of cited work.

Chapter 2

AUTOMATED CLASSIFICATION OF FACIAL EXPRESSIONS

Emotions, their recognition and expression make people to people communication comprehensive, effective and meaningful (Niemic 2002). In order to make human-computer interaction similarly effective, researchers have been investigating the possibilities of developing affective human-computer interaction models (Picard 2000). Hence, automated facial expression classification (AFEC) and automated affect recognition (AAR) have emerged as important research areas during the last three decades (Allanson and Fairclough 2004).

Using the taxonomy proposed in chapter 1, a survey of existing AFEC and automated affect recognition approaches is presented in the following paragraphs. The system design approaches used for developing the AFEC and AAR capable systems are also discussed. The strengths and limitations of the widely used AFEC approaches are also examined in this chapter.

2.1 The need for AFEC and AAR capable systems

Facial expressions are considered a major source of information about emotions, intentions and affective states albeit they are often used together with the non-visual human information such as voice and body movement (Bartneck 2001; Cacioppo et al. 1990; Ekman 1982). A significantly large number of studies suggest that facial expressions provide highly useful visual information about emotions and affective states (Busso 2004; White 1999). Scientists assert that facial visual information are used during any people to people communication (Morishima 2001), help understand cognition and behaviour (Sloan et al. 2002), and are believed to have a significant role in the future HCI systems (Lisetti and Nasoz 2004; Hosseini and Krechowec 2004). Researchers assert that achieving the AFEC functionality in machines may be useful for several professional communities and user groups (Ekman et al. 1993; Gao et al. 2003).

Studies suggest that psychologists primarily rely on facial expressions for clinical investigations (Christie and Friedman 2004). They analyse the facial expressions to interpret emotions, understand intentions, evaluate personality and assess the cognitive conditions (Collet et al. 1997; Critchley et al. 2000; Dimberg 1990b). AFEC capable systems are considered potentially useful in enhancing psychologists' ability to judge an individual's personality and determine his/her personal traits such as shyness and sociability (Dimberg et al. 2000). Some studies also propose use of facial expressions for gathering information about psycho-pathological conditions and nature of behavioural disorders (Sloan et al. 2002).

Security, intelligence and surveillance communities are believed to be potential beneficiaries of AFEC capable systems for detecting and discovering concealed emotions and intentions (Pavlidis and Levine 2002). Automated classification of facial expressions has been proposed for interpreting intentions and emotions in real life situations. For example, (Garbey et al. 2004) used the blood flow rate estimation on the face for facial expression recognition and classification for developing an AFEC capable system. In other recent investigations, thermal images were analysed for scoring polygraph tests in lie-detection. For example, Pavlidis (2000) analysed thermal images acquired using an infrared camera to detect the blood flow rate variation on the face. (Pollina et al. 2006) monitored variations in facial skin temperature to classify people who committed crimes into deceptive and non-deceptive categories.

Scientists assert that emotions, feelings and momentary experiences cause, trigger and influence human facial expressions. Consequently, facial expression monitoring has been proposed for pain measurement, patient monitoring, and patho-physiological diagnosis and condition monitoring (Diakides 1998; Hosseini and Krechowec 2004; Hussein and Granat 2002; Sloan et al. 2002). Studies propose that automated recognition of emotions would be useful in medical diagnosis and allow better patient monitoring (Diakides 1998). Similar studies inspired researchers to propose monitoring patients' health conditions using automated affect recognition (Hosseini and Krechowec 2004). Use of AFEC to interpret affective states for determining patients' well being in remote and on-line health monitoring systems has also been proposed (Hosseini and Krechowec 2004). In an earlier study, the AFEC systems were able to provide a realistic assessment of pain (Herry and Frize 2002). Scientists have also used facial expression analysis for gathering information about patients' intentions (Hussein and Granat 2002). In a closer field, investigators proposed the non-contact measurement of stress and emotional conditions through thermal infrared imaging of the face (Puri et al. 2005).

AFEC was also proposed to the computing and engineering communities for developing smart and adaptive man-machine interaction mechanisms and the HCI interfaces (Ekman et al. 1993; Kearney and Mckenzie 1993; Picard et al. 2001; Reeves and Nass 1996). Some reportedly successful applications include software and appliance usability tests, adaptive system design and operationally critical systems' monitoring (Henderson et al. 1995; Ohnishi and Sugi 1996; Ward et al. 2003; Zaatri and Oussalah 2003). (Ward et al. 2003) confirmed in a recent investigation that tracking the facial expressions might assist in conducting software usability tests.

A number of recent studies reported the growing importance of human factors in engineering. These studies cited a significant increase in the number of manufacturers who perform user-satisfaction tests for satisfaction assessment and product quality audits (Paterno 2005; Wilfong 2006; Zaatri and Oussalah 2003). AFEC methods were reportedly capable of revealing the positive and negative user reaction and thus allowed better assessment of user satisfaction. As a result, the manufacturing sector was also anticipated to benefit from the AFEC and AAR functionality (Paterno 2005; Wilfong 2006; Zaatri and Oussalah 2003).

Multidisciplinary studies in applied psychology and HCI also make a strong case for real time automated recognition of positive and negative expressions of emotions. For example, computer anxiety, a major obstacle in the professional and personal development of many individuals, may be detected through measurement of negative emotions using an AFEC capable system (Bozionelos 2001; Brosnan 1998; Wilfong 2004). Meyer and Rakotonirainy (2003) reported that recognition of positive and negative emotions and their expressions could enable sensors activate life-support systems in context-aware homes.

Non-invasive recognition of negative and positive facial expressions was also desirable for man-machine interaction (Picard 1999). The HCI literature reported that facial expression analysis could help multimodal human-computer interaction devices in controlling critical industrial systems. Investigators were able to use facial expression tracking in sophisticated control systems for adapting and responding to emergencies and critical conditions (Paterno 2005; Zaatri and Oussalah 2003).

A robust AFEC functionality was also considered important for building socially aware systems and sociable intelligent robots (Brooks 2002). Researchers argue that service robots and rehabilitation machines must acquire a reliable AFEC and AAR capability (Arkin et al. 2003; Hara and Kobayashi 1997; Sugimoto et al. 2000). Scientists foresee that systems capable of interpreting emotions from facial expressions may soon be able to respond to various social situations and an individual's personal needs (Brooks 2002).

Theses potential uses of AFEC capable systems seem to have inspired scientific communities to explore new and advanced methods of developing more robust and reliable AFEC and AAR capable systems (Brooks 2002). Significant efforts were made to design and built machines that would recognise facial expression of affective states (Cohen et al. 2003; Fasel and Luettin 2003). These so called socially intelligent machines try to interpret affective states using some form of visual cues gathered from the facial expression of emotive states. Some of these machines also employ carefully selected non-visual signals or a combination of both visual and non-visual cues (Ang et al. 2004; Christie and Friedman 2004; Lisetti and Nasoz 2004; Herry and Frize 2002).

The following sections of this chapter examine the existing AFEC approaches employed for implementing the AFEC and AAR capable systems. High-level architecture and implementation of the AFEC capable systems and their functional components are discussed in the following chapters. The computational methods employed for designing the AFEC capable systems are discussed in Chapter 5.

2.2 Existing AFEC enabling approaches

During the last three decades of the last century Paul Ekman and his group carried out important theoretical and empirical work on human facial expression analysis and representation. They developed a Facial Action Coding System (FACS) to code facial expressions using muscles' movements on the face (Ekman and Friesen 1978). They also discovered evidence for supporting the universality of facial expressions (Ekman 1992). Another coding system, known as the Maximally Discriminative Affect Coding System (MAX), was developed around the same time (Izard 1979). MAX also received considerable attention and was used in a small number of investigations carried out for facial expression analyses (Ekman et al. 1993).

The emergence of FACS and MAX inspired automated analysis of facial expressions using still images and frame sequences such as visible-spectrum video clips (Cohen et al. 2003; Ekman et al. 1993). Researchers were able to track the facial features and measure the amount of facial movement to categorise the universal facial expressions. The work Ekman and Friesen (1978) carried inspired a significant number of recent works on vision-based facial expression recognition and classification. They introduced a method of measuring the facial movements in terms of facial Action Units (AUs) for classifying basic facial expressions (Ekman and Friesen 1978). Several investigators employed FACS for AFEC and AAR (Cohen et al. 2003; Ekman et al. 1993; Fasel and Luettin 2003; Mase 1991; Pantic and Rothkrantz 2000).

Vision-based AFEC capable systems have demonstrated little practical use in real life situations (Baldwin et al. 1998). Limited success of vision-based AFEC capable systems appears to be the major driving force behind the attempts of developing the non-vision-based AFEC (NVAFEC) capable systems. Current literature published in the areas of bioinformatics, pattern recognition, image processing, human information processing and HCI suggested fusion of visual and non-visual cues could enhance the AFEC and AAR functionality (Christie and Friedman 2004; Lisetti and Nasoz 2004; Pham et al. 2000; Sebastiani et al. 2003). Fusion of visual and non-visual cues for AFEC and AAR was proven technically possible and computationally viable in some of the reported systems. Several recently developed systems used a combination of visual and non-visual cues (such as auditory and bio-physiological signals) for achieving the AFEC functionality (Kim et al. 2004; Sugimoto 2000; Yoshitomi 2000).

The theoretical background, empirical results and technical issues involved in design and implementation of both vision-based and non-vision-based AFEC capable systems are available in (Black and Yacoob 1997; Christie and Friedman 2004; Christine and Nasoz 2004; Ekman et al. 1993; Essa and Pentland 1997; Fasel and Luettin 2003; Kim et al. 2004; Pantic and Rothkrantz 2000; Sugimoto 2000; Yoshitomi 2000).

The literature cited above groups the AFEC systems on the basis of the types of signals used and the architectural implications of the employed signals. The AFEC and AAR capable systems can be divided into two broad categories:

1. Vision-based AFEC systems
2. Non-vision based AFEC (NVAFEC) systems

This categorisation is used in this thesis to highlight the significant differences in the measurement and processing of signals and cues used in the AFEC and AAR capable systems.

Though capabilities and limitations of these two different types of AFEC systems differ, they share several common features. For example, the two categories of AFEC systems appear to have similarities in their top-level generic architecture. Both types of AFEC systems work in multiple stages. Both visual and non-visual cues are first measured and then transformed into a particular desired format of usable signals. Once the signals are transformed into usable data (features) they are represented in a decision space using a proper representation approach. Finally, selected facial features are used to classify the faces using a set of discriminant rules. The two approaches of achieving the AFEC and AAR functionality are further examined in the following paragraphs.

2.3 Vision-based AFEC systems

Most vision-based AFEC systems rely on estimation of facial muscular movements and associated physiognomic activities taking place on a human face. The physiognomic data and the information gathered from the facial muscular activities are interpreted and processed using one of the several recognition and classification methods (Ekman et al. 1993). Scientists were able to relate muscular activities with the physiognomic signals observed on a human face (Ekman et al. 1993). It is believed that these signals, individually and collectively, participate in visually noticeable facial activities. The visual signals generated through the facial muscles' movement are generally classified into four general categories (Ekman et al. 1993):

1. Static facial signals representing the permanent features of the face;
2. Gradual and slowly changing facial signals representing changes in the appearance of the face over time;

3. Artificial signals representing exogenous features of the face such as hair, jewelry and glasses; and
4. Rapid facial signals representing changes in neuromuscular activities leading to visually detectable changes in facial appearance, such as opening of mouth or drop of jaws.

The fourth class of signals is considered more related to the facial muscular activity. It is therefore considered a major source of input to the facial expression analysis (Ekman et al. 1993). Scientists have developed the facial action measurement systems such as MAX and FACS using the measurements of the fourth class of signals (Ekman et al. 1993; Ekman and Friesen 1978).

FACS, generally accepted as a comprehensive coding system, is a widely used method of measuring the facial muscular movements. It attempts to measure contraction of each individual muscle alone and in combination with other muscles. The measurements are used to observe any changes in the appearance of a face (Black and Yacoob 1997; Ekman et al. 1993; Essa and Pentland 1997; Fasel and Luettin 2003; Gao et al. 2003).

MAX is a comparatively less popular coding system used for measuring the facial muscle movement. It is regarded as a theory-based system that measures visible changes on the face (Lisetti and Schiano 2000). It was contested for its underlying theoretical basis suggesting that only specific areas of a face should be involved in certain emotions and their (Lisetti and Schiano 2000). Since it employs the observation and measurements of visible changes specific to emotions related appearances for classifying facial expressions, its implementation is considered complex and cumbersome (Ekman et al. 1993). MAX uses the units formulated in terms of appearances relevant to (only) eight specific emotions (Lisetti and Schiano 2000).

In addition to FACS and MAX, two other methods: optical flow analysis and 3D wireframe model analysis are used in AFEC.

Optical flow analysis is a popular techniques employed for facial expression analysis (Gao et al. 2003). The method relies on the measurement of apparent motion of brightness in an image (DeCarlo and Metaxas 2000). Optical flow analysis allows measuring facial muscle activity on an individual’s face. The facial features are represented as patches whose motion in an image sequence is modeled using several low-level polynomials. The optical flow estimate is used to recover the motion of patches. These observed motion parameters are believed to provide a concise description of facial feature motion (Black and Yacoob 1997). Optical flow analysis based AFEC systems reportedly achieved up to 80 % classification accuracy in controlled test environments (Yacoob and Davis 1996). Several investigations extended optical flow analysis approach for achieving the AFEC functionality. For example, the optical flow model was employed to develop a physical model of human face using a recursively refined and improved facial motion estimation (Essa and Pentland 1997). In another study, estimates of forces resulting from the facial muscular movements were used for AFEC (Cabanac and Guillemette 2001). The system claimed to achieve up to 98% accurate classification results (Cabanac and Guillemette 2001). Black and Yacoob (1997) proposed another approach that employed the local parameterised model of image motion and holistic spatial analysis. This approach was considered computationally expensive albeit it provided a plausible method of measuring the facial muscular movements (Gao et al. 2003).

The 3D wireframe face model provides another noteworthy facial expression analysis approach. It involves developing an explicit 3D wireframe face model to track the geometric facial features defined on a face (Tao and Huang 1999). However, developing a 3D face model for tracking the facial muscular movements involves a complex process. Investigators noted that constructing a representative 3D model of certain anatomical areas of a human face was computationally demanding (Gong et al. 2000; Gur et al. 2002).

Vision-based AFEC is usually performed using one of the aforementioned methods. Though each of these methods poses a unique set of problems, they help achieve the AFEC and AAR functionality (Fasel and Luettin 2003; Lisetti and Schiano 2000; Pantic and Rothkrantz 2000; Tina et al. 2001).

Vision-based AFEC systems attempt to measure the temporal facial muscular movement and estimate the resulting energy changes in the facial muscles for AFEC and AAR (Essa and Pentland 1997). Hence, vision-based AFEC systems typically employ either static images or a sequence of images to detect faces, extract features and classify expressions (Fasel and Luettin 2003; Pantic and Rothkrantz 2000). Facial muscular activities are regarded momentary and subtle (Ekman et al. 1993; Essa and Pentland 1997). Measurement of instantaneous contraction of muscles and associated changes in the appearance of the physiognomy involves observing location, intensity and dynamics of the facial muscular actions (Ekman et al. 1993). Vision-based AFEC systems would therefore (typically) work in three high-level stages. These three stages are referred to as; facial feature extraction, feature selection and representation, and expression classification (Fasel and Luettin 2003; Pantic and Rothkrantz 2000). Architectural implementation, functional description of components, and performance evaluation of most popular AFEC systems are available in (Baldwin et al. 1998; Bartneck 2001; Brooks 2002; Cohen et al. 1993; Donato et al. 1999; Ekman et al. 1993; Essa and Pentland 1997; Fasel and Luettin 2003; Huang and Huang 1997; Kearney and McKenzie 1993; Posamentier and Abdi 2003; Yoshitomi 2000).

2.4 Limitations of the vision-based AFEC systems

Affective computing and pattern recognition literature identifies a number of problems that restrict the real life application of AFEC systems. (Picard et al. 2001) argued that a majority of experiments on these AFEC systems was conducted in controlled laboratory environments using pre-segmented data. Hence, it is argued that the real life performance and relevance of reportedly successful AFEC systems need further verification. Participants’ control over intensity of emotions during the experiments was also perceived arbitrary and unrealistic in some studies. Critiques argued that the underlying relationship between the internal feelings and their external expressions was not considered during the experiments and performance tests (Baldwin et al. 1998). The true nature of the association between the physiological muscle movements and facial expressions, it is argued, has not been established yet. (Picard et al. 2001) point out that a universally agreed meaning of the neutral state does not exist and the relationship between the inter and intra emotional states remain unexplained. It is also argued that the mapping of several emotions to a few expressions can also make the AFEC results ambiguous and misleading. (Baldwin et al. 1998), questioning the reliability of vision-based AFEC systems, argue that a person can simultaneously experience more than one affective state and the AFEC systems cannot realise that fact. It is also argued that anger and depression or joy and positive surprise can be experienced together albeit a vision-based AFEC system cannot notice that. The inability to understand any relationship between various intensities of expressions and emotions, and the lack of ability to distinguish between the reactive emotions from pretended ones make relevance of the vision-based AFEC systems questionable (Baldwin et al. 1998; Ekman et al. 1993; Klein et al. 2002; Picard et al. 2001).

The multi-disciplinary literature on AFEC and AAR highlights many core problems that hinder the performance and reliability of the vision-based AFEC systems. For example it is argued that the issues pertaining to the geometric complexity of the facial features need to be resolved. Methods and means to systematically acquire, understand and process facial geometric information have yet to be further developed, tested and validated. The ability to transform geometric variations into useful information (with a high degree of accuracy) has yet to be achieved. Avoiding any misinformation embedded into the geometric variations is not possible yet. Issues concerning the deformability of facial features are also important and need to be addressed. The vision-based AFEC systems lack the understanding of the intensity and valance of deformation as it pertains to an individual face and a broad category of facial expressions. Issues concerning the validity of data are also important. Avoiding noise, usually caused by factors such as variation in the intensity of ambient light, is not possible yet. The intensity of light influences face detection and resulting feature extraction in an AFEC system. The measurement of any correlation between intensity of expressions and underlying emotional estate is also required for a better and reliable affect recognition (Baldwin et al. 1998; Ekman et al. 1993; Gao et al. 2003; Klein et al. 2002; Picard 2001).

In order to achieve a robust and reliable AFEC and AAR functionality, researchers tend to use the human bio-physiological information in the NVAFEC capable systems. Attempts were also made to infuse the auditory signals, the bio-physiological cues, and the visual cues. An overview of the NVAFEC systems is provided in the following section.

2.5 The NVAFEC systems

Investigators were able to discover a strong relationship between the emotion, human autonomic response and bio-physiological signals. Some researchers were able to measure the basic emotions using either autonomic signals or bio-physiological cues (Christie and Friedman 2004; Christine and Nasoz 2004). Details of some successful NVAFEC and AAR capable systems are reported in (Christie and Fireman 2004; Collet et al 1997; Kim et al. 2004; Stemmler 1989; Yoshitomo 2000). These signals, found helpful in achieving the AFEC and AAR functionality, are measured using some purpose-specific measurement devices or sensors. Of the above human information, only the skin temperature measurements could be measured through non-invasive means. Measuring other signals might require a direct contact with the human body.

NVAFEC is a new approach but a growing interest in this approach is evident in recent publications in the domains of computer perception, HCI, biometric recognition, and intelligent systems (Christine and Nasoz 2004). Some existing NVAFEC and AAR systems reportedly employed auditory signals, electrocardiogram readings, electrodermal measurements, skin temperature variations, brain activity measured using the electroencephalograms (EEG), functional magnetic resonance imaging (fMRI) readings, and blood oxygen level dependent (BOLD) data (Christine and Nasoz 2004; Niemic 2002). Of these signals, researchers find sound signals, brain signals, electrodermal measurements, and skin temperature highly feasible for developing the AFEC and AAR capable systems (Christine and Nasoz 2004; Ohnishi and Sugie 1996; Picard 2002; Pollina et al. 2006).

The fusion of auditory signals with the visual and/ or bio-physiological cues for AFEC and AAR was also tried in some recent investigations. Studies propose that auditory signals work in all lighting conditions but the ambient noise can easily influence the sound signals. Such influences and their associated problems warrant use of special signal processing tools and methods. Humans can exercise some control on receiving and transmitting communication signals so control over auditory signals is always possible during people-to-people communication. Inability to exercise similar control in machines may allow deception and may deteriorate the understanding and measurement of the relation between the voice signals and true emotional estates.

Scientific studies suggest that humans do not posses the ability to completely control the autonomic response and bio-physiological reaction to emotions and their expression. Fusion of bio-physiological signals with either visual cues or auditory signals is therefore believed to result in more robust and reliable AFEC and AAR functionality.

The theoretical foundations, methods, tools, and architectural details of NVAFEC capable systems are available in the recent works by (Ang et al. 2004; Boulic and Thalmann 1998; Christie and Friedman 2004; Christine and Nasoz 2004; Critchley et al. 2000; Niemic 2002; Pizzagalli et al. 1998; Posamentier and Abdi 2003; Prokoski and Iedel 1999; Sarto et al. 2005; Schwarz et al. 2002; Sebastian et al. 2003; Socolinsky et al. 2003; Stern et al. 2001; Szabo et al. 2000; Yoshitomi et al. 2000).

A growing interest in using skin temperature measurements for developing the AFEC and AAR capable systems is evident in the literature. In many recent studies, skin temperature was measured using Thermal Infrared Imaging (TIRI) equipment. Successful use of TIRI for developing non-invasive AFEC and AAR capable systems is reported in (Christine and Nasoz 2004; Pavlidis 2000; Pavlidis 2004; Pavlidis and Levine 2002; Pollina 2006; Puri et al. 2005). Attempts were also made to use the TIRI with the visible imagery and voice for AFEC (Sugimoto et al. 2000; Yoshitomi et al. 2000). Previous work suggested that skin temperature measurements through the TIRI would assist in AFEC and AAR (Bolle 2004; Kakadiaris et al. 2005a; Kakadiaris et al. 2005b; Kong et al. 2005; Kunzmann and Gruhn 2005; Kurse et al. 2001; Matsuzaki and Mizote 1996; McGimpsey et al. 2000; Ogasawara et al. 2001; Pavlidis 2004; Pollina 2006; Socolinsky et al. 2003; Stemmler 1989; Sugimoto 2000; Wolf et al. 2005; Yoshitomi et al. 2000). TIRI offers following advantages in the context of AFEC and AAR.

1. Facial skin temperature can be measured from a distance using the infrared cameras. Since no body contact is required, the target person may not notice any thermographic activity though this may result in breach of personal privacy and may raise some ethical issues. Despite these issues, surveillance and security communities require non-contact and secret monitoring of suspects and would benefit from TIRI based AFEC and AAR;
2. Modern infrared equipment allows non-invasive thermographic measurements. This may be particularly useful for medical and

psychological diagnostic applications under conditions when patients are either unable or unwilling to cooperate;

3. TIRI is invariant to light and illumination conditions;
4. TIRI equipment is accessible and is becoming less expensive and affordable;
5. Modern TIRI equipment is light, aesthetically appealing and is easy to handle;
6. The latest infrared cameras are highly sensitive to any thermal variations on the human skin. These cameras are capable of sensing up to ±0.05 ºC thermal variations;
7. TIRIs provide both visual and physiological information for the AFEC and AAR; and
8. TIRI is safe and harmless to both the user of infrared equipment and the target individual.

However, several questions are raised about the potential problems of using TIRI in AFEC and AAR applications. For example, can the facial skin temperature measurements be transformed into useful signals for facial expressions recognition? Is the association between the measured facial skin temperature variations and expression of affective states strong enough to help in AFEC and AAR? How effective would TIRI be in sensing any facial thermal variations? Are appropriate computational methods available for extracting and representing thermal features for AFEC and AAR? Is there any historical evidence of using the TIRI in AFEC and AAR?

The literature reviewed for this thesis and reported in the next chapter provides an insight for answering these and similar questions. The literature reports that factors such as blood volume flow variation, musculo-thermal activities, and body metabolism react to emotive states and cause a change in the facial skin temperature (Briese 1995; Collet et al. 1997; Drummond and Lance 1987; Kistler et al. 1998; Naemura et al. 1993; Sinha and Parson 1996). Effectiveness of TIRI in measuring the facial skin temperature variations was also reported in the literature (Jones and Plassmann 2002; Khan et al. 2006; Otsuka et al. 2002; Pollina et al. 2006).

2.6 Conclusion

The vision-based automated facial expression recognition and classification has been more successfully under perfect and controlled conditions. Several methodological and functional limitations have been identified as detrimental to the real life performance of the vision-based automated facial expression classification systems. Recently, non-visual cues and the bio-physiological signals were employed for developing the automated facial expression classification and automated affect recognition systems. One major problem in employing the bio-physiological signals for automated facial expression classification and automated affect recognition is the invasive nature of the signal measurement mechanisms. Thus, unconscious facial expression monitoring is not possible if the bio-physiological signals are employed for AFEC and AAR.

Thermal infrared measurement of skin temperature promises a non-intrusive and technically apt mechanism for achieving the non-invasive AFEC and AAR functionality. Previous researchers have reported encouraging results in the use of facial skin temperature measurements in the AFEC and AAR. The following chapters examine the recent approaches for employing the facial skin temperature measurements in AFEC and AAR capable systems.

Chapter 3

MEASUREMENT OF EMOTION-SPECIFIC AUTONOMIC AND PHYSIOLOGICAL INFORMATION

This chapter discusses how humans physiologically response to emotions and presents mechanisms available to detect the human response. Previous studies that discovered an association between the core body temperature and the affective states are presented. Methods and tools for measuring the emotion-specific human information in the cited studies are also discussed. Emotion-specific changes in human body and their detection through skin temperature measurement are also discussed. Uses of thermal infrared imaging techniques for non-invasive measurement of skin temperature are also reviewed. This chapter finally proposes using infrared measurement of transient facial skin temperature variations to develop non-invasive automated facial expression classification (AFEC) and automated affect recognition (AAR) systems.

3.1 Emotion-specific autonomic and physiological information

Ancient wisdom, preserved in various fiction forms and literature, believed in an association between emotion and unintentional observable human responses. Modern science supports this ancient wisdom and confirms the existence of a relationship between the emotions and autonomic response (Ang et al. 2004; Allanson and Fairclough 2004; Busso et al. 2004; Christine and Nasoz 2004; Ekman 1982; Ekman et al. 1983; Ekman et al. 2000). Scientists were able to discover and measure the involuntary autonomic and physiological signals generated in response to the affective states using some purpose-specific equipment, sensors or devices (Bradley et al. 2003; Dimberg, U. 1990a; Dimberg, U. 1990b; Palomba et al. 2000; Sinha and Parson 1996; Wright et al. 2004).

(Sinha and Parson 1996) examined the bio-physiological parameters in normal conditions, and in response to situations of anger, fear, joy, and sadness. They examined participants' heart rate, skin conductance level, finger temperature, blood pressure, electro-oculogram readings, and facial EMG recordings. They concluded that emotion-specific physiological response patterns to fear and anger were significantly different than the ones observed in participants under the neutral conditions (Sinha and Parson 1996).

(Ekman et al. 1983) noted that evoked autonomic activity would help in distinguishing between the positive and negative affective states. Using the autonomic activity measurements, they were also able to discern between the other negative emotions (Ekman et al. 1983; Zajonc 1985).

(Christie and Friedman 2004) also investigated the autonomic nervous system response to the experimentally manipulated emotions. Their study confirmed previous findings that had discovered existence of emotion-specific autonomic nervous system activity.

(Collet et al.1997) reported several relevant works that investigated the association between the basic autonomic activities, emotions and expression of emotions. The six autonomic nervous system parameters: skin conductance, skin potential, skin resistance, skin blood flow, skin temperature and instantaneous respiratory frequency were found to be responsive to the *basic emotions* in the studies (Collet et al. 1997) cited. A prominent scientist Paul Ekman is credited for coining the terms *basic emotions* and *basic facial expressions* (Ekman 1992). He proposes that the six basic facial expressions and six basic emotions may be used to represent all major emotional experiences. Paul Ekman suggests that the expressions of anger, disgust, fear, happiness, sadness and surprise can represent all major emotional experiences. Any reference made to the basic emotions or six basic facial expressions in this thesis is based on Paul Ekman's theory of six basic emotions.

A significant number of studies conducted during the last three decades suggest that visual and bio-physiological information may provide useful human information for clinical and medical investigations, biometrics, security and surveillance, criminal investigation and HCI. More recently, few attempts were made to use the bio-physiological signals in AFEC and AAR (Christie and Friedman 2004; Collet et al. 1997; Naemura et al. 1993).

Advanced methods of measuring the bio-physiological signals including some imaging techniques were employed in the reported investigations to acquire the human physiological information (Allanson and Fairclough 2004; Dimberg 1990a; Dimberg et al. 2000; Dimberg and Petterson 2000; Hess et al. 1992; Iwase et al. 2002; Lundqvist 1995; Root and Stephens 2003; Vrana 1993; Varna and Gross 2004; Veldhuizen 2003; Winkielman and Cacioppo 2001).

3.2 Emotion-specific musculo-physiological activities on the face

A significant number of investigations have been carried out to explore the emotion-specific musculo-physiological activities on the human face. Notably, the EMG (electromyogram) measurements were helpful in ascertaining an association between the emotion and the musculo-physiological activities. The EMG measurements rely on the electrical potential measured in the skeletal muscles. The EMG technique is usually employed as a diagnostic test since it records the electrical response of skeletal muscles while at rest and during any voluntary or stimulated facial action. The EMG is considered a robust and reliable technique (Iwase et al. 20002; Vrana and Gross 2004; Wolf et al. 2005). The major facial muscles that are considered responsive to emotions are shown in Figure 3.1.

A recent investigation used EMG measurements to investigate how the two groups of facial muscles: Orbicularis Oculi, and Mentalis and Depressor Anguli Oris would contribute to the facial expression of pain (Wolf et al. 2005). The results concurred with the previous studies and explained the relationship between the facial expression of pain and the two muscle groups (Wolf et al. 2005).

(Vrana and Gross 2004) compared the EMG response in Zygomaticus Major and Corrugator Supercilii EMG response pertaining to the feelings of joy and anger. They reported a greater Zygomaticus Major EMG response to the feeling of joy as compare to that of the Corrugator Supercilii. They also discovered the physiological response to expression of anger was similar to that observed in neutral condition (Vrana and Gross 2004).

(Veldhuizen et al. 2003) studied influence of mental fatigue on facial EMG activity during a simulated workday. They were able to identify an association between mental fatigue and the facial EMG activities (Veldhuizen et al. 2003).

(Root and Stephens 2003) investigated the organisational patterns of the central control of facial expression muscles in men. They recorded the surface EMG on the ipsilateral pairs of facial muscles while participants intentionally smiled, expressed sadness and horror. They discovered peaks in the cross-correlograms of the EMG readings on the Orbicularis Oculi and the Zygomaticus Major during smile, on the Corrugator and the Depressor Anguli Oris while expressing sadness and on the Frontalis and the Mentalis during the horror look (Root and Stephens 2003).

Figure 3.1: Frontal view of the facial muscle map showing all major facial muscles

(Iwase et al. 2002) investigated the neural substrates of the facial expression of induced experiences of joy. They observed a strong correlation between the regional cerebral blood flow in the bilateral supplementary motor area and the left putamen. They also observed significant correlation between the regional cerebral blood flow in the primary motor area and the magnitude of the EMG readings in the bilateral supplementary motor area of the face (Iwase et al. 2002).

(Winkielman and Cacioppo 2001) observed that ease of task (and task processing) could elicit higher EMG activity over the region of the Zygomaticus Major. They inferred existence of a physiological correlation between ease of task and positive emotional response (Winkielman and Cacioppo 2001).

Some relevant studies suggest that Zygomaticus Major is the primary muscle of smile, Orbicularis Oculi is the major muscle of joy and Orbicularis Oris is the smile modifier muscle (Kall 1990). Studies also suggest that Zygomaticus Major, Orbicularis Oculi, Mentalis, Platysma and Orbicularis collectively represent happiness and joy (Kall 1990). Similarly, Corrugator, Masseter, Triangularis, Orbicularis Oculi Palpabraeous, Procerus Nasi, Labii Inferioris, and Platysma are (jointly) considered involved in the expression of aggression and rage (Kall 1990). In a similar way, Frontalis, Palpabraeous Superior and Inferior, Labii Superioris, Orbicularis Oculi, Masseter, Triangularis, Corrugator and Buccinator are believed to play a major role expressing fear and sadness (Kall 1990). Similar musculo-physiological representations of the emotional experiences were reported in (Dimberg 1990; Dimberg et al. 2000; Dimberg and Petterson 2000; Hess et al. 1992; Lundqvist 1995; Vrana 1993).

3.3 Emotion-specific thermal variations in the human body

It may be deduced from the cited literature that humans involuntarily react to emotions. The emotion specific musculo-physiological activities can be measured and recorded using measurement techniques such as EMG. Since the musculo-physiological activities are believed to cause generation of some musculo-thermal cues, one may infer that a change in affective state would also cause a change in the body temperature. Some evidence of emotion-specific variations in the body temperature, drawn from the multidisciplinary literature, is being presented in the following paragraphs.

Human body temperature is considered a useful physiological signal (Fujimasa 1998). Typical core body temperature of a healthy person, under normal conditions, ranges between 35.5 $^{\circ}$C in the morning to 37.7 $^{\circ}$C in the evening (Jones 1998). Human are *homotherms* and are capable of maintaining a constant body temperature (Jones 1998). An increase (or decrease) in the body temperature may cause the body malfunction and even the failure of body organs (Jones and Plassmann 2002). The “naturally” normal core body temperature helps in preserving the homeostasis. The hypothalamus, a temperature-regulating element in the body has a sensing part that senses any increase in the blood temperature level. The hypothalamus, connected to the pituitary gland balances the generation and loss of heat inside the body and controls the core body temperature. Physically, the pituitary gland resides at the base of the brain close to the termination of the brain stem. The hypothalamus acts as a part of the human nervous system and works as a negative feedback circuit in the body (Jones 1998; Jones and Plassmann 2002). Neurons in the hypothalamus constantly monitor blood temperature against a natural thermal value and act as receptors. It is believed that the internal body temperature setting may change for illness, fever, shock, trauma and anxiety. The hypothalamus tries to regulate body temperature under all circumstances and when it fails to do so, the body experiences some abnormal conditions (Jones and Plassmann 2002).

In addition to other factors, contraction of muscles and a change in the body metabolism help in generating the heat inside the human body (Jones and Plassmann 2002). Blood circulation helps transport the heat from within the body core (Jones 1998). When an increased blood temperature is sensed, the hypothalamus signals to release the body heat. Heat from the human body is released through vasodilatation, perspiration, exhalation, and reduction of the metabolic rate (Jones and Plassmann 2002). A small area in the posterior of hypothalamus detects any blood temperature fall (Jones 1998). When the blood temperature drops, the rate of heat loss is reduced through the initiation of vasconostriction (Jones and Plassmann 2002). The rate of heat transfer between the body core and the surroundings is used to determine the skin surface temperature (Jones and Plassmann 2002).

Core body serves as the source of heat generation. It remains almost at constant temperature under normal conditions. The blood flow through the vessels continuously transports heat to the skin through conduction of heat between the vessels and the skin (Jones 1998; Jones and Plassmann 2002). The skin continuously gains heat and then losses it through thermal radiation, thermal conduction, exhalation, natural convection, forced convection and evaporation (Jones and Plassmann 2002). A change in affective state may also cause a variation in the skin temperature (Asthana and Mandal 1997; Christine and Nasoz 2004; Pham 2000; Vrana 1993).

(Briese 1995) conducted a study to determine if the stress induced by an academic examination would raise the core body temperature and if there was a correlation between the examination-induced stress and the test scores. Their work suggested existence of emotional hyperthermia in 108 students who participated in the investigation.

(Kistler et al. 1998) reported several studies suggesting that a change in skin blood flow would serve as an indicator of sympathetic reflex response to the emotion stimuli. Using the thermal infrared measurements, (Kistler et al. 1998) observed that certain stimuli triggered the sympathetic nervous system and caused a decrease in the cutaneous microcirculation, particularly around the fingertips skin surface.

Relevant literature suggests that emotional experiences may influence the body metabolism and may, consequently, cause variations in the core body temperature. It may be deduced that the body temperature, in addition to the other bio-physiological signals, varies with a change in affective states.

A significant number of studies suggest that animals, birds and humans may experience a considerable (though often momentary) core body temperature change as they confront emotional conditions and situations (Drummond and Lance 1987; Nakayama et al. 2005; Sarlo et al. 2005; Vianna and Carrive 2005). Some recent studies provide convincing evidence of a relationship between the affective states and the skin temperature. These studies suggest that emotions influence body metabolism, physiological conditions and skin temperature (Arkin et al.2003; DeSilva et al. 1997; Pham et al. 2000; Plutchik 1980). A number of researchers have also discovered a direct relationship between the skin temperature and the levels of stress, pain and anxiety (Gavhed et al. 2000). Patterns of observed thermal feature variations in the body tissues are believed to result from and represent the heat transfer to and from the body surface (Jones 1998; Jones and Plassmann 2002). The abnormal thermal patterns observed on human body surface are believed to explain some patho-physiological conditions (Fujimasa 1998; Ogasawara et al. 2001; Prkachin and Mercer 1989). Investigators were able to discover a relationship between the respiratory response and variation in the facial skin temperature (LeBlanc 1976; Stroud 1991). Some researchers were able to employ the relationship between emotions and body skin temperature in HCI, AFEC and AAR (Christine and Nasoz 2004; Ekman 1983; Nakayam et al. 2005; Pollina et al. 2006; Puri et al. 2005; Sarlo et al. 2005; Sugimoto 2000; Yoshitomi 2000).

Several bio-physiological parameters are believed to have an influence on the skin temperature distribution (Fujimasa 1998) but the blood flow rate in the cutaneous tissues remains a major contributor to the variations in the skin thermogram patterns in a neutral environment (Fujimasa et al. 2000; Jones 1998). The vasomotor tone acts on the subcutaneous arterioles and controls the local blood flow. This makes it possible to observe the abnormalities and variations of the nervous system using the skin thermograms (Fujimasa 1998; Fujimasa et al. 2000). Skin thermal characteristics are also believed to reflect changes in the metabolism (Jones 1998), and blood flow rate and blood oxygen level (Fujimasa 1998; Fujimasa et al. 2000). Scientists believe that variations associated with the facial muscular movements may cause changes in the blood flow patterns and result in detectable thermal variation on the face (Fujimasa 1998; Fujimasa et al. 2000; Jones 1998).

3.4 Emotion-specific body heat and temperature flow model

The human skin temperature is determined by the amount of heat dissipated from the body as a result of the blood flow, metabolic function, subcutaneous tissue structure and the sympathetic nervous activities (Bales 1989; Fujimasa et al. 2000). Earlier investigators were able to estimate the amount of heat dissipated from the core body and were successful in estimating the emotion-specific temperature variations on the facial skin. A typical body heat and temperature flow model shown in Figure 3.2 describes the flow of heat from the core body through the human skin. The heat generated inside the human body (Q_{BM}) is supposed to set the core body temperature (T_{body}). In a typical human body heat and temperature flow model, the body temperature (T_{body}) and the core body temperature (T_{core}) are assumed to be equal (Bales 1989; Fujimasa et al. 2000; Jones and Plassmann 2002). Three body heat-flux factors and three heat production factors determine the skin temperature. The three body heat-flux factors are convection heat-flux (Q_{CN}), radiation heat-flux (Q_{RD}), and evaporation heat-flux (Q_{EV}). The body heat production depends on the heat conduction from the core body (Q_{TC}), body metabolism (Q_{TM}) and the amount of heat convection due to blood flow (Q_{BC}). Equation 3-1 exhibits how the thermal equilibrium is achieved on the skin surface under neutral conditions (Bales 1989; Fujimasa et al. 2000).

$$(Q_{CN}) + (Q_{RD}) + (Q_{EV}) = (Q_{TC}) + (Q_{TM}) + (Q_{BC}) \qquad \text{3-1}$$

When thermal infrared imaging is employed to detect emotion-specific skin temperature variations, time-sequential thermal images are analysed to determine the regional skin temperature variations or their associated transient changes in physiological functions (Fujimasa et al. 2000). Equation 3-1 allows comparing the amount of heat produced with the amount of dissipated heat in the time-sequential images. An imbalance between the two sides of Equation 3-1 suggests either heat loss or heat gain in the skin regions under investigation (Fujimasa et al. 2000). Time-trend analysis of temperature variations within the regions of interest in the employed thermograms is also frequently used for blood-flow rate change analysis (Bales 1989; Fujimasa et al. 2000).

3.4.1 Emotion-specific variations in the facial skin temperature

(Zajonc 1985) investigated emotional expressions in humans and observed that facial muscles acted as ligatures on the facial blood vessels and regulated the cerebral blood flow. (Zajonc 1985) concluded that subjective feelings influenced regulation of blood flow in the facial blood facials.

Figure. 3.2: The heat flow and skin temperature equilibrium model.

(Drummond and Lance 1987) discovered evidence of sweating and flushing on the forehead and around the cheeks in response to the body heating, embarrassment and strong gustatory stimulation. They concluded that the gustatory vasodilatation became exaggerated under strong emotional conditions.

(Naemura et al.1993) investigated effect of loud noise on the nasal skin temperature. They report that stress conditions, similar to mother-infant separation, cause the nasal skin temperature to drop. They inferred that changes in the skin temperature might have resulted from the bodily reaction associated with the emotional changes. Using the infrared measurements of 26 participants' facial skin temperature variations caused by the loud noise, (Naemura et al. 1993) concluded that facial skin temperature variation provided promising indices for detecting the emotional changes.

Studies suggest that a change in affective state may cause variation in the blood volume flow under the facial skin. Furthermore, it is argued that the facial expression of emotion results in musculo-thermal activities on the face. The blood volume flow variations and the musculo-thermal changes are believed to cause variations in the facial skin temperature (Bales 1989; Dimberg 1990a; Dimberg 1990b; Otsuka et al. 2002). Since the facial expressions change rapidly, the effect of ambient temperature on the facial skin temperature may be ignored. Hence any imbalance observed between the two sides of Equation 3-1 may be attributed to the facial skin temperature gain or loss due to a change in the facial expression of affect.

Assuming C_{skin} is the heat capacity of the facial skin, the facial skin temperature change (ΔT_{skin}) observed over a short time period (Δt) is expressed as:

$$C_{skin}\ \Delta T_{skin} = (Q_{TC}) + (Q_{TM}) + (Q_{BC}) - [(Q_{CN}) + (Q_{RD}) + (Q_{EV})]. \qquad 3\text{-}2$$

Equation 3-2 allows calculating the skin temperature changes over a short time period due to a change in the expression of affective states. Two thermograms, each recorded with a different facial expression may be subtracted to determine the facial skin temperature change on regions of interest in the thermograms (Bales 1989; Fujimasa et al. 2000).

3.5 Skin temperature measurement methods

Ancient physicians used to assess their patients' physical conditions by measuring the body temperature, just by touching them with bare hands (Ring 1998). Successful quantitative measurement of body temperature was made possible in early 18th century (Ring 1998). Several advanced and easy to use methods and tools of recording the body temperature are available now. Body temperature is measured using mercury in glass, sterile thermocouples, radiometers and liquid crystal (Ring 1998). These temperature measurement systems are inexpensive, accurate and easy to use. However, they require direct contact with the body for temperature measurement (Ring 1998). The only method of non-contact body temperature measurement is the detection and quantification of the natural radiation. The radiation measurement technique provides the basis of modern thermal infrared imaging (Ring 1998).

The human body surface, because of its natural composition and structure, is an efficient radiator. It is therefore easy to observe and measure any infrared emission from the skin surface using some well known non-invasive radiation detection methods (Ring 1998). Sophisticated and inexpensive infrared cameras are widely used to investigate human physiological conditions through analysis of patterns of skin temperature variations (Fujimasa 1998).

Infrared imaging is usually performed under a controlled and comfortable environment, usually referred to as the neutral environment (Fujimasa 1998). A neutral environment allows human body to maintain a state of thermal equilibrium (Fujimasa 1998). Generally, the body temperature in a neutral environment varies between 29-31 ºC with light clothing and 25-29 ºC without clothing (Fujimasa 1998). Heat production and losses in a neutral environment are believed to be almost equal. Only the blood flow rate of cutaneous tissues may cause some change in the skin surface temperature in a neutral environment (Fujimasa 1998; Fujimasa et al. 2000).

Since this work proposes use of infrared measurement of skin temperature for achieving the AFEC functionality, it may be useful to examine the available TIRI methods and systems and establish their suitability for AFEC and AAR.

3.6 Thermal Infrared Imaging (TIRI)

TIRI is an old, established and reliable method that has a long history of military and non-military applications (Paul and Lupo 2002). The ever first scientific demonstration of infrared radiation existence by William Herschel dates back to the year 1800 (Phillips 2002). TIRI was previously limited to military applications. Governmental control and regulations were major obstacles in the development of the non-military TIRI systems (Ring 1998). The research and development work in the field of TIRI was focused on military use and was prohibitively expensive (Paul and Lupo 2002; Sayette et al. 2001). A window of opportunity first opened in 1950s when the principles of infrared imaging were declassified and civilian scientists got access to the technical information. Since then TIRI is gradually emerging as a novel civilian technology (Ring 1998).

Early success of non-military TIRI systems seemed to have inspired scientists to explore the possibilities of using TIRI in other disciplines. This new breed of TIRI systems, referred to as “the third generation TIRI systems,” is used in clinical investigations, remote sensing, medical sciences, engineering maintenance and non-destructive testing of materials (Paul and Lupo 2002; Sayette et al. 2001). TIRI is now considered an affordable technology that is easy to acquire, learn and use. What used to be the ugly and bulky TIRI equipment is now available as trendy, aesthetically pleasing and lightweight equipment (Phillips 2002). Some latest TIRI cameras resemble the operation and appearance of the consumer quality digital video cameras.

Thermal radiation, the basis of thermal imaging, is different from the other two modes of heat transfer; conduction and convection. Thermal radiation propagates through the vacuum and is similar to light in behaviour. Both light and radiation take place in the electromagnetic spectrum, both are photonic phenomena and both travel at the same speed. The energy radiated from a surface is proportional to the 4th power of its absolute temperature. The radiant thermal energy transfer that takes place between the two surfaces is proportional to the 3rd power of the temperature difference between the two surfaces. These characteristics of thermal radiation help in distinguishing it from the other two modes of heat transfer (Cantronic Inc. 2002).

Thermal radiation leaving a surface is called the radiant existence or radiosity. It includes three components, emission from the surface (W_e), reflection off the surface (W_r), and transmission from the surface (W_t). The total radiosity, as shown in Figure 3.3, is the sum of these three components (Wilson and Buffa 1990) and is described as,

$$\text{Total radiosity} = W_e + W_r + W_t \qquad 3\text{-}3$$

Thermal radiation measurements from a target surface are used for non-contact surface temperature measurement and thermography (Wilson and Buffa 1990). Location of the infrared measurement region in the electromagnetic spectrum is exhibited in

Figure 3.4. Two physical laws (the Stephan-Boltzmann law and the Wien's displacement law) are used to define and measure the emission of infrared energy from a surface (Wilson and Buffa 1990).

The Stephan-Boltzmann law explains the radiation of heat from a surface as (Wilson and Buffa 1990):

$$W = \varepsilon \sigma T^4 \qquad 3\text{-}4$$

where W is the radiant flux emitted per unit area of the surface in joules per second (or Watts), ε is the Emissivity, σ is the Stephan-Boltzmann constant with a value of $5.67*10^{-8}$ watts cm^{-2} K^{-4}, and T is the absolute temperature of the target surface in *ºK*.

Figure 3.3: Total radiation and radiosity of a surface

Figure 3.4: Infrared measurement and visible spectrum regions in the electromagnetic spectrum

The Wien's displacement law describes the peak wavelength at which a surface radiates the energy as

$$\lambda_m = B/T \qquad 3\text{-}5$$

where, λ_m denotes the wavelength of maximum radiation in meters (μm), B is the Wien's displacement constant measured in meters-^oK (Wien's displacement constant has a value of 2897 μm K), and T represents the surface temperature in ^oK.

The real surfaces are categorised into three broad categories, blackbodies, greybodies and non-greybodies. This classification is based on their power to radiate the heat (Cantronic Inc. 2001). A blackbody is the imaginary and theoretical surface with a high emissivity ratio of 1.00. It can be measured at all wavelengths as it absorbs all the radiant energy impinged upon it. A body whose surface properties are independent of wavelength is said to be a greybody. The emissivity of a greybody is between 0.00 and 1.00 (Cantronic Inc. 2001). Emissivity, usually denoted by ε is defined as the ratio of the radiant energy emitted from a surface to the energy emitted from a blackbody surface at the same temperature (Cantronic Inc. 2001; Wilson and Buffa 1990). The non-grey bodies do not have a fixed emissivity ratio as their emissivity changes with the wavelength (Wilson and Buffa 1990).

Human skin behaves like a blackbody. It has a high emissivity value, close to 1.00 and is therefore regarded as the blackbody (Otsuka et al. 2002). The emissivity of human skin (ε_s) was independently observed in various studies and was found to be in the range of 0.95 to 0.99. In case of a greybody, the emission, reflection and transmission are constant for all the wavelengths within a particular waveband. A greybody therefore neither absorbs nor reflects all radiations impinged upon it (Jones et al. 1988; Otsuka et al. 2002). In the thermal infrared measurement systems, referring to Equation 3-3, only emission from the surface (W_e) is related to the temperature of a target surface. The temperature is calculated by eliminating or compensating the other two components W_r and W_t (Cantronic Inc. 2001).

Infrared radiation from a target surface needs to pass through some transmission medium to reach the infrared lenses. When a perfect vacuum is available, no energy is lost. The available medium is not considered a perfect vacuum as radiations pass through the atmospheric air in usual circumstances. If radiation takes place under the normal conditions, effects of the atmospheric gases are negligible for short distances.

The energy losses may cause some errors in reading if the distance between the target surface and the camera is very large. However, the two highly used spectral intervals, 3-5 μm and 8-14 μm are relatively less prone to energy losses caused by the distance between the camera and the surface (Jones et al. 1988). Almost all infrared cameras operate within these two spectral intervals. The Cantronic infrared 860 camera used for this research operates within 8-14 μm spectral interval and is less prone to errors due to energy losses (Cantronic Inc. 2001).

3.7 Interpreting skin surface temperature from the infrared thermal images

Human skin is described as a complex structure in the literature (Jones 1998; Otsuka et al. 2002). It comprises of the outer epidermis (the outer most layer of skin) and an approximately 1-2 mm thick layer of the epithelial cells under the epidermis that constitutes the inner layer of the skin. The dermis is a thick layer of dense connective tissues that contains the blood and lymph vessels, hair follicles and glands. Figure 3.5 exhibits a cross-sectional view of the skin segment.

Thermal radiation from the epidermis is easy to monitor since it has a high absorption coefficient, 2.5 to 3.0 mm^{-1}at wavelengths between 2.2 and 5.0 μm (Jones 1998; Otsuka et al. 2002). Since the human skin radiates peak infrared signals around 10 μm of the electromagnetic radiation spectrum, the 8-14 μm bandwidth is widely used in infrared imaging of skin (Bales 1998). This makes infrared imaging an effective technique for converting the electromagnetic radiations emanating from a skin surface into a visible image. The temperature distribution patterns on the facial skin may therefore be observed using methods such as pixel-grey level analysis (Fujimasa et al. 2000, Wolff et al. 2005).

Measuring the light in any part of electromagnetic spectrum is referred to as radiometry but the term generally refers to the measurement of infrared, visible and ultraviolet radiations (Ashdown 1994; Otsuka et al. 2002). Infrared imaging works on the principles of radiometry and photometry. Light is radiant energy, denoted as Q and measured in Joules.

Figure 3.5: A Cross-sectional view of the human skin

When absorbed by a physical object, it can be converted into another form of energy. Spectral radiant energy is the amount of radiant energy per unit wavelength interval at a wavelength λ. It is measured in joules per nanometer as $Q_\lambda = dQ/d\lambda$. The time rate of flow of radiant energy, known as the radiant flux ϕ_e $(=dQ/dt)$ is measured in watts and is integrated over all instances of wavelengths as (Ashdown 1994; Otsuka et al. 2002)

$$\phi_e = \int_0^\infty \phi_{e\lambda} d\lambda \,. \qquad 3\text{-}6$$

Spectral radiant flux $\phi_{e\lambda}$ is the power emitted or received as radiation per wavelength interval at a wavelength λ and is measured as

$$\phi_{e\lambda} = \phi_e / d\lambda \,. \qquad 3\text{-}7$$

Radiance is the infinitesimal amount of radiant flux contained in a ray of light arriving at or leaving a point on a surface in any given direction. A ray is conceptually considered an infinitesimally narrow cone having its apex at a point on a surface. This cone is assumed to have a differential solid angle $d\omega$, measured in sterdiance (Ashdown 1994; Otsuka et al. 2002).

Luminous flux, the photometrically weighted radiant flux, is measured in lumens as 1/683 watts of radiant power at a frequency of 540×10^{12} Hertz. It is defined as (Ashdown 1994; Otsuka et al. 2002).

$$\phi_v = 683 \int_{380}^{780} \phi_{e\lambda} V(\lambda) d\lambda \,. \qquad 3\text{-}8$$

The power of radiation to produce visual sensation is represented as luminous efficacy. A human observer sees radiation between the wavelengths of 380-780 nanometers. The maximum luminous efficacy of any radiation is 683 lumens per watt at 555 nanometers (Ashdown 1994; Otsuka et al. 2002).

Human eyes perceive the luminance, an approximate measure of the brightness of the surface being viewed from a particular direction, measured in lumens per square meter per sterdian. Luminance is equivalent to the photometric measure of radiance, related to the sensation of brightness (Ashdown 1994; Otsuka et al. 2002).

The infrared thermal cameras measure the radiance for image construction. The principles of radiometry suggest that distance between a light source and sink does not influence the radiance in the absence of scattering and absorption. Hence radiance of a bundle of rays remains constant when it moves across an optical space. Therefore, the radiance of a source remains same as that of an image. Infrared cameras use focal plane array of detector elements to capture images (Ashdown 1994; Otsuka et al. 2002; Davis and Lettington 1988).

Modern thermal infrared cameras are typically equipped with the Germanium lens to focus thermal radiations onto a focal plane array of microbolometer detectors. A bolometer is a temperature-sensitive electrical resistor. The microbolometer detectors employ a monolithic pixel structure. The arrays of microbolometer elements in such infrared cameras are thermally isolated to prevent thermal losses and reduce the possibilities of adding noise. An external electronic circuit measures its temperature rise caused by the absorption of the incident radiant energy (Ashdown 1994; Bales 1998; Davis and Lettington 1988; Kurse 2001; Otsuka et al. 2002).

When a face is focused on the microbolometer detectors, each pixel undergoes a temperature increase and generates a signal that depends on the irradiance falling on it. Since the irradiance is a product of the radiance of the facial skin and the solid angle subtended by exit pupil at the image, the pixel grey-level depends on the radiant flux per unit area and the detector quantum efficiency. The detector quantum efficiency depends on the detector element area, detector absorption coefficient and its conversion efficiency (Kurse 2001).

When the radiant energy is focused on the microbolometer array, its temperature increases, causes a change in resistance, and allows detecting the radiant power. Assuming R is the resistance and W is the dissipated power, the rate of change in resistance *dR/dW* is proportional to the rate of change in electrical resistance *dR/dT* with the temperature *T* (Ashdown 1994; Bales 1998; Davis and Lettington 1988; Kurse 2001; Otsuka et al. 2002).

Assuming *R** is the intrinsic detector responsivity and *I** is the bias current, the thermal impedance *Z** may be described as

$R^* = I^* (dR/dW) = I^*(dR/dT)|Z^*|$. 3-9

The bolometric detector resistance nonlinearly varies with the facial skin temperature. The skin area focused onto a pixel determines the total radiant flux falling onto the pixel. The grey-level may therefore be interpreted as the average skin temperature.

During the infrared imaging of human face, the radiation arises from the natural thermal radiation from the face in the scene. Equation 3-4 transforms into equation 3-10 and provides an estimate of the heat radiated from a facial skin surface. It is governed by the Stephen-Boltzmann law and is denoted as

$\phi_e = \varepsilon_s \sigma T_s$. 3-10

Where, ϕ_e is the radiant flux per unit area of facial skin, ε_s is the skin emissivity, the proportional amount of energy emission with respect to a perfect absorber, σ is the Stephan-Boltzmann constant (=5.673x10^{-12}) and T_s is the absolute temperature of facial skin in ^oK (Ashdown 1994; Bales 1998; Davis and Lettington 1988; Kurse 2001; Otsuka et al. 2002). The Wien's displacement law describes the peak wavelength at which the facial skin radiates. It is estimated using Equation 3-5.

$\lambda_m = B / T_s$. 3-11

Where λ_m is the wavelength of maximum radiation in μm and B is the Wien's displacement constant (=2897 $\mu m\ ^oK$).

From the infrared radiation point of view, facial skin has a very high emissivity (ε_s), much higher than the common surroundings such as glass or concrete and closer to a perfect blackbody. Studies suggest that the skin emissivity does not vary much with the wavelength and remains almost constant (Bales 1998; Jones and Plassmann 2002; Sloan et al. 2002).

The pixel grey-levels in an infrared image therefore provide a measure of the response of the microbolometer array to the radiant power it absorbs, integrated over all angles. Changes in the grey-level may reflect, theoretically, the changes in facial skin temperature. For this reason radiance, an exponential function of the skin temperature is considered an indicator of the level of blood perfusion in the skin in (Fujimasa et al. 2000; Otsuka et al. 2002; Pavlidis 2004). The pixel grey-level may provide a measure of skin surface radiance and is also used to estimate the skin surface temperature. In a typical thermogram analysis, a grey-level of zero corresponds to the minimum temperature and the maximum grey-level corresponds to the maximum facial skin temperature. The infrared signals measured by the arrays of microbolometer detectors are sent to a frame grabber for image construction, image processing and viewing in a way that the map of 3-D temperature distribution of a face is converted into a 2-D image (Ashdown 1994; Bales 1998; Davis and Lettington 1988; Kurse 2001; Otsuka et al. 2002; Wolff et al. 2005). Figure 3.6 exhibits a typical thermal imaging system. It is important to note that each infrared system manufacturer reportedly uses its own proprietary standards (Paul and Lupo 2002). Figure 3.6 therefore provides a generic high-level representation of a typical infrared imaging system.

3.8 Infrared imaging application in automated affect recognition

TIRI was used in some recent investigations for classifying the human emotions and affective states, typically for a binary classification of affective states. Some attempts were also made to analyse the facial expressions using TIRI, in combination with other cues (Kim et al. 2004; Nakayam et al. 2005; Pavlidis 2004; Pavlidis and Levine 2002; Pollina et al. 2006; Puri et al. 2005; Sugimoto et al. 2000; Yoshitomi et al. 2000). Rationale and motivation for these investigations arise from the fact that certain real life situations and emotional conditions cause a change in blood volume flow under the facial skin (McGimpsey et al. 2000; Ogasawara et al. 2001; Phillips 2002).

In a recent investigation, thermal facial screening was employed to detect attempted deceit using a three-stage system (Pavlidis 2004). During the first stage, thermal images were acquired using an infrared camera. Acquired thermal images were used to transform the facial thermal data into a blood flow model in the second stage. The hæmodynamic model was built upon the premise that significant blood flow redistribution would be taking place vis-à-vis a change in emotional conditions and anxiety. During the third stage, the hæmodynamic model was used to classify people into deceptive or non-deceptive categories. The system reportedly achieved results compatible with the polygraph examination by human experts (Pavlidis 2004).

Figure. 3.6: Schematic representation of a typical thermal imaging system.

(Sugimoto et al. 2000) used TIRI to detect the transitions of emotional states by synthesising certain facial expressions. Facial thermal changes caused by the muscular movements were analysed for the purpose. The system compared a neutral expression face with the test face, and geometrically reformed them to develop a thermal differential model. Results suggested it was possible to detect facial temperature changes caused by transition of emotional states and their associated physiological changes. Results further suggested that detected facial temperature changes could help understand transition of emotional states. The resulting system successfully detected facial temperature changes caused by pleasure or tiredness (Sugimoto et al. 2000).

(Yoshitomi et al. 2000) employed a combination of visual images, thermal features and audio signals for recognising affective states. They examined the possibility of classifying the neutral, happy, sad, angry and surprised faces through the integration of visual, thermal and audio signals. Sound signals were employed to train a Hidden Markov Model (HMM). The visual and thermal features were fused together for training a neural network. The two output values (coming out of the HMM and the neural network) were fused together to recognise the emotive states. This integrated signal classifier performed with 85% accuracy (Yoshitomi et al. 2000).

(Kim et al. 2004) developed a physiological signal based emotion recognition system. The system fused the electrodermal activity data, electrocardiogram readings and the skin temperature measurements together to determine the emotive states. Like several other emotion recognition systems, this system also recognised emotions in three stages. First, the signals were pre-processed, then features were extracted for pattern analysis and finally, emotions were classified. The support vector machine (SVM) based classifier developed in this work reportedly achieved more than 78 % classification accuracy (Kim et al. 2004).

(Puri et al. 2005) reported development of a system for non-contact measurement of computer users’ emotional states through TIRI. They reported a correlation between the users’ stress level and an increased blood flow in the frontal vessels of forehead, probably caused by the dissipation of convective heat. Their system monitored dissipation of convective heat though infrared imaging and helped in identifying users’ emotional states. They discovered that thermal infrared measurements were highly correlated with the real time measurement of energy expenditure (Puri et al. 2005).

Studies reported in this section motivate using thermal variation patterns, measured on the face, for developing an AFEC and AAR capable system. Getting encouraged by the previous investigations, this thesis attempts to use facial skin temperature measurements for achieving the AFEC and AAR functionality.

3.9 Conclusion

Like other muscles of a human body, facial skeletal muscles contract and perform some work in order to bring changes in the facial features. Studies suggest that while on work, the facial muscles are physically active and produce heat for maintaining the body temperature (Netter and Hansen 2002; Starr et al. 2003). Using the facial EMG readings, scientists were able to discover an association between the muscular movements, muscle energy expenditure and the facial expressions of affective states (Allanson and Fairclough 2004; Cacioppo et al. 1990).

The increased blood volume flow under an area of facial skin (as the result of stress) is termed as reactive hyperemia (Ogasawara et al. 2001). Reactive hyperemia includes situations such as mechanical insult to the skin, chemical reactions causing vasodilatation of blood capillaries and thermal stress (like cold water immersion). Infrared imaging is used to diagnose, monitor and quantify hyperemia effects and quantify the dynamic stress on the skin (McGimpsey et al. 2000; Ogasawara et al. 2001; Phillips 2002).

Studies suggest that facial muscles either contract or expand when the facial expressions change (Pessa et al. 1998). Muscular contraction and expansion are believed to cause some fluctuations in the rate and volume of blood flow under the facial skin. A change in the emotional experience is also believed to influence the blood flow rate under the facial skin (Jones 1998; Jones and Plassmann 2002; Otsuka et al. 2002).

An accurate and representative model of estimating the relationship between the changes in facial expression, fluctuation in blood flow rates, contraction and / or expansion of facial muscles and variation in the facial skin temperature is not available yet. Such a model, if developed, will enhance our understanding of the relationship between facial expressions and the facial thermal and physiological characteristics. In the absence of such a model, many approaches are employed for detecting, extracting and interpreting facial expressions. Thermal infrared imaging is one such approach employed for recognition and classification of expressions.

Investigators have been able to successfully discover and analyse the skin thermal variations associated with the positive and negative emotive states using thermal infrared imaging. However, earlier investigators had employed the facial thermal features only for a binary classification of facial expression of affective states. The work in this thesis, for the first time, attempts to classify the most common facial expressions of affective states using the facial thermal features in a non-dichotomous manner.

The scientific information cited in the preceding sections of this chapter encouraged examining the possibilities of using the infrared facial skin temperature measurements for achieving the AFEC and AAR functionality. Details of the algorithmic methods used for interpreting pixel grey-level information to measure facial skin temperature using the TIRIs are presented in the following chapters. Facial features extraction, selection, representation and classification approaches employed in some previous investigations are also reported in the following chapters.

Chapter 4

INITIAL EXPLORATION OF THE PROPOSED AFEC APPROACH

The last two chapters provided scientific evidence and rationale for using the facial skin temperature measurements in AFEC and AAR. This chapter presents the findings from the first exploration of the capture and analysis of thermal facial images. It begins by describing the instruments used and introducing the participants who volunteered for this investigation. Ethical conduct and human protection practices observed during the image acquisition process are briefly reported in this chapter and are further explained in Appendix I. It then introduces the basic image processing needed to remove noise and locate sensing points in thermal infrared images. Pixel grey-level interpretation of infrared images performed to extract the facial thermal features from within the acquired thermal images, led to the discovery of the Facial thermal Feature Points (FTFPs) along the major facial muscles. Thermal Intensity Values (TIVs) recorded at these FTFP sites were analysed to examine if the data were suitable for invoking the multivariate analyses and pattern analyses algorithms.

4.1 Thermal infrared image acquisition

A Cantronic model IR860 thermal infrared imaging camera was used for thermal imaging. The camera was equipped with a storage disk, special-purpose image recording accessories, and the camera-to-computer data transfer peripherals. The IR 860 camera saves up to 62 thermal images on a type III PCMCIA (Personal Computer Memory Card International Association) standard card. It digitises the image information and facilitates communicating the digital information to a personal computer. The infrared camera and its accessories are shown in Figure 4.1. Two closer views of the camera are also shown in Figure 4.2. The camera has a full screen radiometric capability that allows measuring the surface temperature from a target surface. The camera captures an image array of 320 × 240 pixels using an uncooled device (temperature sensitive electrical resistor, called the microbolometer) for measuring the incident radiation. It uses a Focal Plane Array (FPA) detector system that employs several uncooled microbolometer FPA detectors for thermal imaging. The detector has a high thermal sensitivity in the light spectrum wavelength range of 7.50-14.00 μm. The camera is supplied with the proprietary thermal image analysis software, CMView Plus.

Figure 4.1: The IR 860 infrared camera with its accessories

Figure 4.2: Two closer views of the IR 860 infrared camera

The IR 860 camera has a thermal sensitivity of 0.08 °C at 30 °C with an accuracy of ± 2 °C under normal temperature and pressure (Cantronic 2002). It allows adjusting the radiating surface emissivity to ensure an accurate temperature reading from the target surface for analysing the thermal images. The CMView Plus image analysis software that works only under Windows 95 and 98 operating systems was used for image processing in this work.

A DELL Optiplex GX 110 personal computer, fitted with a 500 MHz processor, 128 MB DRAM, and a 40 GB HDD attached to a DELL Multisync high-resolution monitor, was used in this work. The system was equipped with a CD burner and a HP Scanjet 3400C flatbed scanner. Another external 40 GB hard disk was attached with the computer and was used as additional (external) memory during the thermal image processing.

4.2 Experiment design

In order to examine the viability of using the transient facial thermal features in AFEC and AAR, a database of sample visible-spectrum and infrared images was developed. Experimental work for this investigation, involving development of image-database, was carried out in the School of Architecture and Design building at the American University of Sharjah, United Arab Emirates. The infrared images were acquired under a normal, controlled and comfortable building environment. The internal room temperature varied between 19-22 ºC during the thermal infrared image acquisition. The building air conditioning system was equipped with a humidity controller and an air recycling system. Each participant was given at least 20 minutes to acclimatise with the environment. Thermal images were captured in several independent sessions in October 2003, November 2003, April 2004, and September 2004. Images were recorded in the afternoons between 0100 and 0430 PM. A low emissivity background was used to ensure better separation of the background from the desired regions of the TIRI during the image processing (Jones and Plassmann 2002; Otsuka et al. 2002). Therefore, a concrete wall background having low emissivity (ε = 0.54) was selected during the image acquisition process.

The IR 860 thermal infrared camera was set to measure the facial skin temperature range between 0-40 ºC. The skin surface emissivity (ε) was set between 0.97 and 0.99.

In order to accurately capture the frontal view of a participant, the central vertical line on the camera viewfinder was aligned to the center of each participant face. A visible image camera was placed (on the left side) next to the thermal infrared camera. Two volunteers independently operated the two cameras. The camera operators used a visual signal to push the image capture button at the same time. The process was rehearsed to minimise the delay time between the two image shots. Each facial expression image was captured at least twice. Three referees, this author and the two staff members who helped in acquiring the images, selected those pictures from the captured pictures that best described the facial expressions and were most clear. Thermal infrared images corresponding to the selected visible images were used for discovering the temporal thermal features from within the thermal images. After the neutral faces were captured, each participant pretended and expressed happiness, sadness, disgust, surprise, anger and fear.

4.3 Participants and ethical issues

Initially, 16 adult undergraduate students, 12 boys and 4 girls, with a mean self-reported age of 20 years 9 months volunteered for the research experiments. The participants included Africans, Caucasians, Arabs, Iranians, Indians and Pakistanis. All participants allowed use of their visible-spectrum and infrared images in scholarly publications for dissemination of the research information.

At the beginning of each image acquisition session, participants were briefed about the objectives of these experiments, methods, procedures, potential benefits and the probable outcome of the experiments. Ethical experiment design practices and protocols were also explained to the participants at the beginning of each image acquisition session and were observed during the experiments. Participants were given a choice to leave the experiments at any stage. Ethical principles and guidelines for the protection of human participants of biomedical and behavioural research provided in the Belmont report (Belmont Report 1979; DHEW1979) were observed during the experiments. The code of ethics for conducting Psychological research, developed by the Australian Psychological Society was also referred and followed during the design of experiments (The Australian Psychological Society 2003). As outlined in the Belmont report, efforts were made to protect the participants from any physical and/or emotional harm and damage. The human protection practices observed in this investigation are reported in Appendix I.

4.4 Thermal patterns of affective states

Figure 4.3 shows a participant's visible spectrum images, thermal infrared images and their corresponding thermograms. Each thermogram shows the temperature frequency distribution on each pixel in a thermal image. It can be seen in Figure 4.3 that a change in facial expression causes some changes in the thermal characteristics of the face. The varying temperature frequency distributions in the acquired thermal images are obvious in the thermograms shown in Figure 4.3. Each thermogram provides a detailed account of the observed temperature frequency distribution on the participant face. The instances of temperature readings (thermal values) are expressed as percentages in the thermograms. Actual temperature readings are shown along the *x*-axis. The frequency of observing a particular temperature reading is shown along the *y*-axis.

4.5 Thermal infrared image processing

Pixel grey-levels extracted from a thermal image provide a measure of the response of the detector element (such as the microbolometer array installed on the IR 860 thermal camera) to the infrared radiant power absorbed. The radiant power falling on the detector element is considered a function of the radiance of the surface and the solid angle subtended by the exit pupil of the thermal camera. The solid angle, by the virtue of camera design, remains constant thus allowing the grey-levels to change with any changes in the radiance of the object surface (Jones and Plassmann 2002; Otsuka et al. 2002).

Infrared cameras operate in a temperature range of -20 to 500 ºC. Depending upon the bit-depth, between 4096-16384 grey levels of pixels represent the extreme temperature points in a thermogram. A grey-level of zero corresponds to the lowest temperature and the highest level of grey corresponds to the highest temperature in a TIRI (Jones and Plassmann 2002). The camera temperature range was set between 0-40 ºC during the TIRI acquisition hence the grey-level of 0 corresponded to 0 ºC and the highest grey-level corresponded to the highest facial skin temperature observed in an acquired facial TIRI.

Figure 4.3: Visible-spectrum images, infrared images and the corresponding thermograms of a participant are shown in each row. Top row: neutral face and faces with the intentional expressions of happiness; Middle row: sadness and disgust; Bottom row: fear.

The built-in radiance measurement and image digitisation mechanisms in a thermal camera cause addition of undesired noise in the TIRIs. Sources of noise in the TIRIs include skin surface states and drifting of detection element temperature (caused by the fluctuations in the heat exchange).

Many convolution methods are recommended for noise reduction and edge detection to minimise the influence of noise factors in the TIRIs (Jones and Plassmann 2002; Otsuka et al. 2002; Pavlidis and Levine 2002).

To avoid any undesired noise in the TIRIs and to have most accurate thermal measurements, the thermal images are usually enhanced before extraction of the pixel grey level information. Facial infrared images were therefore processed to reduce the built-in noise using CMView Plus, a thermal analysis software. The “median smoothing filter” recognised as one of the best order-statistic filter (Gonzalez and Woods 2002), was invoked on the thermal images for blurring and noise reduction. The median smoothing filter, applies a non-liner solution approach for recovering the original image signals and results in excellent noise reduction with a minimal blurring (Gonzalez and Woods 2002). The filter replaces value of a pixel by the median of the grey levels in the neighborhood of the pixel as

$$\hat{f}(p,q) = \underset{(s,t)\in S_{pq}}{median}\{k(s,t)\} \qquad 4\text{-}1$$

where, $\hat{f}(p,q)$ is the median filter that replaces the value of a pixel (s, t) by the median of the grey levels within a neighborhood. Median smoothing assumes that a small number of corrupted image signals in a grey pixels image would randomly take on a value of white or black. Depending on the density of the prevailing noise in a grey-level thermogram, a median filter computes the removable noise over an appropriate pixel-neighborhood. The tradeoff is that a larger neighborhood leads to a loss of detail, whereas a small neighborhood results in a loss of signal quality (Acharya and Ajoy 2005). Since the acquired thermal images were of good quality, a small neighborhood filter was considered appropriate for enhancing the thermal images.

In a following image enhancement step, the Sobel operator-based edge detection algorithm was invoked for extracting the contours within the infrared images. The Sobel operator is basically a neighborhood-based gradient operator. Two convolution masks for the Sobel operator are separately applied on the input facial image to yield the two gradient components G_s and G_t in the horizontal and vertical directions. The neighborhood kernels define the convolution masks. For the selected 3x3 neighborhood the gradient operators were calculated as (Acharya and Ajoy 2005; Gonzalez and Woods 2002),

$$G_s=[f(i-1, j-1)+2f(i-1, j)+ f(i-1, j+1)] - [f(i+1, j-1)+2f(i+1, j)+ f(i+1, j+1)], \quad 4\text{-}2$$

and

$$G_t=[f(i-1, j-1)+2f(I, j-1)+ f(i+1, j-1)] - [f(i-1, j+1)+2f(I, j+1)+ f(i+1, j+1)]. \quad 4\text{-}3$$

The gradient magnitude was computed as

$$G[f(s,t)] = \sqrt{G_s^2 + G_t^2}. \quad 4\text{-}4$$

4.6 Thermally significant Facial Thermal Feature Points

(Jones and Plassmann 2002) have suggested a method to analyse a series of TIRIs that were separated by a small amount of time (also referred to as time-sequential images in the literature). They compared the temperature measurements at the points of registration within a series of images to discover the temporal changes in the temperature distributions (Jones and Plassmann 2002). Some investigators have used a different approach that would discover and examine the left-right symmetric regions of interest in thermal infrared images to extract the temporal facial features (Jones and Plassmann 2002; Otsuka et al. 2002; Pavlidis and Levine 2002).

This phase of the work began with identification of the left-right symmetric regions within each individual’s seven thermal images to examine the temporal thermal differences in the TIRIs. The CMView Plus software allows automated discovery of the contra-lateral symmetric regions within the TIRIs. The symmetric regions discovered within a neutral face TIRI and within the TIRIs showing a positive or negative facial expression were found to be inconsistent for different participants and different expressions. For some participants, the temperature measurements taken within the symmetric regions inside the neutral face image and the images with pretended facial expressions were almost the same. For others, the temperature measurements taken within the symmetric regions within different TIRIs had different values. It was concluded that the facial thermal symmetries observed in the captured images might not help in distinguishing between the facial expressions of affective states.

In a following step, the sequential subtraction of a series of TIRIs was attempted. Some manually selected, nearly equal symmetric facial regions within the TIRIs were used as the registration points. CMView Plus thermal analysis software has a built-in multi-field temperature measurement option that allows selecting multiple regions of interests (ROI) in a sequence of thermal images and subtract the temperature readings measured at the selected ROI. The multi-field temperature measurement option was invoked to subtract the TIVs at ROI in sequential thermograms. This allowed comparing the TIVs measured at the selected symmetric regions in a series of images. The TIV data gathered from the selected symmetric regions were then statistically analysed. The maximum temperature measurements within the symmetric regions had a high correlation and the variance test statistic was significant ($p<0.01$) suggesting that the TIVs measured at symmetric regions had very little between-facial expression variance. It became obvious from the analysis that temperature intensity values measured at the selected symmetric regions in a series of images would not allow distinguishing between the facial expressions.

Realising the ineffectiveness of the temperature measurements taken within the symmetric regions of the TIRIs, temperature measurements taken at different sets of registration points within the TIRIs were selected to discover any temporal changes in the temperature distributions. Thermal variations at manually selected (multiple) locations on the forehead, around the eyes, on the cheeks and chin were repeatedly analysed. The TIVs were repeatedly measured at different sets of equal points to ensure a minimum correlation among the data and a maximum "between-facial expression" variance. The multivariate analysis and pattern recognition algorithms work better on data that is independent and have little (ideally no) within data correlation and demonstrates a significant variance (Field 2000; Sharma 1996; Rencher 1995). The statistical analyses suggested that the TIV data measured at these manually selected locations were highly correlated and violated the assumption of similarity between-facial expression group variance structure ($p<0.05$).

Obviously, a better feature extraction approach was required to discover the underlying variance in the acquired TIRIs. Earlier researchers have proposed several methods for acquiring thermo-physiological data reflecting any thermo-muscular activities under the body skin. Some studies have suggested that a thermal infrared image with a "neutral face" having all muscles in their natural (and neutral) position would be (thermally) different than the ones that would exhibit the facial expression of an affective state (Garbey et al. 2004; Pavlidis 2000; Sugimoto et al. 2000; Yoshitomi et al. 2000). It was therefore decided to select the facial regions along the major facial muscles as regions of interest for discovering the temporal thermal information within the TIRIs.

In a following step, 16 square segments (each of 36 × 36 pixels) along the facial muscles in each thermogram were selected. The highest TIVs in each of the 16 square segments were recorded and analysed for the two objectives; ($\mathrm{Correlation}_{\min}$ and $\mathrm{Variance}_{\max}$). Observing that the two criteria were not met, the TIRIs were repeatedly divided into an increasing number of square segments along the major facial muscles. The TIRIs were divided into 32 square segments (each of 25 × 25 pixels) and 64 square segments (each of 25 × 25 pixels). Figure 4.4 exhibits the 16, 32 and 64 square segments on a facial thermal map. Each set of resulting TIV data were analysed for correlation and variance. Some sets of the TIVs recorded in the square segments of the individual TIRIs showed significant differences in the thermal intensity values than the others when a change in facial expression occurred.

The process was repeated again and again until significant thermal variations were discovered at 75 physical sites located all over the face along the major facial muscles within the 64 TIRIs. The TIV data gathered from these 75 facial thermal feature points (FTFPs - the square segments of 16 × 16 pixels), also allowed achieving the two objectives; $\mathrm{Correlation}_{\min}$ and $\mathrm{Variance}_{\max}$. Please note that the multi-field temperature measurement option in the CMView Plus thermal analysis software was invoked to compare the TIRIs and to discover the temporal thermal variations at the 75 FTFPs.

Figure 4.5 shows these 75 FTFPs on a neutral human face, exhibits a muscular map of a human face, and represents the geometric profile of the facial thermal feature points. The 75 sites shown in Figure 4.5 showed a consistent and significant variation in the thermal intensity values recorded with a change in facial expression. These significant Facial Thermal Feature Points and their muscular alignments are listed in Table 4.1. It is obvious from Table 4.1 that more than 50 % of these FTFPs are located on the five major facial muscles. The five major facial muscles, Frontalis (16 FTFPs), Orbicularis Oculi Pars Orbital (12 FTFPs), Levator Labii Superioris (6 FTFPs) and Risorious (6 FTFPs) seem to hold 53.33% of the FTFPs on a human face. Figure 3.1 exhibits the frontal view of a facial thermal map.

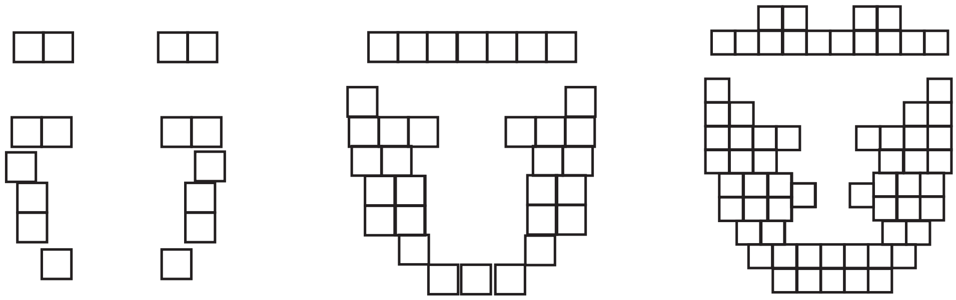

Figure 4.4: Left to right: The highest thermal intensity values were measured within the shown 16, 32 and 64 square segments on the face.

A side view of the major facial muscles is also shown in Figure 4.6. A closer examination of Figures 3.1, 4.5 and 4.6 suggests that the FTFPs are located along the major facial muscles. Previous studies reported in the last two chapters also suggested that these same facial muscles would play a major role in the facial expression of affective states.

The FTFPs were spread all over the face. Of these FTFPs, 10 were found on the forehead, 21 were located around the eyes, 18 were spread on the cheek, 17 were located around mouth, and 9 were located on the chin. Table 4.2 shows the physical location of the FTFPs on the face.

The TIV data recorded at the 75 FTFP sites were used to represent each thermal face as a 75-dimensional thermal feature vector for the subsequent investigation and analyses.

Figure 4.5: Left to right: Geometric profile of FTFPs, FTFPs on a facial muscle map and FTFPs on a human face

TABLE 4.1: FACIAL THERMAL FEATURE POINTS AND THEIR MUSCULAR ALIGNMENT

Facial Muscle	Facial Thermal Feature Point (FTFP)
Frontalis, pars medialis	1, 3, 6, 8, 13, 15
Frontalis, inner center edges of pars medialis and pars lateralis	2, 7
Frontalis, pars lateralis	4, 5, 9, 10, 11, 12, 16, 17
Procerus/ Levator, labii superioris alaquae nasi	21
Depressor, supercilii	14
Orbicularis Oculi, pars orbital	18, 19, 20, 22, 23, 24, 25, 26, 27, 29, 30, 31
Orbicularis Oris	45, 51, 64, 65, 66
Levator, labii superioris alaquae nasi	28, 35, 36
Levator, labii superioris	33, 34, 37, 38, 44, 46
Masseter, superficial	40, 41, 49, 50
Levator, anguli oris	43, 47
Zygomaticus major	32, 39, 42, 48
Risorious/ Platysma	52, 53, 54, 59, 60, 61
Depressor anguli oris	55, 58
Buccinator	56, 57
Platysma	62, 63, 67, 68
Depressor Labii Inferioris	69, 70, 71, 72
Mentalis	73, 74, 75

TABLE 4.2: PHYSICAL LOCATION OF FTFPS ON THE FACE

Part of the face	FTFPs
Forehead	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Around the eyes	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
Cheeks	32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 47, 48, 49, 50, 62, 63
Around mouth	44, 45, 46, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 64, 65, 66
Chin	67, 68, 69, 70, 71, 72, 73, 74, 75

Figure 4.6: A side view of the human face showing the major facial muscles

Figure 4.7: Estimated mean values of the facial skin temperatures for neutral and six basic facial expressions. The curve shows significant thermal variance with a change in the facial expression

Figure 4.7 exhibits the facial expression specific differences in the mean facial skin temperature estimated using the TIV data measured at the 75 FTFPs on the participant faces. It is evident in Figure 4.7 that the TIVs at the FTFPs change with a change in facial expression. These thermal differences at the FTFPs probably resulted form the changes in the blood volume flow and the associated thermo-muscular activities on the participant faces. The observed thermal differences shown in Figure 4.7 were consistent with the results reported in (Nakanishi and Imai-Matsumura, 2007)

4.7 Initial analyses of TIV data

The TIV data were first tested for the assumption of normal distribution. Standard statistical tests suggested that the TIV data within the individual thermal images were normally distributed. Following the successful test of normal distribution of thermal data in the individual TIRIs, infrared images belonging to each particular expression were grouped together and were tested for the normal distribution of the TIV data. Histograms, Q-Q plots, and the analyses of skewness and kurtosis showed that the TIV data on more than 92% FTFP sites were normally distributed. Two other standard tests;

the Kolmogorov-Smirnov test statistic (*$p>0.05$*) and the Shapiro-Wilk test statistic (*$p>0.05$*) were also non-significant and suggested no departure from the assumption of normal distribution.

(Coakes and Steed 1999) and (Sharma 1996) suggested that if the data violate the assumption of sphericity, several undesired biases might surface during the multivariate analysis. Hence the homogeneity of covariance (also referred to as the sphericity in the literature) was tested before invoking the multivariate analysis on the data. The variances in the TIV data were calculated from the estimates of within-facial expression cluster covariances about the cluster means. The non-significant Levene's test statistics [$F(6,10)=2.92$, $p>0.05$)] for the thermal data suggested the homogeneity of variances in the data. It was therefore considered safe to assume that

$\text{Variance}_{(neutral)} \approx \text{Variance}_{(happy)} \approx \text{Variance}_{(sad)} \approx \text{Variance}_{(surprise)} \approx \text{Variance}_{(angry)} \approx \text{Variance}_{(fear)} \approx \text{Variance}(_{disgust)}$

The interaction between the independent variables (facial expressions) and the dependent variables (TIV measurements on the entire face) was significant [$F(296, 4440) = 7.32$, $p<0.01$) suggesting a significant effect of facial expressions on the measured TIV data. It was therefore safe to assume that the data might be used for multivariate analyses and pattern classification (Chatfield Collins1995; Field 2000; Turner and Thayer 2001).

The test statistics pertaining to the 7 facial expressions were encouraging and suggested that the available thermal data were appropriate for invoking the multivariate analysis. Furthermore, the tests suggested that relevant pattern recognition algorithms could be invoked on the acquired TIV data (Everitt and Dunn 1991; Field 2000; Kinnear and Gray 2000; McLachlan 2004).

4.8 Conclusion

A set of appropriate noise reduction and edge detection methods was identified and invoked to enhance the quality of acquired thermal images and extract the most effective facial thermal features from the TIRIs. The TIV data were collected from equal points on the faces within the TIRIs that were acquired at different times when participants pretended facial expression of affective states. The data used in this investigation consisted of the TIRIs of known origin hence each facial expression was considered a separate and independent facial expression group. The thermal images could be clustered together for supervised learning based upon their respective facial expression group memberships. Each thermal image was represented as a feature vector $\mathbf{x}$ having p number of TIV measurements obtained from the grey-level pixel analysis of the TIRIs.

The parametric multivariate analysis of this type of data requires accounting for the statistical parameters such as the overall mean and the overall measure of error in the data (Rencher 1995). Hence, most appropriate statistical tests were identified and invoked for examining the TIV data. The standard test statistics suggested that the TIV measurements taken on the FTFP sites of the participant faces had a multivariate normal distribution with X ~ $N(\mu_i, \Sigma_i)$ and could therefore be treated as Gaussian distribution (Borowski and Borwein 1991; Webb 2002). The relevant standard test statistics also suggested the presence of a similar group covariance structure ($C_1 = C_2 =, \ldots, = C_n$) in the TIV data acquired from the available measurement space.

Chapter 5

FACIAL THERMAL FEATURE EXTRACTION, SELECTION AND CLASSIFICATION

Having measured the emotion-specific facial thermal variations and localised the thermally significant facial thermal feature points (FTFPs) on the face, this work, in its next logical phase, required classification of thermal features into facial expressions of affective states.

This chapter begins by providing an overview of the most common pattern analysis and classification approaches used in the earlier Automated Facial Expression Classification (AFEC) capable systems. The algorithmic approaches proposed for the facial thermal feature selection, representation and classification in this investigation are then presented. Finally, salient features and advantages of the proposed AFEC approach are reported.

In a typical classification problem, the three essential tasks: feature extraction, feature selection, and feature classification need to be carried out in a sequential order (Duda et al. 2001; Fukunaga 1990). These three tasks could be carried out using either a statistical or a neural classifier. However, selecting one of the two classifiers might require trading off the space complexity for the time complexity (Bishop 1995; Duda et al. 2001). The neural network-based classifiers are considered fast learners and easy to implement. However, some space complexity issues detract from their performance. The statistical classifiers, on the other hand, are considered slow but space-parsimonious learners (Blue et al. 1994). Both types of classifier could be set up to learn from the training data.

Neural networks are generally viewed as parallel computing systems comprising of large number of processors with several interconnections. The neural network models employ organisational principles such as learning, generalization and adaptability in a network of weighted directed graphs. The nodes in theses graphs are artificial neurons and the directed edges, with weights, provide connections between the input and output neurons. Since the neural networks are able to learn complex non-linear input-output relationships, employ sequential training procedures and adapt to data, their use is common in domains such as bioinformatics and biometric [Abbas and Fahmy 1994; Jain et al 2000].

The main goal of this investigation was to examine the possibility of classifying the measured temporal facial thermal features into facial expressions. Hence, the work required developing some explicit parametric cost functions using an appropriate parametric model. The neural network approach was found less suitable for this work as it would typically produce a non-parametric and model-free classifier.

A parametric statistical classifier could implicitly estimate the class densities and take the estimate of the *a priori* probabilities of class membership into account. It would allow representing each facial thermal feature pattern in terms of *p* measurements and viewing it as a point in a *p*-dimensional measurement space. Thus, it would help choose those features that allow patter vectors belonging to different facial expressions to occupy compact and disjoint regions in the given feature space. Hence, the statistical pattern analysis approach would allow establishing the viability of using temporal facial thermal features for classifying the facial expressions of affective states (Fukunaga 1990; Jolliffe 2002; Webb 2002). A statistical classification approach was therefore considered more appropriate for classifying the facial thermal features in this investigation.

5.1 Pattern classification approaches for implementing AFEC systems

The three higher-level tasks that a statistical classification network performs in a sequential order are usually referred to as face model acquisition, feature extraction and classification in the context of a vision-based AFEC capable system (Fasel and Luettin 2003; Pantic and Rothkrantz 2000). Face acquisition ensures availability of a face to extract the required features. Facial feature extraction involves either face localisation (when the static images are used) or face tracking (when a sequence of facial images is used). The feature extraction process would typically lead to the development of a representative model of the face. The expression classification requires definition and recognition of the expression categories to complete the AFEC process (Fasel and Luettin 2003; Pantic and Rothkrantz 2000).

Figure 5.1 shows the schematic diagram of a typical vision-based AFEC system. The pattern analysis parts of the system are visible in the grey box in Figure 5.1.

Like their vision-based counterparts, the NVAFEC systems are built using a standard classifier network architecture such as Artificial Neural Network (ANN), Hidden Markov Model (HMM), *K*-nearest Neighbor (*K*NN), Logistic Regression (LR), Naïve Bayes Classifier, Support Vector Machines (SVM), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) (Bartlett et al. 1999; Chen and Huang 2003; Cohn et al. 1999; DeSilva et al. 1997; Essa and Pentland 1997; Tian et al. 2002).

The classifier design usually depends on factors such as the type of available features, sample size, the probability density distribution of the features (given certain conditions), number of features and the availability of *a priori* information (Duda et al. 2001; Everitt and Dunn 1991; Fukunaga 1990; Manly 1994; Turner and Thayer 2001).

5.1.1 Frame-based AFEC systems

The measurement space in this investigation comprises of discrete and static time-sequential thermal infrared images. The classification problem therefore requires extracting temporal information form a set of static images. The AFEC systems that extract temporal features from discrete static images are usually referred to as either image-based or frame-based classifiers in the literature (Baldwin et al. 1998; Cottrell and Metacalfe 1991; Donato et al. 1999; Fasel and Luettin 2003; Haussecker and Fleet 2000; Moriyama et al. 2002; Pantic and Rothkrantz 2000; Terzopoulos et al. 2004). Standard statistical pattern recognition schemes employed in the image sequence-based AFEC systems include LDA, HMM, ANN and rule-based classification (Sung and Poggio 1996; Swets and Weng 1998; Tian et al. 2001; Tian et al. 2002).

Several earlier investigators have proposed a three-stage classification process for classifying features when little temporal information is available. In the first stage, the feature space dimensions are reduced to discover the major directions of variance in the data. During the second stage, the best discriminating principal components are discovered using an appropriate criterion function such as minimum error rate or maximum class separation. It is recommended not to select the classifier-training features using a threshold value for maximizing the inertia. During the third and final stage, the best discriminating features are projected in a compact optimal feature space (Chen and Huang 2003; Cottrell and Metcalfe 1991; Dubuisson et al. 2002; Jolliffe 2002; Kim et al. 2003; Krishnan et al. 1996; Lyons et al. 1999; Webb 2002).

NVAFEC capable systems employ similar parametric approaches for feature extraction, selection and classification. Like their vision-based counterparts, these systems also use the temporal information drawn from a sequence of input sources such as infrared video or a stream of audio signals (Abidi et al. 2004; Ang et al. 2004).

Selection of the employed recognition and classification approaches in these systems depends on the nature and format of the extracted features.

5.2 The proposed AFEC approach

The pattern recognition approach being proposed for this investigation was developed after a careful consideration and analysis of the works reported in (Abidi et al. 2004; Ang et al. 2004; Calder et al. 2001; Dubuisson et al. 2002; Everitt and Dunn 1991; Gottumukkal and Asari 2004; Huang and Y. Huang 1997; Kim et al. 2003; Kirby and Sirovich 1990; Krishnan et al. 1996; Nakayam et al. 2005; Pavlidis 2004; Pavlidis and Levine 2002; Pollina et al. 2006; Puri et al. 2005; Sharma 1996; Turk and Pentland 1991; Turner and Thayer 2001; Sugimoto et al. 2000; Yoshitomi et al. 2000). The proposed computational approach, indicated in Figure 5.2, begins by deriving the principal components to obtain independent linear combinations of the measured TIV data. The principal components are then examined for their contribution in the between-cluster separation and only the most discriminating features are kept to construct an optimal feature space.

The resulting optimal eigenspace is partitioned using the linear discriminant hyperplanes. The partitioning itself classifies new samples of TIV data with unsuitable levels of ambiguity, so a metric-based classification procedure assigning new input to the nearest facial expression cluster using the Mahalanobis distance is developed. Following paragraphs present algorithmic details of the proposed AFEC approach.

5.3 Principal component analysis

The pattern recognition literature suggests that PCA is a computationally inexpensive and robust feature extraction method (Berkey 1991; Jolliffe 2002). In particular, PCA is highly effective in situations when a large number of features and a small number of samples raise complexity and data-integrity issues. PCA reduces the computational risks by extracting such key feature indices that result in a linear combination of the original features and are capable of retaining the maximum information about a particular class (Jolliffe 2002). Furthermore, multi-collinearity in the data increases the chances of compounding computational errors during the parameter estimation (Field 2000). PCA, through the linear transformation, allows forming a smaller number of uncorrelated variables that provide maximum information about the features in a low-dimensional space and yield a stable regression model (Jolliffe 2002; Webb 2002).

Several advantages of invoking PCA on the temporal and non-temporal measurements are reported in (Calder et al. 2001; Dubuisson 2002; Everitt and Dunn 1991; Gottumukkal and Asari 2004; Huang and Huang 1997; Kirby and Sirovich 1990; Sharma 1996; Turk and Pentland 1991; Turner and Thayer 2001; Webb 2002). The algorithmic details of principal component derivation are presented in the following section.

5.3.1 Principal component derivation

In order to derive the principal components, each thermal image was considered a *p*-dimensional random facial thermal vector **x**. There were *n* such thermal feature vectors, $\mathbf{x_i}, (i = 1, 2, \ldots, n)$ having *p* TIV measurements in the learning set. Each thermal vector could be represented as $\mathbf{x_i} = [x_{i1}, x_{i2} \ldots, x_{ip}]^T$. The TIV data were standardised to draw a learning set $G_0 = [\mathbf{x}_1 \mid \mathbf{x}_2 \mid \ldots \mid \mathbf{x}_n]$ containing the *n* number of *p*-dimensional facial thermal vectors. The mean facial thermal vector $\bar{\mathbf{x}}$ of the learning set was obtained as:

$$\bar{\mathbf{x}} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{n=1}^{n} \mathbf{x_i} \qquad 5\text{-}1$$

The mean facial thermal vector $\bar{\mathbf{x}}$ was subtracted from each random facial thermal vector present in the data set to find its difference $\tilde{\mathbf{x}}_\mathbf{i}$ from $\bar{\mathbf{x}}$ as

$$\tilde{\mathbf{x}}_\mathbf{i} = \mathbf{x_i} - \bar{\mathbf{x}}. \qquad 5\text{-}2$$

After off-setting in this way, the learning set was presented as a *pxn* matrix

$$G = [\tilde{\mathbf{x}}_1 \mid \tilde{\mathbf{x}}_2 \mid \ldots \mid \tilde{\mathbf{x}}_n].$$

The *pxp* covariance matrix of the learning set was thus

$$C = GG^T. \qquad 5\text{-}3$$

The covariance matrix *C*, being symmetric and positive-definite was reducible to the form

$$C = H_l D H_l^T \qquad 5\text{-}4$$

Where the linear transformation matrix H_l is an orthogonal non-zero eigenvector matrix of *C* and is represented as columns of eigenvectors

$$H_l = [\upsilon_1 \mid \upsilon_2 \mid \ldots \mid \upsilon_p]. \qquad 5\text{-}5$$

The face recognition and facial expression classification literature often refers to the eigenvectors $\upsilon_i[i = 1, 2, \ldots, p]$ as eigenfaces, and we follow this nomenclature The matrix *D* in Equation 5-4 is the corresponding diagonal eigenvalue matrix of H_l such that

$$D = diag[\lambda_1, \lambda_2, \ldots, \lambda_p] \qquad 5\text{-}6$$

The diagonal elements of the eigenvalue matrix *D* are arranged in a descending order as $[\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p]$. Arranging eigenvectors in this order shows the most important and largest directions of the variance in the data set. By removing the lowest eigenvalues from D, and the corresponding columns from the transformation matrix H_l, suitable data-reduction is achieved, reducing the thermal feature vector-space to a span of only M eigenfaces ($M << p$).

The learning set *G* in this work was pre-classified so it was easy to group together the facial thermal feature vectors into *g* number of facial expression clusters. Thus, the data set *G* could be regarded as a disjoint union of *g* facial expression groups such that

$$G = G_1 \cup G_2 \cup \ldots \cup G_g . \quad \text{5-7}$$

By arrangement, n_j samples of a face with expression *j* were included in the group G_j. Hence the statistical model of the data set *G* could be assumed to take the form

$$\tilde{x}_{ijk} = \mu_i + \tau_{ij} + \varepsilon_{ijk}; \ (1 \leq j \leq p,\ 1 \leq j \leq g,\ 1 \leq k \leq n_j) \quad \text{5-8}$$

In Equation 5-8, $\tilde{x}_{ijk}$ is the i^{th} observation for a face expressing emotive state *j* and is the k^{th} such face with this expression, μ_i is the mean value of all observations at point *i*, τ_{ij} represents the offset of the centre of the j^{th} cluster from μ_i and ε_{ijk} is a residual that is minimised while estimating the other model parameters from the data set. This model can be generalised to represent the expression of an affective state as multiple clusters if some variations are bimodal or of higher modality. However, the small sample of facial thermal images used in this work had no multi-modal variations and did not require such a generalisation.

5.4 Discovering the best discriminating features

Even after data-reduction, the principal components could not be trusted to yield the true discrimination functions for separating the feature space (Chatfield and Collins 1995; Duda et al. 2001; Webb 2002). These cited authors argued that the directions discarded by the PCA might contain the real directions needed for discriminating between the groups under investigation. For this reason, the role of PCA in the classification problems like the AFEC is generally limited to providing the few uncorrelated components for retaining the original information (Dubuisson 2002). For developing a classification system, the PCA is often followed by LDA, since the LDA, not the PCA, seeks the directions which explicitly separate each group from the other groups (Jolliffe 2002; Webb 2002).

There is growing evidence, however, that LDA, when directly invoked on the high eigenvalued principal components, does not produce the most effective discriminant functions (Jolliffe 2002; McLachlan 2004; Webb 2002). The literature suggests that instead of invoking LDA on all principal components remaining after data-reduction, a set of the most influential principal components should be discovered and used for a compact yet holistic and effective representation of the available feature space. These ‘optimal’ principal components would usually capture representative features of the original data set most important for discriminating between the clusters (Webb 2002). Accordingly, based on the relevant information available in (Chatfield and Collins 1995; Dubuisson 2002; Duda et al. 2001; Jolliffe 2002; McLachlan 2004; Webb 2002), an optimal feature selections schema was developed to further the investigation.

Having derived the M number of linear principal components from the set of measurements on P available variables, a subset of K best discriminating principal components was needed to optimally partition the decision space. This could be done by evaluating a pre-defined optimality criterion on the M derived principal components to select a subset of K best discriminating principal components for which the criterion was maximised (Jolliffe 2002; Webb 2002).

In order to choose a set of best discriminating features, a measure of the features’ able to distinguish between the g facial expression groups was needed. It is usually recommended to estimate the overlap between the distributions from which the data are drawn and select those features for which the overlap is minimal. This recommended procedure maximises the between-cluster separation (Webb 2002). Such an optimal feature selection approach would eliminate the less effective features and would retain only the best discriminating features in a resulting optimised space. The Fisher criterion (F), a general class separability measure, is widely used to measure the contribution of a feature set in the between-group separation (Dubuisson 2002; Jolliffe 2002; Liu and Motoda 1998; Webb 2002).

5.4.1 Optimal feature selection algorithm

An optimal facial thermal feature selection algorithm was developed for this investigation. The Fisher’s criterion was adapted in this work for selecting the optimal feature set. It is the ratio between the determinants of the between-class scatter matrix S_B and the within-class scatter matrix S_W and is generally expressed as

$$F = |S_B| / |S_W|. \tag{5-9}$$

The parameters required to compute the F ratio are estimated using relatively standard matrix algebra. The following notations were used to make this algebra clear.

Eigenvalues and eigenvectors of C are denoted, with eigenvalues in decreasing order of size $\lambda_1, \lambda_2, ...\lambda_p \quad and \quad \upsilon_1, \upsilon_2, ...\upsilon_p$ respectively. The p-dimensional column vector of the components of $\tilde{x}_{ijk}$ along the eigenvector directions is denoted $q_i^{(jk)}$, suffix i being omitted when referring to the whole column rather than a component. The cluster centroid of cluster l in this basis is

$$u^{(j)} = \frac{1}{g_j} \sum_k q^{(jk)}. \tag{5-10}$$

The global centroid in this basis is $u = \frac{1}{n} \sum_j \sum_k q^{(jk)}$. 5-11

The within-cluster scatter matrix of the training sample is

$$S_W = \sum_j \sum_k (q^{(jk)} - u^{(j)})^T (q^{(jk)} - u^{(j)}). \tag{5-12}$$

The between-cluster scatter matrix of the training sample is

$$S_B = \sum_j (u^{(j)} - u)^T (u^{(j)} - u). \tag{5-13}$$

Given a set of d distinct eigenfaces $\{\upsilon_{K1}, \upsilon_{K2}, ...\upsilon_{Kd}\}$ the subspace spanned by the set is denoted $X_d(K)$ and the projector $P_d(K)$ that maps onto this space is the symmetric and idempotent matrix which projects onto $X_d(K)$ is given by

$$P_d(K) = \sum_{l=1}^{d} (\upsilon_{Kl})^T (\upsilon_{Kl}). \tag{5-14}$$

In the case where $X_d(K)$ is the first m eigenfaces carrying say 95% of the variance eigenvalue total, the associated projector performs the data-reduction of standard PCA. More generally, the projected $d \times d$ scatter matrices $P_d(K)S_W P_d(K)^T$ and $P_d(K)S_B P_d(K)^T$ represent the within- and between-cluster scatter in subspace $X_d(K)$. The Fisher ratio for the subspace $X_d(K)$ is the ratio of determinants

$$F_d(K) = | P_d(K)S_B P_d(K)^T | / | P_d(K)S_W P_d(K)^T |, \tag{5-15}$$

which, using the cyclic invariance of determinants and the idempotence of projectors reduces to the simpler expression $| P_d(K)S_B | / | P_d(K)S_W |$. With a view to optimising such a ratio, it is helpful to consider the set

$$\Xi = \{X_d(K) \mid d = 1,2,...,m\} \qquad \text{5-16}$$

of all candidate subspaces after data-reduction on which F-ratios can be defined. Then the optimization goal is expressible as one of finding

$$\hat{F} = \max_{X \in \Xi} \left\{ \frac{|P(X)S_B|}{|P(X)S_W|} \right\}. \qquad \text{5-17}$$

Such an optimum F-ratio occurs on a subspace $\hat{X}_d(K)$ specified by

$$\hat{X}_d(K) = \arg\max_{X \in \Xi} \left\{ \frac{|P(X)S_B|}{|P(X)S_W|} \right\}. \qquad \text{5-18}$$

In this investigation, the search for such an optimally discriminating subspace is conducted iteratively. A stepwise forward selection algorithm is considered computationally inexpensive and can select efficiently the sought after optimal features independent of the criterion function used. It can also avoid any overfitting of the data (Webb 2002). As a preparatory step, the eigenvalues of the eigenfaces were used to order them in a descending order.

The start-up is with the highest eigenvalued eigenface, and the iterative step is to include the next eigenface to step up dimension *k* by one unit. In the larger space, if the new F-ratio exceeds that in the former space, the next eigenface is included in a preferred set, the expanded space is returned and iteration re-commences. If there is no improvement in F-ratio, the next eigenface is rejected, the space dimension is stepped down, and the iteration continues with a new next eigenface. The process terminates when all *M* eigenfaces have been incorporated or rejected. A flowchart summarizing this algorithm is given in Fig 5.3 on the next page.

5.5 Facial expression classification

LDA has been successfully used in several related investigations (Sung and Poggio 1996; Swets and Weng 1998). It works at three levels for optimally dividing a Gaussian like feature space into linear regions of interest. At the first level, it identifies the variables that best separate each cluster in a training sample from the rest of the sample. On the second level, LDA uses the identified variables to define and compute new functions of input data. It does so by parsimoniously projecting the between-cluster differences.

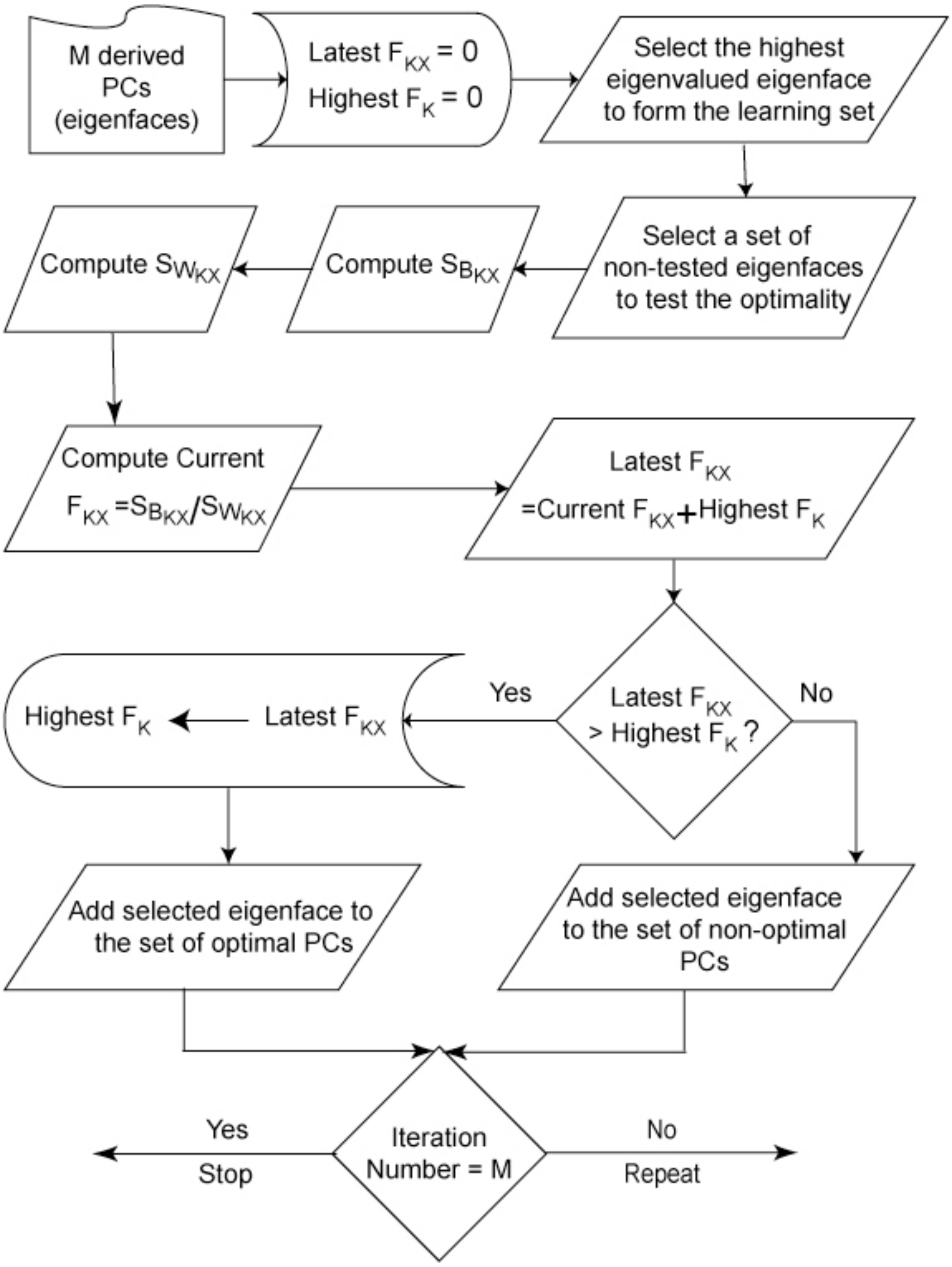

Figure 5.3: Optimal features subset selection algorithm

At the third level, LDA uses the discriminant functions to classify any future observations (Everitt and Dunn 1991; George and Mallery 1995). In essence, LDA seeks a linear space to maximise the between-group separation. Since there were *K* optimal features in the optimal learning set, the between-cluster separation measure J_K would allow quantifying the discrimination power of the training features (Everitt and Dunn 1991; Jolliffe 2002; McLachlan 2004).

LDA essentially seeks a transformation matrix W_p that maximises the ratio of between-cluster scatter S_B to within-cluster scatter S_W. The ratio between the determinants of S_B and S_W is a scalar measure of the scatter. If W_p denotes the optimal projection matrix, the ratio $J_K(W_P)$ estimated using Equation 5-19 measures the hyper-ellipsoidal scattering volume in the decision space and provides an effective discriminant criterion function.

$$J_K(W_P) \;=\; \arg\ \max\{|W_p^T S_B W_p| / |W_p^T S_W W_p|\} \qquad \text{5-19}$$

In a typical forward selection stepwise LDA, this ratio is maximised by resolving the equation

$$S_B W_p \;= \Lambda S_W W_p \qquad \text{5-20}$$

where the matrix W_p contains the eigenvectors of $S_W^{-1} S_B$ and Λ is a diagonal eigenvalue matrix. For a typical g-class distribution, the column vector of W_p (also referred to as the *Fishervector*) provides the basis for the (g-1)-dimensional optimal space.

5.5.1 Classification algorithm

In order to optimally separate a cluster of the training samples from the rest of the samples, a hyperplane is generally needed. With G facial expression clusters, the resulting G hyperplanes partition the observation space into 2^G regions bounded by the hyperplanes, of which G contain only one cluster centre, ${}^G C_2$ contain 2 centres, ${}^G C_3$ contain 3, and so on. A new thermal image vector may be 'classified' by assigning it to one region, using the transformed features associated with W_p above. But if the region contains several centres, the classification becomes ambiguous.

A distance or similarity criterion is usually used to remove the classification ambiguity. The distance criterion help assign a new vector to the nearest or most similar cluster centre in the region. In this work the Mahalanobis distance is used and defined in terms of the pooled within-cluster variance matrix, W_w, of a training sample. If e-e_g is the vector joining a new image vector $\mathbf{x_i}$ to the centre of cluster g, then the Euclidean length $\|(e-e_g)W_w^{-1}\|$ defines the Mahalanobis distance from the input image vector $\mathbf{x_i}$ to cluster g. Thus, the nearest cluster to an unknown facial thermal vector in a region containing several cluster centres is given by

$$arg\ min_g\ \{\|(e-e_g)W_w^{-1}\|\}. \qquad \text{5-21}$$

Mahalanobis distance is a widely used metric for comparing distances from an unknown feature vector to training classes. Others are the k-nearest neighbour metric that averages distances from the unknown vector to a small number k of near neighbours in a training cluster, and mean or median distance metrics that take a statistical ‘average’ of all distances from unknown to the cluster members.

A major advantage of using the Mahalanobis distance as a measure of similarity is that the squared Mahalanobis distance ensures maximum separation among all pairs of the groups. At each step of the forward selection stepwise LDA, the variables that provide the maximum increase in the measure of between-group separation are determined using the squared Mahalanobis distance. These effective variables are then used for further analysis of the between-group separation (Everitt and Dunn 1991; McLachlan 2004).

5.5.2 Cross-validation

Pattern recognition practitioners respond to the multiplicity of choices for feature space and distance metric by empirical testing to estimate the error rates associated with different discriminant rules (e.g.Fukunaga 1990; McLachlan 2004). Ideally, this is done with a validation sample chosen independently from the same population as the training sample. If samples are costly to obtain and process, this may not be feasible, so ways of re-using training sample data have been devised – so-called cross-validation procedures. If the experimental errors in raw data features are well understood, both blurring errors and impulsive errors, simulated forms of these can be injected into training data to produce a reliably-simulated validation sample. But the investigation of experimental errors may involve prohibitively long visits to a physical standards laboratory. Accordingly, several ‘fast and dirty’ cross-validation methods have been devised and employed for testing the performance of statistical classifiers. Two such methods, the split-sample cross-validation (CV) and the leave-one-out (LOO) method have been widely used in multi-class classification problems. Split-sample CV, essentially dividing the sampled data into a training part and a validation part by random selection, runs into small-sampling errors, and if the two parts of the split are allowed to share members, may bias the empirical error rates towards an over-optimistic view of classifier accuracy. However, (McLachlan 2004) asserts that the LOO method avoids such bias. Essentially, the LOO method extracts one of the *n* feature vectors from the training sample and estimates the discriminant functions using the remaining *n-1* samples. The extracted feature vector has a known cluster label and is assigned another label on test. (Fukunaga 1990; McLachlan 2004; Webb 2002). This training-testing cycle in repeated *n* times to test each of the *n* available samples. Since the L-method is believed to yield an unbiased estimate of the classifier performance, it was preferred for estimating the classifier performance in this work.

The L-method begins by assuming $\mathbf{x_i}$ as a *p*-dimensional random thermal feature vector. The training data can be denoted by $G=\{\mathbf{x_i}, i=1, \ldots, n\}$ such that each realised sample $\mathbf{x_i}$ has two parts ; $\{p_i$, i=1,…,p$\}$ denoting the measurements and $\{z_i$, i=1,…,n$\}$ denoting the corresponding class labels. The class labels can be coded as corresponding vectors $(\mathbf{z}_i)_j = 1;\ \mathbf{x}_i \in G_i$ and $(\mathbf{z}_i)_j = 0;\ \mathbf{x}_i \notin G_i$. The corresponding categorical class label can then be represented as $\omega(\mathbf{z}_i)$. Also, the discriminant rule developed using the training data can be represented as $\eta\ (x; G)$; where, η is the class to which **x** is assigned to by the classifier using *G*. Using this notation, a loss function $Q(\omega(z), \eta(x;G))$ can be estimated as

$$Q(\omega(z),\ \eta(x;G)) = \begin{cases} 0 & if \quad \omega(z) = \eta(x;G) - correct \quad classification) \\ 1 & otherwise \end{cases} \qquad \text{5-22}$$

Having the training set G_j with a sample thermal image $\mathbf{x}_j$ removed from it, the cross-validation error can be estimated as

$$e_{CV} = \frac{1}{n}\sum_{j=1}^{n} Q(\omega(z_j),\ \eta(x_j, G_j)). \qquad \text{5-23}$$

Since it needs *n* iterations to test each of the *n* samples in a data set, the computational complexity of L-method becomes its major limitation. Another one that is peculiar to classifying humans is that classifier performance is different for participants inside and outside a training sample. A recogniser will generally be more accurate with samples from the same familiar participants as the training sample. Strangers will not express themselves to exactly the same patterns as familiars. Thus the LOO method does not entirely escape bias, unless the material left out is all the data supplied by one participant (that is all his expressions, so that, on test, the classifier sees a stranger rather than a familiar).

5.6 Classification significance

Loss functions on test, whether ideal tests or simulation tests or cross-validation methods, may be small and insignificant, but if original raw data, or data after reduction and transformation is not sufficient for the classification task attempted, errors of statistical significance may be present. Such errors are of concern in the practical use of automatic classifiers. Of several computational approaches used for determining the statistical and practical significance levels of the classification results, the approach (Huberty 1984; Huberty 1994) suggested was found more suitable for this investigation. Huberty's computational approach, described below, was used to determine the statistical and practical significance levels of the classification results.

5.6.1 Determining the statistical significance of classification results

Based on Cohen's approach (Cohen 1977) of analysing the classifier performance and evaluating the practical significance of the classification results, Huberty developed a test for assessing the statistical significance of the classification results (Huberty 1984; Huberty 1994). He considers a validation sample of size n, in which o_g is the number of correct classifications for group *g*, n_g is the number of observations in group *g*, *o* is the total number of correct classifications. The *z*-statistic for classification of a group is estimated as

$$Z_g^* = (o_g - e_g)\sqrt{n_g} \Big/ \sqrt{e_g(n_g - e_g)}, \qquad \text{5-24}$$

where the expected mean value of correct classifications over Poissonian trials with this sample size is $e_g = n_g^2 / n$. For testing the overall classification results using an expected mean value of correct classifications over all groups $e = 1/n \sum_{g=1}^{G} n_g^2$, the *z*-statistic is given by

$$Z^* = (o - e)\sqrt{n} \Big/ \sqrt{e(n - e)}. \qquad \text{5-25}$$

These test statistics Z_g^* and Z^* follow a distribution that is asymptotically standard normal, and provide an estimate of the extent to which the classification rates are significant.

5.6.2 Determining the practical significance of classification results

The literature suggests using the practical power of a classifier is to be rated by comparison of its correctness rate with the rate likely to be obtained using random assignment of validation sample members to classes (Sharma 1996). A well-tried test for assessing the practical significance of the classification results was used to assess the practical significance of the classifier (Huberty 1984; Huberty 1994). The practical significance index is given by

$$I = {}^{o/n - e/n}\!\big/_{1 - e/n} \times 100 \qquad \text{5-26}$$

where *o* is the number of correct classifications, *e* is the expected correct number of classification, and *n* is the total number of observations. The ratio is one of the improvements on random assignment observed under test, to the maximum possible improvement achievable with an error-free classifier. Equations 5-24 to 5-26 were used throughout this work for determining the statistical and practical significance of the classification results.

5.7 Advantages of the proposed AFEC approach

The proposed algorithmic approach offers the following major advantages:

1. The employed algorithmic approach is based on the multivariate analysis of the TIV data. The two-way multivariate model developed for analysis of the TIV data measured at equal points in the time sequential images helps in separating the multitude of built-in covariance in the acquired TIV data. The multivariate analytical approach might also help in estimating the effect of facial expressions on the facial skin temperature and thus help in minimising the overall error rate (e^{j}_{geneal}). Having several linear spaces in the learning set, each spanned by a selected set of principal components, it was possible to represent the affective states as distinguished and separated clusters. The restriction of the diagonalised matrix *D* to each linear space encoded varying amounts of within-cluster and between-cluster variances in each image vector of the data set.
2. The number of samples in this investigation was smaller than the number of measurements in the available measurement space. Consequently, the inverse of the within-subject matrix could get closer to being singular. Hence it was not viable to apply the LDA directly on the set of available features for the statistical independence of the features in the training set could not be established (Duda 2001). The proposed method reduces the dimensionality of the available feature space through the PCA and produces the transformed features that are the linear combinations of the original features. Thus, no information is lost in effect and the actual discrimination power of the original feature set is retained. During the

feature selection phase, the most discriminant components could be sorted and projected in an optimal space to form the discriminant functions.

3. When PCA is invoked and resulting high valued eigenvectors are used to construct the discriminant pace, retention of the most discriminant features is not guaranteed. The proposed classification approach helped in retaining the maximum discrimination information since it selected the optimal components for a following LDA. As demonstrated in the following chapters, this approach resulted in a much less classification error rate e^{j}_{geneal} as compared to the higher valued eigenvectors based LDA approaches. The proposed approach guaranteed the maximum between-class separation and exhaustively examined the probabilities of allocating the new and unknown TIRIs to a particular facial expression group.
4. One common purpose of using the PCA - LDA combination is to further reduce the dimensions of the feature space by replacing the observation vector **x** by the first *m* (high variance) principal components for deriving the discriminant rules. A common assumption in such uses of the PCA-LDA approach is that the groups have a similar covariance structure. Hence, it is assumed that the PCA is being invoked on an estimate of the common within-group covariance matrix. This procedure often proves to be wrong for two reasons. First, the within-group covariance matrix might be different for each individual expression group. Second, there is no guarantee that the between-group separation follows the direction of the high-variance principal components. The first few principal components provide useful information about the variance when the 'within-group' and 'between-group' variations have the same directions. If that is not the case, the low-variance principal components possess most of the information about the between-group separation (Jolliffe 2002). In such cases, the low-variance principal components may also be highly correlated with the dependent variables. A variation of this problem occurs in situations when one ignores the group structure and calculates an overall covariance matrix based on the raw data (Jolliffe 2002). When the between-group variation is much larger than the within-group variation, the first few principal components define the direction in which there were large between-group differences (Jolliffe 2002). The employed approach takes advantage of the covariance matrix of the mean-centered raw data to derive the principal

components and then apply a criterion-based selection rule to find the best discriminating principal components. In doing so, the ratio between the between-group and the within-group scatter is used to discover the most effective principal components. Therefore, the problem of loosing any information available in the actual data is minimised.

5.8 Conclusion

After an extensive review of the statistical and neural classification networks, a robust algorithmic approach was proposed for classifying the facial thermal features in this investigation.

The proposed algorithmic approach, in a sequential manner, extracts feature, selects the optimal feature, and classifies them for achieving the AFEC functionality. It first discovers the most effective dimensions of variation in the thermal data and reduces the complexity of the available feature space. A set of optimal principal components is then discovered from the linear components of the original measurements. The optimal features are finally used develop the discriminant functions during the LDA. The minimum distance measure, calculated using the Mahalanobis distance, is then used to classify new and unknown thermal faces. The discriminant rules are cross-validated using the leave-one-out cross validation method. Lastly, the statistical and practical significance levels of the classification results are determined to further validate the discriminant functions.

Chapter 6

CLASSIFICATION OF PRETENDED POSITIVE AND NEGATIVE FACIAL EXPRESSIONS

A significant number of emotion theorists assert that a small set of discrete and basic emotions may represent all affective states and emotional experiences. Some scientific studies negate these assertions and describe emotions in terms of continuous dimensions of valance and intensity (Ekman and Friesen 1971; Picard 2000; Plutchik 1980; Tomskin 1984). Scientific evidence reported in the literature supports both theories hence emotions are described and measured in both ways. Several previous works on description and measurement of emotions are discussed in (Smith 1999).

As evident in (Ekman et al. 1993; Fasel and Luettin 2003; Mase 1991; Pantic and Rothkrantz 2000), the affective computing and automated facial expression recognition literature generally describes the affective states with the help of six basic facial expressions using the visual cues. Few recent investigations have attempted to recognise deceit and stress levels using either facial thermal features or facial hæmodynamic measurements.

This chapter reports an attempt to classify the facial expressions of affective states using the facial thermal features along the dimension of valence. First part of this chapter reports classification of neutral, happy and sad facial expressions. The following sections of this chapter present classification of the two positive (happiness and surprise) and two negative (angry and disgusted) facial expressions. The classifier performance and the observed classification results are analysed before concluding this phase of investigation.

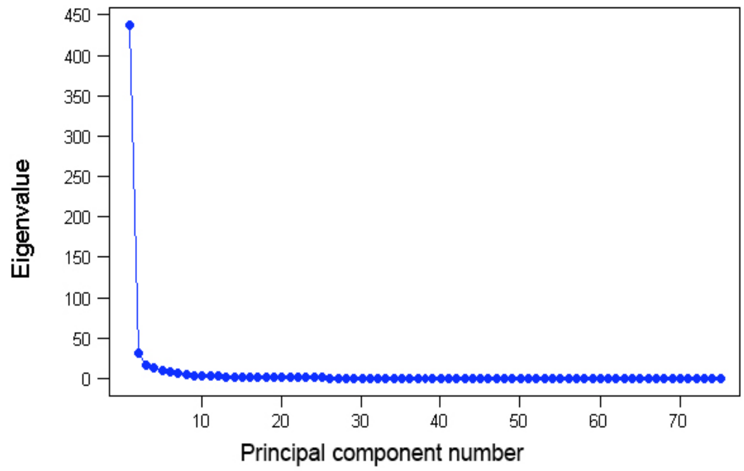

Figure 6.1: Contribution of the 75 principal components in the measured TIV data variance

6.1 Classification of neutral, happy and sad facial expressions

The computational approach proposed in chapter 5 was employed to classify the neutral and pretended happy and sad facial expressions. The TIV data obtained from the thermal images of 16 participants (reported earlier in chapter 4) were used in the analysis.

Initial data analyses results were reported earlier in chapter 4. In addition to the tests reported earlier, the TIV data measured in the TIRIs having pretended neutral, happy and sad expressions were also tested for sphericity. The test of sphericity was significant [F (1.56, 0.892)= 1.743, p<0.01] suggesting some differences in the variance structures of the three facial expression groups. However the two more conservative estimates of sphericity, the Greenhouse-Geisser correction and the lower-bound correction were insignificant (p>0.05) suggesting the three facial expression groups had a similar variance structure (Field 2000; Kinnear and Gray 2000).

6.1.1 Classifier construction

Using the algorithmic approach described in chapter 5, the principal components were derived to reduce the dimensions of the decision space for classifying the neutral and pretended happy and sad facial expressions. The first 28 principal components (PCs) derived from the TIV data accounted for the majority (over 99%) of scatter in the thermal data. Each of the first 28 PCs accounted for at least one percent of the total variance. However, only first seven PCs, shown in the scree plot in Figure 6.1, accounted for over 90% of the scatter in the thermal data. These PCs helped in understanding the dimensions of variance in the thermal data pertaining to various facial expressions.

Figure 6.2 shows the positive and negative values of the variable weights (w) for the first seven PCs. The white dots in Figure 6.2 represent the positive w values and black dots represent the negative w values of the variables for each PC. A closer examination of the first seven PCs reveals the origins and dimensions of the differences between the thermal data. Some useful relationships between the thermo-muscular activities and the facial thermal characteristics might also be drawn by examining the first seven PCs since they accounted for the majority of variance in the data.

The first PC accounted for over 75% variance in the data. As evident in Figure 6.2, it provided an overall negative (thermal) index of the face.

The second PC provided an account of the facial thermal characteristics by adding the variable weights on Orbicularis Oculi Pars Orbital and Levator Labii Superioris Alaquae Nasi and subtracting the variable weights measured on all other facial muscles from the added values of the variable weights.

The third PC added the variable weights on Orbicularis Oculi Pars Orbital and Masseter Superficial, Levator Labii Superioris Alaquae Nasi and subtracted the variable weights on other major muscles. Hence, it seems to be examining the differences between the thermal characteristics of these muscles.

The fourth PC appears to be comparing the differences between the variable weights taken from the muscles on the right and left sides of the face. It seems to be ignoring the variable weights measured on Depressor Labii Inferioris and Mentalis though.

The fifth PC seems to be repeating the same measurements. However, it adds the variable weights on Depressor Labii Inferioris and Mentalis to the discovered thermal difference between the two sides of the face.

The sixth PC adds the variable weights on Frontalis Pars Medialis, Frontalis Pars Lateralis, Levator Labii Superioris Alaquae Nasi, Depressor Superior Silii, Risorious, and Platysma and subtracted the sum from the sum of the variable weights measured at various FTFPs sites.

The seventh PC adds the variable weights on Frontalis Pars Lateralis, Depressor Superior Silii, Levator Anguli Oris, Zygomaticus Major, Depressor Anguli Oris, Buccinator and Depressor Labii Inferioris and subtracts the added values from the added values of TIVs on other facial muscles. This might help in finding the difference between the TIVs measured at the major facial muscles to create a thermal profile of the face.

Furthermore, PCs 2, 3, and 4 appear to be calculating the thermal gradient between the FTFP sites on (1) Frontalis, (2) Levator Labii Superioris, (3) Depressor Anguli Oris, (4) Buccinator and (5) Zygomaticus Major. The other PCs apparently compared the thermal gradient on various facial muscles. It may be noticed that the PC-1 tries to extract a first-order thermal feature set that was based on the direct thermal measurements of the face. All other PCs extract a second-order feature set that is based on the relative thermal features calculated from the first-order thermal features.

Figure 6.2 suggests that the first three PCs, which account for about 82 % variation in the data, constantly keep a (negative) variable weight for the lower face FTFPs. It may also be noted that the valance of the variable weights on the lower part of the face in the other PCs did not change. On the contrary, the valance of the variable weights on the FTFPs located on the upper parts of the face exhibited noticeable variations in almost all major PCs. The visible imagery based facial expression classification results suggest that visual features gathered from the upper part of a face contribute more in facial expression of affects than the features around the lower part of the face (Ekman 1982). This pattern appears to be true in Figure 6.2 as well.

Figure 6.3 exhibits the possible separation between the neutral faces and the faces with intentional facial expression of happiness and sadness in a 2-principal component eigenspace. Though the first principal component (along the abscissa) contributes more to the variance, the separation shown in Figure 6.3 seems to be influenced by the second principal component (along the ordinate).

The eigenvalues of the 75 principal components derived from the TIV data on 16 participant faces is given in Table 6.1. Only 47 of 75 principal components in Table 6.1 contribute to the variance in the data. However, there is no evidence to suggest that only these 47 principal components may contribute to the separation in the facial expression groups (Duda et al. 2001; Jolliffe 2002). It was therefore considered necessary to project the principal components in an optimal subspace for discovering the most discriminating principal components.

Figure 6.2: TIV variations-based representation of the first seven principal components. The white dots represent +ve weights of the variables. The black dots represent the –ve weights of the variables

The resulting optimal principal components were expected to produce the most effective discriminant functions for classifying the unknown faces.

During the step 2 of the classifier construction, the principal components derived earlier, were analysed to select a set of optimal features using the stepwise elimination method. The optimal feature selection algorithm described in chapter 5 was used to select the best discriminating features (principal components) for deriving the discriminating functions. Figure 6.4 shows the recursion involved in the stepwise selection of the optimal principal components. It also shows the corresponding increase in the value of Fisher ratio $F = |S_B|/ |S_W|$. The line joining recursions 1 and 20 in Figure 6.4 explains the overall (recursive) improvement in the F ratio.

Figure 6.3: Separation between the neutral, happy and sad facial expression in a 2-principal component eigenspace

Figure 6.4: Recursive stepwise selection of the optimal components and the corresponding increase in the *F*-ratio ($=S_B/S_W$)

The 20th recursion and the corresponding optimal components in Figure 6.4 resulted in the highest value of the Fisher statistic (*F*). The principal components that helped in increasing the F ratio were not the ones that had the largest eigenvalues. Instead, the stepwise feature selection algorithm recursively discovered a new set comprising of both high and low eigenvalued principal components.

Since the probability density functions and the group memberships were known *a priori* for each group, each group could be given an equal *a priori* during the analyses. Table 6.2 presents the canonical discriminant functions, their relevant statistics and contribution of the two discriminant functions in the between-group separation of the three facial expressions. The canonical discriminant functions reported in Table 6.2 are orthogonal to each other and attempt to maximise the differences between the dependent variables.

TABLE 6.1: THE 75 PRINCIPAL COMPONENTS AND THEIR RESPECTIVE EIGENNVALUES

PC	Eigenvalue	Proportion	Cumulative	PC	Eigenvalue	Proportion	Cumulative
1	437.13	0.763	0.763	***39***	0.24	0.00	0.998
2	32.37	0.057	0.82	***40***	0.23	0.00	0.999
3	16.62	0.029	0.849	***41***	0.19	0.00	0.999
4	13.87	0.024	0.873	***42***	0.18	0.00	0.999
5	10.12	0.018	0.891	***43***	0.16	0.00	0.999
6	7.65	0.013	0.904	***44***	0.12	0.00	1.00
7	6.58	0.011	0.915	***45***	0.09	0.00	1.00
8	4.63	0.008	0.923	***46***	0.06	0.00	1.00
9	3.99	0.007	0.93	***47***	0.05	0.00	1.00
10	3.81	0.007	0.937	***48***	0.00	0.00	1.00
11	3.55	0.006	0.943	***49***	0.00	0.00	1.00
12	3.21	0.006	0.949	***50***	0.00	0.00	1.00
13	2.50	0.004	0.953	***51***	0.00	0.00	1.00
14	2.43	0.004	0.957	***52***	0.00	0.00	1.00
15	2.10	0.004	0.961	***53***	0.00	0.00	1.00
16	1.86	0.003	0.964	***54***	0.00	0.00	1.00
17	1.80	0.003	0.968	***55***	0.00	0.00	1.00
18	1.56	0.003	0.97	***56***	0.00	0.00	1.00
19	1.52	0.003	0.973	***57***	0.00	0.00	1.00
20	1.46	0.003	0.975	***58***	0.00	0.00	1.00
21	1.38	0.002	0.978	***59***	0.00	0.00	1.00
22	1.24	0.002	0.98	***60***	0.00	0.00	1.00
23	1.17	0.002	0.982	***61***	0.00	0.00	1.00
24	1.02	0.002	0.984	***62***	0.00	0.00	1.00
25	1.00	0.002	0.986	***63***	0.00	0.00	1.00
26	0.90	0.002	0.987	***64***	0.00	0.00	1.00
27	0.84	0.001	0.989	***65***	0.00	0.00	1.00
28	0.77	0.001	0.99	***66***	0.00	0.00	1.00
29	0.65	0.001	0.991	***67***	0.00	0.00	1.00
30	0.61	0.001	0.992	***68***	0.00	0.00	1.00
31	0.52	0.001	0.993	***69***	0.00	0.00	1.00
32	0.50	0.001	0.994	***70***	0.00	0.00	1.00
33	0.44	0.001	0.995	***71***	0.00	0.00	1.00
34	0.43	0.001	0.995	***72***	0.00	0.00	1.00
35	0.37	0.001	0.996	***73***	0.00	0.00	1.00
36	0.36	0.001	0.997	***74***	0.00	0.00	1.00
37	0.31	0.001	0.997	***75***	0.00	0.00	1.00
38	0.25	0.00	0.998				

TABLE 6.2: SUMMARY OF CANONICAL DISCRIMINANT FUNCTIONS

Function	Eigenvalue	Percentage of Variance	Cumulative percentage	Canonical Correlation
1	9.024	77.4	77.4	0.949
2	2.637	22.6	100	0.852

TABLE 6.3: CLASSIFICATION SUCCESS RESULTS WITH THE HIGH EIGENVALUED PRINCIPAL COMPONENTS

Classification		Group	Predicted Group Membership			Total
			Neutral	***Happy***	***Sad***	
Original cases [a]	***Count***	***Neutral***	10	4	2	16
		Happy	7	4	5	16
		Sad	4	3	9	16
	Percentage	***Neutral***	62.5	25.0	12.5	100.0
		Happy	43.8	25.0	31.3	100.0
		Sad	25.0	18.8	56.3	100.0
Cross-Validated cases [b]	***Count***	***Neutral***	10	4	2	16
		Happy	7	4	5	16
		Sad	4	4	8	16
	Percentage	***Neutral***	62.5	25.0	12.5	100.0
		Happy	43.8	25.0	31.3	100.0
		Sad	25.o	25.0	50.0	100.0
[a] 47.9 % of original group cases correctly classified						
[b] 45.8 % of cross-validated group cases correctly classified						

Table 6.3 shows the confusion matrix observed when the high eigenvalued principal components were used to train the classifier and classify the facial expressions.

Table 6.4 shows the confusion matrix observed when the optimal principal components were used for discriminating between the three facial expressions. Figure 6.5 provides a visual comparison of the two classifiers' performance. The high eigenvalued components based classification resulted in a higher error rate (e^{j}_{geneal} = 100-45.8=54.2%). The classifier performance significantly improved (e^{j}_{geneal} = 100-83.8=16.2%) when the optimal principal components were used for training the classifier.

Figure 6.6 shows the decision boundaries resulting from the two classifiers. On the left side of Figure 6.6 is the decision boundary resulted from the high eigenvalued principal components. The decision boundary on the right side of Figure 6.6 resulted when the optimal components were used to train the classifier. The two decision boundaries are significantly different and the one resulting from the optimal components appears to be smoother and better separated.

TABLE 6.4: CLASSIFICATION SUCCESS RESULTS WITH THE OPTIMAL COMPONENTS

Classification		Group	Predicted Group Membership			Total
			Neutral	*Happy*	*Sad*	
Original cases [a]	***Count***	***Neutral***	16	0	0	16
		Happy	0	16	0	16
		Sad	0	0	16	16
	Percentage	***Neutral***	100	0	0	100.0
		Happy	0	100	0	100.0
		Sad	0	0	100	100.0
Cross-Validated cases [b]	***Count***	***Neutral***	13	1	2	16
		Happy	0	14	2	16
		Sad	1	2	13	16
	Percentage	***Neutral***	81.3	6.3	12.5	100.0
		Happy	0	87.5	12.5	100.0
		Sad	6.3	12.5	81.3	100.0

[a] 100.0 % of original group cases correctly classified

[b] 83.8 % of cross-validated group cases correctly classified

Figure 6.5: Comparison of the two classifiers' performance

Figure 6.6: (a) The decision boundary with the high-eigenvalued Principal Components and (b) The decision boundary with the Optimal Principal Components

6.1.2 Analysis of the classification results

The stepwise feature selection algorithm recursively selected a set of most influential variables. These most influential variables were used for developing the discriminant functions and constructing the classifier during the linear discriminant analysis. The statistical significance of the two functions developed for differentiating between the three facial expression groups was calculated using the following relationship:

$$\chi^2 = [n - 1 - ((p + G)/2)\sum_{k=1}^{k} \ln(1 + \lambda_k) \qquad 6\text{-}1$$

Having n (=*48*) observations, p (=20 variates), G (=*3* facial expression groups,) k (=*2* discriminant functions) and the corresponding eigenvalues (λ_k) computed earlier and given in column 2 of Table 6.2, the χ^2 value for each discriminant function was calculated using Equation 6-1. For example, the χ^2 value of the first discriminant function in Table 6.5 was found to be 127.663. It was computed as $\chi^2 = (48 - 1 - ((20 + 3)/2))[\ln(1 + 9.024) + \ln(1 + 2.637)]$.

This relationship in Equation 6-1 was recursively used for computing the significance level of each discriminant function, reported in column 5 of Table 6.5. The significance levels of the two discriminant functions (p<0.05) in Table 6.5 suggest the possible separation between the facial expression groups along the two discriminant functions reported in Table 6.5.

Since the derived discriminant scores were linear combinations of the original variables, their mathematical structures could reveal the nature of relationship between the actual variables and the discriminant functions. Table 6.6 presents the structure matrix resulted from the discriminant analysis. The structure coefficients of the discriminant functions allow interpreting the contribution of each variable to the formation of the discriminant functions. In other words, they represent the correlations between a given independent variable and the discriminant scores associated with a discriminant function. The coefficients of a given discriminator variable are therefore the coefficients of the correlation between the discriminant scores and the discriminator variables. Their numeric values range between +1 and -1. These coefficients (given in Table 6.6) were calculated using the formula

$$C_i = \sum_{k=1}^{k} r_{ij} b_j^*, \qquad 6\text{-}2$$

where, C_i is the coefficient of a variable; r_{ij} is the pooled correlation between the variables i and j; and b_j^* is the standardised coefficient of the variable j.

A closer examination of the structure matrix in Table 6.7 suggests that the first discriminant function heavily relies on variates 9, 10, 39 and 44. In effect, the first discriminant function relies on the thermal variations measured on Frontalis Pars Lateralis (9 and 10), Zygomaticus Major (39) and Levator Labii Superioris (44).

The second discriminant variable mostly relies on variates 10, 13, 39, 40 and 44. The second discriminant function therefore relies on the thermal variations on Frontalis Pars Medialis (13), Frontalis Pars Lateralis (10), Zygomaticus Major (39), Masseter Superficial (40) and Levator Labii Superioris (44).

Zygomaticus Major and Mentalis are considered the muscles of positive expressions (Kall 1990). Masseter Superficial and Labii Superioris are considered the muscles of sadness and fear (Kall 1990). The separation between the three facial expressions seems to be based on the transient thermal features measured along these major facial muscles. Figures 3.1 and 4.5 (with Table 4.1) exhibit the physical location of these major muscles.

Table 6.5: Significance of individual discriminant functions

Test of Functions	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	0.027	127.663	40	0.00
2	0.275	45.837	19	0.001

TABLE 6.6: STRUCTURE MATRIX REPRESENTING COMPOSITION OF THE TWO DISCRIMINATING FUNCTIONS

Principal component	Function 1	Function 2
VARIATE-44	0.116(*)	0.107
VARIATE-09	-0.106(*)	-0.022
VARIATE-04	0.088(*)	-0.011
VARIATE-14	-0.078(*)	-0.072
VARIATE-21	-0.076(*)	0.065
VARIATE-45	-0.061(*)	-0.044
VARIATE-29	0.057(*)	-0.001
VARIATE-43	-0.055(*)	0.009
VARIATE-30	-0.052(*)	0.021
VARIATE-12	0.075	0.264(*)
VARIATE-19	0.063	0-.249(*)
VARIATE-24	0.055	-0.217(*)
VARIATE-39	0.101	0.167(*)
VARIATE-10	-0.113	0.152(*)
VARIATE-13	-0.01	0.141(*)
VARIATE-40	0.037	0.133(*)
VARIATE-07	0.081	-0.091(*)
VARIATE-34	-0.065	-.082(*)
VARIATE-46	0.044	-0.078(*)
VARIATE-08	0.063	-0.072(*)
* Largest absolute correlation between each variable and any discriminant function		

6.1.3 Significance of the observed classification results

The classifier performance and the practical significance of the classification results were determined using the statistical tests proposed earlier in section 5.5. Equations 5-24 and 5-25 were used to determine the statistical significance of the classification results.

The *z*-statistic for the neutral, happy, and sad faces and the overall cross-validation results (reported in Table 6.4) were respectively found to be:

$Z^*_{normal} = 1.921$ (significant at alpha-level 0.0274);

$Z^*_{happy} = 1.512$ (significant at alpha-level 0.0655);

$Z^*_{sad} = 1.921$ (significant at alpha-level 0.0274);

$Z^*_{overall} = 3.096$ (significant at alpha-level 0.001).

The overall significance test statistic (*p<0.05*) suggested that classification results were significant. It is therefore safe to assume that the TIV data gathered at the 75 FTFP sites on the participant faces may help classify the neutral and pretended happy and sad facial expressions.

The practical significance of a classifier is usually determined to study its viability for the real-life use. The index (*I*) for the cross-validation results was calculated using Equation 5-26. The resulting index was found to be $I = {}^{40/48-16/48}\!\big/_{1-16/48} \times 100$ = 74.99. The estimated index (*I*) suggests that the employed discrimination method may help reduce the chances of making computational errors by 74.99%.

6.2 Classification of more negative and positive facial expressions

In a following analysis, an attempt was made to classify the two positive pretended (happiness and surprise) and the two negative (anger and disgust) facial expressions using the facial skin temperature measurements.

6.2.1 Classifier construction

Using the algorithmic approach reported earlier in Chapter 5, the TIV data were first transformed into uncorrelated principal components. The first 50 of the 75 derived principal components caused at least 1 % of the variation in the TIV data. However, there was no reason to assume that only these 50 principal components might have contributed to the between-facial expression group separation.

Figure 6.7 shows the possible separation between the two pretended positive (happy and surprise) and the two pretended negative (anger and disgust) facial expressions in a 2-principal component eigenspace. The first principal component (along the abscissa) appears to be contributing more to the variation in the TIV data. The between-group separation in Figure 6.7 seems to be well influenced by the second principal component (along the ordinate) as well. A strong separation between the facial expressions of happiness and anger is evident in Figure 6.7. The two negative facial expressions (disgust and anger) are also well separated along the second principal component. Similarly, the two positive facial expressions, happiness and surprise are also well separated in a 2-principal component eigenspace. However, the first two principal components do not appear to be helpful in distinguishing between the facial expressions of happiness and disgust.

It might help to recall that the facial expression of disgust was also difficult to recognise in the previous studies reported earlier in chapter 3, section 3.2.

Figure 6.7: Separation of positive and negative facial expressions in a 2- component eigenspace

Earlier investigations have concluded that the EMG readings taken on the major facial muscles (that influenced formation of the second Principal component) did not provide enough information about the facial expression of disgust. However, as reported later in this section, the two positive and the two negative facial expressions may be better separated in a higher-dimensional, optimal eigenspace. Probably some other principal components provided additional (and required) information for better separating the negative facial expressions.

The derived principal components were again examined to select a set of optimal principal components using the stepwise feature selection method. Figure 6.8 shows the stepwise elimination of less important principal components and a recursive selection of the most influential principal components. The corresponding increase in the F statistics ($F = |S_B|/ |S_W|$) is evident in Figure 6.8.

Table 6.7 presents the classification results and the confusion matrix obtained using the 36 highest eigenvalued principal components that accounted for more than 99% variance in the thermal data. Table 6.8 shows the confusion matrix obtained with the optimal principal components used for discriminating between the four facial expressions. In Table 6.7 an error rate e^{j}_{geneal} (= 62.5) was observed during the cross-validation tests when the highest eigenvalued features were used for recognition and classification. The classifier trained with the optimal principal components performed at a much lower error rate e^{j}_{geneal} (= 32.8).

Figure 6.8: Recursive stepwise selection of the optimal components and the increasing ratio S_B/S_W

TABLE 6.7: CLASSIFICATION RESULTS FOR FOUR FACIAL EXPRESSIONS USING THE HIGH EIGENVALUED COMPONENTS

Classification		Group	Predicted Group Membership				Total
			Happy	*Disgust*	*Surprise*	*Angry*	
Original cases [a]	***Count***	***Happy***	15	0	1	0	16
		Disgust	0	16	0	0	16
		Surprise	0	0	16	0	16
		Angry	1	0	0	15	16
	Percentage	***Happy***	93.8	0	6.3	0	100.0
		Disgust	0	100	0	0	100.0
		Surprise	0	0	100	0	100.0
		Angry	6.3	0	0	93.8	100
Cross-Validated cases [b]	***Count***	***Happy***	5	3	5	3	16
		Disgust	5	6	1	4	16
		Surprise	2	3	8	3	16
		Angry	4	6	1	5	16
	Percentage	***Happy***	31.3	18.8	31.3	18.8	100.0
		Disgust	31.3	37.5	6.3	25.0	100.0
		Surprise	12.5	18.8	50	18.8	100.0
		Angry	25.0	37.5	6.3	31.3	100.0
[a] 96.6 % of original group cases correctly classified							
[b] 37.5 % of cross-validated group cases correctly classified							

TABLE 6.8: CLASSIFICATION RESULTS FOR FOUR FACIAL EXPRESSIONS USING THE OPTIMAL COMPONENTS

Classification		Group	Predicted Group Membership				Total
			Happy	*Disgust*	*Surprise*	*Angry*	
Original cases [a]	***Count***	***Happy***	16	0	0	0	16
		Disgust	0	16	0	0	16
		Surprise	0	0	16	16	16
		Angry	0	0	0	16	16
	Percentage	***Happy***	100	0	0	0	100.0
		Disgust	0	100	0	0	100.0
		Surprise	0	0	100	0	100.0
		Angry	0	0	0	100	100
Cross-Validated cases [b]	***Count***	***Happy***	10	1	5	0	16
		Disgust	1	9	3	3	16
		Surprise	1	1	14	0	16
		Angry	1	3	2	10	16
	Percentage	***Happy***	62.5	6.3	31.3	0	100.0
		Disgust	6.3	56.3	18.8	18.8	100.0
		Surprise	6.3	6.3	87.5	0	100.0
		Angry	6.3	18.8	12.5	62.5	100.0
[a] 100 % of original group cases correctly classified							
[b] 67.2 % of cross-validated group cases correctly classified							

Figure 6.9 compares the classification results obtained using the two algorithmic approaches. An equal probability of group membership (*a priori*) was used in the analysis. The facial expression of disgust was confused with the other positive facial expressions. As evident in Tables 6.7 and 6.8, the two negative facial expressions (anger and disgust) could also be confused with each other. Most probably for the reason that the TIVs measured at the same muscles (Corrugator, Orbicularis Oculi Superior and Orbicularis Oculi Inferior) were engaged in expressing these two negative facial expressions.

TABLE 6.9: SUMMARY OF CANONICAL DISCRIMINANT FUNCTIONS

Function	Eigenvalue	Percentage of Variance	Cumulative percentage	Canonical Correlation
1	7.972	67.8	67.8	0.943
2	2.290	19.5	87.3	0.834
3	1.490	12.7	1000	0.774

Figure 6.9: Difference between classification success rates for the four expressions using the two algorithmic approaches

TABLE 6.10: SIGNIFICANCE OF INDIVIDUAL DISCRIMINANT FUNCTIONS

Test of Functions	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	0.014	206.272	40	0.00
2 through 3	0.122	100.955	19	0.00
3	0.402	43.792		0.008

Table 6.9 reports the canonical discriminant functions, statistical parameters and contribution of each of the two functions required to discriminate between the four facial expressions. Table 6.10 reports the significance levels of the resulting discriminant functions.

6.2.2 Analysis of classification results

The stepwise feature selection algorithm recursively selected a set of 26 most influential variables (reported in the first column of Table 6.11). These 26 effective variates were used for training the classifier and developing the discriminant functions. Statistical significance of the three functions was calculated using the mathematical relationship given in Equation 6-1. Using n (=64, 16 for each of the 4 facial expression groups) observations, p (=26 in this case) variates, G (=4) facial expression groups, k (=3) discriminant functions and the corresponding eigenvalues computed earlier, the χ^2

value for each discriminant function was calculated using Equation 6-1. The significance levels of the three discriminant functions (p<0.05) in Table 6.10 suggest the possible separation between the facial expression groups along the three discriminant functions.

Since the discriminant scores were linear combinations of the original variates, their mathematical structure provided information about the relationship between the actual variables and the discriminant functions. Table 6.11 presents the resulting structure matrix. The structural coefficients of a discriminant function are used to interpret its contribution in the between-group separation. The coefficients of a discriminator variable represent the correlation between the discriminant scores and the discriminator variables. These coefficients, given in Table 6.11, are calculated using Equation 6-2.

A closer examination of the structure matrix in Table 6.11 suggests that the first discriminant function relies on variates 12, 14, 21 and 47. So, it relies on the TIVs measured on Frontalis Pars Lateralis (12), Depressor Supercilii (14), Procerus/ Levator Labii Superioris Alaquae Nasi (21) and Levator Anguli Oris (47) for recognising and classifying the unknown thermal faces. The second discriminant function relies on variates 19, 26, 29, 31, 34, 35, 58 and 64.

TABLE 6.11: STRUCTURE MATRIX SHOWING COMPOSITION OF THE THREE DISCRIMINATING FUNCTIONS

Principal component	Function 1	Function 2	Function 3
VARIATE-21	.107(*)	-0.027	0.094
VARIATE-22	-.094(*)	0.025	-0.013
VARIATE-66	.082(*)	-0.031	0.075
VARIATE-25	-.081(*)	0.011	0.031
VARIATE-07	-.080(*)	-0.052	0.026
VARIATE-30	-.069(*)	-0.002	0.003
VARIATE-35	0.001	.287(*)	0.088
VARIATE-19	-0.003	.208(*)	-0.198
VARIATE-29	0.034	.151(*)	-0.117
VARIATE-34	0.046	.150(*)	-0.141
VARIATE-26	0.087	.127(*)	-0.026
VARIATE-58	-0.057	.125(*)	-0.037
VARIATE-31	0.053	.118(*)	-0.046
VARIATE-64	0.047	.106(*)	0.043
VARIATE-14	0.042	-.087(*)	0.07
VARIATE-38	0.055	-.066(*)	0.061
VARIATE-17	-0.119	0.188	.370(*)
VARIATE-32	0.032	-0.059	.216(*)
VARIATE-20	-0.064	0.151	.201(*)
VARIATE-15	-0.035	-0.067	.192(*)
VARIATE-54	-0.047	0.107	-.156(*)
VARIATE-40	0.035	0.109	.155(*)
VARIATE-37	-0.069	-0.072	-.125(*)
VARIATE-12	0.108	0.085	.111(*)
VARIATE-47	0.102	0.008	.111(*)
VARIATE-10	-0.066	0.038	.100(*)
* Largest absolute correlation between each variable and any discriminant function			

TABLE 6.12: DISCRIMINANT FUNCTIONS AT GROUP CENTROIDS

Facial Expression	Function 1	Function 2	Function 3
Happy	-3.485	0.657	-1.281
Disgust	2.082	-2.120	-0.675
Surprise	-1.792	-0.406	1.866
Angry	3.195	1.870	0.0897
Unstandardised canonical discriminant functions evaluated at group means.			

Thus, the second discriminant function might be relying on the measurements taken on Orbicularis Oculi Pars Orbital (19, 26, 29, 31), Orbicularis Oris (64), Procerus/ Levator Labii Superioris Alaquae Nasi (35), and Depressor Anguli Oris (58) to recognise and classify the unknown thermal faces.

It could be argued that the third discriminant function relied on thermal variations on Frontalis Pars Medialis (15) Frontalis Pars Lateralis (10, 12, 17), Orbicularis Oculi Pars Orbital (19, 20, 29), Zygomaticus Major (32), Levator Labii Superioris (34), Masseter Superficial (40), Levator Anguli Oris (47) and Platysma (54) for recognising the unknown thermal faces and allocating them to one of the facial expression groups.

As mentioned earlier, each of the three discriminant functions measured the thermal variations at one or more of the known muscles of positive and negative expressions[1]. For example, the first two discriminant functions rely on Procerus/ Levator Labii Superioris Alaquae Nasi (a muscle of frown). The second discriminant function relies on Orbicularis Oris (a muscle of excitement), and the third discriminant function relies on thermal variations along Masseter (a muscle of anger).

The three discriminant functions, in effect, examine the thermal variations that take place on the major facial muscles. Interestingly, the first discriminant function uses the variates coming from the thermal measurements on the muscles of positive experience (Orbicularis Oris and Zygomaticus) and the muscles of negative experience (Corrugator and Masseter). Other discriminant functions use a combination of the thermal variations measured on Frontalis Pars Medialis, Frontalis Pars Lateralis, Levator Labii Superioris, Levator Anguli Oris, Buccinator, Platysma, and Mentalis.

Table 6.12 shows the respective coordinates of the centroid of each facial expression group in the resulting eigenspace. The underlying span of the first discriminant function is much wider than that of the second and third discriminant functions in Table 6.12.

[1] The facial muscles considered responsible for expressing the positive and negative facial expressions of affective states are reported in section 3.2 of chapter 3.

This implies a larger contribution of the first discriminant function in the between-group separation. The underlying contribution of the second and third discriminant functions appears to be almost equal. The fourth column of Table 6.9 presents the contribution of each discriminant function in the between-group separation. Results reported in Table 6.9 graphically concur with the results reported in Table 6.12. Figure 6.10 shows the two positive and the two negative facial expressions at their respective group centroids.

6.2.3 Significance of the classification results

Once again, the practical significance of the classifier performance was calculated using Equation 5-24. The significance levels for the classification of happy, disgust, surprise and angry expressions and for the overall classification results were estimated as:

$Z^*_{happy} = 3.09$, significant at alpha-level 0.001;

$Z^*_{disgust} = 3.615$, significant at alpha-level 0.001;

$Z^*_{surprise} = 1.0328$ not significant at alpha-level 0.05;

$Z^*_{angry} = 3.09$ significant at alpha-level 0.001.

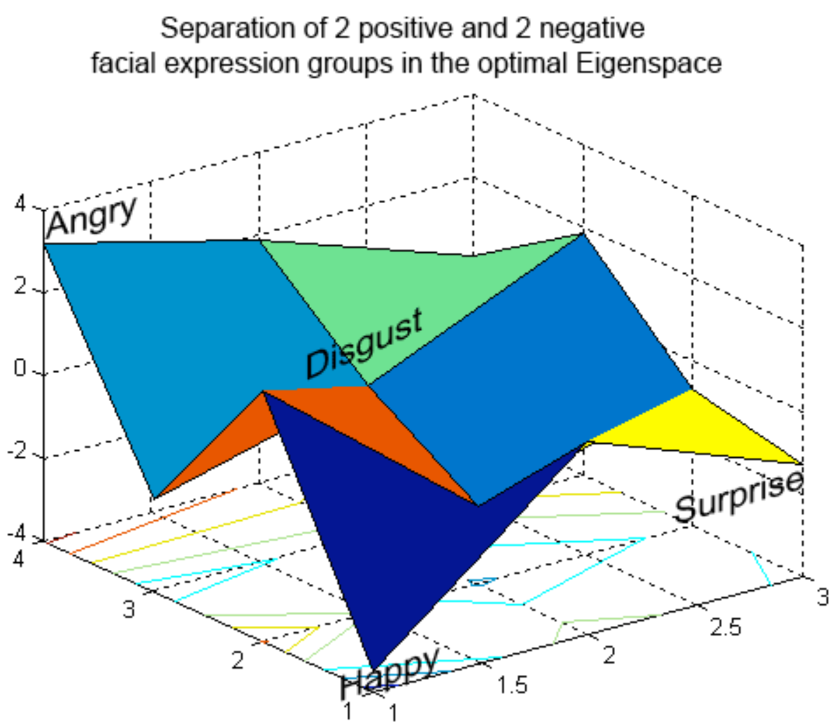

Figure 6.10: The two positive and the two negative facial expression groups at their respective group centroids

The overall significance of the classification results was estimated using Equation 5-25 and was found to be ($Z^*_{overall} = 5.59$) significant at alpha-level 0.001. The overall significance test statistic (*p<0.05)* suggested that classification results were significant.

Huberty's test was used for assessing the practical significance of the classification. Equation 5-26 was used for calculating the index ($I = {}^{43/48-16/48}\!\big/_{1-16/48} \times 100$ = 84.37). The value of *I* suggested that the employed discrimination analysis procedure resulted in 84.37% reduction in the chances of obtaining errors by chance.

It is therefore safe to assume that the TIV data gathered at 75 FTFP sites on the participant faces may help in classifying the pretended facial expressions of happiness, surprise, anger and disgust.

6.3 Conclusion

When the high eigenvalued principal components were used for classifying the neutral faces and the faces with the pretended expressions of happiness and sadness, only 45.8% of the participant faces could be correctly classified during the cross-validation tests. Using a set of optimal discriminating features, 83.8% of the 48 thermal faces were correctly classified during the cross-validation tests. This suggested that the eigenspace constructed with the optimal features might allow a better between-facial group separation as compare to the eigenspace constructed with higher eigenvalued principal components. Similarly, when the two positive and two negative facial expressions were classified, the optimised eigenspace resulted in a better between-group separation (67.20% classification success rate) as compare to the eigenspace constructed using the high eigenvalued components (37.50% classification success rate). The improved classifier performance suggests that the eigenspace optimisation may allow a compact and effective representation of the thermal features and may result in a linear division of the eigenspace.

The proposed eigenspace optimisation approach allowed better between-groups separation albeit some facial expressions were better recognised than the others. This suggested that the complexity of a constructed Gaussian space might influence the classifier performance. Earlier studies showed that a reduction in the classification space complexity would sometimes result in insufficient degrees of freedom for linear discrimination. However, there was no explicit evidence that the linear discrimination in this analysis had insufficient degrees of freedom.

It is important to note that the LDA algorithm achieved excellent classification results on the training features (principal components) but could not generalise to the new and unknown thermal faces with the same level of accuracy. These results were consistent with the classification results (Chellappa 1998; Donato et al. 1999) reported earlier. Previous studies suggested that a large training set might help correctly classify the new and unknown faces and improve the classifier performance in similar scenarios (Chellappa 1998; Donato et al. 1999).

The classifier trained using the thermal data obtained from the two positive and the two negative facial expressions performed at a lower classification success rate. This increased classification error rate might be indicative of an overlap between the thermal features and a similarity of the facial thermal features measured during the intentional expression of multiple positive and negative affective states. Probably, the two positive and the two negative facial expressions engaged the same facial muscles and hence experienced similar thermo-muscular changes, resulting in a similar thermal variation pattern.

The discriminant functions relied (for AFEC) on the thermal variations measured along those muscles that were previously known to be representative of the positive and negative facial expressions. In particular, previous EMG studies reported significant activity at the same facial muscles in the facial expression of positive and negative expressions.

When the discriminant functions were developed to classify the neutral faces and the faces with the facial expression of happiness and sadness, the first discriminant function heavily relied on thermal variations measured on Frontalis Pars Lateralis, Zygomaticus Major and Levator Labii Superioris. The second discriminant variable mostly relied on the TIV data measured at Frontalis Pars Medialis, Frontalis Pars Lateralis, Zygomaticus Major, Masseter Superficial, and Levator Labii Superioris. As mentioned in (Kall 1990), Zygomaticus Major and Mentalis are considered the muscles of positive expressions and Masseter Superficial and Labii Superioris are considered the muscles of sadness and fear.

The three discriminant functions developed during an attempt to distinguish between the facial expressions of happiness, surprise, anger and disgust exhibited a similar tendency. The first discriminant function used thermal measurements taken on the muscles of positive experience (Orbicularis Oris and Zygomaticus) and the muscles of negative experience (Corrugator and Masseter). The other principal components included in the three discriminant functions used a combination of some thermal features measured on Frontalis Pars Medialis, Frontalis Pars Lateralis, Levator Labii Superioris, Levator Anguli Oris, Buccinator, Platysma, and Mentalis. It was obvious that the three discriminant functions employed the thermal variations measured on the “already known” muscles of positive and negative expressions.

The classification results and analyses reported in the preceding paragraphs make it obvious that the facial skin temperature variations may provide useful information about the positive and negative facial expressions of affective states. It might be concluded that an appropriate facial thermal feature extraction, selection and representation and classification approach may allow person-independent classification of the facial expressions of affective states.

Chapter 7

CLASSIFICATION OF BASIC FACIAL EXPRESSIONS

As mentioned earlier in chapters 2 and 3, previous non-vision based automated facial expression classification (NVAFEC) systems appeared to be more focused on distinguishing between the positive and negative affective states. This chapter reports an attempt to classify the neutral and six basic facial expressions using the facial thermal features. As the classification results suggest, a Gaussian mixture model that had seven components; the neutral expression and the six common facial expressions, resulted in a complex Gaussian space. Despite that, this part of the investigation exhibits the viability of using the facial thermal features for classifying the facial expressions in a complex decision space.

7.1 Initial analysis

In addition to the initial data analyses reported in chapter 4, Andrews' curves for the TIV data were plotted to further examine the clusters of seven facial expression groups and explore the underlying variance in the data. Each *p*-variate observation in the data set was represented by a function plotted over the range $-\pi < t < +\pi$ such that for the *n*th observation, the function $f(t)$ was defined as

$$f(t) = z_1/\sqrt{2} + z_2 \sin t + z_3 \cos t + z_4 \sin t + z_5 \cos t + \ldots\ldots z_{17} \cos t + z_{18} \sin t + \ldots + \ldots\ldots \quad \text{7-1}$$

where, $z_1, z_2, \ldots\ldots z_n$ represented the observed numerical values of the TIV data. The Andrews' curves are shown (as $z_1, z_2, \ldots\ldots z_n$) in Figure 7.1. Excluding the first term in Equation 7-1, the function $f(t)$ is a mixture of sine and cosine waves and produces a (visually) representative wave pattern depending on the observed values of the *p* variables.

An Andrews' plot is based on the distance between the two functions that are defined by the expression

$$\int_{-\pi}^{\pi} \left[f_p(t) - f_q(t) \right]^2 dt . \quad \text{7-2}$$

Thus the similarities in the Andrews' plots in Figure 7.1 exhibit the similarities between the underlying functions. The observations that were actually closer in a high-dimensional space had a similar wave pattern in an Andrews' graph space (Chatfield and Collins 1995; Everitt and Dunn 1991; Jolliffe 2002).

A problem that frequently arises with this technique is that only a limited number of observations may be plotted on the same diagram. A large number of observations make the diagram confusing and less useful. Methods of overcoming this problem have been suggested in the literature (Jolliffe 2002). For example, separate plots, one for each set or cluster are plotted and compared against the others. Also, selected quantities of the *n* values of the function *f(t)* are plotted along with the curves of the selected individual observations (Everitt and Dunn 1991). In many earlier studies, Andrews' curves for each class were clustered and plotted in a separate window for further visual analysis (Jolliffe 2002).

Andrews' curves for the seven clusters (facial expressions) of the 16-participants' facial thermal data in Figure 7.1 reveal that the raw TIV data provide little information about the thermal differences between the facial expression groups. The facial expression group clusters had a similar waveform. The profiles of Andrews' curves being similar do not show any (large) between-group variation. However, the within-group homogeneity, even on the basis of the raw TIV data, was evident in each cluster of facial expression in Figure 7.1. The within-group curves fall into moderately narrow bands suggesting an adequate level of similarity in the thermal representation of the individuals' facial expressions. The within-group homogeneity was considered a positive sign. Its presence encouraged exploring the possibilities of distinguishing between the seven facial expressions using the facial thermal features (Chatfield and Collins 1995; Everitt and Dunn 1991; Jolliffe 2002).

Few distant curves identified in the neutral facial expression cluster window in Figure 7.1 belong to the participants with facial hair. The TIVs measured on the faces of the three participants (having facial hair) were slightly different than the ones measured on other participant faces. Facial hair, having a different emissivity ratio (ε) might have resulted in different TIV measurements on those parts of the face that were covered under the hair. The TIVs measured on these parts of the face were therefore different than the TIVs measured on the uncovered parts of the face. These three participants were kept in the experiments to examine the effect of facial hair on the classifier performance and the validity of the resulting discriminant rules.

Some notable differences in the clusters of Andrews' curve profiles are visible between the range $-\pi/2$ and $+-\pi/2$ in Figure 7.1. The curves plotted using the raw TIV data did not provide a practically useful measure of the between-group variation and warrant formation of a set of uncorrelated variables. Such variables are usually formed using a technique like PCA to develop the discriminant rules (McLachlan 2004).

7.2 Classifier construction

The algorithmic approach mentioned in chapter 5 was used to develop a facial expression classifier for distinguishing between the seven facial expressions. Recorded TIV data were transformed into uncorrelated linear combinations of the input vectors using the principal component analysis. These vectors (principal components) are exhibited as a scree plot in Figure 7.2. Only 32 of the 75 derived principal components significantly contributed to the variance in the TIV data. Of these, the first 20 principal components represented over 90% variance in the data albeit there was no reason to believe that only these 20 principal components would contribute to the classification of the facial expressions. Figure 7.3 exhibits the possible between-group separation in the first two-principal component eigenspace.

Comparing Figures 7.1 and 7.3 would reveal that transforming the raw TIV data into the principal components allowed a better between-facial expression group separation. The negative facial expressions (sadness, disgust and fear) are well separated within the 2-principal component eigenspace in Figure 7.3. This trend was consistent with the previous studies that employed either FACS or facial EMG readings to classify the negative facial expressions. Negative facial expressions in previous investigations could be easily separated from each other as compare to the positive facial expressions (Ekman et al. 1993; Kall 1990).

Figure 7.1: Andrews' curve for the seven facial expressions plotted using the raw TIV data. The function *t* spans along the abscissa and ranges between -π and +π. Corresponding values of the TIVs are plotted along the ordinate. The curve profiles for each particular facial expression group show the within-group homogeneity. The similarities between the waveforms of various facial expressions suggest that the facial expressions cannot be classified using the raw TIV data.

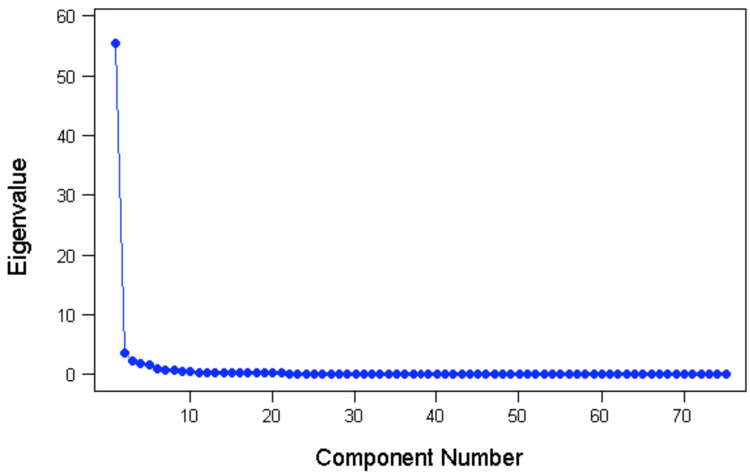

Figure 7.2: Contribution of the 75 principal components in the underlying variance of the 7-facial expression TIV data

Figure 7.3: Between-group separation of the seven facial expressions in a 2-principal component eigenspace

The neutral faces and the faces with intentionally expressed happiness and surprise (positive expression) seem to be comparatively difficult to separate from the other facial expressions in a 2-principal component eigenspace. This pattern of separation was also observed in a previous investigation that Kall (1989) carried out. As reported in chapter 6, the first principal component was the negative index of a thermal face and hence provided an overall thermal measurement of the face. The second principal component provided an account of the facial thermal characteristics as it added the TIVs on Orbicularis Oculi Pars Orbital and Levator Labii Superioris Alaquae Nasi and subtracted the TIVs on all other facial muscles from this added value. One could infer from Figure 7.3 that the uncorrelated linear variables provided some useful information about the variance in the TIV data to classify the facial expressions of affective states.

In the next step, the derived principal components were analysed using the stepwise elimination and feature selection algorithm described earlier in chapter 5. The iterative increase in the value of F ratio ($F = |S_B|/\ |S_W|$) is obvious in Figure 7.4. The optimal feature selection algorithm recursively selected a set of 22 principal components for training the classifier. These 22 selected principal components were not the highest eigenvalued principal components. The feature selection algorithm discovered a new and different set of most discriminating components. Only few of the high eigenvalued principal components were included in this new set of optimal principal components. The algorithm discovered many low-eigenvalued principal components that were also able to contribute to the between-group separation.

The TIV data were normally distributed, the facial expressions groups had a similar variance structure, and each facial expression group had the same numbers of participants. Hence, each facial expression group could be given an equal *a priori* in the analysis during the linear discriminant analysis.

Table 7.1 exhibits a summary of the canonical discriminant functions showing the contribution of each discriminant function in distinguishing between the intentional facial expressions.

Table 7.2 shows the chi-square values and the respective significance levels of the discriminant functions.

TABLE 7.1: SUMMARY OF CANONICAL DISCRIMINANT FUNCTIONS

Function	**Eigenvalue**	**Percentage of Variance**	**Cumulative percentage**	**Canonical Correlation**
1	1.759	32.0	32.0	0.798
2	1.399	25.4	57.4	0.764
3	1.153	20.9	78.3	0.732
4	0.492	8.9	87.3	0.574
5	0.358	6.5	93.8	0.513
6	0.343	6.2	1000	0.505

TABLE 7.2: SIGNIFICANCE OF INDIVIDUAL DISCRIMINANT FUNCTIONS

Test of Functions	Wilks' Lambda	Chi-square	df	Sig.
1 through 6	0.026	352.914	132	0.00
2 through 6	0.071	254.979	105	0.00
3 through 6	0.171	170.517	80	0.00
4 through 6	0.368	96.573	57	0.001
5 through 6	0.549	57.942	36	0.012
6	0.745	28.445	17	0.040

Figure 7.4: Recursive stepwise selection of optimal components with an increasing *F* ratio (=S_B/S_W). The line connecting the first and the last iterations exhibits a significant increase in the F-ratio.

TABLE 7.3: STRUCTURE MATRIX SHOWING COMPOSITION OF THE TWO DISCRIMINATING FUNCTIONS

Contributing variables	Function 1	Function 2	Function 3	Function 4	Function 5	Function 6
VARIATE-19	0.233(*)	0.065	-0.191	0.017	0.185	0.004
VARIATE-18	-0.020	0.268(*)	-0.062	0.088	-0.059	0.183
VARIATE-23	0.153	-0.227(*)	0.052	-0.022	-0.128	-0.184
VARIATE-07	-0.98	0.152(*)	0.064	0.028	-0.138	-0.151
VARIATE-15	0.102	0.141	0.388(*)	-0.290	0.158	-0.238
VARIATE-26	-0.160	-0.076	-0.271	-0.023	0.160	-0.134
VARIATE-37	0.110	-0.082	0.174(*)	0.058	-0.051	0.023
VARIATE-21	-0.060	0.044	0.081	0.366(*)	-0.212	-0.256
VARIATE-22	0.216	0.022	0.026	0.359(*)	0.117	0.024
VARIATE-20	-0.156	0.156	0.016	0.299(*)	0.038	0.077
VARIATE-27	0.059	0.171	-0.088	-0.249	0.111	0.018
VARIATE-11	-0.108	-0.242	-0.004	-0.245	-0.022	0.153
VARIATE-41	-0.088	-0.067	0.286	-0.008	0.400(*)	0.127
VARIATE-30	0.170	0.092	0.098	-0.274	-0.392(*)	0.138
VARIATE-14	0.184	0.101	0.093	0.194	0.312(*)	-0.018
VARIATE-38	-0.038	0.084	-0.143	-0.214	0.263(*)	0.033
VARIATE-09	-0.026	-0.077	0.091	0.039	0.263	0.420(*)
VARIATE-24	-0.146	0.038	0.058	0.098	-0.096	0.408(*)
VARIATE-31	-0.164	-0.129	0.251	0.037	0.061	0.264(*)
VARIATE-75	-0.030	-0.150	0.069	0.073	-0.214	0.260(*)
VARIATE-13	0.209	0.011	0.173	0.024	-0.151	0.253(*)
* Largest absolute correlation between each variable and any discriminant function						

TABLE 7.4: STANDARDISED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

Contributing variables	Function 1	Function 2	Function 3	Function 4	Function 5	Function 6
VARIATE- 07	-0.254	0.342	0.129	0.039	-0.175	-0.190
VARIATE- 09	-0.065	-0.166	0.177	0.052	0.323	0.510
VARIATE- 11	-0.261	-0.510	-0.007	-0.320	-0.026	0.180
VARIATE- 13	0.505	0.024	0.325	0.031	-0.179	0.279
VARIATE- 14	0.446	0.213	0.176	0.255	0.372	-0.021
VARIATE- 15	0.219	0.262	0.648	-0.335	0.166	-0.247
VARIATE- 18	-0.362	0.638	0.080	-0.176	-0.121	0.014
VARIATE- 19	0.557	0.135	-0.355	0.021	0.218	0.004
VARIATE- 20	-0.384	0.334	0.031	0.397	0.046	0.093
VARIATE- 21	-0.147	0.094	0.156	0.487	-0.257	-0.307
VARIATE- 22	0.518	0.045	0.049	0.465	0.138	0.028
VARIATE- 23	0.371	-0.480	0.099	-0.028	-0.153	-0.217
VARIATE- 24	-0.365	0.082	0.112	0.132	-0.117	0.494
VARIATE- 26	-0.384	-0.157	-0.505	-0.029	0.189	-0.156
VARIATE- 27	0.150	0.377	-0.174	-0.340	0.138	0.023
VARIATE- 28	-0.050	0.573	-0.119	0.118	-0.071	0.219
VARIATE- 30	0.400	0.188	0.180	-0.349	-0.454	0.159
VARIATE- 31	-0.386	-0.264	0.463	0.047	0.071	-0.303
VARIATE- 37	0.283	-0.183	-0.351	0.082	-0.064	0.029
VARIATE- 38	-0.096	0.186	-0.285	-0.295	0.330	0.041
VARIATE- 41	-0.205	-0.137	0.523	-0.010	0.462	0.145
VARIATE- 75	-0.077	-0.333	0.137	0.100	-0.269	0.323

Table 7.3 reports the structure matrix showing the pooled within-group correlation between the discriminant variables and the canonical functions.

Table 7.4 shows the standardised canonical discriminant function coefficients. These coefficients were used to calculate the predicted group membership of the unknown thermal faces. The predicted group membership of an unknown face was iteratively calculated using these scores and the coefficients of discriminant functions.

Table 7.5 shows the classification success rate and the confusion matrix that resulted when the high eigenvalued principal components were used to train the classifier. A higher error rate e^{j}_{geneal} (= 61.6%) was observed when the highest eigenvalued principal components were used for recognition and classification.

Table 7.6 shows the classification success results and the confusion matrix observed when the optimal principal components were used to train the classifier. As evident in Table 7.6 the error rate significantly dropped to e^{j}_{geneal} (= 42.9%) when the optimal features were used.

7.3 Classification error analysis

The linear discriminant algorithm discovered the most influential variates for developing the discriminant functions using the optimal principal components. It discovered 22 most effective variates for recognising and classifying the

TABLE 7.5: CLASSIFICATION RESULTS WITH THE HIGHEST EIGENVALUED COMPONENTS

Classification		Group	Predicted Group Membership							Total
			Neutral	*Happy*	*Sad*	*Disgust*	*Surprise*	*Angry*	*Fear*	
Cross-Validated cases [b]	***Count***	***Neutral***	5	2	3	1	3	2	0	16
		Happy	1	3	4	3	3	2	0	16
		Sad	4	1	6	0	2	1	2	16
		Disgust	0	0	1	8	1	5	1	16
		Surprise	0	2	1	2	6	3	2	16
		Angry	2	3	3	1	1	4	2	16
		Fear	1	0	0	1	3	0	11	16
	Percentage	***Neutral***	31.3	12.5	18.8	6.3	18.8	12.5	0	100.0
		Happy	6.3	18.8	25.0	18.8	18.8	12.5	0	100.0
		Sad	25.0	6.3	37.5	0	12.5	6.3	12.5	100.0
		Disgust	0	0	6.3	12.5	37.5	18.8	12.5	100.0
		Surprise	0	12.5	6.3	6.3	6.3	25.0	12.5	100.0
		Angry	12.5	18.8	18.8	6.3	6.3	25.0	12.5	100.0
		Fear	6.3	0	0	6.3	18.8	0	68.8	100.0
[b] 38.4 % of cross-validated group cases correctly classified										

TABLE 7.6: CLASSIFICATION RESULTS OBSERVED USING THE OPTIMAL FEATURES

Classification		Group	Predicted Group Membership							Total
			Neutral	*Happy*	*Sad*	*Disgust*	*Surprise*	*Angry*	*Fear*	
Cross-Validated cases [b]	***Count***	***Neutral***	5	2	3	2	3	1	0	16
		Happy	2	10	0	0	3	0	1	16
		Sad	4	00	11	0	0	1	0	16
		Disgust	2	1	0	10	0	1	2	16
		Surprise	3	0	2	1	7	2	1	16
		Angry	3	3	0	0	1	7	2	16
		Fear	0	1	0	0	1	0	14	16
	Percentage	***Neutral***	31.3	12.5	18.8	12.5	18.8	6.3	0	100.0
		Happy	12.5	62.5	0	0	18.8	0	6.3	100.0
		Sad	25.0	0	68.8	0	0	6.3	0	100.0
		Disgust	12.5	6.3	0	62.5	0	6.3	12.5	100.0
		Surprise	18.8	0	12.5	6.3	43.8	12.5	6.3	100.0
		Angry	18.8	18.8	0	0	6.3	43.8	12.5	100.0
		Fear	0	6.3	0	0	6.3	0	87.5	100.0
[b] 57.1 % of cross-validated group cases correctly classified										

unknown thermal faces. The mathematical compositions of these variates are listed in Tables 7.3 and 7.4.

A set of six discriminant functions was developed by the linear discriminant algorithm to classify the thermal images into seven facial expressions. The statistical significance levels (p<0.05) of the discriminant functions, given in Table 7.2, were calculated using Equation 6-1. Results reported in Tables 7.1, 7.2 and 7.3 suggested some significant differences between the facial thermal features pertaining to the seven facial expressions.

The mathematical structure of a discriminant function also provides useful information about the relationship between the actual variables and the discriminant functions. The numerical coefficients of the discriminant functions, calculated using Equation 6-2, allowed interpreting the contribution of each variable in the formation of discriminant functions. The coefficients of a given discriminator were the coefficients of correlation between the discriminant scores and the discriminator variables.

A closer examination of the structure matrix in Table 7.3 suggests that the first discriminant function relies on variates measured on Frontalis (11, 13, 14, 15), Orbicularis Oculi (19, 22, 23, 26, 30, 31) and Levator Labii Superioris (37). The first discriminant function therefore measured the thermal features on the upper part of the face to classify the unknown thermal faces.

The second discriminant function relies on the variates measured on Frontalis (7, 11, 13, 15), Orbicularis Oculi Pars Orbital (18, 20, 23, 27, 31), and Mentalis (75) for allocating the unknown thermal faces to a facial expression group.

The third discriminant function uses the variates measured on Frontalis (13, 15), Orbicularis Oculi Pars Orbital (18, 19, 20, 23, 26, 27, 31), Levator Labii Superioris (37, 38), and Masseter superficial (41) for recognition and classification of the new and unknown faces.

The fourth discriminant function employs the variates measured at Frontalis Pars Lateralis (11), Frontalis Pars Medialis (15), Orbicularis Oculi Pars Orbital (20, 21, 22, 27, 30) and Levator Labii Superioris (38) for discerning between the facial expressions.

The fifth discriminant function uses the variates measured on Frontalis (7, 9, 13, 14, 15), Orbicularis Oculi Pars Orbital (19, 21, 22, 23, 26, 27,30), Levator Labii Superioris (38), and Masseter superficial (41) for recognition and classification of the unknown faces.

Sixth and last discriminant function relies on the variates measured around Frontalis (7, 9, 11, 13, 15), Orbicularis Oculi Pars Orbital (18, 21, 23, 24, 26, 30, 31), Masseter superficial (41) and Mentalis (75) for classifying the unknown faces.

Interestingly, the major facial muscles along which the variates could be aligned were physically located around the upper part of the face. It could be assumed that the six discriminant functions exploited the thermal variations that took place on the upper parts of the face for recognition and classification of the seven facial expressions. It might be useful to state that the FACS based AFEC systems and studies on facial EMG readings found the muscular movements on the upper parts of the face helpful for automated facial expression classification (Kall 1990; Puri et al. 2005). Some successful

AFEC systems (cited earlier in chapters 2 and 3) also relied on the signals extracted from the upper parts of the face for facial expression classification.

Figure 7.5 shows the six discriminant functions at their respective group centroids. The varying span of the six discriminant functions implies varying influence of each discriminant function in the between-group separation and classification of unknown thermal images.

The leave-one-out cross validation tests was also invoked at each stage of this investigation. The validation results shown in the confusion matrix in Table 7.6 revealed the complexity of the underlying eigenspace. For example, the neutral faces were (equally) confused with the intentional expressions of sadness and surprise. In a 2-principal component eigenspace shown in Figure 7.3 above, the overlap between neutral and surprised faces was obvious. The similarities (and overlap) between the thermal characteristics of the neutral faces and the facial expressions of sadness and surprise were also obvious in the eigenspace. The confusion matrix in Table 7.6 also showed some similarities between the facial expressions of happiness and surprise.

Given the small sample size and a large number of measured thermal features, the overall error rate (42.9%) observed during the leave-one-out cross validation tests could be considered encouraging. The observed error rate was suggestive of (1) the potential effectiveness of the thermal features for AFEC, and (2) the aptness of the employed algorithmic approach.

7.4 Significance of the classification results

The practical significance of the classifier performance was calculated using Equations 5-24 and 5-25. The statistical significance levels for classification results are given in Table 7.7. The overall significance test statistic (*$p<0.01$*) suggested that the classification results were significant.

Equation 5-26 was used to determine the extent to which the classification results were better than the ones that could be observed by chance. The estimated index $I = {}^{64/112 - 16/112}\!\big/_{1-16/112} \times 100$ = 50.0 suggested that the employed discriminant rules reduced the probability of by chance getting the classification error rate by 50.0%.

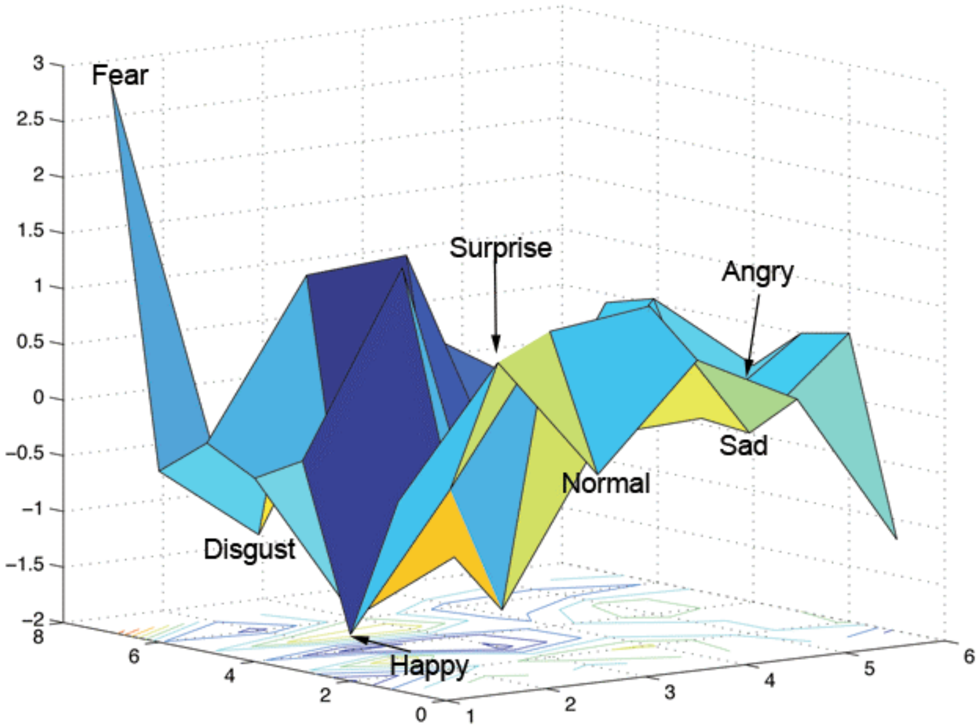

Figure 7.5: The facial expression groups at their respective group centroids in a 6-discriminant function eigenspace

TABLE 7.7: SIGNIFICANCE OF THE CLASSIFICATION RESULTS

Statistic	Value	Level of significance (α)
Z^*_{normal}	5.933	0.001
Z^*_{happy}	2.45	0.0071
Z^*_{sad}	2.236	0.0132
$Z^*_{disgust}$	2.45	0.0071
$Z^*_{surprise}$	3.0	0.0013
Z^*_{angry}	3.0	0.0013
Z^*_{fear}	1.41	0.0655
$Z^*_{overall}$	6.93	0.001

7.5 Conclusion

The classification results presented in Table 7.6 reveal certain similarities between the thermal characteristics of a neutral face and the six facial expressions. During the cross validation tests, the neutral faces were often confused with the pretended sad (18.8%) and surprised (18.8%) expressions. (Cohen et al. 2003) attempted person-independent facial expression classification in a vision-based AFEC system and observed a similar classification and confusion pattern. Neutral faces in their investigation were confused with expressions of sadness and surprise. Another vision-based AFEC capable system that (Calder et al. 2001) developed did confuse some 23% of neutral faces with the expression the sad faces and confused around 07% neutral faces with the faces of surprise. Hence some parallels might be drawn between the confusion patterns observed in this investigation and the ones observed in some earlier vision-based AFEC systems.

The pretended expressions of happiness in this investigation were confused with the neutral faces (12.5%) and with the faces of surprise (18.8%). The system (Calder et al. 2001) developed confused only 02% of the neutral faces with the faces expressing happiness. (Huang and Huang 1997) also used visual cues for AFEC and claimed that their system confused only 02% of neutral faces with the facial expression of happiness.

The pretended expression of anger was often confused with the neutral (18.8%) faces. The angry faces were equally confused with the happy (18.8%) faces. In a system that (Cohen et al. 2003) reported, around 2.04% expressions of anger were confused with neutral faces and 4.76% of angry faces were confused with the happy faces. In another system (Huang and Huang 1997) designed, only 04% faces with the expression of anger were confused with the faces of smile. This suggests that the facial expressions of anger and happiness cause different musculo-thermal and hæmodynamic activities on the face. Consequently, these two facial expressions result in a different temperature variation pattern along the major facial muscles. Thus, the linear division of eigenspaces did not allow much overlap and the classifier could well separate the two facial expressions.

Interestingly, the classification of the pretended expression of fear was highly successful and was rarely confused with the other pretended facial expressions.

Though the confusion patterns observed in this study had some resemblance with the confusion patterns observed in the earlier vision-based AFEC systems, the observed classification success rates were comparatively lower than the ones reported in (Calder et al. 2001; Cohen et al. 2003; Huang and Huang 1997). However, our classification results cannot be directly compared with those of the vision-based AFEC systems.

It could be inferred that the facial expressions of affective states would influence the facial thermal features. Consequently, the facial skin temperature at the identified physical locations (FTFPs) changed with a change in facial expression. The analyses reported in the preceding sections suggest that the non-contact thermal infrared measurement of facial skin temperature might help in recognising and classifying the basic facial expressions. Statistical and practical significance of the observed classification results suggest that the employed algorithmic approach and the developed discriminant rules were effective in recognising and classifying the facial expression of basic affective states.

The classifier, when trained with the high eigenvalued principal components performed less effectively than the classifier trained with a set of optimal discriminating features. It might therefore be inferred that the eigenspace constructed with the optimal features would allow better between-facial group separation than the eigenspace constructed with some higher eigenvalued principal components. It might also be inferred that the proposed algorithmic approach allowed a developing an effective set of discriminant rules.

Chapter 8

CLASSIFICATION OF COVERED AND OCCLUDED FACES

Studies reported earlier in chapters 2 and 3 suggest that the vision based automated facial expression classification (AFEC) systems perform well when either an entire face or a complete frontal view of the face is available for feature extraction. Previous investigations reported that when it was not possible to extract features from an entire face for factors such as facial hair, glasses, lighting conditions, pose and occlusion, the vision-based AFEC systems could not perform well.

The existing literature did not cite any significant work on investigating the efficacy and relevance of the non-vision based automated facial expression classification (NVAFEC) systems when an entire face was not available for feature extraction. The NVAFEC capable systems, therefore, have yet to be tested for their effectiveness in situations when the faces are covered or occluded. It was therefore considered prudent to investigate the possibilities of recognising facial expressions using thermal data gathered from selected parts of the face assuming that the other parts of the face were either covered or occluded. This work, to the best of this author's knowledge and belief, is the first attempt to investigate the possibilities of classifying facial expressions using the thermal data gathered from selected parts of the face.

This chapter first reports a facial muscle grouping approach for extracting facial thermal features from the selected parts of the face. Details of the initial analysis and classifier construction are then presented. Finally the classification results are discussed and analysed.

8.1 Facial muscles grouping and preliminary data analysis

Human face is considered a complex and information-rich part of the body. Researchers have employed different approaches for dividing the human face into anatomical and muscular regions for extracting features of interest from the face. To examine the potential of classifying the facial expressions of affective states using the TIV data gathered from the selected parts of the face, division of a thermal face into muscular regions was vital. Based on the reviewed literature reported in chapters 2 and 3, the human face was divided into four facial regions. Each facial region had a number of facial muscles included in it. This division of a face allowed grouping various facial muscles together. The division of a face and the facial muscle grouping employed in this work were similar to the ones (Huang and Yan 2002) used in their investigation. They used the facial anatomy and a representation of the facial mesh geometry to model the face and simulate the facial muscle features. Their work helped in dividing the face into various regions for grouping the facial muscles in this investigation. The facial muscle grouping and FTFPs within each of facial region are visible in Figure 8.1.

Figure 8.1: Division of a face into four (muscular) facial regions

As evident in Figure 8.1, the forehead was labeled as facial region 1 (R1). The areas around the eyes and the cheeks were (jointly) considered facial region 2 (R2). The facial region 3 (R3) included parts of the face around the mouth. The facial areas around the chin were considered facial region 4 (R4).

One major advantage of using this muscle grouping was that each region could be separately examined to evaluate the effect of facial hair, glasses, pose, lighting conditions or objects such as microphone on the AFEC functionality. For example, if parts of the forehead (R1) and the area around eyes and nose (R2) were occluded under the glasses, the TIV data measured around the mouth (R3) and chin (R4) could be used for AFEC and AAR. Similarly, if the facial regions R3 and R4 were covered under the facial hair, the TIV data from the facial regions R1 and R2 could be used for the facial expression classification.

Table 8.1 presents the major facial muscles and the numbers of FTFPs located within each facial region. The TIV data in each facial region were analysed for the assumptions of normal distribution and the symmetry of covariance matrices. The 17 TIVs in the facial region R1 (forehead) were considered a unique vector. Having 34 TIVs, the facial region R2 (around eyes and cheeks) was represented as another vector. Similarly, the TIVs in the facial region R3 (around lips) were represented as a vector in a 5-dimensional space and the TIVs in the facial region R4 (around the chin) were represented as a 19-dimensional vector.

TABLE 8.1: FACIAL REGIONS, MUSCULAR GROUPING AND THE FTFP SITES WITHIN EACH FACIAL REGION

Facial region	**Muscles in the region**	**FTFPs in the region**
R1	Frontalis, pars medialis Frontalis, inner center edges of pars medialis and pars lateralis Frontalis, pars lateralis Procerus/ Levator, labii superioris alaquae nasi	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 21
R2	Depressor, supercilii Orbicularis Oculi, pars orbital Levator, labii superioris alaquae nasi Levator, labii superioris Levator, anguli oris Zygomaticus major Masseter, superficial Buccinator	14, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 50, 56, 57
R3	Orbicularis Oris	45, 51, 64, 65, 66
R4	Risorious/ Platysma Depressor Labii Inferioris Mentalis	52, 53, 54, 55, 58, 59, 60, 61, 62, 63, 67, 68, 69, 70, 71, 72, 73, 74, 75

The significant test result [F (6,10)=3.842, p<0.05] for the TIV data measured on the entire face suggested homogeneity of the TIV data. However, the test statistics for the TIV data from various facial regions, analysed separately, were insignificant. The test statistics for R1 [F (6,10)=3.102, p>0.05], R2 [F (6,10)=1.896, p>0.05], R3 [F (6,10)=2.568, p>0.05], and R4 [F (6,10)=1.607, p>0.05] did not suggest similarity of variance. Figure 8.2 shows the estimated mean values for the thermal data within the selected facial regions. A comparison of the estimated mean temperature values in Figure 8.2 would reveal the changing patterns of thermal features within the various facial regions with a change in facial expressions.

8.2 AFEC using TIVs measured on forehead (R1), around eyes and on cheeks (R2)

An attempt was made to classify the facial expressions using the TIV data measured on the forehead, around eyes and on the cheeks. The facial expression classifier was constructed using the algorithmic approach described earlier in chapter 5. The TIV data were first transformed into uncorrelated linear variables and a set of 51 principal components was derived using the TIV data measured from within regions 1 and 2.

Figure 8.2: The mean facial skin temperature for the 7 facial expressions estimated using the TIV data measured at the FTFP sites within each facial region

First 24 of the 51 transformed principal components significantly contributed to the variance in data. However, there was no reason to believe that these 24 principal components would significantly contribute to the between-facial expression group separation. Figure 8.3 exhibits a scree plot showing the contribution of each principal component in the underlying variance of the TIV data collected from the FTFP sites in facial regions 1 and 2.

In the next step, the derived principal components were analysed to discover a set of optimal principal components. Only 18 principal components were found to be significantly contributing to the between-group separation. Some of these principal components did not have a higher eigenvalue.

Figure 8.4 exhibits the Andrews' curves for the individual clusters of 7 facial expression groups drawn using the 18 optimal principal components. These18 principal components were assumed to provide the most useful information for explaining the underlying variation in the TIV data measured around the forehead, eyes and cheeks. The clusters of observations, each shown in a different window, were different. These obvious visual differences in the clusters of the TIV data were indicative of the between-facial expression group separation.

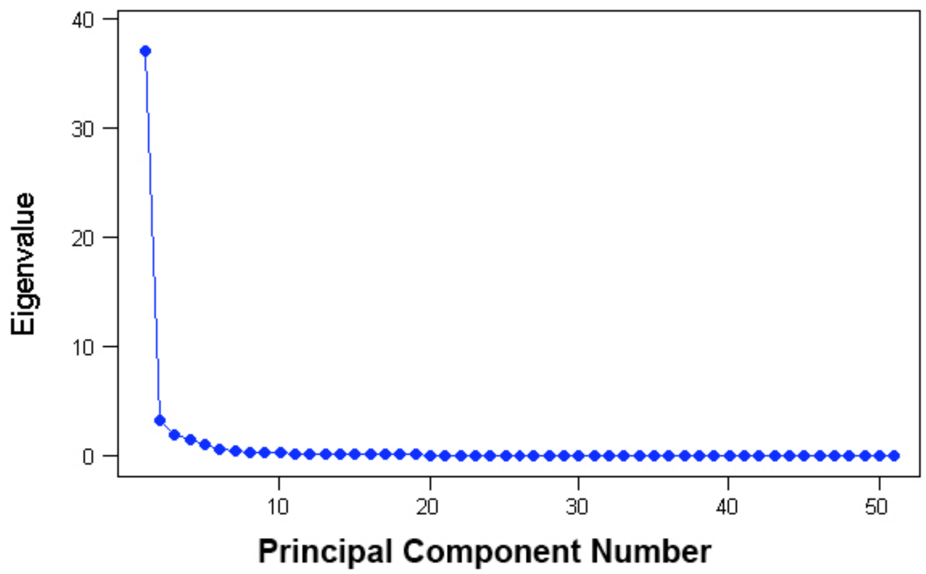

Figure 8.3: Contribution of the 75 principal components in the 7 facial expression group variance. The principal components were derived using the TIV data gathered from the facial regions R1 and R2

Figure 8.4: Andrews' Curves for the 7 facial expressions drawn using the PCA scores derived from the TIV data gathered from the facial regions R1 (forehead) and R2 (around eyes and cheeks). Similar profiles of the curves within each cluster represent the within-group homogeneity. Apparent minor differences in the profiles of the 7 clusters of facial expressions exhibit the between-group variance.

Figure 8.5: Mixed plot of Andrews' Curves for the 7 facial expressions drawn using the TIVs measured on the facial regions R1 (forehead) and R2 (around eyes and the cheeks). The between-group variance is obvious through the shift in Andrews' curves along the ordinate. The within-group homogeneity is also obvious from the narrow-band of Andrews' curves in each facial expression group.

Andrews' curves for the 7 separate clusters of the 16-participant facial expressions, shown in Figure 8.4 reveal the underlying differences between the facial expression groups. Each facial expression group had a slightly different profile of the Andrews' curves along the x-axis. This highlighted the between-group variance (obvious within the ellipses in Figure 8.5) in the TIV data that could help in classifying the facial expressions. The profiles of the Andrews' curves, being somewhat similar, do not show an apparently large between-group variation. However, the curves within each cluster window in Figures 8.4 and 8.5 fall into moderately narrow bands suggesting an adequate level of within-group homogeneity (Chatfield and Collins 1995; Everitt and Dunn 1991; Jolliffe 2002).

The three distinct curves, highlighted with arrows in the neutral face cluster in Figure 8.4 belong to the three participants with facial hair. The TIVs measured on the faces of these three participants were slightly different than those measured for others, particularly within the facial regions covered under hair. It seems likely that the facial hair, having a different emissivity ratio (ε) might have resulted in some slightly different TIV measurements on the parts of the face covered under the facial hair.

Some significant differences in the Andrews' curve profiles are visible between the range $-\pi$ and $-3\pi/4$ in Figure 8.4. A mixed plot of the 7 expression groups is also shown in Figure 8.5. The mixed plot effectively exhibits the prevailing "within-group" similarities and the "between-group" differences in the TIV data.

The discriminant analysis algorithm was invoked on the optimal thermal features selected from within the principal components. The algorithm recursively selected the most appropriate variates for developing the discriminant functions and constructing a classifier. The invoked algorithm picked a set of 18 most effective variates for developing the discriminant functions. As evident in Table 8.2, only 3 of 6 discriminant functions were significant (p<0.05) suggesting a low probability of the between-facial expression group separation.

TABLE 8.2: SIGNIFICANCE OF THE DISCRIMINANT FUNCTIONS DERIVED USING THE TIV DATA FROM THE FACIAL REGIONS R1 AND R2

Test of Functions	Wilks' Lambda	Chi-square	df	Sig.
1 through 6	0.059	278.920	108	0.00
2 through 6	0.185	166.040	85	0.00
3 through 6	0.360	100.586	64	0.002
3 through 6	0.570	55.351	45	0.139
3 through 6	0.794	22.689	28	0.748
6	0.940	6.121	13	0.942

TABLE 8.3: THE STRUCTURE MATRIX FOR THE SIX DISCRIMINATING FUNCTIONS DERIVED USING THE TIV DATA FROM FACIAL REGIONS R1 AND R2

Variate	Function 1	Function 2	Function 3	Function 4	Function 5	Function 6
Variate-15	-.481(*)	-.279	.020	.138	.014	.052
Variate-22	.020	.376(*)	.051	.062	.167	.014
Variate-32	.077	-.023	.447(*)	.099	.005	-.029
Variate-17	.031	-.241	.304(*)	.281	-.068	.204
Variate-18	.204	-.179	.227(*)	.178	-.103	.013
Variate-27	.063	.038	-.015	-.462(*)	.420	.058
Variate-19	-.077	.039	.183	-.319(*)	-.108	.195
Variate-20	.120	.162	-.113	.296(*)	-.003	-.132
Variate-13	-.041	.060	.379	.016	.473(*)	.001
Variate-12	.077	-.012	-.178	.428	.428(*)	.123
Variate-28	-.111	.138	.174	.057	.332(*)	-.123
Variate-25	.121	.140	.054	-.014	-.214(*)	-.088
Variate-21	-.042	.220	.165	.006	-.131	.624(*)
Variate-29	-.033	.208	.234	-.055	-.279	-.390(*)
Variate-11	-.032	.292	-.124	.165	-.094	.328(*)
Variate-38	.164	-.180	-.090	-.185	.139	.327(*)
Variate-4	.143	-.053	.048	-.066	-.124	.274(*)
Variate-8	.175	-.071	.043	-.094	.104	-.242(*)
(*)Largest absolute correlation between each variable and any discriminant function						

Figure 8.6: Classification results observed when the TIV data from facial regions R1 (forehead) and R2 (eyes and cheeks) were used to train the classifier

TABLE 8.4: CLASSIFICATION RESULTS WHEN OPTIMAL FEATURES (COMPONENTS) USED IN ANALYSIS

Classification		**Group**	**Predicted Group Membership**							**Total**
			Neutral	***Happy***	***Sad***	***Disgust***	***Surprise***	***Angry***	***Fear***	
Original Cases	***Count***	***Neutral***	8	0	2	1	1	4	0	16
		Happy	1	10	0	1	3	1	0	16
		Sad	0	0	14	1	0	1	0	16
		Disgust	0	0	0	12	1	3	0	16
		Surprise	0	1	1	0	13	1	0	16
		Angry	2	0	1	0	1	12	0	16
		Fear	1	0	0	0	0	0	15	16
	%	***Neutral***	50	0	12.5	6.3	6.3	25	0	100.0
		Happy	6.3	62.5	0	6.3	18.8	6.3	0	100.0
		Sad	0	0	87.5	6.3	0	6.3	0	100.0
		Disgust	4	0	0	0	75	6.3	18.8	100.0
		Surprise	0	6.3	6.3	0	81.3	6.3	0	100.0
		Angry	12.5	0	6.3	0	6.3	75	0	100.0
		Fear	6.3	0	0	0	0	0	93.8	100.0
Cross-Vali-dated Cases	***Count***	***Neutral***	3	3	2	2	2	4	0	16
		Happy	2	5	0	2	5	1	1	16
		Sad	1	1	9	1	2	2	0	16
		Disgust	5	0	1	6	1	3	0	16
		Surprise	0	4	2	0	6	3	1	16
		Angry	3	0	2	3	2	5	1	16
		Fear	0	1	1	0	1	0	13	16
	%	***Neutral***	18.8	18.8	12.5	12.5	12.5	25	0	100.0
		Happy	12.5	31.3	0	12.5	31.3	6.3	6.3	100.0
		Sad	6.3	6.3	56.3	6.3	12.5	12.5	0	100.0
		Disgust	31.3	0	6.3	37.5	6.3	18.8	0	100.0
		Surprise	0	25	12.5	0	37.5	18.8	6.3	100.0
		Angry	18.8	0	12.5	18.8	12.5	31.3	6.3	100.0
		Fear	0	6.3	6.3	0	6.3	0	81.3	100.0

As evident in Table 8.3, the first discriminant function uses the discriminant variables derived from the TIVs measured on the left and right sides of the two major muscles, the Frontalis pars medialis and the Orbicularis oris. The second function (Function 2) uses the discriminant variables derived from the TIV data measured on the left and side of the Orbicularis oris. The last significant function (Function 3) uses the TIVs measured on the Frontalis pars medialis, Frontalis pars lateralis, and Zygomaticus major. Figures 3.1, 4.4 and 4.5 and Tables 4.1, 4.2 and 8.1 exhibit the physical location and muscular alignment of the FTFP sites on the face.

Figure 8.6 and Table 8.4 present the classification results obtained when the optimal principal components were used for developing the discriminant rules to distinguish between the seven facial expressions. As shown in Table 8.4, a high error rate (e^{j}_{geneal} = *58.0%*) was observed. Figure 8.7 shows the eigenspace and the profiles of these discriminant functions around the group centroids. A complex eigenspace obvious in Figure 8.7 suggests an overlap between the newly created smaller linear spaces within the eigenspace. It further suggests that between-group separation was not effective using the available thermal data. One could interpret that the facial thermal features measured on the upper parts of the face (R1 and R2) did not provide efficient discriminant functions for developing a robust classifier.

8.2.1 Significance of the classification results

The practical significance of the classifier performance was calculated using Huberty's statistical significance test. The statistics for the neutral, pretended happy, sad, disgusted, surprised, angry and fearful faces are given in Table 8.5. The overall significance test statistic ($p<0.05$) suggests that the classification results were statistically significant.

Huberty's test was also invoked for assessing the practical significance of classification results. The resulting index value I (*=32)* suggested that the employed computational procedure reduced the chances of classification errors by 32.29 %. The computational methods for estimating the significance of classification results were presented in chapter 5.

8.3 AFEC using the TIVs measured around mouth (R3) and on chin (R4)

In a following analysis, the TIVs recorded on the FTFP sites within the facial regions 3 and 4 (around mouth and chin) were used for training the classifier. Once again, the TIVs data were first transformed into the uncorrelated principal components. Only first 8 of 24 derived principal components significantly contributed to the variance in the

TIV data. Figure 8.8 exhibits a scree plot showing the contribution each principal component made in the underlying variance of the TIVs data collected at FTFP sites within the facial regions 3 and 4.

In the next step of classifier construction, the derived principal components were analysed to discover the optimal principal components. Only 11 of 24 derived principal components were found helpful in the between-class separation.

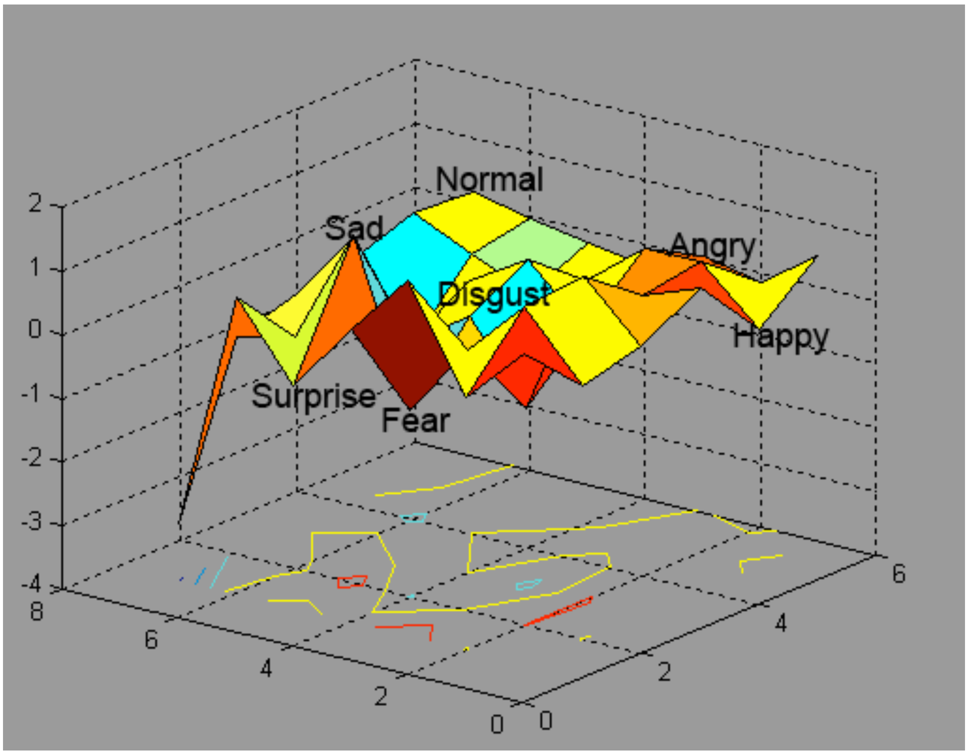

Figure 8.7: Approximate location of the 7 group centroids within the eigenspace. TIV data from the facial regions R1 (forehead) and R2 (eyes and cheeks) were used to train the classifier

TABLE 8.5: SIGNIFICANCE OF CLASSIFICATION RESULTS

Statistic	Value	Level of significance (α)
Z^*_{normal}	8.33	<0.005
Z^*_{happy}	5.93	<0.005
Z^*_{sad}	4.32	<0.005
$Z^*_{disgust}$	5.14	<0.005
$Z^*_{surprise}$	5.14	<0.005
Z^*_{angry}	5.93	<0.005
Z^*_{fear}	1.92	0.0274
$Z^*_{overall}$	8.06	<0.005

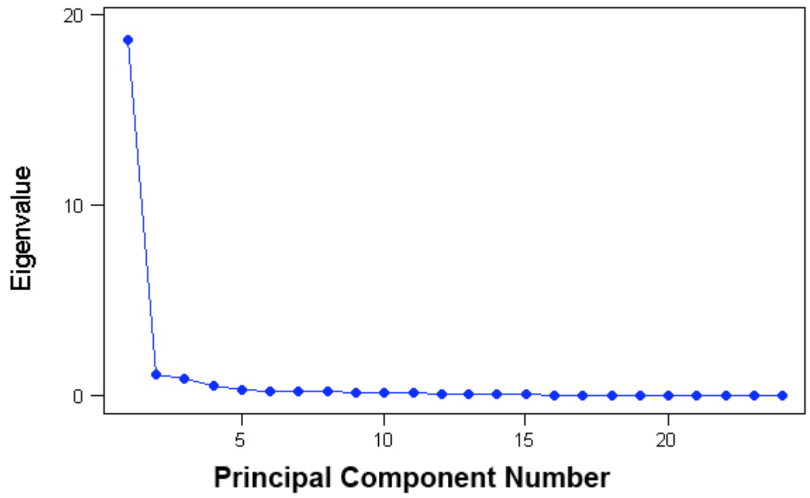

Figure 8.8: Contribution of the 24 principal components in the data variance for the 7 facial expression measured using the TIV data from facial regions R3 and R4

These components contributed in increasing the *F* ratio when the optimal principal components selection algorithm was invoked. As anticipated, some of the 11 most discriminating components did not have a high eigenvalue. Figure 8.9 presents the Andrews' curves drawn for the 11 optimal principal components that contributed to the underlying variation in the TIV data measured around the lips and chin. The Andrews' curves for the 7 facial expression group clusters in Figure 8.9 explain the between-facial expression group differences. Each facial expression has a distinct profile of curves. The curves in each cluster window fall into a moderately narrow band suggesting homogeneity in the variance structure.

As was the case in the previous analysis, the three distinct curves, visible in the clusters of facial expression groups in Figure 8.9 belong to the three participants with facial hair. The TIVs measured on the faces of these participants were different than those measured on other participant faces. Compared to Figure 8.4, these three faces with facial hair appeared far from the remaining faces within their respective facial expression group windows in Figure 8.9. Excluding these three participants, the curves appear to have similar profiles suggesting presence of a reasonable amount of within-group similarities. The Andrews' curves between the range $-\pi$ and $-3\pi/8$ show major between-group variation in Figure 8.9.

Figure 8.9: Andrews' curves drawn using the TIV data gathered from the facial regions R3 (mouth) and R4 (chin)

After the discovery of optimal facial thermal features, the linear discriminant algorithm was invoked on them for developing the discriminant functions. The respective statistical significance levels of the resulting discriminant functions were calculated using Equation 6-1. Only first two discriminant functions were significant (p<0.05) suggesting a low probability of between-facial expression group separation within the eigenspace.

Table 8.7 presents the structure matrix resulting from the discriminant analysis. The numeric coefficients in the structure matrix allow interpreting the contribution of each variable in the formation of the discriminant functions.

The structure matrix in Table 8.7 suggests that the first discriminant function uses the discriminant variables derived from the TIVs measured on Orbicularis Oris (please refer to Figure 8.1 and Table 8.1). The second significant function, Function 2 uses the discriminant variable derived from the TIVs measured on the left hand side of Orbicularis Oris, on Depressor Anguli Oris and on the Mentalis.

Figure 8.10 and Table 8.8 show the classification success results observed when the optimal principal components (derived from around mouth and chin) were used as input vectors to train the classifier. A higher classification error rate ($e^{j}_{geneal} = 66.1\%$) was observed when the optimal features were used for classifying the unknown TIRIs.

TABLE 8.6: SIGNIFICANCE OF INDIVIDUAL DISCRIMINANT FUNCTIONS DERIVED USING THE TIV DATA MEASURED ON FACIAL REGIONS R3 AND R4

Test of Functions	Wilks' Lambda	Chi-square	df	Sig.
1 through 6	0.222	153.460	66	0.00
2 through 6	0.386	97.081	50	0.00
3 through 6	0.656	42.992	36	0.197
3 through 6	0.821	20.159	24	0.688
3 through 6	0.914	09.140	14	0.822
6	0.985	1.539	06	0.957

TABLE 8.7: STRUCTURE MATRIX FOR THE SIX DISCRIMINANT FUNCTIONS DERIVED USING THE TIV DATA MEASURED ON FACIAL REGIONS R3 AND R4

Principal compone nt	Function 1	Function 2	Function 3	Function 4	Function 5	Function 6
C-13	0.426	0.137	-0.062	-0.134	-0.316	-0.174
C-08	0.273	0.243	-0.050	0.060	-0.167	-0.160
C-04	-0.219	0.230	0.527	0.219	-0.010	0.042
C-10	-0.258	-0.073	0.044	0.576	-0.214	0.181
C-16	0.355	0.066	0.453	0.496	0.148	-0.305
C-18	0.059	-0.294	0.155	-0.075	0.493	-0.210
C-02	0.121	-0.166	-0.437	0.436	0.469	-0.085
C-14	-0.008	0.283	0.185	-0.275	0.465	-0.017
C-20	0.340	-0.206	0.063	0.129	-0.164	0.569
C-22	0.066	0.478	-0.230	0.140	0.341	0.568
C-07	0.142	-0.250	0.337	-0.231	0.136	0.491

Figure 8.10: Classification results observed when the TIV data from the facial regions R3 (mouth) and R4 (chin) were used to train the classifier

TABLE 8.8: CLASSIFICATION RESULTS OBSERVED WHEN OPTIMAL FEATURES WERE USED FOR CLASSIFICATION

Classification		**Group**	**Predicted Group Membership**							**Total**
			Neutral	***Happy***	***Sad***	***Disgust***	***Surprise***	***Angry***	***Fear***	
Original Cases	***Count***	***Neutral***	4	2	1	2	4	1	2	16
		Happy	1	5	1	3	2	2	2	16
		Sad	3	0	10	0	0	1	2	16
		Disgust	1	2	0	9	1	3	0	16
		Surprise	1	1	0	1	13	0	0	16
		Angry	3	1	1	4	0	7	0	16
		Fear	1	0	1	1	2	0	11	16
	%	***Neutral***	25	12.5	6.3	12.5	25.0	6.3	12.5	100.0
		Happy	6.3	31.3	6.3	18.8	12.5	12.5	12.5	100.0
		Sad	18.8	0	62.5	0	0	6.3	12.5	100.0
		Disgust	6.3	12.5	0	56.3	6.3	18.8	0	100.0
		Surprise	6.3	6.3	0	6.3	81.3	0	0	100.0
		Angry	18.8	6.3	6.3	25.0	0	43.8	0	100.0
		Fear	6.3	0	6.3	6.3	12.5	0	68.8	100.0
Cross-Validated Cases	***Count***	***Neutral***	1	2	1	2	4	2	4	16
		Happy	2	1	1	2	5	3	2	16
		Sad	4	0	8	0	0	2	2	16
		Disgust	1	4	1	6	1	3	0	16
		Surprise	2	2	1	1	10	0	0	16
		Angry	4	3	1	5	0	3	0	16
		Fear	2	1	1	1	2	0	9	16
	%	***Neutral***	6.3	12.5	6.3	12.5	25	12.5	25	100.0
		Happy	12.5	6.3	6.3	12.5	31.3	18.8	12.5	100.0
		Sad	25	0	50	0	0	12.5	12.5	100.0
		Disgust	6.3	25	6.3	37.5	6.3	18.8	0	100.0
		Surprise	12.5	12.5	6.3	6.3	62.5	0	0	100.0
		Angry	25	18.8	6.3	31.3	0	18.8	0	100.0
		Fear	12.5	6.3	6.3	6.3	12.5	0	56.3	100.0

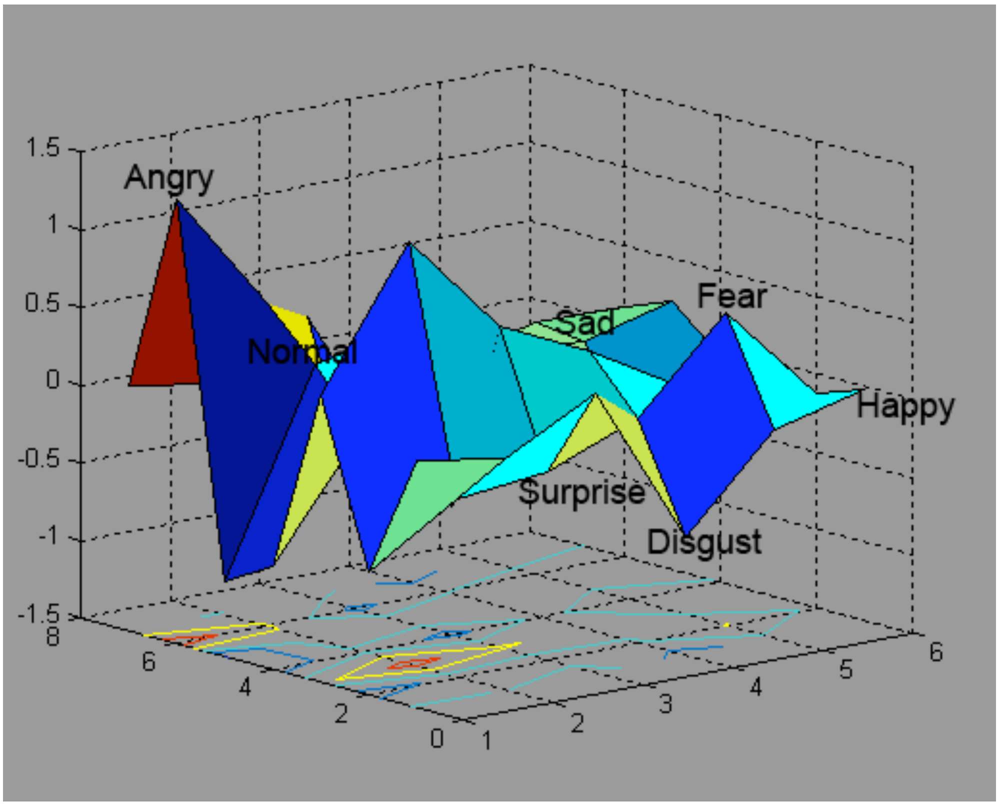

Figure 8.11: The 7 facial expression groups at their respective group centroids. The TIV data from around mouth (R3) and on chin (R4) were used to train the classifier

TABLE 8.9: SIGNIFICANCE OF CLASSIFICATION RESULTS

Statistic	Value	Level of significance (α)
Z^*_{normal}	15.49	*<0.005*
Z^*_{happy}	15.49	*<0.005*
Z^*_{sad}	4.0	*<0.005*
$Z^*_{disgust}$	5.14	*<0.005*
$Z^*_{surprise}$	2.45	*<0.005*
Z^*_{angry}	8.32	*<0.005*
Z^*_{fear}	4.32	*<0.005*
$Z^*_{overall}$	8.66	*<0.005*

Figure 8.11 exhibits the discriminating functions at their group centroids. A complex discriminating space is evident in Figure 8.11 suggesting an overlap between the smaller linear spaces within the eigenspace.

8.3.1 Significance of the classification results

The significance levels for the neutral, pretended happy, sad, disgusted, surprised, angry and fearful faces, and the overall significance level are given in Table 8.9. The test significance levels reported in Table 8.9 ($p<0.05$) were obtained using Equations 5-24 and 5-25. The resulting statistics suggested that classification results were significant.

Huberty's test for assessing the practical significance of classification results was also invoked to determine the extent to which the classification results were better than those that could be obtained by chance alone. The index *I*, calculated using Equation 5-26, was found to be 21.87 suggesting that the discrimination procedure reduced the chances of classification errors by 21.87 %.

8.4 Discussion

A previous investigation suggested that the increased blood volume flow around the upper parts of the face might result in dissipating more amount of heat from the upper parts of the face and might allow detection of negative emotional experiences (Puri et al. 2005). During this investigation, the classifier trained using the thermal features extracted from the upper parts of the face (R1 and R2) performed comparatively better than the classifier trained using the thermal features extracted from the lower parts of the face (regions 3 and 4). Figure 8.12 compares the differences between the AFEC potential of thermal facial features extracted from the upper and lower parts of participant faces.

The classification results reported in Figure 8.12 were also consistent with some other previous investigations. For example, (Kobayashi and Tagami 2004) studied the differences in the biophysical functions at various physical locations on the facial skin. Their work focused on examining the poor functional properties of the stratum corneum epidermis. The term "*stratum corneum*" is used for the outermost layer of the skin (also called epidermis). The stratum corneum is made up of the dead and usually flat skin cells.

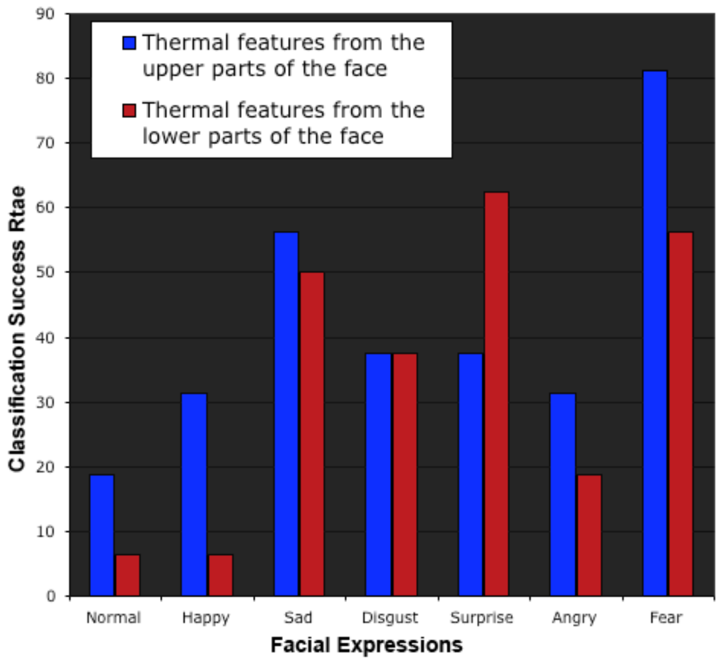

Figure 8.12: Performance of the two classifiers compared. The blue bars indicate the performance of the classifier trained using the TIV data measured from facial regions R1 (forehead) and R2 (eyes and cheeks). The red bars indicate the performance of the classifier trained using the TIV data measured from facial regions R3 (mouth) and R4 (chin).

These dead cells shed frequently, generally about after every 14-15 days (Skin Anatomy and Physiology 2005). It was also discovered in (Kobayashi and Tagami 2004) that the cheeks had the lowest epidermis temperature as compare to the other anatomical parts of the face. However, the cheeks and facial regions around the nose had a much higher blood flow rate as compare to the other anatomical parts of the face (Kobayashi and Togami 2004). The epidermis temperature measurements around various anatomical parts of the face (shown in Figure 8.1) in this investigation were very similar to the ones (Kobayashi and Togami 2004) observed and reported in their paper. The facial temperature distribution patterns derived from the TIV data concurred with the epidermis temperature distribution patterns (Kobayashi and Togami 2004) observed.

(Kobayashi and Togami 2004) also observed a positive correlation between the blood-flow rate and skin temperature measured on the epidermis (Kobayashi and Togami 2004). It is argued that an increased blood flow results in the dissipation of more heat from the epidermis of the face (Puri et al. 2005).

The similarities in emotion-specific blood volume flow variations observed at various anatomical locations of the face and the positive correlation between the blood volume flow and the skin temperature measurements might have contributed to the classification errors and confusion patterns reported earlier in Tables 8.4 and 8.8. As evident in Table 8.4, a lower classification error rate was attained when TIVs on the forehead and from around the eyes and cheek (high blood volume flow areas) were used for training the classifier. Classification error rates reported in Tables 8.4 and Table 8.8 show that the classifier trained with the thermal data around mouth and chin (low blood volume flow areas) was less efficient. A comparison of the confusion matrices in Tables 8.4 and 8.8 suggest that the discriminant rules developed using the thermal data form the high blood flow areas (forehead, eyes and cheeks) of the face were comparatively more efficient than the discriminant rules developed using the TIV data measured from around the lower parts of the face.

(Partala et al. 2006) estimated the real time emotional experiences using the electromyographic (EMG) activity measured on the two major facial muscles: Zygomaticus Major and the Corrugator Supercilli. Zygomaticus Major is believed to be active during smile whereas Corrugator Supercilli is believed to be active during frowning. As evident in Figures 4.5 and 4.6 (in Chapter 4), Zygomaticus Major is located around the mouth and cheeks whereas the Corrugator Supercilli is physically located on the upper part of the face. (Partala et al. 2006) estimated negative and positive facial expression of evoked affective states using the EMG activity measured on Zygomaticus and Corrugator Supercilli (Partala et al. 2006). The average Zygomaticus Major and Corrugator Supercilli responses in the study showed a significant influence of stimulus on these two muscles. However, the Corrugator Supercilli measurements were found to be more effective than those of the Zygomaticus Major for estimating the positive and negative facial expressions of affective states (Partala et al. 2006). These facial expression classification results were also consistent with the AFEC results observed in this investigation (presented earlier in Tables 8.4 and Table 8.8).

The classification results observed in this work were also consistent with those realised by (Lien et al. 1998) in a previous study. They developed a computer vision system to differentiate between the subtly different facial expressions. They employed three different facial feature extraction methods: facial feature point tracking method, dense flow tracking with PCA, and high-gradient component analysis method for developing the facial expression classification system. Only 15 % classification error rate was observed when the high-gradient component analysis method was used on the features extracted from the upper part of the face. When the same method was used on the facial features extracted from the lower part of the face, a higher (19%) error rate was observed (Lien et al. 1998). However, their classification results were much different when the facial feature point tracking method was used for extracting the facial features (Lien et al. 1998). The facial feature point tracking based classification system developed using the facial features from the upper part of the face was a little less effective than the one that employed facial features extracted from the lower part of the face. The dense flow tracking with PCA based method (Lien et al. 1998) also resulted in a 93% classification success rate when the features extracted from the upper parts of the face were used to train the classifier. It is important to remember that (Lien et al. 1998) used the visual signals for developing the AFEC capable system so their classification results may not be directly compared with the ones observed in this investigation.

As evident in the Andrews' curves plot in Figures 8.4, 8.5 and 8.9, the TIV measurements on the participant faces covered under the facial hair were significantly different than the ones measured on the facial skin surface. This difference probably represents a difference in the emissivity (ε) of the two surfaces. Therefore, it would be appropriate to develop a person independent classifier for people having facial hair or wearing glasses.

These AFEC results encourage exploring the possibilities of classifying the facial expressions using a combination of the TIV data measured along the major facial muscles. For example, lighting and pose conditions sometimes allow extracting the facial (or thermal) features from only one side of the face. Hence, it might be useful to measure TIV data along the facial muscles of positive and negative expressions on just one side of the face for classifying the positive and negative facial expressions.

8.5 Conclusion

This part of the investigation aimed at exploring the possibilities of recognising and classifying facial expressions using some regional thermal features under the assumption that the other parts of the face were not available for feature extraction. Thermal images of participant faces were divided into four regions. The forehead was considered region 1(R1). The areas around the eyes and cheek were considered region 2 (R2). The area around the mouth was considered region 3 (R3) and the part of the face around the chin was considered region 4 (R4). The TIVs recorded at the FTFP sites within regions 1 and 2 were grouped together. Facial regions 3 and 4 were separately grouped together. The TIVs measured within each of these two groupings were separately used to train the classifiers.

When the TIVs around regions 1 and 2 (upper parts of the face) were used for AFEC, a high error rate (58.0%) was observed during the cross-validation tests. The classifier performed at an even higher error rate (66.1%) when the TIVs around regions 3 and 4 (lower parts of the face) were used to train the classifier. In both cases, a subset of optimal features was used to train the classifier. These results suggest that variations in the TIVs data measured on the FTFP sites within the selected regions of a series of thermal images do not allow dividing the eigenspace along the discriminant function boundaries. Hence, one would assume that the thermal features gathered from the FTFP sites on the selected parts of the face did not help in classifying the facial expressions of affective states. Also, it became obvious from this investigation that the regional facial thermal data might not allow constructing a set of effective discriminant rules for achieving the AFEC functionality.

The classification results therefore suggest that the entire frontal views of the faces would be required to classify the facial expressions of affective states using the facial thermal features.

Chapter 9

CLASSIFICATION OF EVOKED FACIAL EXPRESSIONS

The investigations and analyses reported in chapters 6 and 7 suggest that the TIV data gathered from the FTFP sites on an entire face might help in classifying the simulated facial expressions of affective states. In a life like situation, the AFEC would be performed on naturally occurring spontaneous and evoked or reactive expressions of affective states. Therefore, in this part of the investigation, an attempt was made to measure and classify the temporal thermal features on the faces with naturally evoked facial expressions of affective states.

Earlier research suggests that an ability to distinguish between the subtle and reactive facial expressions would help affective systems perform in life like situations (Dautenhahn and Billard 1999). In some earlier works, scientists were able to extract useful human information from the bio-physiological signals for recognising the subtle and finer expressions of affects (Dautenhahn and Billard 1999). However, the viability of using the temporal facial thermal features in classifying the subtle and evoked facial expressions of affective states has yet to be explored.

Previous attempts of using the human bio-physiological information in AFEC and AAR have been critically analysed in the literature. For example, in concluding their recent work on problems and prospects of affective computing, (Ward and Marsden 2004) wrote *"If physiological measurement is to be useful in human–computer interaction, in the ways currently envisaged in the literature, it has to be able to identify reactions to subtle events, not just major failures of interaction. Similarly, physiological measurement has to be able to detect these reactions in loosely controlled naturalistic situations representative of real computer use, rather than tightly controlled laboratory settings. Psychophysiological data is very noisy, making cause and effect difficult to demonstrate. Even where there is clear cause and effect, interpretation in terms of users' internal mental processes and experiences presents serious further problems."*

This part of the investigation provides an opportunity to address some of these concerns. From the affective computing perspective, this part of the thesis, for the first time, compares the thermal differences between the simulated and naturally occurring facial expressions of affective states. The differences in a classifier's ability to recognise the pretended and evoked facial expressions using the facial skin temperature measurements are also examined. Differentiating between the evoked and reactive facial expressions with the help of facial thermal features might supplement the existing approaches for classifying subtly different facial expressions. Since hiding the bio-physiological responses to emotions is difficult, this investigation might assist researchers in finding a novel way of discovering the concealed emotions.

Following paragraphs present details of the evoked thermal data acquisition protocols. The statistical analyses of the evoked expression data and the evoked facial expression classification results are also reported in the following paragraphs. The observed classification results are discussed at the end of this chapter to conclude this last phase of investigation.

9.1 Equipment, software and participants

The physical facilities, equipment, hardware, software and accessories reported in chapter 4 were again used during this phase of investigations. The image acquisition process reported in chapter 4 was repeated again to acquire the thermal and visible-spectrum images of 10 undergraduate students. The participants, 7 male and 3 female had a mean self-reported age of 21 years 2 months. Only 3 of the 10 participants also participated in the previous experiments when the intentional facial expressions were recorded. Participants included Arabs, Iranians and Indians. All participants allowed using their visible-spectrum and infrared images for publication and dissemination of information.

9.2 Evoking expressions and acquiring thermal images

The Psychology and Cognitive Studies literature cited several methods of prompting and stimulating emotions and affective states. Earlier researchers reported that inducing genuine and authentic emotions was a difficult job and suggested that care should be taken in making judgments about the observed emotions and their expression. It was suggested that the desired emotive states be evoked in a way that participants would not simulate the emotion either intentionally or unintentionally (Dror et al. 2005; Hirsch and Mathews 1997; Murphy and Zajonc 1993; Niedenthal et al. 2000; Wild et al. 2001).

Some widely used methods of evoking emotions include affective picture viewing, emotive text reading and storytelling (Partala et al. 2006; Whiteside 1998). Participants in (Dror et al. 2005; Toivanen 2004; Wild et al. 2001) were exposed to one or more of these stimuli for invoking the desired positive and negative emotions. In some experiments, emotions were evoked by letting the participants read stories or emotive text. The literature highlighted a need for specialised expertise to select the most effective text and stories. In a typical setting, qualified and trained psychologists are involved in selecting effective stimuli as the task required special skills (Dror et al. 2005; Partala et al. 2006; Toivanen 2004; Wild et al. 2001).

Some specially collected and well-tested databases of images and photographs that are used for evoking emotions are available on the World Wide Web. These databases have been employed in many recent studies for invoking the affective states. One such database, International Affective Picture System (IAPS), was developed at the Center for the Study of Emotion and Attention at the University of Florida (Dror et al. 2005; Toivanen 2004). The IAPS database was used in several recent investigations for invoking emotions (Mikels et al. 2005). It contains a large set of different pictures that allow evoking emotions along the dimensions of valance and arousal (Wild et al. 2001). The IASP site managers were contacted to access the IASP database but they never acknowledged so the IASP image database could not be accessed. As an alternate, the relevant literature was reviewed for developing an appropriate approach in this investigation.

A set of still images and video clips was selected and used for evoking expressions of happiness, sadness, disgust and anger in this investigation. The contents of these stimuli were similar to the contents of the pre-categorised IASP images available on the web. The selected stimuli were available at the official web sites of some prestigious publishers such as BBC, MSNBC, and CNN. Selected images were compared and matched with the categorised pictures (available online) in the IAPS database. Extremely violent and disturbing images and images with unethical contents were avoided. The images and contents of selected video clips were no more extreme that those shown on mainstream television news and feature stories. Sources of the employed image and video clip sources are listed below.

Resources for evoking happiness and its facial expression were taken from:

- ‘Allo ‘Allo (BBC.COM)
- Bread (BBC.COM)
- Porridge (BBC.COM)

Resources for evoking sadness and its facial expression were taken from:

- Ferry Disaster in Pictures (BBC.COM)
- Pictures from the Disaster zone, Galtuer, Austria (BBC.COM)
- Turkey 1999 earthquake pictures (The New York Times)
- Caracas, Venezuela, pictures of 1999 floods and disasters (CNN.COM)
- Oklahoma City Bombing destruction pictures (CNN.COM)
- Pictures of June 2001 Missile fire destruction in Iraq (CNN.COM)

Expression of disgust was evoked using the following resources:

- Shock posters (BBC.COM)
- Shock Ads (BBC.COM)
- Stop Litter pictures (BBC.COM)
- Caracas, Venezuela, pictures of 1999 floods and disasters (CNN.COM)
- Oklahoma City bombing destruction pictures (CNN.COM)
- Pictures of June 2001 Missile fire destruction in Iraq (CNN.COM)

Expression of anger was evoked using the following resources:

- 2002 Riots in Gujarat, India (BBC.COM)
- In pictures: Anger and anguish in Iraq (BBC.COM)
- In pictures: Argentina anger boils over (BBC.COM)
- Iraqi prisoner abuse (Washington Post)
- Abu Ghoraib prison pictures (The New Yorker online)

The employed images and video clips had both high-emotion and low-emotion evoking contents. Participants’ visible and infrared images were simultaneously recorded when an evoked emotion was realised. After capturing the images with neutral faces, each participant was made to experience and express happiness, sadness, disgust and anger.

Recordings were often repeated for developing a good set of desired images. Of several recorded images, the ones that best represented the emotive states were selected and used in the investigation. This ensured using the best practices for prompting the affective states and recording the realistic facial expression of affects.

A team of three referees was requested to select the best visible pictures that truly reflected the desired emotions. The referees were requested to use their collective judgment for selecting the best expression of affective states. The referees, by consensus, agreed on the realistic expressions in the pictures and selected the most natural looking pictures of evoked expressions. With each selected visible picture, the corresponding thermal image was selected and used in the further analysis and classification. The evoked facial expressions of two participants are shown in Figures 9.1 and 9.2.

Figure 9.1: A female participant with evoked facial expressions

Figure 9.2: A male participant with evoked facial expressions

As in the previous experiments, participants were briefed about the objectives of experiments, methods and procedures and the probable outcome of the experiments. All participants were adults. They were informed about the potential benefits of the research and related experiments. Participants were given an opportunity to either continue or quit after the briefing. Efforts were made to protect participants from any physical and/or emotional harm and damage.

9.3 Analyses of evoked expression thermal data

The thermal data with evoked facial expressions were first tested for the assumption of normal distribution. Following the successful test of normal distribution of the data in individual TIRIs, participants' images showing a particular facial expression were grouped together and the data were tested again for the assumption of normal distribution. Based on the standard test results, The TIVs measured at only 4 of 75 (*5.33%*) FTFP sites violated the assumption of normal distribution.
Similarly, the TIVs measured at 16 of 75 (*21.33%*) FTFP sites had kurtosis statistics that suggested violation of the assumption of normal distribution. The assumption of normal distribution was further examined through the visual inspection of Histograms and Q-Q plots. The Kolmogorov-Smirnov and Shapiro-Wilk tests were also invoked on the data. Based on the initial analyses, it was considered safe to assume that the evoked expressions data were normally distributed. Figure 9.3 exhibits a few distributions of the TIV data measured at various FTFP sites.

The facial expression groups were also tested for the similarity of covariance structure before invoking the multivariate analyses and pattern recognition algorithms. The groups of TIRIs with neutral faces and faces with four evoked expression (happiness, sadness, disgust and anger) were compared. The test of sphericity was insignificant at ($p > 0.05$) suggesting symmetry of variance structures in the facial expression groups.

Figure 9.3: A typical distribution of the TIV data measured on six randomly selected FTFP sites

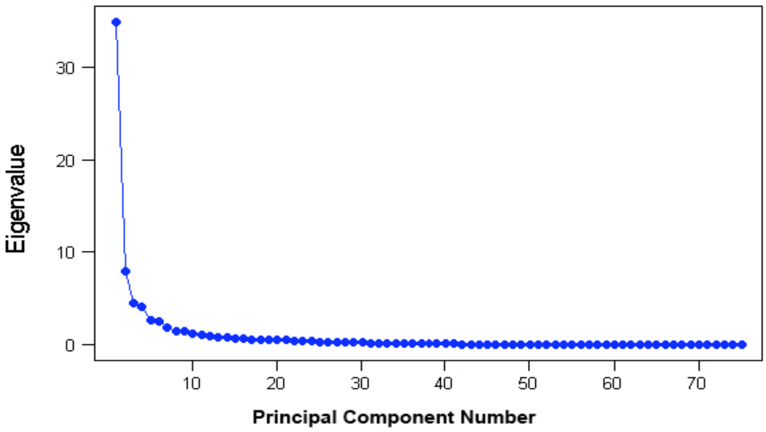

Figure 9.4: Contribution of the 75 principal components in the measured evoked TIV data variance

The conservative estimates of sphericity: Greenhouse-Geisser correction, Huynh-Fieldt correction and lower-bound test also suggested that the TIV data met the assumption of compound symmetry and the facial expression groups had a similar structure of variance. The initial data analyses suggested that the TIV data were suitable for multivariate transformations and the pattern recognition techniques might be invoked on the data.

9.4 Classifying the evoked facial expressions

The algorithmic approach presented earlier in Chapter 5 was employed for classifying the neutral faces and the faces with evoked facial expressions. As was done while classifying the intentional facial expressions, the TIV data were first transformed into uncorrelated principal components using the singular value decomposition algorithm. Figure 9.4 exhibits a scree plot showing the principal components and their contribution in the underlying variance of the TIV data. Table 9.1 provides the eigenanalysis results for all 75 principal component derived from the 75 TIVs on the 10 participant faces. Only 24 of 75 principal components in Table 9.1 significantly contributed to the variance in the TIV data. There was no evidence to believe that these 24 principal components would significantly contribute to the between-group separation.

TABLE 9.1: THE 75 PRINCIPAL COMPONENTS AND THEIR RESPECTIVE EIGENNVALUES

Principal Component	Eigenvalue	Proportion	Cumulative	Principal Component	Eigenvalue	Proportion	Cumulative
PC- 1	34.903	0.465	0.465	***PC- 39***	0.095	0.001	0.993
PC- 2	7.942	0.106	0.571	***PC- 40***	0.088	0.001	0.994
PC- 3	4.530	0.06	0.632	***PC- 41***	0.085	0.001	0.995
PC- 4	4.057	0.054	0.686	***PC- 42***	0.069	0.001	0.996
PC- 5	2.689	0.036	0.722	***PC- 43***	0.065	0.001	0.997
PC- 6	2.511	0.033	0.755	***PC- 44***	0.063	0.001	0.998
PC- 7	1.829	0.024	0.779	***PC- 45***	0.049	0.001	0.998
PC- 8	1.483	0.02	0.799	***PC- 46***	0.039	0.001	0.999
PC- 9	1.438	0.019	0.818	***PC- 47***	0.032	0	0.999
PC- 10	1.193	0.016	0.834	***PC- 48***	0.026	0	1
PC- 11	1.123	0.015	0.849	***PC- 49***	0.019	0	1
PC- 12	0.956	0.013	0.862	***PC- 50***	0	0	1
PC- 13	0.831	0.011	0.873	***PC- 51***	0	0	1
PC- 14	0.787	0.01	0.884	***PC- 52***	0	0	1
PC- 15	0.726	0.01	0.893	***PC- 53***	0	0	1
PC- 16	0.64	0.009	0.902	***PC- 54***	0	0	1
PC- 17	0.599	0.008	0.91	***PC- 55***	0	0	1
PC- 18	0.585	0.008	0.918	***PC- 56***	0	0	1
PC- 19	0.508	0.007	0.924	***PC- 57***	0	0	1
PC- 20	0.49	0.007	0.931	***PC- 58***	0	0	1
PC- 21	0.484	0.006	0.937	***PC- 59***	0	0	1
PC- 22	0.437	0.006	0.943	***PC- 60***	0	0	1
PC- 23	0.391	0.005	0.948	***PC- 61***	0	0	1
PC- 24	0.352	0.005	0.953	***PC- 62***	0	0	1
PC- 25	0.329	0.004	0.958	***PC- 63***	0	0	1
PC- 26	0.314	0.004	0.962	***PC- 64***	0	0	1
PC- 27	0.285	0.004	0.965	***PC- 65***	0	0	1
PC- 28	0.273	0.004	0.969	***PC- 66***	0	0	1
PC- 29	0.235	0.003	0.972	***PC- 67***	0	0	1
PC- 30	0.224	0.003	0.975	***PC- 68***	0	0	1
PC- 31	0.196	0.003	0.978	***PC- 69***	0	0	1
PC- 32	0.181	0.002	0.98	***PC- 70***	0	0	1
PC- 33	0.175	0.002	0.983	***PC- 71***	0	0	1
PC- 34	0.159	0.002	0.985	***PC- 72***	0	0	1
PC- 35	0.144	0.002	0.987	***PC- 73***	0	0	1
PC-36	0.134	0.002	0.988	***PC- 74***	0	0	1
PC- 37	0.127	0.002	0.99	***PC- 75***	0	0	1
PC- 38	0.109	0.001	0.992				

Figure 9.5 shows the separation between the neutral and invoked facial expressions of happiness in a 2-principal component eigenspace. The neutral faces could be easily separated from the happy faces. The underlying between-group separation is obvious in Figure 9.5.

Figure 9.6 exhibits how the neutral and sad faces were separated in a 2-principal component eigenspace.

Figure 9.5: Separation of the neutral faces and evoked expression of happiness in a 2-PC eigenspace

Figure 9.6: Separation of neutral faces and evoked expression of sadness in a 2-PC eigenspace

Figure 9.7: Separation of neutral faces and faces with evoked expression of disgust in a 2-PC eigenspace

Figure 9.8: Separation of the neutral faces and evoked expression of anger in a 2-PC eigenspace

Figure 9.7 shows how the neutral faces and the faces with the evoked expression of disgust were separated in a 2-principal component eigenspace.

Figure 9.8 exhibits the separation between the neutral faces and faces with the evoked facial expression of anger in the 2-principal component eigenspace.

Figures 9.5, 9.6, 9.7 and 9.8 provide convincing information about the differences in the thermal profiles of the neutral faces and the faces with evoked facial expressions. It is obvious in the four figures that a neutral face, with all facial muscles in their natural position, is thermally different from a face that expresses an evoked positive or negative effective state.

In a following analysis, the patterns of thermal differences between the positive and negative facial expressions were examined. The first 2 principal component scores computed for the facial expressions of happiness, sadness, disgust and anger were plotted in a 2-dimensional eigenspace.

As evident in Figure 9.9, facial expressions of happiness and sadness were found difficult to distinguish in a more complex 2-principal component eigenspace.

The separation between the evoked negative expression of sadness, disgust and anger also appeared difficult in a 2-dimensional eigenspace. Figure 9.10 shows a complete overlap between facial expression of sadness and disgust on 8 of 10 participant faces. It is obvious in Figure 9.10 that the first two principal components might not provide sufficient information for effectively classifying the three facial expressions.

Figure 9.9: Evoked facial expressions of happiness and sadness in a 2-PC eigenspace

Figure 9.10: Evoked facial expressions of sadness, disgust and anger in a 2-PC eigenspace

Figure 9.11: Evoked facial expressions of happiness, sadness, disgust and anger in a 2-PC eigenspace

Figure 9.11 exhibits a positive and three evoked negative facial expressions in a 2-principal component eigenspace. Figure 9.11 also demonstrates an overlap between the facial expressions of the three negative affective states. However, the positive facial expression of happiness appears to be easily separable from the three negative facial expressions.

Figure 9.12: Neutral faces and faces with evoked facial expressions of happiness, sadness, disgust and anger in a 2-PC eigenspace

The facial expressions of negative emotions; sadness, disgust and anger in Figure 9.12, appear to have similar eigenscores in a 2-dimensional eigenspace. However, the neutral faces and faces with evoked expression of happiness seem to have different eigenscores and are distinguishable from the negative facial expressions in a 2-principal component eigenspace.

The thermal profiles of the neutral faces and the faces with evoked facial expressions of positive and negative experiences appear to be different, even within a 2-dimensional eigenspace. However, the first two principal components might not be expected to explain all the variance between the neutral and the four evoked facial expressions.

To further examine the variance structure of the TIV data, the Andrews´ plots explained earlier in Chapter 7, were drawn using the principal component scores derived from the TIV data. Figure 9.13 shows principal component scores Andrews´ plot for each facial expression in a separate window. The plots for neutral and happy facial expressions were different than those for the faces with evoked expressions of sadness, disgust and anger. This difference in thermal profiles implied the underlying differences between the thermal features pertaining to the neutral, positive and negative facial expressions.

The individual facial expression groups in each of the 5 windows exhibit a visually consistent within-group profile of curves. Hence, the within-group variance in the individual facial expression group clusters in Figure 9.13 seems encouraging. On the contrary, the graphical appearance of the patterns of between-group variance seems visually different and suggests that the TIV data might help in developing the discriminant functions. The distinct curves within the individual facial expression clusters, representing the within-group variation, might have resulted from the varying intensities of the facial expressions (arousal factor). Natural differences in the facial skin temperature of the participants or an underlying difference in the facial skin emissivity (ε) due to factors such as facial hair and the skin colour could also contribute to such variations (Jones 1998). The "*stratum corneum epidermis*" factor discussed earlier could also influence the thermal profile of an individual participant's face (Kobayashi and Tagami 2004; Skin Anatomy and Physiology 2005).

Examining the actual influence of one or more of these factors on the facial skin temperature and the TIV measurements is beyond the scope of this investigation. However, the principal component scores plotted in an eigenspace and the graphical profiles of the Andrews´ curves suggest that the neutral faces and the faces with evoked facial expressions may be thermally distinguished.

The derived principal components were tested for their contribution to the variance in the thermal data. The optimal feature selection algorithm was invoked for discovering the optimal facial features. Figure 9.14 shows the recursive discovery of the optimal principal components and the corresponding increase in the resulting *F* ratio. Only 30 of the derived principal components shown in Figure 9.14 were found significantly helpful in increasing the *F* ratio. These most discriminating principal components were used to train the classifier.

Since the data were normally distributed, the facial expression groups had a similar variance structures and each facial expression group had the same number of participants, an equal *a priori* was assumed during the discriminant analysis.

Figure 9.13: Andrews' curves drawn using the principal components scores for the neutral and evoked facial expressions of happiness, sadness, disgust and anger

TABLE 9.2: SUMMARY OF CANONICAL DISCRIMINANT FUNCTIONS

Function	**Eigenvalue**	**Percentage of Variance**	**Cumulative percentage**	**Canonical Correlation**
1	46.767	78.5	78.5	0.989
2	7.115	11.9	90.4	0.936
3	3.770	06.3	96.7	0.889
4	1.943	03.3	100.0	0.813

Figure 9.14: Stepwise selection of optimal components and corresponding increase in the *F* ratio

TABLE 9.3: SIGNIFICANCE OF INDIVIDUAL DISCRIMINANT FUNCTIONS

Test of Functions	**Wilks' Lambda**	**Chi-square**	**df**	**Sig.**
1 through 4	0.000	279.564	112	0.00
2 through 4	0.009	153.908	81	0.00
3 through 4	0.071	85.861	52	0.002
4	0.340	35.082	25	0.087

Table 9.2 summarises the canonical discriminant functions and the pertinent statistics. It also provides an estimate of the underlying contribution of each discriminant function in between-group separation.

Table 9.3 shows the chi-square statistics and the respective significance of each discriminant function.

Table 9.4 reports the structure matrix showing the pooled within-groups correlations between the discriminating variables.

Table 9.5 presents the standardised canonical discriminant function coefficients. The group membership of an unknown thermal face is determined using these numeric coefficients during the leave-one-out cross validation.

TABLE 9.4: STRUCTURE MATRIX FOR THE FOUR DISCRIMINATING FUNCTIONS

Variate	Function 1	Function 2	Function 3	Function 4
VARIATE-55	0.005	-.161(*)	0.051	0.029
VARIATE-74	0.053	.127(*)	0.071	0.102
VARIATE-42	-0.025	.113(*)	0.044	-0.096
VARIATE-03	-0.002	.082(*)	0.004	-0.067
VARIATE-07	0.017	-0.094	.193(*)	-0.132
VARIATE-24	0.018	0.109	.127(*)	0.024
VARIATE-40	0.025	-0.095	-.123(*)	-0.051
VARIATE-14	-0.025	-0.076	-.120(*)	-0.009
VARIATE-71	0.053	0.006	.118(*)	0.031
VARIATE-19	0.025	0.013	.098(*)	0.003
VARIATE-41	0.006	-0.048	.086(*)	-0.03
VARIATE-46	-0.065	-0.063	-.080(*)	0.046
VARIATE-17	-0.033	0.024	.074(*)	-0.03
VARIATE-23	-0.025	-0.024	.048(*)	-0.011
VARIATE-62	0.031	0.026	-0.003	-.307(*)
VARIATE-09	-0.02	-0.021	0.116	.216(*)
VARIATE-31	-0.019	0.051	0.13	.193(*)
VARIATE-38	-0.021	0.076	0.136	-.180(*)
VARIATE-36	0.031	0.066	-0.071	.158(*)
VARIATE-32	0.016	-0.107	0.13	-.134(*)
VARIATE-08	0.009	-0.048	0.121	.127(*)
VARIATE-22	-0.035	0.101	0.057	-.125(*)
VARIATE-10	0.045	-0.003	0.06	.125(*)
VARIATE-54	-0.016	-0.08	-0.065	.123(*)
VARIATE-16	0.078	0.037	-0.034	-.087(*)
VARIATE-30	0.013	-0.012	0.012	.067(*)
VARIATE-35	0.016	0.044	-0.019	-.062(*)
VARIATE-47	0.013	0.023	-0.053	.056(*)

TABLE 9.5: STANDARDISED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

Variates	Function 1	Function 2	Function 3	Function 4
VARIATE-03	0.05	0.756	-0.021	-0.069
VARIATE-07	0.716	-0.545	0.68	-0.366
VARIATE-08	0.334	-0.455	0.581	0.362
VARIATE-09	-0.665	-0.161	0.456	0.496
VARIATE-10	2.193	0.061	0.226	0.251
VARIATE-14	-1.259	-0.567	-0.466	0.177
VARIATE-16	3.052	0.235	-0.066	-0.163
VARIATE-17	-1.578	0.104	0.227	-0.048
VARIATE-18	0.763	0.021	0.468	0.093
VARIATE-24	-0.864	0.644	0.262	-0.538
VARIATE- 23	-0.965	-0.336	0.177	0.074
VARIATE-24	0.476	0.515	0.541	0.001
VARIATE-30	0.859	-0.257	0.141	0.061
VARIATE-31	-1.036	0.427	0.34	0.163
VARIATE-32	1.322	-0.567	0.494	-0.138
VARIATE-35	1.381	0.499	-0.102	0.082
VARIATE-36	1.509	0.557	-0.431	0.384
VARIATE-38	-1.578	0.735	0.729	-0.468
VARIATE- 40	2.506	-0.25	-1.055	-0.564
VARIATE-41	0.743	-0.973	0.559	-0.112
VARIATE-42	-1.128	0.771	0.305	-0.263
VARIATE-46	-2.238	0.391	-0.202	0.48
VARIATE-47	0.67	-0.437	-0.12	0.081
VARIATE-54	-1.747	0.208	0.398	-0.069
VARIATE-55	-2.131	-1.344	0.586	0.374
VARIATE-63	-1.02	-0.017	-0.059	-0.965
VARIATE-71	2.806	0.077	0.505	-0.264
VARIATE-74	-1.278	1.223	-0.011	1.105

Table 9.6 shows the confusion matrix and the classification success results observed when the classifier was trained using the high eigenvalued principal components to discriminate between the evoked facial expressions. As shown in Table 9.6, a high error rate (e^{j}_{geneal} = *84.0%*) was observed when the highest eigenvalued principal components were used for classification of unknown TIRIs.

Table 9.7 presents the classification results and the confusion matrix observed when the classifier was trained using the optimal principal components. The classification results in Table 9.7 show a significant reduction in the error rate (e^{j}_{geneal} = *28.0%*) as compare to the classification results shown in Table 9.6.

9.5 Classification error analysis

The linear discriminant algorithm discovered a set of 28 most effective variates for classifying the unknown thermal faces during the leave-one-out classification tests. These influential variates were listed in Tables 9.4 and 9.5.

TABLE 9.6: CLASSIFICATION RESULTS WITH HIGHEST EIGENVALUED FEATURES (COMPONENTS)

Classification		**Group**	**Predicted Group Membership**					
			Neutral	***Happy***	***Sad***	***Disgust***	***Angry***	***Total***
Original cases	***Count***	***Neutral***	9	0	0	0	1	10
		Happy	1	8	1	0	0	10
		Sad	0	1	9	0	0	10
		Disgust	1	0	0	9	0	10
		Angry	1	1	0	0	8	10
	%	***Neutral***	90.0	0	0	0	10.0	100.0
		Happy	10.0	80.0	10.0 1	0	0	100.0
		Sad	0	10.0	90.0	0	0	100.0
		Disgust	10.0	0	0	90.0	0	100.0
		Angry	10.0	10.0	0	0	80.0	100.0
Cross-Validated cases	***Count***	***Neutral***	3	0	2	2	3	10
		Happy	1	2	3	2	2	10
		Sad	3	1	3	2	1	10
		Disgust	2	2	1	0	5	10
		Angry	4	3	1	2	0	10
	%	***Neutral***	30.0	0	20.0	20.0	30.0	100.0
		Happy	10.0	20.0	30.0	20.0	20.0	100.0
		Sad	30.0	10.0	30.0	20.0	10.0	100.0
		Disgust	20.0	20.0	10.0	0	50.0	100.0
		Angry	40.0	30.0	10.0	20.0	0	100.0

TABLE 9.7: CLASSIFICATION RESULTS OBSERVED WITH THE OPTIMAL COMPONENTS

Classification		Group	Predicted Group Membership					
			Neutral	*Happy*	*Sad*	*Disgust*	*Angry*	*Total*
Original cases	***Count***	***Neutral***	10	0	0	0	0	10
		Happy	0	10	0	0	0	10
		Sad	0	0	10	0	0	10
		Disgust	0	0	0	10	0	10
		Angry	0	0	0	0	10	10
	%	***Neutral***	100.0	0	0	0	0	100.0
		Happy	0	100.0	0	0	0	100.0
		Sad	0	0	100.0	0	0	100.0
		Disgust	0	0	0	100.0	0	100.0
		Angry	0	0	0	0	100.0	100.0
Cross-Validated cases	***Count***	***Neutral***	7	0	0	1	2	10
		Happy	0	7	2	1	0	10
		Sad	0	1	9	0	0	10
		Disgust	1	1	0	7	1	10
		Angry	2	0	0	2	6	10
	%	***Neutral***	70.0	0	0	10.0	20.0	100.0
		Happy	0	70.0	20.0	10.0	0	100.0
		Sad	0	10.0	90.0	0	0	100.0
		Disgust	10.0	10.0	0	70.0	10.0	100.0
		Angry	20.0	0	0	20.0	60.0	100.0

The statistical significance levels of the four discriminant functions developed for differentiating between the neutral and four evoked facial expression groups were estimated using Equation 6-1. The significance (p<0.05) of the first three discriminant functions in column 5 of Table 9.3 suggested a possible separation between the facial expression groups along the first three discriminant functions. However, the fourth discriminant function was non-significant (p>0.05) and could not be assumed to have a significant role in classifying the unknown thermal faces.

Table 9.4 reported the structural coefficients of discriminant functions. These numeric coefficients were used to interpret the contribution each variable made in formulating the discriminant functions. Method of computing these coefficients was presented earlier in chapter 6.

The discriminant function coefficients in the structure matrix (Table 9.4) suggest that the first discriminant function derived constituent values from the TIVs measured on Frontalis Pars Medialis and Frontalis Pars Laterals (3, 7, 9, 14, 16, 17), Orbicularis Oculi (22, 31) and Levator Labii Superioris Alaquae Nasi (36), Levator Labii Superioris (46), Depressor Labii Inferioris (71) and Mentalis (74). The first discriminant function therefore computed the thermal features on the selected sites around the upper and lower parts of the face for classifying a new and unknown thermal face.

The second discriminant function relied on the variates measured from Frontalis (8, 14, 16), Orbicularis Oculi Pars Orbital (22, 24, 31), Levator Labii Superioris Alaquae

Nasi (35, 36), Levator Labii Superioris (38, 46), Zygomaticus Major (42), Risorious/ Platysma (54), Depressor Anguli Oris (55) and Mentalis (74) for classifying the new and unknown faces.

The third discriminant function relied on the variates measured at several locations on Frontalis (7, 8, 9, 14, 16, 17) Orbicularis Oculi Pars Orbital (19, 23, 24, 31), Masseter Superficial (40, 41), Zygomaticus Major (32, 42), Levator Labii Superioris (38, 46), Depressor Anguli Oris (55), Depressor Labii Inferioris (71) and Mentalis (74) for allocating an unknown thermal face to one of the evoked facial expression groups.

The fourth discriminant function relied on the variates measured at several locations on Frontalis (3, 8, 9, 10, 16), Frontalis Pars Medialis (7), Orbicularis Oculi Oars Orbital (22, 30, 31), Levator Labii Superioris (38, 46), Zygomaticus Major (32), Levator Labii Superioris Alaquae Nasi (35, 36), Levator Anguli Oris (47), Risorious/ Platysma (54), and Platysma (62) for allocating the unknown thermal faces to one of the facial expression groups.

Figure 9.15 shows the five facial expressions at their respective group centroids in a 3-discriminant function eigenspace. The varying spans of these discriminant functions highlight their respective influence in the between-group separation.

The leave-one-out cross validation test results presented in Table 9.7 exhibit the confusion patterns and the similarities between the evoked facial expressions in the underlying eigenspace. Some problems in the classification of unknown thermal faces are also evident in the confusion matrix in Table 9.7. For example, the evoked expression of anger was confused with the neutral faces and with the evoked expression of disgust.

The evoked expressions of happiness and sadness appeared to be overlapping in a 2-dimensional eigenspace in Figure 9.11. However, the confusion matrix in Table 9.7 suggested they were separated in a high dimensional eigenspace. Therefore, the evoked expression of happiness appears to be well separated from the other facial expressions in Table 9.7.

Given a small sample size and a large number of measured features (TIVs), the overall error rate ($e^{j}_{geneal} = 28.0\%$) observed during the leave-one-out cross validation tests seems encouraging.

Figure 9.15: The neutral faces and the faces with four evoked facial expressions at their respective group centroids

TABLE 9.8: SIGNIFICANCE OF CLASSIFICATION RESULTS

Statistic	Value	Level of significance (α)
Z^{*}_{normal}	2.070	0.0197
Z^{*}_{happy}	2.070	0.0197
Z^{*}_{sad}	1.054	0.1496
$Z^{*}_{disgust}$	2.070	0.0197
Z^{*}_{angry}	2.5819	0.0049
$Z^{*}_{overall}$	3.741	0.0011

The observed error rate was suggestive of the potential effectiveness of the thermal features for achieving the AFEC and AAR functionality. The observed classification results also suggested the effectiveness of the employed algorithmic approach.

9.6 Significance of the classification results

Equations 5-26 and 5-27 were used to determine the practical significance of the classification results. The significance levels for the facial expressions classification and the overall significance test statistics in Table 9.8 were significant at ($p<0.05$).

The practical significance of the observed classification results was estimated using Equation 5-28. The resulting index ($I = (36/50 - 10/50)/(1 - 10/50) \times 100$ = 65.0) suggested a 65.0% reduction in the error rate. It could be assumed that the TIV data gathered from the 75 FTFP sites on the participant faces could help distinguish between the neutral and four evoked facial expressions.

9.7 Conclusion

Attempts were made to distinguish between the neutral faces and faces with evoked facial expressions of happiness, sadness, anger and disgust using the facial skin temperature measurements. The TIV data obtained from the participant faces were found to be statistically suitable for invoking the relevant multivariate analysis. Results suggest that the facial thermal features might help in classifying the facial expressions of naturally evoked affective states.

The classifier trained with the high eigenvalued principal components could not distinguish between the evoked facial expressions suggesting that the employed dimension reduction techniques were not able to reduce the complexity of the underlying Gaussian model. However, when a set of optimal features was used to train the classifier, the overall error rate significantly reduced and the classifier performance improved. The reduced classification error rate suggested several possibilities.

First possibility arises from a study carried out in early 1960s that sparked several investigations and lead to important discoveries about the nature and processing of data measured at equal points at different times (Rao 1964). These studies suggested that the overall covariance matrix represents a mixture of contributions from within-group and between-group treatments, and variances (Rao 1964; Jolliffe 2002). For the discriminant analysis, one may prefer to separate the various types of covariances prevailing in the data. It is argued that even though the derived principal components are uncorrelated overall, they cannot be assumed as completely uncorrelated with respect to the between-group and within-group variations (Rao 1964; Jolliffe 2002). This problem is frequently encountered and warrants careful analysis of the data (Jolliffe 2002; McLachlan 2004). Instead of relying on the principal components that explain major variations in the data, a set of principal components having a combination of the low and high eigenvalued components might prove more useful in such situations (Jolliffe 2002). May be, the combination of low and high eigenvalued optimal principal components used for training the classifier was helpful in improving the classifier performance.

As evident in the reported confusion patterns (Tables 9.6 and 9.7), the optimal features based classifier performed better than the one trained using the higher eigenvalued principal components. The classifier performance differences may also be attributed to the set of variables selected through the optimal feature selection algorithm. The higher eigenvalued principal components might have preserved all the variation in the TIVs data but they probably could not find the most effective dimensions of the within-group separation. The optimal features on the other hand, might not have preserved the maximum within-group variation but most probably were able to keep track of the dimensions of within-group variation (Jolliffe 2002; McLachlan 2004; Rao 1964).

Second, the leave-one-out cross-validation results reported in Table 9.7 demonstrate that up to 72% unknown TIRIs may be correctly classified using the proposed computational approach. These results make it obvious that the employed algorithmic approach divided the eigenspace into smaller and linear spaces and there was very little overlap between the divided linear spaces.

Third, the confusion patterns observed during the AFEC of the evoked facial expressions are consistent with the previous studies carried out to investigate the relationship between the emotions and the facial musculo-physiological activities. Like the previous investigations, these classification results and confusion patterns also suggest some similarities between the thermal measurements of some of the evoked facial expressions. Interestingly, previous studies, albeit using the visual cues or measurements of different bio-physiological signals, discovered similar facial expression recognition and confusion patterns. These initial results suggest that infrared measurement of facial thermal features may help in classifying the subtly different evoked or reactive facial expressions.

9.8 A comparison of the intentional and evoked expression classifiers

Examining the relative performance of the 4 classifiers constructed in this thesis (chapters 6, 7, 8 and 9) might help realise the potential of using the facial skin temperature in AFEC and AAR capable systems. Table 9.9 compares the 4 classifiers, highlights their underlying Gaussian mixture models, and presents their respective classification success rates.

When the Gaussian space was constructed using the TIV data taken from the neutral and simulated happy and sad faces (row 2 of Table 9.9), 83.3% of the TIRIs were successfully classified.

When the TIV data measured on the neutral faces and simulated happy, disgusted, positively surprised and angry faces were used to construct the Gaussian space (row 3 of Table 9.9), the complexity of the Gaussian increased and the classification success rate reduced to 67.2%.

When the TIV data measured on the neutral faces and faces with (six) basic intentional expressions were used for constructing the classifier (row 4 of Table 9.9), only 57.1% new and unknown faces were successfully classified.

When the Gaussian space was constructed using the TIV data measured on the neutral faces and the faces with four evoked expressions (last row of Table 9.9), the classifier performed at 72% success rate.

As evident in Table 9.9, the four evoked expression classifier (last row of Table 9.9) performed better than the four intentional expressions classifier (row 3 of Table 9.9). This difference in the performance of these two classifiers highlights the linear division patterns of the two eigenspaces. It is likely that the complexity of a Gaussian mixture model, in addition to the effectiveness of the optimal features, influences the classifier performance.

Extending this discussion and comparing the observed confusion patterns may help understand the differences between the thermal measurements of the evoked and intentional facial expressions.

Table 9.9: Construction of the Gaussian space, employed training features and the classifier performance

Facial expressions and TIV data used for constructing the Gaussian space	**Training features employed**	**Overall classification success rate observed**
Neutral and Intentionally happy and sad expressions	Optimal features recursively drawn from among the derived principal components	*83.3 %*
Intentional happy, disgust, Surprise and angry expressions	-do-	67.2%
Neutral and Intentionally happy, sad, disgust, surprise, angry and fear expressions	-do-	57.1%
Neutral and Evoked happy, sad, disgust, and angry expressions	-do-	72.0%

When the neutral faces and the faces with intentional expressions of happiness and sadness were classified using the optimal features (Chapter 6), 12.5 % of neutral faces were confused with the intentionally sad faces. During an attempt to classify the neutral faces and the faces with basic intentional expressions (Chapter 7), the neutral faces were more frequently confused with the intentional expression of sadness (18.8%). During the classification of evoked facial expressions, the neutral faces were not at all confused with the evoked expressions of happiness or sadness. This confusion pattern might be understood in the light of the discussion on emotion-specific musculo-physiological activities reported in section 3.2 of chapter 3. As reported in section 3.2, (Wolf et al. 2005) discovered that Orbicularis Oculi, Mentalis, and Depressor Anguli Oris contribute to the expression of negative emotions. Probably, simulating the emotion of sadness did not allow enough musculo-physiological and hæmodynamic activities along Orbicularis Oculi, Mentalis, and Depressor Anguli Oris. So the simulated sadness was confused with the neutral face. It seems that the facial muscles were more engaged when the facial expressions were evoked. Consequently, the quantitative differences in the TIV data measured on the neutral faces and on the faces with evoked sadness were different and more helpful in distinguishing between the neutral and sad faces.

When the neutral faces and faces with intentional expressions of sadness and happiness were classified using the optimal features (Chapter 6), 6.5% of neutral faces were confused with intentionally happy faces. During an attempt to classify the neutral faces and the faces with six basic intentional expressions (Chapter 7), the neutral faces were more frequently confused with the happy faces (12.6 %). The neutral faces were not confused with the evoked expressions of happiness. Again, the discussion in section 3.2 of Chapter 3 might help understand this confusion pattern. A significant number of previous studies found that Zygomaticus Major, Orbicularis Oris, Orbicularis Oculi, Mentalis and Platysma contribute to the expression of positive emotional experiences. These muscles were probably not fully activated when the expressions were being simulated. Probably, for this reason, the intentionally happy faces were confused with the neutral faces. However, when emotions were evoked, some significant musculo-physiological and hæmodynamic activities took place along these muscles. The evoked expressions of happiness were therefore not confused with the neutral faces.

The intentional expression of sadness was confused with the expression of happiness (18.8%) when the neutral faces and the faces with intentionally positive and negative expressions were classified using the optimal features (Chapter 6). During an attempt to classify the neutral faces and the faces with simulated basic expressions (Chapter 7), the happy faces were not confused with the sad faces. Around 20% of the evoked expressions of happiness were confused with the evoked sad expressions in chapter 9 so a higher confusion rate was observed. However, only 10% of (evoked) sad faces were confused with the (evoked) happy faces (Chapter 9). This confusion pattern might also be examined in the light of emotion-specific musculo-physiological activities reported earlier in Section 3.2. Studies cited in Chapter 3 suggested that musculo-physiological activities along Zygomaticus Major, Orbicularis Oris, Mentalis and Platysma represent the facial expression of positive emotional experiences. The facial expression of negative emotions in the reported studies involved Corrugator, Masseter, Triangularis, Orbicularis Oculi Palpabraeous, Platysma, and Bucccinator. The cited studies reported some musculo-physiological activities around the Orbicularis during the facial expression of both positive and negative emotions. The musculo-physical activities along this particular muscle might have caused the classifiers to confuse the facial expression of happiness with sadness.

When the intentionally positive and negative facial expressions were classified using the optimal features, 18.8% of the disgusted faces were confused with the angry faces (Chapter 6). Similarly, 18.8% of angry faces were confused with the disgusted faces. During an attempt to classify the neutral and the six basic facial expressions, only 6.3 % faces with disgust were confused with the angry faces but the angry faces were confused with the faces showing disgust in chapter 7. When evoked expressions were classified, 10% of the disgusted faces were confused with the angry faces (Chapter 9). However, more angry faces (20%) were confused with the faces showing disgust. An examination of the negative emotion-specific musculo-physiological activities would explain this confusion pattern. As the studies reported in Chapter 3 suggest, the facial expression of aggression and rage may involve Corrugator, Masseter, Triangularis, Orbicularis Oculi, Palpabraeous, Procerus Nasi, Labii Inferioris and Platysma. The expressions of sadness and fear reportedly involve Frontalis, Palpabraeous Superior and Inferioris, Labii Superioris Orbicularis Oculi, Masseter, Triangularis and Bucccinator.

Since some of these muscles were involved in the expression of sadness, rage, anger and fear, the similarities between the musculo-physiological and hæmodynamic activities around these muscles probably produced the observed confusion pattern.

The confusion patterns observed and reported in Chapters 6, 7, 8 and 9 were realised to be, to a large extent, consistent with the previous studies carried out to investigate the association between emotion and musculo-physiological activities.

Chapter 10

DISCUSSIONS, FUTURE RESEARCH DIRECTIONS AND CONCLUSIONS

This chapter first presents a summary of the investigations and reports the observed classification results. Some important inferences are then drawn from the results. The inferences are analysed to propose dimensions of future work on the use of facial skin temperature measurements in AFEC and AAR capable systems.

10.1 Summary of investigations

Prior to this work, facial thermal features were mainly used for binary detection of stress levels, deceit and anxiety. This thesis, building upon the previous investigations, explores the possibilities of using the facial skin temperature measurements for classifying the facial expressions of most common affective states. This work is based on the scientific findings that suggest that facial expression of emotions would cause changes in the blood volume flow, would influence the musculo-thermal characteristics of the face, and would consequently cause variations in the facial skin temperature. In essence, this investigation focuses on developing an efficient facial thermal feature extraction, selection, representation and classification approach.

This work began by capturing 224 visible and thermal infrared images of 16 participant faces with neutral expression and intentional expressions of happiness, sadness, disgust, surprise, anger and fear were initially acquired for this investigation. At a later stage, 100 visible and thermal infrared images of 10 participants with neutral faces and faces with evoked facial expressions of happiness, sadness, disgust and anger were separately recorded. Hence, a database of 324 discrete, visible-spectrum and infrared images was developed for this investigation.

The acquired TIRIs were first segmented. The edge detection tools available within the thermal analysis software were applied to detect the faces within the acquired TIRIs. The affective-state-specific temperature distributions on the facial skin were examined through the pixel grey-level analysis. As reported in Chapter 4, examining the temperature variations within the selected regions of interest within the TIRIs led to the discovery of some significant facial thermal feature points (FTFPs) along the major facial muscles. The TIVs measured on the FTFPs were then tested for normal distribution and the similarity of the facial expression groups' variance structure.

A purpose specific algorithmic approach, reported in Chapter 5, was developed and employed for classifying the facial expressions. The acquired TIRIs were represented along the principal components (PCs) of a covariance matrix using the singular value decomposition based principal component analysis. The resulting PCs were ranked in the order of their effectiveness in the between-cluster separation. Only the most effective PCs were retained to construct an optimised eigenspace. A supervised learning algorithm was then invoked for linear subdivision of the optimised eigenspace. The statistical significance levels of the classification results were estimated for validating the discriminant functions.

In a series of analyses, several Gaussian mixture models, having a varying number of components, were constructed. The facial thermal variances prevailing in the TIV data were first used to discern between the neutral and pretended happy and sad facial expressions. As reported earlier in Chapter 6, the TIV data in the second analysis were used for classifying the two positive (happy and surprise) and two negative (disgust and angry) facial expressions. A third analysis was carried out to classify the six common pretended facial expressions (reported in Chapter 7). During the fourth analysis, attempts were made to classify the pretended facial expressions assuming that parts of the face were covered or occluded (reported in Chapter 8). In the fifth analysis, evoked facial expressions of happiness, sadness, disgust, and anger were classified using the TIV data gathered from the 10 participant faces (reported in Chapter 9).

10.2 Observations and results

The facial thermal feature extraction, selection, representation and classification approach employed in this work was helpful in distinguishing between the facial expressions of affective states. The employed algorithmic approach achieved excellent classification results on the feature vectors used for training the classifier. However, the developed classifiers could not generalize to the new and unknown thermal faces with the same level of high accuracy.

The inconsistencies observed in the classifiers' performance were traced and reviewed the relevant literature to better understand the observed classification results. Some earlier studies provided insight and explained the rationale for the variations in the performance of different classifiers developed for this investigation. The literature suggested that the relevance and aptness of extracted features, the size and complexity of a Gaussian mixture model, and the composition of the underlying eigenspace might influence the classifier performance. These factors varied at each stage of this investigation and probably caused inconsistencies in the classifier performance.

The literature further suggested that the nature of the variates included in a reduced and optimised eigenspace might also influence the classifier performance. A significant number of earlier studies suggest that a reduction in the feature space dimensions would occasionally produce insufficient degrees of freedom for a consequent linear division of the space. The employed algorithmic approach involved a two-level dimension reduction of the discriminant space. This reduction in the discriminant space dimensions might have contributed to the variations in the classifier performance.

Despite some variations in the classifier performance, the proposed algorithmic approach, in general, allowed developing a set of effective discriminant functions at each stage of this work. The classification results reported in chapters 6, 7, 8 and 9 provide a convincing evidence of the effectiveness of employed optimal feature selection, representation and classification approach.

During the first analysis reported in section 6.1 of chapter 6, the neutral, pretended happy, and sad facial expressions were classified. When the high eigenvalued principal components were used for classifying the neutral and pretended expressions of happiness and sadness, only 45.8% faces could be correctly classified during the cross-validation test. In a continuing analysis, 83.8% of the faces were correctly classified during the cross-validation test when the optimal principal components were used to train the classifier. This significant reduction in the classification error rate suggested that optimal principal components would help in developing a compact feature representation scheme that resulted in a better between-group separation.

Section 6.2 of chapter 6 reported an attempt to distinguish between the two pretended positive (happy and surprise) and two pretended negative (angry and disgust) facial expressions. The classifier trained with the optimal features proved to be more effective (67.20% success rate) than the classifier trained using the high eigenvalued principal components (37.50% success rate). The discriminant eigenspace constructed with the optimal feature resulted in 29.70% reduction in the classification error rate. This variation in the classifier performance might have resulted from an overlap between the thermal features of different expressions represented in the Gaussian mixture model. Also, some degree of similarity between the musculo-physiological activities that took place during the expression of positive and negative affects might have influenced the classifier performance.

Chapter 7 presented an attempt to classify the neutral faces and the faces with six pretended basic facial expressions. When the high eigenvalued principal components were used for classifying the seven facial expressions, only 38.4% faces were successfully classified during the cross-validation tests. The classifier trained with the optimal principal components could correctly classify 57.1 % thermal faces. This ever first attempt of classifying the six basic facial expressions using the facial skin temperature was encouraging. Results suggest that the thermal infrared measurements of facial skin temperature may help in distinguishing between the six basic facial expressions, provided an appropriate feature selection and representation approach is employed.

The possibilities of classifying facial expressions under the pose and illumination conditions and under the assumed occlusion were explored and reported in chapter 8. Each thermal face was divided into four regions: forehead was considered region 1, areas around eyes and cheek were considered region 2, area around the mouth was considered region 3 and area around the chin was considered region 4. Thermal intensity values recorded at the FTFP sites within regions 1 and 2 were grouped together. Thermal intensity values recorded at the FTFP sites within regions 3 and 4 were separately grouped together. The TIVs within the two groups were separately used as input vectors to the facial expression classifiers. When the TIVs in regions 1 and 2 were used for AFEC, a low classification success rate was observed during the cross-validation results. The classifier performed at an even lower success rate when the TIVs in the second grouping were used for training the classifier. Results suggested that the TIV data measured on some selected FTFP sites of the face might not allow distinguishing between the facial expressions. Thus, in order to distinguish between the facial expressions, the TIV should be measured at all the FTFP sites on the face.

The results of classifying the evoked facial expressions, reported in Chapter 9 suggested that the evoked and reactive facial expression might be recognised using the facial thermal features. During the cross-validation tests when optimal features were used to train the classifier, the classifier correctly classified 72% unknown thermal faces. These results therefore suggested that the evoked facial expressions of happiness, sadness, disgust and anger could be classified using the facial thermal features. Results also suggest that the optimal thermal features, when projected in an optimised eigenspace, may reveal the distinguishable facial thermal characteristics. These facial thermal characteristics were transient and were made available through changes in the facial expressions. Though no attempt was made to distinguish between the evoked and pretended facial expressions on the basis of the facial skin temperature, observed results suggested that the two conditions might be realised on the basis of facial skin temperature distributions.

Results of the evoked facial expression classification suggested that the transient facial thermal features generated as a result of evoked emotions might maintain some degree of thermal similarities between the facial expressions. However, these initial results suggested that the subtly different evoked or reactive facial expressions might result in different and distinguishable facial skin temperature measurements at the FTFP sites on the face.

Despite the inconsistencies observed in the classifier performance, the observed classification results, to a large extent, were consistent with the classification results reported in some previous investigations. For example, the facial expression classification results reported in (Chellappa 1998) and (Donato et al. 1999) were similar to the classification and confusion patterns observed in this investigation.

10.3 Some possible inferences

The literature cited earlier in chapters 2, 3, and 5 and the reported results of this investigation suggested that the facial skin temperature measurements might contribute in achieving the AFEC and AAR functionality. The observations reported in chapters 6, 7, 8 and 9 suggested that the facial skin temperature variations caused by a change in affective state could be measured using a non-contact thermal infrared camera. The pixel grey level analyses of the acquired thermal images provided further evidence of previously reported findings that like other bio-physiological signals, the facial skin temperature might change with the changing affective states.

Thermal measurements taken around any particular part of the face or along any particular facial muscle did not provide enough information for classifying the facial expressions. The physical locations of the identified discriminator variables suggested that the AFEC and AAR might require monitoring the thermal changes at multiple locations on the face along the major facial muscles.

Thermal analysis of the TIV data gathered from the participant faces suggested presence of some degree of correlation among the TIV data recorded along the major facial muscles. This variation might have resulted from the contraction and /or expansion of the muscles during the expression of emotions. Perhaps, the human metabolic reaction to the changes in the affective states was influential in changing the rate of heat transfer from the core body to the facial skin surface. Perhaps a change in core body temperature caused some measurable thermal changes at several locations on the face.

The observed classification results also reflect upon the computational efficacy of the proposed algorithmic approach. The statistical classifiers that were trained using the optimal principal components consistently out performed the classifiers that were trained using the high eigenvalued principal components. The linear discriminant algorithm consistently achieved excellent classification results on the training vectors. However, the linear discriminant algorithm consistently generalized to new and unknown thermal faces with a comparatively lesser efficiency. Also, some facial expressions were better classified than the others in each classifier test. These observations highlight the influence of a constructed Gaussian mixture model on the performance of the classifier.

This work, in a broad perspective, suggests the viability of using the facial skin temperature measurements in security and surveillance, clinical diagnosis, criminal investigations and human-computer interaction applications. The future AFEC and

AAR capable systems might possibly be able to employ and rely upon the non-invasive facial skin temperature measurements taken within the infrared light spectrum.

The reported observations and the inferences drawn from this work resulted in realising the following key research directions for the future work.

10.4 Suggestions for future research

For more than three decades, scientists have been investigating the possibilities of using the bio-physiological cues to recognise the affective states and their expressions. This work, building upon the previous investigations, has demonstrated the viability of using the temporal facial thermal features for classifying the affective states and their facial expressions. However, development of a reliable facial features-based AFEC and AAR capable system requires further investigations and validation. The work reported in the previous chapters of this thesis allowed setting an agenda for the future research on use of facial skin temperature in AFEC and AAR capable systems. This work led to identification of the following research dimensions for the future work.

To progress from this point, future work should focus on developing procedures and methods for finding a finer distinction between the expressions of affects. This may warrant accounting for the individual differences while classifying the expressions of affective states. This work also warrants extension of the thermal image database for further verification of the observed results. An extended database will also warrant further validation and cross-validation of the facial expressions related musculo-thermal behaviour. Such validation exercises may help further establish the relevance and reliability of the employed AFEC and AAR approach. A different direction of future research may involve construction of the hybrid AFEC-capable systems through fusion of thermal features with some other visual and non-visual cues.

These proposed key research directions are further elaborated in the following paragraphs.

10.4.1 Finer distinction between the facial expressions

The true nature and extent of association between the intensity of emotion and level of musculo-thermal activities are the areas of active investigations. However, their true relationships have not been discovered and understood yet. Therefore, this important relationship could not be taken into consideration during the thermal image acquisition. The issue of finer distinction between the facial expressions is critical in designing a robust AFEC capable system. The issue should therefore be dealt with more carefully in a future investigation.

The understanding and perception of facial expressions in this work were based on the understanding of physiognomy and the visual appearance of a face. Referees used their own understanding of the facial expressions and facial display of emotions. The effect of intensity of emotion on the facial expressions may influence the nature of musculo-thermal activities and may consequently cause some fluctuations in the facial skin temperature. Also, the observers' interpretation of a facial expression would have influenced the data acquisition process. Though three referees were requested to examine and agree on the facial expressions of affective states, the underlying relationship between the intensities of emotions and their facial expressions could not be taken into account. A future investigation should therefore pay more attention to the association between the intensity of emotions and their facial expressions.

From an application point of view, an AFEC system should also be able to distinguish between the pretended and evoked facial expressions of affective states. An investigation in this direction may result in a better AFEC and AAR functionality. This may also provide a better and reliable AFEC tool to the other scientific communities such as physicians, psychologists, and criminal investigators.

An increased thermal sensitivity of the infrared detectors mounted on the thermal infrared cameras might help better discover the variances in the acquired thermal images. Recent developments in sensors and micro-machine technologies have resulted in the availability of more sensitive thermal infrared detectors (Phillips 2002). Their better thermal sensitivity might help in extracting the finer features and selecting the most effective features. Such a capability will also allow effective measurement of variance in the thermal data. A high-sensitivity thermal camera might also allow acquiring more precise and accurate thermal facial data. A higher accuracy of measurement might also help in finer definition and classification of the facial expressions of affective states.

10.4.2 Individual differences

Several researchers have highlighted the importance of examining the underlying differences between the muscular construction of faces of various racial, ethnic, and geographic groups of people. Different groups of people were also reported to have varying muscular and bio-physiological responses to affective states. Such differences should be understood and accounted for in designing the AFEC and AAR capable systems.

Studies suggest that muscular structure of people belonging to different ethnic and racial backgrounds would differ. Also, certain muscles present in some groups of people might not be present in the other groups of people (Pessa et al. 1998).

Several studies have also reported the differences between the bio-physiological reactions to emotions in men and women. For example, during a study that investigated the gender differences in fingertip response to music, female population's fingertip temperature decreased more significantly than that of the male population (McFarland and Kadish 1991).

Another dimension of the individual differences arises from the emotion-specific response differences between various age groups. Studies suggest that bio-physiological reaction to emotions varies and decreases with the age. Influence of age on emotional response, autonomic responses to emotion and bio-physiological expression of emotions is not fully understood yet (Kunzmann and Gruhn 2005). Mentally ill people and psychopaths are also believed to have slow and different response to emotion stimuli as compare to normal people (Pham et al. 2000). Hence, the age factor should also be considered in a future investigation.

Yet another skin related individual difference arises from the facial skin condition. It has been established that the skin diseases more severely and more frequently influence certain parts of the facial skin. Diseases such as dermatitis, contact urticaria and seborrheic dermatitis influence the biophysical function of the facial skin at some particular parts of the face (Kobayashi and Tagami 2004). Also, factors such as stratum corneum, reported in chapter 8, should be considered in developing and extending the database of infrared images for further investigations.

A second skin health related issue might also have implications for the future work. The issue arises from the fact that when people imagine happy, sad and angry situations, varying patterns of facial-muscular activities are observed through the electromyography (Schawrtz et al. 1976). These facial expression patterns particularly differ between depressed and non-depressed or cheerful people. This difference in facial expression patterns ads an additional dimension to a future investigation.

10.4.3 Extended database of thermal images

To accommodate for the individual differences between various population groups, and to validate the observed classification results, an extended database of sample thermal images should be developed for a future investigation. Sample thermal images should include various ethnic, regional, racial and age groups. People of different skin colours should be sampled separately as well as pooled together to examine the underlying influence of the facial skin colour on the composition of the discriminant rules and the resulting performance of a classifier.

An extended database will allow addressing the issue of individual differences and so, will provide more reliable parametric estimates of the thermal data.

10.4.4 Further validation

It would be beneficial to compare the affective states related patterns of facial skin temperature variations with other human information patterns. The patterns of emotion-specific Energy Expenditure (EE) and the EMG measurement patterns on the face should be studies and compared with the thermal variation patterns. A comparison of the thermal data with other bio-physiological measurements might encourage fusion of multiple physiological signals for AFEC and AAR. Such an investigation might also reveal the possibilities of using local facial information for classifying the covered or occluded faces using the bio-physiological signals. Such a comparison might also provide insight into the similarities (or dissimilarities) between the patterns of EMG, EE and thermal measurements along the major facial muscles.

Another way of validating the classification approach and observed classification results would be to apply some competing pattern recognition methods on the measured TIV data and compare the resulting classification and confusion patterns. The physiological pattern analysis and classification is a comparatively new and less explored domain. Little work has been done on bio-physiological feature extraction and selection for classifying the emotions and affective states. Generally, common pattern recognition techniques are applied on the physiological signals. These techniques do not attempt to overcome the issues pertaining to the nature of bio-physiological features and their measurements. The inherent overlap between the bio-physiological indicators of affects and affective states would usually cause some confusion between the new patterns and a variant of any different pattern. Hence the competing pattern recognition algorithms should be invoked on facial thermal features to further improve the classification results and validate the reported results.

10.4.5 Data fusion

The fusion of multiple bio-physiological, visual and auditory signals should also be explored for developing more robust and reliable AFEC and AAR capable systems. However, fusion of multiple bio-physiological, visual and auditory cues may require a different feature extraction, selection, representation and classification approach. Combining multiple classifiers may also pose some unique implementation related challenges.

10.5 Conclusions

This investigation suggests that the digital infrared imaging of facial thermal features within the 8-14 μm bandwidth of electromagnetic radiations may be used to measure the affective-state-specific thermal variations on the human face. Furthermore, this work makes it obvious that the pixel-grey level analysis of the thermal infrared images may allow localising the thermally significant FTFPs along the major facial muscles of the face.

The uncorrelated principal components of the facial thermal features, when ranked in the order of their effectiveness in the between facial expression group separation, were able to reveal the most effective dimensions of variances in the facial thermal features. The higher eigenvalued input vectors (principal components) were less successfully classified into the facial expressions of affective states. However, the transient thermal feature allowed an effective classification of facial expressions of affective states in an optimized eigenspace of input feature vectors. Consequently, some effective discriminant functions could be developed for the person-independent recognition of the expressions of affective states. The input feature vectors used for training the classifier were more successfully recognised than the new and unknown thermal faces. Furthermore, the Gaussian mixture model with one cluster per affect worked better for some facial expressions than others in this investigation. The observed classification patterns highlighted the influence of a Gaussian mixture model structure on the accuracy of the classification results.

The classification results highlight the efficacy of the novel facial feature extraction, selection, representation, and classification approach proposed in this thesis for achieving the AFEC and AAR functionality. The proposed computational approach, for the first time, was able to classify the facial thermal features for recognising the facial expression of the most common affective states.

This work provided new and convincing evidence that the transient facial thermal features were effective in automated classification of expression of most common affective states. The observed results in this investigation were consistent with the ones reported in several earlier studies. However, further validation of the observed classification patterns may help realise their practical significance and relevance.

APPENDIX I

HUMAN PROTECTION PRACTICES

Since these experiments were conducted at the American University of Sharjah, United Arab Emirates, an American accredited and US incorporated institution, the human protection practices outlined in the US Government's Belmont report (Belmont report 1979, DHEW 1979) were followed during the design of experiments, thermal image acquisition, and dissemination of the resulting information. Details of the observed human protection practices are provided in the following paragraphs.

1.1 Selection of participants

Participants were invited through written notices that were posted on the designated public places (such as student notice boards) within the School of Architecture & Design, School of Engineering and College of Arts & Science of the American University of Sharjah. Participation in this investigation was entirely voluntary. Efforts were made to select participants without any religious, cultural, ethnic, gender or age discrimination. The female population of the university was particularly encouraged to participate in the experiments. The participation of male and female participants with diversified cultural, ethnic backgrounds was deemed necessary for having an appropriate sample of participant faces.

During the first phase of experiments, 16 undergraduate students, 12 boys and 4 girls with a mean reported age of 20 years and 9 months volunteered to participate in the experiments. Participants came from different academic, cultural, racial and ethnic backgrounds. All participants were adult undergraduate students. Volunteering participants included Caucasians, Arabs, Iranians, Indians and Pakistanis.

During the second phase of this work, participants were recruited through new public announcements. Only 3 male students who participated in the previous image acquisition exercise volunteered again to have their thermal images taken with evoked facial expressions. Seven (7) new participants volunteered to have their thermal images recorded. Of these 7 new volunteers, 3 were continuing education students and were not enrolled in a regular academic program at the American University of Sharjah. Mean reported age of these 10 participating volunteers (7 boys and 3 girls) was 21 years 2 months. This group of 10 volunteers included Arabs, Iranians and Indians.

1.2 Compensation and costs

No monetary rewards were offered to participants. The participants were offered digital copies of their recorded infrared and visible spectrum images free of cost.

1.3 Briefing and debriefing

Participants were briefed before the start of each image acquisition session. The purpose and scope of the study were explained and questions were answered to let participants understand the benefits and outcome of these experiments. Participants were told that they were not being put to any short or long term physical risk. Equipment, image acquisition procedure and post-acquisition data analysis methods were explained.

During the second phase of this work when expressions were invoked, participants were briefed as before but they were also informed about the nature and content of emotion invoking images and video clips. They were given an option to discontinue their participation if the content and nature of imagery were not acceptable to them. Sources of imagery were also revealed to the participants. Images and video clips similar to the ones used for actually invoking the expressions were shown to the participants before start of the image acquisition session.

1.4 Procedure for obtaining informed consent

Consent for using each participant's images and thermal data was obtained either during the briefing sessions or before the image acquisition session. All participants allowed publication of numerical and statistical data emanating from their respective thermal images. All male and female participants except 3 female participants allowed use of their respective thermal and visible images for dissemination of information through scholarly journals, periodicals and conference proceedings.

1.5 Risks to participants

Current literature on thermal infrared imaging and infrared imaging equipment (including the manuals that accompanied Cantronix IR 860 thermal infrared camera) suggested no short or long-term physical risk in recording thermal infrared images. Thermal infrared imaging has been used, without consent, upon members of the public at airports, for example to screen for passengers with high temperature possibility caused by contagious respiratory diseases such as bird flu. No physical harm or other risks of thermal infrared imaging were reported in the literature cited in previous chapters of this thesis.

Information pertaining to the suitability and safety of thermal imaging techniques in the context of human protection is available in the cited literature. Chapters 2 and 3 of this thesis refer to the information on human related application of thermal infrared imaging.

1.6 Methods and Procedures

This work aimed to examine the possibilities of distinguishing between the facial expressions using variations in facial skin temperature. Participants' thermal faces were acquired when they showed a normal and neutral face (with all facial muscles in their resting position). During the first phase of this work, participants were asked to intentionally express happiness, sadness, anger, surprise, fear and disgust. Their thermal (and visible spectrum) images were captured while they pretended and expressed emotions. The same procedure was repeated in the second phase of this study when expressions were evoked.

Infrared images were acquired under a normal, controlled and comfortable building environment. Internal temperature of rooms used for conducting experiments varied between 19-22 ºC during the image acquisition. The building air conditioning systems were equipped with a humidity controller and an air recycling system. Each participant was given at least 20 minutes to acclimatize with the environment. Thermal images were captured in several sessions in October 2003, November 2003, April 2004, and September 2004. Images were recorded between 0100 and 0430 pm on the working days.

During image acquisition, visible-spectrum cameras were placed about 2.5 to 3.5 meters away from the participants. To avoid any loss of thermal information, the infrared camera was always placed 1.8 to 2.4 meters away from the participants.

During the first phase of this work when the pretended expressions were recorded, the participants were trained on acting and intentionally expressing the emotions. They were shown still images and video clips to get a better understanding of how to facially express different affective states.

During the second phase of this work when the expressions were evoked, participants were shown still images and video clips to invoke emotions. All video clips and still images were taken from established and ethically responsible organizations such as British Broadcasting Corporation (BBC), the American TV channel CNN, and the print media sources such as New Yorker and Washington Post.

A high quality set of casual and comfortable chairs was offered to participants while they waited for their turn before their images were recorded. Image acquisition time varied for various reasons including participants' ability to express emotions, technical problems in recording images and participant requested breaks. Image recording time varied between 45 minutes and 120 minutes. This variation in time was, at times, also caused by the investigators failure to capture acceptable quality images. During the pretended expression image acquisition, when needed, participants were demonstrated how to express different emotions. Two senior level (final year) undergraduate students were recruited to help during image acquisition. They were trained for using the digital and thermal cameras and recognizing facial expressions

1.7 Data processing and storage

The thermal analysis software CMView Plus was used for reading and analyzing the acquired thermal images. Numerical data obtained from the thermal images were saved in allowed file formats. The data were further analysed using the statistical analysis software SPSS and mathematical analysis software Matlab. All data were stored on a computer hard disk and a backup copy of the data was stored on compact disks.

1.8 Public release of data

Actual data was kept confidential but statistical analyses and relevant results were published in scholarly publications emanating from this work. As mentioned earlier, publication of data and results were discussed with the participants and their consents obtained.

1.9 Description and sources of secondary data

As evident from the reported analytical approaches employed in this investigation and their respective results, no secondary data pertaining to participants were required in this work. The mathematical and statistical analyses carried out during the reported investigations (or any inferences made using the involved analyses) did not require secondary data about the participants. Participants' names, related personal identifiers and their respective departments of study were noted and kept confidential.

APPENDIX II

LIST OF PUBLICATIONS EMANATED FROM THIS WORK

1. Khan, M.M., Ward, R.D., and Ingleby, M. (accepted in 2007- to appear). “Classifying pretended and evoked facial expression of positive and negative affective states using infrared measurement of skin temperature,” *ACM Transactions on Applied Perception*, ACM Press, NY, ISSN: 1544-3558.

2. Khan, M.M., Ward, R.D., Ingleby, M. (2007). *“Automated Classification of Affective States using Facial Thermal Features,"* in S. Singh and M. Singh (Eds.), Progress in Pattern Recognition, IWAPR’07, International Workshop on Advances in Pattern Recognition, Advances in Pattern Recognition Series, London: Springer, pp. 138-144, ISBN: 978-1-84628-944-6.

3. Khan, M.M., Ingleby, M., and Ward, R.D. (2006). “Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variation,” *ACM Transactions on Autonomous and Adaptive Systems*, Vol. 1, No. 1, pp. 91-113, ISSN: 1556-4665.

4. Khan, M.M., Ward, R.D., and Ingleby, M. (2006). *“Infrared thermal sensing of positive and negative facial expressions,”* in the proceedings of the IEEE 2006 Conference on Robotics, Automation and Mechatronics, Bangkok, Thailand, June 2006, pp. 406-411, ISBN: 1-4244-0025-2.

5. Khan, M.M., Ward, R.D. and Ingleby M. (2005). *“Distinguishing facial expressions by thermal imaging using facial thermal feature points,”* in L. Mackinnon, O. Bertelsen and N. Bryan-Kinns (Eds.), The bigger picture, proceedings of HCI 2005, 19th British HCI group Annual Conference, September 2005, Edinburgh: British Computer Society, vol. 2, pp. 10-14, ISBN 1-902505-69-7.

6. Khan, M.M., Ward, R.D. and Ingleby M. (2004). *“Automated classification and recognition of facial expressions,”* in the proceedings of the IEEE 2004 Conference on Cybernetics and Intelligent Systems, pp. 202-206, 1-3 December 2004, Singapore, ISBN 0-7803-8644-2.

References

Abidi, B., Huq, S. and Abidi M. 2004. “Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition,” *In the Proceedings of the International Carnahan Conference on Security Technology*, IEEE 38th Annual 2004 International Carnahan Conference on Security Technology, pp. 325-330.

Acharya, T., and Ajoy R. 2005. *Image Processing: Principles and Applications*, London: Wiley Interscience.

Allanson J. and Fairclough, S.H. 2004. “A research agenda for physiological computing,” *Interacting with computers,* vol. 16, pp. 857-878.

Ang, L.B.P., Belen, E.F., Bernardo, R.A., Boongaling Jr., E.R., Briones, G.H.H., and Coronel, J.B. 2004. Facial expression recognition through pattern analysis of facial muscle movements utilizing electromyogram sensors. *In the proceedings of TENCON 2004-the 2004 IEEE Region 10 Conference - Analog and Digital Techniques in Electrical Engineering,* vol*. 3*, Chiang Mai, Thailand, November 2004, 600-603.

Arkin, R.C., Fujita, M., Takagi T., and Hasegawa, R. 2003. “An ethological and emotional basis for human-computer interaction,” *Robotics and Autonomous Systems*, vol. 42, pp. 191-201.

Baldwin, J.F., Case S.J., and Martin, T. P. 1998. “Machine interpretation of facial expressions,” *BT technology Journal*, vol. 16, no. 3.

Bales, M. 1998. “High-resolution infrared technology for soft –tissue injury detection,” *IEEE Engineering in Medicine and Biology,”* vol. 17, pp. 56-59.

Bartlett, M.S., Hager, J.C., Ekman P., and Sejnowski, T.J. 1999. “Measuring Facial Expressions by Computer Image Analysis,” *Journal of Psychophysiology*, vol. 36, no. 2.

Bartneck, C. 2001. “How convincing is Mr. Data’s smile: Affective expressions of machines,” *User Modeling and User-Adapted Interaction*, vol. 11, pp. 279-295.

Belhumeur, P.N., Jespanha J.P., and Kriegman, D.J. 1997. “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*,” vol. 19, no. 7, pp. 711-720.

Belmont Report. 1979. *Ethical principles and guidelines for the protection of human subjects of research*, available online: http://ohsr.od.nih.gov/guidelines/belmont.html, United States Government, National Institute of Health, Washington D.C.

Berkey, C.S., Laird, N.M., Valadian, I., and Gardner, J. 1991. “Modeling adolescent blood pressure patterns and their prediction of adult pressure,” *Biometrics*, vol. 47, no. 3, pp. 1005-1018.

Bishop, C.M. 1995. *Neural networks for pattern recognition*, England: Oxford University Press.

Black, M.J., and Yacoob, Y. 1997. “Recognizing facial expressions in image sequences using local parameterized models of image motion,” *International Journal of Computer Vision*, vol. 25, pp. 23-48.

Blue, J.L., Candela, G.T. Grother, P.J., Chellappa R., and Wilson, C.L. 1994. “Evaluation of pattern classifiers for fingerprint and OCR applications,” *Pattern Recognition*, vol. 27, no. 4, pp. 485-501.

Blum, A.L., and Langley, P. 1997. “Selection of relevant features and examples in machine learning,” *Artificial Intelligence,* vol. 97, pp. 246-271.

Bolle, R.M., Connell, J.H., Pankanti, S., Ratha, N.K., Senior, A.W. 2004. *Guide to Biometrics*, New York: Springer.

Borowski, E.J., and Borwein, J.M. 1991. *The Harper Collins Dictionary of Mathematics*, New York: Harper Collins Publishers.

Boulic, L.E.R., and Thalmann, D. 1998. “Interacting with virtual humans through body actions,” *IEEE Computer Graphics and Applications*,” January/February 98, pp. 8-11.

Bozionelos, N. 2001. “The relationship of instrumental and expressive traits with computer anxiety,” *Journal of Personality and Individual differences,* vol. 31, no.6, pp. 955-974.

Bradley, M.M., Sabatinelli, D., Lang, P.J., Fitzsimmons, J.R., King, W., Desai, P. 2003. “Activation of the visual cortex in motivated attention,” *Behavioral Neuroscience,* vol. 117 pp. 369–380.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and Regression Trees,* California: Wadsworth Publishing.

Briese, E. 1995. “Emotional hyperthermia and performance in humans,” *Physiological Behavior*, vol. 58, no. 3, pp. 615-618.

Brooks, R.A. 2002. *Flesh and Machines: How Robots will change us,* New York: Pantheon Books.

Brosnan, M.J. 1998. “The impact of computer anxiety and self-efficacy upon performance,” *Journal of Computer Assisted Learning*, vol. 14, no. 3, pp. 223-234.

Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C.-M., Kazemzadeh, A., Lee, S., Neumann, U., Narayanan, S. 2004. “Analysis of emotion recognition using facial expressions, speech and multimodal information,” *In the Proceedings of 6th International Conference on Multimodal Interface*, ICMI’04, PA, pp. 205-211.

Cabanac, A.J., and Guillemette, M. 2001. “Temperature and heart rate as stress indicators of handled common eider,” *Physiological Behavior,* vol. 74, nos. 4-5, pp. 475-479.

Cacioppo, J.T., Bush L.K., and Tassinary, L.G. 1990. “Microexpressive facial actions as function of affective stimuli: replication and extension, “ *Personality and Social Psychology Bulletin*, vol. 18, pp. 515-526.

Calder, A.J., Burton, A.M., Miller, P., Young, A.W., Akamatsu, S. 2001. “A principal component analysis of facial expressions,” *Vision Research,* vol. 41, pp. 1179-1208.

CANTRONIC Systems Inc. 2001. *IR 860 User Manual*, BC, Canada Coquitlam: Cantronic Systems Inc.

CANTRONIC Systems Inc. 2001[a]. *CMView Plus Imaging Software Manual*, BC, Canada, Coquitlam: Cantronic Systems Inc.

CANTRONIC Systems Inc. 2002. IR 860 Product Data Sheet, available at http://www.cantronix.com/ir860_spec.html, last visited 18 February 2002.

Chang, W.C. 1983. “On using principal components before separating a mixture of two multivariate normal distributions,” *Applied Statistics*, vol. 32, pp. 267-275.

Chatfield C., and Collins, A.J. 1995. *Introduction to Multivariate Analysis,* London, Chapman & Hall.

Chellappa, R. 1998. “Discriminant analysis for face recognition,” in *Face recognition: From theory to applications*, Wechsler, S., Phillips, J., Bruce, V. Fogelman-Soulie, F., and Huang, T., (Eds.), Springer-Verlag, London, pp. 564-571.

Chen, X., and Huang, T. 2003. “Facial expression recognition: A clustering-based approach,” *Pattern Recognition Letters*, vol. 24, pp. 1295-1302.

Choi, S.C. 1972. “Classification of multiple observed data,” *Biomedical Journal*, vol. 14, pp. 8-11.

Christie, I. C. and Friedman, B. H. 2004. “Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach,” *International Journal of Psychophysiology*, vol. 51, pp. 143-153.

Coakes, S.J., and Steed, L.G. 1999. *SPSS: Analysis without anguish,* New York: John Wiley & Sons.

Cohen, I., Sebe, N., Garg, A., Chen, L.S., and Huang, T.S. 2003. "Facial expression recognition from video sequences: temporal and static modeling, "*Journal of Computer Vision and Image Understanding*, no. 91, pp. 160-187.

Cohen, J. 1977. *Statistical Power Analysis for Behavioral Sciences,* New York: Academic Press.

Cohn, J., Zlochower, A., Lien J., and Kanade, T. 1999. "Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding," *Psychophysiology*, vol. 36, pp. 35-43.

Collet, C., Vernet-Maury, E., Delhomme, G., and Dittmar, A. 1997."Autonomic nervous system response patterns specificity to basic emotions," *Journal of Autonomic Nervous System*, vol. 62, pp. 45-57.

Costanza, M.C., and Affifi, A.A. 1979. "Comparison of stopping rules in forward stepwise discriminant analysis," *Journal of American Statistical Association*, vol. 74, pp. 777-785.

Cottrell G., and Metacalfe, J. 1991. "EMPATH: Face, gender and emotion recognition using holons," in *Advances in Neural Information Processing Systems*, Lippman, R., Moody J., and Touretzky D., (Eds.), Morgan Kaufmann Pub., CA, vol. 3, pp. 564-571.

Critchley, H.D, Daly, E.M., Bullmore, E.T., Williams, S.C.R., Amelsvoort, T.V., Robertson, D.M., Rowe, A., Phillips, M., McAlonan, G., Howllin P., and Murphy, D.G.M. 2000. "The functional neuroanantomy of social behavior: changes in cerebral blood flow when people with autistic disorder process facial expressions," *Brain*, no. 123, pp. 2203-212.

Darabi A., and Maldague, X. 2002. "Neural network based defect detection and depth estimation in TNDE," *NDT&E International*, no. 35, pp. 165-175.

Dautenhahn, K., and Billard, A. 1999. "Bringing up robots or psychology of socially intelligent robots: from theory to implementation," *In the Proceedings of 3rd 2002 International Conference on Autonomous Agents*, Seattle, WA, pp. 366-367.

DeCarlo, D., and Metaxas, D. 2000. "Optical flow constraints to deformable models with applications to face tracking," *International Journal of Computer Vision*, vol. 38, no. 2, pp. 99-127.

DeSilva, L.C., Miyasato, T. and Nakatsu, R. 1997. "Facial emotion recognition using multimodal information," *In the Proceedings of IEEE information, Communication and Signal Processing Conference 1997*, pp. 397-401.

DHEW. 1979. Department of Health, Education and Welfare. Publication Number (OS) 78-0014. US Government Printing Office, Washington, D.C.

Diakides, N.A. 1998. "Infrared Imaging: An emerging technology in medicine," *IEEE Engineering in Medicine and Biology*, pp. 17-18.

Dimberg U., and Petterson, M. 2000. "Facial reactions to happy and angry facial expressions: evidence for right hemisphere dominance," *Psychophysiology*, vol. 37, no. 5, pp. 693-696.

Dimberg, U. 1990a. "Facial electromyography and emotional reactions," *Psychophysiology*, vol. 27, no. 5, pp. 481-494.

Dimberg, U. 1990b. "Facial reactions to auditory stimuli: sex differences," *Scandinavian Journal of Psychology*, vol. 31, no. 3, pp.228-233.

Dimberg, U., Thunberg, M., and Elmehed, K. 2000. "Unconscious facial reactions to emotional facial expressions," *Psychology Science*, vol. 11, no. 1, pp. 86-89.

Donato, G. 1999. Bartlett, M.S., Hager, J.C., Ekman, P., and Sejnowski, T.J., "Classifying Facial Expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974-989.

Dror, I.E., Péron, A.E., Hind, S.L., and Charlton, D. 2005. When emotions get better of us: The effect of contextual top-down processing on matching fingertips," *Applied Cognitive Psychology*, vol. 19, pp. 799-809.

Drummond, P.D., and Lance, J.W. 1987. "Facial flushing and sweating mediated by the sympathetic nervous system," *Brain*, vol. 110, pp. 793-803.

Du, Y. and Lin, X. 2003. "Emotional facial expression model building," *Pattern Recognition Letters*, vol. 24, pp. 2923-2934.

Dubuisson, S., Davoine F., and Masson, M. 2002. "A solution for facial expression representation and recognition," *Signal Processing: Image Communication*," vol. 17, pp. 657-673.

Duda, R.O., Hart, P.E., Stork, D.G. 2001. *Pattern Classification*, New York: Wiley Interscience.

Ekman, P. 1992. "An argument for basic emotions," *Cognition and Emotion*, vol.6, pp. 169-200.

Ekman, P., and Friesen, W.V. 1971. "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, pp. 124-129.

Ekman, P., and Friesen, W.V. 1978. Facial Action Coding System: A technique for the measurement off facial movement, Pal Alto, CA, *Consulting Psychology Press*.

Ekman, P., Davidson, R.J. and Friesen, W.V. 2000. "Duchenne's smiles: Emotional expression and their brain physiology II," *Journal of Personality and Social Psychology*, vol. 58, pp.342-353.

Ekman, P. 1982. *Emotion in the Human Face*, England: Cambridge University Press.

Ekman, P., Friesen, W.V., and O'Sullivan, M. 1988. "Smiles when lying," *Journal of Personality and Social Psychology*, vol. 54, pp.414-420.

Ekman, P., Huang, T.S., Sejnowski, T.J., and Hager, J.C. 1993. "Final report to NSF of the planning workshop on facial expression understanding," Human-Computer Interaction Laboratory, University of California, San Francisco, (also available on http://face-emotion.com/dataface/nsfrept/nsf_contents.html, last visited 13 December 2004).

Ekman, P., Levenson, R.W. and Friesen, W.V. 1983. “Autonomic nervous system activity distinguishes among emotions,” *Science*, vol. 221, pp. 1208-1210.

Essa, I., Pentland, A. 1997. “Coding, analysis, interpretation and recognition of facial expressions,” *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 19, no. 7, pp. 757-763.

Eveland, C.K., Socolinsky, D.A., Wolf, L.B. 2003. “Tracking human faces in infrared video, *Image and Vision Computing*, vol. 21, no. 7, pp. 579-590.

Everitt B.S., and Dunn, G. 1991. *Applied Multivariate Data Analysis*, London: John Wiley and Sons.

Fasel B., and Luettin, J. 2003. “Automatic facial expression analysis: a survey,” *Pattern Recognition*, vol. 36, pp. 259-275.

Field, A. 2000. *Discovering Statistics using SPSS for Windows*. London: Sage Publications.

Fried, L.A. 1976. *Anatomy of the head, neck, face and jaws*, Philadelphia: Lea and Febiger.

Friedman, S. M. 1970. *Visual Anatomy*, vol. I, Head and Neck, New York: Harper and Row.

Fujimasa, I. 1998. “Pathophysiological expression and analysis of infrared thermal images,” *IEEE Engineering in Medicine and Biology*, vol. 17, no. 4, pp. 34-42.

Fujimasa, I., Chinzei, T., and Saito, I. 2000. “Converting far infrared image information to other physiological data,” *IEEE Engineering in Medicine and Biology*, vol. 19, no. 3, pp. 71-76.

Fukunaga, K. 1990. *Statistical Pattern Recognition*, London: Academic Press.

Gao, Y., Leung, M.K.H., Hui S.C., and Tananda, M.W. 2003. “Facial expression recognition form line-based caricatures,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 33, no. 3, pp. 407-412.

Garbey, M., Merla, A. and Pavlidia I. 2004. “Estimation of blood flow speed and vessel location from thermal video,” *In the Proceedings of CVRP 2004, IEEE 2004 Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1356-1363, Washington DC.

Gavhed, D., Makinen, T., Holmer I., and Rintamaki, H. 2000. “Face temperature and cardiorespiratory responses to wind in thermoneutral and cool subjects exposed to –10 ºC,” *European Journal of Applied Physiology*, vol. 83, pp. 449-456.

George D., and Mallery, P. 1995. *SPSS/ PC+ Step by Step: A simple guide and reference*. London: Wadsworth Publishing Co.

Gong, S., McKenna S.J., and Psarrou, A. 2000. *Dynamic Vision: From Images to Face Recognition*, London: Imperial College Press.

Gonzalez, R.C., and Woods, R.E. 2002. *Digital Image Processing*. New York: Addison-Wesley Publishing Inc.

Gottumukkal R., and Asari, V.K. 2004. “An improved face recognition technique based on modular PCA approach,” *Pattern Recognition Letters*, vol. 25, pp. 429-436.

Gupta, A.K., and Logan, T.P. 1990. “On a multiple observations model in discriminant analysis,” *Journal of Statistics and Computer Simulation,* vol. 34, pp. 119–132.

Gur, R.C., Sara, R., Hagendoorn, M., Maron, O., Hughett, P., Macy, L., Turner, T., Bajcsy, R., Posner A., and Gur, R.E. 2002. “A method of obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies,” *Journal of neuroscience methods*, no. 115, pp. 137-143.

Hager, J. C. 1985. “A comparison of units for visually measuring facial actions,” *Journal of Behavior Research, Methods, Instruments and Computers*, vol. 17, pp. 450-468.

Hara, F. and Kobayashi, H. 1997. “State of Art in component development for interactive communications with humans,” *Advanced Robotics*, vol. 11, no. 6, pp. 585-604.

Haussecker, H.W., and Fleet, D.J. 2000. “Computing optical flow with physical models of brightness variation,” *In the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000*, vol. 2, IEEE Computer Society, pp. 760-767.

Haykin, S. 1994. *Neural Networks: A comprehensive foundation*, New York: McMillian Publishing.

Head, M., and Dyson, S. 2001. “Talking temperature of equine thermography,” *The Veterinary Journal*, no. 162, pp. 166-167.

Heijden, F., Duin, R.P. W., Ridder D., and Tax, D.M.J. 2004. *Classification, Parameter Estimation and State Estimation*, Western Sussex, England: John Wiley & Sons.

Henderson, R., Podd, J., Smith M., and Varela-Alvarez, H. 1995. “An examination of four user-based software evaluation methods,” *Interacting with Computers*, vol. 7, no. 4, pp. 412-432.

Herry, C.L., and Frize, M. 2002. “Digital processing techniques for the assessment of pain with infrared thermal imaging,” *In the Proceedings of the 2002 IEEE International Conference on Engineering in Medicine and Biology, IEEE EMBS 2002*, pp. 1157-1158, Houston, October 23-26.

Hess, U., Kappas, A., McHugo, G.J., Lanzetta, J.T., and Kleck, R.E. 1992. “The facilitative effect of facial expression on self-generation of emotion,” *Psychophysiology*, vol. 12, no.3, pp. 251-265.

Hirsch, C. and Mathews, A. 1997. “Interpretative inferences when reading about emotional events,” *Behavioral Research Therapy*, vol. 35, no. 12, pp. 1123-32.

Holt, R.J., Huang, T.S., Netravali, A.N., and Qian, R.J. 1997. “Determining articulated motion from perspective views,” *Pattern Recognition*, vol. 30, no. 6, pp. 585-604.

Hosseini, H.G., and Krechowec, Z. 2004. “Facial expression analysis for estimating patients emotional states in RPMS,” *In the Proceedings of the 2004 IEEE International Conference on Engineering in Medicine and Biology, IEEE EMBC 2004*, vol. 2, pp. 1517-1520, San Francisco, September 1-5.

Huang D., and Yan, H. 2001. “Modeling and animation of human expressions using NURBS curves based on facial anatomy,” *Signal Processing: Image Communication*, vol. 17, pp. 457-465.

Huang, C., and Huang, Y. 1997. “Facial expression recognition using model-based feature extraction and action parameters classification,” *Journal of Visual Communication and Image Representation*, vol. 8, no. 3, pp. 278-290.

Huberty C.J. 1984 “Issues in the use and interpretation of discriminant analysis,” *Psychological Bulletin*, vol. 95, no. 1, pp. 156-171.

Huberty, C.J. 1994. *Applied Discriminant Analysis*, New York: Wiley.

Hussein, S.E., and Granat, M.H. 2002. “Intention detection using a neuro-fuzzy EMG classifier,” *IEEE Engineering in Medicine and Biology,* vol.21, no. 6, pp. 123-129.

IAPS. 1997. “International affective picture system technical manual and affective ratings,” available online. http://www.unifesp.br/adap/instructions.pdf.

Iwase, M., Ouchi, Y., Okada, H., Yokohama, C., Nobezawa, S., Yoshikawa, E., Tsukada, H., Takeda, M., Yamagguti, K., Kuratsune, H., Shimizu A., and Watanabe, Y. 2002. “Neural substrates of human facial expression of pleasant emotion induced by comic films: a PET study,” *Neuroimage*, vol. 17, no.2, pp. 758-768.

Izard, C.E. 1979. “The maximally discriminative facial movement coding system (MAX),”unpublished manuscript available at the University of Delaware Library through Instructional Resource Center. Inter-library loan services. Delaware.

Jolliffe, I.T. 2002. *Principal Component Analysis,* New York: Springer-Verlag.

Jones B.F., and Plassmann, P. 2002. "Digital infrared thermal imaging of human skin," *IEEE Engineering in medicine and biology*, vol. 21, no.6, pp. 41-48.

Jones, B.F. 1998. "A reappraisal of use of infrared thermal image analysis in medicine," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1019-1027.

Jones, C.H., Ring E.F.J., and Clark, R.P. 1988. "Medical Thermography" in *Applications of Thermal Imaging* Burnay, S. G., Williams, T. L. and Jones, C. H. N. (Eds.), Adam Hilger, Bristol.

Kakadiaris, I.A., Passalis, G., Theoharis, T., Toderici, G., Konstantinidis, I. and Murtuza, N. 2005a. "8D-THERMO CAM: Combination of geometry with physiological information for face recognition," *In the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 2, pp. 1183.

Kakadiaris, I.A., Passalis, G., Theoharis, T., Toderici, G., Konstantinidis, I., and Murtuza, N. 2005b. "Multimodal face recognition: Combination of geometry with physiological information," *In the Proceedings of the2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 2, IEEE Computer Society, pp. 1022-1029.

Kall, R. 1990. "Emotional self regulation and facial expression muscle measurement and training," in *Clinical EMG for surface recording*, vol. *2,* J.R. Cram and J.V. Basmalian (Eds.), 1990, Clinical Resources, Nevada City, CA.

Karat, C. M., Halverson, C., Horn, D., Karat, J. 1999. "Patterns of entry and correction in large vocabulary continuous speech recognition system," *In the Proceedings of 1999 International conference on Computer-Human Interaction*, pp. 568-575.

Kearney G.D., and McKenzie, S. 1993. "Machine Interpretation of Emotion: Design of Memory-Based Expert Systems for Interpreting Facial Expressions in Terms of Signaled Emotions," *Cognitive Science*, vol. 17, no. 4, pp. 589-622.

Khan, M.M., Ward, R.D. and Ingleby, M. 2004. "Automated classification and recognition of facial expressions using infrared thermal imaging," *In the Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems*, Singapore, December 2004, pp. 202-206.

Khan, M.M., Ward, R.D., and Ingleby, M. 2005. "Distinguishing facial expressions by thermal imaging using facial thermal feature points," *In the Proceedings of 19th British HCI Group Annual Conference HCI 2005*, vol. 2, September 2005, Edinburgh, UK, L. MacKinnon, O. Bertelsen and N. Bryan-Kinns (Eds.), British Computer Society, Scotland, pp. 10-14.

Kim, H.K., Bang S.W., and Kim, S.R. 2004. "Emotion recognition system using short term monitoring of physiological signals," *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 419-427.

Kim, H-C., Kim, D., Bang, S.Y. 2003. "An efficient model order for PCA mixture model," *Pattern Recognition Letters*, vol. 24, pp. 1385-809, 1393.

Kinnear, P.R., and Gray, C.D. 2000. *SPSS for Windows made simple*, East Sussex, England: Psychology Press Ltd.

Kirby, M., Sirovich, L. 1990. "Applications of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*," vol. 12, pp. 103-108.

Kistler, A., Mariauzouls, C., von Berlepsch, K. 1998. "Fingertip temperature as an indicator for sympathetic responses," *International Journal of Psychophysiology*, vol. 29, no. 1, pp. 35-41.

Klein, J., Moon, Y., Picard, R.W. 2002. "This computer responds to user frustration: Theory, design and results," *Interacting with computers*, vol. 14, no. 2, pp.119-140.

Kobayashi, H. and Tagami, H. 2004. "Distinct locational differences observable in biophysical functions of the facial skin: with special emphasis on the poor functional properties of the stratum corneum of the perioral region," *International Journal of Cosmetic Science*, vol. 26, pp. 91-101.

Kong, S.G., Heo, J., Abidi, B.R., Paik, J. and Abidi, M.A. 2005. "Recent Advances in visual and infrared face recognition – a review," *Computer Vision and Image Understanding*, vol. 97, no. 1, pp. 103-135.

Koyama, N., Hirata, K., Hori, K., Dan K., and Yokota, T. 2002. "Biphasic vascomotor reflex responses of the hand skin following intradermal injection of melittin into the forearm skin," *European Journal of Pain*, no. 6, pp. 447-453.

Krishnan, Samudravijaya, K., and Rao, P.V.S. 1996. "Feature selection for pattern classification with Gaussian mixture models: a new objective criterion," *Pattern Recognition Letters*, vol. 17, pp. 803-809.

Kunzmann, U., and Gruhn, D. 2005. "Age differences in emotional reactivity: the sample case of sadness," *Psychology and Aging*, vol. 20, no. 1, pp. 47-59, 2005.

Kurse, P.W. 2001. *Uncooled Thermal Imaging: Analysis, Systems and Applications*, SPIE Press, Washington.

Lanitis, A., Taylor, C., Cootes, T. 1997. "Automatic Interpretation and Coding of face images using flexible models," *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 19, no. 7, pp. 743-756.

Layman, D. 2006. *Medical Terminology Demystified*, New York: McGraw-Hill Inc.

LeBlanc, J., Blais, B., Barabe B., and Cote, J. 1976. "Effects of temperature and wind on facial temperature, heat rate and sensation," *Journal of Applied Physiology*, vol. 40, pp. 127-131.

Lee, H.K., and Kim, J.H. 1999. "An HMM based threshold model approach for gesture recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 961-973.

Lien, J., Kanade, T., Cohn, J., and Ching-Chung, L. 1998. “Subtly different facial expression recognition and expression intensity estimation, *In the Proceedings of IEEE 1998 conference on computer vision and pattern recognition*, Santa Barbara, CA.

Lindgren, D. and Spångéus, P. 2004. “A novel feature extraction algorithm for asymmetric classification,” *IEEE Sensors Journal*, vol. 5, no. 5, pp. 643-650.

Lisetti, C. L, and Nasoz, F. 2004. “Using non-invasive wearable computers to recognize human emotions from physiological signals,” *EURASIP Journal of Applied Signal Processing,* no. 11, pp. 1672-1687.

Lisetti, C.S., and Schiano, D.J. 2000. “Automatic facial expression interpretation: Where human-computer interaction, artificial intelligence and cognitive science intersect,” *Pragmatics and cognition*, vol. 8, no. 1, pp. 185-235.

Liu, H., and Motoda, H. 1998. *Feature Extraction, Construction and Selection: A data mining perspective*, London: Kluwer Academic Publishers.

Lundqvist, L.O. 1995. “Facial EMG reactions to facial expressions: a case of facial emotional contagion,” *Scandinavian Journal of Psychology*, vol. 36, no. 2, pp. 130-141.

Lyons, M., Budynek J., and Akamatsu, S. 1999. “Classifying images of facial expression using a Gabor wavelet representation,” *In the Proceedings of 2nd International Conference on Cognitive Science*, Tokyo, Japan, 1999, pp. 113-118.

Manly, B.F.J., *Multivariate Statistical Methods*, London: Chapman and Hall, 1994.

Mase, K. 1991. “Recognition of facial expressions from optical flow,” *IEICE Transactions: Special Issue Computer Vision and its Applications*, vol. 74, pp. 3474-3483.

Matsuzaki, H., and Mizote, M. 1996. “Measurement of facial temperature fluctuations by thermal image analysis,” *Progress in Biophysics and Molecular Biology*, vol. 65, supplement 1, pp. 185-186.

McFarland, R.A., and Kadish, R. 1991. “Sex differences in finger temperature response to music,” *International Journal of Psychophysiology*, vol. 11, no. 3, pp. 295-298.

McGimpsey, J.G., Vaidya, A., Biagioni P.A., and Lamey, P.J. 2000. “Role of thermography is the assessment of infraorbital nerve injury after malar fractures,” *British journal of Oral and Maxillofacial Surgery*,” no. 38, pp. 581-584.

McLachlan, G. J., 2004. *Discriminant Analysis and Statistical Pattern Recognition*, New Jersey: Wiley & Sons.

Meyer S., and Rakotonirainy, A. 2003. “A survey of research on context-aware homes” in *Conferences in Research and Practice in Information Technology*, vol. *21*, Australian Computer Society, Sydney, 2003, C. Johnson, P. Montague and C. Steketee (Eds.), ACS, Canberra.

Mikel, J.A., Fredrickson, B.L., Larkin, G.R., Lindberg, C.M., Mglio, S., and Reuter-Lorenz, P.A. 2005. “Emotional category data on images from the International Affective Picture System,” *Behavior Research Methods*, vol. 37, no. 4, pp. 626-630.

Morishima, S. 2001. “Face analysis and synthesis,” *IEEE Signal Processing Magazine*, vol. 18, no. 3, pp. 26-34.

Moriyama, T., Kanade, T., Cohn, J., Xiao, J., Ambadar, Z., Gao, J., and Imanura, M. 2002. “Automatic recognition of eye blinking in spontaneously occurring behavior,” *In the Proceedings of the 16th International Conference on Pattern Recognition*, ICPR’ 2002, vol. 4, pp. 78-81.

Murphy, S.T. and Zajonc, R.B. 1993. *“Affect, cognition and awareness: affective priming with optimal and suboptimal stimulus exposures,” Journal of Personality and Social Psychology*, vol. 64, pp. 723-739.

Naemura, A., Tsuda, K., Suzuki, N. 1993. “*Effects of loud noise on nasal skin temperature*,” originally published in Japanese, Shingrigaku Kenkyu, vol. 64, no. 1, pp. 51-54.

Nakayam, K., Goto, S., Kuraoka K., and Nakamura, K. 2005. “Decrease in nasal temperature of rhesus monkeys (*Macaca mulatta*) in negative emotional state,” *Journal of Physiology and Behavior*, vol. 84, pp. 783-790.

Nanavati, S., Thieme, M., Nanavati, R. 2002. *Biometrics: Identity Verification in a Networked World*, New York: John Wiley & Sons.

Netter F.H., and Hansen, J.T. 2002. *Atlas of Human Anatomy*, 3rd edition, California: ICON Learning Systems.

Niedenthal, P.M., Halberstadt, J.B., Margolin, J. and Innes-Ker, A.H. 2000. *“Emotional state and the detection of change in facial expression of emotion,” European Journal of Social Psychology*, vol. 30, pp. 211-222.

Niemic, C.P. 2002. *“Studies of Emotions: A theoretical and empirical review of psychophysiological studies of emotion,”* Journal of Undergraduate Research, vol.1, no. 1, pp. 15-18.

Ogasawara, T., Kitagawa, Y., Ogawa, T., Yamada, T., Kawamura Y., and Sano, K. 2001. “*MR Imaging and Thermography of facial angioedema: A case report*,” Oral Surgery Oral Medicine Oral Pathology, vol. 92, no. 4, pp. 473-476.

Ohnishi, N., and Sugie, N. 1996. “Visual-auditory interfaces for machines that serve humans,” *Robotics and Autonomous Systems*, vol. 18, pp. 243-249.

Olivier, B., Zethof, T., Pattij, T., Van Boogaert, M., Van Oorshot, R., Leahy, C., Oosting, R., Bouwknecht, A., Veening, J., Van der, Gugten, and Groenink, L. 2003. “Stress-induced hypothermia and anxiety: pharmacological validation,” *European Journal of Pharmacology*, vol. 28, no. 463(1-3), pp. 117-132.

Otsuka, K., Okada, S., Hassan, M., Togawa, T. 2002. “Imaging of skin thermal properties with estimation of ambient radiation,” *IEEE Engineering in Medicine and Biology*, vol. 21, no. 6, pp. 49-55.

Palomba, D., Sarlo, M., Angrilli, A., Mini, A., Stegagno, L. 2000. "Cardiac responses associated with affective processing of unpleasant film stimuli," *International Journal of Psychophysiology*, vol. 36 pp. 45–57.

Pantic, M., and Rothkrantz, L.J.M. 2000. "Automatic analysis of facial expressions: The state of the art," *IEEE Transactions on Pattern analysis and Machine Understanding*," vol. 22, no. 12, pp. 1424-1445.

Partala, T., Surakka, V., and Vanhala, T. 2006. "Real-time estimation of emotional experience from facial expressions," *Interacting with Computers,* vol. 18, pp. 208-226.

Paterno, F. 2005. "Model-based tools for pervasive usability," *Interacting with computers*, vol. 17, pp. 291-315.

Paul J.L., and Lupo, J.C. 2002. "From tanks to tumors," *IEEE Engineering in Medicine and Biology*, vol. 21, no. 6, pp. 34-35.

Pavlidis I. and Levine, J. 2002. "Thermal image analysis for polygraph testing," *IEEE Engineering in Medicine and Biology*, vol. 21, no.6 pp. 56-64.

Pavlidis, I. 2000. "Lie detection using thermal imaging," *In Thermal facial screening: Proceedings of SPIE – Thermosense XXVI, Annual conference of the International Society for Optical Engineering*, April 2004, Bellingham, USA, pp. 270-279.

Pavlovic, V.I., Sharma R., and Huang, T.S. 1997. "Visual interpretation of hand gestures for Human-Computer Interaction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 677-695.

Pentland A., and Choudhury, T. 2000. "Face recognition for smart environments," *IEEE Computer*, February 2000, pp. 50-55.

Pessa, J., Zadoo, V., Gerza, P., Adrian, E.J., Dewitt, A., and Garza, J. 1998. "Double or bifid zygomatcius major muscle: anatomy, incidence and clinical correlation, *Journal of Clinical Anatomy*, vol. 11, pp. 310-313.

Pham, T.H., Philippot, P., Rime, B. 2000. “Subjective and autonomic responses to emotion induction in psychopaths,” *Encaphale*, vol. 26, no. 1, pp. 45-51.

Phillips, T.J. 2002. “High performance thermal imaging technology,” *Advanced Semiconductor Magazine*, vol. 15, no. 7, pp. 32-36.

Picard, R.W. 1999. “Affective computing in HCI,” in Human Computer Interaction: Ergonomics and User Interfaces, *In the Proceedings of HCI International ’1999, 8th International Conference on Human Computer Interaction*, vol. 1, pp. 829-833.

Picard, R.W. 2000. *Affective Computing*, MA: MIT Press.

Picard, R.W., and Klein, J. 2002. “Computers that recognize and response to user emotion: theoretical and practical implications,” *Interacting with Computers*, vol. 14, pp. 141-169.

Picard, R.W., Vyzas, E., and Healey, J. 2001. Toward machine emotional intelligence: Analysis of affective physiological state,” *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 23, no. 10, pp. 1175-1191.

Pizzagalli, D., Koenig, T., Regard M., and Lehmann, D. 1998. “Faces and emotion: brain electric field sources during covert emotional processing,” *Neuropsychologia*, vol. 36, no. 4, pp. 323-332.

Pollina, D.A., Dollins, A.B., Senter, S.M., Brown, T.E., Pavlidis, I., Levine, J.A., and Ryan, A.H. 2006. “Facial skin surface temperature changes during a concealed information test,” *Annals of Biomedical Engineering*, vol. 34, no. 7, pp. 1182-1189.

Plutchik, R. 1980. “A general psychoevolutionary theory of emotion,” in *Emotion: Theory, research and experience,* vol*. 1,* R. Plutchik and H. Kellerman (Eds.), Hillsdale, New Jersey, pp.3-31.

Posamentier M.T., and Abdi, H. 2003. “Processing faces and facial expressions,” *Neuropsychology Review*, no. 13, pp. 113-143.

Prkachin, K.M. and Mercer, S.R. 1989. "Pain expression in patients with shoulder pathology: validity properties and relationship to sickness impact, *Brain,* vol. 39, no. 3, pp. 257-265.

Prokoski, F.J., and Iedel, R. 1999. "Infrared identification of faces and body parts," in A.K. Jain, R.M. Bolle, and S., Pankanti (Eds.), *Biometrics: Personal Identification in Networked Society*, Kulwer Academic Press, Boston.

Puri, C., Olson, L., Pavlidis, I., Levine J., and Starren, J. 2005. "StressCam: Non-contact measurement of users' emotional states through thermal imaging," *In the Proceedings of the CHI 2005*, April 2005, Portland, Oregon, pp. 1725-1728.

Rao, C. R. 1964. "The use and interpretation of principal component analysis in applied research," *Sankhya Series A*, vol.26, pp. 329-357.

Redford, P. 2000. "Conference Report: Social goals and emotions," *The Psychologist*, vol. 13, no. 6, pp. 290-291.

Reeves, B., and Nass, C. 1996. *The media equation: How People Treat Computers, Television and New Media Like Real People and Places*. London: Cambridge University Press.

Rencher, A.C. 1995. *Methods of multivariate analysis,* New York: Wiley and Sons.

Ring, E.F.J. 1998. "Progress in the measurement of human body temperature," *IEEE Engineering in Medicine and Biology*, vol. 17, no. 4, pp. 19-24.

Root, A.A., and Stephens, J.A. 2003. "Organization of the central control of muscles of facial expression in man," *The Journal of Physiology*, vol. 549, pp. 289-298.

Sarlo, M., Buodo, G., Poli, S., and Palomba, D. 2005. "Changes in EEG alpha power to different disgust elicitors: the specificity of mutilations," *Neuroscience Letters*, vol. 382, pp. 291-296.

Sayette, M.A., Cohn, F.C., Wertz, J.M., Perott, M.A., and Parrott, D.J. 2001. “A psychometric evaluation of the facial action coding system for assessing spontaneous expressions,” *Journal of Nonverbal Behavior*, vol. 25, no. 3, pp. 167-185.

Schwartz, G.E., Fair, P. L., Salt, P., Mandel, M.R., and Klerman, G.L. 1976. “Facial Muscle Patterning to Affective Imagery in Depressed and Nondepressed Subjects,” *Science*, vol. 192, no. 4238, pp. 489-491.

Schwarz, S., Hofmann, M.H., Gutzen, C., Schlax S., and Emde, Von der G. 2002. “VIEWER: a program for visualizing, recording, and analyzing animal behavior, *Computer Methods and Programs in Biomedicine*, no. 67, pp. 55-66.

Sebastiani, L., Simoni, A., Gemignani, A., Ghelarducci, B., and Santarcangelo, E.L. 2003. “Autonomic and EEG correlates of emotional imagery in subjects with different hypnotics susceptibility,” *Brain Research Bulletin*, vol. 60, pp. 151-160.

Setiono, R., and Liu, H. 1998. “Feature Extraction via Neural Networks,” in *Feature Extraction, Construction and Selection*, H. Liu and H. Motoda (Eds.), Londod: Kluwer Academic Publishers, pp. 191-204.

Sharma, S. 1996. *Applied Multivariate Techniques*. New York: Wiley.

Sinha, R. and Parson, O. 1996. “Multivariate response patterning of fear and anger,” *Cognition and Emotion*, vol. 10, no. 2, pp. 173-198.

Skin Anatomy and Physiology. 2005. *Online @ www.essentialdayspa.com.* Last visited on 25 April 2006.

Sloan, D.M., Bradley, M.M., Dimoulas E., and Lang, P.J. 2002. “Looking at facial expressions: Dysphoria and facial EMG,” *Journal of Biological Psychology*, vol. 60, pp. 79-90.

Smith, W.W. 1999. *The measurement of Emotion*, London: Routledge.

Socolinsky, D.A., Selinger, A., and Neuheisel, J.D. 2003. “Face recognition with visible and thermal infrared imagery,” *Computer Vision and Image Understanding*, vol. 91, pp. 72-114.

Solomon, R.C. 2003. *What is an emotion?*, New York: Oxford University Press.

Starr, C., and McMillan, B. 2003. *Human Biology*, California: Books Cole Thomson Learning.

Stemmler, G. 1989. “The autonomic differentiation of emotions revisited: convergent and discriminant validation.” *Psychophysiology*, vol. 26, no. 6, pp. 617-632.

Stern, R.M., Ray W.J., and Quigley, K.S. 2001. Muscles: *Psychophysiological Recording*, New York: Oxford University Press Inc.

Stroud, M.A. 1991. “Effects on energy expenditure of facial cooling,” *European Journal of Applied Physiology*, vol. 63, pp. 376-380.

Sugimoto, Y., Yoshitomi Y., and Tomita, S. 2000. “A method of detecting transitions of emotional states using a thermal facial image based on a synthesis of facial expressions,” *Robotics and Autonomous Systems*, no. 31, pp. 147-160.

Suhm, B., Myers, B., and Waibel, A. 2001. “Multi-modal error correction for speech user interface,” *ACM Transactions on Human-Computer Interaction*, vol.8, no. 1, pp. 60-98.

Sung, K.-K. and Poggio, T. 1996. “Example-based learning for view based human face detection”, *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 7, pp. 831-836.

Surakka, V., Illi M., and Isokoski, P. 2004. “Gazing and frowning as a new human-computer interaction technique,” *ACM Transaction on Applied Perception*, vol. 1, no. 1, pp. 40-56.

Swets, D., and Weng, J. 1998. “Using discriminant eigenfeatures for image retrieval”, *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 20, pp. 39-51.

Szabo, T., Horkay, F., Fazekas, L., Geller, L., Gyongy, T., Juhasz-Nagy, A. 2000. “Thermographic evaluation of myocardial protection,” *IEEE Engineering in Medicine and Biology*, vol. 19, no. 3, pp. 49-55.

Takeuchi A., and Nagao, K. 1993. “Communicative facial displays as a new conversational modality,” *In the Proceedings of ACM Intergraph 1993*, pp. 187-193.

Tao, H., and Huang, T. 1999. “Explanation based motion tracking using a piecewise Bezier volume deformation model, *In the proceedings of IEEE 1999 International Conference on Computer Vision and Pattern Recognition, IEEE CVPR 99,* June 1999, vol. 1, IEEE Computer Society, pp. 23-25.

Terzopoulos D., Lee, Y., and Vasilescu, M.A.O. 2004. “Model-based and image-based methods for facial image synthesis, analysis and recognition,” *In the Proceedings of the6th IEEE International Conference on Automatic Face and Gesture Recognition, FG2004*, Seoul, May 2004, vol. 2, IEEE Computer Society, pp. 3-8.

The Australian Psychological Society. 2003. Code of Ethics. Melbourne, Australia: The Australian Psychological Society Limited.

Tian Y.-L., Kanade K., and Cohn, J. 2001. “Recognizing action units for facial expression analysis,” *IEEE Transactions on Pattern Analysis, Machine Intelligence*, vol. 23, no. 2, pp. 97-115.

Tian, Y.-L., Kanade, T. and Cohn, J. 2002. “Evaluation of Gabor-Wavelet based facial action unit recognition in image sequence of increasing complexity,” *In the Proceedings of the 5th IEEE International Conference on Automatic Face and Expression Recognition (FG’02)*, 2002, Washington DC.

Toivanen, J., Våyrynen, E., and Seppånen, T. 2004. “Automatic discrimination of emotions in spoken Finnish,” *Language and Speech* vol. 47, no. 4, pp. 383-412.

Tomskin, S. 1984. “Affect theory,” in *Approaches to Emotions*, K.R. Scherer and P. Ekman (Eds.), New York: Academic Press, pp. 163-195.

Turk M., and Pentland, A. 1991. “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience,* vol. 3, pp. 71-86.

Turner J.R., and Thayer, J.F. 2001. *Introduction to Variance Analysis*. London: Sage Publications.

Ukai, K., Saida S., and Ishikawa, N. 2001. “Use of infrared TV cameras built into head-mounted displays to measure torsional eye movements,” *Japanese Journal of Ophthalmology,* no. 45, pp. 5-12.

Vrana, S.R. 1993. “The psychophysiology of disgust: differentiating negative emotional contexts with facial EMG,” *Psychophysiology,* vol. 30, no.3, pp. 279-286.

Vrana, S.R., and Gross, D. 2004. “Reactions to facial expressions: effects of social context and speech anxiety on response to neutral, anger and joy expressions,” *Biological Psychology*, vol. 66, no. 1, pp. 63-78.

Veldhuizen, I.J., Gaillard A.W., and De Vries, J. 2003. “The influence of mental fatigue on facial EMG activity during a simulated workday,” *Biological Psychology*, vol. 63, no. 1, pp. 59-78.

Vianna, D.M., and Carrive, P. 2005. “Changes in cutaneous and body temperature during and after conditioned fear to context in the rat,” *European Journal of Neuroscince,* vol 21, no. 9, pp. 2505-2512.

Ward, R., Bell, D. and Marsden, P. 2003. “An exploration of facial expression tracking in affective HCI,” *In the Proceedings of HCI2003*, University of Bath, England, P. Palanque, P. Johnson and E. O'Neill (Eds.), People and Computers XVII: Designing for Society, British Computer Society, England, vol. 1, pp. 383-399.

Ward, R.D., and Marsden, P.H. 2004. “Affective computing: problems, reactions and intentions”, *Interacting with Computers*, vol. 16, no. 4, pp. 707-713.

Webb, A.R. 2002. *Statistical Pattern Recognition*, London: John Wiley.

White, M. 1999. “Representation of facial expression of emotions,” *American Journal of Psychology*, vol. 112, no. 3, pp-371-381.

Whiteside, S. 1998. *In the Proceedings of ICSLP 1998:International Conference on Spoken Language Processing,* November-December, 1998, Sydney, Australia, pp.699-703.

Wild, B., Erb, M., and Bartels M. 2001. “Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences,” *Psychiatry Research,* vol. 102, pp. 109-124.

Wilfong, J.D. 2006. “Computer anxiety and anger: the impact of computer use, computer experience, and self-efficacy beliefs,” *Computers in Human Behavior*, vol. 22, no. 6, pp. 1001-1011.

Wilson J.D., and Buffa A. J. 1990. *College Physics.* New Jersey: Prentice Hall.

Wilson, A.D., and Bobick, A.F. 1999. “Parametric Hidden Markov Models for gesture recognition,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 884-900.

Winkielman, P., and Cacioppo, J.T. 2001. “Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect,” *Journal of Personality and Social Psychology*, vol. 81, no. 6, pp. 989-1000.

Wolf, L.B., Socolinsky, D.A., and Eveland, C.K. 2005. “Face recognition in thermal infrared,” in B. Bhanu and I. Pavlidis (Eds.), *Computer vision beyond visible spectrum*, Springer-Verlag, London.

Wolf, K., Raedler, T., Henke, K., Kiefer, F., Mass, R., Quante, M., and Wiedemann, K. 2005. “The face of pain- a pilot study to validate the measurement of facial pain expression with improved electromyogram method,” *Pain Research and Management*, vol. 10, no. 1, pp. 9-14.

Wright, P., He, G., Shapira, Goodman, W.K., and Liu, Y. 2004. "Disgust and the insula: fMRI responses to pictures of mutilation and contamination, " *NeuroReport*, vol. 15 pp. 2347–2351.

Wyse, N., Dubes, R., and Jain, A.K. 1980. "A critical evaluation of intrinsic dimensionality algorithms," in Gelsema, E.S. and Kanal, L.N. (Eds.) *Pattern Recognition in Practice*, London: Morgan Kaufmann Inc., pp. 415-425.

Xiao, R, Li, M.-J., and Zhang, H.-J. 2004. "Robust multipose face detection in images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 31-41.

Yacoob Y., and Davis, L. 1996. "Recognizing human facial expression from long image sequences using optical flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 636-642.

Yoshitomi, Y., Kim, S-I., Kawano T., and Kitazoe, T. 2000. "Effects of sensor fusion for recognition of emotional states using voice, face image and thermal image of face," *In the Proceedings of the IEEE International workshop on Robotics and Human Interactive Communication*, Osaka, Japan, September 2000, pp. 178-183.

Zaatri, A., and Oussalah, M. 2003. "Integration and design of multimodal interfaces for supervisory control systems," *Information Fusion*, vol. 4, pp. 135-150.

Zajonc, R.B. 1985. "Emotion and facial efference: a theory reclaimed," *Science*, vol. 228, pp. 15-21.

Zhai, J., Barreto, A.B., Chin, C., and Li, C. 2005. "User stress detection in human-computer interaction," *Journal of Biomedical Sciences Instrumentation*, vol. 41, pp. 277-282.