Toward Use of Facial Thermal Features in Dynamic Assessment of Affect and Arousal Level

Masood Mehmood Khan, *Member, IEEE*, Robert D. Ward, and Michael Ingleby

Abstract—Automated assessment of affect and arousal level can help psychologists and psychiatrists in clinical diagnoses; and may enable affect-aware robot-human interaction. This work identifies major difficulties in automating affect and arousal assessment and attempts to overcome some of them. We first analyze thermal infrared images and examine how changes in affect and/or arousal level would cause hæmodynamic variations, concentrated along certain facial muscles. These concentrations are used to measure affect/arousal induced facial thermal variations. In step-1 of a 2-step pattern recognition schema, 'between-affect' and 'between-arousal-level' variations are used to derive facial thermal features as Principal Components (PCs) of the facial thermal measurements. The most influential of these PCs are used to cluster the feature space for different affects and subsequently assign a set of thermal features to an affect cluster. In step-2, affect clusters are partitioned into high, medium and mild arousal levels. The distance between a test face vector and the centroids of sub-clusters at three arousal levels belonging to a single affective state, identified from step-1, is used to determine the arousal level of the identified affective state.

Index Terms—Affect-aware agents, affect classification, affect arousal assessment, human-robot interaction

1 Motivation

An affective state is regarded as a temporal, multi-component, transient and reactive response pattern to a particular stimulus, event or situation. An affect is therefore viewed as an outcome that is dependent on an external stimulus or situation [1], [2]. Visual-sign-based facial expression recognition systems such as FACS (Facial Action Coding System) [3] and MAX (maximally discriminative facial movement coding system) [29] are entrenched in this view. Studies suggest that affect elicitation stimuli produce changes in heart rate; blood pressure; pupil dilation; flow of blood to body muscles; respiration; and release of epinephrine and norepinephrine from the adrenal medulla through sympathetic activation [4], [5], [6]. Concurring with the relevant literature, automated affect and arousal assessment systems posit that measurements of cognitive, biophysiological, neurological, facial or bodily response signals would allow assessing affect and associated arousal level [7].

Though significant progress has been made toward building automated facial expression and affective-state recognition systems, little progress is made toward automating the dynamic assessment of affect-arousal. Recently, efforts were made to use visual, psychophysiological and neural signals for dynamic affect and arousal assessment. These efforts aim to facilitate affect-aware psychiatric diagnosis and intervention; psychological assessment; forensic investigation; and human-robot interaction [8, 9, 10, 88]. Affective science literature identifies problems hindering the development of dynamic affect arousal assessment tools. Some of these problems, discussed in Section 2, are addressed in this work.

- M.M. Khan is with the Mechanical Engineering Department, Faculty of Science and Engineering, Curtin University, Bentley Campus, Perth, Western Australia, Australia. E-mail: masood.khan@curtin.edu.au.
- R.D. Ward was with the School of Human and Health Sciences and School of Computing and Engineering, University of Huddersfield, Huddersfield, West Yorkshire, U.K. E-mail: r.d.ward@btinternet.com.
- M. Ingleby is with the University of Huddersfield, Huddersfield, West Yorkshire, U.K. and the King Abdulaziz University, Saudi Arabia. E-mail: inglebymeltham@btinternet.com.

Manuscript received 23 Aug. 2015; revised 15 Jan. 2016; accepted 7 Feb. 2016. Date of publication 25 Feb. 2016; date of current version 12 Sept. 2017.
Recommended for acceptance by P. Morency.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TAFFC.2016.2535291

2 Problems in Affect-Arousal Assessment

1. For reliable assessment of affect/associated affect-arousal, the subject shouldn't get an opportunity to simulate/ dissimulate the affective experience [11], [12], [13], [14]. This is difficult for participants of "designed" experiments. When people are aware that an affect/ arousal level is being elicited or assessed, they naturally tend to simulation or dissimulation: so there are 'experimenter effects'. There are also noise effects: inter and intra personal differences in humans' (affective) response patterns are sensitive to experimental conditions such as laboratory set up and the fact that emotions cannot be produced *de novo* but show hysteresis depending on past experiences of similar stimuli and the time evolution of present experience [18].
2. Stimuli need to be compliant with the strict ethical frameworks enforced by the relevant professional and governmental organizations [15], [16], [17]. Such compliance related constraints may prohibit eliciting the desired directions and intensities of affective experiences and may limit assessment of affect-arousal level.
3. Connection between a stimulus and human temporal factors like affect latency, rise time, affect magnitude and affect duration is difficult to model. Incorporating

such interactions would add to the design-complexities of affect-arousal level assessment systems [18].

4. Separating intrinsic and extrinsic elements of affective state assessment is not easy. Distinguishing intrinsic from extrinsic factors requires multi-dimensional analysis of human factors and stimulus characteristics. Particularly, unexpected variations and temporal transformations caused by factors such as pose change, intrapersonal movement of facial muscles, shade and illumination conditions are all difficult to model as statistical noise. Hence, intrinsic and extrinsic elements aren't usually separated while designing the affect/ arousal assessment tools.

Implementing an affect-arousal assessment tool would require eliciting pre-defined levels of affective experiences. One or more stimuli such as: hypnosis, story-telling, imagined scenes and situations, paintings, illustrations, images (surreal), photographs (real), video clips, music clips, medicine, drugs, odors, sound effects, narratives, self-statements, facial expressions, gestures, body movements, head movements, and various forms of social interaction must be used for affect elicitation [18], [19]. Affective science literature suggests that it is difficult to assess how and when a stimulus can produce the desired intensity of affective experience [18], [20], [21], [22], [23], [24], [25], [26].

Many attempts have been made for establishing a set of measurable, universally applicable, repeatable and useful parameters for dynamically assessing affect and arousal level. We exploit the fact that facial muscle movements and their associated biophysiological cues can help in determining affect and arousal level and attempt to use ethically sound, remotely acquired, non-visual (thermal) data for the task. The aim is to overcome the affect and arousal level assessment related problems, reported above.

3 Automated Assessment of Affect/ Arousal Level

Since the publication of Charles Darwin's seminal work [85], facial muscle movement patterns are considered corollary to common affective states. Systems of coding facial motion and deformations into affective states are discussed in [3], [29], [32], [33]. Of these, FACS, a sign-based system for coding facial motion and deformation in affect classes [3] in widely used. Despite the popularity of facial coding systems, their use in arousal level assessment is not common. Rather, retrospectively self-reported descriptions of affective experiences are frequently used in assessing affect arousal levels [27].

Use of cognitive somatic, visceral and bodily response signals in affect and affect arousal level assessment is also evident now. Contact-based physiological measurements like electrodermal or psychogalvanic activity, cardiovascular activity and brain signals are also used for the purpose. Recent works also attempted to relate biophysiological signals to with emotive experiences [28], [30], [31], [34], [35], [36].

Facial musculo-electrical measurements were employed for real time assessment of emotional experiences [89]. Higher correlations ($r > 0.9$) between participants' self-reported emotive ratings and machine- measured estimates of musculo-electrical activities were observed [89]. In [31], brain signals were employed for affect and arousal level assessment. Relation between cardiovascular response and facial skin temperature measurements was studied to suggest a relationship between respiratory responses and variation in facial skin temperature [38].

A growing interest in using non-contact facial skin temperature measurements for affect and arousal assessment is evident in the literature. In [39], stress-induced neurophysiological responses, measured on the perinasal area of the face (manifested as a momentary perspiration pattern), were used for remotely quantifying human stress levels. In [41], stress-caused blood volume flow variations were measured along the orbital muscles to detect deceit. The binary-classifier used facial thermal features to distinguish non-deceptive from deceptive subjects [41]. In [42], young female subjects, presented with stressful incentives, experienced significant temperature changes on hand and facial skin. Skin temperature variations have also been used for identifying levels of stress, pain, anxiety and affect in [17], [37], [38], [39], [43], [44], [45], [46], [47], [48], [49]. In [44] a system was developed for non-contact measurement of computer users' affective state using facial thermal variations. The employed algorithmic approach was similar to the one used in [43]. In [46] infrared imaging of facial thermal features was used for detecting transitions of emotional states. In [47] a combination of visual images, thermal features and audio signals were used to classify affective states. A set of optimal principal components (PCs) (derived from the thermal variation measurements observed along the facial muscles) was used to classify neutral, happy, sad, angry, pleasantly surprised and horrified facial expressions [48], [49], [50], [51], [52]. In [54] affect arousal level was evaluated using facial skin temperature variations around the nasal region. The relationships between the arousal levels and the observed facial skin temperatures were used to establish an index of sympathetic nervous system activity [54].

In another investigation, thermal patterns concomitant to specific facial action units, proposed in [60], were investigated. A spatial pattern approach (based on PC analytic decomposition of the thermal signals) was employed. The thermal fluctuations were found to be specific to the activated AUs and were sensitive to the kinetics and intensities of AU production [87].

In [97], thermal imaging based facial expression classifier was found to be more effective than visual images based classifier. The classifier could also overcome illumination and skin complexions related problems [97]. Temporal, spatial and spatio-temporal response factors were extracted from the facial thermal infrared images in [55] for binary classification of six affective states and self-reported affect arousal levels. The authors used images from International Affective Pictures System (IAPS) database [56] as visual stimuli for eliciting affective states and affect-arousal levels. Using the interpersonal (known as circumplex) model in which conclusions about the affective states depend on the degree of arousal and valance, a pattern classification scheme was implemented [55]. The training and classification features were selected using a genetic algorithm, and the selected features were classified through linear discriminants (LDs) of the feature space [55].

It's beyond the scope of this paper to cite and review all relevant works but the cited literature suggests that visual, biophysiological and neural signals can be used for affect and arousal level appraisal.

Following the previous works we propose a novel facial-thermal-feature-supported classification approach for dynamic of assessment affect and arousal level. Section 4 describes our experimental design including equipment details, elicitation procedure, and imaging protocols. The employed computational approach is presented in Section 5. In section 6, we demonstrate how facial thermal features change under the influence of affective state and arousal level. Section 7 reports the classification results. Section 8 concludes this work, highlights its significance and proposes directions for future work.

4 Experimental Design

4.1 Participants

Thermal infrared images showing participants' evoked facial expressions of affective states were acquired under controlled conditions. Thermal images of twelve male and seven female students (mean self-reported age of 20 years 3 months) were recorded and analyzed. Participants included Africans, Caucasians, Arabs, Iranians, Indians and Pakistanis.

4.2 Eliciting Involuntary Affects

A set of carefully selected stimuli (containing still images and video clips) was used for evoking expressions of happiness, and sadness in this investigation. In order to maintain the authenticity of the stimuli, their contents were aligned with the contents of the pre-categorized IAPS images [56]. The selected stimuli were taken from reputed publishers like BBC, MSNBC, and CNN. Any violent, unethical or extremely disturbing images were avoided. The image and video clip contents were no more extreme or excessive than those of the productions usually shown on the mainstream television. The employed images and video clips were capable of producing low to high affect intensities.

4.3 Equipment and Experimental Design

In order to develop a database of visible-spectrum and thermal facial expression images a commercial quality high-resolution video camera and an uncooled-microbolometer (FPA detector- mounted) thermal infrared camera were used [45], [46], [47], [48], [49], [50]. The employed thermal infrared camera has a high thermal sensitivity (0.07 °C at 30 °C) and has thermal accuracy of $\pm$ 2 °C in the infrared wavelength range of 7.50-14.00 μm. The camera had a pixel resolution of 320 $\times$ 240. A low emissivity ($\varepsilon = 0.54$) concrete wall background was used to ensure emissivity contrast, better image segmentation and effective separation of background from the desired regions of the thermal images [57], [58].

Through air-conditioning, the internal room temperature was maintained between 19-22 °C during the image acquisition sessions. All participants were given at least 20 minutes to acclimatize with the environment and get their respective body temperatures steadily set.

In order to ensure high quality of visual-spectrum and thermal infrared images, the image acquisition process was monitored and the acquired images were carefully analyzed.

Participants' visible and corresponding thermal infrared images were simultaneously recorded. Images with neutral faces were recorded before each participant was made to experience an affective state. Images were assessed for quality and, when needed, imaging was repeated for developing a set of good quality images (clearly exhibiting the desired affective state and arousal level). Fig. 1 shows how facial skin temperature varies with a change in affective state and affect arousal level. A participant's neutral face (top), medium happy face (middle) and very happy face (bottom) are shown in Fig. 1. It is obvious that facial skin temperatures, as measured along *orbicularis oculi* (eyes), change with a change in affective state and arousal level.

4.4 Affect Arousal Assessment

Images in our thermal image database were analyzed for estimating the arousal levels, evident in them. The affect arousal levels were determined on the basis of observed variations in visual signs [18], [27]. For this investigation, visual-signs-supported criteria were developed and used to assess arousal levels (as evident) in participants' visual images (photographs). The rules, norms and guidelines of 'affect perception' from facial expression images were derived from the wider literature on affective science [3], [18], [26], [27], [59], [60], [61], [62], [63]. In particular, the '*crude-review*' approach [27] was followed to develop the assessment criteria. Three trained and experienced human assessors were asked to analyze visual-signs through judgment of facial features' appearance, interplay of the facial features, head movement, body movement and the interplay of head and body movements in participants' photographs. The developed criteria is reported in Table 1. Using this criteria participants' affective experiences were categorized into: high arousal level, medium arousal level and mild arousal level. In order to evaluate intensity of affect arousal as evident in the images, first the visible-spectrum images were examined and ranked. From these ranked images, the ones that provided the most reliable representation of affective state were used to determine the level of affect arousal. The corresponding thermal infrared images of high-ranking visible images were used to analyze how variations in affective states would correspond to the facial thermal variations along the facial muscles. This ensured selecting the most realistic images of facial expressions.

Our assessment of arousal level is rooted in Russell's circumplex model [64]. Using a modified version of Russell's model, shown in Fig. 2, the arousal level in an image could be classified as very obvious (VO), Medium obvious (Med-O) and Mildly obvious (Mild-O) in the backdrop of a low to high presence of valance (LV to HV). The medium obvious (Med-O) case served as the baseline in expression classification. Such modifications in Russell's model have also been employed in [55], [65]. Table 1 shows how the developed affect assessment criteria would help in concluding high level of happiness arousal in a participant.

Through application of the assessment criteria, thermal variations pertaining to arousal levels of happy and sad affective states were earmarked. This resulted in selection of the facial thermal images containing: neutral; evoked very happy; evoked medium-happy; evoked mildly happy; evoked very sad; evoked medium-sad and; evoked mildly sad expressions, used for analyses in this work. The employed assessment criteria ensured approaching the facial expression

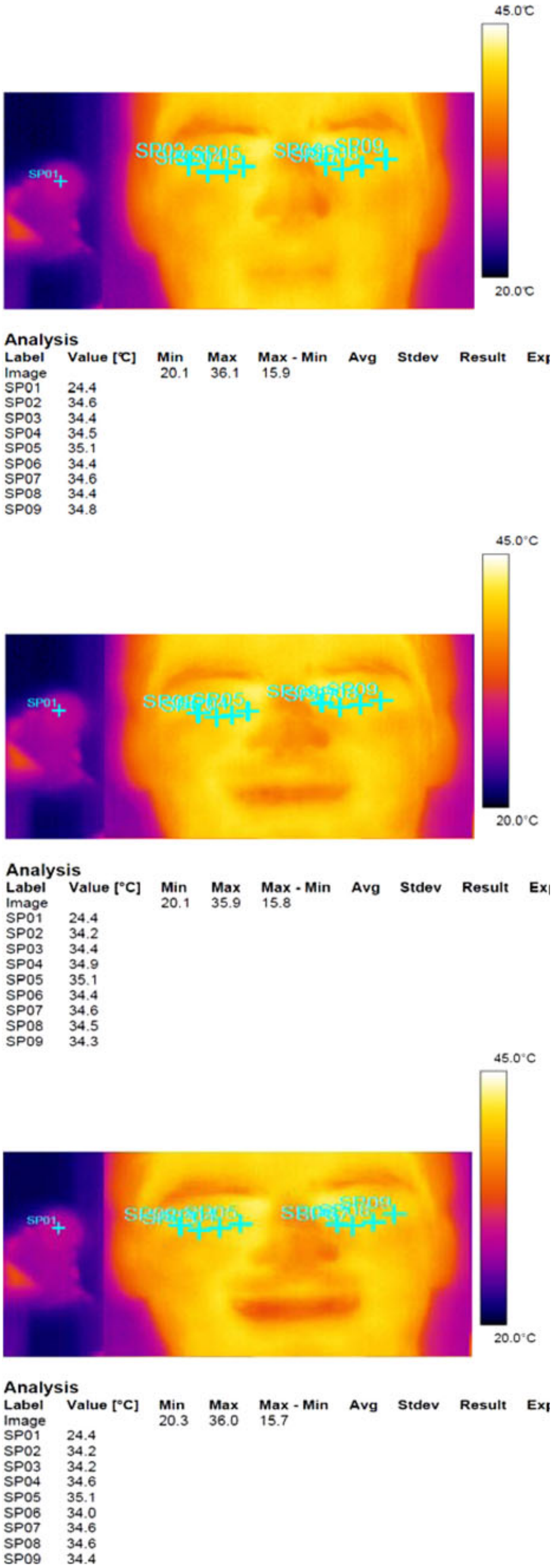

Fig. 1. Variations in facial skin temperatures measured along the eyes (Orbicularis Oculi). Top—Neutral, Middle–medium-happy, Bottom—very happy. Point SP01 in these images measures the temperature of a cavity-type blackbody surface. These images were recorded using a FLIR Thermovision—A series Infrared camera at Curtin University's Bentley campus.

TABLE 1
Criteria for Analyzing Facial Expression of Affects in Visible Spectrum Images (Level of Happiness Evaluated)

Factor Description	Observed Scale/ Level of Evidence[a]				
	0	*1*	*2*	*3*	*4*
Orbicularis oculi in movement					X
Eyes widened				X	
Eyebrows stretch		X			
Eyebrows down	X				
Eyebrows lifted			X		
Orbicularis oris in movement			X		
Massetter area stretched			X		
Mouth opening				X	
Mouth closing	X				
Lips tensing				X	
Lips relaxing	x				
Zygomaticus major inward movement	X				
Zygomaticus major outward movement					X
Compressor naris deformation		X			
Forehead skin stretching/ tensing		X			
Forward head movement	X				
Backward head movement				X	
Sideway head movement		X			
Body movement		X			
Intensity of expressed affect ~~neutral~~ / **happy**/~~sad~~ (strike out what is not applicable)					X
Expressed affect visible quality					X

[a] *0 = No evidence; 1 = Low evidence; 2 = Visible evidence; 3 = High evidence; 4 = Strong Evidence.*

recognition in a systematic way and avoiding any self-reported assessments of affect and arousal level.

5 Computational Methodology

Building upon the previous works, we employed a self-justifying approach to recognition and avoided binary-classification method. We opt to partition a multi-affect and multi-arousal level discriminant space is into affect clusters. The proximity of a test face (a test-image feature-vector) to a cluster centroid is used to allocate the test face to a particular affect-cluster. The affect arousal assessment is carried out in a smaller mixture space constructed with the available distributions of participants' thermal faces representing only the recognized affective state. The distance between a

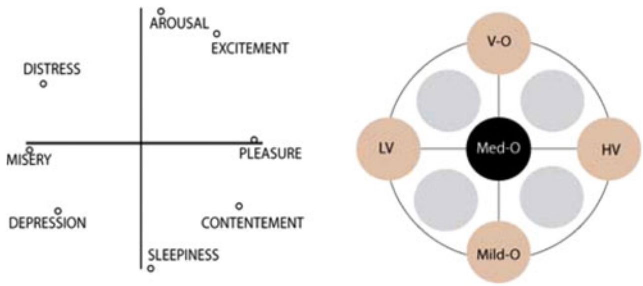

Fig. 2. Left—A basic representation of Russell's eight-affect concept model [64]. Right— Modified version of Russell's circumplex model with only three arousal levels; very obvious (V-O), medium obvious (Med-O) and mild obvious (Mild-O).

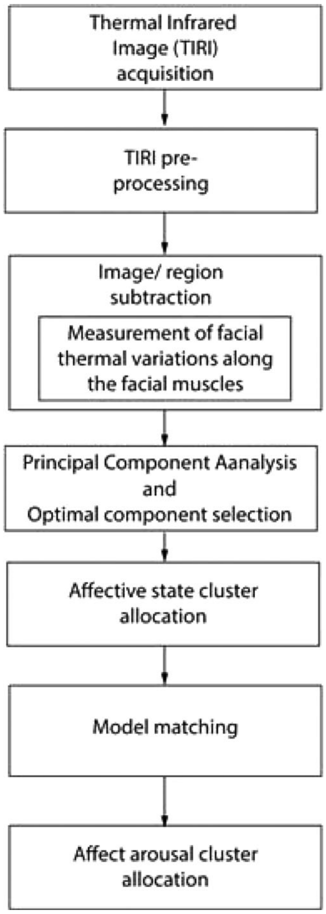

Fig. 3. The employed affect recognition and arousal level assessment algorithm.

test face and a pre-defined level of affect arousal is used to label the test face as showing high, medium or mild level of (recognized affect) arousal. Thus, in our proposed schema shown in Fig. 3, the affective state is identified before assessing the arousal level.

Our employed computational method is based on several (affect and expression) classification algorithms used in previous works. For example, in [98], Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were applied in two steps. First the face image (original vector space) was projected to a face subspace using PCA. Then, LDA was used to implement a linear classifier. The PCA and LDA combination improved the generalization ability of classifier trained only on a few samples of each class. In several other relevant works, stepwise classification is used when limited yet high-dimensional temporal data are available. The feature space dimensions are reduced first to discover most influential directions of data variance. A best set of discriminating PCs is discovered then. An appropriate criterion function such as minimum error rate or maximum class separation is employed in the second step. Use of a threshold value (for maximizing the inertia) for selecting the classifier-training features is avoided. In the final step, the best discriminating features are projected in a compact yet optimal feature space [66], [75], [90], [91], [92], [93], [94], and [95]. A stepwise classification approach results in a robust, model-based classifier for temporal, high-dimension data [75]. We detail the employed algorithmic approach below.

Fig. 4. Left to right: The highest thermal intensity values were measured within the shown 16, 32 and 64 square segments on the face.

5.1 Step-1: Recognition of Affective States

Transient thermal features can be detected from within a series of thermal images sampled at an appropriate rate [58]. Time-series thermal images are first registered, then, the temperature variations in the regions of interest are estimated by subtracting the registered images [43], [54], [57], [58]. The observed temperature variations in the regions of interest represent temporal changes in the facial thermal distribution. However, such thermal variation detection approaches cannot guarantee statistical aptness of the data for invoking multivariate analyses and relevant statistical. It is necessary to avoid a high correlation between the facial thermal measurements and discover the maximum between-facial-expression variance. We needed an appropriate feature extraction approach for discovering thermal variance in the image data.

For image analysis, the facial locations going through the transient thermal changes were the regions of interest (ROI). The ROI were expected to appear along the major facial muscles. The ROI were identified for discovering the temporal thermal variations and testing the measured data for statistical aptness. For discovering the ROI, thermal infrared images were first divided in 16 square segments of 36×36 pixels (along the facial muscles). The maximum Thermal Intensity Values (TIVs) in each of the 16 square regions of interest were recorded and tested for (possessing) $\mathrm{Correlation}_{\mathrm{min}}$ and $\mathrm{Variance}_{\mathrm{max}}$.

After this normative study, thermal images were repeatedly divided into increasing numbers of square segments along the major facial muscles. We progressively analysed 16, 32 and then 64 square segments of 25×25 pixels. Fig. 4 shows locations of 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 thermal infrared images showed significant differences in the thermal intensity values compared with the others.

The process was repeated until all locations experiencing significant thermal variations (and possessing the required statistical characteristics) were discovered. Through much iteration, we eventually noticed that 75 physical sites, along the major facial muscles, would experience significant thermal variations with a change in affect/ arousal level. We refer to them as Facial Thermal Feature Points (FTFPs). The TIV data gathered from these 75 FTFPs, each a square segment of 16×16 pixels, also met the conditions of $\mathrm{Correlation}_{\mathrm{min}}$ and $\mathrm{Variance}_{\mathrm{max}}$ [48], [49], [50], [53]. Fig. 5 shows these 75 FTFPs on a neutral human face, exhibits a muscular map of a human face, and portrays the geometric profile of the FTFPs. The 75 sites shown in Fig. 5 showed near consistent and significant variation in the TIV's

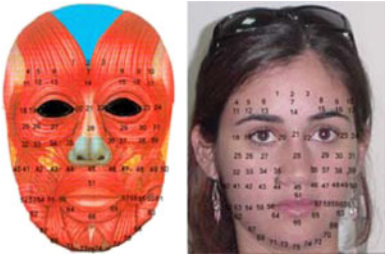

Fig. 5. Left to right— FTFPs on a facial muscle map and FTFPs on a human face.

measured in male and female participants' thermal images. These temperature measurements were used to discover affect-induced changes in facial skin temperature. One participant's visual and thermal faces with various facial expressions are shown in Fig. 6. PCA was invoked on the facial thermal variation data for dimension reduction [48], [49], [68], [69], [70].

We regarded each facial thermal image as a p-dimensional random facial thermal vector $\mathbf{x}$. The acquired image database had n thermal image vectors; $\mathbf{x_i}, (i = 1, 2, \ldots, n)$. Each $\mathbf{x}$ had p components (coming from the TIV measurements) as $\mathbf{x_i} = [x_{i1}, x_{i2} \ldots, x_{ip}]^T$. The TIV data were standardized and a learning set, drawn as $G_0 = [\mathbf{x}_1|\mathbf{x}_2|\ldots|\mathbf{x}_n]$, contained n number of p-dimensional facial thermal vectors. Using a conventional approach, the mean facial thermal vector $\bar{\mathbf{x}}$ of the learning set was obtained as:

$$\bar{\mathbf{x}} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{n=1}^{n} \mathbf{x_i} \tag{1}$$

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

$$\tilde{\mathbf{x}}_\mathbf{i} = \mathbf{x_i} - \bar{\mathbf{x}}. \tag{2}$$

This way of offsetting would shape the learning set as a $p \times n$ matrix $G = [\tilde{\mathbf{x}}_1|\tilde{\mathbf{x}}_2|\ldots|\tilde{\mathbf{x}}_n]$. The $p \times p$ sized covariance matrix C of the learning set was obtained as

$$C = GG^T. \tag{3}$$

Being symmetric and positive-definite, the covariance matrix, was reduced to the form

$$C = H_l D H_l^T. \tag{4}$$

Here, H_l is the linear transformation matrix. It is an orthogonal non-zero eigenvector matrix of C made up of columns of eigenvectors as

$$H_l = [\upsilon_1|\upsilon_2|\ldots|\upsilon_p]. \tag{5}$$

The matrix D is the diagonal eigenvalue matrix of H_l of the form,

$$D = diag\ [\lambda_1, \lambda_2, \ldots, \lambda_p]. \tag{6}$$

The eigenvalues are arranged in a descending order as $[\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p]$. The eigenvectors $\upsilon_i[i = 1, 2, \ldots, p]$, also referred to as eigenfaces in the literature, contain all important features required for pattern classification purposes. The ordering of eigenvectors of D shows the direction of the largest variance in the data. By rating the PCs (eigenvectors) for affect discrimination power and removing the less discriminating eigenvalues (λ_s) from D, together with the corresponding columns from the transformation matrix H_l, suitable data reduction is achieved. This reduction results in a smaller thermal feature space that is spanned by only M eigenfaces ($M \ll p$). The reduction process used for optimal thermal feature selection is outlined later.

Fig. 6. Top row: Visible and thermal images of neutral expression. Second row: Visible and thermal images of highly happy expression. Third row: Visible and thermal images of medium-happy expression. Fourth row: Visible and thermal images of mildly happy expression. Facial temperature variations are evident in the above images. These images were edited for publication [50].

The learning set G is pre-classified so it is 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 as,

$$G = G_1 \cup G_2 \cup \ldots \cup G_g. \tag{7}$$

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

$$\tilde{x}_{ijk} = \mu_i + \tau_{ij} + \varepsilon_{ijk}; (1 \leq \mathrm{j} \leq \mathrm{p}, 1 \leq \mathrm{j} \leq \mathrm{g}, 1 \leq \mathrm{k} \leq \mathrm{nj}). \tag{8}$$

In Equation (8), $\tilde{x}_{ijk}$ is the ith observation of a face expressing affective state j and is the kth such face with this expression, $\boldsymbol{\mu}_i$ is the mean value of all observations at point i, $\boldsymbol{\tau}_{ij}$ represents the offset of the center of the jth cluster from $\boldsymbol{\mu}_i$ and $\boldsymbol{\varepsilon}_{ijk}$ is a residual that is minimized while estimating the other model parameters from the data set.

5.1.1 Optimal Thermal Feature Selection

In order to use a set of optimal thermal features, a subset of M best discriminating (most influential) PCs from within

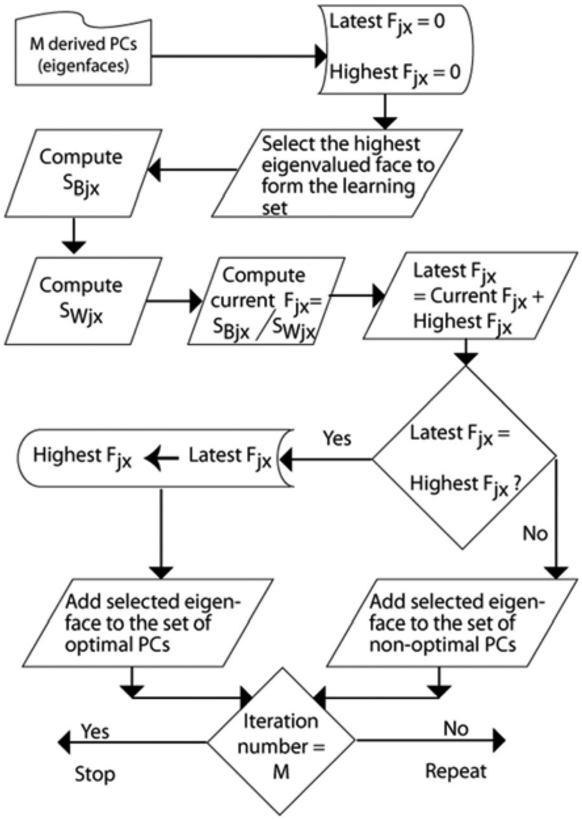

Fig. 7. The iterative algorithm used for optimal feature selection.

the set $\{\zeta^{(l)}|l = 1, \ldots, p\}$ of the PCs is discovered iteratively. The method is based on the stepwise test process in which less effective PCs are eliminated, one after another. Only the best discriminating PCs are retained in an optimized subspace. The process begins with ranking all PCs by the size of a Fisher ratio $F = \frac{|S_B|}{|S_W|}$ for the distribution of components x_{ijk}. The training sample observation vectors are projected along the $\boldsymbol{\zeta}^{(l)}$ direction. The distributions of projected components had variances $S^{(l)}$, which were sums of a within-cluster part S^l_W, and a between-cluster part S^l_B. The Fisher ratio for $\boldsymbol{\zeta}^{(l)}$ is given as $F^{(l)} = S^l_B/S^l_W$ and is greater for the PCs that better discriminate between the clusters. A similar ratio could be defined for subspaces E spanned by several PCs, except that the variances $S^{(E)}$, $S^{(E)}_W$, and $S^{(E)}_B$ were generalized to square matrices. For such an E, the Fisher ratio, $F^{(E)}$, can be defined as the ratio of determinants S^l_B, and S^l_W of these variance matrices. Thus, the optimized subspace that our procedure seeks out can be expressed as

$$M = \arg\max_E \{F^{(E)}\}. \tag{9}$$

The space E in equation (9) varies over all possible spans of two, three or more PCs, and M is the argument of function $F^{(E)}$ that corresponds to its the maximum value. Since first 23 eigenvectors had 99.7 percent of variation in the data, it was appropriate to select a corresponding value of M. Finding influential PCs using Fisher's criterion is an established and well-tested approach that was used in several similar studies [73], [74], [75], [84]. The influential component selection algorithm is described in Fig. 7.

5.1.2 Classification for Affect Recognition

The use of LDA on the PCs of training data has been successful in many complex classification scenarios [68], [69], [98]. A linear discriminant is a hyperplane that optimally separates a cluster of the training sample from the rest of the sample [66]. With J affect clusters, the resulting J hyperplanes partition the observation space into 2^J regions bounded by hyperplanes, of which J contain only one cluster center, ${}^JC_2$ contain 2 centers, ${}^JC_3$ contain 3, and so on. A new thermal image vector can be 'classified' by assigning it to one region, but if the region contains several centers the classification is ambiguous.

To resolve ambiguities, we associated each center with the 'inter' and 'intra' personal facial thermal variances about it. These are caused by the combined effect of an affective state and arousal level. Then, we calculate the within-cluster variance matrix and apply a well-known similarity criterion to discover which center was most similar to the new thermal image vector of various similarity measures. We use the distance, defined in terms of the pooled within-cluster variance matrix W of a training sample. If the x-x_j is the vector joining a new image vector to the center of cluster j, then the Euclidean length $||(x - x_j)W^{-1}||$ defines the distance from x to cluster j. Hence, the nearest cluster to x in a region (that might contain several cluster centers) is determined by estimating $\arg\min_j \{\|(x - x_i)W^{-1}\|\}$. This process of developing discriminant rules in a classification problem has a long history [67], [73]. We implemented it as a critical foil to our optimized subspace method, operating in a subspace spanned by the optimal PCs of training samples that contribute more than 1 percent to total variance measured by the trace of D and used it as a check on our construction of an optimized subspace [48], [49].

5.2 Step-2: Affect-Arousal Level Assessment

Classifying affect arousal is a more challenging task than affect classification for separating intrinsic from extrinsic features is difficult and can be computationally expensive. However, as evident from the reported results, a reduced feature space (to the three arousal levels of an affective state) would help in arousal level assessment.

We first allocate an affective state to a test face in step-1. The optimal thermal feature vector of a test face is then taken to a smaller, single-affect, three-arousal-level feature space for assessing arousal level. This smaller feature space is constructed by the optimal facial thermal feature vectors derived from the thermal images of high, medium and mild expressions of the affective state, identified in step-1. The smaller feature space effectively incorporates within-affect arousal-class and between-affect arousal-class facial thermal variations. Hence, step-2 of our pattern recognition schema can handle both intrinsic and extrinsic aspects of affect-induced facial thermal variations.

Step-2 of our schema begins by dividing the identified single-affect space into α-clusters of affect arousal levels $\alpha =$ [1(high), 2 (medium), 3 (mild)]. Each α-cluster centroid m_α is computed. An initial arousal cluster is then randomly assigned to the test optimal feature vector (ζ_t), whose arousal level is to be assessed. The optimal feature vector (ζ_t) to be tested, is projected to the next arousal level cluster space spanned by all of the optimal features such that

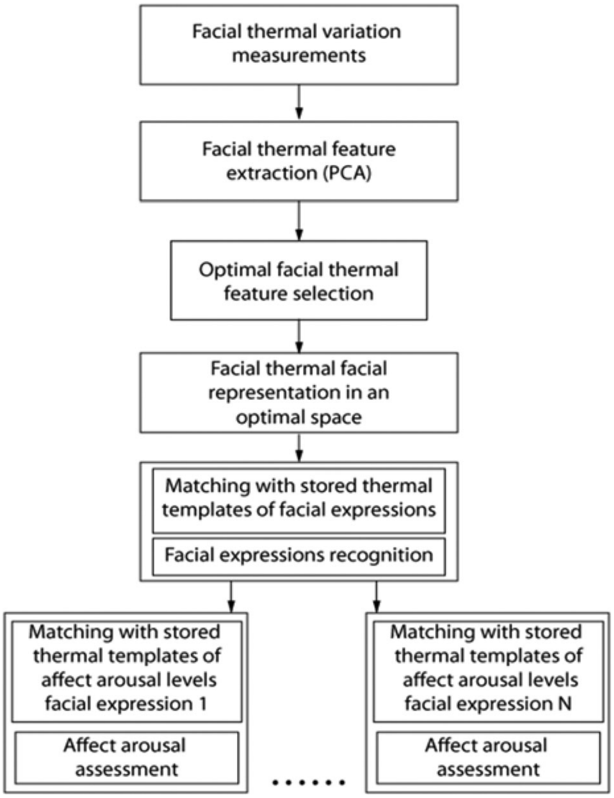

Fig. 8. The employed classification approach.

$W^n = \lfloor \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \rfloor$. The arousal assessment error, ε' is then calculated as:

$$(\varepsilon')^2 = \left[(\zeta - m_\alpha) - W^n (W^n)^T (\zeta - m_\alpha) \right]^2. \tag{10}$$

Using the minimum arousal assessment error (ε'), the test vector (ζ_t) is assigned to a new cluster. The assignment error minimizing process is terminated when no smaller ε' is found. Hence, the distance between a test optimal feature vector and the centroids of arousal-level clusters determines the most plausible arousal-level cluster to which a test face vector belongs [76], [77], [78].

This approach guarantees inclusion of all 'between and within class thermal variations belonging to a particular affective state' during arousal level assessment. Our smaller mixture model comprises of a single affect related multiple subspaces (of arousal levels). The mixture therefore includes maximum possible facial thermal variations observed during the expression of different arousal levels of a single affect. Also, instead of using some statistical distance, we use the distance between the optimal feature vectors and the centroids of arousal levels for affect assessment. Hence, we effectively incorporate all facial thermal variations measured on FTFPs that contribute to the facial thermal expression of an affect and arousal level. Our results confirm this [76], [77], [78]. The employed computational approach is illustrated in Fig. 8.

5.3 Computational Method's Efficacy Evaluation

In order to test the efficacy of the proposed classification method, we invoked this algorithm on a publicly available IRIS Thermal/Visible Face Database (IEEE OTCBVS Benchmark). The IRIS database [96] contains high quality, 8-bit grey-scale JPEG thermal infrared images. The IEEE

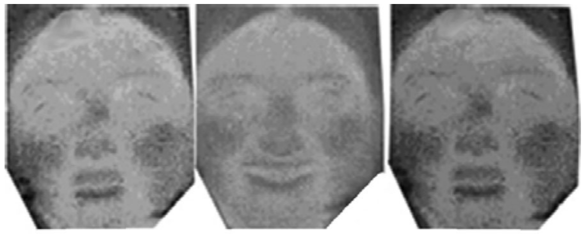

Fig. 9. Left: Destination image with neutral facial expression. Middle: image with very happy facial expression. Right: the registered image. The images were edited for publication.

OTCBVS images do not contain any thermal data but provide pixel grey-level information. We invoked our algorithm on the IEEE OTCBVS images hoping that affect-induced facial skin temperature variations would cause pixel grey-level changes and would lead to cluster-analytic classification of the affective states.

We also compared our IEEE OTCBVS images classification results with the classification results reported in another well-cited work [99]. The results and comparisons are reported in Section 8.

Since the proposed algorithm was designed to classify thermal information carrying frontal facial images, we were unable to test its performance in classifying the benchmarked visible spectrum images like those included in [100].

6 Facial Hæmodynamic Thermal Response to Affective States and Arousal Levels

We examined how Facial Hæmodynamic and Thermal Features (FHTFs) change with a change in either affective state or affect arousal level. Figs. 1 and 6 have earlier shown how facial skin temperature varies with expression of affects and with associated arousal levels. Fig. 9 shows registration of an image with very happy facial expression to the corresponding neutral image. Fig. 10 shows a difference image obtained by subtracting the images of neutral and very happy expressions (registered in Fig. 9).

In our database of male and female participants, the very happy facial expression would cause noticeable thermal variations along *frontalis*, orbicularis oculi, compressor naris, *levator labii superioris alaque nasi*, depressor *anguli oris* and *orbicularis oris*. These variations were found to be consistent in the acquired images. Noticeable (but smaller in magnitude) thermal variations were also observed along *zygomaticus* major, a well-established muscle of joy and happy expression. Probably, muscular movements and associated variations in blood volume flow caused the observed thermal variations along *frontalis* [48], [49], [53].

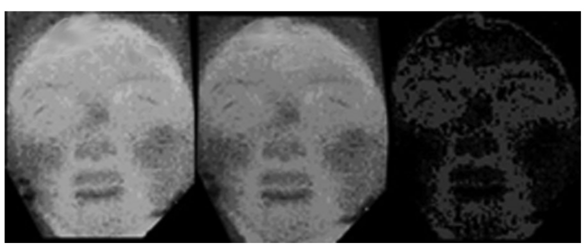

Fig. 10. From left to right: Neutral expression, registered image; and the difference image obtained by subtracting the two images. The images were edited for publication.

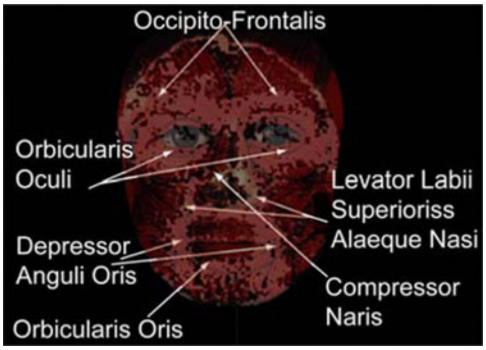

Fig. 11. A facial muscle image superimposed on a difference image obtained by subtracting images of neutral and very happy facial expressions. Facial thermal variations along muscles are made obvious by editing the picture.

In the acquired images, some thermal variations were also observed along *mentalis* or 'pouting muscle' during the expression of very happy state. However, in our database, variations along *mentalis* were inconsistent in appearance, location and magnitude. Probably, participants' facial anatomy and physiognomy contributed to these inconsistencies.

Fig. 11 shows a difference image obtained by subtracting the image of medium happy expression from the neutral expression image. Locations of the facial muscles, engendering significant thermal variations, are shown in Fig. 11. In the difference image of the same individual obtained by subtracting the image of mildly happy expression from the neutral expression image, neither *frontalis* nor orbicularis oculi experienced any significant thermal variations.

Through comparing participants' images with neutral and high, medium and mild facial expressions of affects, a reasonably consistent thermal response pattern was deduced from the observed TIVs.

6.1 Quantitative Measurement of Facial Thermal Variations

A number of statistical tests were invoked on measurements of thermal intensity values recorded along the major facial muscles. The acquired data had normal (Gaussian) distribution. Homogeneity of variance and sphericity were also observed in the data. The mean TIVs for three facial muscles; *zygomaticus major*, *risorious* and *orbicularis oculi*, were selected for preliminary analysis and comparison across the expressions of affective states. Figs. 12 and 13 report the observed facial thermal variations in the muscles of upper and lower face, respectively.

The five facial muscles included in our analyses of upper face, showed different thermal intensity patterns across the facial expressions. A '7expressions $\times$ 5muscles' ANOVA resulted in an overall difference in temperature variations along these facial muscles across the expressions ($F=3.65$, $p=0.02$). A significant interaction effect ($F=30.35$, $p<=0.004$), indicated that the variations in thermal intensity values measured along the major facial muscles had different magnitudes. Similar statistical results were also reported in [30]. In the following analysis, thermal intensity values were examined and compared for exploring the facial thermal response patterns under the influence of various affective states. In some previous works, *orbicularis oris triangularis, depressor labii inferioris, frontalis pars medialis* and *orbicularis oculi pars lateralis* experienced noticeable thermal variations during the involuntary expression of emotive states. However, during the voluntary expressions of the same affects, comparatively less magnitudes of thermal variations are observed [30], [48], [49], [50], [52], [79]. When images in our dataset were compared with other datasets [48], [49], [50], [52], [79], *risorius* experienced lower thermal variations during the voluntary expression of positive affects than what was experienced during involuntary expression of the same affects. Increased thermal variations were also observed during the voluntary expression of negative and positive emotive states along *frontalis pars lateralis.*

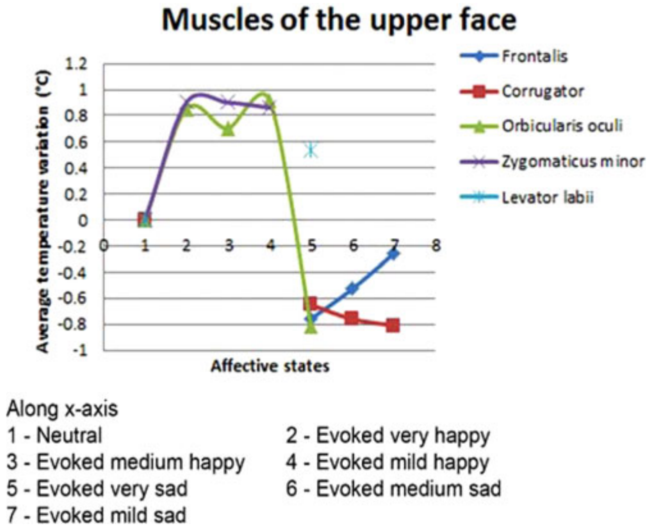

Fig. 12. Average variation in thermal intensity values (°C) measured along five muscles of upper faces.

An earlier analysis [52] showed that the two well-recognized muscles of disgust; *lavatory labii superioris* and *orbicularis oculi* experienced higher thermal variations when disgust was expressed. The two muscles experienced much lower thermal variations under the influence of other emotive states (Fig. 14). The thermal variations observed on *corrugator* under the influence of disgust were much higher than the thermal variations observed under the influence of

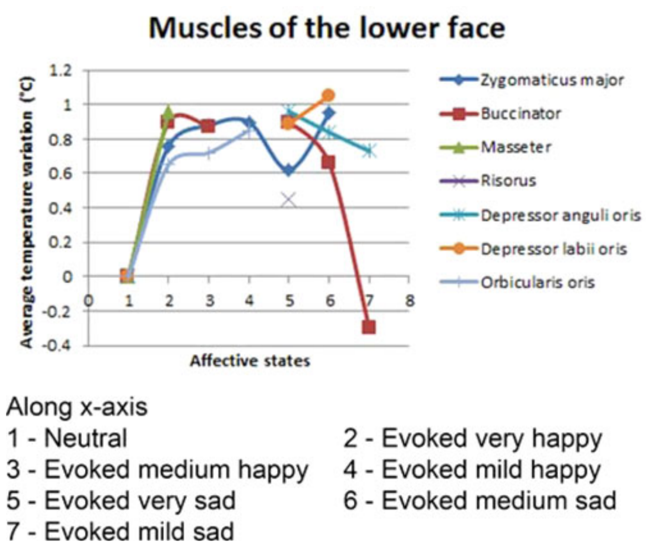

Fig. 13. Average variation in thermal intensity values (°C) measured along seven muscles of lower faces.

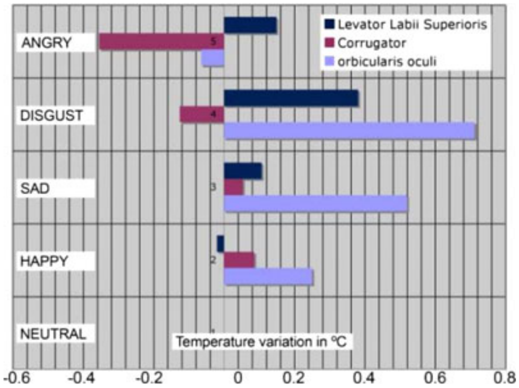

Fig. 14. Thermal variations (in °C) observed along the three muscles known to be involved in expression of disgust. Thermal variations were observed under the influence of negative and positive expressions of affective states.

sadness and happiness. The temperature measurements reported in this paper and in previous works [48], [49], [50], [52], [79] made it obvious that involuntary expressions of emotive states would cause varying facial thermal variations along the major facial muscles. The facial thermal response patterns evident in Figs. 12 and 13 also suggested that the nature and intensity of affect experience would also influence the facial thermal response patterns.

7 Results

Facial expressions of positive emotional experiences were found to cause significant (measurable) thermal variations along major facial muscles; *zygomaticus* major, *orbicularis oris*, *mentalis*, and *platysma* [80], [81], [82]. In previous studies, facial expression of negative emotional experiences would cause deformations and musculo-thermal variations along *corrugator*, *masseter*, *triangularis*, *orbicularis oculi palpabraeous*, *platysma*, and *bucccinator*. Studies have also found that orbicularis is engaged in facial expression of both; positive and negative emotional experiences [80], [81], [82]. Our results were consistent with previous findings. In our dataset, the observed facial thermal variations would cause varying degrees of thermo-muscular activities along the facial muscles. The results show how the skin temperature would vary under the influence of positive and negative affects at various locations along the facial muscles.

Table 2 shows the affect-classification matrix. It highlights that the neutral emotive condition and expressions of happiness and sadness can be separated in a three-affect, multi-arousal-level classification space. Table 3 shows arousal level can be identified after recognition of a single affect. As obvious in Table 3, classification of arousal levels of a single affect into high, medium and mild was quite successful. Arousal classification results suggest that our classifier was more successful in distinguishing between the arousal levels of happiness as compared to those of sadness.

Overall, these preliminary results suggest that measurements of facial thermal variations can help in recognizing affective states and affect arousal levels.

7.1 Comparison with Other Systems

Visual-sign based affect and arousal assessment methods have shown a wide range of successes. For example, in [36] classification success of a classifier varied between 62.5 percent (for expression of disgust) to 100 percent (for expressions of anger and surprise). Around 96 percent of the happy and sad facial expressions were correctly classified. The expression of fear was also classified with 76 percent accuracy [36].

In [31], the frontal brain EEG (electroencephalogram) signals were used to assess the effectiveness of emotion elicitation. The reported classification success varied between 75.0 and 93.3 percent for one individual and between 45.6 and 61.3 percent for another individual. It could be concluded that within-group and between-group variances had a major effect on the classification results.

Our classification results, reported in Tables 2 and 3, varied between 68.4-84.2 percent for affect recognition and between 57.9-84.2 percent for arousal level assessment. As was the case with visual-sign based and EEG-supported affect classifiers, our thermal data had some built-in within and between group variances. Despite that, our classification schema was able to produce promising affect and arousal classification results. These results are comparable with those obtained using other visual and psychophysiological cues. Overall, these results show that non-contact-based facial skin temperature measurements can provide psychophysiological bases for affect and arousal level recognition.

In order to test the efficacy of our employed classification algorithm, we invoked it on 10 individuals' images from the IEEE OTCBVS Benchmark dataset of neutral, happy and angry expressions. The confusion matrix in Table 4 shows that our algorithmic approach was able to successfully classify more than 96 percent images showing different affective states. The IEEE OTCBVS dataset images were also classified in [99] and excellent results were observed; 70 percent of happy facial expression images and 84 percent of angry facial

TABLE 2
Person-Independent Classification of Neutral, and Evoked (Happy and Sad) Facial Expressions

Classification Type	Facial Expression	Predicted Group Membership			
		Neutral	*Happy*	*Sad*	*Total*
Cross validation [a,b] (%)	Neutral	84.2 % (16)	10.5% (02)	5.2% (1)	100% (19)
	Happy	15.8% (03)	73.7% (14)	10.5% (02)	100% (19)
	Sad	21.1% (4)	10.5% (2)	68.4% (13)	100% (19)

[a] *100 percent of original groups were correctly classified.* [b] *73.03 percent of cross-validated group cases correctly classified.*

TABLE 3
Recognition Rate for Person-Independent Classification of Arousal Levels: Happy and Sad Facial Expressions

Facial Expression	Arousal Levels		
	High Arousal	*Medium Arousal*	*Mild Arousal*
Happy	84.2% (16)	68.4% (13)	73.7% (14)
Sad	73.7% (14)	57.9% (11)	68.4% (13)

[a] *Each arousal level was assessed on 19 facial thermal images.*

TABLE 4
Results of Classifying Neutral, Happy and Angry Facial Expression Images of The IEEE OTCBVS Dataset**

Classification Results[a]						
			Predicted Group Membership			Total
		Express	1	2	3	
Original	Count	1	10	0	0	10
		2	0	10	0	10
		3	0	1	9	10
	%	1	100.0	.0	.0	100.0
		2	.0	100.0	.0	100.0
		3	.0	10.0	90.0	100.0

[a] *96.7 percent of original grouped cases correctly classified.*
** *Arabic numerals 1, 2 and 3 respectively show neutral, happy and angry expressions.*

expression images were correctly classified. Our algorithm, as shown in Table 4, was able to correctly classify 100 percent happy and 90 percent angry facial expression images. In [99], 16 percent of angry facial expression images were interpreted as happy expressions. Our algorithm, confused only 10 percent angry expression images with happy expression images. Though a direct comparison between the two algorithms looks infeasible, our algorithmic approach was able to do a little better than the one employed in [99].

8 Conclusion and Discussion

These results highlight that: (1) It is possible to use thermal infrared imaging in automated affect recognition and dynamic affect-arousal assessment; (2) Affect induced facial thermal variations may help in determining the specificity and extent of discrete negative and positive affective states and affective experiences; (3) The employed human information and algorithmic approach may help in overcoming some of the problems associated with dynamic assessment of affect-arousal level, reported in Section 2.

The proposed approach, requiring non-contact-based facial thermal variation measurement, can meet any stringent ethical framework. It relies on a harmless psychophysiological signal that is also used in medical diagnostic procedures.

The proposed thermal data measurement protocols can be repeated and are less likely to influence human affective responding patterns.

The proposed approach provides a set of data acquisition, processing and analyses methods that can be applied in a variety of real life situations. Thus the process of dynamic affect and arousal assessment can be standardized. The proposed approach may also be extended for examining affect latency, rise time and affect duration through careful examination of thermal infrared video clips.

The visual-sign-supported assessment criteria proposed for identifying affect arousal can be validated and standardized. The proposed affect-arousal assessment criteria can also be used for vision supported affect arousal-level assessment in a variety of situations.

The empirical data presented in this work provides a detailed and comprehensive account of how emotive states would cause variations in facial skin temperature along the major muscles. The reported data corroborates with a number of data presented in earlier studies.

Certain techniques, such as image enhancement, feature extraction; selection and image classification are essential in performing image analyses and pattern classification. Some improvements to these techniques, as this work demonstrates, can allow localizing and detecting facial thermal changes and ascertaining the nature of affective experiences.

The pattern analysis approach used in step-1 was developed to include both: affective state related within and between-group thermal variations; and arousal related within and between group thermal variations. The idea was to train the classifier for dealing with uncertainties caused by affective state and affect-arousal. Thus a complex and larger discrimination space was included in the classification. Our results demonstrated that the classifier would attain an acceptable classification rate without compromising the degree of recognition accuracy.

The classification approach in step-2 was conservative. A smaller mixture model of arousal levels of a single affective state was employed for arousal assessment. Since arousal induced “within and between” group thermal variations were used in the mixture, a very good level of arousal level assessment was observed. This conservative approach enabled judging the arousal levels using measurements of a single physiological cue. For the future, it would be helpful to look at arousal effects from a more temporal point of view instead of the three-level exploratory approach used here. There exist appropriate methods in multivariate time-series analysis to look at the evolution of affect intensity; for example autocorrelation of time-sampled responses at suitable sampling rates.

In order to classify the facial thermal data as affect/arousal classes, several well-established algorithmic methods were considered for designing a classifier. The employed 2-stage affect and arousal classification schema is a modified version of methods widely used in model-based pattern classification. The 2-stage classification schema resulted in a computationally efficient system; had a swift flow of execution; was helpful in reducing the data dimensions; and enabled reducing the built-in data biases and noise.

In our view, this first exploratory study indicates that affect assessment in the infrared spectrum offers great application promise - forensic and security-related in the detection of dissimulation by suspect individuals. Perhaps, the greatest promise lies in human interaction with IR sensory robots. A robot, aware of the fact that its behavior or gestures were causing concern in a human interlocutor could embody new concepts in artificial affective intelligence.

8.1 Limitations

The thermal images for this work were acquired in a highly controlled and comfortable building environment. All the participants in reported experiments were young and healthy students. In the future, it would be beneficial to include a larger number of people with broader range of ethnic and cultural backgrounds, skin colors and health conditions.

Our proposed approach won’t allow easy and quick separation between the history-extrinsic and stimulus-intrinsic elements of affective states. It should be noted that almost all prevailing methods of affect/arousal assessment suffer from this inability to distinguish between the extrinsic and intrinsic elements of emotive experiences.

REFERENCES

[1] K. R. Scherer, "What are emotions and how they can be measured," *Social Sci. Inform.*, vol. 4453, no. 4, pp. 695–729, 2005.

[2] K. R. Scherer, "Appraisal considered as a process of multi-level sequential checking," in *Appraisal Process in Emotion*, K. R. Scherer, A. Schorr and T. Johnston, Eds. New York, NY, USA: Oxford Univ. Press, 2001, pp. 92–120.

[3] P. Ekman, W. V. Friesen and S. Ancoli, "Facial signs of emotional experience," *J. Personality Social Psychol.*, vol. 39, pp. 1125–1134, 1980.

[4] J. P. J. Pinel, *Biopsychology*. Boston, MA, USA: Allyn & Bacon, 1990, pp. 573–578.

[5] X. Ax, "The physiological differentiation between fear and anger in humans," *Psychosomatic Med.*, vol. 15, pp, 433–442, 1955.

[6] R. Plutchik, "The nature of emotion: Clinical implications," in *Emotion Elicitation and Psychopathology*, M. Clynes and J. Panksepp, Eds. New York, NY, USA: Plenum, 1988, pp. 1–20.

[7] J. P. Lang, "Anxiety: Towards a psychophysiological definition," in *Psychiatric Diagnosis: Exploration of Biological Criteria*, H. S. Akiskai, and W. L. Webb, Eds. New York, NY USA: Spectrum Press, 1978, pp. 265–389.

[8] C. Breazeal and B. Scassellati, "How to build robots that make friends and influence people," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 1999, pp. 858–863.

[9] D. Kulic and E. A. Croft, "Affective state estimation for human-robot interaction," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 991–1000, Oct. 2007.

[10] A. Kleinsmith, N. Bianc hi-Berthhouze and A. Steed, "Automatic recognition of non-acted affective states," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 4, pp. 1027–1038, Aug. 2011.

[11] E. Dror, A. E. Péron, S. L. Hind, and D. Charlton, "When emotions get better of us: The effect of contextual top-down processing on matching fingertips," *Appl. Cognitive Psychol.*, vol. 19, pp. 799–809, 2005.

[12] C. Hirsch, and A. Mathews, "Interpretative inferences when reading about emotional events," *Behavioral Res. Therapy*, vol. 35, no. 12, pp. 1123–32. 1997.

[13] S. T. Murphy, and R. B. Zajonc, "Affect, cognition and awareness: affective priming with optimal and suboptimal stimulus exposures," *J. Personality Social Psychol.*, vol. 64, pp. 723–739, 1993.

[14] B. Wild, M. Erb, and M. Bartels, "Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: Quality, quantity, time course and gender differences," *Psychiatry Res.*, vol. 102, pp. 109–124, 2001.

[15] *The Australian Psychological Society, Code of Ethics*. Melbourne, Australia: The Australian Psychological Society Limited, 2003.

[16] Belmont Report. (1979). Ethical principles and guidelines for the protection of human subjects of research. National Institutes of Health, U.S. Government, Washington, DC, USA, [Online]. Available: http://ohsr.od.nih.gov/guidelines/belmont.html

[17] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen, X. Wang, "A natural visible and infrared facial expression database for expression recognition and emotion inference," *IEEE Trans. Multimedia*, vol. 12, no. 7, pp. 682–690, Nov. 2010.

[18] J. Rottenberg, R. D. Ray, and J. J. Gross, "Emotion elicitation using films," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 9–28.

[19] J. Humrichouse, M. Chmielewski, E. A. McDade-Montez, and D. Watson, "Affective assessment through self-report methods," in *Emotion and Psychopathology, Bridging Affective and Clinical Science*, J. Rottenberg, and S. L. Johnson, Eds. Washington, DC, USA: American Psychological Assoc., 2007, pp. 13–34.

[20] E. K. Gray and A. D. Watson, "Assessing positive and negative affect via self-report," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 171–183.

[21] E. Harmon-Jones, D. M. Amodio, and L. R. Zinner, "Social psychological methods of emotion elicitation," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 106–123.

[22] E. Eich, J. T. W. Ng, D. Macaulay, A. D. Percy, and I. Grebneva, "Combining music with thought to change mood," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 124–136.

[23] E. Wilson, C. MacLeod, and L. Campbell, "The information processing approach to emotion research," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Ed. New York, NY, USA: Oxford Univ. Press, 2007, pp. 184–202.

[24] J. F. Cohn, Z. Ambadar, and P. Ekman, "Observer-based measurement of facial expression with the facial action coding system," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 203–221.

[25] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 2000, pp. 123–135.

[26] J. M. Gottman and R. W. Levenson, "A valid procedure for obtaining self-report of affect in marital interaction," *J. Consulting Clin. Psychol.*, vol. 53, no. 1, pp. 151–160, 1985.

[27] E. L. Rosenberg and P. Ekman, "Coherence between expressive and experimental systems in emotion," *Cognition Emotion*, vol. 8, pp. 201–229, 1994.

[28] R. J. Davidson, P. Ekman, C. D. Saron, J.A. Sanulis and W.V. Friesen, "Approach-withdrawal and cerebral asymmetry: Emotional experience and brain physiology," *J. Personality Social Psychol.*, vol. 58, pp. 330–341, 1990.

[29] C. E. Izard, *The Maximally Discriminative Facial Movement Coding System (MAX)*. Newark, DE, USA: Instructional Resources Centre, University of Delaware, 1979.

[30] M. M. Khan, R. D. Ward and M. Ingleby, "Distinguishing facial expressions by thermal imaging using facial thermal feature points," in *Proc. 19th Brit. HCI Group Annu. Conf.*, Sep. 2005, vol. 2, pp. 10–14.

[31] P. C. Petrantonakis and L. J. Hadjileontiadis, "A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition," *IEEE Trans. Inform. Technol. Biomed.*, vol. 15, no. 5, pp. 737–746, Sep. 2011.

[32] B. Fasel, and J. Luettin, "Automatic facial expression analysis: a survey," *Pattern Recog.*, vol. 36, pp. 259–275, 2003.

[33] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Trans. Pattern Anal. Mach. Understanding*, vol. 22, no. 12, pp. 1424–1445, Dec. 2000.

[34] K. R. Scherer, H. G. Wallbot, D. Matsumoto and T. Kudoh, "Emotional experience in cultural context: A comparison between Europe, Japan and the United States," in *Facets of Emotion*, K. R. Scherer, Ed. NJ, USA: Lawrence Erlbaum, 1998, pp. 5–30.

[35] Z. Zeng, M. Pantic, G. I. Roisman and T. S. Huang, "A survey of affect recognition methods: Audio, visual and spontaneous expressions," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 31, no. 1, pp. 39–58, Jan. 2009.

[36] M. Yeasin, B. Bullot and R. Sharma, "Recognition of facial expressions and measurement of level of interest from video," *IEEE Trans. Multimedia*, vol. 8, no. 3, pp. 500–508, 2006.

[37] D. Gavhed, T. Makinen, I. Holmer and H. Rintamaki, "Face temperature and cardiorespiratory responses to wind in thermoneutral and cool subjects exposed to –10 °C," *Euro. J. Appl. Physiol.*, vol. 83, pp. 449–456, 2000.

[38] J. LeBlanc, B. Blais, B. Barabe, and J. Cote, "Effects of temperature and wind on facial temperature, heat rate and sensation," *J. Appl. Physiol.*, vol. 40, pp. 127–131, 1976.

[39] D. Shastri, M. Papadakis, P. Tsiamyrtzis, B. Bass, and I. Pavlidis, "Perinasal imaging of physiological stress and its affective potential," *IEEE Trans. Affective Comput.*, vol. 3, no. 3, pp. 366–378, Jul.-Sep. 2012.

[40] M. A. Stroud, "Effects on energy expenditure of facial cooling," *Euro. J. Appl. Physiol.*, vol. 63, pp. 376–380, 1991.

[41] P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. Pavlidis, M. G. Frank, and P. Ekman, "Imaging facial physiology for the detection of deceit," *Int. J. Comput. Vision*, vol. 71, no. 2, pp. 197–214, Feb. 2007.

[42] S. E. Rimm-Kaufman and K. Kagan, "The physiological significance of changes in skin temperature," *Motivation Emotion*, vol. 20, no. 1, pp. 63–78, 1996.

[43] I. Pavlidis and J. Levine, "Thermal image analysis for polygraph testing," *IEEE Eng. Med. Biol.*, vol. 21, no. 6, pp. 56–64, Nov./Dec. 2002.

[44] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, "StressCam: Non-contact measurement of users' emotional states through thermal imaging," in *Proc. Comput.-Human Interaction*, Portland, Oregon, 2005, pp. 1725–1728.

[45] I. Pavlidis, "Lie detection using thermal imaging," *Proc. SPIE*, vol. 5405, pp. 270–279, 2004.

[46] Y. Sugimoto, Y. Yoshitomi, and S. Tomita, "A method of detecting transitions of emotional states using a thermal facial image based on a synthesis of facial expressions," *Robot. Autonom. Syst.*, no. 31, pp. 147–160, 2000.

[47] Y. Yoshitomi, S-I, Kim, T. Kawano, and T. Kitazoe, "Effects of sensor fusion for recognition of emotional states using voice, face image and thermal image of face," in *Proc. IEEE Int Workshop Robot. Human Interactive Commun.*, Osaka, Japan, 2000, pp. 178–183.

[48] M. M. Khan, M. Ingleby and R. D. Ward, "Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variation," *ACM Trans. Autonom. Adaptive Syst.*, vol. 1, no. 1, pp. 91–113, 2006.

[49] M. M. Khan, R. D. Ward and M. Ingleby, "Classifying pretended and evoked facial expression of positive and negative affective states using infrared measurement of facial skin temperature," *ACM Trans. Appl Perception*, vol. 6, no. 1, pp. 6:1–6:22, 2009.

[50] M. M. Khan, "Cluster-analytic classification of facial expressions using infrared measurements of facial thermal features," Ph.D. Thesis, Dept. Comput. Eng., Univ. of Huddersfield, Huddersfield, U.K., 2008.

[51] G. Koukiou, G. Panagopoulos and V. Anastassopoulos, "Drunken person identification using thermal infrared images," in *Proc. Digit. Signal Process.*, 2009, pp. 1–4.

[52] M. M. Khan, "Cluster analytic detection of disgust arousal" in *Proc. 9th Int. Conf. Intell. Syst. Design Appl.*, Rome, Italy 2009, pp. 641–647.

[53] M. M. Khan, R. D. Ward, M. Ingleby, "Automated classification and recognition of facial expressions using infrared thermal imaging," in *Proc. IEEE Conf. Cybern. Intell. Syst.*, 2004, pp. 202–206.

[54] A. Nozawa and M. Tacano, "Correlation analysis on alpha attenuation and nasal skin temperature," *J. Statist. Mech., Theory Exp.*, no. 1, Jan. 2009, article no. 01007.

[55] B. R. Nhan and T. Chau, "Classifying affective states using thermal infrared imaging of the human face," *IEEE Trans. Biomed. Eng.*, vol. 57, pp. 979–987, Apr. 2010.

[56] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (IAPS): Affective ratings of pictures and instruction manual," Univ. Florida, Gainesville, FL, USA, *Tech. Rep. A-6*, 2005.

[57] K. Otsuka, S. Okada, M. Hassan, T. Togawa, "Imaging of skin thermal properties with estimation of ambient radiation," *IEEE Eng. Med. Biol.*, vol. 21, no. 6, pp. 49–55, Nov./Dec. 2002.

[58] B. F. Jones and P. Plassmann, "Digital infrared thermal imaging of human skin," *IEEE Eng. Med. Biol.*, vol. 21, no. 6, pp. 41–48, Nov./Dec. 2002.

[59] *Emotion in the Human Face*, P. Ekman, Ed., 2nd ed. Cambridge, U.K. Cambridge Univ. Press, 1982.

[60] P. Ekman and H. Oster, "Facial expressions of emotion," *Ann. Rev. Psychol.*, vol. 30, pp. 527–554, 1979.

[61] J. A. Russell, J. Bachorowski, and J. Fernandez-Dols, "Facial and vocal expressions of emotion," *Annu. Rev. Psychol.*, vol. 54, pp. 329–349, 2003.

[62] K. R. Scherer, "Appraisal theory," *Handbook of Cognition and Emotion*. T. Dalgleish and M.J. Power, Eds. Hoboken, NJ, USA: Wiley, 1999, pp. 637–663.

[63] L. Chen, T. S. Huang, T. Miyasato, and R. Nakatsu, "Multimodal human emotion/expression recognition," in *Proc. IEEE Int. Conf. Automat. Face Gesture Recog.*, 1998, pp. 396–401.

[64] J. A. Russel, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

[65] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[66] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer-Verlag, 2002.

[67] B. S. Everitt and G. Dunn, *Applied Multivariate Data Analysis*. London, U.K.: Wiley, 1991.

[68] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Vis. Pattern Recog.*, HI, USA, Jun. 1999, pp. 586–591.

[69] V. Belhumeur, J. Hespanda and D. Kiregeman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[70] X. Wang and X. Tang, "Bayesian face recognition based on Gaussian mixture models," in *Proc. 17th IEEE Int. Conf. Pattern Recogn.*, vol. 4, 2004, pp. 142–145.

[71] R Gross, J Yang, A Waibel, "Growing Gaussian mixture models for pose invariant face recognition," in *Proc. 15th Int. Conf. Pattern Recogn.*, 2000, pp. 1088–1091.

[72] H. Kim, D. Kim and S. Y. Bang, "Face recognition using LDA mixture model," in *Proc. Int. Conf. Pattern Recogn.*, 2002, pp. 486–489.

[73] G. J. McLachlan. *Discriminant Analysis and Statistical Pattern Recognition*. NJ, USA: Wiley, 2004.

[74] C. J. Huberty, *Applied Discriminant Analysis*. New York, NY, USA: Wiley, 1994.

[75] D. Séverine, F. Davoine, and M. Masson. "A solution for facial expression representation and recognition." *Signal Process, Image Commun.*, vol. 17, no. 9, pp. 657–673, 2002.

[76] Gupta, A.K., and Logan, T.P. "On a multiple observations model in discriminant analysis," *J. Statist. Comput. Simul.*, vol. 34, pp. 119–132, 1990.

[77] B Moghaddam, A Pentland, "Probabilistic visual learning for object detection," in *Proc. 5th Int. Conf. Comput. Vis.*, 1995, pp. 786–793.

[78] S. Lawrence, P. Yianilos, and I. Cox, "Face recognition using mixture-distance and raw images," in *Proc. IEEE Int. Conf. Syst. Man, Cybern.*, NJ, USA, 1997, pp. 2016–2021.

[79] M. M. Khan, R. D. Ward and M. Ingleby, "Capturing physiology of emotion along facial muscles: A method of distinguishing feigned from involuntary expressions," in *Proc. 13th Int. Conf. Comput. Anal. Images Patterns*, Sep. 2009, pp. 1196–1203.

[80] R. Kall, 1990, "Emotional self regulation and facial expression muscle measurement and training," in J. R. Cram and J. V. Basmalian (Eds). *Clinical EMG for Surface Recording*, vol. 2, Nevada City, CA, USA: Clinical Resources, 1990.

[81] A. A. Root and J. A. Stephens, "Organization of the central control of muscles of facial expression in man," *J. Physiol.*, no. 549, pp. 289–298, 2003.

[82] S. R. Vrana and D. Gross, "Reactions to facial expressions: Effects of social context and speech anxiety on response to neutral, anger and joy expressions," *Biological Psychol.*, vol. 66, no. 1, pp. 63–78, 2004.

[83] G. E. Schwartz, P. L. Fair, P. Salt, M. R. Mandel and G. L. Klerman. "Facial muscle patterning to affective imagery in depressed and non-depressed subjects," *Science*, vol. 192, pp. 489–491, 1976.

[84] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.

[85] C. Darwin, *The Expression of Emotion in Man and Animals*. London, U.K.: Murray, 1872.

[86] J. A. Coan and J. J. B. Allen, *Handbook of Emotion Elicitation and Assessment*. London, U.K.: Oxford Univ. Press, 2007.

[87] S. Jarlier, et al., "Thermal analysis of facial muscles contractions," *IEEE Trans. Affective Comput.*, vol. 2, no. 1, pp. 2–9, Jan.-Mar. 2011.

[88] E. I. Barakova, R. Gorbunov, M., Rauterberg, "Automatic interpretation of affective facial expressions in the context of interpersonal interaction," *IEEE Trans. Human-Mach. Syst.*, vol. 45, no. 4, pp. 409–418, Aug. 2015.

[89] T. Partala, V. Surakka, T. Vanhala, "Real-time estimation of emotional experiences from facial expressions," *Interacting Comput.*, no. 18, vol. 2, pp. 208–226, 2006.

[90] X. Chen, and T. Huang, "Facial expression recognition: A clustering-based approach," *Pattern Recogn. Lett.*, vol. 24, pp. 1295–1302, 2003.

[91] M. Lyons, J. Budynek, and S. Akamatsu, "Classifying images of facial expression using a Gabor wavelet representation," in *Proc. 2nd Int. Conf. Cognitive Sci.*, Tokyo, Japan, 1999, pp. 113–118.

[92] A. R. Webb, *Statistical Pattern Recognition*. London, U.K.: Wiley, 2002.

[93] S. K. Krishnan and P. V. S. Rao, "Feature selection for pattern classification with Gaussian mixture models: A new objective criterion," *Pattern Recogn. Lett.*, vol. 17, pp. 803–809, 1996.

[94] G. Cottrell and J. Metacalfe, "EMPATH: Face, gender and emotion recognition using holons," in *Proc. Conf. Adv. Neural Inf. Process. Syst.*, 1991, vol. 3, pp. 564–571.

[95] H-C, Kim, D. Kim, S. Y. Bang, "An efficient model order for PCA mixture model," *Pattern Recogn. Lett.*, vol. 24, pp. 1385–1393, 2003.

[96] OTCBVS. (2007). OTCBVS Benchmark Dataset Collection. Object Tracking and Classification in and Beyond the Visible Spectrum, Oklahoma State University. Accessed July 3, 2012, [Online]. Available: http://www.cse.ohiostate.edu/otcbvs-bench/

[97] A. Wesley, P. Buddharaju, R. Pienta, and I. Pavlidis. "A comparative analysis of thermal and visual modalities for automated facial expression recognition," *Adv. Vis. Comput.*, vol. 7432, pp. 51–60, 2012.
[98] W. Zhao, A. Krishnaswamy, R. Chellappa, D. Swets, and J. Weng, "Discriminant analysis of principal components for face recognition," in *Face Recognition-NATO ASI Series*, H. Wechsler, P. Phillips, V. Bruce, F. Soulié, and T. Huang, Eds. Berlin, Germany: Springer, 1998, vol. 163, pp 73–85.
[99] B. Hernández, G. Olague, R. Hammoud, L. Trujillo, and E. Romero, "Visual learning of texture descriptors for facial expression recognition in thermal imagery," *Comput. Vis. Image Understand*, vol. 106, no. 2, pp. 258–269, 2007.
[100] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recogn. Workshops*, 2010, pp. 94–101.

Masood Mehmood Khan received the BE, MS, and PhD degrees from NED University of Engineering and Technology, Karachi; Colorado State University, CO.; and the University of Huddersfield, United Kingdom, respectively. He had taught at the National University of Computer and Emerging Sciences, Karachi; Jefri Bolkiah College of Engineering Kuala Belait, Brunei Darussalam; and the American University of Sharjah, UAE before joining the Mechanical Engineering Department at Curtin University, Australia. His research activities revolve around affective computing, machine vision and perception, human-computer interaction and bio-medical imaging. He is member of the IEEE.

Robert D. Ward received degree in psychology from Hull University and the master's degree in computing from Manchester. He returned to Hull to research computer-based developmental language remediation, and received the PhD degree in 1988. Following further research into intelligent tutoring systems at Aberdeen University, he spent several years in the computing industry. He has recently retired from the University of Huddersfield where he taught human-computer interaction, information systems and multimedia.

Michael Ingleby received the BA degree from Cambridge University (where he was a W.S. Churchill Prize Scholar) and the PhD degree from McMaster University (where he was a NATO Exchange Fellow). He has taught and won international research grants at the Universities of McMaster, Sussex, Huddersfield, Yaoundé and the École Supérieure Interafricaine de l'Électricité. He is now retired but continues research as a visiting Professor and Research Fellow at several Universities. In the 1970s, he changed his specialization from mathematical physics to AI/cognitive science. His focus is on robust pattern recognition, with applications to automatic understanding of images and speech, to forensic and biometric identification and to financial market dynamics, where he uses hidden Markov models, Brownian motion models, Bayesian nets, cluster analysis and concept lattices.