

Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals

Christine Lætitia Lisetti

Department of Multimedia Communications, Institut Eurecom, 06904 Sophia-Antipolis, France
Email: lisetti@eurecom.fr

Fatma Nasoz

Department of Computer Science, University of Central Florida, Orlando, FL 32816-2362, USA
Email: fatma@cs.ucf.edu

Received 30 July 2002; Revised 14 April 2004

We discuss the strong relationship between affect and cognition and the importance of emotions in multimodal human computer interaction (HCI) and user modeling. We introduce the overall paradigm for our multimodal system that aims at recognizing its users' emotions and at responding to them accordingly depending upon the current context or application. We then describe the design of the emotion elicitation experiment we conducted by collecting, via wearable computers, physiological signals from the autonomic nervous system (galvanic skin response, heart rate, temperature) and mapping them to certain emotions (sadness, anger, fear, surprise, frustration, and amusement). We show the results of three different supervised learning algorithms that categorize these collected signals in terms of emotions, and generalize their learning to recognize emotions from new collections of signals. We finally discuss possible broader impact and potential applications of emotion recognition for multimodal intelligent systems.

Keywords and phrases: multimodal human-computer interaction, emotion recognition, multimodal affective user interfaces.

1. INTRODUCTION

The field of human-computer interaction (HCI) has recently witnessed an explosion of adaptive and customizable human-computer interfaces which use cognitive user modeling, for example, to extract and represent a student's knowledge, skills, and goals, to help users find information in hypermedia applications, or to tailor information presentation to the user. New generations of intelligent computer user interfaces can also adapt to a specific user, choose suitable teaching exercises or interventions, give user feedback about the user's knowledge, and predict the user's future behavior such as answers, goals, preferences, and actions. Recent findings on emotions have shown that the mechanisms associated with emotions are not only tightly intertwined neurologically with the mechanisms responsible for cognition, but that they also play a central role in decision making, problem solving, communicating, negotiating, and adapting to unpredictable environments. Emotions are now therefore considered as organizing and energizing processes, serving important adaptive functions.

To take advantage of these new findings, researchers in signal processing and HCI are learning more about the unsuspectedly strong interface between affect and cognition

in order to build appropriate digital technology. Affective states play an important role in many aspects of the activities we find ourselves involved in, including tasks performed in front of a computer or while interacting with computer-based technology. For example, being aware of *how* the user receives a piece of provided information is very valuable. Is the user satisfied, more confused, frustrated, amused, or simply sleepy? Being able to know *when* the user needs more feedback, by not only keeping track of the user's actions, but also by observing cues about the user's emotional experience, also presents advantages.

In the remainder of this article, we document the various ways in which emotions are relevant in multimodal HCI, and propose a multimodal paradigm for acknowledging the various aspects of the emotion phenomenon. We then focus on one modality, namely, the autonomic nervous system (ANS) and its physiological signals, and give an extended survey of the literature to date on the analysis of these signals in terms of signaled emotions. We furthermore show how, using sensing media such as noninvasive wearable computers capable of capturing these signals during HCI, we can begin to explore the automatic recognition of specific elicited emotions during HCI. Finally, we discuss research implications from our results.

2. MULTIMODAL HCI, AFFECT, AND COGNITION

2.1. Interaction of affect and cognition and its relevance to user modeling and HCI

As a result of recent findings, emotions are now considered as associated with adaptive, organizing, and energizing processes. We mention a few already identified phenomena concerning the interaction between affect and cognition, which we expect will be further studied and manipulated by building intelligent interfaces which acknowledge such an interaction. We also identify the relevance of these findings on emotions for the field of multimodal HCI.

Organization of memory and learning

We recall an event better when we are in the same mood as when the learning occurred [1]. Hence eliciting the same affective state in a learning environment can reduce the cognitive overload considerably. User models concerned with reducing the cognitive overload [2]—by presenting information structured in the most efficient way in order to eliminate avoidable load on working memory—would strongly benefit from information about the affective states of the learners while involved in their tasks.

Focus and attention

Emotions restrict the range of cue utilization such that fewer cues are attended to [3]; driver's and pilot's safety computer applications can make use of this fact to better assist their users.

Perception

When we are happy, our perception is biased at selecting happy events, likewise for negative emotions [1]. Similarly, while making decisions, users are often influenced by their affective states. Reading a text while experiencing a negatively valenced emotional state often leads to very different interpretation than reading the same text while in a positive state. User models aimed at providing text tailored to the user need to take the user's affective state into account to maximize the user's understanding of the intended meaning of the text.

Categorization and preference

Familiar objects become preferred objects [4]. User models, which aim at discovering the user's preferences [5], also need to acknowledge and make use of the knowledge that people prefer objects that they have been exposed to (incidentally even when they are shown these objects subliminally).

Goal generation and evaluation

Patients who have damage in their frontal lobes (cortex communication with limbic system is altered) become unable to feel, which results in their complete dysfunctionality in real-life settings where they are unable to decide what is the next action they need to perform [6], whereas normal emotional arousal is intertwined with goal generation and decision-making, and priority setting.

Decision making and strategic planning

When time constraints are such that quick action is needed, neurological shortcut pathways for deciding upon the next appropriate action are preferred over more optimal but slower ones [7]. Furthermore people with different personalities can have very distinct preference models (Myers-Briggs Type Indicator). User models of personality [8] can be further enhanced and refined with the user's affective profile.

Motivation and performance

An increase in emotional intensity causes an increase in performance, up to an optimal point (inverted U-curve Yerkes-Dodson Law). User models which provide qualitative and quantitative feedback to help students think about and reflect on the feedback they have received [9] could include affective feedback about cognitive-emotion paths discovered and built in the student model during the tasks.

Intention

Not only are there positive consequences to positive emotions, but there are also positive consequences to negative emotions—they signal the need for an action to take place in order to maintain, or change a given kind of situation or interaction with the environment [10]. Pointing to the positive signals associated with these negative emotions experienced during interaction with a specific software could become one of the roles of user modeling agents.

Communication

Important information in a conversational exchange comes from body language [11], voice prosody, facial expressions revealing emotional content [12], and facial displays connected with various aspects of discourse [13]. Communication will become ambiguous when these are accounted for during HCI and computer-mediated communication.

Learning

People are more or less receptive to the information to be learned depending on their liking (of the instructor, of the visual presentation, of how the feedback is given, or of who is giving it). Moreover, emotional intelligence is learnable [14], which opens interesting areas of research for the field of user modeling as a whole.

Given the strong interface between affect and cognition on the one hand [15], and given the increasing versatility of computers agents on the other hand, the attempt to enable our tools to acknowledge affective phenomena rather than to remain blind to them appears desirable.

2.2. An application-independent paradigm for modeling user's emotions and personality

Figure 1 shows the overall paradigm for multimodal HCI, which was adumbrated earlier by Lisetti [17]. As shown in the first portion of the picture pointed to by the arrow *user-centered mode*, when emotions are experienced in humans, they are associated with physical and mental manifestations.

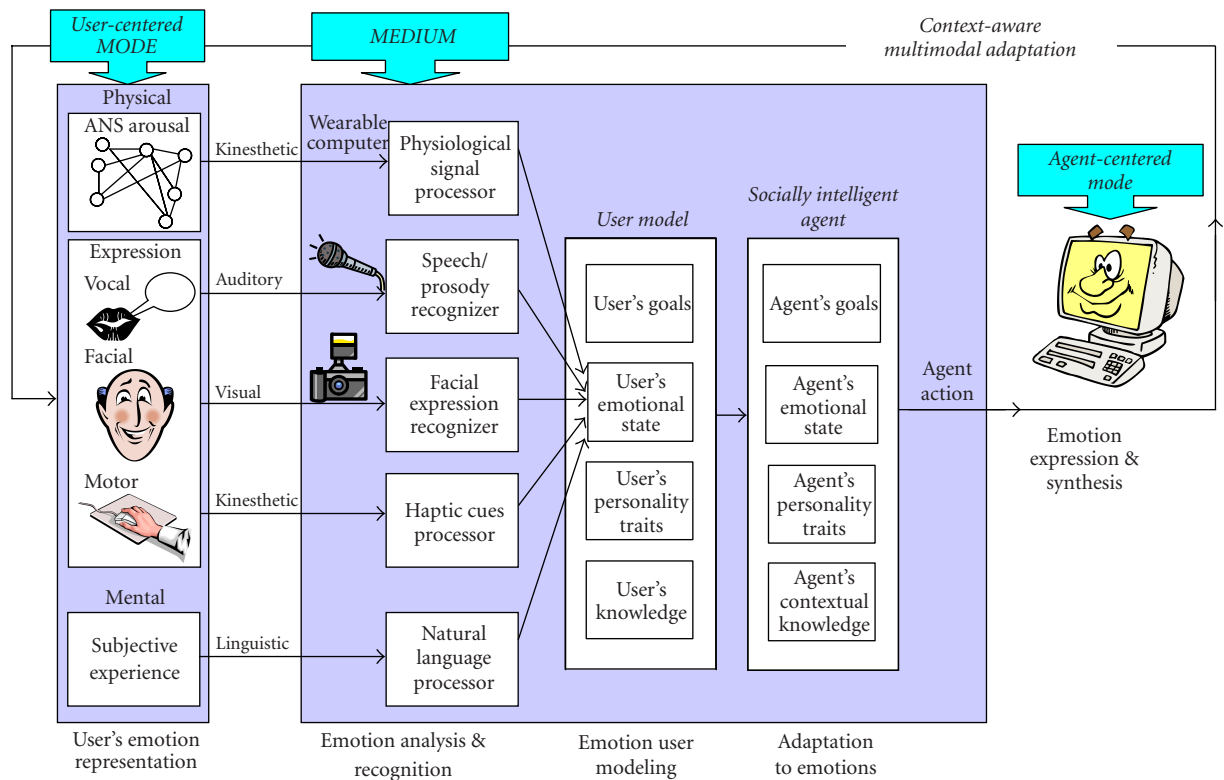


FIGURE 1: The MAUI framework: multimodal affective user interface [16].

The physical aspect of emotions includes ANS arousal and multimodal expression (including vocal intonation, facial expression, and other motor manifestations). The mental aspect of the emotion is referred to here as subjective experience in that it represents what we tell ourselves we feel or experience about a specific situation.

The second part of the Figure 1, pointed to by the arrow *medium*, represents the fact that using multimedia devices to sense the various signals associated with human emotional states and combining these with various machine learning algorithms makes it possible to interpret these signals in order to categorize and recognize the user's almost probable emotions as he or she is experiencing different emotional states during HCI.

A *user model*, including the user's current states, the user's specific goals in the current application, the user's personality traits, and the user's specific knowledge about the domain application can then be built and maintained over time during HCIs.

Socially intelligent agents, built with some (or all) of the similar constructs used to model the user, can then be used to drive the HCIs, adapting to the user's specific current emotional state if needed, knowing in advance the user's personality and preferences, having its own knowledge about the application domain and goals (e.g., help the student learning in all situations, assist in insuring the driver's safety).

Depending upon the application, it might be beneficial to endow our agent with its own personality to best adapt to

the user (e.g., if the user is a child, animating the interaction with a playful or with different personality) and its own multimodal modes of expressions—the *agent-centered mode*—to provide the best adaptive personalized feedback.

Context-aware multimodal adaptation can indeed take different forms of embodiments and the chosen user feedback need to depend upon the specific application (e.g., using an animated facial avatar in a car might distract the driver whereas it might raise a student's level of interest during an e-learning session). Finally, the back-arrow shows that the multimodal adaptive feedback in turn has an effect on the user's emotional states—hopefully for the better and enhanced HCI.

3. CAPTURING PHYSIOLOGICAL SIGNALS ASSOCIATED WITH EMOTIONS

3.1. Previous studies on mapping physiological signals to emotions

As indicated in Table 1, there is growing evidence indeed that emotional states have their corresponding specific physiological signals that can be mapped respectively. In Vrana's study [27], personal imagery was used to elicit disgust, anger, pleasure, and joy from participants while their heart rate, skin conductance, and facial electromyogram (EMG) signals were measured. The results showed that acceleration of heart rate was greater during disgust, joy, and anger imageries than during pleasant imagery; and disgust could be discriminated from anger using facial EMG.

TABLE 1: Previous studies on emotion elicitation and recognition.

Reference	Emotion elicitation method	Emotions elicited	Subjects	Signals measured	Data analysis technique	Results
[18]	Personalized imagery	Happiness, sadness, and anger	20 people in 1st study, 12 people in 2nd study	Facial EMG	Manual analysis	EMG reliably discriminated between all four conditions when no overt facial differences were apparent
[19]	Facial action task, relived emotion task	Anger, fear, sadness, disgust, and happiness	12 professional actors and 4 scientists	Finger temperature, heart rate, and skin conductance	Manual analysis	Anger, fear, and sadness produce a larger increase in heart rate than disgust. Anger produces a larger increase in finger temperature than fear. Anger and fear produce larger heart rate than happiness. Fear and disgust produce larger skin conductance than happiness
[20]	Vocal tone, slide of facial expressions, electric shock	Happiness and fear	60 undergraduate students (23 females and 37 males)	Skin conductance (galvanic skin response)	ANOVA	Fear produced a higher level of tonic arousal and larger phasic skin conductance
[21]	Imagining and silently repeating fearful and neutral sentences	Neutrality and fear	64 introductory psychology students	Heart rate, self report	ANOVA Newman-Keuls pairwise comparison	Heart rate acceleration was more during fear imagery than neutral imagery or silent repetition of neutral sentences or fearful sentences
[22]	Easy, moderately, and extremely difficult memory task	Difficult problem solving	64 undergraduate females from Stony Brook	Heart rate, systolic, and diastolic blood pressure	ANOVA	Both systolic blood pressure (SBP) and goal attractiveness were nonmonotonically related to expected task difficulty
[23]	Personalized imagery	Pleasant emotional experiences (low-effort vs. high effort, and self-agency vs. other-agency)	96 Stanford University undergraduates (48 females, 48 males)	Facial EMG, heart rate, skin conductance, and self-report	ANOVA and regression	Eyebrow frown and smile are associated with evaluations along pleasantness dimension, heart rate measure offered strong support between anticipated effort and arousal. Skin conductance offers further support for that but not as strong as heart rate
[24]	Real life inductions and imagery	Fear, anger, and happiness	42 female medical students (mean age = 23)	Self-report, Gottschalk-Gleser affect scores, back and forearm extensor EMG activity, body movements, heart period, respiration period, skin conductance, skin temperature, pulse transit time, pulse volume amplitude, and blood volume	ANOVA, planned univariate contrasts among means, and pairwise comparisons by using Hotelling's T^2	Planned multivariate comparisons between physiological profiles established discriminant validity for anger and fear. Self-report confirmed the generation of affective states in both contexts

TABLE 1: Continued.

Reference	Emotion elicitation method	Emotions elicited	Subjects	Signals measured	Data analysis technique	Results
[25]	Contracting facial muscles into facial expressions	Anger and fear	12 actors (6 females, 6 males) and 4 researchers (1 female, 3 male)	Finger temperature	Manual analysis	Anger increases temperature, fear decreases temperature
[26]	Contracting facial muscles into prototypical configurations of emotions	Happiness, sadness, disgust, fear, and anger	46 Minangkabau men	Heart rate, finger temperature, finger pulse transmission, finger pulse amplitude, respiratory period, and respiratory depth	MANOVA	Anger, fear, and sadness were associated with heart rate significantly more than disgust. Happiness was intermediate
[27]	Imagery	Disgust, anger, pleasure, and joy	50 people (25 males, 25 females)	Self-reports, heart rate, skin conductance, facial EMG	ANOVA	Acceleration of heart rate was greater during disgust, joy, and anger imageries than during pleasant imagery. Disgust could be discriminated from anger using facial EMG
[28]	Difficult task solving	Difficult task solving	58 undergraduate students of an introductory psychology course	Cardiovascular activity (heart rate and blood pressure)	ANOVA and ANCOVA	Systolic and diastolic blood pressure responses were greater in the difficult standard condition than in the easy standard condition for the subjects who received high-ability feedback, however it was the opposite for the subjects who received low-ability feedback
[29]	Difficult problem solving	Difficult problem solving	32 university undergraduates (16 males, 16 females)	Skin conductance, self-report, objective task performance	ANOVA, MANOVA correlation/regression analyses	Within trials, skin conductance increased at the beginning of the trial, but decreased by the end of the trials for the most difficult condition
[30]	Imagery script development	Neutrality, fear, joy, action, sadness, and anger	27 right-handed males between ages 21–35	Heart rate, skin conductance, finger temperature, blood pressure, electro-oculogram, facial EMG	DFA, ANOVA	99% correct classification was obtained. This indicates that emotion-specific response patterns for fear and anger are accurately differentiable from each other and from the response pattern for neutrality
[31]	Neutrally and emotionally loaded slides (pictures)	Happiness, surprise, anger, fear, sadness, and disgust	30 people (16 females and 14 males)	Skin conductance, skin potential, skin resistance, skin blood flow, skin temperature, and instantaneous respiratory frequency	Friedman variance analysis	Electrodermal responses distinguished 13 emotion pairs out of 15. Skin resistance and skin conductance ohmic perturbation duration indices separated 10 emotion pairs. However, conductance amplitude could distinguish 7 emotion pairs

TABLE 1: Continued.

Reference	Emotion elicitation method	Emotions elicited	Subjects	Signals measured	Data analysis technique	Results
[32]	Film showing	Amusement, neutrality, and sadness	180 females	Skin conductance, inter-beat interval, pulse transit times and respiratory activation	Manual analysis	Interbeat interval increased for all three states, but for the neutrality it was less than the amusement and sadness. Skin conductance increased after the amusement film, decreased after the neutrality film, and stayed the same after the sadness film
[33]	Subjects were instructed to make facial expressions	Happiness, sadness, anger, fear, disgust, surprise	6 people (3 females and 3 males)	Heart rate, general somatic activity, GSR and temperature	DFA	66% accuracy in classifying emotions
[34]	Unpleasant and neutrality film clips	Fear, disgust, anger, surprise, and happiness	46 undergraduate students (31 females, 15 males)	Self-report, electrocardiogram, heart rate, T-wave amplitude, respiratory sinus arrhythmia, and skin conductance	ANOVA, Greenhouse-Geisser correction. Post hoc means comparisons and simple effects analyses	Films containing violent threats increased sympathetic activation, whereas the surgery film increased the electrodermal activation, decelerated the heart rate, and increased the T-wave
[35]	11 auditory stimuli mixed with some standard and target sounds	Surprise	20 healthy controls (as a control group) and 13 psychotic patients	GSR	Principal component analysis clustered by centroid method	78% for all, 100% for patients
[36]	Arithmetic tasks, video games, showing faces, and expressing specific emotions	Attention, concentration, happiness, sadness, anger, fear, disgust, surprise and neutrality	10 to 20 college students	GSR, heart rate, and skin temperature	Manual analysis	No recognition found, some observations only
[37]	Personal imagery	Happiness, sadness, anger, fear, disgust, surprise, neutrality, platonic love, romantic love	A healthy graduate student with two years of acting experience	GSR, heart rate, ECG and respiration	Sequential floating forward search (SFFS), Fisher Projection (FP) and hybrid (SFFS and FP)	81% for by hybrid SFFS and Fisher method with 40 features 54% rate with 24 features
[38]	A slow computer game interface	Frustration	36 undergraduate and graduate students	Skin conductivity and blood volume pressure	Hidden Markov models	Pattern recognition worked significantly better than random guessing while discriminating between regimes of likely frustration from regimes of much less likely frustration

In Sinha and Parsons' study [30], heart rate, skin conductance level, finger temperature, blood pressure, electrooculogram, and facial EMG were recorded while the subjects were visualizing the imagery scripts given to them to elicit neutrality, fear, joy, action, sadness, and anger. The results indicated that emotion-specific response patterns for fear and anger are accurately differentiable from each other and from the response pattern neutral imagery conditions.

Another study, which is very much related to one of the applications we will discuss in Section 5 (and which therefore we describe at length here), was conducted by Jennifer Healey from Massachusetts Institute of Technology (MIT) Media Lab [39]. The study answered the questions about how affective models of users should be developed for computer systems and how computers should respond to the emotional states of users appropriately. The results showed that people do not just create preference lists, but they use affective expression to communicate and to show their satisfaction or dissatisfaction. Healey's research particularly focused on recognizing stress levels of drivers by measuring and analyzing their physiological signals in a driving environment.

Before the driving experiment was conducted, a *preliminary emotion elicitation experiment* was designed where eight states (anger, hate, grief, love, romantic love, joy, reverence, and no emotion: neutrality) were elicited from participants. These eight emotions were Clynes' [40] emotion set for basic emotions. This set of emotions was chosen to be elicited in the experiment because each emotion in this set was found to produce a unique set of finger pressure patterns [40]. While the participants were experiencing these emotions, the changes in their physiological responses were measured.

Guided imagery technique (i.e., the participant imagines that she is experiencing the emotion by picturing herself in a certain given scenario) was used to generate the emotions listed above. The participant attempted to feel and express eight emotions for a varying period of three to five minutes (with random variations). The experiment was conducted over 32 days in a single-subject-multiple-session setup. However only twenty sets (days) of complete data were obtained at the end of the experiment.

While the participant experienced the given emotions, her galvanic skin response (GSR), blood volume pressure (BVP), EMG, and respiration values were measured. Eleven features were extracted from raw EMG, GSR, BVP, and respiration measurements by calculating the mean, the normalized mean, the normalized first difference mean, and the first forward distance mean of the physiological signals. Eleven-dimensional feature space of 160 emotions (20 days \times 8 emotions) was projected into a two-dimensional space by using Fisher projection. Leave-one-out cross validation was used for emotion classification. The results showed that it was hard to discriminate all eight emotions. However, when the emotions were grouped as being (1) anger or peaceful, (2) high arousal or low arousal, and (3) positive valence or negative valence, they could be classified successfully as follows:

- (1) anger: 100%, peaceful: 98%,
- (2) high arousal: 80%, low arousal: 88%,
- (3) positive: 82%, negative: 50%.

Because of the results of the experiment described above, the scope of the driving experiment was limited to recognition of levels of only one emotional state: emotional stress.

At the beginning of the *driving experiment*, participants drove in and exited a parking garage, and then they drove in a city and on a highway, and returned to the same parking garage at the end. The experiment was performed on three subjects who repeated the experiment multiple times and six subjects who drove only once. Videos of the participants were recorded during the experiments and self-reports were obtained at the end of each session. Task design and questionnaire responses were used to recognize the driver's stress separately. The results obtained from these two methods were as follows:

- (i) *task design analysis* could recognize driver stress level as being *rest* (e.g., resting in the parking garage), *city* (e.g., driving in Boston streets), or *highway* (e.g., two-lane merge on the highway) with 96% accuracy;
- (ii) *questionnaire analysis* could categorize four *stress* classes as being lowest, low, higher, or highest with 88.6% accuracy.

Finally, video recordings were annotated on a second-by-second basis by two independent researchers for validation purposes. This annotation was used to find a correlation between stress metric created from the video and variables from the sensors. The results showed that physiological signals closely followed the stress metric provided by the video coders.

The results of these two methods (videos and pattern recognition) coincided in classifying the driver's stress and showed that stress levels could be recognized by measuring physiological signals and analyzing them by pattern recognition algorithms.

We have combined the results of our survey of other relevant literature [18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38] into an extensive survey-table. Indeed, Table 1 identifies many chronologically ordered studies that

- (i) analyze different body signal(s) (e.g., skin conductance, heart rate),
- (ii) use different emotion elicitation method(s) (e.g., mental imagery, movie clips),
- (iii) work with varying number of subjects,
- (iv) classify emotions according to different method(s) of analysis,
- (v) show their different results for various emotions.

Clearly, more research has been performed in this domain, and yet still more remains to be done. We only included the sources that we were aware of, with the hope to assist other researchers on the topic.

TABLE 2: Demographics of subject sample aged 18 to 35 in pilot panel study.

Classification	Gender		Ethnicity			
	Female	Male	Caucasian	African American	Asian American	Hispanic American
Number of subjects	7	7	10	1	2	1

TABLE 3: Movies used to elicit different emotions (Gross and Levenson [41]).

Emotion	Movie	N	Agreement	Mean Intensity*
Sadness	Bambi	72	76%	5.35
	The Champ	52	94%	5.71
Amusement	When Harry Met Sally	72	93%	5.54
Fear	The Shining	59	71%	4.08
	Silence of the Lambs	72	60%	4.24
Anger	My Bodyguard	72	42%	5.22
Surprise	Capricorn One	63	75%	5.05

3.2. Our study to elicit emotions and capture physiological signals data

After reviewing the related literature, we conducted our own experiment to find a mapping between physiological signals and emotions experienced. In our experiment we used movie clips and difficult mathematics questions to elicit targeted emotions—*sadness*, *anger*, *surprise*, *fear*, *frustration*, and *amusement*—and we used BodyMedia SenseWear Armband (BodyMedia Inc., www.bodymedia.com) to measure the physiological signals of our participants: *galvanic skin response*, *heart rate*, and *temperature*. The following subsections discuss the design of this experiment and the results gained after interpreting the collected data. The data we collected in the experiment described below was also used in another study [42]; however in this article we describe a different feature extraction technique which led to different results and implications, as will be discussed later.

3.2.1. Pilot panel study for stimuli selection: choosing movie clips to elicit specific emotions

Before conducting the emotion elicitation experiment, which will be described shortly, we designed a pilot panel study to determine the movie clips that may result in high subject agreement in terms of the elicited emotions (sadness, anger, surprise, fear, and amusement). Gross and Levenson's work [41] guided our panel study and from their study we used the movie scenes that resulted in high subject agreement in terms of eliciting the target emotions. Because some of their movies were not obtainable, and because anger and fear movie scenes evidenced low subject agreement during our study, alternative clips were also investigated. The following sections describe the panel study and results.

Subject sample

The sample included 14 undergraduate and graduate students from the psychology and computer science departments of University of Central Florida. The demographics are shown in Table 2.

Choice of movie clips to elicit emotions

Twenty-one movies were presented to the participants. Seven movies were included in the analysis based on the findings of Gross and Levenson [41] (as summarized in Table 3). The seven movie clips extracted from these seven movies were same as the movie clips of Gross and Levenson's study.

Additional 14 movie clips were chosen by the authors, leading to a set of movies that included three movies to elicit sadness (*Powder*, *Bambi*, and *The Champ*), four movies to elicit anger (*Eye for an Eye*, *Schindler's List*, *American History X*, and *My Bodyguard*), four to elicit surprise (*Jurassic Park*, *The Hitchhiker*, *Capricorn One*, and a homemade clip called *Grandma*), one to elicit disgust (*Fear Factor*), five to elicit fear (*Jeepers Creepers*, *Speed*, *The Shining*, *Hannibal*, and *Silence of the Lambs*), and four to elicit amusement (*Beverly Hillbillies*, *When Harry Met Sally*, *Drop Dead Fred*, and *The Great Dictator*).

Procedure

The 14 subjects participated in the study simultaneously. After completing the consent forms, they filled out the questionnaires where they answered the demographic items. Then, the subjects were informed that they would be watching various movie clips geared to elicit emotions and between each clip, they would be prompted to answer questions about the emotions they experienced while watching the scene. They were also asked to respond according to the emotions they experienced and not the emotions experienced by the actors in the movie. A slide show played the various movie scenes and, after each one of the 21 clips, a slide was presented asking the participants to answer the survey items for the prior scene.

Measures

The questionnaire included three demographic questions: age ranges (18–25, 26–35, 36–45, 46–55, or 56+), gender, and ethnicity. For each scene, four questions were asked. The first question asked, “Which emotion did you experience from this

TABLE 4: Agreement rates and average intensities for movies to elicit different emotions with more than 90% agreement across subjects.

Emotion	Movie	Agreement	Mean Intensity	SD
Sadness	Powder	93%	3.46	1.03
	Bambi	100%	4.00	1.66
	The Champ	100%	4.36	1.60
Amusement	Beverly Hillbillies	93%	2.69	1.13
	When Harry Met Sally	100%	5.00	0.96
	Drop Dead Fred	100%	4.00	1.21
	Great Dictator	100%	3.07	1.14
Fear	The Shining	93%	3.62	0.96
Surprise	Capricorn One	100%	4.79	1.25

$N = 14$

TABLE 5: Movie scenes selected for the our experiment to elicit five emotions.

Emotion	Movie	Scene
Sadness	The Champ	Death of the Champ
Anger	Schindler's List	Woman engineer being shot
Amusement	Drop Dead Fred	Restaurant scene
Fear	The Shining	Boy playing in hallway
Surprise	Capricorn One	Agents burst through the door

video clip (please check one only)?," and provided eight options (anger, frustration, amusement, fear, disgust, surprise, sadness, and other). If the participant checked "other" they were asked to specify which emotion they experienced (in an open choice format). The second question asked the participants to rate the intensity of the emotion they experienced on a six point scale. The third question asked whether they experienced any other emotion at the same intensity or higher, and if so, to specify what that emotion was. The final question asked whether they had seen the movie before.

Results

The pilot panel study was conducted to find the movie clips that resulted in (a) at least 90% agreement on eliciting the target emotion and (b) at least 3.5 average intensity.

Table 4 lists the agreement rates and average intensities for the clips with more than 90% agreement.

There was not a movie with a high level of agreement for anger. Gross and Levenson's [41] clips were most successful at eliciting the emotions in our investigation in terms of high intensity, except for anger. In their study, the movie with the highest agreement rate for anger was *My Bodyguard* (42%). In our pilot study, however, the agreement rate for *My Bodyguard* was 29% with a higher agreement rate for frustration (36%), and we therefore chose not to include it in our final movie selection. However, because anger is an emotion of interest in a driving environment which we are particularly interested in studying, we did include the movie with the highest agreement rate for anger, *Schindler's List* (agreement rate was 36%, average intensity was 5.00).

In addition, for *amusement*, the movie *Drop Dead Fred* was chosen over *When Harry Met Sally* in our final selection due to the embarrassment experienced by some of the subjects when watching the scene from *When Harry Met Sally*.

The final set of movie scenes chosen for our emotion elicitation study is presented in Table 5. As mentioned in Section 3.2.1, for the movies that were chosen from Gross and Levenson's [41] study, the movie clips extracted from these movies were also the same.

3.2.2. Emotion elicitation study: eliciting specific emotions to capture associated body signals via wearable computers

Subject sample

The sample included 29 undergraduate students enrolled in a computer science course. The demographics are shown in Table 6.

Procedure

One to three subjects participated simultaneously in the study during each session. After signing consent forms, they were asked to complete a prestudy questionnaire and the noninvasive BodyMedia SenseWear Armband (shown in Figure 2) was placed on each subject's right arm.

As shown in Figure 2, *BodyMedia SenseWear Armband* is a noninvasive wearable computer that we used to collect the physiological signals from the participants. SenseWear Armband is a versatile and reliable wearable body monitor created by BodyMedia, Inc. It is worn on the upper arm and includes a galvanic skin response sensor, skin temperature sensor, two-axis accelerometer, heat-flux sensor, and a near-body ambient temperature sensor. The system also includes polar chest strap which works in compliance with the armband for heart rate monitoring. SenseWear Armband is capable of collecting, storing, processing, and presenting physiological signals such as GSR, heart rate, temperature, movement, and heat flow. After collecting signals, the SenseWear Armband is connected to the Innerwear Research Software (developed by BodyMedia, Inc.) either with a dock station or wirelessly to transfer the collected data. The data can either

TABLE 6: Demographics of subject sample in emotion elicitation study.

Classification	Gender		Ethnicity				Age range	
	Female	Male	Caucasian	African American	Asian American	Unreported	18 to 25	26 to 40
Number of subjects	3	26	21	1	1	6	19	10



FIGURE 2: BodyMedia SenseWear Armband.

be stored in XML files for further interpretation with pattern recognition algorithms or the software itself can process the data and present it using graphs.

Once the BodyMedia SenseWear Armbands were worn, the subjects were instructed on how to place the chest strap. After the chest straps connected with the armband, the in-study questionnaire were given to the subjects and they were told (1) to find a comfortable sitting position and try not to move around until answering a questionnaire item, (2) that the slide show would instruct them to answer specific items on the questionnaire, (3) not to look ahead at the questions, and (4) that someone would sit behind them at the beginning of the study to time-stamp the armband.

A 45-minute slide show was then started. In order to establish a *baseline*, the study began with a slide asking the participants to relax, breathe through their nose, and listen to soothing music. Slides of natural scenes were presented, including pictures of the oceans, mountains, trees, sunsets, and butterflies. After these slides, the first movie clip played (sadness). Once the clip was over, the next slide asked the participants to answer the questions relevant to the scene they watched. Starting again with the slide asking the subjects to relax while listening to soothing music, this process continued for the anger, fear, surprise, frustration, and amusement clips. The frustration segment of the slide show asked the participants to answer difficult mathematical problems without using paper and pencil. The movie scenes and frustration exercise lasted from 70 to 231 seconds each.

Measures

The prequestionnaire included three demographic questions: age ranges (18–25, 26–35, 36–45, 46–55, or 56+), gender, and ethnicity.

The in-study questionnaire included three questions for each emotion. The first question asked, “*Did you experience SADNESS (or the relevant emotion) during this section of the experiment?*,” and required a yes or no response. The second question asked the participants to rate the intensity of the emotion they experienced on a six-point scale. The third question asked participants whether they had experienced any other emotion at the same intensity or higher, and if so, to specify what that emotion was.

Finally, the physiological data gathered included heart rate, skin temperature, and GSR.

3.2.3. Subject agreement and average intensities

Table 7 shows subject agreement and average intensities for each movie clip and the mathematical problems. A two-sample binomial test of equal proportions was conducted to determine whether the agreement rates for the panel study differed from the results obtained with this sample. Participants in the panel study agreed significantly more to the target emotion for the sadness and fear films. On the other hand, the subjects in this sample agreed more for the anger film.

4. MACHINE LEARNING OF PHYSIOLOGICAL SIGNALS ASSOCIATED WITH EMOTIONS

4.1. Normalization and feature extraction

After determining the time slots corresponding to the point in the film where the intended emotion was most likely to be experienced, the procedures described above resulted in the following set of physiological records: 24 records for anger, 23 records for fear, 27 records for sadness, 23 records for amusement, 22 records for frustration, and 21 records for surprise (total of 140 physiological records). The differences among the number of data sets for each emotion class are due to the data loss for the data of some participants during segments of the experiment.

In order to calculate how much the physiological responses changed as the participants went from a relaxed state to the state of experiencing a particular emotion, we normalized the data for each emotion. Normalization is also important for minimizing the individual differences among participants in terms of their physiological responses while they experience a specific emotion.

Collected data was normalized by using the average value of corresponding data type collected during the relaxation period for the same participant. For example, we normalized the GSR values as follows:

$$\text{normalized_GSR} = \frac{\text{raw_GSR} - \text{raw_relaxation_GSR}}{\text{raw_relaxation_GSR}}. \quad (1)$$

TABLE 7: Agreement rates and average intensities for the elicited emotions.

Emotion	Stimulus: movie or math problem	N	Agreement	Mean intensity	SD
Sadness	The Champ	27	56%	3.53	1.06
Anger	Schindler's List	24	75%	3.94	1.30
Fear	The Shining	23	65%	3.58	1.61
Surprise	Capricorn One	21	90%	2.73	1.28
Frustration	Math problems	22	73%	3.69	1.35
Amusement	Drop Dead Fred	23	100%	4.26	1.10

After data signals were normalized, features were extracted from the normalized data. Four features were extracted for each data signal type: *minimum*, *maximum*, *mean*, and *variance* of the normalized data. We stored the data in a three-dimensional array of real numbers: (1) the subjects who participated in the experiment, (2) the emotion classes (sadness, anger, surprise, fear, frustration, and amusement) and (3) extracted features of data signal types (minimum, maximum, mean, and variance of GSR, temperature, and heart rate).

Each slot of the array consists of one specific feature of a specific data signal type, belonging to one specific participant while s/he was experiencing one specific emotion. (e.g., a slot contains the mean of normalized skin temperature value of, say, participant number 1 while s/he was experiencing anger, while another slot, for example, contains the variance of normalized GSR value of participant number 5 while s/he was experiencing sadness).

As mentioned, four features were extracted for each data type and then three supervised learning algorithms were implemented that took these 12 features as input and interpreted them. Following subsections describe the algorithms implemented to find a pattern among these features.

4.2. *k*-nearest neighbor algorithm

k-nearest neighbor (KNN) algorithm [43] uses two data sets: (1) the training data set and (2) the test data set. The training data set contains instances of minimum, maximum, mean, and variance of GSR, skin temperature, and heart rate values, and the corresponding emotion class. The test data set is similar to the training data set.

In order to classify an instance of a test data into an emotion, KNN calculates the distance between the test data and each instance of training data set. For example, let an arbitrary instance x be described by the feature vector $\langle a_1(x), a_2(x), \dots, a_n(x) \rangle$, where $a_r(x)$ is the r th feature of instance x . The distance between instances x_i and x_j is defined as $d(x_i, x_j)$, where,

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2}. \quad (2)$$

The algorithm then finds the k closest training instances to the test instance. The emotion with the highest frequency among k emotions associated with these k training instances is the emotion mapped to the test data. In our study KNN was tested with leave-one-out cross validation.

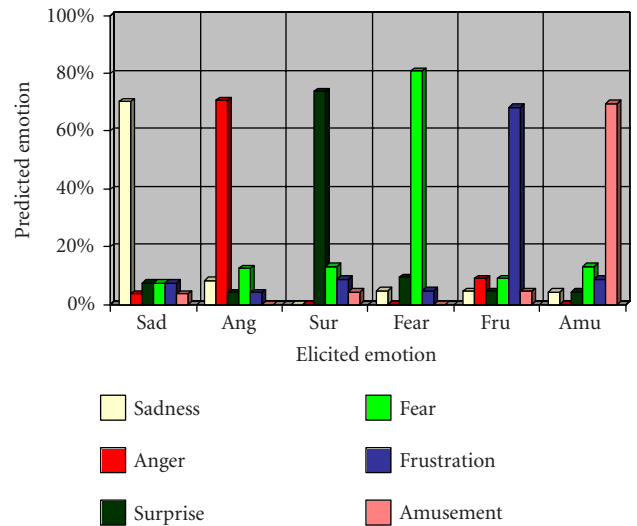


FIGURE 3: Emotion recognition graph with KNN algorithm.

Figure 3 shows the emotion recognition accuracy rates with KNN algorithm for each of the six emotions. KNN could classify sadness with 70.4%, anger with 70.8%, surprise with 73.9%, fear with 80.9%, frustration with 78.3%, and amusement with 69.6% accuracy.

4.3. Discriminant function analysis

The second algorithm was developed using *discriminant function analysis* (DFA) [44], which is a statistical method to classify data signals by using linear discriminant functions. DFA is used to find a set of linear combinations of the variables, whose values are as close as possible within groups and as far as possible between groups. These linear combinations are called discriminant functions. Thus, a discriminant function is a linear combination of the discriminating variables. In our implication of discriminant analysis, the groups are the emotion classes (sadness, anger, surprise, fear, frustration, and amusement) and the discriminant variables are the extracted features of data signals (minimum, maximum, mean, and variance of GSR, skin temperature, and heart rate).

Let x_i be the extracted feature of a specific data signal. The functions used to solve the coefficients are in the form of

$$f = u_0 + u_1 * x_1 + u_2 * x_2 + u_3 * x_3 + u_4 * x_4 + u_5 * x_5 + u_6 * x_6 + u_7 * x_7 + u_8 * x_8 + u_9 * x_9 + u_{10} * x_{10} + u_{11} * x_{11} + u_{12} * x_{12} + u_{13} * x_{13}. \quad (3)$$

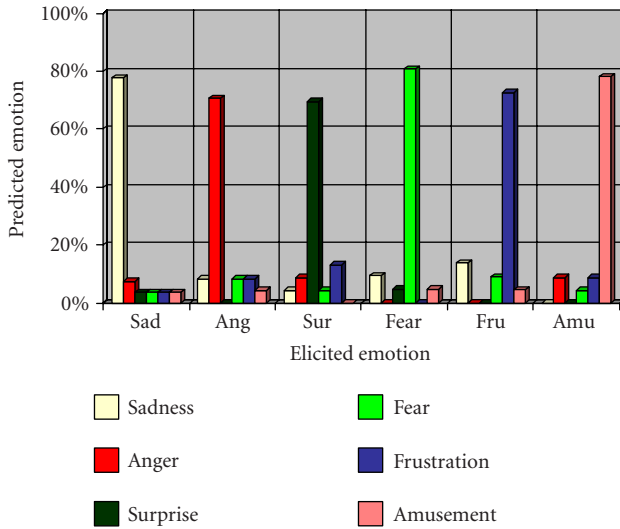


FIGURE 4: Emotion recognition graph with DFA algorithm.

The objective of DFA is to calculate the values of the coefficients $u_0 - u_{13}$ in order to obtain the linear combination. In order to solve for these coefficients, we applied the generalized eigenvalue decomposition to the between-group and within-group covariance matrices. The vectors gained as a result of this decomposition were used to derive coefficients of the discriminant functions. The coefficients of each function were derived in order to get a maximized difference between the outputs of group means and a minimized difference within the group means.

As can be seen in Figure 4, the DFA algorithm's recognition accuracy was 77.8% for sadness, 70.8% for anger, 69.6% for surprise, 80.9% for fear, 72.7% for frustration, and 78.3% for amusement.

4.4. Marquardt backpropagation algorithm

The third algorithm used was a derivation of a backpropagation algorithm with Marquardt-Levenberg modification called Marquardt backpropagation (MBP) algorithm [45]. In this technique, first the Jacobian matrix, which contains the first derivatives of the network errors with respect to the weights and biases, is computed. Then the gradient vector is computed as a product of the Jacobian matrix ($J(\underline{x})$) and the vector of errors ($\underline{e}(\underline{x})$), and the Hessian approximation is computed as the product of the Jacobian matrix ($J(\underline{x})$) and the transpose of the Jacobian matrix ($J^T(\underline{x})$) [45].

Then the Marquardt-Levenberg modification to the Gauss-Newton method is given by

$$\Delta \underline{x} = [J^T(\underline{x})J(\underline{x}) + \mu I]^{-1} J^T(\underline{x})\underline{e}(\underline{x}). \quad (4)$$

When μ is 0 or is equal to a small value, then this is the Gauss-Newton method that is using the Hessian approximation. When μ is a large value, then this equation is a gradient descent with a small step size $1/\mu$. The aim is to make the μ converge to 0 as fast as possible, and this is achieved by decreasing μ when there is a decrease in the error function and

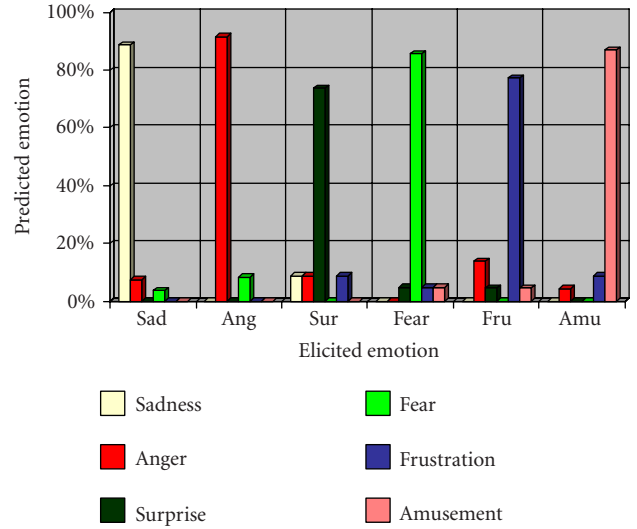


FIGURE 5: Emotion recognition graph with MBP algorithm.

increasing it when there is no decrease in the error function. The algorithm converges when gradient value reaches below a previously determined value.

As stated in Section 4.1, a total of 140 usable (i.e., without data loss) physiological records of GSR, temperature, and heart rate values were collected from the participants for six emotional states and 12 features (four for each data signal type) were extracted for each of the physiological record. As a result, a set of 140 data instances to train and test the network was obtained. The neural network was trained with MBP algorithm 140 times.

The recognition accuracy gained with MBP algorithm is shown in Figure 5, which was 88.9% for sadness, 91.7% for anger, 73.9% for surprise, 85.6% for fear, 77.3% for frustration, and finally 87.0% for amusement.

Overall, the DFA algorithm was better than the KNN algorithm for sadness, frustration, and amusement. On the other hand, KNN performed better than DFA for surprise. MBP algorithm performed better than both DFA and KNN for all emotion classes except for surprise and frustration.

5. DISCUSSION AND FUTURE WORK

5.1. Discussion

There are several studies that looked for the relationship between the physiological signals and emotions, as discussed in Section 3.1, and some of the results obtained were very promising. Our research adds to these studies by showing that emotions can be recognized from physiological signals via noninvasive wireless wearable computers, which means that the experiments can be carried out in real environments instead of laboratories. Real-life emotion recognition hence becomes closer to achieve.

Our multimodal experiment results showed that emotions can be distinguished from each other and that they can be categorized by collecting and interpreting physiological

signals of the participants. Different physiological signals were important in terms of recognizing different emotions. Our results show a relationship between galvanic skin response and frustration. When a participant was frustrated, her GSR increased. The difference in GSR values of the frustrated participants was higher than the differences in both heart rate and temperature values. Similarly, heart rate was more related to anger and fear. Heart rate value of a feared participant increased, whereas it decreased when the participant was angry.

Overall, three algorithms, KNN, DFA, and MBP, could categorize emotions with 72.3%, 75.0%, and 84.1% accuracy, respectively. In a previous study [42] where we interpreted the same data set without applying feature extraction, the overall recognition accuracy was 71% with KNN, 74% with DFA, and 83% with MBP. The results of our latest study showed that implementing a feature extraction technique slightly improved the performance of all three algorithms.

Recognition accuracy for some emotions was higher with the pattern recognition algorithms than the agreement of the subjects on the same emotions. For example, fear could be recognized with 80.9% accuracy by KNN and DFA and with 85.6% accuracy by MBP, although the subject agreement on fear was 65%. This might be understood from Feldman Barrett et al.'s study [46]: the results of this study indicate that individuals vary in their ability to identify the specific emotions they experience. For example, some individuals are able to indicate whether they are experiencing a negative or a positive emotion, but they cannot identify the specific emotion.

5.2. Applications and future work

Our results are promising in terms of creating a multimodal affective user interface that can recognize its user's affective state, adapt to the situation, and interact with her accordingly, within given context and application, as discussed in Section 2.1 and depicted in Figure 1.

We are specifically looking into *driving safety* where intelligent interfaces can be developed to minimize the negative effects of some emotions and states that have impact on one's driving such as anger, panic, sleepiness, and even road rage [47]. For example, when the system recognizes the driver is in a state of frustration, anger, or rage, the system could suggest the driver to change the music to a soothing one [47], or suggest a relaxation technique [48], depending on the driver's preferred style. Similarly, when the system recognizes that the driver is sleepy, it could suggest (maybe even insist) that she/he rolls down the window for awakening fresh air.

Our future work plans include designing and conducting experiments where *driving-related* emotions and states (frustration/anger, panic/fear, and sleepiness) are elicited from the participating drivers while they are driving in a driving simulator. During the experiment, physiological signals (GSR, temperature, and heart rate) of the participants will be measured with both BodyMedia SenseWear (see Figure 2) and ProComp+ (see Figure 6). At the same time, an ongoing video of each driver will be recorded for annotation and

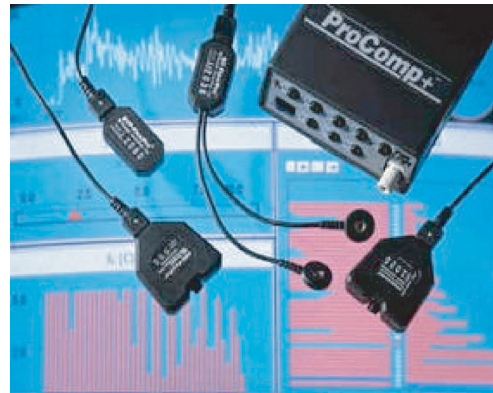


FIGURE 6: ProComp+.

facial expression recognition purposes. These measurements and recordings will be analyzed in order to find unique patterns mapping them to each elicited emotion.

Another application of interest is *training/learning* where emotions such as frustration and anxiety affect the learning capability of the users [49, 50, 51]. In an electronic learning environment, an intelligent affective interface could adjust the pace of training when it recognizes the frustration or boredom of the student, or it can provide encouragement when it recognizes the anxiety of the student.

One other application is *telemedicine* where the patients are being remotely monitored at their home by health-care providers [52]. For example, when the system accurately recognizes repetitive sadness (possibly indicating the reoccurrence of depression) of telemedicine patients, the interface could forward this affective information to the health-care providers in order for them to be better equipped and ready to respond to the patient.

Those three applications, driver safety, learning, and telemedicine, are the main ones that we are investigating, aiming at enhancing HCI via emotion recognition through multimodal sensing in these contexts. However using the generic overall paradigm of recognizing and responding to emotions in a user-dependent and context-dependent manner discussed in Section 2.2 and shown in Figure 1, we hope that other research efforts might be able to concentrate on different application areas of affective intelligent interfaces.

Some of our future work will focus on the difficulty to recognize emotions by interpreting a single (user mode), or modality. We are therefore planning on conducting multimodal studies on facial expression recognition and physiological signal recognition to guide the integration of the two modalities [16, 53, 54]. Other modalities, as shown in Figure 1, could include vocal intonation and natural language processing to obtain increased accuracy.

6. CONCLUSION

In this paper we documented the newly discovered role of affect in cognition and identified a variety of human-computer interaction context in which multimodal affective information could prove useful, if not necessary. We also

presented an application-independent framework for multimodal affective user interfaces, hoping that it will prove useful for building other research efforts aiming at enhancing human-computer interaction with restoring the role of affect, emotion, and personality in human natural communication.

Our current research focused on creating a multimodal affective user interface that will be used to recognize users' emotions in real-time and respond accordingly, in particular, recognizing emotion through the analysis of physiological signals from the autonomic nervous system (ANS). We presented an extensive survey of the literature in the form of a survey table (ordered chronologically) identifying various emotion-eliciting and signal-analysis methods for various emotions.

In order to continue to contribute to the research effort of finding a mapping between emotions and physiological signals, we conducted an experiment in which we elicited emotions (sadness, anger, fear, surprise, frustration, and amusement) using movie clips and mathematical problems while measuring certain physiological signals documented as associated with emotions (GSR, heart rate, and temperature) of our participants. After extracting minimum, maximum, mean, and variance of the collected data signals, three supervised learning algorithms were implemented to interpret these features. Overall, three algorithms, KNN, DFA, and MBP, could categorize emotions with 72.3%, 75.0%, and 84.1% accuracy, respectively.

Finally, we would like to emphasize that we are well aware that full-blown computer systems with multimodal affective intelligent user interfaces will only be applicable to real use in telemedicine, driving safety, and learning once the research is fully mature and results are completely reliable within restricted domains and appropriate subsets of emotions.

ACKNOWLEDGMENTS

The authors would like to thank Kaye Alvarez for her precious help in setting up the emotion elicitation experiment. They would also like to acknowledge that part of this research was funded by a grant from the US Army STRICOM. Part of this work was accomplished when C. L. Lisetti was at the University of Central Florida.

REFERENCES

- [1] G. Bower, "Mood and memory," *American Psychologist*, vol. 36, no. 2, pp. 129–148, 1981.
- [2] S. Kalyuga, P. Chandler, and J. Sweller, "Levels of expertise and user-adapted formats of instructional presentations: a cognitive load approach," in *Proceedings of the 6th International Conference on User Modeling (UM '97)*, A. Jameson, C. Paris, and C. Tasso, Eds., pp. 261–272, New York, NY, USA, 1997.
- [3] D. Derryberry and D. Tucker, "Neural mechanisms of emotion," *Journal of Consulting and Clinical Psychology*, vol. 60, no. 3, pp. 329–337, 1992.
- [4] R. Zajonc, "On the primacy of affect," *American Psychologist*, vol. 39, no. 2, pp. 117–123, 1984.
- [5] G. Linden, S. Hanks, and N. Lesh, "Interactive assessment of user preference models: the automated travel assistant," in *Proceedings of the 6th International Conference on User Modeling (UM '97)*, A. Jameson, C. Paris, and C. Tasso, Eds., pp. 67–78, New York, NY, USA, 1997.
- [6] A. Damasio, *Descartes' Error*, Avon Books, New York, NY, USA, 1994.
- [7] J. Ledoux, "Emotion and the amygdala," in *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*, J. P. Aggleton, Ed., pp. 339–351, Wiley-Liss, New York, NY, USA, 1992.
- [8] P. Paranagama, F. Burstein, and D. Arnott, "Modeling the personality of decision makers for active decision support," in *Proceedings of the 6th International Conference on User Modeling (UM '97)*, pp. 79–81, Sardinia, Italy, June 1997.
- [9] S. Bull, "See yourself write: a simple student model to make students think," in *Proceedings of the 6th International Conference User Modeling (UM '97)*, pp. 315–326, New York, NY, USA, 1997.
- [10] N. H. Frijda, *The Emotions*, Cambridge University Press, New York, NY, USA, 1986.
- [11] R. Birdwhistle, *Kinesics and Context: Essays on Body Motion and Communication*, University of Pennsylvania Press, Philadelphia, Pa, USA, 1970.
- [12] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1975.
- [13] N. Chovil, "Discourse-oriented facial displays in conversation," *Research on Language and Social Interaction*, vol. 25, pp. 163–194, 1991.
- [14] D. Goleman, *Emotional Intelligence*, Bantam Books, New York, NY, USA, 1995.
- [15] H. Leventhal and K. Sherer, "The relationship of emotion to cognition: a functional approach to a semantic controversy," *Cognition and Emotion*, vol. 1, no. 1, pp. 3–28, 1987.
- [16] C. L. Lisetti and F. Nasoz, "MAUI: a multimodal affective user interface," in *Proceedings of the ACM Multimedia International Conference*, Juan les Pins, France, December 2002.
- [17] N. Bianchi-Berthouze and C. L. Lisetti, "Modeling multimodal expression of user's affective subjective experience," *User Modeling and User-Adapted Interaction*, vol. 12, no. 1, pp. 49–84, 2002.
- [18] G. E. Schwartz, P. L. Fair, P. S. Greenberg, Friedman M. J., and G. L. Klerman, "Facial EMG in the assessment of emotion," *Psychophysiology*, vol. 11, no. 2, pp. 237, 1974.
- [19] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.
- [20] J. T. Lanzetta and S. P. Orr, "Excitatory strength of expressive faces: effects of happy and fear expressions and context on the extinction of a conditioned fear response," *Journal of Personality and Social Psychology*, vol. 50, no. 1, pp. 190–194, 1986.
- [21] S. R. Vrana, B. N. Cuthbert, and P. J. Lang, "Fear imagery and text processing," *Psychophysiology*, vol. 23, no. 3, pp. 247–253, 1986.
- [22] R. A. Wright, R. J. Contrada, and M. J. Patane, "Task difficulty, cardiovascular response and the magnitude of goal valence," *Journal of Personality and Social Psychology*, vol. 51, no. 4, pp. 837–843, 1986.
- [23] C. A. Smith, "Dimensions of appraisal and physiological response in emotion," *Journal of Personality and Social Psychology*, vol. 56, no. 3, pp. 339–353, 1989.
- [24] G. Stemmler, "The autonomic differentiation of emotions revisited: convergent and discriminant validation," *Psychophysiology*, vol. 26, no. 6, pp. 617–632, 1989.
- [25] R. W. Levenson, P. Ekman, and W. V. Friesen, "Voluntary facial action generates emotion-specific autonomic nervous system activity," *Psychophysiology*, vol. 27, pp. 363–384, 1990.

- [26] R. W. Levenson, P. Ekman, K. Heider, and W. V. Friesen, "Emotion and autonomic nervous system activity in the Minangkabau of west Sumatra," *Journal of Personality and Social Psychology*, vol. 62, no. 6, pp. 972–988, 1992.
- [27] S. Vrana, "The psychophysiology of disgust: differentiating negative emotional contexts with facial EMG," *Psychophysiology*, vol. 30, no. 3, pp. 279–286, 1993.
- [28] R. A. Wright and J. C. Dill, "Blood pressure responses and incentive appraisals as a function of perceived ability and objective task demand," *Psychophysiology*, vol. 30, no. 2, pp. 152–160, 1993.
- [29] A. Pecchinenda and C. Smith, "Affective significance of skin conductance activity during difficult problem-solving task," *Cognition and Emotion*, vol. 10, no. 5, pp. 481–503, 1996.
- [30] R. Sinha and O. Parsons, "Multivariate response patterning of fear and anger," *Cognition and Emotion*, vol. 10, no. 2, pp. 173–198, 1996.
- [31] C. Collet, E. Vernet-Maury, G. Delhomme, and A. Dittmar, "Autonomic nervous system response patterns specificity to basic emotions," *J. Auton. Nerv. Syst.*, vol. 62, no. 1-2, pp. 45–57, 1997.
- [32] J. J. Gross and R. W. Levenson, "Hiding feelings: the acute effects of inhibiting negative and positive emotion," *Journal of Abnormal Psychology*, vol. 106, no. 1, pp. 95–103, 1997.
- [33] W. Ark, D. C. Dryer, and D. J. Lu, "The emotion mouse," in *Human-Computer Interaction: Ergonomics and User Interfaces*, H. J. Bullinger and J. Ziegler, Eds., pp. 818–823, Lawrence Erlbaum, London, UK, 1999.
- [34] D. Palomba, M. Sarlo, A. Angrilli, and A. Mini, "Cardiac responses associated with affective processing of unpleasant film stimuli," *International Journal of Psychophysiology*, vol. 36, no. 1, pp. 45–57, 2000.
- [35] M. P. Tarvainen, A. S. Koistinen, M. Valkonen-Korhonen, J. Partanen, and P. A. Karjalainen, "Analysis of galvanic skin responses with principal components and clustering techniques," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 10, pp. 1071–1079, 2001.
- [36] M. E. Crosby, B. Auernheimer, C. Aschwanden, and C. Ikehara, "Physiological data feedback for application in distance education," in *Proceedings of Workshop on Perceptive User Interfaces (PUI '01)*, Orlando, FL, USA, November 2001.
- [37] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: analysis of affective physiological state," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.
- [38] J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard, "Frustrating the user on purpose: a step toward building an affective computer," *Interacting With Computers*, vol. 14, no. 2, pp. 93–118, 2002.
- [39] J. Healey, *Wearable and automotive systems for affect recognition from physiology*, Ph.D. thesis, Massachusetts Institute of Technology, Mass, USA, May 2000.
- [40] D. M. Clynes, *Sentics: The Touch of Emotions*, Anchor Press, New York, NY, USA, 1977.
- [41] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition and Emotion*, vol. 9, no. 1, pp. 87–108, 1995.
- [42] F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals for presence technologies," *Cognition, Technology, and Work*, vol. 6, no. 1, pp. 4–14, 2004.
- [43] T. M. Mitchell, *Machine Learning*, McGraw-Hill, New York, NY, USA, 1997.
- [44] A. A. Nicol and P. M. Pexman, *Presenting Your Findings: A Practical Guide for Creating Tables*, American Psychological Association, Wash, DC, USA, 1999.
- [45] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.
- [46] L. Feldman Barrett, J. J. Gross, T. Conner Christensen, and M. Benvenuto, "Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation," *Cognition and Emotion*, vol. 15, no. 6, pp. 713–724, 2001.
- [47] L. James, *Road Rage and Aggressive Driving: Steering Clear of Highway Warfare*, Prometheus Books, Amherst, NY, USA, 2000.
- [48] J. Larson and C. Rodriguez, *Road Rage to Road-Wise*, Tom Doherty Associates, New York, NY, USA, 1999.
- [49] V. E. Lewis and R. N. Williams, "Mood-congruent vs. mood-state-dependent learning: implications for a view of emotion," *The Journal of Social Behavior and Personality*, vol. 4, no. 2, pp. 157–171, 1989, Special Issue on Mood and Memory: Theory, Research, and Applications.
- [50] J. J. Martocchio, "Effects of conceptions of ability on anxiety, self-efficacy, and learning in training," *Journal of Applied Psychology*, vol. 79, no. 6, pp. 819–825, 1994.
- [51] P. Warr and D. Bunce, "Trainee characteristics and the outcomes of open learning," *Personnel Psychology*, vol. 48, no. 2, pp. 347–375, 1995.
- [52] C. L. Lisetti, F. Nasoz, C. LeRouge, O. Ozyer, and K. Alvarez, "Developing multimodal intelligent affective interfaces for tele-home health care," *International Journal of Human-Computer Studies*, vol. 59, no. 1-2, pp. 245–255, 2003, Special Issue on Applications of Affective Computing in Human-Computer Interaction.
- [53] C. L. Lisetti and D. J. Schiano, "Facial expression recognition: where human-computer interaction, artificial intelligence and cognitive science intersect," *Pragmatics and Cognition*, vol. 8, no. 1, pp. 185–235, 2000.
- [54] C. L. Lisetti and D. Rumelhart, "Facial expression recognition using a neural network," in *Proceedings of the 11th International Florida Artificial Intelligence Research Society Conference (FLAIRS '98)*, pp. 328–332, AAAI Press, Menlo Park, Calif, USA, 1998.

Christine Lætitia Lisetti is a Professor at the Institut Eurecom in the Multimedia Communications Department, Sophia-Antipolis, France. Previously, she lived in the United States where she was an Assistant Professor in the Department of Computer Science at the University of Central Florida. From 1996 to 1998, she was a Postdoctoral Fellow at Stanford University in the Department of Psychology and the Department of Computer Science. She received a Ph.D. in computer science in 1995, from Florida International University. She has won multiple awards including a National Institute of Health Individual Research Service Award, the AAAI Nils Nilsson Award for Integrating AI Technologies, and the University of Central Florida COECS Distinguished Research Lecturer Award. Her research involves the use of artificial intelligence techniques in knowledge representation and machine learning to model affective knowledge computationally. She has been granted support from federally funded agencies such as the National Institute of Health, the Office of Naval Research, and US Army STRICOM as well as from industries such as Interval Research Corporation and Intel Corporation. She is a Member of IEEE, ACM, and AAAI, is regularly invited to serve on program committees of international conferences, and has cochaired several international workshops on affective computing.



Fatma Nasoz is a Ph.D. candidate in the Computer Science Department of the University of Central Florida, Orlando, since August 2001. She earned her M.S. degree in computer science from the University of Central Florida and her B.S. degree in computer engineering from Bogazici University in Turkey, in 2003 and 2000, respectively. She was awarded the Center for Advanced Transportation System Simulation (CATSS)



Scholarship in 2002 to model emotions of drivers for increased safety. Her research area is affective computing and she specifically focuses on creating adaptive intelligent user interfaces with emotion recognition abilities that adapt and respond to the user's current emotional state by also modeling their preferences and personality. Her research involves eliciting emotions in a variety of contexts, using noninvasive wearable computers to collect the participants' physiological signals, mapping these signals to affective states, and building adaptive interfaces to adapt appropriately to the current sensed data and context. She is a Member of the American Association for Artificial Intelligence and of the Association for Computing Machinery, and she has published multiple scientific articles.

Special Issue on Image Perception

Call for Papers

Perception is a complex process that involves brain activities at different levels. The availability of models for the representation and interpretation of the sensory information opens up new research avenues that cut across neuroscience, imaging, information engineering, and modern robotics.

The goal of the multidisciplinary field of perceptual signal processing is to identify the features of the stimuli that determine their “perception,” namely “a single unified awareness derived from sensory processes while a stimulus is present,” and to derive associated computational models that can be generalized.

In the case of vision, the stimuli go through a complex analysis chain along the so-called “visual pathway,” starting with the encoding by the photoreceptors in the retina (low-level processing) and ending with cognitive mechanisms (high-level processes) that depend on the task being performed.

Accordingly, low-level models are concerned with image “representation” and aim at emulating the way the visual stimulus is encoded by the early stages of the visual system as well as capturing the varying sensitivity to the features of the input stimuli; high-level models are related to image “interpretation” and allow to predict the performance of a human observer in a given predefined task.

A global model, accounting for both such bottom-up and top-down approaches, would enable the automatic interpretation of the visual stimuli based on both their low-level features and their semantic content.

Among the main image processing fields that would take advantage of such models are feature extraction, content-based image description and retrieval, model-based coding, and the emergent domain of medical image perception.

The goal of this special issue is to provide original contributions in the field of image perception and modeling.

Topics of interest include (but are not limited to):

- Perceptually plausible mathematical bases for the representation of visual information (static and dynamic)
- Modeling nonlinear processes (masking, facilitation) and their exploitation in the imaging field (compression, enhancement, and restoration)

- Beyond early vision: investigating the pertinence and potential of cognitive models (feature extraction, image quality)
- Stochastic properties of complex natural scenes (static, dynamic, colored) and their relationships with perception
- Perception-based models for natural (static and dynamic) textures. Theoretical formulation and psychophysical validation
- Applications in the field of biomedical imaging (medical image perception)

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	December 1, 2005
Acceptance Notification	April 1, 2006
Final Manuscript Due	July 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Gloria Menegaz, Department of Information Engineering, University of Siena, Siena, Italy; menegaz@dii.unisi.it

Guang-Zhong Yang, Department of Computing, Engineering Imperial College London, London, UK; gzy@doc.ic.ac.uk

Maria Concetta Morrone, Università Vita-Salute San Raffaele, Milano, Italy; concetta@in.cnr.it

Stefan Winkler, Genista Corporation, Montreux, Switzerland; stefan.winkler@genista.com

Javier Portilla, Department of Computer Science and Artificial Intelligence (DECSAI), Universidad de Granada, Granada, Spain; javier@decsai.ugr.es

Special Issue on Music Information Retrieval Based on Signal Processing

Call for Papers

The main focus of this special issue is on the application of digital signal processing techniques for music information retrieval (MIR). MIR is an emerging and exciting area of research that seeks to solve a wide variety of problems dealing with preserving, analyzing, indexing, searching, and accessing large collections of digitized music. There are also strong interests in this field of research from music libraries and the recording industry as they move towards digital music distribution. The demands from the general public for easy access to these music libraries challenge researchers to create tools and algorithms that are robust, small, and fast.

Music is represented in either encoded audio waveforms (CD audio, MP3, etc.) or symbolic forms (musical score, MIDI, etc.). Audio representations, in particular, require robust signal processing techniques for many applications of MIR since meaningful descriptions need to be extracted from audio signals in which sounds from multiple instruments and vocals are often mixed together. Researchers in MIR are therefore developing a wide range of new methods based on statistical pattern recognition, classification, and machine learning techniques such as the Hidden Markov Model (HMM), maximum likelihood estimation, and Bayes estimation as well as digital signal processing techniques such as Fourier and Wavelet transforms, adaptive filtering, and source-filter models. New music interface and query systems leveraging such methods are also important for end users to benefit from MIR research.

Although research contributions on MIR have been published at various conferences in 1990s, the members of the MIR research community meet annually at the International Conference on Music Information Retrieval (ISMIR) since 2000.

Topics of interest include (but are not limited to):

- Automatic summarization (succinct representation of music)
- Automatic transcription (audio to symbolic format conversion)
- Music annotation (semantic analysis)
- Music fingerprinting (unique identification of music)
- Music interface
- Music similarity metrics (comparison)

- Music understanding
- Musical feature extraction
- Musical styles and genres
- Optical music score recognition (image to symbolic format conversion)
- Performer/artist identification
- Query systems
- Timbre/instrument recognition

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	December 1, 2005
Acceptance Notification	April 1, 2006
Final Manuscript Due	July 1, 2006
Publication Date	3rd Quarter, 2006

GUEST EDITORS:

Ichiro Fujinaga, McGill University, Montreal, QC, Canada, H3A 2T5; ich@music.mcgill.ca

Masataka Goto, National Institute of Advanced Industrial Science and Technology, Japan; m.goto@aist.go.jp

George Tzanetakis, University of Victoria, Victoria, BC, Canada, V8P 5C2; gtzan@cs.uvic.ca

Special Issue on Visual Sensor Networks

Call for Papers

Research into the design, development, and deployment of networked sensing devices for high-level inference and surveillance of the physical environment has grown tremendously in the last few years.

This trend has been motivated, in part, by recent technological advances in electronics, communication networking, and signal processing.

Sensor networks are commonly comprised of lightweight distributed sensor nodes such as low-cost video cameras. There is inherent redundancy in the number of nodes deployed and corresponding networking topology. Operation of the network requires autonomous peer-based collaboration amongst the nodes and intermediate data-centric processing amongst local sensors. The intermediate processing known as in-network processing is application-specific. Often, the sensors are untethered so that they must communicate wirelessly and be battery-powered. Initial focus was placed on the design of sensor networks in which scalar phenomena such as temperature, pressure, or humidity were measured.

It is envisioned that much societal use of sensor networks will also be based on employing content-rich vision-based sensors. The volume of data collected as well as the sophistication of the necessary in-network stream content processing provide a diverse set of challenges in comparison with generic scalar sensor network research.

Applications that will be facilitated through the development of visual sensor networking technology include automatic tracking, monitoring and signaling of intruders within a physical area, assisted living for the elderly or physically disabled, environmental monitoring, and command and control of unmanned vehicles.

Many current video-based surveillance systems have centralized architectures that collect all visual data at a central location for storage or real-time interpretation by a human operator. The use of distributed processing for automated event detection would significantly alleviate mundane or time-critical activities performed by human operators, and provide better network scalability. Thus, it is expected that video surveillance solutions of the future will successfully utilize visual sensor networking technologies.

Given that the field of visual sensor networking is still in its infancy, it is critical that researchers from the diverse disciplines including signal processing, communications, and electronics address the many challenges of this emerging field. This special issue aims to bring together a diverse set of research results that are essential for the development of robust and practical visual sensor networks.

Topics of interest include (but are not limited to):

- Sensor network architectures for high-bandwidth vision applications
- Communication networking protocols specific to visual sensor networks
- Scalability, reliability, and modeling issues of visual sensor networks
- Distributed computer vision and aggregation algorithms for low-power surveillance applications
- Fusion of information from visual and other modalities of sensors
- Storage and retrieval of sensor information
- Security issues for visual sensor networks
- Visual sensor network testbed research
- Novel applications of visual sensor networks
- Design of visual sensors

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	December 1, 2005
Acceptance Notification	April 1, 2006
Final Manuscript Due	July 1, 2006
Publication Date	3rd Quarter, 2006



GUEST EDITORS:

Deepa Kundur, Department of Electrical Engineering,
Texas A&M University, College Station, Texas, USA;
deepa@ee.tamu.edu

Ching-Yung Lin, Distributed Computing Department,
IBM TJ Watson Research Center, New York, USA;
chingyung@us.ibm.com

Chun Shien Lu, Institute of Information Science, Academia
Sinica, Taipei, Taiwan; lcs@iis.sinica.edu.tw

Special Issue on Multirate Systems and Applications

Call for Papers

Filter banks for the application of subband coding of speech were introduced in the 1970s. Since then, filter banks and multirate systems have been studied extensively. There has been great success in applying multirate systems to many applications. The most notable of these applications include subband coding for audio, image, and video, signal analysis and representation using wavelets, subband denoising, and so forth. Different applications also call for different filter bank designs and the topic of designing one-dimensional and multidimensional filter banks for specific applications has been of great interest.

Recently there has been growing interest in applying multirate theories to the area of communication systems such as, transmultiplexers, filter bank transceivers, blind deconvolution, and precoded systems. There are strikingly many dualities and similarities between multirate systems and multicarrier communication systems. Many problems in multicarrier transmission can be solved by extending results from multirate systems and filter banks. This exciting research area is one that is of increasing importance.

The aim of this special issue is to bring forward recent developments on filter banks and the ever-expanding area of applications of multirate systems.

Topics of interest include (but are not limited to):

- Multirate signal processing for communications
- Filter bank transceivers
- One-dimensional and multidimensional filter bank designs for specific applications
- Denoising
- Adaptive filtering
- Subband coding
- Audio, image, and video compression
- Signal analysis and representation
- Feature extraction and classification
- Other applications

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their

complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	January 1, 2006
Acceptance Notification	May 1, 2006
Final Manuscript Due	August 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Yuan-Pei Lin, Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan; ypl@mail.nctu.edu.tw

See-May Phoong, Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan; smp@cc.ee.ntu.edu.tw

Ivan Selesnick, Department of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201, USA; selesi@poly.edu

Soontorn Oraintara, Department of Electrical Engineering, The University of Texas at Arlington, Arlington, TX 76010, USA; oraintar@uta.edu

Gerald Schuller, Fraunhofer Institute for Digital Media Technology (IDMT), Langewiesener Strasse 22, 98693 Ilmenau, Germany; shl@idmt.fraunhofer.de

Special Issue on Multisensor Processing for Signal Extraction and Applications

Call for Papers

Source signal extraction from heterogeneous measurements has a wide range of applications in many scientific and technological fields, for example, telecommunications, speech and acoustic signal processing, and biomedical pattern analysis. Multiple signal reception through multisensor systems has become an effective means for signal extraction due to its superior performance over the monosensor mode. Despite the rapid progress made in multisensor-based techniques in the past few decades, they continue to evolve as key technologies in modern wireless communications and biomedical signal processing. This has led to an increased focus by the signal processing community on the advanced multisensor-based techniques which can offer robust high-quality signal extraction under realistic assumptions and with minimal computational complexity. However, many challenging tasks remain unresolved and merit further rigorous studies. Major efforts in developing advanced multisensor-based techniques may include high-quality signal extraction, realistic theoretical modeling of real-world problems, algorithm complexity reduction, and efficient real-time implementation.

The purpose of this special issue aims to present state-of-the-art multisensor signal extraction techniques and applications. Contributions in theoretical study, performance analysis, complexity reduction, computational advances, and real-world applications are strongly encouraged.

Topics of interest include (but are not limited to):

- Multiantenna processing for radio signal extraction
- Multimicrophone speech recognition and enhancement
- Multisensor radar, sonar, navigation, and biomedical signal processing
- Blind techniques for multisensor signal extraction
- Computational advances in multisensor processing

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their

complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	January 1, 2006
Acceptance Notification	May 1, 2006
Final Manuscript Due	August 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Chong-Yung Chi, National Tsing Hua University, Taiwan; cychi@ee.nthu.edu.tw

Ta-Sung Lee, National Chiao Tung University, Taiwan; tslee@cc.nctu.edu.tw

Zhi-Quan Luo, University of Minnesota, USA; luozq@ece.umn.edu

Kung Yao, University of California, Los Angeles, USA; yao@ee.ucla.edu

Yue Wang, Virginia Polytechnic Institute and State University, USA; yuewang@vt.edu

Special Issue on

Search and Retrieval of 3D Content and Associated Knowledge Extraction and Propagation

Call for Papers

With the general availability of 3D digitizers, scanners, and the technology innovation in 3D graphics and computational equipment, large collections of 3D graphical models can be readily built up for different applications (e.g., in CAD/CAM, games design, computer animations, manufacturing and molecular biology). For such large databases, the method whereby 3D models are sought merits careful consideration. The simple and efficient query-by-content approach has, up to now, been almost universally adopted in the literature. Any such method, however, must first deal with the proper positioning of the 3D models. The two prevalent-in-the-literature methods for the solution to this problem seek either

- Pose Normalization: Models are first placed into a canonical coordinate frame (normalizing for translation, scaling, and rotation). Then, the best measure of similarity is found by comparing the extracted feature vectors, or
- Descriptor Invariance: Models are described in a transformation invariant manner, so that any transformation of a model will be described in the same way, and the best measure of similarity is obtained at any transformation.

The existing 3D retrieval systems allow the user to perform queries by example. The queried 3D model is then processed, low-level geometrical features are extracted, and similar objects are retrieved from a local database. A shortcoming of the methods that have been proposed so far regarding the 3D object retrieval, is that neither is the semantic information (high-level features) attached to the (low-level) geometric features of the 3D content, nor are the personalization options taken into account, which would significantly improve the retrieved results. Moreover, few systems exist so far to take into account *annotation* and *relevance feedback* techniques, which are very popular among the corresponding content-based image retrieval systems (CBIR).

Most existing CBIR systems using knowledge either annotate all the objects in the database (full annotation) or

annotate a subset of the database manually selected (partial annotation). As the database becomes larger, full annotation is increasingly difficult because of the manual effort needed. Partial annotation is relatively affordable and trims down the heavy manual labor. Once the database is partially annotated, traditional image analysis methods are used to derive semantics of the objects not yet annotated. However, it is not clear “how much” annotation is sufficient for a specific database and what the best subset of objects to annotate is. In other words how the knowledge *will be propagated*. Such techniques have not been presented so far regarding the 3D case.

Relevance feedback was first proposed as an interactive tool in text-based retrieval. Since then it has been proven to be a powerful tool and has become a major focus of research in the area of content-based search and retrieval. In the traditional computer centric approaches, which have been proposed so far, the “best” representations and weights are fixed and they cannot effectively model high-level concepts and user’s perception subjectivity. In order to overcome these limitations of the computer centric approach, techniques based on *relevant feedback*, in which the human and computer interact to refine high-level queries to representations based on low-level features, should be developed.

The aim of this special issue is to focus on recent developments in this expanding research area. The special issue will focus on novel approaches in 3D object retrieval, transforms and methods for efficient geometric feature extraction, annotation and relevance feedback techniques, knowledge propagation (e.g., using Bayesian networks), and their combinations so as to produce a single, powerful, and dominant solution.

Topics of interest include (but are not limited to):

- 3D content-based search and retrieval methods (volume/surface-based)
- Partial matching of 3D objects
- Rotation invariant feature extraction methods for 3D objects

- Graph-based and topology-based methods
- 3D data and knowledge representation
- Semantic and knowledge propagation over heterogeneous metadata types
- Annotation and relevance feedback techniques for 3D objects

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	February 1, 2006
Acceptance Notification	June 1, 2006
Final Manuscript Due	September 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Tsuan Chen, Carnegie Mellon University, Pittsburgh, PA 15213, USA; tsuhan@cmu.edu

Ming Ouhyoung, National Taiwan University, Taipei 106, Taiwan; ming@csie.ntu.edu.tw

Petros Daras, Informatics and Telematics Institute, Centre for Research and Technology Hellas, 57001 Thessaloniki, Greece; daras@iti.gr

Special Issue on Robust Speech Recognition

Call for Papers

Robustness can be defined as the ability of a system to maintain performance or degrade gracefully when exposed to conditions not well represented in the data used to develop the system. In automatic speech recognition (ASR), systems must be robust to many forms of signal degradation, including speaker characteristics (e.g., dialect and accent), ambient environment (e.g., cellular telephony), transmission channel (e.g., voice over IP), and language (e.g., new words, dialect switching). Robust ASR systems, which have been under development for the past 35 years, have made great progress over the years closing the gap between performance on pristine research tasks and noisy operational data.

However, in recent years, demand is emerging for a new class of systems that tolerate extreme and unpredictable variations in operating conditions. For example, in a cellular telephony environment, there are many nonstationary forms of noise (e.g., multiple speakers) and significant variations in microphone type, position, and placement. Harsh ambient conditions typical in automotive and mobile applications pose similar challenges. Development of systems in a language or dialect for which there is limited or no training data in a target language has become a critical issue for a new generation of voice mining applications. The existence of multiple conditions in a single stream, a situation common to broadcast news applications, and that often involves unpredictable changes in speaker, topic, dialect, or language, is another form of robustness that has gained attention in recent years.

Statistical methods have dominated the field since the early 1980s. Such systems tend to excel at learning the characteristics of large databases that represent good models of the operational conditions and do not generalize well to new environments.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Channel and microphone normalization
- Stationary and nonstationary noise modeling, compensation, and/or rejection
- Localization and separation of sound sources (including speaker segregation)

- Signal processing and feature extraction for applications involving hands-free microphones
- Noise robust speech modeling
- Adaptive training techniques
- Rapid adaptation and learning
- Integration of confidence scoring, metadata, and other alternative information sources
- Audio-visual fusion
- Assessment relative to human performance
- Machine learning algorithms for robustness
- Transmission robustness
- Pronunciation modeling

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	February 1, 2006
Acceptance Notification	June 1, 2006
Final Manuscript Due	September 1, 2006
Publication Date	4th Quarter, 2006

GUEST EDITORS:

Herve Bourlard, IDIAP Research Institute, Swiss Federal Institute of Technology at Lausanne (EPFL), 1920 Martigny, Switzerland; herve.bourlard@idiap.ch

Mark Gales, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK; mjfg@eng.cam.ac.uk

Maurizio Omologo, ITC-IRST, 38050 Trento, Italy; omologo@itc.it

S. Parthasarathy, AT&T Labs - Research, NJ 07748, USA; sps@research.att.com

Joe Picone, Department of Electrical and Computer Engineering, Mississippi State University, MS 39762-9571, USA; picone@cavs.msstate.edu

Special Issue on Signal Processing Technologies for Ambient Intelligence in Home-Care Applications

Call for Papers

The possibility of allowing elderly people with different kinds of disabilities to conduct a normal life at home and achieve a more effective inclusion in the society is attracting more and more interest from both industrial and governmental bodies (hospitals, healthcare institutions, and social institutions). Ambient intelligence technologies, supported by adequate networks of sensors and actuators, as well as by suitable processing and communication technologies, could enable such an ambitious objective.

Recent researches demonstrated the possibility of providing constant monitoring of environmental and biomedical parameters, and the possibility to autonomously originate alarms, provide primary healthcare services, activate emergency calls, and rescue operations through distributed assistance infrastructures. Nevertheless, several technological challenges are still connected with these applications, ranging from the development of enabling technologies (hardware and software), to the standardization of interfaces, the development of intuitive and ergonomic human-machine interfaces, and the integration of complex systems in a highly multidisciplinary environment.

The objective of this special issue is to collect the most significant contributions and visions coming from both academic and applied research bodies working in this stimulating research field. This is a highly interdisciplinary field comprising many areas, such as signal processing, image processing, computer vision, sensor fusion, machine learning, pattern recognition, biomedical signal processing, multimedia, human-computer interfaces, and networking.

The focus will be primarily on the presentation of original and unpublished works dealing with ambient intelligence and domotic technologies that can enable the provision of advanced homecare services.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Video-based monitoring of domestic environments and users
- Continuous versus event-driven monitoring
- Distributed information processing

- Data fusion techniques for event association and automatic alarm generation
- Modeling, detection, and learning of user habits for automatic detection of anomalous behaviors
- Integration of biomedical and behavioral data
- Posture and gait recognition and classification
- Interactive multimedia communications for remote assistance
- Content-based encoding of medical and behavioral data
- Networking support for remote healthcare
- Intelligent/natural man-machine interaction, personalization, and user acceptance

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscripts through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	March 1, 2006
Acceptance Notification	July 1, 2006
Final Manuscript Due	October 1, 2006
Publication Date	1st Quarter, 2007

GUEST EDITORS:

Francesco G. B. De Natale, Department of Information and Communication Technology, University of Trento, Via Sommarive 14, 38050 Trento, Italy; denatale@ing.unitn.it

Aggelos K. Katsaggelos, Department of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3118, USA; aggk@ece.northwestern.edu

Oscar Mayora, Create-Net Association, Via Solteri 38, 38100 Trento, Italy; oscar.mayora@create-net.it

Ying Wu, Department of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3118, USA; yingwu@ece.northwestern.edu

Special Issue on Spatial Sound and Virtual Acoustics

Call for Papers

Spatial sound reproduction has become widespread in the form of multichannel audio, particularly through home theater systems. Reproduction systems from binaural (by headphones) to hundreds of loudspeaker channels (such as wave field synthesis) are entering practical use. The application potential of spatial sound is much wider than multichannel sound, however, and research in the field is active. Spatial sound covers for example the capturing, analysis, coding, synthesis, reproduction, and perception of spatial aspects in audio and acoustics.

In addition to the topics mentioned above, research in virtual acoustics broadens the field. Virtual acoustics includes techniques and methods to create realistic percepts of sound sources and acoustic environments that do not exist naturally but are rendered by advanced reproduction systems using loudspeakers or headphones. Augmented acoustic and audio environments contain both real and virtual acoustic components.

Spatial sound and virtual acoustics are among the major research and application areas in audio signal processing. Topics of active study range from new basic research ideas to improvement of existing applications. Understanding of spatial sound perception by humans is also an important area, in fact a prerequisite to advanced forms of spatial sound and virtual acoustics technology.

This special issue will focus on recent developments in this key research area. Topics of interest include (but are not limited to):

- Multichannel reproduction
- Wave field synthesis
- Binaural reproduction
- Format conversion and enhancement of spatial sound
- Spatial sound recording
- Analysis, synthesis, and coding of spatial sound
- Spatial sound perception and auditory modeling
- Simulation and modeling of room acoustics
- Auralization techniques
- Beamforming and sound source localization
- Acoustic and auditory scene analysis
- Augmented reality audio

- Virtual acoustics (sound environments and sources)
- Intelligent audio environments
- Loudspeaker-room interaction and equalization
- Applications

Authors should follow the EURASIP JASP manuscript format described at <http://www.hindawi.com/journals/asp/>. Prospective authors should submit an electronic copy of their complete manuscript through the EURASIP JASP manuscript tracking system at <http://www.mstracking.com/asp/>, according to the following timetable:

Manuscript Due	May 1, 2006
Acceptance Notification	September 1, 2006
Final Manuscript Due	December 1, 2006
Publication Date	1st Quarter, 2007

GUEST EDITORS:

Ville Pulkki, Helsinki University of Technology, Espoo, Finland; ville@acoustics.hut.fi

Christof Faller, EPFL, Lausanne, Switzerland; christof.faller@epfl.ch

Aki Harma, Philips Research Labs, Eindhoven, The Netherlands; aki.harma@philips.com

Tapio Lokki, Helsinki University of Technology, Espoo, Finland; ktlokki@cc.hut.fi

Werner de Bruijn, Philips Research Labs, Eindhoven, The Netherlands; werner.de.bruijn@philips.com



NEWS RELEASE

Nominations Invited for the Institute of Acoustics

2006 A B Wood Medal

The Institute of Acoustics, the UK's leading professional body for those working in acoustics, noise and vibration, is inviting nominations for its prestigious A B Wood Medal for the year 2006.

The A B Wood Medal and prize is presented to an individual, usually under the age of 35, for distinguished contributions to the application of underwater acoustics. The award is made annually, in even numbered years to a person from Europe and in odd numbered years to someone from the USA/Canada. The 2005 Medal was awarded to Dr A Thode from the USA for his innovative, interdisciplinary research in ocean and marine mammal acoustics.

Nominations should consist of the candidate's CV, clearly identifying peer reviewed publications, and a letter of endorsement from the nominator identifying the contribution the candidate has made to underwater acoustics. In addition, there should be a further reference from a person involved in underwater acoustics and not closely associated with the candidate. Nominees should be citizens of a European Union country for the 2006 Medal. Nominations should be marked confidential and addressed to the President of the Institute of Acoustics at 77A St Peter's Street, St. Albans, Herts, AL1 3BN. The deadline for receipt of nominations is **15 October 2005**.

Dr Tony Jones, President of the Institute of Acoustics, comments, "A B Wood was a modest man who took delight in helping his younger colleagues. It is therefore appropriate that this prestigious award should be designed to recognise the contributions of young acousticians."

Further information and an nomination form can be found on the Institute's website at www.ioa.org.uk.

A B Wood

Albert Beaumont Wood was born in Yorkshire in 1890 and graduated from Manchester University in 1912. He became one of the first two research scientists at the Admiralty to

work on antisubmarine defence. He designed the first directional hydrophone and was well known for the many contributions he made to the science of underwater acoustics and for the help he gave to younger colleagues. The medal was instituted after his death by his many friends on both sides of the Atlantic and was administered by the Institute of Physics until the formation of the Institute of Acoustics in 1974.

PRESS CONTACT

Judy Edrich

Publicity & Information Manager, Institute of Acoustics
Tel: 01727 848195; E-mail: judy.edrich@ioa.org.uk

EDITORS NOTES

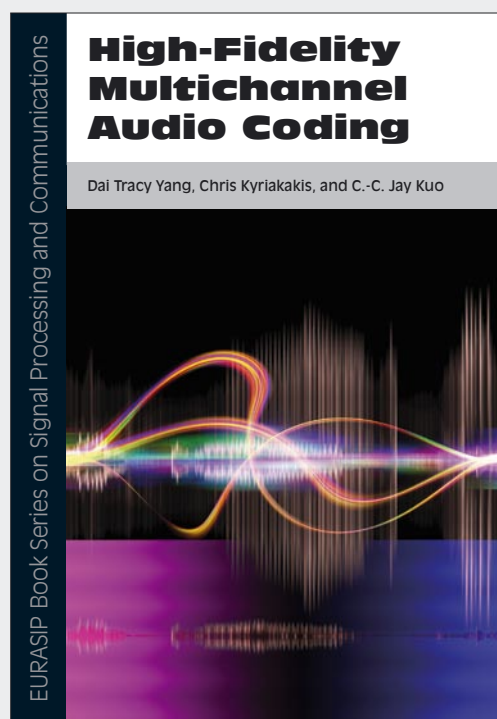
The Institute of Acoustics is the UK's professional body for those working in acoustics, noise and vibration. It was formed in 1974 from the amalgamation of the Acoustics Group of the Institute of Physics and the British Acoustical Society (a daughter society of the Institution of Mechanical Engineers). The Institute of Acoustics is a nominated body of the Engineering Council, offering registration at Chartered and Incorporated Engineer levels.

The Institute has some 2500 members from a rich diversity of backgrounds, with engineers, scientists, educators, lawyers, occupational hygienists, architects and environmental health officers among their number. This multidisciplinary culture provides a productive environment for cross-fertilisation of ideas and initiatives. The range of interests of members within the world of acoustics is equally wide, embracing such aspects as aerodynamics, architectural acoustics, building acoustics, electroacoustics, engineering dynamics, noise and vibration, hearing, speech, underwater acoustics, together with a variety of environmental aspects. The lively nature of the Institute is demonstrated by the breadth of its learned society programmes.

For more information please visit our site at www.ioa.org.uk.

HIGH-FIDELITY MULTICHANNEL AUDIO CODING

Dai Tracy Yang, Chris Kyriakakis, and C.-C. Jay Kuo



This invaluable monograph addresses the specific needs of audio-engineering students and researchers who are either learning about the topic or using it as a reference book on multichannel audio compression. This book covers a wide range of knowledge on perceptual audio coding, from basic digital signal processing and data compression techniques to advanced audio coding standards and innovative coding tools. It is the only book available on the market that solely focuses on the principles of high-quality audio codec design for multichannel sound sources.

This book includes three parts. The first part covers the basic topics on audio compression, such as quantization, entropy coding, psychoacoustic model, and sound quality assessment. The second part of the book highlights the current most prevalent low-bit-rate high-performance audio coding standards—MPEG-4 audio. More space is given to the audio standards that are capable of supporting multichannel signals, that is, MPEG advanced audio coding (AAC), including the original MPEG-2 AAC technology, additional MPEG-4 toolsets, and the most recent aacPlus standard. The third part of this book introduces several innovative multichannel audio coding tools, which have been demonstrated to further improve the coding performance and expand the available functionalities of MPEG AAC, and is more suitable for graduate students and researchers in the advanced level.

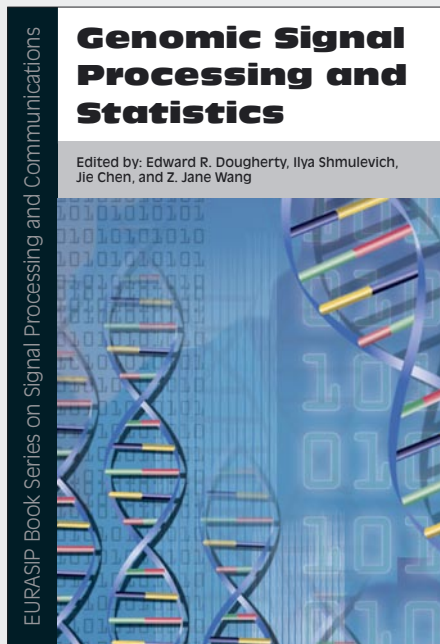
Dai Tracy Yang is currently Postdoctoral Research Fellow, Chris Kyriakakis is Associated Professor, and C.-C. Jay Kuo is Professor, all affiliated with the Integrated Media Systems Center (IMSC) at the University of Southern California.

For more information and online orders please visit: <http://www.hindawi.com/books/spc/volume-1/>
For any inquiries on how to order this title please contact books.orders@hindawi.com

The EURASIP Book Series on Signal Processing and Communications publishes monographs, edited volumes, and textbooks on Signal Processing and Communications. For more information about the series please visit: <http://hindawi.com/books/spc/about.html>

GENOMIC SIGNAL PROCESSING AND STATISTICS

Edited by: Edward R. Dougherty, Ilya Shmulevich, Jie Chen, and Z. Jane Wang



Recent advances in genomic studies have stimulated synergetic research and development in many cross-disciplinary areas. Genomic data, especially the recent large-scale microarray gene expression data, represents enormous challenges for signal processing and statistics in processing these vast data to reveal the complex biological functionality. This perspective naturally leads to a new field, genomic signal processing (GSP), which studies the processing of genomic signals by integrating the theory of signal processing and statistics. Written by an international, interdisciplinary team of authors, this invaluable edited volume is accessible to students just entering this emergent field, and to researchers, both in academia and industry, in the fields of molecular biology, engineering, statistics, and signal processing. The book provides tutorial-level overviews and addresses the specific needs of genomic signal processing students and researchers as a reference book.

The book aims to address current genomic challenges by exploiting potential synergies between genomics, signal processing, and statistics, with special emphasis on signal processing and statistical tools for structural and functional understanding of genomic data. The book is partitioned into three parts. In part I, a brief history of genomic research and a background introduction from both biological and signal-processing/statistical perspectives are provided so that readers can easily follow the material presented in the rest of the book. In part II, overviews of state-of-the-art techniques are provided. We start with a chapter on sequence analysis, and follow with chapters on feature selection, clustering, and classification of microarray data. The next three chapters discuss the modeling, analysis, and simulation of biological regulatory networks, especially gene regulatory networks based on Boolean and Bayesian approaches. The next two chapters treat visualization and compression of gene data, and supercomputer implementation of genomic signal processing systems. Part II concludes with two chapters on systems biology and medical implications of genomic research. Finally, part III discusses the future trends in genomic signal processing and statistics research.

For more information and online orders please visit: <http://www.hindawi.com/books/spc/volume-2/>
For any inquiries on how to order this title please contact books.orders@hindawi.com

The EURASIP Book Series on Signal Processing and Communications publishes monographs, edited volumes, and textbooks on Signal Processing and Communications. For more information about the series please visit: <http://hindawi.com/books/spc/about.html>