

# Analysis of Physiological Responses and Use of Fuzzy Information Granulation–Based Neural Network for Recognition of Three Emotions

Byoung-Jun Park, Eun-Hye Jang, Kyong-Ho Kim, and Sang-Hyeob Kim

**In this study, we investigate the relationship between emotions and the physiological responses, with emotion recognition, using the proposed fuzzy information granulation–based neural network (FIGNN) for boredom, pain, and surprise emotions. For an analysis of the physiological responses, three emotions are induced through emotional stimuli, and the physiological signals are obtained from the evoked emotions. To recognize the emotions, we design an FIGNN recognizer and deal with the feature selection through an analysis of the physiological signals. The proposed method is accomplished in premise, consequence, and aggregation design phases. The premise phase takes information granulation using fuzzy *c*-means clustering, the consequence phase adopts a polynomial function, and the aggregation phase resorts to a general fuzzy inference. Experiments show that a suitable methodology and a substantial reduction of the feature space can be accomplished, and that the proposed FIGNN has a high recognition accuracy for the three emotions using physiological signals.**

**Keywords:** Emotion recognition, autonomic nervous system response, physiological signal, feature selection, fuzzy inference, information granulation.

Manuscript received Aug. 8, 2014; revised Sept. 9, 2015; accepted Sept. 16, 2015.

This work was funded by the Industrial Strategic Technology Development Program of MOTIE (10040927, Driver-oriented vehicle augmented reality system based on head up display for the driving safety and convenience).

Byoung-Jun Park (bj\_park@etri.re.kr), Eun-Hye Jang (cleta4u@etri.re.kr), Kyong-Ho Kim (kkh@etri.re.kr), and Sang-Hyeob Kim (corresponding author, shk1028@etri.re.kr) are with the IT Convergence Technology Research Laboratory, ETRI, Daejeon, Rep. of Korea.

## I. Introduction

In the field of vehicle research, studies on safety have been conducted to ensure the safety and security of automobiles and passengers. Vehicle-to-vehicle communication; electronic stability control; warning and emergency braking systems; blind spot monitoring; lane support systems; speed alert; and other features have been studied. Recently, systems with an in-vehicle heads-up display have been introduced to recognize driving-safety information, such as other vehicles, pedestrians, and obstacles with the risk of collision, and offer information to the driver from the driver's viewpoint [1]. European car safety researchers have developed a camera-based system that monitors a driver's facial expressions as they are driving along and that uses highly accurate emotion detection algorithms to determine whether the driver is suffering from road rage [2]. As this example indicates, in-vehicle technologies are playing an increasingly important role in car safety; in particular, where the emotional state of the driver is of concern.

One of the interesting topics in human–computer interaction (HCI) research is how to produce humanlike devices or machines for an intelligent user interface. Emotion plays an important role in the contextual understanding of messages from others in speech or visual form. For affective communication between a user and a computer, it is necessary to consider how emotions can be recognized and expressed during HCI, and emotion recognition is one of the key steps toward emotional intelligence in advanced human–machine interactions [3]. Many emotion-related physiological signals have been used to recognize human emotions because of the

strong relationship between physiological reactions and human emotional states [4].

Various approaches on emotion have reported a correlation between basic emotions and physiological responses [4]–[7]. Emotion recognition using physiological signals has recently been conducted using various methodologies; for example, neural networks, Fisher’s linear discriminant (FLD), the  $k$ -Nearest Neighbors ( $k$ -NN) algorithm, and a support vector machine (SVM) [7]–[9]. Neural networks have been widely used to deal with pattern recognition problems. It was shown that neural networks can be trained to approximate complex discriminant functions [10]. However, a neural network (NN) requires a large number of parameters to be determined, particularly in the case of a multilayer network topology [11]. The appropriate neural architecture, the number of hidden layers, and the number of neurons in each hidden layer are the important design issues that can affect the prediction accuracy [12].

For this study, we have focused on the relationship between emotions and physiological responses, as well as the design of a fuzzy information granulation–based NN recognizer for three emotions (boredom, pain, and surprise). In the study, we hypothesize and prove that these emotions can be characterized by certain physiological signals and that the proposed recognizer can correctly categorize such signals. To analyze the physiological responses of these emotions, we achieve physiological signals by inducing emotions. The induction of emotions is done through stimuli with the proper appropriateness and effectiveness. To obtain the necessary physiological signals for the above emotions, we recruited the help of a number of subjects to participate in our study. The physiological signals of a subject are measured for the induced emotions. The electrodermal activity (EDA), skin temperature (SKT), photoplethysmography (PPG), and electrocardiogram (ECG) are acquired as physiological signals. From these signals, 27 features are extracted for a steady state (baseline), emotion state, and difference between the baseline and emotion state. In addition, we select some feature groups throughout the statistical analysis of the physiological responses.

With the selected features, we investigate the development of a fuzzy information granulation–based NN (FIGNN) recognizer driven by a fuzzy  $c$ -means (FCM) clustering paradigm. The proposed FIGNN embraces three phases (namely, premise, consequence, and aggregation design phases) from the viewpoint of linguistic analysis expressed as a collection of “if-then” rules. In its design, the premise phase of the rule relates to the use of FCM clustering as a tool of information granulation. In the conclusion phase of the network, we develop three types of polynomials as the corresponding local models, while the aggregation involves

mechanisms of fuzzy inference. This study provides an algorithmic framework and demonstrates the effectiveness of the proposed approach. To demonstrate the usefulness of the proposed recognizer for the three emotions, we discuss the comparative results of emotion recognition using some machine learning algorithms such as C4.5,  $k$ -NN, FLD, SOM, and SVM.

## II. Experiments on Emotion Induction and Feature Extraction for Emotion Recognition

In this section, we deal with experiments on the induction of the three emotions (boredom, pain, and surprise) using stimuli, and the feature extraction of the acquired physiological signals induced by emotional stimuli. A total of 217 college students (mean age of  $22.3 \pm 2.04$  years) participated in our experiments. They reported no history of medical illness from heart disease, respiration, or central nervous system disorders or psychotropic medications. They were introduced to the experiment protocols and filled out a written consent before the beginning of the experiments. In addition, they were paid USD 30 per session to compensate for their participation.

The laboratory used for the experiment was a sound-proof room (lower than 35 dB), 5 m  $\times$  2.5 m in size, where all outside noise and artifacts were completely blocked. In the laboratory, a comfortable chair, a 38-inch monitor, an intercommunication device, and a CCTV camera were installed to observe and record the behavior of the subjects. We introduced the experimental procedures and indications in detail to the subjects. Each subject had an adaptation time of about 30 min to become comfortable in the laboratory environment, and electrodes were then attached to the subject’s wrist, finger, and ankle to measure their physiological signals.

For the three emotions, we used 1 min- to 3 min-long audio-visual stimuli. The boredom stimulus was a combination of the presentation of a “+” symbol on the screen and the repeated sound of numbers being counted from 1 to 10 during a 3 min period. The stimulus for provoking pain combined a “+” symbol on the screen and an increase in pressure using a blood pressure cuff during a 1 min period. The surprise provoking stimulus was the sudden presentation of the boredom stimulus image and a hog-caller, the sound of breaking glass, and thunder as the subjects tried to concentrate on a task during an 1 min period. Audio-visual film clips are widely used because they have the desirable properties of being readily standardized, involve no deception, and are dynamic rather than static. They also have a relatively high degree of ecological validity, in so far as emotions are often evoked by dynamic visual and auditory stimuli that are external to the individual [13]–[14].

The stimuli were used to induce emotions and test their

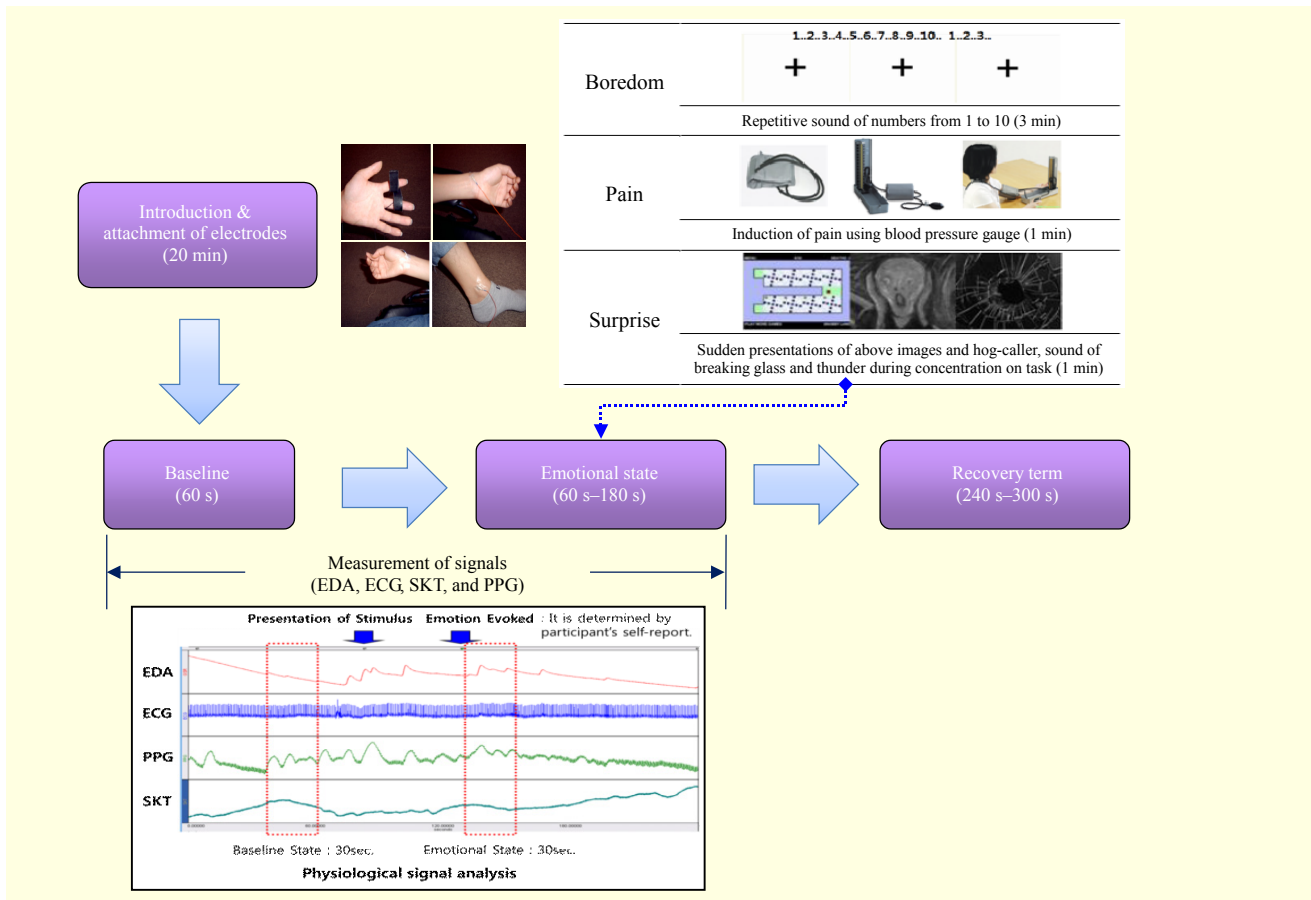


Fig. 1. Experimental procedure for emotion induction.

Table 1. Appropriateness and effectiveness of emotion stimuli.

	Boredom	Pain	Surprise	Average
Appropriateness	86.0	97.3	94.1	92.5
Effectiveness	5.23	4.96	6.12	5.43

appropriateness and effectiveness. The appropriateness of emotional stimuli means the consistency between the target emotions designed to induce each emotion and the categories of a subject's experienced emotion. The effectiveness was determined by the intensity of the emotions reported and rated by the subjects on a 1- to 7-point Likert-type scale, with 1 being the "least bored" or "not bored" and 7 being the "most bored." Table 1 shows the results of the appropriateness and effectiveness of emotional stimuli obtained from a preliminary study. The emotional stimuli showed an appropriateness of 92.5% and an effectiveness of 5.43 points on average.

For the induced emotions, the physiological signals (ECG, EDA, SKT, and PPG) of the subject were measured for both the baseline and the emotional states. An ECG is a signal used

to detect the electrical activity of the heart through electrodes attached to the outer surface of the skin. An EDA represents the activity of the autonomic nervous system using the activity of the sweat glands. SKT is an important and effective indicator of an emotional state, and reflects activity in the autonomic nervous system. Variations in SKT mainly arise from localized changes in blood flow caused by vascular resistance or arterial blood pressure. PPG aims to observe the mechanical movement of the heart and the kinetics of the blood flow, and manifests the pulsation of the chest wall and great arteries followed by a heartbeat and wave form. Electrodes were attached on the first joint of the ring finger and the last joint of the thumb finger of the non-dominant hand for measuring the SKT and PPG, respectively. The EDA was measured through two Ag/AgCl electrodes attached to the middle joint of the index and middle fingers of the non-dominant hand. For the ECG, electrodes were placed on both wrists and the left ankle (reference) using a two-electrode method based on lead I. The electrodes were filled with a 0.05 molar isotonic NaCl paste to provide a continuous connection between the electrodes and skin.

For the induced emotions, these physiological signals were measured for 60 s prior to the emotional stimuli (baseline), for

Table 2. Features extracted from physiological signals.

Signals	Features
EDA	bSCL, bSCR, eSCL, eSCR, dSCL, dSCR
SKT	bmeanSKT, emeanSKT, dmeanSKT
PPG	bBVP, bPPT, eBVP, ePPT, dBVP, dPPT
ECG	bHR, bLF, bHF, bHRV, eHR, eLF, eHF, eHRV, dHR, dLF, dHF, dHRV

1 min to 3 min during the presentation of the stimuli (emotional state), and for 4 min to 5 min after presentation of the emotional stimuli as the recovery period. The subjects then rated their own emotion they experienced during the presentation of the stimuli based on the emotion assessment scale. Figure 1 shows the experimental procedure for the emotion induction. The physiological signals obtained were analyzed for 30 s from the baseline and emotional state, as shown in Fig. 1. The baseline was selected and analyzed for 30 s before the emotional stimulus was presented. The emotional states were determined based on the results of the subject’s self-reporting, in which an emotion was most strongly expressed during the presentation of a stimulus.

For the three emotions, 27 features were extracted from the analysis of the physiological signals and are summarized in Table 2. These features are well known and are generally used to analyze the autonomic nervous system. The 27 features can be categorized into three groups. The nine features of the first group are denoted by “b,” which represents “baseline signals,” and the second group has nine features marked by an “e,” which represents “extracted from emotional states.” The last group is denoted by “d” and involves nine features based on the difference between the baseline and emotion signals.

### III. Design of Fuzzy Information Granulation–Based NN Recognizer

An NN is a computational intelligence model inspired by the structure and functional aspects of biological neurons [15]. This approach has been widely used to deal with pattern recognition problems. The generic topology of an NN generally consists of three layers. A neuron in the input layer is connected to a layer of hidden neurons, and a hidden neuron is connected to an output neuron. The activity of the input neurons represents the raw information fed into the network; the activity of each hidden neuron is determined based on the activities of the input neuron and the weights on the connections between the input and hidden neurons; and the behavior of the output depends on the activity of the hidden neurons and connection weights

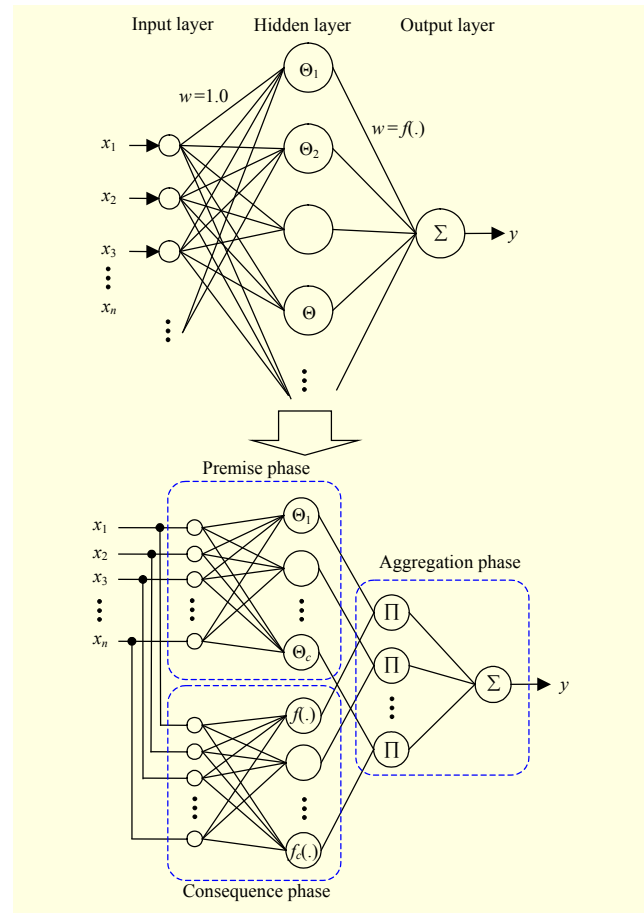


Fig. 2. Topology of FIGNN.

between the hidden and output layers.

The proposed FIGNN exhibits a similar topology as the one encountered in a simple NN, as shown in Fig. 2. However, the functionality and associated design process exhibit some evident differences. In particular, the receptive fields do not assume any explicit functional form (for example, Gaussian or ellipsoidal), but are directly reflective of the nature of the data and come as a result of fuzzy clustering.

Let us consider a set of prototypes,  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c$ , which have been formed by the FCM clustering method. Then, the receptive fields can be expressed in the following way:

$$\Theta_i(\mathbf{x}) = \frac{1}{\sum_{j=1}^c \left( \frac{\|\mathbf{x} - \mathbf{v}_i\|^2}{\|\mathbf{x} - \mathbf{v}_j\|^2} \right)}. \quad (1)$$

In addition, the weights between the output layer and the hidden layer are not constants but come in the form of polynomials of the input variables; namely,

$$w_i = f_i(\mathbf{x}). \quad (2)$$

The neurons located at the output layer complete a linear

combination of the activation levels of the corresponding receptive fields as follows:

$$y(\mathbf{x}) = \sum_{i=1}^c w_i \Theta_i(\mathbf{x}) = \sum_{i=1}^c f_i(\mathbf{x}) \Theta_i(\mathbf{x}). \quad (3)$$

The above structure of the FIGNN can be represented through a collection of fuzzy rules, as follows:

$$\text{If } \mathbf{x} \text{ is } \Theta_i, \text{ then } f_i(\mathbf{x}), \quad (4)$$

where  $\Theta_i$  is a fuzzy set resulting from the  $i$ -cluster (membership function) of the  $i$ th fuzzy rule,  $f_i(\mathbf{x})$  is a polynomial function generalizing a numeric weight used in a simple NN, and  $c$  is the number of fuzzy rules (clusters). The FIGNN employs a partition function created by FCM clustering as an activation function in the hidden layer, and polynomial weights between the hidden and output layers.

Let us discuss an extension of the network by considering the fuzzy rules described by (4) in terms of fuzzy inference. Figure 2 illustrates the architecture of the FIGNN. Here, the premise phase includes input and hidden layers of the NN, and connection weights of the NN are reflected in the consequence phase. The aggregation phase embraces the output layer and activity ( $\Pi$ ) of the hidden neurons and weights between the hidden and output layers of the NN. All connections are 1.0.

From the viewpoint of a linguistic analysis, the network is implemented in three design phases (that is, the premise, consequence, and aggregation design phases), which are reflected in the IF part of the fuzzy rule, the THEN part of the fuzzy rule, and fuzzy inference, respectively. The premise phase relates to the partition function of the input space using FCM clustering. In the consequence phase, a polynomial function carries out the presentation of a partitioned local space. Finally, the output of the network is obtained by fuzzy inference in the aggregation phase.

The premise phase of the FIGNN is formed by means of FCM clustering. The objective function,  $Q$ , guiding the clustering process of FCM is expressed as a sum of the distances of individual data from  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c$ .

$$Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|\mathbf{x}_k - \mathbf{v}_i\|^2, \quad (5)$$

$$\|\mathbf{x}_k - \mathbf{v}_i\|^2 = \sum_{j=1}^n \frac{(x_{kj} - v_{ij})^2}{\sigma_j^2}. \quad (6)$$

Here,  $m$  represents a fuzzification factor,  $m > 1.0$ . The commonly used value of  $m$  is 2. In addition,  $N$  is the number of patterns (data) and  $\sigma_j$  is the standard deviation of the  $j$ th variable. The minimization of  $Q$  is realized in successive iterations by adjusting both the prototypes and entries of the partition matrix,  $\min Q(\mathbf{U}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c)$ .

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i\|}{\|\mathbf{x}_k - \mathbf{v}_j\|} \right)^{\frac{2}{m-1}}}, \quad (7)$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^N u_{ik}^m \mathbf{x}_k}{\sum_{k=1}^N u_{ik}^m}. \quad (8)$$

The properties of the optimization algorithm are well documented in the literature [16]–[17]. In the context of our investigations, we note that the resulting partition matrix is a clear realization of  $c$  fuzzy relations with the membership functions  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_c$  forming the corresponding rows of the partition matrix  $\mathbf{U}$ ; that is,  $\mathbf{U} = [\mathbf{u}_1^T \ \mathbf{u}_2^T \ \dots \ \mathbf{u}_c^T]$ .

Polynomial functions are dealt with in the consequence phase. In Fig. 2 and (4),  $f_i(\mathbf{x})$  is represented as a polynomial of the following forms:

$$\text{Constant: } f_i(\mathbf{x}) = a_{i0}, \quad (9)$$

$$\text{Linear: } f_i(\mathbf{x}) = a_{i0} + \sum_{j=1}^n a_{ij} x_j, \quad (10)$$

$$\text{Quadratic: } f_i(\mathbf{x}) = a_{i0} + \sum_{j=1}^n a_{ij} x_j + \sum_{j=1}^n \sum_{k=j}^n a_{ijk} x_j x_k. \quad (11)$$

These functions are activated by a partition matrix and lead to local regression models for the consequence phase in each linguistic rule. Interestingly, to improve the performance (accuracy) of the FIGNN, we typically require a substantial number of receptive fields, which amounts to an increased level of granularity of the resulting construct. In the architecture proposed here, we achieve a low classification rate by forming a network at a lower level of granularity (higher generality) by making the connections nonlinear. As the experimental evidence presented later on in this study demonstrates, we retained a fairly low level of granularity in comparison with the FIGNN.

Let us consider the FIGNN structure by considering the fuzzy partition realized in terms of the FCM clustering, as shown in Fig. 2. In this figure, the node denoted by  $\Pi$  is realized as a product of the corresponding fuzzy set and polynomial function. The family of fuzzy sets,  $\Theta_i$ , forms a partition (such that the sum of the membership grades sums up to 1 at each point of the input space). The  $\Sigma$  neuron is described by a linear sum, as shown in (3). The output of the network can be obtained through a general fuzzy inference based on a collection of “if-then” fuzzy rules. More specifically, we obtain

$$y = g(\mathbf{x}) = \sum_{i=1}^c \frac{u_i f_i(\mathbf{x})}{\sum_{j=1}^c u_j} = \sum_{i=1}^c u_i f_i(\mathbf{x}), \quad (12)$$

where  $u_i = \Theta_i(\mathbf{x})$ , and these membership degrees sum up to 1. In addition,  $g(\mathbf{x})$  is a representation of the FIGNN as a discriminant function. Based on the local representation schemes (polynomials), the global characteristics of the networks result through the composition of their local relationships during the aggregation phase.

We consider the use of the FIGNN as a human emotion recognizer, and the discriminant function assigns  $\mathbf{x}$  to  $\omega_i$  if  $g_i(\mathbf{x}) \geq g_j(\mathbf{x})$  for all  $j \neq i$ . The final output of the networks (that is, the result of (12)) is used as a discriminant function and can be rewritten as a linear combination as follows:

$$g(\mathbf{x}) = \mathbf{a}^T \mathbf{f}\mathbf{x}, \quad (13)$$

where  $\mathbf{a}$  is a vector of coefficients of polynomial functions used in the consequence layer of the rules in (9)–(11). More specifically, we have the following:

- Constant:  $\mathbf{a}^T = [a_{10}, \dots, a_{c0}]$   
 $\mathbf{f}\mathbf{x} = [u_1, \dots, u_c]^T$
- Linear:  $\mathbf{a}^T = [a_{10}, \dots, a_{c0}, a_{11}, \dots, a_{c1}, \dots, a_{cn}]$   
 $\mathbf{f}\mathbf{x} = [u_1, \dots, u_c, u_{1 \times 1}, \dots, u_c x_1, \dots, u_c x_n, \dots, u_c x_n x_n]^T$
- Quadratic:  $\mathbf{a}^T = [a_{10}, \dots, a_{c0}, a_{11}, \dots, a_{c1}, \dots, a_{cn}, \dots, a_{cm}]$   
 $\mathbf{f}\mathbf{x} = [u_1, \dots, u_c, u_{1 \times 1}, \dots, u_c x_1, \dots, u_c x_n, \dots, u_c x_n x_n]^T$

The coefficient vector,  $\mathbf{a}$ , in the consequence phase of the FIGNN recognizer can be determined based on the standard method of least squares.

#### IV. Results of Physiological Responses Induced by Emotional Stimuli

For the physiological signals obtained as the emotional responses to stimuli, verification of the differences between the steady state (baseline) and the emotion states was carried out. To do so, we used a statistical significance measure (namely, a paired t-test) to assure the difference of the response. Figure 3 illustrates the results of a paired t-test between the baseline and emotional states.

For a state of boredom, there was a significant increase in the meanSKT ( $p < 0.5$ ). SKT serves as a surrogate marker of blood flow changes resulting from vascular reactivity. SKT is influenced mainly by sympathetic adrenergic vasoconstrictor nerves, and an increased SKT in our results indicates the occurrence of vasodilation through the withdrawal of neural activity. SKT shows an extreme decrease during mental load, stress, fear, and so on, and an increase during relaxation, boredom, and sleep. In particular, SKT under emotional stress changes significantly [18]–[19]. We were only able to verify

that pain and surprise are associated with a mild increase in SKT, as the change in meanSKT during a 30 s emotional state was used in the analysis. However, considering that SKT is a relatively slow indicator of changes in an emotional state, we need to analyze the change of SKT during an emotional state over time to confirm whether a significant change in SKT exists, such as in previous studies.

Physiological responses induced by pain showed a strong decrease in BVP, a mild decrease in PTT, and mild increases in SCL and SCR ( $p < 0.001$ ). BVP is a measure used to determine the amount of blood currently running through the vessels (for example, in the finger of the test subject) and serves as information of vasoconstriction. A decrease in BVP amplitude from the baseline in response to a stimulus implies a peripheral vasoconstriction in the finger, and is known to be associated with arousal due to the stimulus [20]–[21]. An increased PTT means a suppression of the sympathetic nervous system activation, and thus a strong decrease of PPT during a surprised emotion reflects a sympathetic activation [22]. In addition, an increased SCL and SCR indicate that the skin is sweaty and the sympathetic nervous system is activated. In particular, the SCL and SCR are related to the sympathetic-adrenal-medullary activation, which indicates the progression of pain.

The surprise response was typified by a significant increase in SCL, SCR, and HR, and a strong decrease in BVP and PTT ( $p < 0.001$ ). Kreibig reported that surprise is associated with a short-term duration of SCR with a medium response size and characterized by a rapid increase and rapid return; an increase in SCL; an increase in HR; and a decrease or increase in finger temperature [23].

For the emotion recognition, we selected some features with a significant difference between the baseline and emotional state throughout the results of autonomic nervous system responses. Firstly, nine features with a difference between the baseline and emotion state were selected from the 27 features; that is, dSCL, dSCR, dmeanSKT, dBVP, dPPT, dHR, dLF, dHF, and dHRV. Figure 4 shows the relation between three emotions and three features (SCR, SCL, and SKT) of each group (baseline, emotion, and difference). In the baseline state shown in Fig. 4(a), we can see that the three emotions are blended. Surprise, boredom, and pain emotions in the emotion state are more distinguished than those of the baseline state, as shown in Fig. 4(b). However, for the pain emotion, it is difficult to draw a line between the other emotions. Figure 4(c) shows the three types of emotions formed into emotion classes with a difference between the emotional and baseline states.

Second, we selected six features with significantly different values; namely, dSCL, dSCR, dmeanSKT, dBVP, dPPT, and dHR. Finally, three features (SCL, SCR, and BVP) with

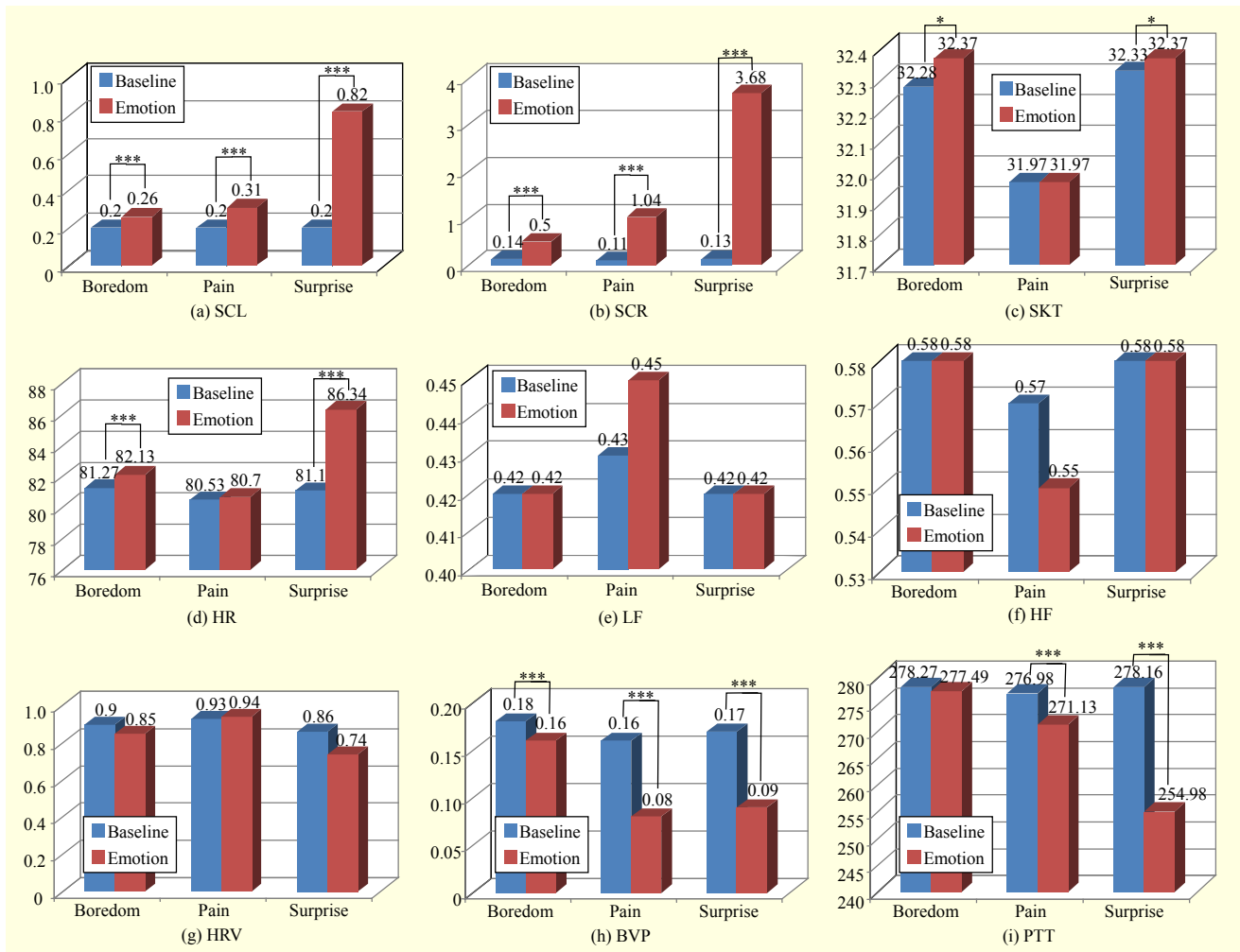


Fig. 3. Results of paired t-test between steady state (baseline) and emotional state (\*  $p < 0.05$ , \*\*\*  $p < 0.001$ ).

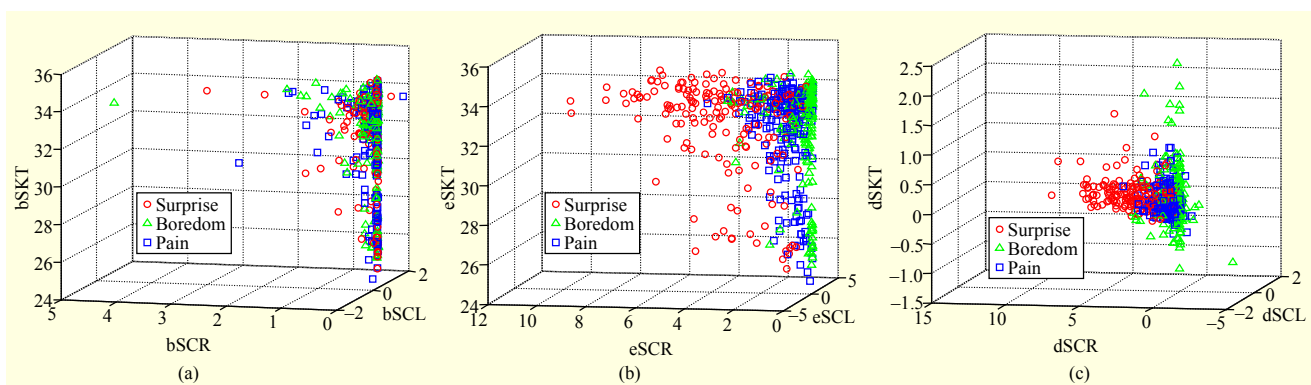


Fig. 4. Relation between three emotions (surprise, boredom, and pain) and features (SCR, SCL, and SKT): (a) baseline state, (b) emotional state, and (c) difference between baseline and emotional states.

significantly difference values for three emotions were selected.

Comparative results of the emotion recognition applied to the proposed FIGNN with the selected features are shown in the next section.

## V. Results of Emotion Recognition

In this section, the FIGNN recognizer is applied to the recognition of the three emotions with the selected features described in the previous section. Our objective is to quantify

**Table 3.** Recognition accuracy of FIGNN for three emotions.

No. of features	$f(x)$	CV	RRSV
3	Constant	71.3	70.4 ± 3.2
	Linear	72.8	71.0 ± 1.9
	Quadratic	72.1	70.7 ± 3.2
6	Constant	77.3	62.6 ± 4.1
	Linear	78.8	75.0 ± 2.8
	Quadratic	78.2	78.1 ± 1.7
9	Constant	71.7	58.4 ± 17.0
	Linear	76.0	71.6 ± 4.5
	Quadratic	71.5	70.8 ± 2.0

the performance of the proposed FIGNN recognizer and compare it with the performance of some other machine learning algorithms reported in the literature [24]–[28]. Here, we consider several well-known methods concerning recognition problems; that is, C4.5,  $k$ -NN, FLD, SOM, and SVM. In the assessment of the performance of the proposed recognizer, we use the recognition accuracy for three emotions. The experiments completed in this study are reported for a ten-fold cross-validation (CV) and ten-fold repeated random subsampling validation (RRSV) for assessing how the results of a statistical analysis will generalize to an independent dataset. In the case of RRSV, 70% of all emotional patterns were selected randomly for training, and the remaining patterns used for testing purposes.

For the recognition results of boredom, pain, and surprise, Table 3 summarizes the recognition accuracy (%) of the proposed FIGNN. The experiments were completed for a number of features (see Table 3). As mentioned in the previous chapter, we extracted 27 features and selected nine, six, and three features throughout the analysis of the autonomic nervous system responses using the paired  $t$ -test. The feature selection constitutes a fundamental development phase of pattern recognition and predetermines the effectiveness of the overall recognition schemes to a significant extent [29]. It has become apparent that this is essential, both to reduce the overall computational overload and to possibly enhance the discriminatory capabilities of the reduced feature space.

The quality of the reduced feature space is quantified with the use of the recognition accuracy produced by the FIGNN recognizer on the testing set. Namely, the accuracy with nine features is lower than with three and six features. In other words, the use of all features drops the recognition accuracy of the three emotions. We achieve substantially higher recognition accuracy with six features, which is 66.7% of the space

**Table 4.** Recognition accuracy of FIGNN with six features and linear type of polynomial functions for each emotion.

	Surprise	Boredom	Pain
Surprise	81.82	1.14	17.05
Boredom	3.55	75.15	21.30
Pain	8.85	11.98	79.17

**Table 5.** Recognition accuracy of well-known methods for three emotions.

Methods	No. of features	CV	RRSV
C4.5 [24]	3	64.2	61.2 ± 6.7
	6	64.8	62.4 ± 4.3
	9	64.6	60.4 ± 4.4
$k$ -NN [25]	3	62.8	63.6 ± 3.4
	6	71.3	69.8 ± 3.3
	9	68.3	66.4 ± 3.2
FLD [14]	3	68.3	66.9 ± 3.0
	6	70.8	70.8 ± 4.1
	9	66.5	61.0 ± 10.2
SOM [27]	3	65.7	63.7 ± 4.0
	6	68.7	69.6 ± 2.7
	9	64.7	64.1 ± 3.1
SVM [14]	3	66.3	67.4 ± 4.2
	6	71.1	70.1 ± 2.6
	9	68.3	67.0 ± 2.5

dimensionality. In the case of three features, the recognition accuracy is similar with nine features, despite the dimensionality of the space being reduced to 66.7%.

Further detailed results of the emotion recognition accuracy by the FIGNN with six features and a linear type of polynomial functions are shown in Table 4. This provides an accuracy of 78.8% when recognizing all emotions and when the accuracy of each emotion has a range of 75% to 82%. Namely, the proposed methodology successfully recognized pain (81.8%), boredom (75.2%), and surprise (79.2%).

In addition, the experimental results show that the proposed approach outperforms the existing methods in terms of better recognition capabilities on three, six, and nine feature spaces, as shown through a comparison of Tables 3 and 5. The comparative analysis illustrated in Table 5 contrasts the proposed method, shown in Table 3, with other methods, and the proposed models are preferred as the architecture in the recognizer of the three emotions in general. The generalization (recognition accuracy on the testing set) of the well-known



methods is 62% to 71%, as shown in Table 5. This means that some algorithms shown in the table are not useful as a recognizer for the recognition of the three emotions. The values of the testing performance are good indicators of the generalization capabilities of the constructed methods. When selecting a method, if the approximation capability of a trained model is only considered, then the selected model has great recognition accuracy; however, it has a deteriorated generalization (prediction) capability and cannot be applied to a real system. In particular, this is conspicuous in a nonlinear problem. The proposed method leads to better recognition results for a reduced feature space than the other methods, as shown in Tables 3 and 5. Here, C4.5 comes from the Classification Toolbox of MATLAB. For  $k$ -NN, FLD, and SVM, we used Duda's Toolbox ([www.yom-tov.info/toolbox.html](http://www.yom-tov.info/toolbox.html)). The SOM toolbox available in MATLAB has offered SOM algorithms, as can be found at [www.cis.hut.fi/projects/somtoolbox/](http://www.cis.hut.fi/projects/somtoolbox/).

## VI. Conclusion

In this study, we dealt with an analysis of physiological responses and the proposed FIGNN recognizer for three emotions (boredom, pain, and surprise).

For the analysis of physiological responses, the subjects' emotions were induced by emotion stimuli, and emotional physiology signals, such as EDA, SKT, PPG, and ECG, were acquired from the subjects. The emotion stimuli used to induce a subject's emotion were evaluated for their appropriateness and effectiveness. The results showed that the emotional stimuli have a suitability of 92.5% and an effectiveness of 5.43 points (on a 7-point Likert-type scale) on average. The appropriateness of the emotional stimuli means the consistency between the target emotions designed to induce each emotion and the categories of a subject's experienced emotion. The effectiveness is determined based on the intensity of the emotions.

From the obtained physiological signals, 27 features were extracted on the baseline and emotion states, and the difference between both; in addition, three, six, and nine features were selected through an analysis of the physiological responses using a paired  $t$ -test. For boredom, higher SCL, SCR, SKT, and HR, and lower BVP, of the emotion state were exhibited significantly as compared with those of the baseline. Pain shows a significant difference between the baseline and emotion for SCL, SCR, BVP, and PTT. In the case of surprise, a significant difference between the baseline and emotion states was shown in all features except LF, HF, and HRV.

In addition, we proposed the FIGNN recognizer based on fuzzy information granulation for emotion recognition. The

proposed recognizer is expressed as "if-then" rules and accomplished using three phases (that is, premise, consequence, and aggregation design phases) from the viewpoint of a linguistic analysis. The premise phase took the fuzzy  $c$ -means clustering as information granulation, the consequence phase adopted a polynomial function, and the aggregation phase resorted to a general fuzzy inference to recognize patterns. Using two types of polynomial functions in the consequence phase can help improve the characteristics of a basic NN recognizer and carry out the presentation of a partitioned local space. These phases contribute directly to "if-then" rules that provide intuitional interpretation using a linguistic analysis for the given problem.

This study has the following limitations: (a) it needs to investigate a broader range of emotions that can be induced while driving, such as anger, sadness, and so on, for an intelligent driving assistance service within a vehicle, and (b) it needs to analyze emotions obtained from subjects who are tested under real driving conditions. In this study, we used physiological signals obtained from subjects operating within an indoor test environment (that is, a laboratory), where all outside noise and artifacts were completely blocked. To acquire physiological signals of a driver in a vehicle, sensors would need to be installed within the vehicle in places such as the steering wheel, the driver's seat, buttons, and so on, and a monitoring system and robust sensors would have to be developed to oppose noise, such as the physical motions of the driver.

The experimental results provided sound evidence behind the selection process, showing that the reduced feature spaces lead to better recognition results than those obtained through other methods in terms of generalization. The proposed recognizer will lead to a better chance to recognize human emotions through physiological signals during the emotional interaction between humans and machines.

## References

- [1] H.S. Park et al., "In-Vehicle AR-HUD System to Provide Driving-Safety Information," *ETRI J.*, vol. 35, no. 6, Dec. 2013, pp. 1038–1047.
- [2] A. Yüce, N.M. Arar, and J.-P. Thiran, "Multiple Local Curvature Gabor Binary Patterns for Facial Action Recognition," *LNCS*, vol. 8212, Oct. 2013, pp. 136–147.
- [3] J. Wagner, J. Kim, and E. Andre, "From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification," *IEEE Int. Conf. Multimedia Expo*, Amsterdam, Netherland, July 2005, pp. 940–943.
- [4] O. Alaoui-Ismaili et al., "Basic Emotions Evoked by Odorants:

- Comparison between Autonomic Responses and Self-Evaluation,” *Physiology Behavior*, vol. 62, no. 4, Oct. 1997, pp. 713–720.
- [5] D. Palomba et al., “Cardiac Responses Associated with Affective Processing of Unpleasant Film Stimuli,” *Int. J. Psychophysiology*, vol. 36, no. 1, Apr. 2000, pp. 45–57.
- [6] G. Stemmler, “Physiological Processes during Emotion,” in *The Regulation Emotion*, Mahwah, NJ, USA: Erlbaum, 2004, pp. 33–70.
- [7] C.L. Stephens, I.C. Christie, and B.H. Friedman, “Autonomic Specificity of Basic Emotions: Evidence from Pattern Classification and Cluster Analysis,” *Biol. Psychology*, vol. 84, no. 3, July 2010, pp. 463–473.
- [8] B.-J. Park et al., “Design of Prototype-Based Emotion Recognizer Using Physiological Signals,” *ETRI J.*, vol. 35, no. 5, Oct. 2013, pp. 869–879.
- [9] E.-H. Jang et al., “Classification of Three Negative Emotions Based on Physiological Signals,” *INTELLI*, Venice, Italy, Apr. 2013, pp. 75–78.
- [10] R.P. Lippmann, “An Introduction to Computing with Neural Nets,” *IEEE ASSP Mag.*, vol. 4, no. 1, Apr. 1987, pp. 4–22.
- [11] A. Patrikar and J. Provenge, “Pattern Classification Using Polynomial Networks,” *Electron. Lett.*, vol. 28, no. 12, 1992, pp. 1109–1110.
- [12] F. Ros, M. Pintore, and J.R. Chretien, “Automatic Design of Growing Radial Basis Function NNs Based on Neighborhood Concepts,” *Chemometrics Intell. Laboratory Syst.*, vol. 87, no. 2, June 2007, pp. 231–240.
- [13] D. Palomba et al., “Cardiac Responses Associated with Affective Processing of Unpleasant Film Stimulus,” *Int. J. Psychophysiology*, vol. 36, no. 1, Apr. 2000, pp. 45–57.
- [14] J.J. Gross and R.W. Levenson, “Emotion Elicitation Using Films,” *Cognition Emotion*, vol. 9, no. 1, 1995, pp. 87–108.
- [15] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, 2nd ed., New York, USA: Wiley-Interscience, 2000.
- [16] F. Akiyama, “An Example of Software System Debugging,” *Inf. Process.*, vol. 71, 1971, pp. 353–379.
- [17] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York, USA: Plenum Press, 1981.
- [18] H. Helson and L. Quantius, “Changes in Skin Temperature Following Intense Stimulation,” *J. Experimental Psychology*, vol. 17, no. 1, Feb. 1934, pp. 20–35.
- [19] F. Talbot, “Skin Temperatures of Children,” *American J. Diseases Children*, vol. 42, 1931, pp. 965–967.
- [20] C. Iani, D. Gopher, and P. Lavie, “Effects of Task Difficulty and Invested Mental Effort on Peripheral Vasoconstriction,” *Psychophysiology*, vol. 41, no. 5, Sept. 2004, pp. 789–798.
- [21] V.N. Salimpoor et al., “The Rewarding Aspects of Music Listening are Related to Degree of Emotional Arousal,” *PlosOne*, vol. 4, no. 10, 2009, p. e7487.
- [22] J.M. Gottman et al., “The Relationship between Heart Rate Reactivity, Emotionally Aggressive Behavior and General Violence in Batterers,” *J. Family Psychology*, vol. 9, no. 3, Sept. 1995, pp. 227–248.
- [23] S.D. Kreibig, “Autonomic Nervous System Activity in Emotion: a Review,” *Biol. Psychology*, vol. 84, no. 3, July 2010, pp. 394–421.
- [24] L. Breiman et al., “*Classification and Regression Trees*,” Monterey, CA, USA: Wadsworth, Inc., 1984.
- [25] J.R. Quinlan, “*C4.5 Programs for Machine Learning*,” San Mateo, CA, USA: Morgan Kaufmann, 1992.
- [26] J.M. Keller, M.R. Gray, and J.A. Givens, Jr., “A Fuzzy K-Nearest Neighbor Algorithm,” *IEEE Trans. Syst., Man, Cybern.*, vol. 15, no. 4, 1985, pp. 580–585.
- [27] P.D. Wasserman, “*Advanced Methods in Neural Computing*,” New York, USA: Van Nostrand Reinhold, 1993, pp. 35–55.
- [28] T. Kohonen, “Self-Organizing Maps,” *Springer Series Inf. Sci.*, vol. 30, Heidelberg, Berlin: Springer, 2001, p. 501.
- [29] C.M. Bishop, “*Neural Networks for Pattern Recognition*,” Oxford, UK: Oxford Univ. Press, 1995.



**Byoung-Jun Park** received his BS, MS, and PhD degrees in control and instrumentation engineering from Wonkwang University, Iksan, Rep. of Korea, in 1998, 2000, and 2003, respectively. From 2005 to 2006, he held the position of postdoctoral fellow with the Department of Electrical and Computer Engineering, University of Alberta, Canada. Since 2008, he has worked as a senior researcher at ETRI. His research interests encompass computational intelligence; pattern recognition; granular and relational computing; and IT Convergence.



**Eun-Hye Jang** received her BA, MA, and PhD degrees in experimental & biological psychology from Chungnam National University, Daejeon, Rep. of Korea, in 2000, 2002, and 2009, respectively. Since May 2009, she has worked as a senior researcher at ETRI.

Her research interests include psychophysiology of emotion, emotion recognition, and IT-cognition/emotion convergence.



**Kyong-Ho Kim** is a principal researcher and director of the Smart Driving Assistance Research Section, ETRI. He received his BS and MS degrees in electronic engineering from Kyungpook National University, Daegu, Rep. of Korea, in 1993 and 1995, respectively, as well as receiving his PhD degree in computer

science from the Korea Advanced Institute of Science and Technology, Daejeon, Rep. of Korea, in 2010. Since 1994, he has been with ETRI. His current research topics include intelligent vehicles, human-computer interaction, head-up displays, and augmented reality applications in vehicles.



**Sang-Hyeob Kim** received his MS and BS degrees in material science from Jeonbuk National University, Jeonju, Rep. of Korea, in 1984 and 1986, respectively, as well as receiving his PhD degree in material science and engineering from Tohoku University, Sendai, Japan, in 1994. From 1994 to 1997, he worked as a post-doc. at KRISS, Daejeon, Rep. of Korea. Since 2000, he has worked both as a senior researcher and principal researcher at ETRI. His research interests include organic/inorganic hybrid devices, IT-BT-NT-ET convergence devices, and synthesis of nano-structured materials. He has authored or co-authored over 50 papers and holds 10 US patents as well as 30 Korean patents.