A Review on the Computational Methods for Emotional State Estimation from the Human EEG

Min-Ki Kim, 1 Miyoung Kim, 2 Eunmi Oh, 2 and Sung-Phil Kim 3

1 Department of Brain and Cognitive Engineering, Korea University, Seoul 136701, Republic of Korea
2 Samsung Electronics, DMC R&D Center, Suwon 443742, Republic of Korea
3 Research and Business Foundation, Korea University, Seoul 136701, Republic of Korea

Correspondence should be addressed to Sung-Phil Kim; spkim@korea.ac.kr

Received 11 January 2013; Accepted 18 February 2013

Academic Editor: Chang-Hwan Im

Copyright © 2013 Min-Ki Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A growing number of affective computing researches recently developed a computer system that can recognize an emotional state of the human user to establish affective human-computer interactions. Various measures have been used to estimate emotional states, including self-report, startle response, behavioral response, autonomic measurement, and neurophysiologic measurement. Among them, inferring emotional states from electroencephalography (EEG) has received considerable attention as EEG could directly reflect emotional states with relatively low costs and simplicity. Yet, EEG-based emotional state estimation requires well-designed computational methods to extract information from complex and noisy multichannel EEG data. In this paper, we review the computational methods that have been developed to deduct EEG indices of emotion, to extract emotion-related features, or to classify EEG signals into one of many emotional states. We also propose using sequential Bayesian inference to estimate the continuous emotional state in real time. We present current challenges for building an EEG-based emotion recognition system and suggest some future directions.

1. Introduction

An emotional state refers to a psychological and physiologic state in which emotions and behaviors are interrelated and appraised within a context [1]. From the psychological aspects, the space of the emotional state can be built from the discrete model or the dimensional model. In the discrete model, an emotional state is defined as a set of a finite number of discrete states corresponding to one of core emotions, including anger, fear, disgust, surprise, happiness, and sadness, or a combination of them [2]. The dimensional model defines an emotional state spatially with the basic dimensions of emotion such as valence and arousal and interprets an emotion through the levels of each dimension [3]. These emotion models have been used for systematical and multilateral analyses of emotion [3]. Based on the emotion models, neurophysiologic mechanisms under the emotional state have been vigorously investigated. Broadly, it has been documented that the emotional processes performed at the ventral and dorsal systems in the human brain are functionally different [4]. The ventral system, including ventral anterior cingulate gyrus and some ventral areas of prefrontal cortex (ventromedial prefrontal cortex and medial orbitofrontal cortex), is involved in the production of emotional states and the regulation of affective responses, whereas the dorsal system, including dorsal anterior cingulate gyrus, some dorsal areas of prefrontal cortex (dorsolateral, posterior dorsolateral, and mid-dorsolateral prefrontal cortex), and hippocampus, is involved in effortless emotion regulation and subsequent behavior [4, 5].

Recently, affective computing (AC) has emerged as a converging technology blending emotion into human computer interaction (HCI) [6]. AC, often called emotion aware computing, builds emotional interactions between a human and a computer by measuring the emotional state through behavioral and physiological signals and developing computational models for the emotional state [6, 7]. One of the key elements in AC is emotion recognition that estimates the emotional state of users from their behavioral and physiological responses [7]. Emotion recognition aims to advance
the intelligence of computer for creating affective user interfaces and to enhance the quality of psychiatric health care.

A variety of measures have been used for emotion recognition including self-report, startle response, behavioral response, autonomic measurement, and neurophysiologic measurement [3]. Self-report readily acquires emotional responses according to the emotion modeling framework but makes it difficult to track rapid affective changes and needs to rely on the outcome from self-estimation of the emotional state [3, 8]. The startle response magnitude using electromyography (EMG) measures unconscious myoneural responses but assesses only partial aspects of emotion (e.g., arousal level) [3, 9]. Behavioral measurement detects changes in facial and/or whole-body behavior using EMG or video image, but needs an assumption that EMG signals directly correspond to a specific emotional state [3, 10]. Autonomic measurement can objectively detect emotion-related physiological responses of autonomic nervous system (ANS), such as skin conductance responses (SCRs) and heart rate variability (HRV), but only access the subspaces of the emotional state [3, 11]. Neurophysiologic measurement based on electrophysiological and neuroimaging techniques can detect a wide range of dynamics of the emotional state by directly accessing the fundamental structure in the brain from which an emotional state emerges [3, 12]. Hence, neurophysiologic measurements clearly provide the most direct and comprehensive means for emotion recognition.

A large body of research has investigated neural correlates of emotion in humans using many noninvasive sensor modalities, each presenting unique characteristics with respect to spatiotemporal resolution and mobility. Functional magnetic resonance imaging (fMRI) has been used to find cortical and subcortical structures implicated in emotional states [13]. MEG has also been used to find emotion-related neural signals from specific sources in a timely manner with fine spatial and temporal resolutions [14]. But the cost and immobility of fMRI and MEG prevents these modalities from being used for practical emotion recognition systems [15, 16]. EEG, although suffering from its poor spatial resolution and high susceptibility to noise, has been widely used to investigate the brain dynamics relative to emotion as it enables the detection of immediate responses to emotional stimuli with an excellent temporal resolution [17–21]. Being developed to become more cost-effective and mobile with increased practicability and less physical restriction [22], EEG, not without its downsides, still carries critical advantages in practical usage and therefore has been a primary option for the development of online emotion recognition systems. In fact, there have been a growing number of efforts to recognize a person’s emotion in real time using EEG. For example, EmoRate developed as a commercial product (Emotiv Corp., CA, USA) detects the flow of the emotional state while user is watching a film [23]. Brown et al. proposed an EEG-based affective computer system that measures the state of valence and transmits it via a wireless link [24].

The development of an EEG-based emotion recognition system requires computational models that describe how the emotional state is represented in EEG signals and how one can estimate an emotional state from EEG signals. Despite a long history of searching for EEG indices of emotion, less attention has been paid to the computational models for emotional state estimation. Hence, we feel needs for a review of the state-of-the-art computational models for emotional state estimation to subserve the development of advanced emotion recognition methods. This paper will review the current computational methods of emotional state estimation from the human EEG with discussion on challenges and some future directions.

This paper will particularly focus on the following aspects of EEG-based emotional state estimation models. First, it will start with a quick review on EEG correlates of emotion, including definition of the emotional state space, the design of emotional stimuli, and the EEG indices of emotion. Then, it will revisit the computational methods to extract EEG features relative to emotional states and to estimate emotional states from EEG. We will also propose a mathematical approach to the estimation of continuous emotional state based on Bayesian inference.

2. EEG Correlates of Emotion

Finding EEG correlates of emotional states should begin with how to define the emotional state space. The emotional state space can be largely categorized into a discrete space and a continuous space. The discrete state space draws upon the discrete emotion model and contains a set of discrete experiential emotional states. The discrete emotional state comprises seven to ten core emotions such as happiness, surprise, sadness, anger, disgust, contempt, and fear [2, 25] and sometimes expands to contain a large number of emotions with the synonyms of these core emotions [25]. The continuous state space is built from the dimensional emotion model and represents an emotional state as a vector in a multidimensional space. This vector space of the continuous emotional state depends on the definition a basis. For instance, the circumplex model, developed by Russell, describes an emotional state in a two-dimensional circular space with the arousal and valence dimensions [26]. Various psychological models define emotional dimensions that subsequently constitute the basis for the emotional state space [25, 27–30].

Based on the construction of the emotional state space, the investigation of EEG correlates of emotion should also address how to determine experimental stimuli to induce emotions. Typically, emotional stimuli are selected to cover desired arousal levels and valence states, and presented in different modalities including the visual, auditory, tactile, or odor stimulation. The ground truth of the emotional state induced by a stimulus is secured by exploiting the self-ratings of subjects or using the standard stimulus sets such as the international affective picture system (IAPS) or the international affective digitized sound system (IADS). The IAPS provides a set of normative pictures for emotional stimuli to induce emotional changes and attention levels [31]. The IADS embodies acoustic stimuli to induce emotions, sometimes together with the IAPS [32]. These international
affective systems are known to be independent of culture, sex, and age [33].

A number of neuropsychological studies have reported EEG correlates of emotion. These EEG features can be broadly placed in one of two domains: time domain and frequency domain. In the time domain, several components of event-related potentials (ERPs) reflected underlying emotional states [34]. These ERP components can be encapsulated in a chronological order: P1 and N1 components generated in a short latency from stimulus onset, N2 and P2 in a middle latency, and P3 and slow cortical potential (SCP) in a long latency. The ERP components of short to middle latencies have been shown to correlate with valence [34–37], whereas with the ERP components of middle to long latencies have been shown to correlate with arousal [38–41]. Basically, the computation of ERPs requires averaging EEG signals over multiple trials, rendering ERP features inappropriate for online computing. However, recent developments of the single-trial ERP computation methods increase a possibility to use ERP features for online emotional state estimation [42–46].

In the frequency domain, the spectral power in various frequency bands has been implicated in the emotional state. The alpha power varied with the valence state [47] or with discrete emotions such as happiness, sadness, and fear [18]. Specifically, the frontal asymmetry of the alpha power has been repeatedly reported as a steady correlate of valence [48]. The subsequent studies have suggested that the frontal alpha asymmetry may reflect the approach/avoidance aspects of emotion, rather than valence per se [49]. The event-related synchronization (ERS) and desynchronization (ERD) of the gamma power has been related to some emotions such as happiness and sadness [50–52]. The ERS of the theta power has also been modulated during transitions in the emotional state [18, 53–55].

Besides the waveforms and the spectral power, the interactive properties between a pair of EEG oscillations such as phase synchronization and coherence have also been implicated in emotional processes. For instance, the phase synchronization level between the frontal and right temporoparietal areas varied with the emotional states of phase synchronization level between the frontal and right postfrontal areas [56]. The EEG coherence across the prefrontal and posterior beta oscillations was increased by viewing high arousal images [57]. Also, increases in the gamma phase synchronization index were induced by unpleasant visual stimuli [58]. As the emotional process engages a large-scale network of the neural structures in the brain, these multichannel analyses of EEG across the brain will reveal more signatures of emotion as they do for other cognitive functions [59–64]. In short, a brief summary of the EEG correlates of emotion is presented in Table 1.

### 3. Computational Methods to Estimate Emotional States

The computational methods to estimate the emotional state have been designed based on various EEG features related to emotional processes. As most EEG analysis methods are accompanied by preprocesses for reducing the artifacts, so is the emotional state estimation method. Figure 1 illustrates the overall processing steps to estimate the emotional state from EEG signals. The recorded EEG signals in response to affective stimuli pass through the preprocessing step in which noise reduction algorithms and spatiotemporal filtering methods are employed to enhance the signal-to-noise power ratio (SNR). Then, the feature extraction step determines specific band powers, ERPs, and phase coupling indices correlated with the target emotional states. Usually, this feature selection process is optimized by mathematical methods to achieve maximum emotional estimation accuracy. The classification step estimates the most probable emotional state from the selected EEG features. The number of class depends on the definition of the emotional state space, such as the continuous state of arousal and valence, or the discrete states.

As the preprocessing methods are relatively general to a variety of EEG signal processing applications, here we focus on the feature extraction and emotion classification methods. We first review the computational methods to extract emotion-related features from EEG, followed by the classification algorithms used to estimate the emotional state from the EEG features. The feature extraction methods usually build a computational model to find emotion-related features based on neurophysiologic and neuropsychological knowledge. Unlike the feature extraction methods, the classification methods draw more upon signal processing theories such as machine learning and statistical signal processing. It has been of interest how each of these two steps impact on estimation accuracy. On one hand, the feature extraction

![Figure 1: Overall emotional state estimation process. The overall emotional state estimation procedure. EEG signals are recorded during emotional situations and passed through the preprocessing step including noise reduction and spatial and temporal filtering. The features related with the emotional states such as spectral power, ERP, and phase synchronization are extracted from the preprocessed EEG signals. These features are used to estimate emotional states by classification methods.](image-url)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>#subjects</th>
<th>Stimulus (duration)</th>
<th>#EEG channels</th>
<th>Channel location</th>
<th>Emotional state</th>
<th>EEG features</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davidson et al. [48]</td>
<td>1990</td>
<td>37</td>
<td>Emotional film clips (60 sec)</td>
<td>8</td>
<td>F3, F4, C3, C4, T3, T4, P3, and P4</td>
<td>Happiness, disgust</td>
<td>Alpha power</td>
<td>Left-frontal: happiness &lt; disgust&lt;br&gt;Right-frontal: happiness &gt; disgust&lt;br&gt;Left-anteriortemporal: happiness &lt; disgust&lt;br&gt;Right-anteriortemporal: happiness &gt; disgust</td>
</tr>
<tr>
<td>Aftanas et al. [54]</td>
<td>2001</td>
<td>22</td>
<td>IAPS (7 sec)</td>
<td>128</td>
<td>IS</td>
<td>Valence (+/−)</td>
<td>Theta power (ERD/ERS)</td>
<td>Anterior temporal region (left-hemisphere)&lt;br&gt;Negative: left &lt; right&lt;br&gt;Positive: left &gt; right&lt;br&gt;Parietotemporal region&lt;br&gt;Negative: left &lt; right&lt;br&gt;Positive: left &lt; right</td>
</tr>
<tr>
<td>Keil et al. [51]</td>
<td>2001</td>
<td>10</td>
<td>IAPS (1 sec)</td>
<td>128</td>
<td>IS</td>
<td>Arousal</td>
<td>Gamma power</td>
<td>Gamma power (46–65 Hz, 500 ms): arousing ↑</td>
</tr>
<tr>
<td>Kemp et al. [104]</td>
<td>2002</td>
<td>16</td>
<td>IAPS (13 Hz)</td>
<td>64</td>
<td>IS</td>
<td>Valence (+/−)</td>
<td>SSVEP amplitude</td>
<td>Negative: SSVEP ↓ at the bilateral anterior frontal area</td>
</tr>
<tr>
<td>Pollatos et al. [105]</td>
<td>2005</td>
<td>44</td>
<td>IAPS (6 sec)</td>
<td>61</td>
<td>IS</td>
<td>Arousal</td>
<td>ERP/ECG</td>
<td>Good heartbeat perceivers show higher P300 peak</td>
</tr>
<tr>
<td>Balcon &amp; Lucchiari [18]</td>
<td>2006</td>
<td>20</td>
<td>Ekman's picture set (500 ms)</td>
<td>14</td>
<td>Fz, Cz, Pz, Oz, F3, F4, C3, C4, T3, T4, P3, P4, O1, and O2</td>
<td>Neutral versus emotions (happy, sad, angry, and, fearful)</td>
<td>ERD (alpha, beta, delta, and theta)</td>
<td>ERD% of theta (150–250 ms) at the anterior regions&lt;br&gt;Emotional state &gt; neutral state</td>
</tr>
<tr>
<td>Baumgartner et al. [106]</td>
<td>2006</td>
<td>24</td>
<td>IAPS + music (16 pictures/4.375 s)</td>
<td>16</td>
<td>F7, F3, FT7, FC3, F4, F8, FC4, FT8, TP7, CP3, P7, P3, CP4, TP8, P4, and P8</td>
<td>Happiness, sadness, and fear</td>
<td>Alpha power</td>
<td>Combining music with pictures evokes more intensive emotional experience</td>
</tr>
<tr>
<td>Sammler et al. [55]</td>
<td>2007</td>
<td>18</td>
<td>Music (22–44 sec)</td>
<td>63</td>
<td>IS</td>
<td>Valence (+/−)</td>
<td>Theta power</td>
<td>Frontal midline theta power is increased by positive emotion</td>
</tr>
<tr>
<td>Balcon &amp; Mazza [47]</td>
<td>2009</td>
<td>19</td>
<td>Ekman's picture set (30 ms, 200 ms)</td>
<td>32</td>
<td>IS</td>
<td>Anger, fear, surprise, disgust, and happiness</td>
<td>Alpha power (ERD)</td>
<td>Right-frontal activity increase for negative emotions</td>
</tr>
<tr>
<td>Li and Lu [107]</td>
<td>2009</td>
<td>10</td>
<td>Picture (6 sec)</td>
<td>62</td>
<td>IS</td>
<td>Happiness, sadness</td>
<td>Gamma ERD</td>
<td>Gamma ERD for emotional stimuli</td>
</tr>
<tr>
<td>Lithari et al. [108]</td>
<td>2010</td>
<td>28</td>
<td>IAPS (2 sec)</td>
<td>19</td>
<td>IS</td>
<td>Arousal</td>
<td>ERP (N100, N200)</td>
<td>ERP peaks increase for unpleasant stimuli in female</td>
</tr>
<tr>
<td>Petrantonakis [69]</td>
<td>2011</td>
<td>16</td>
<td>IAPS (5 sec)</td>
<td>8</td>
<td>F3, F4, C3, C4, T3, T4, P3, P4</td>
<td>Arousal/valence</td>
<td>Alpha and beta power</td>
<td>Asymmetry index can detect arousal levels</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>#subjects</td>
<td>Stimulus (duration)</td>
<td>#EEG channels</td>
<td>Channel location</td>
<td>Emotional state</td>
<td>EEG features</td>
<td>Effects</td>
</tr>
<tr>
<td>--------------------</td>
<td>------</td>
<td>-----------</td>
<td>--------------------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>----------------------------------</td>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Park et al. [109]</td>
<td>2011</td>
<td>34</td>
<td>Video: comic, horror, sadness, and peaceful (10 min)</td>
<td>32</td>
<td>IS</td>
<td>Fear, sadness, peace, and happiness</td>
<td>Alpha, beta, and gamma</td>
<td>Fear emotion: beta wave ↑ at the left temporal lobe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>happy emotion: alpha wave ↓ at the C4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>peaceful state: gamma wave ↑ at the T5, alpha wave ↓ at the CP5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sadness emotion: alpha wave ↓ at the left temporal lobe</td>
</tr>
<tr>
<td>Degabriele et al. [110]</td>
<td>2011</td>
<td>18</td>
<td>Emotional go/no-go inhibition task (300 ms)</td>
<td>24</td>
<td>IS</td>
<td>Happiness, sadness</td>
<td>ERP (P100 and N170)</td>
<td>P100 amplitude: happiness &gt; sadness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N170 amplitude: bipolar disorder is less at all emotional state than normal subjects</td>
</tr>
</tbody>
</table>
seems to be more closely tied to estimation performance since without pointing to the very features correlated with emotion, it is implausible to build a correct model. On the other hand, the classification algorithms should also be carefully designed to fit to the characteristics of the feature space; for instance, using a linear classifier for highly nonlinear feature structures would not make much sense. In general, one should weigh coherence between a feature space and a classifier for increasing estimation accuracy.

3.1. Feature Extraction Methods. As for valence-related features, it has been shown that positive and negative emotions induce asymmetric modulations in the frontal alpha power of EEG, leading to a relative decrease in the left frontal alpha power for positive emotions and a decrease in the right for negative emotions [65]. This frontal alpha asymmetry provides an effective index for valence by computing a difference between the left and right alpha powers, here denoted as $L$ and $R$ respectively, divided by the sum of both:

$$\text{Index} = \frac{L - R}{L + R}.$$ \hspace{1cm} (1)

The computation of the spectral power in the alpha band has been executed by a number of methods, including the squares of the EEG amplitude filtered through an alpha bandpass filter [53], Fourier transform [66], power spectral density [18, 21], and wavelet transform [7, 67, 68]. Most of these methods are well established and can readily be implemented in real time.

As for arousal-related features, one can extract the spectral power features such as the frontal midline theta power similar to the alpha power. Recently, more advanced computational methods have been proposed to evaluate emotional arousal. For instance, Asymmetry index (AsI) assesses emotion elicitation by computing a multidimensional directed information (MDI) between EEG channels [69]:

$$S_n = S_n^{XY} + S_n^{YX},$$

$$S_f = S_f^{XY} + S_f^{YX},$$ \hspace{1cm} (2)

$$\text{AsI} = (S_n - S_f) \times \sqrt{2}.$$  

$S_n^{XY}$ indicates the total amount of information flowing from left hemisphere signals, $X$, to right hemisphere signals, $Y$, when the subject has emotional feelings. $S_f$ refers to the total bidirectional information with emotion. $S_n^{XY}$ indicates the same directional information from $X$ to $Y$ but when the subject does not have emotional feelings, and $S_n$ refers to the total bidirectional information without emotion. AsI can effectively indicate whether an emotional state is elicited or not [69]. Besides AsI, the variance of potentials from a specific channel over different EEG channels has been used as an emotion-related feature [68]. Also, the entropy of EEG signals has been used to extract information related to emotion from intrusive noise [68].

As for individual discrete emotions, a typical approach is to search through all the possible EEG channels, spectral bands, and time segments for a set of features that maximizes the accuracy of emotional state estimation. This approach adopts a greedy search method with supervised learning, often resulting in different optimal feature sets for each individual. To overcome this issue of subject-by-subject variability, a higher order crossing (HOC) analysis was developed to implement a user-independent emotion recognition system [70]. The HOC analysis aims to find EEG features with respect to six affective traits, including surprise, disgust, anger, fear, happiness, and sadness [70]. The HOC model is given as:

$$D_n = \sum_{i=2}^{m} (X_i[n] - X_{i-1}[n])^2,$$

$$i = 1, \ldots, m; \quad n = 1, 2, 3, \ldots,$$

$$FV_{\text{HOC}} = [D_1, D_2, \ldots, D_L], \quad 1 < L \leq \text{Maximum order}.$$ \hspace{1cm} (3)

$D_n$ is the simplified version of the HOC feature that counts the number of zero-crossing from a high-pass filtered, standardized EEG time series. Zero-crossing indicates an event at which the signal amplitude passed through a zero-line with the change of polarity. The zero-crossing counts often represent oscillation properties more robustly than the spectral power. A vector of the simple HOCs is constructed to contain the features related to emotion. A higher value of $n$ means decreases in the discrimination power of the simple HOC because different processes can yield almost the same $D_n$. $X_i[n]$ indicates a binary time series with zeros and ones: at time instant $i$ where $X_i[n] = 0$ if the amplitude of the filtered signal is negative and $X_i[n] = 1$ otherwise. $m$ indicates the length of the time series $X_i$. The EEG feature vector is defined as $FV_{\text{HOC}}$ that consists of multiple simple HOCs [70]. The computational methods to extract emotional features from EEG are summarized in Table 2.

3.2. Emotion Classification Methods. The EEG feature vector provides observations from which an emotional state can be inferred. Commonly, a classifier has been used for decoding the feature vector into one of possible emotional states. A number of classification methods have been used for emotional state estimation, including discriminant analysis (DA), support vector machine (SVM), k-nearest neighbor (k-NN), and the Mahalanobis distance (MD) based method. DA performs dimensionality reduction in a high-dimensional feature space onto a low-dimensional space with an aim to maximize the Fisher discriminant ratio, $S$, of between-class scatter, $C_B$, to within-class scatter, $C_W$, [42, 71–76].

$$S = \text{tr} \left( C_B C_W^{-1} \right).$$ \hspace{1cm} (4)

A larger $S$ value indicates greater separation between classes. The dimensionality of the low-dimensional space varies from one up to the number of classes minus one.

SVM is derived from DA but determines a decision boundary in a kernel space instead of the original feature space. SVM finds an optimal hyperplane, $H(x)$, and the hypermargin of the decision boundary in the feature space.
Table 2: Emotional state estimation model.

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Stimulus (duration)</th>
<th>#channels</th>
<th>Channel location</th>
<th>Emotional states</th>
<th>EEG features</th>
<th>Feature extraction methods</th>
<th>Classifier</th>
<th>Accuracy (%)/classes</th>
<th>Online versus offline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li and Lu [107]</td>
<td>Picture (6 sec)</td>
<td>62</td>
<td>IS</td>
<td>Happiness, sadness, surprise, fear, disgust, and neutral</td>
<td>Gamma ERD, Delta, theta, alpha, beta, and gamma</td>
<td>CSP, Wavelet transform</td>
<td>SVM</td>
<td>93.5/2</td>
<td>Offline</td>
</tr>
<tr>
<td>Murugappan et al. [81]</td>
<td>Video (−)</td>
<td>62</td>
<td>IS</td>
<td>Joy, anger, sadness, and pleasure</td>
<td>Delta, theta, alpha, beta, and gamma</td>
<td>Short-time Fourier transform</td>
<td>kNN, LDA</td>
<td>83.04/5 (62 channels)</td>
<td>Offline</td>
</tr>
<tr>
<td>Lin et al. [111]</td>
<td>Music (30 sec)</td>
<td>24</td>
<td>Fp1-Fp2, F7-F8, F3-F4, FT7-FT8, FC3-FC4, T7-T8, P7-P8, C3-C4, TP7-TP8, CP3-CP4, P3-P4, O1-O2 (bipolar)</td>
<td>Joy, anger, sadness, and pleasure</td>
<td>Delta, theta, alpha, beta, and gamma</td>
<td>Short-time Fourier transform</td>
<td>SVM</td>
<td>82.29/4</td>
<td>Offline</td>
</tr>
<tr>
<td>Petrantonakis and Hadjileontiadis [70]</td>
<td>Ekman's picture set (−)</td>
<td>3</td>
<td>Fp1, Fp2, F3/F4 (bipolar)</td>
<td>Happiness, sadness, surprise, anger, fear, and disgust</td>
<td>ERD/ERS, Delta, theta, alpha, beta, and gamma</td>
<td>Higher-order crossing, Wavelet coefficients, Higuchi's algorithm</td>
<td>QDA, SVM</td>
<td>62.3/6 (QDA)</td>
<td>Offline</td>
</tr>
<tr>
<td>Hosseini et al. [112]</td>
<td>IAPS (3 sec)</td>
<td>5</td>
<td>FP1, FP2, T3, T4, Pz</td>
<td>Calm-neutral, negative-excited</td>
<td>Delta, theta, alpha, beta, and gamma</td>
<td>Wavelet coefficients, Higuchi's algorithm</td>
<td>Elman ANN</td>
<td>82.7/2</td>
<td>Offline</td>
</tr>
<tr>
<td>Wang et al. [113]</td>
<td>Video (4 min)</td>
<td>62</td>
<td>IS</td>
<td>Joy, relax, sad, and fear</td>
<td>Alpha, beta, and gamma</td>
<td>Minimum redundancy maximum relevance method</td>
<td>SVM</td>
<td>66.5/4</td>
<td>Offline</td>
</tr>
<tr>
<td>Brown et al. [24]</td>
<td>IAPS (6 sec)</td>
<td>8</td>
<td>Fp1, Fp2, F3, F4, F7, F8, C3, C4</td>
<td>Positive, negative, and neutral</td>
<td>Alpha1 (6–8 Hz), Alpha2 (8–10 Hz), Alpha3 (10–12 Hz)</td>
<td>Spectral power features</td>
<td>kNN</td>
<td>85.0/3</td>
<td>Online</td>
</tr>
<tr>
<td>Petrantonakis [69]</td>
<td>IAPS (5 sec)</td>
<td>8</td>
<td>F3, F4, C3, C4, T3, T4, P3, P4</td>
<td>Arousal, valence</td>
<td>Alpha and beta power</td>
<td>Asymmetry index</td>
<td>SVM</td>
<td>94.4/2</td>
<td>Offline</td>
</tr>
<tr>
<td>Stelios and Hadjidimitriou [114]</td>
<td>Music (60 sec)</td>
<td>14</td>
<td>AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4</td>
<td>Arousal, neutral</td>
<td>Beta and gamma</td>
<td>Time-frequency (TF) analysis</td>
<td>SVM, KNN, QDA, MDA</td>
<td>86.52/2</td>
<td>Offline</td>
</tr>
<tr>
<td>Konstantinidis et al. [22]</td>
<td>IAPS (2D emotional space)</td>
<td>19</td>
<td>IS</td>
<td>Arousal, valence</td>
<td>ERP (N100, N200)</td>
<td>Amplitude of ERP</td>
<td>SVM</td>
<td>81.3/4</td>
<td>Online</td>
</tr>
</tbody>
</table>

(i) SVM: support vector machine; kNN: k-nearest neighbor; LDA: linear discriminant analysis; QDA: quadratic discriminant analysis; MDA: Mahalanobis distance based discriminant analysis; SOM: self-organization map; ANN: artificial neural networks.
(ii) IS: channel configuration followed the 10/20 International System.
using a supervised learning method. The classifier $u(x_i)$ classifies a new input feature vector $x_i$ using a classification rule given by

$$u (x_i) = \text{sgn} \left( \sum_{v \in v} a_t F (x_i, x_j) + b \right), \quad t_i \in \text{target space}. \quad (5)$$

Here, $v$ indicates a set of the support vectors that are used to determine the maximum margin hyperplane, and $F(x_i, x_j)$ denotes the kernel function of the SVM classifier. $b$ denotes an offset parameter, $x_i$ does training input vectors, and $a_t$ does nonzero weights on each support vector [7, 77–80]. Various kernel functions have been proposed such as the Gaussian function or polynomials. SVM offers advantages of good generalizability for nonlinear feature spaces.

The k-NN algorithm determines the class of a new feature vector according to the number of nearest vectors in the training set surrounding a new feature vector [73, 81]. $k$ is a parameter determining the encircled boundary. The k-NN algorithm depends on how to define a distance between feature vectors, which is subject to be affected by the curse of dimensionality [81, 82].

The MD-based method, has been widely used in the clustering analysis, not only for distance, but also with correlation coefficient and the standard deviation [83, 84]:

$$\text{MD} = (x - \mu_c)^T C_c^{-1} (x - \mu_c). \quad (6)$$

$C_c^{-1}$ and $\mu_c$ indicate the inverse of the covariance matrix and the mean vector of a class $c$, respectively. MD converges to Euclidean distance when the covariance matrix of feature vectors becomes the identity matrix [84]. Basically, when a new feature vector arrives, the MD-based classifier compares the distance of the vector to each class using MD and chooses the class with the smallest distance. The classification methods that have been used for emotional state estimation are summarized in Table 2.

4. A Generative Model for Online Tracking of Emotional States

As described earlier, most computational models estimating emotional states have focused on the discrete state space and classified EEG features into one of a finite number of emotional states. This approach generally suits well to the case of a static determination of which emotion is induced by a given stimulus. Yet, for the development of an online emotion recognition system, where continuous tracking of the emotional state may play an important role, the current approach might be suboptimal because they do not take temporal dynamics of the emotional state into account. Another downside of the current approach originates from their direct modeling framework. A model in this framework builds a direct input-output mapping from the observed EEG signal to the emotional state. Although this framework may be able to provide a reasonable solution just for the purpose of improving classification accuracy, it does not exploit prior information of the emotional state as well as dynamics of the emotional state. These shortcomings make it difficult to gain useful insights on the neural mechanism of emotion. Also, it is often desirable to incorporate prior information of the dynamics of the emotional state within a model, especially for tracking emotional state continuously over time.

To address these issues, we propose a computational modeling approach based on the generative modeling framework [85–87]. Our approach focuses on tracking the change of the emotional state over time from EEG signal. In this approach, a generative model depicts how EEG signal is generated from a hidden emotional state. Also, a prior model explains how the emotional state changes over time. Integrating these two models, we infer a most likely emotional state from an observed EEG signal. Differences between the generative and direct models can be illustrated in a probabilistic view where a goal is to estimate a conditional probability of emotional state variables given EEG observations as accurately as possible. Suppose that a random vector $x$ denotes hidden emotional states and a random vector $y$ denotes observed EEG data (e.g., an EEG feature vector). An estimation model aims to optimize a parameter set, $\theta$, for the following conditional probability:

$$p (x | y, \theta). \quad (7)$$

A direct model forms a functional relationship from $y$ to $x$ with $\theta$, the parameter set of a function $f(\cdot)$,

$$x = f (y; \theta) + e, \quad (8)$$

where $e$ is a residual vector. In many cases, the residual vector is assumed to follow the Gaussian distribution. Parameter estimation of $\theta$ can be accomplished by many standard solutions such as maximum likelihood [88]. On the other hand, a generative model uses maximum a posteriori (MAP) or the Bayesian inference to estimate the conditional probability,

$$p (x | y, \theta) = \frac{1}{D} p (y | x, \theta) p (x), \quad (9)$$

where $D$ represents a constant representing the integral of $p(y)$. The posterior $p(x | y, \theta)$ is estimated by the product of $p(y | x, \theta)$, the likelihood of observation $y$ given a state $x$, and $p(x)$, the prior of the state $x$. The parameter set $\theta$ is used to model a generative relation from $x$ to $y$. In terms of the EEG correlate of the emotional state, the likelihood describes how the observed EEG signal is generated from an emotional state, the prior describes a probability of each emotional state, and the posterior describes which emotional state most likely elicits the observed EEG signal.

Here, we extend this generative approach to take into account the temporal dynamics of the emotional state. We use sequential Bayesian inference to track a time-varying emotional state from EEG signal [89]. To this end, we first assume that the emotional state is defined in a continuous space. An example of a continuous state space consists of two emotional dimensions, such as valence and arousal. The valence dimension ranges from negative values to positive values. The arousal dimension ranges from low to high.
arousal levels. A key point is that an emotional state varies over a continuous space, instead of altering between discrete values. This does not mean that we need to assign an explicit emotion to every possible point in the emotional state space. A specific area or volume in the state space can represent a single emotion.

The generative model is then formulated as follows. Let $x_t$ be an emotional state vector and $y_t$ an EEG signal vector at time instant $t$. $x_t$ contains a set of emotional state variables (e.g., $x_t = [x_{1t}, x_{2t}, x_{3t}]$, where $x_{1t}$ is the valence dimension, $x_{2t}$ is the arousal dimension, and $x_{3t}$ is the dominance dimension). $y_t$ contains a set of EEG features selected to be related to emotion (e.g., the power of certain frequency band at a selected channel). The goal of the model is to find the most probable emotional state given a series of observation from the beginning, $y_{1t}, \ldots, y_t$ (assuming observation begins at $t = 1$). The posterior is formed as

$$p(x_t | y_{1t}, \ldots, y_t).$$

The posterior can be rewritten as a recursive equation,

$$p(x_t | y_{1t}, \ldots, y_t) = p(y_t | x_t) \int p(x_t | x_{t-1}) p(x_{t-1} | y_{1t}, \ldots, y_{t-1}) dx_{t-1}. \tag{11}$$

Note that the likelihood, $p(y_t | x_t)$, depends only on the current time $t$. The prior, $p(x_t | y_{1t}, \ldots, y_{t-1})$, represents state transition from $t - 1$ to $t$, assuming the first-order Markov process. The dynamics of emotional state is embedded in the prior, whereas the generative process of the EEG features from an emotional state is modeled by the likelihood. The integral can be approximately computed by a number of methods with different model assumptions [89].

As this approximation relies on the recursion of the posterior, inference of an emotional state from EEG signal operates sequentially over time. This property enables our model fit well to the purpose of tracking emotional states continuously. In fact, the sequential Bayesian inference model (or called a Bayesian filter) has been widely adopted for many neuroengineering studies (e.g., see [90–94]). Our model may provide an effective way for online emotion aware computing, especially when we need to keep track of changes in the emotional state from EEG measurements continuously over time, for instance, tracking emotional changes while a subject is watching movies [95].

5. Discussion

In this paper, we overviewed the computational methods used for emotional state estimation. We first briefly gave an overview of the EEG correlates of emotion. Then, we revisited the computational methods to extract EEG features correlated with the continuous and discrete emotional states. We also described the classification methods to discriminate a particular emotional state from EEG features. Finally, we proposed a computational approach based on the generative modeling framework that may suit well to tracking the emotional state over time. These computational methods for emotional state estimation will serve as a key element for practical online emotion recognition systems for affective computing.

While affective computing has attracted attentions in the HCI field with a promise to develop a novel user interface, the development of the computational methods to estimate the emotional state still requires further understanding of emotion processes and their neurophysiologic substrates [96]. Especially, the estimation of emotional states from the human EEG has been posed only as a relatively simple classification problem with a few discrete emotions. The development of a real-time emotional state tracking system would require a more rigorous definition of the emotional state space suitable for estimation models.

Exploration of the EEG signatures of emotion that can span a broad area of the emotional state space or represent a number of different discrete emotions should continue. Such investigations may need to overcome many existing challenges. In particular, finding such EEG signatures of emotion that are invariant across individuals will be important for general emotion recognition systems [69]. As the emotion-related features have been mostly found in the frontal EEG signal, online algorithms to overcome the eye movement artifact should be continuously developed [97–99]. Also, bringing the EEG-based emotion recognition system out to the normal users would require a simple yet efficient EEG sensor. A new EEG sensor should meet some criteria such as stabilization of a signal to noise ratio (SNR), reduction of noise elicited from hair, optimization of active dry electrodes, development of multi-channel wireless communication, and sustainment of the quality of EEG signals over a long period [100–103]. Many previous studies have estimated the emotional state by analyzing the EEG responding to specific emotional stimuli. However, this emotion-induction paradigm has a limitation that the EEG signals can be modulated by the stimulus properties irrelevant to emotion [21]. Hence, a computational model that can predict the emotional state with various stimuli may be required for real-world applications.

The computational methods to estimate the emotional state may improve further with several advances in computational models. First, a model that can associate the dynamics of EEG signal with the dynamics of cognitive emotional process will provide a basis for constructing a novel emotional state estimation method. The current methods only capture the static properties in the EEG pattern in response to emotional stimuli. If a new model can embrace the temporal dynamics of emotional information processing in the human cognitive system and find EEG correlates of those dynamical properties, it will estimate the emotional state more precisely. Second, the quest for novel EEG signatures of the emotional state should be pursued. In particular, interactive properties between EEG signals such as cross-frequency coupling and effective connectivity pattern may be worth exploring to find novel EEG correlates of emotion. Third, inference of emotion-related information from high-dimensional and nonlinear EEG data poses an interesting problem to develop and apply the state-of-the-art
Conflict of Interests
The authors declare that there is no conflict of interests.

Acknowledgments
This work was supported by the National Research Foundation of Korea (NRF) Grants funded by the Korean government (2012047239 and 20120006588) and was funded by the Samsung Electronics Grant (R1210241).

References


