

Research Article

Multimodality Inferring of Human Cognitive States Based on Integration of Neuro-Fuzzy Network and Information Fusion Techniques

G. Yang,¹ Y. Lin,² and P. Bhattacharya³

¹College of Information Engineering, Central University for Nationalities, Beijing 100081, China

²Department of Mechanical and Industrial Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA

³Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada H3G 1M8

Correspondence should be addressed to Y. Lin, yilin@coe.neu.edu

Received 11 December 2006; Revised 25 April 2007; Accepted 9 August 2007

Recommended by Dimitrios Tzovaras

To achieve an effective and safe operation on the machine system where the human interacts with the machine mutually, there is a need for the machine to understand the human state, especially cognitive state, when the human's operation task demands an intensive cognitive activity. Due to a well-known fact with the human being, a highly uncertain cognitive state and behavior as well as expressions or cues, the recent trend to infer the human state is to consider multimodality features of the human operator. In this paper, we present a method for multimodality inferring of human cognitive states by integrating neuro-fuzzy network and information fusion techniques. To demonstrate the effectiveness of this method, we take the driver fatigue detection as an example. The proposed method has, in particular, the following new features. First, human expressions are classified into four categories: (i) casual or contextual feature, (ii) contact feature, (iii) contactless feature, and (iv) performance feature. Second, the fuzzy neural network technique, in particular Takagi-Sugeno-Kang (TSK) model, is employed to cope with uncertain behaviors. Third, the sensor fusion technique, in particular ordered weighted aggregation (OWA), is integrated with the TSK model in such a way that cues are taken as inputs to the TSK model, and then the outputs of the TSK are fused by the OWA which gives outputs corresponding to particular cognitive states under interest (e.g., fatigue). We call this method TSK-OWA. Validation of the TSK-OWA, performed in the Northeastern University vehicle drive simulator, has shown that the proposed method is promising to be a general tool for human cognitive state inferring and a special tool for the driver fatigue detection.

Copyright © 2008 G. Yang 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.

1. INTRODUCTION

Broadly speaking, any machine system involves human-machine interaction, for example, the vehicle system where the driver interacts with the vehicle in driving. In order to maintain an effective and save operation of the machine system, there is a need for the machine to understand the human state, especially cognitive state, when the human's operation task demands an intensive cognitive activity. To achieve this need is a complex task, warranting research. This is because the human being behaves in an extremely uncertain manner in terms of the correspondence between expressions and inferred cognitive states. For example, a person's smiling facial expression may not necessarily imply that the person is happy. Therefore, a new paradigm for techniques to under-

stand and measure the human cognitive state is to consider multimodality features of the human operator with a particular idea that both a feature and its context needs to be integrated in any inferring method. In this paper, we present a method for multimodality inferring of human cognitive states by integrating neuro-fuzzy network and information fusion techniques. To demonstrate the effectiveness of this method, we take the driver fatigue detection as an example due to its important social significance.

It is well known that the driver fatigue is responsible for a relatively high proportion of road traffic accidents. The United States National Highway Traffic Safety Administration (NHTSA) estimates that there are about 100 000 crashes every year caused by the fatigue that have led to more than 1 500 fatalities and 71 000 injuries [1]. Some other statistics

reported that drowsiness (a kind of fatigue) accounts for 16% of all kinds of crashes and over 20% of motorway crashes [2]. The driver fatigue has been notoriously called as the “Silent Killer” on the roads. Existing techniques for the driver fatigue detection can be classified into several categories according to literature [3], such as (1) causal/contextual feature, (2) physiological feature, (3) performance feature, and (4) combination of the above categories.

1.1. Casual/contextual features only

These features include (i) individual physical states such as sleep quality (SQ), and circadian rhythm; (ii) working conditions such as noises, and driving hours (DH); and (iii) environment conditions such as monotony of road (MR), and the number of lanes (NL). The inferring of fatigue based on these features is developed by first collecting feature data through questionnaire and then performing classifications. A questionnaire, including the required hours of sleep, difficulties in falling asleep at night, waking up tiredness, and waking up occasionally during the night, was designed for military truck drivers with the objective of finding a relation between fatigue and SQ [4]. This research concluded that the better SQ will lead to the less fatigue. In another study, twenty-six features in accident records were selected, and a neural network model was proposed by taking these features as inputs, and fatigue and nonfatigue as outputs [5]. A multistage evaluation method was applied in [6] using fuzzy set theory, in which fatigue was described as three states, namely, no fatigue, a bit fatigue, and complete fatigue. These studies [5, 6] need to be extended by including more levels of the fatigue.

1.2. Physiological features only

The physiological features are further grouped into the contact and contact-less features. The contact features mainly includes the brain activity, heart rate variability, and skin conductance which can be detected by electroencephalogram (EEG), electrocardiograph (ECG), and electromyogram (EMG). The contact-less features mainly include the eye movement (EM), head movement, and facial expressions which can be obtained from the dynamic images provided by the CCD camera. It is noted that the classification of the EM under the physiological features may be controversial; however, our interpretation of physiology here seems to be broader such that physiological features are those governed by the brain on a continuously updating basis. Nevertheless, this classification does not affect the main result of this research.

The classification of these two groups leads to two general methods: contact-feature-based method (CFBM) and contact-less-feature-based method (CLFBM), respectively. In the case of CFBM, an algorithm based on changes in all major EEG bands (delta, theta, alpha, and beta bands) during fatigue was developed in [7, 8]. Further, a combination of the EEG power spectrum estimation, principal component analysis, and fuzzy neural network model was used to predict the driver’s drowsiness in [8]. The associated wavelet representation of EEG at different scales was applied as system inputs

to detect the starting time the driver begins to feel fatigue in [9].

Besides EEG, the heart rate variability also contains abundant information about fatigue. Several ECG features such as low frequency (LF), very low frequency (VLF), high frequency (HF), and the LF/HF ratio were applied in [4] to classify sleep into wake, rapid eye movement (REM), and non-REM stages. By taking Hermite polynomial coefficients of ECG as input [10] of a neuro-fuzzy network, an approach [11] was proposed to classify the heart rate variation. Selecting the means, the standard deviations, the first differences, and the second difference of EMG, blood volume pulse (BVP), galvanic skin response (GSR), and respiration from the chest expansion as the physiological features, an algorithm was proposed which combines the sequential floating forward search and the fisher projection approaches [12, 13]. Although EEG and ECG have been thought to be accurate and objective to measure fatigue, it is very difficult to apply these two physiological signals in the real driving situation because electrodes and wires are used to contact a driver obtrusively in order to obtain EEG and ECG signals. It is noted that there have been some efforts in developing nonobtrusive EEG and ECG technologies, but they are not on the market yet.

In the case of CLFBM, the visual cues were almost exclusively employed. These visual cues mainly include mouth shape, head position, and eye movements (e.g., changes in the eye gaze direction, eyelid activity, and blinking rate, etc.) which can be extracted from a series of dynamic images provided by a CCD camera [14]. A driver fatigue detection algorithm has been proposed based on the eye tracking and dynamic template matching [15]. The detection of the gaze direction using the time-varying image processing has been studied in [16] where the facial direction and the gaze direction were detected separately, and then they were integrated into a final gaze direction. Taking the openness of mouth and eye, respectively, and the vertical distance between eyebrows and eyes as inputs, a fuzzy neural network model was constructed for detecting fatigue [17]. Percent eye closure (PERCLOS) methodology is a reliable technique for the determination of a driver’s alertness level. Grace et al. in Carnegie Mellon Research Institute developed a video-based system that measures PERCLOS [18]. Optalert patented technology, using the reflectance of invisible light to monitor the movements of eye and eyelids, is also a reliable technique for the determination of a driver’s alertness level [19].

1.3. Performance features only

There is an emerging consensus that fatigue will contribute to deterioration in performance, which may lead to errors and increase the risk of accidents [20]. This is true for driving. It is due to such a viewpoint that the method in this category is defined as being able to infer the fatigue onset by observing driver’s performance, mainly including the operational reaction time, lane position deviation, and hand movement of controlling the steering wheel. A method was proposed in [21–23] to model the driver’s motion behavior when controlling the steering wheel by using the fuzzy theory.

1.4. Combination of 1.1~1.3 using the multiple feature fusion technique

Each of methods in (1), (2), and (3) categories only focuses on certain aspects. While they may succeed in their own “perfect” conditions, unfortunately, these “perfect” conditions may not be practical, which therefore challenges the measurement reliability. For example, inferring driver’s fatigue from facial expression is not always reliable because of the two limitations. One is that current techniques of image processing cannot always ensure the recognition precision, the other is that an introverted person might have tendency of controlling his/her display of emotions, especially in the presence of people he/she is not well-acquainted with [24]. The performance-based measurement technique can easily be challenged because deterioration in driving performance may also be related to such factors as driver’s age, overtaking, or giving way to other cars.

The fundamental principle for solutions to these challenges is to “fuse” multiple kinds of signals of information about persons’ contexts, situations, goals, and preferences [12]. Along this line of thinking, a few studies have been reported. Considering the contextual information and visual cues at a single time instant, a static Bayesian net (SBN) has been constructed [1] to infer and predict the fatigue of human operators. Though their method does enhance measurement reliability, it was unable to model fatigue dynamically [25, 26]. The dynamic Bayesian network (DBN) has been developed to overcome this limitation. Considering the evidence and beliefs of contextual information and visual cues from multiple time slices, a probabilistic framework based on DBN has been introduced in [25]. However, it remains to see how the contact features affect the accuracy of measurement. There is a further general difficulty with the BN or DBN in determining the prior probability and conditional probability which are the important parameters in these models.

From the above analysis, a conclusion is perhaps made that the inferring of human cognitive states based on the fusion of multiple features is an effective way, especially for getting reliable fatigue estimation. In line with this conclusion, a method based on neuro-fuzzy network and information fusion techniques for inferring human mental states with a particular attention to the driver fatigue was proposed in a study to be presented in this paper. There are three salient features with the proposed method. First, the neuro-fuzzy network technique is employed for two reasons: (1) the behavior associated with fatigue is often vaguely described, for example, very tired, very sleepy, and so forth, to which the fuzzy logic is extremely suitable; (2) the neural network brings the low-level learning and computational power to a decision system for capturing the nonlinearity in the system behavior [27]. Second, the information fusion technique is employed in such a way that the cues are taken as inputs to the TSK model which gives outputs, and then they are fused by a particular fusing method which gives outputs corresponding to particular cognitive states under interest (e.g., fatigue). There are fruitful methods [28–36] available for aggregation of multiple features. Ordered weighted aggregation (OWA) method

[36] was selected in this study because of the following reason. There are many features related to fatigue; some have more contribution to the fatigue, while others have less contribution to the fatigue. In information fusion, it is natural that the feature with more contribution to the fatigue should have higher weight, and vice versa. OWA method does work well for this situation because the basic idea of the OWA is that the weights of aggregating variables are not fixed by the absolute values of the variables but by their relations. Third, the three categories of cues are employed, namely, (i) contextual category, (ii) contact category, and (iii) contact-less category. The proposed method is called TSK-OWA.

In addition to the new feature with the proposed method, that is, a combination of neuro-fuzzy network and information fusion techniques, another major difference of the proposed method other than other methods commented before is that none of them has considered the three categories together. In a closely related work [8], the neuro-fuzzy TSK model was employed for measuring fatigue; however, that work only considered the EEG signal. Further in that work, the final aggregation of several channels of information sources into one state has not considered the contribution variation of individual channels of information to that state.

The remainder of this paper is organized as follows. Section 2 will present a general architecture of the proposed method by taking the driver fatigue diction as an example. Section 3 presents the model based on the neuro-fuzzy theory with the features (SQ, DH, EEG, ECG, EM). In Section 4, the method for aggregating the outputs from the neural-fuzzy model is presented. Section 5 presents an experiment validation to the proposed method. Section 6 concludes the paper and discusses future work.

2. THE ARCHITECTURE OF THE PROPOSED METHOD

We take the driver fatigue diction as an example. As mentioned previously, there are many features related to fatigue. Some features may have more contribution to fatigue, while others may have less. In this study, we proposed that each category at least comes up with one feature that contributes to fatigue most. Having this idea in mind, in the following we discuss the section of features in relation to the degree of their relevance with fatigue.

2.1. SQ analysis

SQ is an important contextual feature that has an immediate relation with fatigue [4]. The driver’s SQ is further associated with such quantities as required sleep hours, difficulties in falling asleep at night, waking up tiredness, waking up occasionally during the night, waking up too early in the morning without being able to fall asleep again [4], and other social factors such as the economic burden of a family. Among them, the required sleep hour is taken as a key contributor to SQ because of its relatively high relevance to the degree of fatigue. It is known that an average human being requires 6 to 8 hours sleep per day for his or her normal operation. Another important reason to select the sleep hour as an indicator of

SQ is that the sleep hour is a crisp value and thus easy to obtain in a precise manner.

The hour of sleep is denoted as z_1 and normalized to the range of $[0,1]$ (i.e., $z_1 \in [0, 1]$) which is derived from the time interval $[0, 8]$ hours. Further, the SQ in this case is defined as a probabilistic variable, denoted as $y_1 \in [0, 1]$ corresponding to z_1 . In particular, $y_1 = 0$ means that the probability that a driver is fatigue is 0; that is to say that the driver is not fatigue at all. While $y_1 = 1$ means that a driver is completely or absolutely fatigue; in other words, the probability that the driver is fatigue is 1. The definition of the variable y applies, hereafter, to subsequent discussions in this paper.

2.2. DH analysis

As studies demonstrated, many factors such as long hours, time of day, sleep-related problems, the characteristics of road structure and roadside environment had impacts on driver's state when performing a driving task. However, not all variables can be controlled or examined in any single study [37]. Furthermore, the relevance of DH to the driver fatigue leading to traffic accidents has been already demonstrated by many studies (e.g., [6]). For example, it was pointed out that DH is not only one of the major contributors to fatigue but also one of the potential sources of inferring fatigue in a recent study [38]. Therefore, DH is adopted as a feature to describe fatigue in this paper without considering other factors such as the road structure and roadside environment (e.g., the road monotony). Just the same as the SQ analysis, denote the continuous driving hour z_2 normalized to $[0,1]$ (i.e., $z_2 \in [0, 1]$) derived from the time interval $[0, 12]$ hours). Denote y_2 as the probabilistic variable corresponding to z_2 .

2.3. EEG analysis

EEG is an important feature that has an immediate relation with fatigue; but EEG signals have to be preprocessed because of some artifacts and noises in the raw signals. In this study, the EEG signals first was smoothed by use of a simple low-pass filter with a cutoff frequency of 50 Hz to remove the line noise and other high-frequency noise mainly caused by muscle activity, and then the independent component analysis was employed to remove the artifacts such as EOG mainly created by the eye movement [8]. Finally, the smoothed signals are transformed into the frequency domain by use of the Fast Fourier Transform (FFT) algorithm [9]. The frequency domain includes delta band (0.5–4 Hz) corresponding to sleep activity, theta band (4–7 Hz) related with drowsiness, alpha band (8–13 Hz) corresponding to relaxation and creativity, and beta band (13–25 Hz) corresponding to activity and alertness [7, 8, 20, 39, 40]. Note that among these bands only the theta and alpha bands have strong associations with fatigue. Further, it is the decrease in the alpha and theta rhythms that shows a driver is at the fatigue state. The EEG contains signals from different channels.

In this study, two of these channels (i.e., two different EEG sites on the brain) were chosen [20]. Under a vigorous stage, the driver's average magnitudes of the signal within

the alpha and theta bands are taken as the standard baselines symbolized with \bar{z}_3 and \bar{z}_4 , respectively. In the fatigue situation, obvious changes of the alpha and theta signals around the standard baseline always take place. In this study, the differences denoted as z_3 (for the alpha band) and z_4 (for the theta band) between the baselines and the current magnitudes of the alpha and theta signals are taken as the features to describe fatigue. Given that there are P participants, and their magnitudes within the alpha and theta bands under the vigorous stage are \bar{z}_{ij}^3 and \bar{z}_{ij}^4 ($i = 1, 2, j = 1, 2, \dots, P$), respectively; the standard baselines are calculated with the following equations:

$$\begin{aligned}\bar{z}_3 &= \frac{1}{2} \sum_{i=1}^2 \frac{1}{P} \sum_{j=1}^P \bar{z}_{ij}^3, \\ \bar{z}_4 &= \frac{1}{2} \sum_{i=1}^2 \frac{1}{P} \sum_{j=1}^P \bar{z}_{ij}^4.\end{aligned}\tag{1}$$

The differences z_3 and z_4 are calculated with the following equations:

$$\begin{aligned}z_3 &= \frac{1}{2} \sum_{i=1}^2 \hat{z}_i^3 - \bar{z}_3, \\ z_4 &= \frac{1}{2} \sum_{i=1}^2 \hat{z}_i^4 - \bar{z}_4,\end{aligned}\tag{2}$$

where items \hat{z}_i^3 and \hat{z}_i^4 represent the alpha and the theta current magnitudes of the i th channel, respectively. Denote y_3 as the probabilistic variable corresponding to z_3 and z_4 .

2.4. ECG analysis

Heart rate variability (HRV) differs significantly for the same individual in different states such as alertness and fatigue. This is the primary reason why HRV is often used to detect driver's states. HRV spectrum shows 3 main components: LF, VLF, and HF. Among them is the LF/HF ratio which has a strong relation to driver's fatigue. It was pointed out in [41] that LF/HF ratio will decrease progressively when passing from the awake state to the fatigue state. To calculate the LF/HF ratio, it is necessary to detect the R-wave (the first positive (upward) deflection of the QRS complex in the electrocardiogram) peaks of the driver's ECG signal. In this study, we adopted wavelet transform (WT) to analyze the ECG signal because WT can provide a description of the signal both in the time and frequency domains. Especially, WT can characterize the local regularity of the ECG signal, which is useful to distinguish real signals from noises, artifacts, and drifts produced by vibration and muscle movements in realtime measurement. To apply WT, specifically, first, the quadratic spline wavelet function with WT was performed on the digital ECG signal. The QRS complex (the deflections in the tracing of the electrocardiogram, comprising the Q, R, and S waves, that represent the ventricular activity of the heart) of the digital ECG signal produces two modulus maxima with opposite signs among WT coefficients, which leads to a zero

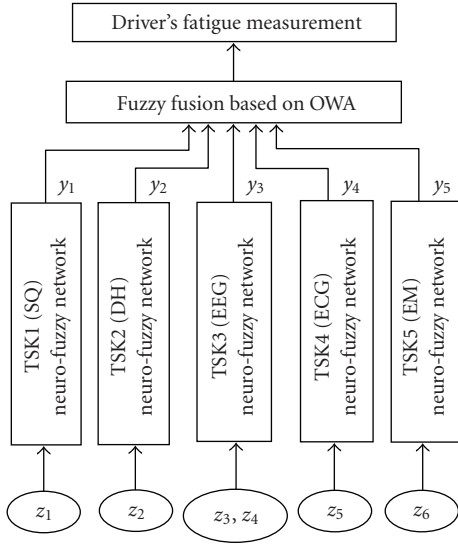


FIGURE 1: Structure of the proposed neuro-fuzzy fatigue recognition model.

crossing point between the two modulus maxima at each scale [42–44]. Consequently, the zero crossing point at the scale 2^4 is taken as the R-wave peak point [42–44], which results in HRV. Then, WT with a Haar wavelet function was performed on HRV, and the result is such that the sum of wavelet decomposition coefficients at 1 and 2 levels corresponds to LF, and the sum of wavelet decomposition coefficients at 3 and 4 levels corresponds to HF [45]. Therefore we can get the LF/HF ratio.

Under a normal condition, the LF/HF ratio is calculated as the standard baseline, and the differences between the baseline and the current LF/HF ratio is calculated, symbolized as z_5 . Denote y_4 as the driver's probabilistic state corresponding to z_5 .

2.5. EM analysis

Eye activity which can be characterized by the percentage of eye closure over a given time is one of the visual behaviors that reflect a driver's fatigue level. This can be demonstrated by the previous studies [1, 46] that the driver maybe is in fatigue as the eyes are at least 80 percent closed in a given time, and that PERCLOS has been found to be the most valid ocular parameter for monitoring fatigue. Therefore, the running average of PERCLOS instead of PERCLOS (to ensure the robustness of the PERCLOS measurement) is accepted as a feature to describe fatigue in this study. We use the normalized variable $z_6 \in [0, 1]$ to denote the running average of PERCLOS, and make the probabilistic variable y_5 correspond to z_6 .

To obtain z_6 , a CCD camera is fixed on the dashboard of the Northeastern University's virtual environments driver simulator to focus on the driver's face for detecting the multiple visual behaviors. The program continuously tracks the driver's pupil shape at each 2 seconds sampling time instance to determine the eye state (openness/closure) (for details,

please refer to [1]). In a given time (e.g., 30 sec), if the driver's eyes are closed continuously for p ($p = 0, 1, \dots, 15$) sampling time instances, and then $z_6 = 2 * p / 30$.

2.6. Summary of the proposed structure

In the above analysis, the SQ and DH fall into the contextual category, the EEG and ECG fall into the contact category, and the EM falls into the contact-less category. As such, there are five pair relations, namely, (z_i, y_i) ($i = 1, 2, 3, 4, 5$), and they are gathered into the architecture of the neuro-fuzzy TSK (Takagi-Sugeno-Kang) model [47] proposed in this study; see Figure 1. Each output y_i only partially reflects driver's fatigue from a certain aspect, which is not reliable to the fatigue measurement. OWA method is chose in this study to fuse the five fuzzy output variables in order to make the final fatigue measurement $y \in [0, 1]$ more reliable.

3. THE NEURO-FUZZY TSK NETWORK MODEL

3.1. Neuro-fuzzy TSK structure

Figure 1 shows that there are 5 neuro-fuzzy TSK subnetworks (named from TSK1 to TSK5) with different parameters but the same structure. Each of them is viewed as a multi-input and single output (MISO) fuzzy system (if a system has only one input and one output, the system is viewed as a special case of the MISO fuzzy system). Let us take one of the five MISO fuzzy systems as an example to explain the structure of the neuro-fuzzy TSK system.

Denote

$$\begin{aligned} y &= y_i, \\ \mathbf{x} &= z_i = [x_1, x_2, \dots, x_N]^T, \\ i &= 1, 2, 3, 4, 5 \end{aligned} \quad (3)$$

as the output value and input vector, respectively, where N is the number of the inputs, and i denotes the i th TSK model; $i = 1, 2, 3, 4, 5$ in this case. Suppose that M inference rules are available for the system. The general form of the k th ($k = 1, 2, \dots, M$) TSK inference rule can be stated as follows [27, 48–50],

$$\text{Rule } k : \text{ If } \mathbf{x} \text{ is } A_k \text{ then } y = f_k(\mathbf{x}), \quad (4)$$

where $f_k(x_1, \dots, x_N)$ is a crisp output function, and A_k is a fuzzy set labeled by a linguistic description (e.g., small, medium, or large).

The first question regarding (4) is how to specify the fuzzy set A_k . Generally speaking, the clustering techniques such as the fuzzy c -means (FCM) algorithm [50], the mountain method [51], and the hybrid clustering and gradient descent (HCGD) approach [52] are effective methods to get A_k from the input-output data available. In this study, HCGD with some modifications is taken because it can automatically generate a number of clusters and classify all input data points into different clusters without requiring any assumptions about the data points. The modified HCGD method works as follows.

Suppose that there are Q samples. Denote the i th input-output pair of samples as $\mathbf{s}_i = (x_1(i), x_2(i), \dots, x_N(i), y(i))^T$ ($i = 1, 2, \dots, Q$). We have the following steps.

Step 1. Define Q number of vectors \mathbf{v}_i ($i = 1, 2, \dots, Q$), and let $\mathbf{v}_i = \mathbf{s}_i$ (i.e., \mathbf{s}_i is the initial value of \mathbf{v}_i).

Step 2. Compute the potential function $h_{ij}(\mathbf{v}_i, \mathbf{v}_j)$ between \mathbf{v}_i and \mathbf{v}_j with the following equation:

$$h_{ij}(\mathbf{v}_i, \mathbf{v}_j) = \exp\left(-\frac{\|\mathbf{v}_i - \mathbf{v}_j\|^2}{2\alpha^2}\right), \quad (5)$$

$$i = 1, 2, \dots, Q, \quad j = 1, 2, \dots, Q,$$

where $\|\mathbf{v}_i - \mathbf{v}_j\|^2$ represents the Euclidean distance between \mathbf{v}_i and \mathbf{v}_j , and α is the width of the Gaussian function which is fixed by experiments.

Step 3. Calculate $\hat{\mathbf{v}}_i$ ($i = 1, 2, \dots, Q$) with the following equation:

$$\hat{\mathbf{v}}_i = \frac{\sum_{j=1}^Q h_{ij} \mathbf{v}_j}{\sum_{j=1}^Q h_{ij}}, \quad (6)$$

and check whether $\hat{\mathbf{v}}_i$ is close enough to \mathbf{v}_i for $i = 1, 2, \dots, Q$, that is,

$$\|\mathbf{v}_i - \hat{\mathbf{v}}_i\| \leq \varepsilon, \quad i = 1, 2, \dots, Q, \quad (7)$$

where ε is a very small positive number which has strong relations with the number of fuzzy sets and the computation load. Generally speaking, the number of fuzzy sets and the computation load increase with the decrease of ε . In most applications, ε is chosen empirically or experimentally. If (7) is satisfied, then go to the next step; otherwise, let $\mathbf{v}_i = \hat{\mathbf{v}}_i$ and go to Step 2.

Step 4. The original data with the same convergent vector is clustered into a cluster, and the number of convergent vectors is equal to the number of clusters. The convergent vector is the cluster center and expressed as

$$\bar{\mathbf{c}}_k = [c_{k1}, c_{k2}, \dots, c_{kN}]^T, \quad k = 1, 2, \dots, M. \quad (8)$$

Compared to the original HCGD [52], the modified HCGD as presented above has the following unique features.

- (1) In the whole iterative process, all of the potential function h_{ij} is taken into account in (6) and (7) no matter how big or small it is. In this way we could avoid the situation where contribution of particular h_{ij} to the convergent vector is excluded when h_{ij} is very small.
- (2) A somewhat ‘‘hard’’ stop criterion is imposed (see (7)) so that any dead-loop in the algorithm can be avoided.

Given that each cluster is associated with one independent inference rule, the centroid of each cluster is automatically assigned to the center of the premise of the rule. After the number of clusters is determined, one needs to specify the membership degree to which variable \mathbf{x} belongs to

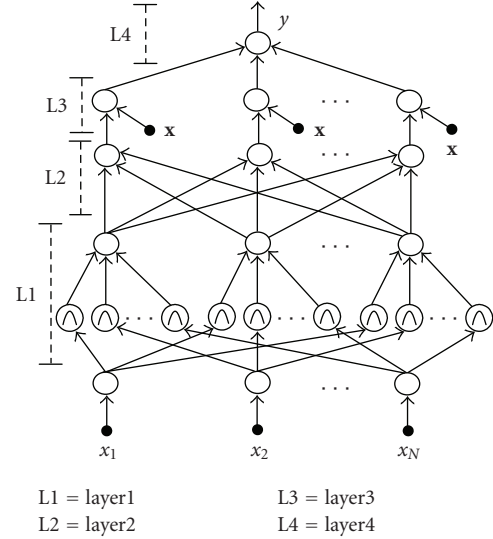


FIGURE 2: One-order neuro-fuzzy TSK network.

the fuzzy set A_k . There are many types of membership functions such as triangle-shape, trapezoidal-shape, bell-shape, and Gaussian membership functions. In this study, the Gaussian membership function was chosen because of its universal approximation and simple multidimensional decomposition [27, 49]. Thus, the premise (if \mathbf{x} is A_k) is described as

$$\mu_n^k(x_n) = \exp\left[-\frac{(x_n - c_{kn})^2}{2\sigma_{kn}^2}\right], \quad n = 1, 2, \dots, N, \quad (9)$$

where σ_{kn} is the width of the Gaussian membership function, which is further determined by the following equation [52]:

$$\sigma_{kn} = \sqrt{\frac{-\sum_{m=1}^N (x_m^* - c_{km})^2}{2 \ln(u)}}, \quad n = 1, 2, \dots, N, \quad (10)$$

where x^* is the farthest data point from the cluster center $\bar{\mathbf{c}}_k$, and $u \in [0.1, 0.3]$ [52]. The procedure as described above was implemented by the fuzzification corresponding to the first layer of the neuro-fuzzy subnetwork, as shown in Figure 2.

The second question regarding (4) is to determine the firing strength of the corresponding fuzzy rule. Let one node represent one fuzzy logic rule in the second layer and the output of the node represent the firing strength corresponding to the fuzzy rule. In this study, the AND operator [27] is chosen to determine the firing strength $\eta_i(\mathbf{x})$, that is,

$$\eta_k(\mathbf{x}) = \prod_{n=1}^N \mu_n^k(x_n) = \exp[-(D_k(\mathbf{x} - \bar{\mathbf{c}}_k))^T (D_k(\mathbf{x} - \bar{\mathbf{c}}_k))], \quad (11)$$

where $D_k = \text{diag}(1/\sigma_{k1}, 1/\sigma_{k2}, \dots, 1/\sigma_{kN})$, and $\bar{\mathbf{c}}_k = (c_{k1}, c_{k2}, \dots, c_{kN})$. The procedure as described above was implemented by the second layer of the neuro-fuzzy subnetwork, as shown in Figure 2.

The first-order TSK crisp output function is often employed to get the result of $f_k(x_1, \dots, x_N)$, which has the following form [49]:

$$f_k(x_1, \dots, x_N) = p_{k0} + \sum_{n=1}^N p_{kn}x_n, \quad (12)$$

where $p_{k0}, p_{k1}, \dots, p_{kN}$, are crisp numbers adjusted at the learning process. After having generated TSK functions f_k , the next step is to calculate the summation of f_k with a normalization procedure to produce the output y of TSK; see the following equations below [27, 49],

$$\begin{aligned} y(\mathbf{x}) &= \sum_{k=1}^M \omega_k f_k(\mathbf{x}) \\ &= \sum_{k=1}^M \omega_k \left[p_{k0} + \sum_{n=1}^N p_{kn}x_n \right], \quad (13) \\ \omega_k &= \frac{\eta_k(\mathbf{x})}{\sum_{m=1}^M \eta_m(\mathbf{x})}. \end{aligned}$$

The procedure as described above was implemented by the third and fourth layers of the neuro-fuzzy subnetwork, as shown in Figure 2.

3.2. Parameter identification of the neuro-fuzzy TSK network

After the structure of the neuro-fuzzy network model as described above is generated from the given input-output data pattern, the network parameters (i.e., the parameters in the TSK functions and the parameters in the Gaussian function) from the same input-output data pattern need to be determined. At this point, both feed-forward network and recurrent neural network can be used to achieve this purpose. The recurrent neural network is more suitable for the problems with highly non-linear dynamics, but it is computationally overhead. The feed-forward network (e.g., the back-propagation network) has extensively been used in the field of function approximation, pattern recognition, and pattern classification because of its computational efficiency, but it may have more chances to get a local minimum. The local minimum problem can usually be resolved by carefully selecting the initial weights of the neural network. Given that the nature of our application, discussed in this paper, is largely about the clustering and pattern recognition and the application demands a fast response, the back-propagation method is employed for learning in this study. In the following, several key steps of back-propagation algorithm for learning are presented.

Denote $y^d(t)$ and $y(t)$ as the desired and current outputs of the network at time t , respectively. In order to obtain the network parameters through learning, define a goal function E as follows:

$$E = \frac{1}{2} [y^d(t) - y(t)]^2. \quad (14)$$

For the convenience of description, denote h_k^ξ as the output of the ξ th node in the ζ th layer of the neuro-fuzzy network. In the last layer (the fourth layer), denote $h_1^4 = y(t)$ because there is only one node in this layer. According to the back-propagation method, the minimum of E corresponds to the determination of the network parameters, which is done iteratively with the following equations [27]:

$$\begin{aligned} p_{kn}(t+1) &= p_{kn}(t) + \alpha [h_1^4(t) - y^d(t)] h_k^2(t) x_n, \\ p_{k0}(t+1) &= p_{k0}(t) + \alpha [h_1^4(t) - y^d(t)] h_k^2(t), \\ c_{kn}(t+1) &= c_{kn}(t) - \alpha \frac{\partial E}{\partial h_1^4} \frac{\partial h_1^4}{\partial h_k^3} \frac{\partial h_k^3}{\partial h_k^2} \frac{\partial h_k^2}{\partial h_k^1} \frac{\partial h_k^1}{\partial c_{kn}}, \quad (15) \\ \sigma_{kn}(t+1) &= \sigma_{kn}(t) - \alpha \frac{\partial E}{\partial h_1^4} \frac{\partial h_1^4}{\partial h_k^3} \frac{\partial h_k^3}{\partial h_k^2} \frac{\partial h_k^2}{\partial h_k^1} \frac{\partial h_k^1}{\partial \sigma_{kn}}, \end{aligned}$$

where α is the learning rate.

4. SENSOR FUSION TECHNIQUE

4.1. Features available

As shown in Figure 1, SQ, DH, EEG, ECG, and EM are fed into neuro-fuzzy networks of TSK1, TSK2, TSK3, TSK4, and TSK5, respectively, resulting in the network outputs y_i ($i = 1, 2, \dots, 5$), denoted as $\mathbf{o} = [y_1, y_2, y_3, y_4, y_5]^T$. Let $\mathbf{w} = [w_1, w_2, w_3, w_4, w_5]^T$ denote the associated weight vector. Construct $\mathbf{b} = [b_1, b_2, b_3, b_4, b_5]^T$ such that b_i ($i = 1, 2, \dots, 5$) is the i th largest element of the collection of y_1, y_2, y_3, y_4 , and y_5 . According to the OWA method [33], y can be calculated by

$$\begin{aligned} y &= \mathbf{w}^T \mathbf{b} = \sum_{i=1}^5 w_i b_i, \\ 0 \leq w_i \leq 1, \quad i &= 1, 2, \dots, 5, \quad (16) \\ \sum_{i=1}^5 w_i &= 1. \end{aligned}$$

A number of techniques [28, 50, 53–55] are available to determine the weight vector \mathbf{w} of (16). In this study, we take a combined technique from the literature [53, 55].

Let $\hat{\mathbf{w}} = \{\hat{w}_i; (i = 1, 2, \dots, 5)\}$ be the estimation of \mathbf{w} , and specify [53]

$$\hat{w}_i = \frac{e^{\lambda_i}}{\sum_{j=1}^5 e^{\lambda_j}}, \quad i = 1, 2, \dots, 5. \quad (17)$$

In order to ensure the constraints of $0 \leq \hat{w}_i \leq 1$ ($i = 1, 2, \dots, 5$) and $\sum \hat{w}_i = 1$, λ_i is taken as the unknown parameter to be determined in the learning process. There are k outputs of the neuro-fuzzy TSK network, denoted by $\mathbf{o}_k = [y_{k1}, y_{k2}, y_{k3}, y_{k4}, y_{k5}]^T$ ($k = 1, 2, \dots, K$). According to OWA [33], we will reorder \mathbf{o}_k to $\mathbf{b}_k = [b_{k1}, b_{k2}, b_{k3}, b_{k4}, b_{k5}]^T$, where b_{ki} is the i th largest element of the collection of $y_{k1}, y_{k2}, y_{k3}, y_{k4}, y_{k5}$. Let \hat{y}_d^k be the current estimated

aggregated values corresponding to \mathbf{b}_k and $\hat{\mathbf{w}}$. Then, \hat{y}_d^k can be calculated by

$$\begin{aligned}\hat{y}_d^k &= \hat{\mathbf{w}}^T \mathbf{b}_k = \sum_{i=1}^5 \hat{w}_i b_{ki} \\ &= \frac{b_{k1} e^{\lambda_1}}{\sum_{j=1}^5 e^{\lambda_j}} + \frac{b_{k2} e^{\lambda_2}}{\sum_{j=1}^5 e^{\lambda_j}} + \cdots + \frac{b_{k5} e^{\lambda_5}}{\sum_{j=1}^5 e^{\lambda_j}}.\end{aligned}\quad (18)$$

Let y_d^k be the expected aggregated values corresponding to \mathbf{o}_k , then the error e_k between y_d^k and \hat{y}_d^k can be calculated by

$$\begin{aligned}e_k &= \frac{1}{2} (\hat{y}_d^k - y_d^k)^2 \\ &= \frac{1}{2} \left(\sum_{i=1}^5 \hat{w}_i b_{ki} - y_d^k \right)^2.\end{aligned}\quad (19)$$

Using the steepest gradient descent method [53], the parameters λ_i ($i = 1, 2, \dots, 5$) are updated with the following equation:

$$\lambda_i(k+1) = \lambda_i(k) - 2\beta w_i (b_{ki} - \hat{y}_d^k) e_k, \quad (20)$$

where β is the learning rate. Consequently, parameters w_i are calculated at each iteration step for the current values of parameters $\lambda_i(k)$ ($i = 1, 2, \dots, 5$).

4.2. Features unavailable

We consider two situations where some features are not available: (1) one feature is not available, and (2) two features are not available. In Situation (1), suppose that a particular feature τ ($1 \leq \tau \leq 5$) is not available. Then, (18) can be rewritten as

$$\hat{y}_d^k = (\hat{\mathbf{w}}')^T \mathbf{b}'_k = \sum_{i=1, i \neq \tau}^5 \hat{w}'_i b'_{ki}, \quad (21)$$

where $\hat{\mathbf{w}}' = \{\hat{w}'_i (i = 1, 2, \dots, 5, \text{ and } i \neq \tau)\}$ which should be obtained through retraining, $\mathbf{b}'_k = \{b'_{ki} (i = 1, 2, \dots, 5, \text{ and } i \neq \tau)\}^T$; and at last, the final estimated output \hat{y}_d^k of the system can be calculated by

$$\hat{y}_d^k = \bar{y}_d^k * (1 - \hat{w}_\tau), \quad (22)$$

where $\hat{w}_\tau \in \{\hat{w}_i (i = 1, 2, \dots, 5)\}$, and $(1 - \hat{w}_\tau)$ stands for the belief function in the case that one feature is not available.

In Situation (2), suppose that two features τ and ξ ($1 \leq \tau, \xi \leq 5$, and $\tau \neq \xi$) are not available. Then, (18) can be rewritten as

$$\hat{y}_d^k = (\hat{\mathbf{w}}'')^T \mathbf{b}''_k = \sum_{i=1, i \neq \tau, i \neq \xi}^5 \hat{w}''_i b''_{ki}, \quad (23)$$

where $\hat{\mathbf{w}}'' = \{\hat{w}''_i (i = 1, 2, \dots, 5, \text{ and } i \neq \tau, i \neq \xi)\}$ which should be obtained through retraining, $\mathbf{b}''_k = \{b''_{ki} (i = 1, 2, \dots, 5, \text{ and } i \neq \tau, i \neq \xi)\}^T$; and at last, the final estimated output \hat{y}_d^k of the system can be calculated by

$$\hat{y}_d^k = \bar{y}_d^k * (1 - \hat{w}_\tau - \hat{w}_\xi), \quad (24)$$

where $\hat{w}_\tau, \hat{w}_\xi \in \{\hat{w}_i (i = 1, 2, \dots, 5)\}$, and $(1 - \hat{w}_\tau - \hat{w}_\xi)$ stands for the belief function in the case that two features are not available. Note that if more than two features are not available, the same procedure can be designed.

5. THE SIMULATION-BASED EXPERIMENT

In order to demonstrate the validity of the TSK-OWA method, we first perform training on a set of data obtained from the subjects who participated in an experiment to determine both the structure and parameters of the TSK-OWA. Then, another set of data obtained from the subjects under different simulation situations is obtained and performed on the TSK-OWA with the trained structure and parameters to illustrate the effectiveness of the TSK-OWA approach.

5.1. Experiment setup

Referring to the experimental conditions for producing the contact-feature datasets of ECG and EEG [7, 8, 20, 39–45, 54], and the contact-less-feature dataset of EM [1, 56], we designed an experiment environment to acquire necessary data based on Northeastern's virtual environments driver simulator. The simulator is equipped with the instruments such as CCD camera, eye gaze tracking, and one for acquiring EEG and ECG signals.

5.2. Data acquisition

To get the dataset of SQ, we designed a questionnaire according to the experimental conditions for producing the casual dataset of SQ [4, 6, 38], mainly concerning the effective required sleep hours. The questionnaires are distributed among the 9 driver participants and query them to answer the question of how many effective hours they sleep at night before participating the experiment.

To get the datasets of EEG, ECG, and EM, the 9 driver participants are asked to participate in the experiment. Each of them sat in front of the monitor with his hands on the steering wheel to control the car running at the speed of 80 kilometer/hour and staying in the center of the simulated freeway. At the same time, EEG and ECG signals of each participant are measured at the sampling rate of 250 HZ, and his/her dynamical facial image is obtained at the sampling rate of 2 seconds. EEG and ECG signals and a series of dynamical facial image are processed with the method presented in Section 2. As a result, nice datasets of EEG, ECG, EM, and DH are obtained and normalized. Seven drivers were randomly selected from the nine participants, along with their datasets, are used for training, and the remaining two drivers are for the algorithm evaluation.

5.3. Implementation of the neuro-fuzzy TSK network model

In this study, 7 datasets are taken as the inputs of TSK1, TSK2, TSK3, TSK4, and TSK5, and α^2 and ε are set to be 0.08 and 0.01, respectively. Under these conditions, each input

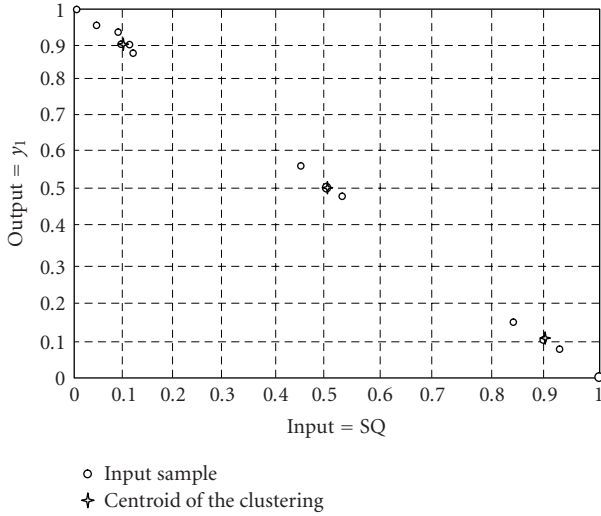


FIGURE 3: SQ input space partition for TSK1.

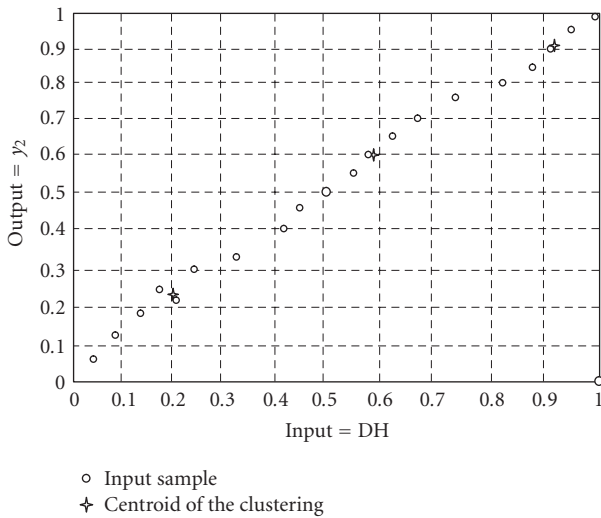


FIGURE 4: DH input space partition for TSK2.

space for TSK1, TSK2, TSK3, TSK4, and TSK5 is partitioned, as shown in Figures 3–7.

From Figure 3, it can be seen that the SQ input space is automatically partitioned into three fuzzy sets. Thus, the neuro-fuzzy TSK1 network has three fuzzy inference rules corresponding to the three fuzzy sets. The premise and consequent parameters of the inference, denoted as c_i^1 ($i = 1, 2, 3$) and p_{ij}^1 ($i = 1, 2, 3, j = 0, 1$), respectively, are determined by training with the same given training samples, and they are listed in Table 1.

From Figure 4, it can be seen that the DH input space is automatically partitioned into three fuzzy sets. Thus, the neuro-fuzzy TSK2 network has three fuzzy inference rules corresponding to the three fuzzy sets. The premise and consequent parameters of the inference, denoted as c_i^2 ($i = 1, 2, 3$) and p_{ij}^2 ($i = 1, 2, 3, j = 0, 1$), respectively, are determined by training with the same given training samples, as shown in Table 2.

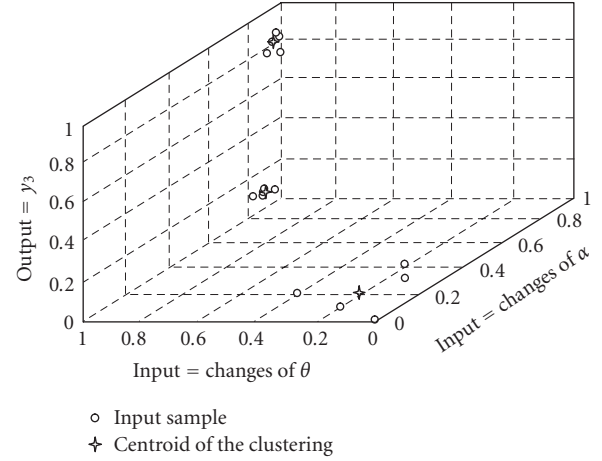


FIGURE 5: EEG input space partition for TSK3.

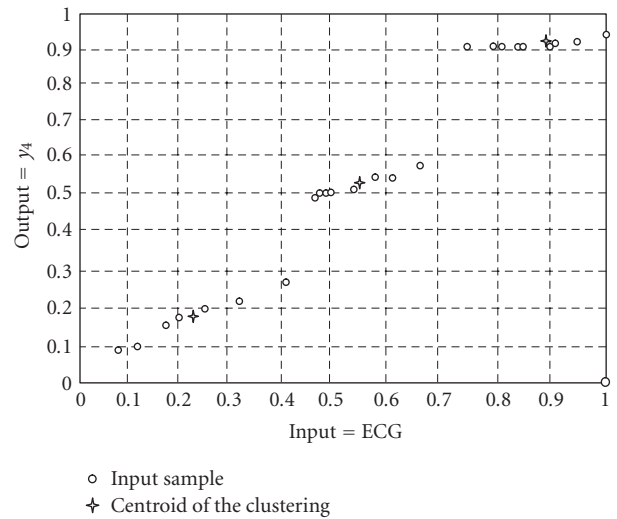


FIGURE 6: ECG input space partition for TSK4.

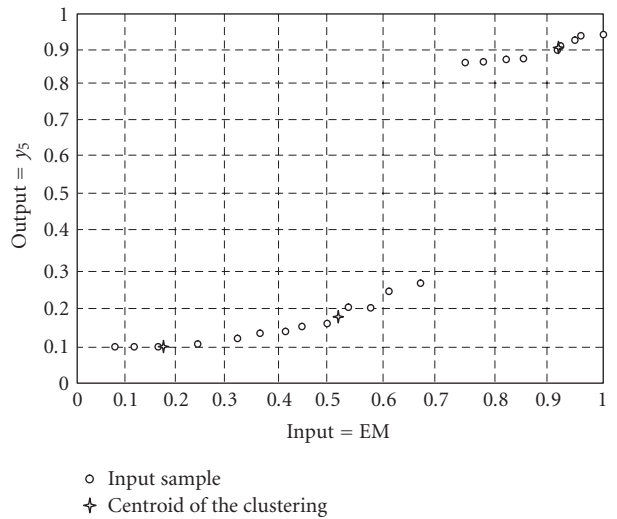


FIGURE 7: EM input space partition for TSK5.

TABLE 1: Parameters for TSK1.

c_1^1	c_2^1	c_3^1
0.9046	0.5007	0.0970
p_{10}^1	p_{20}^1	p_{30}^1
1.0036	0.9504	0.9947
p_{11}^1	p_{21}^1	p_{31}^1
-1.0028	-0.8934	-0.9915

TABLE 2: Parameters for TSK2.

c_1^2	c_2^2	c_3^2
0.2035	0.5907	0.9217
p_{10}^2	p_{20}^2	p_{30}^2
0.0498	-0.0481	-0.1812
p_{11}^2	p_{21}^2	p_{31}^2
0.9216	1.1005	1.1814

TABLE 3: Parameters for TSK3.

c_{11}^3	c_{12}^3	c_{21}^3	c_{22}^3	c_{31}^3	c_{32}^3
0.202	0.182	0.492	0.482	0.846	0.852
P_{10}		P_{11}		P_{12}	
0.01		0		0	
P_{20}		P_{21}		P_{22}	
0.3443		0.0957		0.2087	
P_{30}		P_{31}		P_{32}	
0.8476		0.0324		0.0364	

From Figure 5, it can be seen that the EEG input space is automatically partitioned into three fuzzy sets. Thus the neuro-fuzzy TSK3 network has three fuzzy inference rules corresponding to the three fuzzy sets. The premise and consequent parameters of the inference, denoted as c_{ik}^3 ($i = 1, 2, 3, k = 1, 2$) and p_{ij}^3 ($i, j = 1, 2, 3, j = 0, 1, 2$), respectively, are determined by training with the same given training samples, as shown in Table 3.

From Figure 6, it can be seen that the ECG input space is automatically partitioned into three fuzzy sets. Thus, the neuro-fuzzy TSK4 network has three fuzzy inference rules corresponding to the three fuzzy sets. The premise and consequent parameters of the inference, denoted as c_i^4 ($i = 1, 2, 3$) and p_{ij}^4 ($i = 1, 2, 3, j = 0, 1$), respectively, are determined by training with the same given training samples, as shown in Table 4.

From Figure 7, it can be seen that the EM input space is automatically partitioned into three fuzzy sets. Thus, the neuro-fuzzy TSK5 network has three fuzzy inference rules corresponding to the three fuzzy sets. The premise and consequent parameters of the inference, denoted as c_i^5 ($i = 1, 2, 3$) and p_{ij}^5 ($i = 1, 2, 3, j = 0, 1$), respectively, are determined by training with the same given training samples, as shown in Table 5.

TABLE 4: Parameters for TSK4.

c_1^4	c_2^4	c_3^4
0.2305	0.5634	0.8925
p_{10}^4	p_{20}^4	p_{30}^4
0.0233	-0.1656	0.8339
p_{11}^4	p_{21}^4	p_{31}^4
0.06737	1.2597	0.092

TABLE 5: Parameters for TSK5.

c_1^5	c_2^5	c_3^5
0.179	0.5204	0.9209
p_{10}^5	p_{20}^5	p_{30}^5
0.0435	-0.0617	0.6533
p_{11}^5	p_{21}^5	p_{31}^5
0.2834	0.4755	0.2767

TABLE 6: Training samples for OWA.

y_1	y_2	y_3	y_4	y_5	y_d
0.1	0.2	0.2	0.3	0.1	0.18
0.3	0.5	0.45	0.5	0.2	0.39
0.2	0.3	0.2	0.1	0.4	0.24
0.92	0.85	0.8	0.9	0.95	0.884
0.8	0.7	0.65	0.73	0.9	0.756
0.92	0.96	0.94	0.9	0.91	0.926
...

TABLE 7: Parameters for OWA.

w_1	w_2	w_3	w_4	w_5
0.1769	0.1955	0.2161	0.2161	0.1955

5.4. Implementation of the OWA method

When Outputs of TSK1, TSK2, TSK3, TSK4, and TSK5 ($y_i, i = 1, 2, \dots, 5$) are available, they are taken as the inputs of OWA and fed into OWA to be fused into the final decision (i.e., fatigue estimation). In this study, training data were selected to have a large coverage of possible cases. Some training data pairs (i.e., y_i and the expected aggregated value y_d) are shown in Table 6.

The parameters of OWA are obtained through training with the data as shown in Table 6. The training results are listed in Table 7.

When some outputs of TSK1, TSK2, TSK3, TSK4, and TSK5 ($y_i, i = 1, 2, \dots, 5$) are not available, the structure and parameters of OWA should be adjusted through retraining with the dataset of the features not available. Some training data pairs with features not available are shown in Tables 8, 9, and 10, and the training results are listed in Tables 11, 12, and 13.

TABLE 8: Training samples for OWA with SQ not available.

y_2	y_3	y_4	y_5	y_d
0.96	0.94	0.9	0.91	0.9272
0.5	0.45	0.5	0.2	0.3625
0.2	0.2	0.3	0.1	0.2
0.85	0.8	0.9	0.95	0.875
0.3	0.2	0.1	0.4	0.25
0.7	0.65	0.73	0.9	0.745
0.2	0.2	0.3	0.5	0.3
...

TABLE 9: Training samples for OWA with EM not available.

y_1	y_2	y_3	y_4	y_d
0.2	0.3	0.2	0.4	0.275
0.92	0.85	0.8	0.9	0.8675
0.3	0.5	0.45	0.5	0.4375
0.1	0.2	0.2	0.3	0.2
0.8	0.7	0.65	0.73	0.756
0.65	0.51	0.32	0.78	0.565
0.92	0.96	0.94	0.9	0.93
0.25	0.4	0.87	0.65	0.5425
...

TABLE 10: Training samples for OWA with SQ and EM not available.

y_2	y_3	y_4	y_d
0.7	0.65	0.73	0.693
0.3	0.2	0.1	0.20
0.5	0.45	0.5	0.483
0.85	0.8	0.9	0.85
0.96	0.94	0.9	0.90
0.2	0.2	0.3	0.233
0.65	0.78	0.63	0.687
0.55	0.69	0.34	0.527
...

TABLE 11: Parameters for OWA with SQ not available.

w_2	w_3	w_4	w_5
0.2375	0.2625	0.2625	0.2375

5.5. Results and discussions

In order to test the structure and parameters of the proposed TSK-OWA method, the remaining two drivers of the 9 participants did experiments under different conditions. For the first driver, he had an insufficient sleep (e.g., 2 hours) prior to driving and was asked to drive along a straight and flat freeway in the simulated driving experiment for 2 hours without stop. For the second driver, he had a sufficient sleep (e.g., 7 hours) prior to driving, and was asked to drive along a straight and flat freeway in the simulated driving experiment for 2 hours without stop. At last, the first driver was in fatigue state, while the second driver was in no-fatigue state

TABLE 12: Parameters for OWA with EM not available.

w_1	w_2	w_3	w_4
0.2199	0.2430	0.2686	0.2686

TABLE 13: Parameters for OWA with SQ and EM not available.

w_2	w_3	w_4
0.3115	0.3443	0.3442

TABLE 14: Features and simulation results for the first driver.

Input features	SQ	DH	ECG	EEG	EM
	0.25	0.5	0.83	0.81	0.85
TSK output	y_1	y_2	y_3	y_4	y_5
	0.9046	0.5907	0.8925	0.849	0.89
OWA fused result 0.8258					
Final fused result 0.8258					

TABLE 15: Features and simulation results for the second driver.

Input features	SQ	DH	ECG	EEG	EM
	0.875	0.167	0.33	0.38	0.41
TSK output	y_1	y_2	y_3	y_4	y_5
	0.075	0.21	0.29	0.33	0.41
OWA fused result 0.2598					
Final fused result 0.2598					

after finishing the driving experiment. In the whole driving experiment, EEG and ECG signals, and a series of dynamical facial image of the two drivers were recorded and processed with the method presented in Section 2 to obtain datasets of EEG, ECG, EM, and DH.

All datasets are fed into TSK1, TSK2, TSK3, TSK4, and TSK5, and the outputs (y_i , $i = 1, 2, \dots, 5$) of the 5 TSK networks can be calculated by use of the parameters shown in Tables 1–5.

The outputs (y_i , $i = 1, 2, \dots, 5$) are fused into the final output of the system by use of the parameters shown as Table 7. All simulated experiment results including the intermediate and final computing are shown in the Table 14 (for the first driver) and Table 15 (for the second driver).

When SQ feature is not available, outputs of TSK2, TSK3, TSK4, and TSK5 (y_i , $i = 2, \dots, 5$) are fused into the output of the system by use of the parameters shown in Table 11, and then the final output of the system is calculated with (22). All experiment results, including the intermediate and final computing for the first driver, are shown in the Table 16.

When EM feature is not available, Outputs of TSK1, TSK2, TSK3, and TSK4 (y_i , $i = 1, 2, \dots, 4$) are fused into the output of the system by use of the parameters shown in Table 12, and then the final output of the system is calculated with (22). All experiment results, including the intermediate and final computing for the first driver, are shown in the Table 17.

When SQ and EM features are not available, Outputs of TSK2, TSK3, and TSK4 (y_i , $i = 2, \dots, 4$) are fused into the output of the system by use of the parameters shown in

TABLE 16: Features and simulation result for the first driver when SQ is not available.

Input features	SQ	DH	ECG	EEG	EM
	—	0.5	0.83	0.81	0.82
TSK output	y_1	y_2	y_3	y_4	y_5
	—	0.5907	0.8925	0.859	0.89
OWA fused result 0.8113					
Final fused result 0.6677					

TABLE 17: Features and simulation result for the first driver when EM is not available.

Input features	SQ	DH	ECG	EEG	EM
	0.25	0.5	0.83	0.81	0.85
TSK output	y_1	y_2	y_3	y_4	y_5
	0.9046	0.5907	0.8925	0.859	—
OWA fused result 0.8052					
Final fused result 0.6478					

TABLE 18: Features and simulation result for the first driver when SQ and EM are not available.

Input features	SQ	DH	ECG	EEG	EM
	—	0.5	0.83	0.81	0.85
TSK output	y_1	y_2	y_3	y_4	y_5
	—	0.5907	0.8925	0.859	—
OWA fused result 0.7771					
Final fused result 0.4877					

Table 12, and then the final output of the system is calculated with (24). All experiment results including the intermediate and final computing for the first driver are shown in Table 18.

From Table 14, it can be obviously seen that the final output of the system is 0.8258, which means the probability of the driver who is in the fatigue state is 82.58%. In other words, it is obvious that the driver is in the most fatigue state. This is consistent with the fact that the first driver is in complete fatigue state after finishing the driving experiment. From Table 15, it can be obviously seen that the final output of the system is 0.2598, which means the probability of the driver who is in the fatigue state is 25.98%. In other words, it is obvious that the driver is in the nonfatigue state. This is consistent with the fact that the second driver is in nonfatigue state after finishing the driving experiment. The results obtained as above demonstrate the effectiveness of the TSK-OWA method.

From Tables 16–18, it can be also seen that the probability of the driver fatigue state for the same driver in the same situation decreases with the decrease in the number of features, which means that the recognition reliability of the driver fatigue state decreases with the decrease in the number of features. This implies that it is necessary to fuse multiple features as many as possible in order to make fatigue recognition more reliable when dealing with the driver's fatigue recognition problem.

6. CONCLUSIONS

This paper proposed a new method for inferring human cognitive states based on multimodality cues. The method is based on the integration of the neuro-fuzzy TSK network and the multifeature fusion OWA. This new method is called TSK-OWA. We presented an experimental validation in a virtual driving simulator. The study can conclude.

- (1) The classification of features into three different categories, namely, (1) contextual, (2) contact, and (3) contact-less, adds value to the accuracy of inferring the driver fatigue.
- (2) A high coverage of features over these three categories tends to improve the reliability of the measurement for the driver fatigue.
- (3) More cues appear to be more accurate in inferring the drive fatigue.

One limitation with this work is that all the experimental data were drawn from the simulator in the laboratory environment instead from the real driving environment. Therefore, a further experiment in a real driving environment is one interesting future work. Another limitation is that still only a few features are considered; more features need to be studied in order to have a complete picture of the driver fatigue state—which is an interesting future work. Furthermore, there is a need to perform sensitivity analysis with regard to adding or dropping features. Finally, although it seems feasible to generalize the conclusions drawn for inferring the driver fatigue to any other cognitive state, including emotion and mental workload, a future study seems to be necessary for applying the proposed method to infer some other cognitive and emotion states.

ACKNOWLEDGMENTS

The authors would like to acknowledge the generous financial support from Northeastern University (through a start-up fund) and Natural Sciences and Engineering Research Council (NSERC) of Canada (through a discovery grant and University Faculty Award program) awarded to the second author.

REFERENCES

- [1] Q. Ji, Z. Zhu, and P. Lan, "Real-time non-intrusive monitoring and prediction of driver fatigue," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1052–1068, 2004.
- [2] B. Fasel and J. Luetttin, "Automatic facial expression analysis: a survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259–275, 2003.
- [3] Q. Ji, P. Lan, and C. G. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," *IEEE Transactions on Systems, Man and Cybernetics A*, vol. 36, no. 5, pp. 862–875, 2006.
- [4] T. Oron-Gilad and D. Shinar, "Driver fatigue among military truck drivers," *Transportation Research Part F*, vol. 3, no. 4, pp. 195–209, 2000.
- [5] G. Hamouda and F. F. Saccomanno, "Neural network model for truck driver fatigue accident detection," in *Proceedings of*

- the Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 362–365, Montreal, Quebec, Canada, September 1995.
- [6] C. D. He and C. C. Zhao, “Evaluation of the critical value of driving fatigue based on the fuzzy sets theory,” *Environmental Research*, vol. 61, no. 1, pp. 150–156, 1993.
- [7] S. K. L. Lal, A. Craig, P. Boord, L. Kirkup, and H. Nguyen, “Development of an algorithm for an EEG-based driver fatigue countermeasure,” *Journal of Safety Research*, vol. 34, no. 3, pp. 321–328, 2003.
- [8] R.-C. Wu, C.-T. Lin, S.-F. Liang, T.-Y. Huang, Y.-C. Chen, and T.-P. Jung, “Estimating driving performance based on EEG spectrum and fuzzy neural network,” in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 1, pp. 585–590, Budapest, Hungary, July 2004.
- [9] B. J. Wilson and T. D. Bracewell, “Alertness monitor using neural networks for EEG analysis,” in *Proceedings of the IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing (NNSP ’00)*, vol. 2, pp. 814–820, Sydney, Australia, December 2000.
- [10] T. H. Linh, M. Stodolski, and S. Osowski, “On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network,” *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 4, pp. 1224–1231, 2003.
- [11] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sörnmo, “Clustering ECG complexes using Hermite functions and self-organizing maps,” *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 838–848, 2000.
- [12] R. W. Picard, E. Vyzas, and J. A. Healey, “Toward machine emotional intelligence: analysis of affective physiological state,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.
- [13] E. Vyzas and R. W. Picard, “Affective pattern classification,” in *Emotional and Intelligent: the Tangled Knot of Cognition, AAAI Fall Symposium Series*, Orlando, Fla, USA, October 1998.
- [14] Q. Wang, J. Yang, M. Ren, and Y. Zheng, “Driver fatigue detection: a survey,” in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA ’06)*, vol. 2, pp. 8587–8591, Dalian, China, June 2006.
- [15] W.-B. Horng, C.-Y. Chen, Y. Chang, and C.-H. Fan, “Driver fatigue detection based on eye tracking and dynamic template matching,” in *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, vol. 1, pp. 7–12, Taipei, Taiwan, March 2004.
- [16] Y. Norimatsu, S. Mita, K. Kozuka, T. Nakano, and S. Yamamoto, “Detection of the gaze direction using the time-varying image processing,” in *Proceedings of the 6th IEEE Annual Conference on Intelligent Transportation Systems*, vol. 1, pp. 74–79, Shanghai, China, October 2003.
- [17] D.-J. Kim, Z. Bien, and K.-H. Park, “Fuzzy neural networks(FNN)-based approach for personalized facial expression recognition with novel feature selection method,” in *Proceedings of the 12th IEEE International Conference on Fuzzy Systems (FUZZ ’03)*, vol. 2, pp. 908–913, St. Louis, Mo, USA, May 2003.
- [18] R. Grace, V. E. Byrne, D. M. Bierman, et al., “A drowsy driver detection system for heavy vehicles,” in *Proceedings of the 17th Digital Avionics Systems Conference (DASC ’98)*, vol. 2, pp. 136/1–136/8, Bellevue, Wash, USA, October–November 1998.
- [19] <http://www.optalert.com/optalertTM.technology.html>.
- [20] S. K. L. Lal and A. Craig, “A critical review of the psychophysiology of driver fatigue,” *Biological Psychology*, vol. 55, no. 3, pp. 173–194, 2001.
- [21] P. Vysoký, “Changes in car driver dynamics caused by fatigue,” *Neural Network World*, vol. 14, no. 1, pp. 109–117, 2004.
- [22] P. Vysoký, “Central fatigue identification of human operator,” *Neural Network World*, vol. 11, no. 5, pp. 525–535, 2001.
- [23] Y. Lin, P. Tang, W. J. Zhang, and Q. Yu, “Artificial neural network modelling of driver handling behaviour in a driver-vehicle-environment system,” *International Journal of Vehicle Design*, vol. 37, no. 1, pp. 24–45, 2005.
- [24] C. Conati, “Probabilistic assessment of user’s emotions in educational games,” *Applied Artificial Intelligence*, vol. 16, no. 7–8, pp. 555–575, 2002.
- [25] X. Li and Q. Ji, “Active affective state detection and user assistance with dynamic Bayesian networks,” *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 35, no. 1, pp. 93–105, 2005.
- [26] Y. Zhang, Q. Ji, and C. G. Looney, “Active information fusion for decision making under uncertainty,” in *Proceedings of the 5th International Conference on Information Fusion*, vol. 1, pp. 643–650, Annapolis, Md, USA, July 2002.
- [27] C.-F. Juang and C.-T. Lin, “An on-line self-constructing neural fuzzy inference network and its applications,” *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 12–32, 1998.
- [28] G. Beliakov, “Methods of construction of OWA operators from data,” in *Proceedings of the 10th IEEE International Conference on Fuzzy Systems*, vol. 1, pp. 184–187, Melbourne, Australia, December 2001.
- [29] T. Calvo, R. Mesiar, and R. R. Yager, “Quantitative weights and aggregation,” *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 1, pp. 62–69, 2004.
- [30] D. Filev and R. R. Yager, “On the issue of obtaining OWA operator weights,” *Fuzzy Sets and Systems*, vol. 94, no. 2, pp. 157–169, 1998.
- [31] M. O’Hagan, “Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic,” in *Proceedings of the 22nd Asilomar Conference Signals, Systems and Computers*, vol. 2, pp. 681–689, Pacific Grove, Calif, USA, October–November 1988.
- [32] V. Torra, “Learning weights for weighted OWA operators,” in *Proceedings of the 26th Annual Conference of the IEEE Industrial Electronics Society (IECON ’00)*, vol. 4, pp. 2530–2535, Nagoya, Japan, October 2000.
- [33] R. R. Yager, “On ordered weighted averaging aggregation operators in multicriteria decision making,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988.
- [34] R. R. Yager, “OWA neurons: a new class of fuzzy neurons,” in *Proceedings of the International Joint Conference on Neural Networks (IJCNN ’92)*, vol. 1, pp. 226–231, Baltimore, Md, USA, June 1992.
- [35] R. R. Yager, “Modeling prioritized multicriteria decision making,” *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 34, no. 6, pp. 2396–2404, 2004.
- [36] R. R. Yager, “OWA aggregation over a continuous interval argument with applications to decision making,” *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 34, no. 5, pp. 1952–1963, 2004.
- [37] A. Smiley, “Fatigue management: lessons from research,” in *Managing Fatigue in Transportation*, L. Hartley, Ed., pp. 1–23, Elsevier, Oxford, UK, 1998.
- [38] P. Thiffault and J. Bergeron, “Monotony of road environment and driver fatigue: a simulator study,” *Accident Analysis & Prevention*, vol. 35, no. 3, pp. 381–391, 2003.

- [39] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the EEG power spectrum," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 1, pp. 60–69, 1997.
- [40] J. C. Principe, S. K. Gala, and T. G. Chang, "Sleep staging automaton based on the theory of evidence," *IEEE Transactions on Biomedical Engineering*, vol. 36, no. 5, pp. 503–509, 1989.
- [41] G. Calcagnini, G. Biancalana, F. Giubilei, S. Strano, and S. Cerutti, "Spectral analysis of heart rate variability signal during sleep stages," in *Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, pp. 1252–1253, Baltimore, Md, USA, November 1994.
- [42] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp. 21–28, 1995.
- [43] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ECG delineator evaluation on standard databases," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 4, pp. 570–581, 2004.
- [44] J. S. Sahambi, S. N. Tandon, and R. K. P. Bhatt, "Using wavelet transforms for ECG characterization. An on-line digital signal processing system," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 1, pp. 77–83, 1997.
- [45] K. Murai, Y. Hayashi, and N. Wakabayashi, "Analysis of heart rate variability of navigator at in/from ports by wavelet transform," in *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM '01)*, vol. 2, pp. 681–685, Victoria, BC, Canada, August 2001.
- [46] W. W. Wierwille, L. A. Ellsworth, S. S. Wreggit, R. J. Fairbanks, and C. L. Kim, "Research on vehicle-based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness," Administration Final Report: DOT HS 808 247, National Highway Traffic Safety, 1994.
- [47] S. Barada and H. Singh, "Generating optimal adaptive fuzzy-neural models of dynamical systems with applications to control," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 28, no. 3, pp. 371–391, 1998.
- [48] S. Abe and R. Thawonmas, "A fuzzy classifier with ellipsoidal regions," *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 3, pp. 358–368, 1997.
- [49] S. Osowski, T. H. Linh, and K. Brudzewski, "Neuro-fuzzy TSK network for calibration of semiconductor sensor array for gas measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 53, no. 3, pp. 630–637, 2004.
- [50] M. Sugeno and T. Yasukawa, "Fuzzy-logic-based approach to qualitative modeling," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 7–31, 1993.
- [51] R. R. Yager and D. P. Filev, "Approximate clustering via the mountain method," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 8, pp. 1279–1284, 1994.
- [52] C.-C. Wong and C.-C. Chen, "A hybrid clustering and gradient descent approach for fuzzy modeling," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 29, no. 6, pp. 686–693, 1999.
- [53] D. Filev and R. R. Yager, "Learning OWA operator weights from data," in *Proceedings of the 3rd IEEE Conference on Fuzzy Systems, IEEE World Congress on Computational Intelligence*, vol. 1, pp. 468–473, Orlando, Fla, USA, June 1994.
- [54] J. A. Healey, "Wearable and automotive systems for affective recognition from physiology," Doctoral thesis, Massachusetts Institute of Technology, Mass, USA, 2000.
- [55] V. Torra, "Learning weights for weighted OWA operators," in *Proceedings of the 26th Annual Conference of the IEEE Industrial Electronics Society (IECON '00)*, vol. 4, pp. 2530–2535, Nagoya, Japan, October 2000.
- [56] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 699–714, 2005.