

Using Encephalography to Classify Mind States Associated with Cycling Activities

Wen-Lin Chu,¹ Chih-Jer Lin,^{2*} and Po-Wei Hsiao²

¹Department of Mechanical Engineering, National Chin-Yi University of Technology,
No.57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung 41170, Taiwan (R.O.C.)

²Graduate Institute of Automation Technology, National Taipei University of Technology,
1, Sec. 3, Zhongxiao E. Rd., Taipei 10608 Taiwan (R.O.C.)

(Received January 12, 2019; accepted December 17, 2019)

Keywords: brainwave, electroencephalogram, discrete wavelet transformation, support vector machine, general regression neural network

Cycling is a very popular activity worldwide and cyclists often listen to music while riding. This usually involves the use of a portable music player or a cellphone, which may present safety problems. Since these devices may easily distract the attention of the cyclist. This is a matter that has not received much attention but has great relevance to traffic safety. In this study, the attention level of cyclists was measured and recorded as electroencephalograms (EEGs), and discrete wavelet transforms (DWTs) based on Daubechies wavelets were used to extract the EEG features. Six different cycling activity patterns were investigated and eigenfunctions were used to classify the attention level. After feature extraction by the DWTs, support vector machines (SVMs) and general regression neural networks (GRNNs) were employed to recognize different states of mind associated with specific cycling activities. In Case I, the recognition rates of the SVM and GRNN were used to determine the state of mind associated with two different cycling activities, riding in a straight line and riding around obstacles. In Case II, rider vigilance was investigated using the SVM and GRNN for eight different cycling scenarios. The experimental results validated the proposed method and showed the brainwave patterns to be clearly associated with different cycling activities. The experimental results showed that looking at a cellphone screen or engaging in a call caused riders to miss very obvious peripheral stimuli and the use of a cellphone by a cyclist is a clear danger to the rider and traffic safety.

1. Introduction

Cycling is a very popular activity worldwide and cyclists often listen to music while riding. This usually involves the use of a portable music player or a cellphone and this may present safety problems since devices may easily distract the attention of the cyclist. This matter has not received much attention^(1–5) because most studies have concentrated on the effects of listening to music while driving a vehicle or a motorcycle,^(6–12) and very few studies have investigated the effects of listening to music or operating a cellphone on cycling behavior.^(13–17) In countries such as the Netherlands, the use of a smartphone and other MP3 players while

*Corresponding author: e-mail: cjlin@ntut.edu.tw
<https://doi.org/10.18494/SAM.2020.2601>

riding a bicycle is very common and not forbidden; however, listening to music while cycling is illegal in Germany and New Zealand.⁽¹⁷⁾ De Waard *et al.* observed that 7.7% of the cyclists were listening to an MP3 player while cycling in Groningen.⁽¹³⁾ On the basis of an internet survey of 2500 cyclists, Goldenbeld and Ehlers reported that 40% of cyclists between 12 and 17 and 15% of those between 18 and 34 years of age always listened to music while they rode. The percentage of cyclists who sometimes listened to music was found to be 76% for 12 to 17 year-olds, 54% for cyclists between 18 and 34, 23% for those between 35 and 50, and 14% for the 50+ age group.⁽¹⁴⁾ This report clearly showed that young cyclists listened to music while cycling more frequently than older riders.

Listening to music while riding is a factor that may be highly relevant to cycling accidents and has received little attention. Music can influence human behavior, and the potential of music to influence mood is described as one of the most important functions of music.⁽¹⁵⁾ Hyman *et al.* found that pedestrians who were talking on a cellphone missed very clear peripheral stimuli. However, those who were listening to music did not miss the stimuli as frequently as those not doing so.⁽¹⁸⁾ Music can benefit driving performance according to the mood-arousal hypothesis, since listening to music can lead to a more optimal arousal level.^(19,20) However, the distraction hypothesis suggests that music can divert attention from the task of driving and affect driving safety.⁽¹⁹⁾ However, for cyclists, the effects of listening to music are not very well known. In cycling, loud music may decrease the reaction time of the rider to central stimuli and increase the response time to peripheral stimuli at the same time.^(7,21) In the internet survey by Goldenbeld and Ehlers, 55% of music-listening cyclists reported using two earbuds or over-ear headphones, 23% used only one ear bud, and the rest reported using a loudspeaker or different options from time to time.⁽¹⁴⁾ De Waard *et al.* found that auditory perception was negatively affected, with less traffic auditory information being processed by riders listening to music or when engaged in a phone call, particularly when earbuds were being used. Most cyclists rely on sound to detect the approach or closeness of other vehicles. The main effect of music or a phone conversation would be to dull sound and the cyclist could miss warning signals relevant to cycling safety.⁽¹⁵⁾

However, few studies have been made using electroencephalograms (EEGs) to investigate changes in brainwaves related to cycling behavior. Mental state is the most important factor in cycling safety, and a bad mood, distraction, or fatigue may be frequent causes of traffic accidents involving cyclists. Five different electrical patterns, or brainwaves, known to be associated with mental state can be detected across the human cortex. These brainwaves can be observed with an EEG device and in decreasing order of frequency are gamma, beta, alpha, theta, and delta. Each wave is associated with a specific mental function and all five waves are displayed by the normal conscious human brain. However, one particular wave will be dominant depending on the state of consciousness. Nithiya *et al.* used EEGs to study the alpha waves in patients with attention deficit and used this information in the design of relaxation therapy.⁽²²⁾ Crowley *et al.* studied attention and meditation using the Tower of Hanoi problem and an EEG detection system to determine degrees of attention.⁽²³⁾ Lin *et al.* proposed a real-time driver vigilance system to detect fatigue and sleepiness in an attempt to find ways of reducing traffic accidents.⁽²⁴⁾

This study has been organized as follows. Section 2 presents the feature extraction approach. First, discrete wavelet transforms (DWTs) based on Daubechies wavelets was used to extract the EEG features. Second, six eigenfunctions were used to compute eigenvalues for classification. As reported in Sect. 3, the classifiers of the brainwaves were studied to determine associations with different cycling activities. Two types of classifier, support vector machines (SVMs) and general regression neural networks (GRNNs), are introduced briefly. In Case I, the mind states of two types of cycling activity were studied, riding in a straight line and riding to avoid obstacles. As the experiment progressed, the EEG signals from the subjects were captured using a MindWave headset. The classifiers, SVMs and GRNNs, were then studied to recognize brainwave patterns for the two types of cycling activity. Rider vigilance with respect to the related activities was then investigated using the SVM and GRNN data from Case II as reported in Sect. 4. For each specific cycling movement, types of cycling-related activity were studied. It was assumed that there were two different classes of mind state, “relaxed” and “attentive”, and pattern recognition was needed to define the classes of samples to train the classifiers. Once the classifiers had been trained, it was necessary to evaluate their performance on an independent (labeled) test set. The recognition performance data from the SVM and GRNN were then studied and discussed. In Sect. 5, the results are discussed and conclusions are drawn.

2. EEG Signal Analysis and Feature Extraction

2.1 Principle of WT

The EEG signals collected from the subjects were nonstationary signals that varied with emotion. Fast Fourier transforms (FFTs) of the EEG signals in the time and frequency domains did not clearly reveal their features. Two feature capture methods, the short-time Fourier transform (STFT) and DWT, were used to analyze the changes in EEG signals associated with different cycling activities. The STFT, proposed by Gabor,⁽²⁵⁾ was used, whose equation is given by Eq. (1). Gabor used window-type functions to improve the FT analysis of non-stationary signals, because the FFT has the drawback of only providing a frequency information component without frequency changes over time. The STFT uses a window-type function to divide the signal into a number of time intervals, and the signal sequences for different FT segments are shown in Fig. 1.

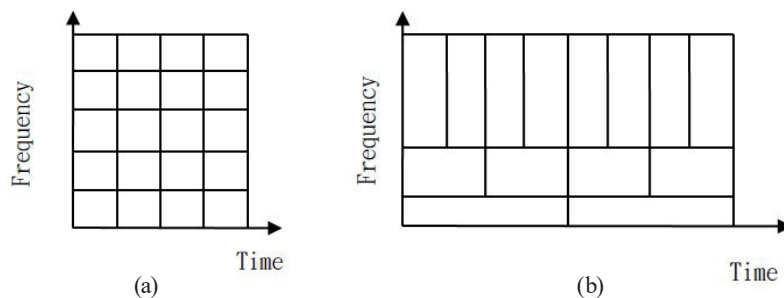


Fig. 1. (a) STFT analysis and (b) wavelet analysis.^(26,27)

$$STFT(\tau, \omega) = \int_{-\infty}^{\infty} x(t)g(t - \tau)e^{i\omega t} dt \quad (1)$$

Here, $g(t - \tau)$ is the window function used as the displacement function, τ is the shift factor, and ω is the modulation factor.

Because the STFT uses a fixed width for the window function (Fig. 1), the resolution of the frequency domain depends on the size of the time frame. To obtain good resolution in the time domain, a narrow time frame can be used, but this results in poor resolution in the frequency domain. On the other hand, choosing a wider time frame for better resolution in the frequency domain gives poor resolution in the time domain. In 1984, Goupillaud *et al.* proposed the WT to overcome the problem of the STFT fixed time window.⁽²⁶⁾ The WT not only retains the STFT characteristic of localization analysis, but also overcomes the problem of the window function with different time intervals. The WT allows the adjustment of the size of the time frame for different frequency bands. At low frequencies, a long time frame is chosen to obtain good resolution, and a short time frame gives good resolution at high frequencies. Figure 1 shows the analysis characteristic of the WT compared with the STFT methods.⁽²⁶⁾ In 1988, Daubechies proposed the WT, a hierarchy of wavelets that have orthogonal and compact support. Daubechies wavelets are widely used in solving signal analysis problems such as self-similarity properties of a signal and abnormal signal discontinuities.⁽²⁷⁾ Generally, the properties of a wavelet are effectively limited in duration, have an average value of zero, and can be expressed as

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty, \quad (2)$$

where $\psi(t)$ is the wavelet function.

2.2 Principle of DWT

Before the classification of the attention level based on the measurements of brainwaves, the DWT was used to extract the features of the EEG signals. As a general rule, WTs are divided into continuous wavelet transforms (CWTs) and DWTs. A CWT involves the signal conversion method proposed by Meyer *et al.* and was first used in 1980. A CWT provides multiple resolutions in the time-frequency analysis of the instantaneous features of non-stationary signals.⁽²⁸⁾ The CWT shown in Eq. (3) transforms a continuous signal into a highly redundant signal with two continuous variables, translation and scaling. For a continuous function $x(t)$, the CWT coefficient relative to the real-valued wavelet $\psi(t)$ is given by Eq. (4).

$$W_{\psi}(s, \tau) = \int_{-\infty}^{\infty} x(t)\psi_{s,\tau}^*(t) dt \quad (3)$$

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t - \tau}{s}\right) \quad (4)$$

Here, s and τ are the scale and translation parameters, respectively. $W_\psi(s, \tau)$ is the WT coefficient and $\psi(t)$ is the fundamental mother wavelet. However, the CWT requires a long time for calculation because it generates large wavelet coefficients.

The first DWT was derived by Haar in 1909.⁽²⁹⁾ He proposed what has become known as the Haar sequence. This is a sequence of wavelets of rescaled “square-shaped” functions, which together form a wavelet family or base. The mother wavelet function $\psi(t)$ of the Haar wavelet can be described as

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

with the scaling function

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ 0, & \text{otherwise} \end{cases} . \quad (6)$$

The most commonly used DWT set was formulated by Daubechies in 1988.⁽²⁸⁾ This formulation is based on the use of recurrence relations to generate progressively finer discrete samplings of an implicit mother wavelet function; each sampling has twice the resolution of the previous scale. In her seminal paper, Daubechies derived a family of wavelets, the first of which is the Haar wavelet, also known as Db1. Interest in this field has grown enormously since then, and many variations of the original Daubechies wavelets have been developed. The DWT was introduced to overcome the long computing times needed for a CWT. It does not generate large amounts of redundant information and allows the system to decompose signals to a subband level.

In the DWT, the original signal is decomposed into different band characteristics using high-pass and low-pass filters. The high-pass filters are designed to allow the signal in a specified band to pass, but signals below a given threshold are filtered. The low-pass filters can also allow signals below a specified threshold to pass. After filtration, the original signal is decomposed into approximate and detailed signals at different levels. In the DWT, the mother wavelet is shifted and scaled in powers of two. This means that s and τ can be defined as $s = 2^j$ and $\tau = 2^j k$, where $j, k \in \mathbb{Z}$. The filter bank implementation of wavelets computes the wavelet coefficients of a discrete set of child wavelets for a given mother wavelet $\psi(t)$. The wavelet function of the DWT is

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j k}{2^j}\right), \quad (7)$$

where j is the scale parameter and k is the shift parameter. Therefore, the DWT can be defined as

$$DWT(a,t) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - 2^j k}{2^j} \right) dt . \tag{8}$$

Consider an original signal $x(t)$ that has been decomposed by the DWT by being passed through a series of filters. The low-pass filter has an impulse response $g[n]$ and the signal also passes through a high-pass filter with an impulse response $h(t)$.

The resulting outputs can be decomposed into approximation coefficients (from the low-pass filter) and detailed coefficients (from the high-pass filter). The original signal passes through both filters and the output is the convolution of the two signals given as

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k], \tag{9}$$

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k], \tag{10}$$

where $x[n]$ is the input sequence in the discrete-time system and $g[n]$ and $h[n]$ are the impulse responses of the low-pass and high-pass filters in the discrete-time domain, respectively. The original signal can be decomposed using $g[n]$ and $h[n]$, which are quadrature mirror filters. According to the Nyquist rule, half the samples can be discarded and the filtered output of the above is then down-sampled by a factor of 2. The resulting output is further filtered by new low-pass and high-pass filters with half the cutoff frequency of the previous ones. This filtration increases the frequency resolution, and the approximation coefficients are decomposed by further filtration until the desired level is reached. Figure 2 is known as a filter bank, where the binary trees represent a subspace with a different frequency localization. The DWT

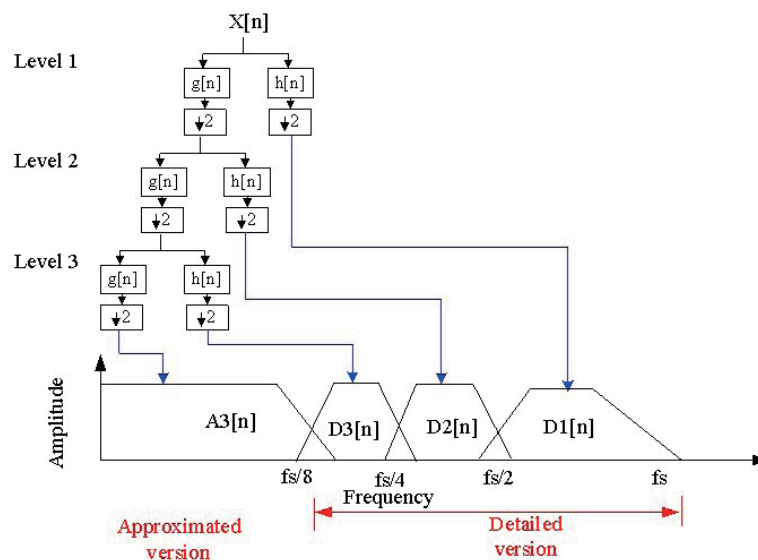


Fig. 2. (Color online) Basic decomposition of DWT.⁽²⁶⁾

carries out hierarchically organized decomposition and the level j is chosen according to the desired cutoff frequency. Figure 2 shows an implementation of a three-level forward DWT based on a two-channel recursive filter bank, where $g[n]$ and $h[n]$ are low-pass and high-pass analysis filters, respectively. In the figure, the block $\downarrow 2$ represents the factor 2 down-sampling operator. The input signal $x[n]$ is recursively decomposed to a total of four subband signals: an approximation version $A_3[n]$ and detailed versions, $D_1[n]$, $D_2[n]$, and $D_3[n]$, of three resolutions.

2.3 DWT analysis of EEG signals

The human brain contains many billions of neurons and the structure is divided into regions such as the prefrontal cortex, thalamus, cerebellum, and corpus callosum. An EEG system is used to detect prefrontal cortex activity. This region is associated with interpretation, judgment, attention, and relaxation.⁽²²⁾ In general, EEG signals can be divided into five bands: delta, theta, alpha, beta, and gamma. Delta waves usually appear during deep sleep and their frequency band is from 0.1 to 3 Hz. Brainwaves within the frequency band from 4 to 7 Hz are theta waves, which usually appear during drowsiness. The frequency band of alpha waves is from 8 to 12 Hz and alpha waves usually appear during relaxation or rest. Beta waves, which range in frequency from 13 to 30 Hz, are associated with concentration and active thinking. Gamma waves, from 31 to 50 Hz, appear during periods of high mental activity. The significance of each brainwave band is summarized in Table 1.⁽³⁰⁾ In this study, the Neurosky MindWave™ EEG-monitoring headset was used to retrieve the brainwave data. The headset can measure brainwave signals safely and easily and its Sky ThinkGear ASIC chip has been designed for health and wellness, as well as educational and popular EEG technology use.⁽³¹⁾ According to the international 10–20 system, the NeuroSky headset is designed for use in position Fp1 to retrieve EEG signals. Position A1 is the reference ground (as shown in Fig. 3). Signals from the NeuroSky headset were transferred to a PC via Bluetooth and the raw data was analyzed using the following DWT method.

The sampling frequency of the NeuroSky EEG headset is 512 Hz and, in this study, the MATLAB wavelet toolbox was used to generate the Daubechies Db4 wavelet components of the EEG signals for decomposition analysis. The EEG signals were decomposed into five DWT levels based on the significance of the brainwave types (see Table 1). There are five dyadic wavelet levels and the input signal $x(n)$ was recursively decomposed into six subband signals: an approximation signal $A_5(n)$ and five detailed signals, $D_1(n)$, $D_2(n)$, $D_3(n)$, $D_4(n)$, and $D_5(n)$. Each of these wavelet levels corresponds to f with its band given in Table 2.

Table 1
Frequency of brainwave types and significance.⁽²⁶⁾

Brainwave type	Frequency (Hz)	Significance
Delta	0.1–3	Deep sleep
Theta	4–7	Drowsiness or early sleep
Alpha	8–12	Relaxation or rest
Beta	13–30	Mental concentration and active thinking
Gamma	30–50	High mental activity

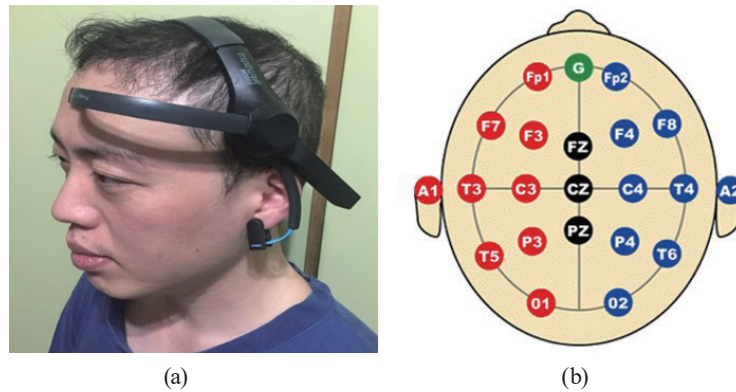


Fig. 3. (Color online) NeuroSky MindWave headset and 10–20 International EEG layout.⁽³²⁾

Table 2
Decomposition of EEG signals into different brainwave frequency bands.

Wavelet level	Frequency (Hz)	Brainwave type
$A_5 (j = 0)$	0–4	Delta
$D_5 (j = 1)$	4–8	Theta
$D_4 (j = 2)$	8–16	Alpha
$D_3 (j = 3)$	16–32	Beta
$D_2 (j = 4)$	32–64	Gamma

$$f = 2^j \left(\frac{f_S}{N} \right), \quad (11)$$

where f_S is the sampling frequency ($f_S = 512$) and N represents the number of data points in the original signal and is also the high frequency limit of the frequency band represented by level j . If the number of data points is chosen as $N = 128$, then the five detailed signals of the DWT can be mapped into the brainwave bands shown in Table 1. For example, the corresponding delta band is 0.1–3 Hz, which can be mapped into A_5 ; the theta band, 4–7 Hz, can be mapped into D_5 and the alpha band, 8–12 Hz, can be mapped into D_4 . Each of these wavelet levels corresponds to a frequency band as shown in Table 2. After DWT decomposition, it is necessary to find the specific eigenfunctions to compute the eigenvalues for classification.

2.4 EEG signal feature extraction using DWT obtained from eigenfunctions

Multiresolution DWT analysis was used to convert the original signals to an approximation signal $A_5(n)$ and five detailed signals, $D_1(n)$, $D_2(n)$, $D_3(n)$, $D_4(n)$, and $D_5(n)$. Then, the DWT for the EEG signals was used to observe the key features associated with the attention level of the cycling subject. However, EEG signals are non-stationary signals that vary with the emotion and the features of the signals could not be identified easily by only WDT analysis. Therefore, six eigenfunctions were applied to analyze the feature patterns for classification as shown in Eqs. (12)–(17).

Maximum (MAX):

$$D_{N,Max} = \max(D_N(n)), \quad (12)$$

Minimum (MIN):

$$D_{N,min} = \min(D_N(n)), \quad (13)$$

Range (RAN):

$$D_{N,Range} = D_{N,Max} - D_{N,min}, \quad (14)$$

Standard deviation (SD):

$$D_{N,\sigma} = \sqrt{\frac{1}{N} \sum_{n=1}^N (D_N(n) - D_{N,\mu})^2}, \quad (15)$$

Median absolute deviation (MAD):

$$D_{N,MD} = \text{median}(|D_N(n) - \text{median}(D_N(n))|), \quad (16)$$

Mean absolute value (MAV):

$$D_{N,mean} = \frac{1}{N} \sum_{n=1}^N |D_N(n)|. \quad (17)$$

To gain an understanding of the effects of different cycling-related activities, optimal pattern recognition was needed, and so the brainwaves were classified using the different feature patterns of the DWT components.

3. Classification of Extracted Brainwave Features Using SVM and GRNN

3.1 Experimental setup

The EEG signals were captured using the NeuroSky headset. The detectors were located at position PF1 on the subject's head to retrieve the raw EEG signals (see Fig. 3) and position A1 was used as the reference ground. Signals from the headset were transmitted to the computer via Bluetooth. The experimental arrangement and system block diagram are shown in Fig. 4. The experiments in case I were designed to investigate the prevailing brainwave patterns for two different types of cycling activity: riding along a straight path and taking a path around obstacles. Eight different cycling-related scenarios were used and the brainwaves of the subject were recorded under different cycling conditions, as shown in Fig. 5, to gain an understanding of the mind state associated with the two activities. The EEG signals were processed and analyzed to determine the relationship between the mental state and the riding situation. The experiment was carried out by six different subjects three times each. Each experiment, both along a straight path and along a path with obstacles, lasted between 16 and 20 s to ensure a

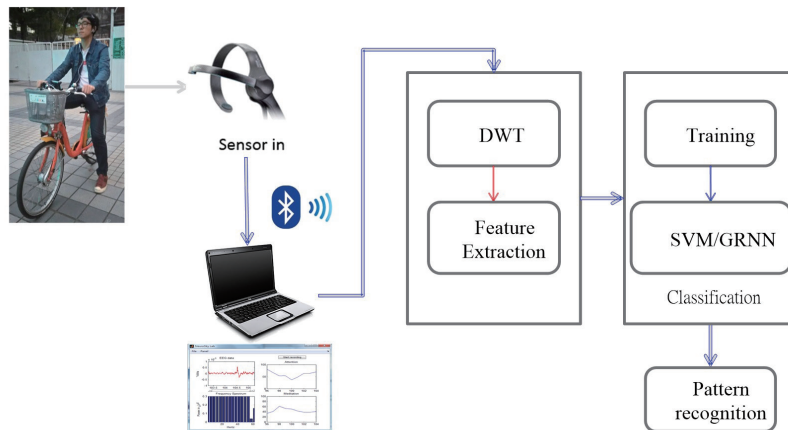


Fig. 4. (Color online) Experimental platform for the recognition of brainwaves.

Ride along a straight path

-without any additional behavior

-While listening music

- While watching a cellphone

- While communication by cellphone

Ride while a bypassing obstacles

-without any additional behavior

- While listening to music

- While watching a cellphone

-While communication by cellphone

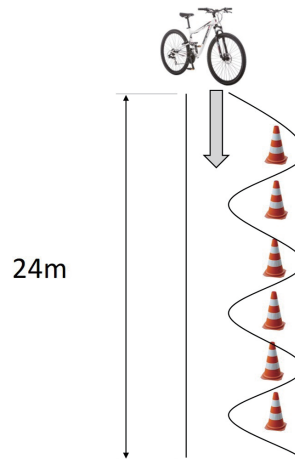


Fig. 5. (Color online) Scenarios for eight different cycling-related activities in Cases I and II.

steady riding speed. The distance traveled along the straight path was about 24 m. There were six obstacles on the path with obstacles that were 3 m apart. The stored EEG brainwave records from the subjects were analyzed and Figs. 6 and 7 respectively show the DWT decomposition of EEG signals collected from a subject riding along the straight path and a subject riding along the path with obstacles carrying out any other activity.

The theta, alpha, and beta waves are associated with attention (see Tables 1 and 2), and therefore the DWT was used to extract the feature bands of D_3 , D_4 , and D_5 for the computation of eigenvalues. Table 2 shows that D_5 covers the theta band, D_4 the alpha band, and D_3 the beta band. To choose a suitable eigenfunction for the DWT components for the classification of the attention level, the eigenvalues of D_3 , D_4 , and D_5 were obtained using the eigenfunctions shown in Eqs. (12)–(17). To understand the effect of the cycling-related activity on the attention level, the SVMs and GRNNs were both employed to classify the brainwaves of the subjects from the EEG data collected during the riding experiments. Because the subjects necessarily

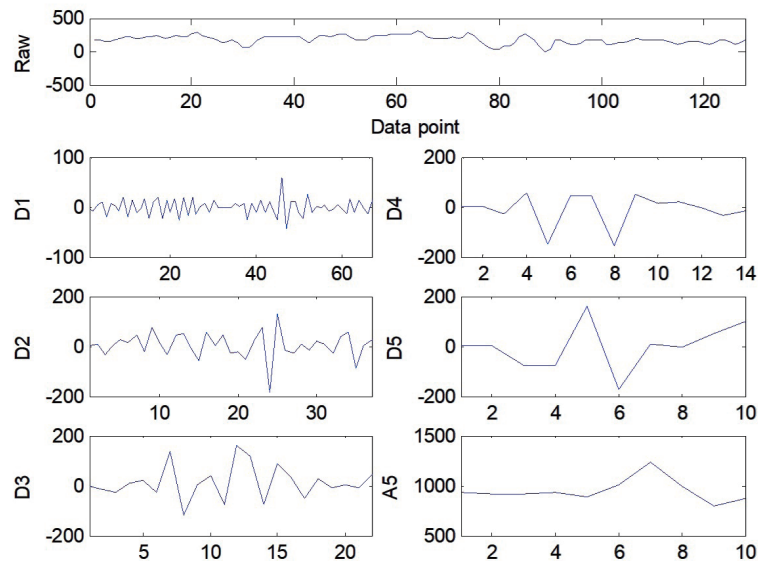


Fig. 6. DWT decomposition of an EEG signal for a subject riding along a straight path.

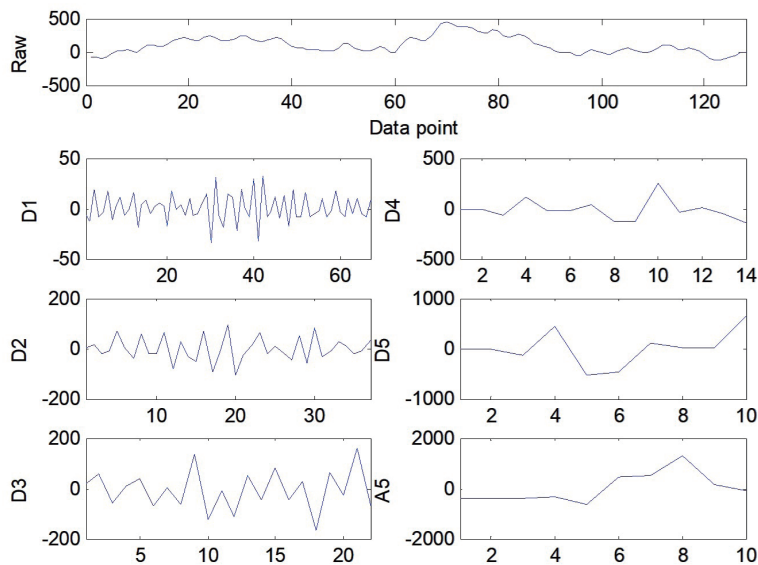


Fig. 7. DWT decomposition of an EEG signal for a subject riding along a path with obstacles.

paid more attention to riding when bypassing obstacles than when riding in a straight line, two classification algorithms were used to distinguish activity requiring a high level of attention from that with a relaxed mind. The relationship between the two mind states revealed by the experiments can help us determine what type of activity would cause distraction detrimental to driving safety.

3.2 Classification of mind state using the SVM

Recently, kernel-based analysis has become popular in data mining applications because it provides a surprisingly rich way to interpolate between pattern analysis, signal processing, and pattern recognition methods.⁽³³⁾ One successful example of a kernel-based learning algorithm is the SVM, which has been employed in many classification problems.^(34–38) SVMs are machine-learning algorithms proposed by Cortes and Vapnik to categorize data into two groups.⁽³⁹⁾ The main basic concept of the SVM is to find a hyperplane that can separate two or more different classes of data. The SVM must be trained using a set of samples that are marked as belonging to one category or another. After training, using a group of given data, the SVM can classify new data according to the established hyperplane.

Given a training set $\{(x_i, y_i)\}_{i=1}^N$, where N is the size of the training set and y_i is the class label of data x_i , and assuming only two classes for classification, $y_i = \{-1, +1\}$, if the data x_i can be linearly separated by hyperplanes, the best hyperplane should separate the two classes with the maximum margin as shown in Fig. 8, i.e., the distance from the hyperplane to the closest samples is maximal. In Fig. 8, $x \cdot w + b = 0$ is the hyperplane with the maximal margin of $2 / \|w\|$ and the circled samples are known as support vectors. In this study, a linear hyperplane cannot be found; therefore, a nonlinear mapping ϕ (see Fig. 9) is introduced to map the data into a linear space H where they are linearly separable. The SVM maps the training set into a higher-dimensional feature space via the mapping ϕ , then a separating hyperplane $w^T \phi(x) + b = 0$ that maximizes the margin of separation between the classes can be found. One of the most important properties of SVMs is that only the inner product $K(x_1, x_2) = \langle \phi(x_1), \phi(x_2) \rangle$, i.e., the kernel, is needed. The choice of the kernel is a specific problem for classification and the common kernel functions are linear, polynomial, the radial-based function (RBF), and the sigmoid function, which are described as follows.

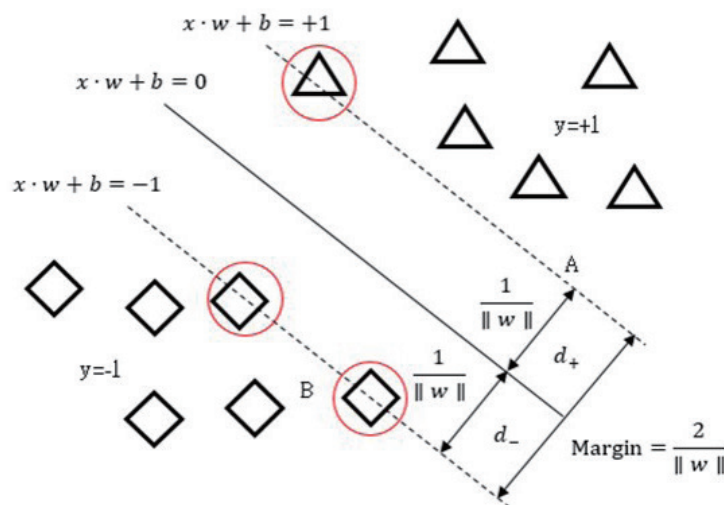


Fig. 8. (Color online) SVM separates two classes of data on a hyperplane with the maximum margin.

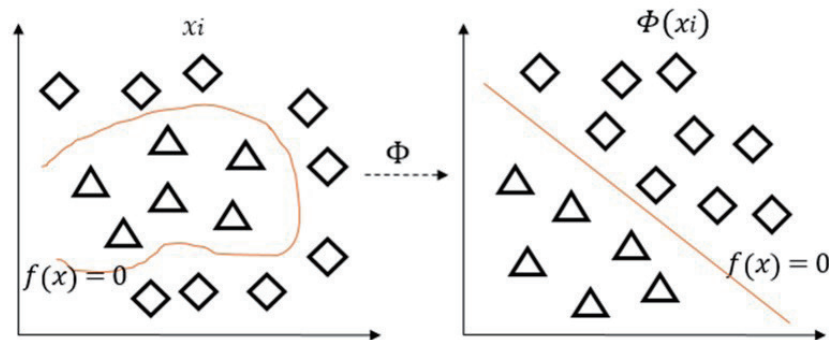


Fig. 9. (Color online) Mapping a training set into a higher-dimensional feature space via a kernel.

Linear:

$$K(x_1, x_2) = x_1 \cdot x_2, \quad (18)$$

Polynomial:

$$K(x_1, x_2) = (x_1 \cdot x_2 + 1)^d, \quad (19)$$

RBF:

$$K(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2), \quad (20)$$

Sigmoid:

$$K(x_1, x_2) = \tanh(kx_1 \cdot x_2 - \delta). \quad (21)$$

In this study, two of the above kernel functions, the RBF and polynomial, were used for the SVM classifier to investigate the accuracy of classification.

Case I. Classification of the mind states for different cycling movements

In the first study, three subjects performed the two cycling conditions three times. Figure 5 shows the details of these two conditions. The first involved riding in a straight line and the second involved avoiding obstacles. Each experiment lasted between 16 and 20 s to ensure a steady riding speed. The distance traveled along the straight path was about 24 m and there were six obstacles on the path, with 3 obstacles that were apart. At the same time, the EEG signals of the subjects were captured as 30 sequences of signals for each subject, giving 90 samples from each scenario and a total of 180 samples for classification. The DWT was used to extract the feature sequences of D_3 , D_4 , and D_5 , which were used to obtain the eigenvalues from Eqs. (12)–(17).

There were two brainwave patterns to be classified: the first was “relaxed”, associated with cycling in a straight line, and the second was “attentive”, associated with the task of avoiding obstacles. The following three testing modes were used: internal training, the half-external test, and the external test. In the half-external test, half the samples were used to train the SVM,

but all the samples were used for testing. In the external test, half the samples were used for training and the other half were used for testing. Table 3 presents the SVM classification results using two different kernel functions: the polynomial and RBF. In the half-external test, half the samples (90/180) were used for training. In the external test, there was no repetition of the training and all 90 samples were tested. The recognition rate was defined as

$$\text{Recognition rate} = \frac{\text{Number of correct classified samples}}{\text{Number of total testing samples}} \times 100\% . \quad (22)$$

The examination of Table 3 (SVM recognition rate) shows that the RBF kernel function is better than the polynomial kernel function. Therefore, the RBF kernel was used for SVM recognition in the study. Table 4 shows the recognition rates for the SVMs using the different eigenfunctions to compute the features D_3 , D_4 , and D_5 . It can be seen that the MAV gives the best recognition rate. Figure 10 shows the resulting hyperplane, the testing data, the classified data, and the support vectors obtained by the proposed method, where the first scenario “relaxed” is denoted by “Med”, and “Att” represents the second scenario “attentive”. The coordinates in Fig. 10 are the values obtained by the eigenfunction MAV of D_3 , D_4 , and D_5 of the DWT. To compare the SVM recognition rate with those of other methods, the same problem was investigated using the GRNN in Sect. 3.3.

3.3 Classification of mind state using GRNN

There are three common types of artificial neural network (ANN), the supervised, unsupervised, and associated learning networks. The GRNN is a supervised learning network that evolves from a probabilistic neural network (PNN).⁽⁴⁰⁾ It is also an antecedent-supervised learning network with a layered neuron structure, as shown in Fig. 11. The GRNN architecture has four layers: input, pattern, summation, and output. It can handle linear and nonlinear regression problems, as well as classification.

Table 3
SVM recognition rates for Case I.

Classifier type	Internal training	Half-external test	External test
(Training/Test)	180/180	90/180	90/90
SVM (RBF)	98.33%	80.00%	60.00%
SVM (Polynomial)	98.33%	78.89%	57.78%

Table 4
SVM recognition rates obtained using different eigenfunctions.

Classifier	Internal training	Half-external test	External test
(Training/Test)	180/180	90/180	90/90
MAX	98.89%	78.33%	58.89%
MIN	96.67%	77.78%	60.00%
RAN	94.44%	72.78%	48.89%
SD	90.56%	77.22%	61.11%
MAD	96.67%	72.78%	48.89%
MAV	96.11%	83.89%	72.22%

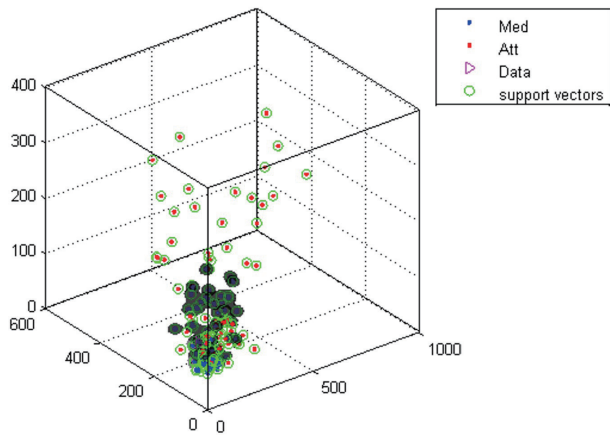


Fig. 10. (Color online) Hyperplane, testing data, classified data, and support vectors obtained by the proposed method.

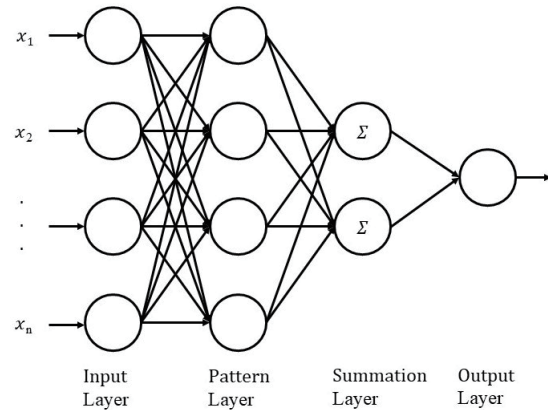


Fig. 11. GRNN network architecture.

In a classification with m -dimensional input vector training samples, the GRNN can use the samples to estimate the function values of unknown samples with a weighted average method as follows:

$$y_j = \frac{\sum_k y_{kj} w_k}{\sum_k w_k}, \quad k = 1, 2, \dots, q. \quad (23)$$

Here y_j is the j th function value of an unknown sample, w_k is the k th weighted value of a given sample, y_{kj} is the j th function value of the k th given sample, and q is the total number of given samples. The weighted w_k value can be obtained as follows:

$$w_k = \exp\left(-\frac{\sum_i (x_i - x_{ik})^2}{2\delta^2}\right), \quad i = 1, 2, \dots, m, \quad (24)$$

where x_i is the i th-dimensional spatial coordinate of the unknown sample, x_{ik} is the i th-dimensional coordinate of the k th sample for the given sample space, δ is the smoothing factor, and m is the sample space dimension. When the given samples are input to the GRNN for training, the learning rule used by the network will search for an appropriate smoothing parameter. Table 5 presents the recognition rates for the GRNN using the different eigenfunctions for the SWT features D_3 , D_4 , and D_5 . The recognition rates obtained by the GRNN for the eigenfunctions of MIN, RAN, SD, and MAD are much better than those obtained by the SVM. However, in the external test, the SVM had better recognition rates than the GRNN using the MAX and MAV eigenfunctions. To summarize, for the external test, the SVM using MAV is clearly the better classifier a recognition rate of 72.22%, compared with 70.00% for the GRNN.

Table 5
Recognition rates obtained by GRNN using different eigenfunctions.

Classifier eigenfunction	Internal training (%)	Half-external test (%)	External test (%)
MAX	100.00	77.78	55.56
MIN	100.00	80.56	61.11
RAN	94.44	83.33	66.67
SD	95.56	80.00	63.33
MAD	85.56	72.78	56.67
MAV	99.44	85.00	70.00

Table 6
Comparison of recognition rates of SVM and GRNN using different features.

Classifier	SVM (%)	GRNN (%)
$D_3 + D_4$	71.11	63.33
$D_3 + D_5$	66.67	60.00
$D_4 + D_5$	61.11	62.22
$D_3 + D_4 + D_5$	72.22	70.00

In these experiments, the features D_3 , D_4 , and D_5 were used to compute the eigenvalues because the theta, alpha, and beta brainwaves are associated with an attentive state. However, to explore the overall significance and relative importance of the EEG components of D_3 , D_4 , and D_5 , classification was carried out using the SVM and GRNN with only two DWT components (see Table 6). The results show that D_3 and D_4 are more important than D_5 , but the SVM and GRNN classifiers using the $D_3 + D_5$ components of the SWT showed a better recognition rate than the other two cases. In a comparison of the SVM and GRNN, using the $D_3 + D_4$ components, the SVM using $D_3 + D_4 + D_5$ showed a small improvement of the recognition rate and is better than SVM using only $D_3 + D_4$. However, the recognition rate of the GRNN using $D_3 + D_4 + D_5$ is much better than GRNN using $D_3 + D_4$. As shown in Table 1, the human attentive state is closely associated with alpha and beta brainwaves. The experimental results shown in Table 6 are clearly consistent with Table 1.

4. Classification of Rider Vigilance Using SVM and GRNN

In the Case I experiments, the brainwaves of a rider following a straight path and one with obstacles were analyzed using the SVM and GRNN. In Case II, rider vigilance was investigated during various cycling-related activities, as shown in Fig. 5, also using the SVM and GRNN. Four different cycling-related activity patterns were studied: (1) riding without any other activity, (2) listening to music, (3) looking at a cellphone screen, and (4) communicating using a cellphone.

Case II. Classification of mind states for the different cycling-related activities

In Case II, six subjects performed eight different scenarios with the different cycling-related activities described in Fig. 5. Scenarios 1–4 involve riding along a straight path while engaging in the following activities: (1) no additional activity, (2) listening to music, (3) looking at the screen of a cellphone, and (4) while communicating with a cellphone. Scenarios 5–8 involve following a path that has obstacles to be avoided and included the same activities as in scenarios 1–4.

Similarly, this pattern recognition problem required that the classes of samples to train the SVM and GRNN classifiers be defined. Once the classifiers were trained, it was necessary to evaluate their performance on an independent (labeled) test set assuming that there were two different mind states, “relaxed” and “attentive”. Table 7 shows the specified label for the eight scenarios. The EEG signals of the six riders were captured using the MindWave headset as they rode; 30 sequences of the EEG signal were captured for each incident. There were 180 samples of EEG signals for each scenario, making a total of 1440 samples for classification. The performance was evaluated as in Case I. There were three testing modes, internal training, the half-external test, and the external test. For the half-external test, half the samples were used to train the classifiers, but all were used for testing. Table 8 shows the experimental results for the SVM and GRNN. The experimental results showed that the GRNN had higher performance than the SVM in the internal training and half-external test, but the SVM performed slightly better than the GRNN in the external test.

The experimental results for the internal training and half-external test in Cases I and II showed that the GRNN had better recognition performance than the SVM. However, the SVM classifier had slightly better performance than the GRNN in the external test. Overall, the GRNN had superior performance to the SVM, and the following comments can be made:

- (1) Riding around obstacles requires more attention than riding along a straight path.
- (2) Looking at the screen of a cellphone distracts the rider’s attention and this can be dangerous in traffic.
- (3) Communicating with a cellphone requires action on the riders part. This diverts attention from the task of riding, which can be dangerous in traffic.
- (4) Riding along a straight path while listening to music may be safe because it does not require much attention.

In summary, looking at a cellphone screen or communicating with a cellphone can cause a rider to miss important peripheral stimuli and can cause traffic accidents. Cellphones should not be used while cycling.

Table 7
Mind state: relaxed/attentive for eight scenarios.

Subjects 1–6			
Riding along a straight path			
Scenario 1	Scenario 2	Scenario 3	Scenario 4
*	*	⊙	⊙
Riding while avoiding obstacles			
Scenario 5	Scenario 6	Scenario 7	Scenario 8
⊙	⊙	⊙	⊙

Relaxed: *, Attentive: ⊙

Table 8
SVM and GRNN recognition rates: Case II.

Subjects 1–6			
Classifier	Internal training	Half-external test	External test
(Training/Test)	1440/1440	720/1440	720/720
SVM	86.04%	79.03%	71.11%
GRNN	99.58%	85.28%	70.69%

5. Conclusions

The NeuroSky EEG headset was used to capture the EEG brainwave signals of subjects as they rode along a predetermined track. To extract the features of the EEG signals, the Daubechies Db4 wavelet components were used in a decomposition analysis of the EEG signals. The brainwave types are shown in Table 1, and on their basis, the EEG signals were decomposed into five DWT levels. The input signal $x(n)$ was recursively decomposed into six subband signals: an approximation signal $A_5(n)$ and five detailed signals, $D_1(n)$, $D_2(n)$, $D_3(n)$, $D_4(n)$, and $D_5(n)$. Human attention is related to the theta, alpha, and beta brainwaves, and the D_3 , D_4 , and D_5 DWT bands were used to compute the eigenvalues. The best eigenvalues for classification were found by determining the recognition rates of the SVM and GRNN using six different eigenfunctions, as shown in Tables 4 and 5. In Case I, the SVM and GRNN were used to recognize the mind states of two different cycling activities, riding along a straight path and riding along a path with obstacles. The SVM had a better recognition rate than the GRNN with the Max and MAV eigenfunctions. To summarize, for the external test, the SVM using MAV was the better classifier with a recognition rate of 72.22%, compared with 70.00% for the GRNN. In Case II, rider vigilance was investigated using the SVM and GRNN for eight different cycling scenarios. The experimental results showed that the GRNN exhibited better classification than the SVM in the internal training and half-external test, but the SVM had slightly better performance than the GRNN in the external test. From the experimental results, we conclude that cyclists should not use cellular phones when riding, because looking at the screen or actually communicating with a cellphone can reduce a rider's ability to notice obvious peripheral stimuli that occur in real traffic situations.

Acknowledgments

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financial support of this research under Contract Nos. MOST 106-2221-E-027-058 and MOST 108-2218-E-167-006 and the APC was funded by National Taipei University of Technology.

References

- 1 W. Brodsky: Transportation Res. Part F: Traffic Psychol. Behav. **4** (2001) 219. [https://doi.org/10.1016/s1369-8478\(01\)00025-0](https://doi.org/10.1016/s1369-8478(01)00025-0).
- 2 W. Consiglio, P. Driscoll, M. Witte, and W. P. Berg: Accid. Anal. Prev. **35** (2003) 495. [https://doi.org/10.1016/s0001-4575\(02\)00027-1](https://doi.org/10.1016/s0001-4575(02)00027-1)
- 3 V. Cantin, M. Lavalliere, M. Simoneau, and N. Teasdale: Accid. Anal. Prev. **41** (2009) 763. <https://doi.org/10.1016/j.aap.2009.03.019>
- 4 K. A. Brookhuis and D. de Waard: Accid. Anal. Prev. **42** (2010) 898. <https://doi.org/10.1016/j.aap.2009.06.001>
- 5 W. Brodsky and Z. Slor: Accid. Anal. Prev. **59** (2013) 382. <https://doi.org/10.1016/j.aap.2013.06.022>
- 6 B. Dalton, D. Behm, and A. Kibele: Occup. Ergon. **7** (2007) 153.
- 7 N. Dibben and V. J. Williamson: Psychol. Music **35** (2007) 571. <https://doi.org/10.1177/0305735607079725>
- 8 D. B. Bellinger, B. M. Budde, M. Machida, G. B. Richardson, and W. P. Berg: Transport Res. F-Traf **12** (2009) 441. <https://doi.org/10.1016/j.trf.2009.08.007>
- 9 C. Pêcher, C. Lemerrier, and J.-M. Cellier: Safety Sci. **47** (2009) 1254. <https://doi.org/10.1016/j.ssci.2009.03.011>

- 10 J. K. Caird, C. R. Willness, P. Steel, and C. Scialfa: *Accid. Anal. Prev.* **40** (2008) 1282. <https://doi.org/10.1016/j.aap.2008.01.009>
- 11 O. Carsten and K. Brookhuis: *Transport Res. F-Traf* **8** (2005) 75. <https://doi.org/10.1016/j.trf.2005.04.003>
- 12 C. Collet, A. Guillot, and C. Petit: *Ergonomics* **53** (2010) 602. <https://doi.org/10.1080/00140131003769092>
- 13 D. de Waard, P. Schepers, W. Ormel, and K. Brookhuis: *Ergonomics* **53** (2010) 30. <https://doi.org/10.1080/00140130903381180>
- 14 C. Goldenbeld, M. Houtenbos, E. Ehlers, and S. W. O. Verkeersveiligheid: *Gebruik Van Draagbare Media-Apparatuur En Mobiele Telefoons Tijdens Het Fietsen: Resultaten Van Een Grootschalige Internetenquête* (Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, 2010).
- 15 D. de Waard, K. Edlinger, and K. Brookhuis: *Transportation Res. Part F: Traffic Psychol. Behav.* **14** (2011) 626. <https://doi.org/10.1016/j.trf.2011.07.001>
- 16 C. Goldenbeld, M. Houtenbos, E. Ehlers, and D. De Waard: *J. Safety Res.* **43** (2012) 1. <https://doi.org/10.1016/j.jsr.2011.08.007>
- 17 D. De Waard, B. Lewis-Evans, B. Jelijs, O. Tucha, and K. Brookhuis: *Transportation Res. Part F: Traffic Psychol. Behav.* **22** (2014) 196. <https://doi.org/10.1016/j.trf.2013.12.003>
- 18 I. E. Hyman, S. M. Boss, B. M. Wise, K. E. McKenzie, and J. M. Caggiano: *Appl. Cognit. Psychol.* **24** (2009) 597. <https://doi.org/10.1002/acp.1638>
- 19 V. Shek and E. Schubert: *Proc. 2nd Int. Conf. Music Communication Science (ICoMCS2)* (2009). <https://doi.org/10.1177/1029864909013002051>
- 20 A. North and D. Hargreaves: *The Social and Applied Psychology of Music* (OUP Oxford, 2008). <https://doi.org/10.1093/acprof:oso/9780198567424.003.0001>
- 21 M. D. van der Zwaag, C. Dijksterhuis, D. de Waard, B. L. Mulder, J. H. Westerink, and K. A. Brookhuis: *Ergonomics* **55** (2012) 12. <https://doi.org/10.1080/00140139.2011.638403>
- 22 A. S. Nithiya and K. J. I. Saraladevi: **3** (2013). <https://doi.org/10.26634/jpr.3.4.13538>
- 23 K. Crowley, A. Sliney, I. Pitt, and D. Murphy: *2010 10th IEEE Int. Conf. Advanced Learning Technologies* (2010) 276. <https://doi.org/10.1109/ICALT.2010.81>
- 24 C. T. Lin, C. H. Chuang, C. S. Huang, S. F. Tsai, S. W. Lu, Y. H. Chen, and L. W. Ko: *IEEE Trans. Biomed. Circuits Syst.* **8** (2014) 165. <https://doi.org/10.1109/TBCAS.2014.2316224>
- 25 D. Gabor: *J. Inst. Electr. Eng. Part III: Radio Commun. Eng.* **93** (1946) 429. <https://doi.org/10.1049/ji-3-2.1946.0074>
- 26 P. Goupillaud, A. Grossmann, and J. Morlet: *Geoexploration* **23** (1984) 85. [https://doi.org/10.1016/0016-7142\(84\)90025-5](https://doi.org/10.1016/0016-7142(84)90025-5)
- 27 I. Daubechies: *Commun. Pure Appl. Math.* **41** (1988) 909. <https://doi.org/10.1002/cpa.3160410705>
- 28 R. X. Gao and R. Yan: *Wavelets: Theory and Applications for Manufacturing* (Springer Science & Business Media, 2010). <https://doi.org/10.1007/978-1-4419-1545-0>
- 29 A. Haar: *Mathematische Annalen* **69** (1910) 331. <https://doi.org/10.1007/bf01456326>
- 30 C. Parameswariah and M. Cox: *IEEE Trans. Power Delivery* **17** (2002) 800. <https://doi.org/10.1109/TPWRD.2002.1022806>
- 31 B. I. Morshed: *J. Bioeng. Biomed. Sci.* **04** (2014) 1. <https://doi.org/10.4172/2155-9538.1000128>
- 32 F. Wang, Z. Zhen, B. Wang, and Z. Mi: *Appl. Sci.* **8** (2017). <https://doi.org/10.3390/app8010028>
- 33 J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis* (Cambridge University Press, 2004). <https://doi.org/10.1017/CBO9780511809682.002>
- 34 C.-J. Lin, H. C. Chuang, C. W. Hsu, and C. S. Chen: *Sens. Mater.* **29** (2017) 1625. <https://doi.org/10.18494/sam.2017.1736>
- 35 A. Bulling, J. A. Ward, H. Gellersen, and G. Troster: *IEEE Trans. Pattern Anal. Mach. Intell.* **33** (2011) 741. <https://doi.org/10.1109/TPAMI.2010.86>
- 36 S. Y. Dai, D. X. Niu, and Y. R. Han: *Appl. Sci.-Basel* **8** (2018). <https://doi.org/ARTN 63610.3390/app8040636>
- 37 S. U. Jan, V. H. Vu, and I. Koo: *Appl. Sci.-Basel* **8** (2018). <https://doi.org/ARTN 42110.3390/app8030421>
- 38 Z. D. Mu, J. F. Hu, and J. L. Min: *Appl. Sci.-Basel* **7** (2017). <https://doi.org/ARTN 15010.3390/app7020150>
- 39 C. Cortes and V. Vapnik: *Mach. Learning* **20** (1995) 273. <https://doi.org/10.1007/Bf00994018>
- 40 D. F. Specht: *IEEE Trans. Neural Netw.* **2** (1991) 568. <https://doi.org/10.1109/72.97934>

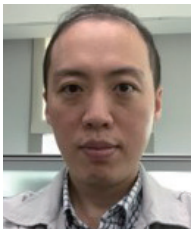
About the Authors



Wen-Lin Chu received her B.S. degree from the Department of Mechanical and Computer-Aided Engineering, Feng Chia University in 2009, her M.S. degree from the Graduate Institute of Automation Control from the National Taiwan University of Science and Technology in 2011, and her Ph.D. degrees from the Department of Biomedical Engineering, National Cheng Kung University in 2017. She is currently an assistant professor with the Department of Mechanical Engineering, National Chin-Yi University of Technology, Taichung, Taiwan. Her current research interests include image processing, physiological signal processing, and automatic analysis techniques. (wlchu@ncut.edu.tw)



Chih-Jer Lin received his B.S., M.S., and Ph.D. degrees in mechanical engineering from National Cheng Kung University, Tainan, Taiwan, in 1992, 1994, and 1998, respectively. He is currently a professor and the director of the Graduate Institute of Automation and Technology, National Taipei University of Technology, Taipei, Taiwan. His current research interests include mechatronics, vibration control, motion control, system identification, sliding mode control, robotics, and evolutionary algorithms. Professor Lin has been a guest editor of Applied Sciences and a member of the editorial board of the Journal of Chinese Society of Mechanical Engineers. (cjlin@ntut.edu.tw)



Po-Wei Hsiao received his M.S. degree in automation technology from National Taipei University of Technology, Taiwan, in 2017. He is currently an engineer in Chunghwa TeleCom. His research interests are in mechatronics and neural networks. (pwhsiao99@gmail.com)