Modular-neural-network-based Decision Fusion for Steady-state Visually Evoked Potential-based Brain–Computer Interfaces

Yeou-Jiunn Chen,[1,2] Shih-Chung Chen,[1] and Chung-Min Wu[3*]

[1]Department of Electrical Engineering, Southern Taiwan University of Science and Technology,
No. 1, Nan-Tai Street, Yungkang Dist., Tainan City, 710, Taiwan
[2]Allied AI Biomedical Research Center, No. 1, Nan-Tai Street, Yungkang Dist., Tainan City, 710, Taiwan
[3]Department of Intelligent Robotics Engineering, Kun Shan University,
No. 195, Kunda Rd., Yungkang Dist., Tainan City, 710, Taiwan

It is very difficult for a patient with severe disabilities to communicate with others or devices, greatly reducing the quality of their lives. In this study, a steady-state visually evoked potential (SSVEP)-based brain–computer interface (BCI) is proposed to make it easy for patients with severe disabilities to communicate. To precisely represent the characteristic of an elicited SSVEP, the four features extracted by fast Fourier transform, canonical correlation analysis, magnitude-squared coherence, and power cepstrum analysis are used. To fuse the decision results obtained by using the different features, a modular neural network (MNN), which includes input models and a decision model, is adopted to improve the recognition performance. To balance the recognition performance and computational complexity, an artificial neural network based on multilayer perceptrons is selected as the basic unit of the MNN. On the basis of the different features, the input model of the MNN can quickly find the decision results. To effectively fuse the decision results, the decision model of the MNN is adopted to obtain a precise decision. The experimental results demonstrated that the MNN has a higher accuracy than other approaches. Therefore, the proposed SSVEP-based BCI with the MNN can effectively help patients interact with their surroundings.

1. Introduction

The communication of a patient with severe disabilities is greatly degraded by the decrease in the sizes of their muscles.[1] A patient with amyotrophic lateral sclerosis, which is characterized by stiff muscles, muscle twitching, and gradually increasing weakness, finds it difficult to speak, swallow, and eventually breathe. However, the sensory nerves and the autonomic nervous system are unaffected. Therefore, patients maintain their hearing, sight, touch, smell, and taste, and their cognitive function is generally spared. Developing an assistive communication system would greatly help these patients in their daily lives and maintain their dignity.

To help patients communicate with others or devices, augmentative and alternative communication (AAC) systems have been developed.[2–5] Unfortunately, these AAC systems, which need to use the voice, mouth movement, and tongue, or require the voluntary control of the user's limbs, are unsuitable communication mechanisms for patients with severe disabilities. Recently, brain–computer interfaces (BCIs) have been successfully applied to implement AAC systems by electroencephalography (EEG) analysis.[6–12] BCIs enable patients to directly communicate with others or send commands to an external device by measuring the brain activity. The brain activity patterns used in BCIs can be categorized into P300 potentials,[6,7] steady-state visually evoked potentials (SSVEPs),[8,9] and motor imagery.[10–12] For these approaches, SSVEP-based BCIs, which are stimulated by a visual stimulus with a specific frequency, can achieve an excellent signal-to-noise ratio, making them very suitable for practical applications. Therefore, the development of SSVEP-based BCIs would greatly benefit patients in their daily lives.

Accurately representing SSVEP signals is one of the effective ways of improving the performance of SSVEP-based BCIs.[13,14] To find the frequencies corresponding to the visual stimulus frequencies, the fast Fourier transform (FFT) and Hilbert–Huang transform were developed and used to obtain the corresponding commands or messages.[15] Canonical correlation analysis (CCA) and magnitude-squared coherence (MSC) were proposed to find the corresponding correlation between the stimulus and a set of predefined pure sine and cosine reference templates in the time and frequency domains, respectively.[16,17] Power cepstrum analysis (PCA) was also adopted to represent the characteristics of SSVEP signals.[18] Therefore, fusing the results of different features should be able to effectively improve the performance of SSVEP-based BCI systems.

Over the last decade, artificial neural networks have been widely applied to many applications.[18–20] Neural networks have been applied to pattern recognition with superior performance to traditional methods. Researchers have started to build hybrid systems that combine the advantages of modular neural networks (MNNs). MNNs have the capability to learn different tasks simultaneously and reduce the complexity of systems. They have robustness and can be made fault-tolerant. Hence, MNNs can be adopted to integrate the decisions obtained by different features, enabling the accuracy of SSVEP-based BCI systems to be greatly improved.

To help patients with severe disabilities, an SSVEP-based BCI is developed as an AAC system in this study. To precisely represent the characteristics of EEG signals, four features are extracted by FFT, CCA, MSC, and PCA. To fuse the decision results determined by using these different features, an MNN consisting of two layers including input modules and an additional decision module is proposed to find a precise decision, which is used as a command or message for an SSVEP-based BCI system.

The remainder of this paper is organized as follows. Section 2 describes the SSVEP signal acquisition, feature extraction, and MNN. Section 3 presents the results of a series of experiments to evaluate the performance of our approach. Finally, conclusions and possible improvements for the future development of this system are given in Sect. 4.

2.   Materials and Methods

In this study, an MNN is adopted to develop the SSVEP-based BCI system whose block diagram is shown in Fig. 1. The system involves visual stimulation, EEG acquisition, feature extraction, and the MNN. First, the EEG signals for the corresponding responses are stimulated by visual stimulation. Second, four features are adopted to precisely represent the characteristics of the frequency responses. Third, for each feature, the input modules in the first layer of the MNN are used to recognize the corresponding results. Finally, the recognized results identified in the first layer of the MNN are fused by an additional decision module in the second layer of the MNN. These procedures are described in detail as follows.

2.1   Visual stimulation and EEG acquisition

To elicit SSVEPs, five blinking boxes are displayed on a 20" LCD screen. These blinking boxes, which flicker at 6.00, 6.67, 7.50, 8.57, and 10.00 Hz, are used to represent five different commands.[20] To reduce the interference between the blinking boxes, each box is arranged as shown in Fig. 1. Subjects are asked to sit in front of the LCD screen at a distance of 55 cm.

When the visual stimulation elicits EEG signals in the visual cortex of the brain, the EEG signals are acquired with an A/D converter with a 22-bit resolution and a 1 kHz sampling rate by using a NuAmps EEG amplifier supplied by Neuroscan Company. The EEG amplifier provides 37 channels for EEG acquisition, and the electrode placement follows the international 10-20 electrode placement system. In this study, the EEG signals were acquired from the Oz channel with reference and ground electrodes placed at the A1 and A2 channels, respectively.

2.2   Feature extraction

When an EEG signal, x, with N samples is acquired, four features are extracted in this study. The first one is the frequency magnitude estimated by FFT. The FFT is adopted to estimate the log magnitude of the spectrum, $X(k)$. The second one is estimated by CCA, which is used to explore the underlying correlation between an input signal, x, and a set of reference signals, y.[17] For CCA, a pair of linear transforms, w_x and w_y, is used such that the correlation ρ between linear $\tilde{x} = w_x^T x$ and $\tilde{y} = w_y^T y$ combinations is maximized. Then, the correlation can be defined as
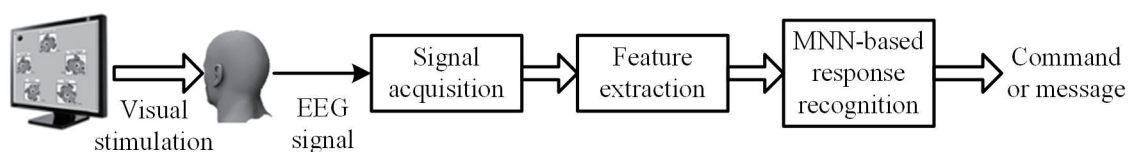

Fig. 1.   Block diagram of proposed SSVEP-based BCI.

$$\rho = \max_{w_x, w_y} \frac{w_x^T \sum_{xy} w_y}{\sqrt{w_x^T \sum_{xx} w_x w_y^T \sum_{yy} w_y}}, \tag{1}$$

where $\sum_{xx} = xx^T$ and $\sum_{yy} = yy^T$ are the within-set covariance matrices and $\sum_{xy} = xy^T$ is the between-set covariance matrix. Any rescaling of w_x and w_y does not affect the correlation maximization, so Eq. (1) is equivalent to

$$\rho = \max_{w_x, w_y} w_x^T \sum_{xy} w_y, \tag{2}$$

such that $w_x^T \sum_{xx} w_x = w_y^T \sum_{yy} w_y = 1$. Lagrange multipliers are then applied to transform Eq. (2) into the following generalized eigenvalue problems:

$$\sum_{xy} \sum_{yy}^{-1} \sum_{yx} w_x = \rho^2 \sum_{xx} w_x \tag{3}$$

and

$$\sum_{yx} \sum_{xx}^{-1} \sum_{xy} w_y = \rho^2 \sum_{xx} w_x . \tag{4}$$

The linear transforms w_x and w_y resulting in the largest canonical correlation between x and y are given by the eigenvectors corresponding to the largest generalized eigenvalues. The set of reference signals is formed by a series of sine-cosine waves.

The third feature is extracted by MSC, which is used to determine the correlation between x and y at each frequency.[16] The MSC at each stimulus frequency k, $M_{xy}(k)$, can be defined as

$$M_{xy}(k) = \frac{\left| G_{xy}(k) \right|^2}{G_{xx}(k) G_{yy}(k)}, \tag{5}$$

where G_{xx} and G_{yy} are the auto-spectral densities of x and y, respectively, and G_{xy} is the cross-spectral density between x and y. The reference signal y is a pure sinusoidal wave and its frequency is equal to the stimulus frequency.

The last type of feature used in this study is extracted by PCA.[18] PCA can effectively transform the EEG signals in low-dimensional space and keep most of the information by discrete cosine transform. PCA is applied to linearly transform the log magnitude of the spectrum, $X(k)$, into the cepstrum domain and obtain the power cepstrum coefficient A, which is derived as

$$A(k) = \sum_{b=1}^{B-1} X(m)\cos\left[\frac{\pi}{B}\left(b+\frac{1}{2}\right)k\right],\qquad\qquad(6)$$

where B is the number of frequency bins in $X(m)$.

2.3 MNN

The structure of the MNN used in this study has two layers of neural models.[21] The first layer is the aggregation of several separately trained subnetworks, which are used as input models. Each input model is trained using a particular feature subset. The outputs of these input models are treated as the inputs of a subnetwork, which is a decision model, in the second layer of the MNN. The decision model is applied to fuse the decisions obtained using the different features. The input models and the decision model are designed as multilayer perceptrons (MLPs).

Figure 2 shows the architecture of the MNN for the SSVEP-based BCI system. The features estimated by FFT, CCA, MSC, and PCA are separately fed into the FFT-based, CCA-based, MSC-based, and PCA-based MLPs, respectively. These MLPs are individually trained using the different features and then each MLP outputs its individual decision. Finally, the aggregation of these decisions is implemented by a decision MLP and then a final decision can be obtained by the MNN.

The MLPs used in this study are feedforward neural networks wherein the connections between the nodes do not form a cycle. An MLP consists of at least three layers of nodes, including an input layer, a hidden layer, and an output layer. Except for the input nodes, each node is a neuron that uses a sigmoid function as the nonlinear activation function in this study. To train the MLP, backpropagation is used to calculate the gradient needed in the calculation of the weights, which are used in the network.
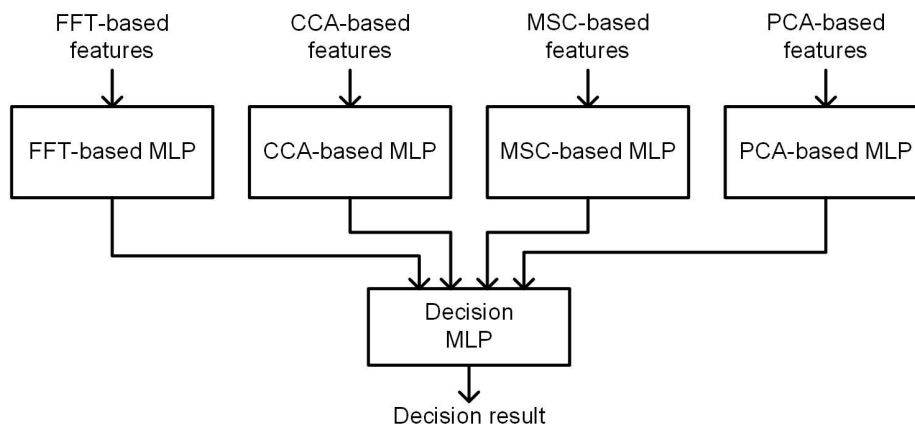
Fig. 2.   Architecture of MNN for SSVEP-based BCI system.

3.　　Experimental Results and Discussion

In this study, 15 subjects were asked to participate in the experiments, and the leave-one-out method was applied to evaluate the proposed approach. The sampling rate and the frame size were set to 1 kHz and 1 s, respectively. Subsequently, a low-pass filter with a cutoff frequency of 50 Hz was used to obtain the downsampled signal with a sampling rate of 100 Hz. Each subject had 30 epochs for each specific frequency. Using a 1024-point FFT, the log magnitude spectrum was estimated and a triangle filter with a 2 Hz bandwidth was applied to estimate the log magnitude of specific frequencies. The order of PCA and the number of harmonics for CCA were set to 8 and 4, respectively. To compare the proposed approach, the decision method of identifying the maximum magnitude of the corresponding frequencies was selected as the baseline system and the results are shown in Table 1. For the feature of PCA, the maximum magnitude of PCA cannot be directly transformed to the frequencies; thus, it was not compared in this experiment.

3.1　　Experimental results of MLP

As reported in this subsection, a feedforward artificial neural network was selected as the MLP and used to examine the different features. The number of hidden layers was set to 1 and the number of hidden nodes examined ranged from 4 to 9. The experimental results of the MLP obtained by FFT, CCA, MSC, and PCA are shown in Table 2 and denoted as NN(FFT), NN(CCA), NN(MSC), and NN(PCA), respectively. When the number of hidden nodes is 5, the recognition rates for each type of feature achieve acceptable results. It is also clear that NN(CCA) outperforms Baseline(CCA), but the performance characteristics of NN(FFT) and NN(MSC) are slightly degraded. When the number of hidden nodes is 7, the recognition rates

Table 1
Recognition rates for baseline system obtained by FFC, CCA, and MSC, denoted as Baseline(FFT), Baseline(CCA), and Baseline(MSC), respectively.

	Stimulus frequency (Hz)					
	6.00	6.67	7.50	8.57	10.00	Average
Baseline(FFT)	94.22	91.33	86.89	82.67	82.00	87.42
Baseline(CCA)	95.11	92.00	91.11	89.33	84.22	90.35
Baseline(MSC)	94.44	84.67	75.11	62.44	57.11	74.75

Table 2
Experimental results (%) of MLP using different features.

	Number of hidden nodes					
	4	5	6	7	8	9
NN(FFT)	83.58	86.78	87.22	87.58	87.58	87.71
NN(CCA)	91.37	91.80	92.03	92.13	92.28	92.43
NN(MSC)	77.72	78.35	78.61	79.10	78.97	78.70
NN(PCA)	84.02	90.40	90.64	90.64	90.65	90.67

for NN(FFT) and NN(MSC) are 87.58 and 79.10%, respectively. According to the experimental results, the features of FFT and MSC are imprecise for identifying the complex responses of EEG signals.

3.2 Experimental results of MNN

In this subsection, we present the results of the proposed approach and the effects of combining it with different features. In accordance with the previous experimental results, the number of hidden nodes in the first layer of the MNN was set to 5. The correlation between pairs of features was examined first. The number of hidden nodes in the second layer of the MNN was set from 4 to 9 and the experimental results are shown in Fig. 3. It is clear that the recognition rates can be effectively improved when decisions decided using two features are fused. Since CCA achieves the best performance when using a single MLP, CCA can precisely represent the EEG signals. The features fused using CCA features are improved from 86.78, 78.35, and 90.40 to 92.18, 88.04, and 90.93% for NN(FFT), NN(MSC), and NN(PCA), respectively. The results clearly show that CCA improves the recognition rates of the other features.

The experimental results obtained using three and four features are shown in Fig. 4. The comparison of the results in Figs. 3 and 4 shows that fusing any three features outperforms fusing any two features and that fusing four features achieves the best performance. Therefore,
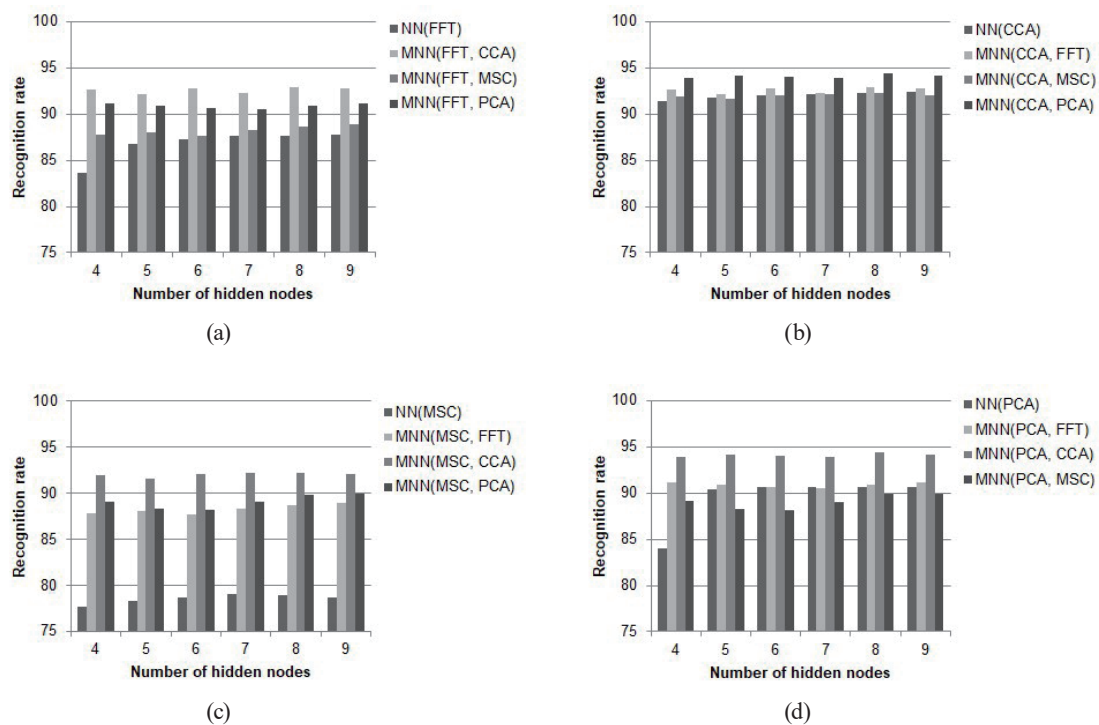

Fig. 3.	Experimental results for fusing (a) FFT, (b) CCA, (c) MSC, and (d) PCA with other features.
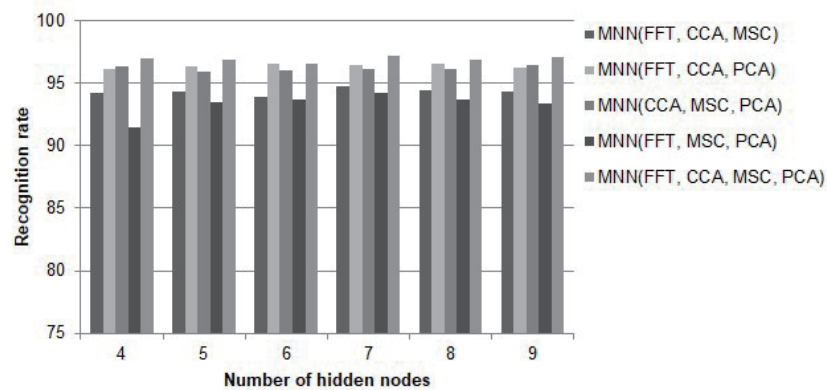
Fig. 4.   Experimental results of MNN with different feature combinations.

Table 3
Experimental results for MNNs and MLPs with different combinations of features.

	FFT, CCA, MSC	FFT, CCA, PCA	CCA, MSC, PCA	FFT, MSC, PCA	FFT, CCA, MSC, PCA
MNNs	94.31	96.36	95.91	93.51	96.88
MLPs	92.87	92.24	93.58	90.71	93.21

these features used in the MNN can effectively represent the characteristics of EEG signals and thus be used to make appropriate decisions. By using five hidden nodes for the second layer of the MNN, the recognition rate is improved from 96.36% for MNN(FFT, CCA, PCA) to 96.88% for MNN(FFT, CCA, MSC, PCA), reducing the rate of error by 14.33%. The computational complexity should be considered when the recognition rates are similar. Hence, the MNN that uses fewer features would greatly increase the efficiency of practical applications.

3.3   Comparison with other approaches

The characteristic of the decision fusion for MNNs with different features is examined in this subsection. To investigate the effects of different features, the different features were combined into a new feature vector and treated as the input of MLPs. When the numbers of hidden layers and hidden nodes in the MLPs were set to 2 and 5, respectively, the numbers of hidden nodes in the MNNs and MLPs are similar. The experimental results for the MNNs and MLPs are shown in Table 3. It is clear that the recognition rates of the MNNs with different combinations of features are higher than those of the MLPs.

In Table 3, this result of the MLP is different from that of the MNN. In the MLP, the connections of the hidden nodes are full, and thus, the effects of different features are inferred in the hidden layers. Moreover, the recognition rate of MLP(CCA, MSC, PCA) is similar to that of MLP(FFT, CCA, MSC, PCA). Since the identification processes of FFT and MSC are in the frequency domain, the experimental results of MLP(FFT, CCA, PCA) and MLP(FFT, CCA, MSC, PCA) show that the identification ability of FFT is similar to that of MSC. Hence, the effect of FFT is small.

4.   Conclusions

In this study, an MNN is proposed to fuse the features of FFT, CCA, MSC, and PCA to improve the efficiency of SSVEP-based BCI systems. The features of FFT, CCA, MSC, and PCA can precisely represent the characteristics of EEG signals. Moreover, the MNN can effectively fuse the decisions obtained by the different features. By using features with distinct properties, the SSVEP-based BCI system can achieve a satisfactory performance for decision fusion. Using the features of CCA and PCA, the MNN can achieve an acceptable result for practical applications. The experimental results showed that the MNN outperforms an MPL with full connections. To achieve a similar result to the MNN, the number of hidden layers should be considerably increased in the MLP such as by using a deep learning approach, which greatly increases the computational complexity. Hence, the proposed approach is very suitable for practical applications and can effectively help patients interact with their surroundings. In the future, the performance of the MNN can be improved by selecting a suitable neural network in the MNN or by fusing the decisions obtained by using different features with distinct properties.

Acknowledgments

References

1   L. D. Usanova, A. D. Usanova, and A. V. Skripal: Biomed. Eng. **46** (2012) 25. https://doi.org/10.1007/s10527-012-9258-9

2   G. Hornero, D. Conde, M. Quilez, S. Domingo, M. P. Rodriguez, B. Romero, and O. Casas: IEEE Access **3** (2015) 1288. https://doi.org/10.1109/ACCESS.2015.2466110

3   A. Jafari, N. Buswell, M. Ghovanloo, and T. Mohsenin: IEEE Trans. Biomed. Circuits Syst. **12** (2018) 58. https://doi.org/10.1109/TBCAS.2017.2757031

4   M. Anila, and P. Radhika: Proc. Int. Conf. Wireless Communications, Signal Processing and Networking (IEEE, 2017) 1182–1187.

5   E. Radici, S. Bonacina, and G. D. Leo: Proc. 38th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (2016) 2574–2577.

6   V. Guy, M. H. Soriani, M. Bruno, T. Papadopoulo, D. Desnuelle, and M. Clerc: Ann. Phys. Rehabil. Med. **61** (2018) 5. https://doi.org/10.1016/j.rehab.2017.09.004

7   Y. Okahara, K. Takano, T. Komori, M. Nagao, Y. Iwadate, and K. Kansaku: Clin. Neurophysiol. Pract. **2** (2017) 147. https://doi.org/10.1016/j.cnp.2017.06.004

8   S. C. Chen, Y. J. Chen, I. A. E. Zaeni, and C. M. Wu: Int. J. Fuzzy Syst. **19** (2017) 553. https://doi.org/10.1007/s40815-016-0289-3

9   Y. J. Chen, S. C. Chen, I. A. E. Zaeni, and C. M. Wu: Appl. Sci. **6** (2016) 270. https://doi.org/10.3390/app6100270.

10 O. R. Pinheiro, L. R. G. Alves, and J. R. D. Souza: IEEE Lat. Am. Trans. **16** (2018) 254. https://doi.org/10.1109/TLA.2018.8291481

11 K. K. Ang and C. Guan: IEEE Trans. Neural Syst. Rehabil. Eng. **25** (2017) 392. https://doi.org/10.1109/TNSRE.2016.2646763

12 J. Jiang, Z. Zhou, E. Yin, Y. Yu, Y. Liu, and D. Hu: Comput. Biol. Med. **66** (2015) 11. https://doi: 10.1016/j.compbiomed.2015.08.011

13 O. Trigui, W. Zouch, M. B. Slima, and M. B. Messaoud: Multimedia Tools Appl. **78** (2019) 12865. https://doi.org/10.1007/s11042-018-6029-y

14 C. C. Lee, C. C. Chuang, C. H. Yeng, Y. J. Chen, and B. S. Lin: IEEE Signal Process. Lett. **24** (2017) 1783. https://doi.org/10.1109/LSP.2017.2761193

15 X. Ruan, K. Xue, and M. Li: Proc. 11th World Congr. Intelligent Control and Automation (2014) 2418–2423.

16 Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki: IEEE Trans. Neural Syst. Rehabil. Eng. **21** (2013) 887. https://doi.org/10.1109/TNSRE.2013.2279680

17 S. Tantisatirapong, P. Dechwechprasit, W. Senavongse, and M. Phothisonothai: Proc. 2017 9th Int. Conf. Knowledge and Smart Technology (2017) 279–283.

18 Y. J. Chen, A. R. A. See, and S. C. Chen: Electron. Lett. **50** (2014) 735. https://doi.org/10.1155/2015/234260

19 M. H. Ahn and B. K. Min: Proc. 2018 6th Int. Conf. Brain-Computer Interface (2018) 1–3.

20 S. C. Chen, C. M. Wu, I. A. E. Zaeni, and Y. J. Chen: Microsyst. Technol. **24** (2018) 199. https://doi.org/10.1007/s00542-016- 3229-0

21 R. Rojas: Neural Networks: A Systematic Introduction (Springer Science and Business Media, Germany, 2013).

About the Authors

Yeou-Jiunn Chen received his B.S. degree in mathematics from Tatung Institute of Technology, Taipei, Taiwan and his Ph.D. degree from the Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, in 1995 and 2000, respectively. He was with the Advanced Technology Center, Computer and Communications Laboratories, Industrial Technology Research Institute, from 2001 to 2005 as a researcher. He is currently a professor at the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan. His research interests include biomedical signal processing, spoken language processing, and artificial intelligence. Dr. Chen is a member of the Biomedical Engineering Society, Taiwan Rehabilitation Engineering and Assistive Technology Society, and Association for Computational Linguistics and Chinese Language Processing. (chenyj@stust.edu.tw)

Shih-Chung Chen received his B.S. degree from the Department of Electrical Engineering, Feng Chia University, Taichung, Taiwan, his M.S. degree from the Institute of Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, and his Ph.D. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1982, 1988, and 2000, respectively. He is a professor and has been with the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, for 32 years. His research interests include brain–computer interfaces, biomedical signal processing, system integration, and assistive device implementation. He is a member of the Taiwanese Society of Biomedical Engineering and Taiwan Rehabilitation Engineering and Assistive Technology Society. (chung@stust.edu.tw)

Chung-Min Wu received his B.S. degree in automatic control engineering from Feng Chia University, Taichung, Taiwan, his M.S. degree in biomedical engineering from National Cheng Kung University, Tainan, Taiwan, and his Ph.D. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1994, 1998, and 2004, respectively. He is an associate professor in the Department of Intelligent Robotics Engineering, Kun Shan University. His research interests include fuzzy control, biomedical signal processing, and assistive tool implementation. He is a member of the Taiwanese Society of Biomedical Engineering and Taiwan Rehabilitation Engineering and Assistive Technology Society. (cmwu@mail.ksu.edu.tw)