Fuzzy-controlled Image Morse Code Input System

Chung-Min Wu,[1] Cheng-Fa Yen,[2] Yeou-Jiunn Chen,[3]
Shih-Chung Chen,[3*] and Cheng-Chi Tai[2]

[1]Department of Intelligent Robotics Engineering, Kun Shan University,
No. 195, Kunda Rd., Yongkang Dist., Tainan City 710303, Taiwan
[2]Department of Electrical Engineering, Nation Cheng Kung University,
No. 1, University Road, Tainan City 701, Taiwan
[3]Department of Electrical Engineering, Southern Taiwan University of Science and Technology,
No. 1, Nan-Tai Street, Yungkang Dist., Tainan City 710301, Taiwan

An image assistive communication system with a real-time fuzzy-controlled image Morse code input (FCIMCI) function was developed to help severely disabled people, such as those with cerebral palsy (CP) or spinal cord injury (SCI), communicate with other people or machines by using lip shape image recognition without using any apparatus. The real-time FCIMCI system offers the functions of text input and mouse control, replacing the traditional keyboard and mouse. The self-developed image recognition techniques include skin color detection, face-shape confirmation, lip location, and opened/closed lip status identification for Morse code conversion. In this paper, we apply the adjustable fuzzy recognition algorithm to tune the threshold values automatically for identifying Morse code. The experimental results of real-time lip image recognition show an average accuracy rate higher than 94% under various environmental conditions. Meanwhile, the average accuracy rate of image Morse code recognition experiments is more than 93.85%. According to the experimental results and a case study of SCI, we demonstrate that the FCIMCI system has real-time performance and good recognition efficiency. It is a novel assistive communication system based on image Morse code techniques and has the practicality and feasibility to help severely disabled people communicate with others.

1. Introduction

The rapid growth of IT has led to the widespread use of PCs. Keyboards and mice are the necessary peripherals of a PC. However, many quadriplegics still cannot operate computers by using traditional keyboards or mice since these devices are not designed for disabled people. Therefore, it is extremely important to design a simple computer input device to replace the traditional keyboard and mouse so that disabled people, especially quadriplegics, can operate their computers. Many assistive input systems for PCs have been developed for disabled people over the past decades. In the past, disabled people used to hit a switch to scan a matrix of letters,

symbols, words, or phrases for inputting text and other data. Later, a dynamically adapting matrix row and column scan method was proposed to increase the text entry rate of users. However, it was still not fast enough to meet user demands. To further help disabled people, specialized assistive computer input devices have been proposed in recent years, including voice-controlled devices,[1] tongue computer interfaces,[2] eye-controlled devices,[3] inductive head-controlled devices,[4] IR-controlled devices,[5] and brain–computer interface (BCI)-controlled devices.[6]

An eye-controlled device[3] based on electrooculography (EOG) was developed to replace mouse control. However, it caused discomfort to users and induced noise interference since there were five electrodes mounted on the user's face. A head-controlled device[4] employed two tilt sensors placed in a headset to determine the head position. It functioned as a simple head-operated computer mouse but it caused dizziness to users. An IR-controlled device[5] with two large keyboard panels and one mouse control panel utilized an IR-transmitting module to select the characters and control the mouse from the three panels. These complicated devices were also unsuitable for quadriplegics since they cannot move their head easily. A voice-controlled device[1] can control a mouse or a keyboard by voice recognition. This device is user friendly, but the user's voice has to be clear with little environmental noise. For inductive tongue computer interfaces,[2] the inductor must be mounted on the palatal plate and the magnetic material has to be placed directly on the tongue, which is neither comfortable nor hygienic. A BCI-controlled device[6] based on an electroencephalogram (EEG) produces a tri-state Morse code scheme, which can be translated into the English alphabet. The power spectral density (PSD) value of mental tasks in EEG can be used to classify the tri-state Morse code. This device can be used by people regardless of their disability, provided they have no intellectual disability. However, this device has several drawbacks, including complexity, a slow response time, and a high cost. It is also not easy to control the device using an EEG without practice.

As mentioned above, some of the devices are unsuitable and inconvenient for quadriplegics since most of the assistive systems require sensors or apparatus to be attached to disabled people. In this paper, we propose and develop a system that can realize real-time image Morse code input and compensate for the weaknesses of previous devices.

Morse code is useful in assistive technology (AT), augmentative and alternative communication (AAC), rehabilitation, and education.[7] Since 1996, several Morse code recognition algorithms have been proposed to raise the recognition rate in fields such as adaptive signal processing (adaptive variable-ratio threshold prediction algorithm,[8] the least mean squares algorithm[9]) and neural networks (backpropagation network,[10] fuzzy theory,[11] support vector machine[12] statistics, and learning vector quantization[13]). Traditional Morse code input is given using a switch or other related hardware components. In this paper, we propose a novel image Morse code input system that does not require any appurtenance to be worn for image recognition.

We describe image tracking and analysis techniques of the face and lips, which identify the opened or closed status of lips. After performing lip image recognition, we transfer the opened/closed status of the lips into image Morse code. We apply an adjustable fuzzy recognition algorithm to modify the threshold values of identifying a dot and dash in Morse code. This fuzzy

algorithm follows the rate of change of the user's lips to input text under non-familiar Morse code conditions automatically and efficiently. We integrate image recognition technology and the fuzzy recognition algorithm of Morse code to effectively implement a PC-based assistive communication system, which we call a fuzzy-controlled image Morse code input (FCIMCI) system.

Quadriplegics suffer from conditions such as spinal muscular atrophy (SMA), amyotrophic lateral sclerosis (ALS), and SCI. People with SMA cannot voluntarily control movement. People with ALS have a group of progressive neurological disorders that destroy cells controlling essential muscular activities such as speaking, walking, breathing, and swallowing. People with SCI suffer from paralysis of the arms, hands, trunk, and legs as their high cervical nerves (C1–C4) are damaged. We also discuss a case study of SCI in this paper.

The FCIMCI system architecture, as shown in Fig. 1, includes the following three major parts: face image processing, lip image recognition and translation into Morse codes, and a Morse code fuzzy recognition algorithm. After the Morse code fuzzy recognition is completed, the Morse codes are translated into ASCII codes using Windows API calls to complete text input or mouse control functions. The rest of the paper describes these parts in more detail.

2. Methods

The FCIMCI system offers real-time functions such as text input and mouse control by replacing the traditional keyboard and mouse. The self-developed image recognition techniques include skin color detection, face recognition, lip location, and opened/closed lip status recognition for Morse code conversion. We apply the adjustable fuzzy recognition algorithm to modify the threshold values of identifying Morse code. A flowchart of the image processing and recognition procedure of the FCIMCI system is shown in Fig. 2.
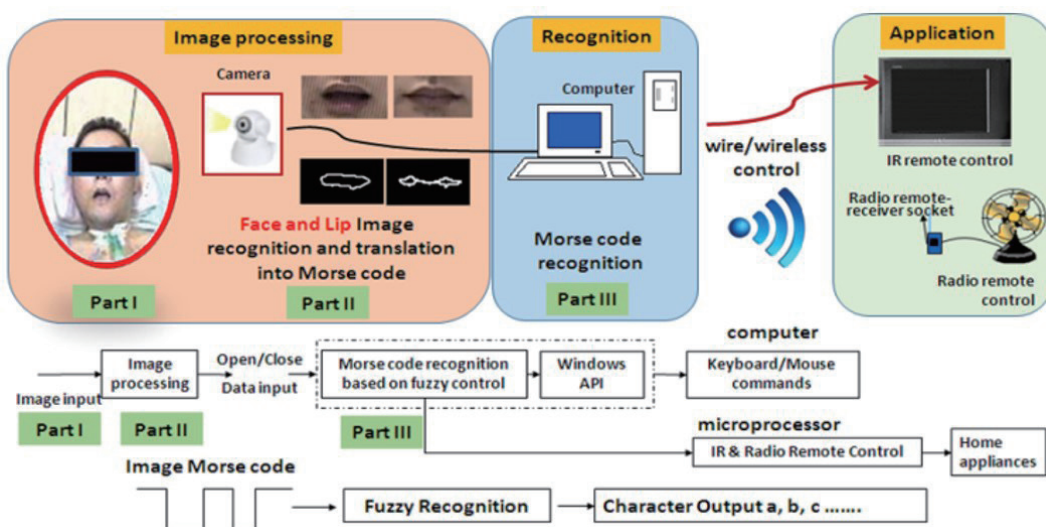

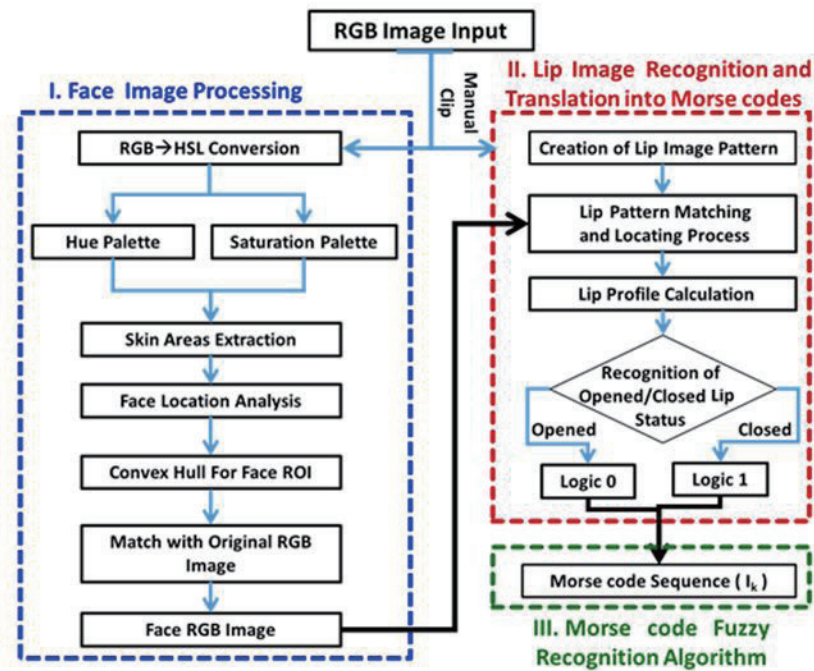Fig. 1.   (Color online) FCIMCI system architecture.

Fig. 2.   (Color online) Image processing and recognition procedure of FCIMCI.

2.1 Face image processing

In the facial image detection and tracking process, skin color area is a key parameter in the facial image recognition algorithm. However, if any object in the background contains similar skin colors, there is a possibility of a system failure. Consequently, other facial features in addition to skin color need to be taken into account (such as making sure the skin color is that of a human face). Moreover, the efficiency of the image recognition algorithm is an important factor in enhancing the effectiveness of optimal facial feature extraction and real-time facial image tracking.

If an algorithm has a high accuracy of facial feature extraction and a long computation time, it may result in a low frame rate and the algorithm being unable to operate in real time. On the basis of the above considerations, system adaptability and system execution time are key constraints to consider in face tracking and recognition systems.

There are four important conditions for system adaptability: the size of the face region, the face position, the background complexity, and the brightness (luminance) change in the environment. In addition, there are three important conditions to consider for the system execution time constraints: the input image size, the computer performance, and the algorithm efficiency for image processing.

As shown in Fig. 2, the algorithm of face detection and tracking for the image processing architecture in the FCIMCI system requires the face area to be located, which is achieved by the following five steps (a)–(e):

(a) RGB to hue, saturation, and lightness (HSL) conversion

The brightness information in an RGB image is affected by the light irradiation angle and intensity, which may result in recognition failure during the image processing procedure. In general, an image in RGB color space format is converted into a non-light-sensitive image,[14] for example, an image in the normalized RGB format [Eq. (1)] or HSL format [Eqs. (2)–(4)]. Here we adopted the HSL format since the image is less impacted by light in the HSL color space and is suitable for morphological logic operations.

$$r_n = \frac{R}{R+G+B} \quad , \quad g_n = \frac{G}{R+G+B} \tag{1}$$

$$H = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[\left[(R-G)^2+(R-B)(G-B)\right]^{\frac{1}{2}}\right]} \right\} \tag{2}$$

$$S = 1 - \frac{3}{(R,G,B)}\left[\min(R,G,B)\right] \tag{3}$$

$$L = \frac{R+G+B}{3} \tag{4}$$

(b) Skin area extraction

In the HSL image model with less influence of brightness, every color in the RGB color space has a set of corresponding hue, saturation, and lightness values in the HSL color space. Any region of interest (ROI) can be extracted by considering the hue and saturation palettes, regardless of the luminance palette, which includes the skin color. Therefore, the skin color, *Skin Color$_j$*, of an object can be extracted by using only hue and saturation palettes limited by the threshold values and is defined as

$$Skin\ Color_j = \begin{cases} 255, & \text{if } H_l < H_j < H_u \ \& \ S_l < S_j < S_u, \\ 0, & \text{Otherwise}, \end{cases} \tag{5}$$

where H_j and S_j are the hue value and saturation value of skin color pixel j, respectively. H_l, H_u, S_l, and S_u are the lower and upper threshold values of the skin color for hue and saturation, respectively. In experiments conducted under a luminance of approximately 360 lux, we observed that the optimal threshold ranges of hue and saturation are from 24 to 243 and from 11 to 142, respectively. These values are used throughout the paper. After color space conversion and threshold operation, the binary hue palette and saturation palette of the original image are generated. The inverted hue palette is operated with the binary saturation palette by using logic AND operations to extract the initial skin color ROI of the original image.

(c) Face location analysis

In the feature-invariant approach, the skin color distribution is the basic condition used to extract the facial area and determine whether the input image contains facial features. Since the image color can be easily affected by environmental illumination, the face color of the subject can be similarly affected. The original RGB image captured by a camera is often converted to another color space that has less illumination sensitivity, for example, normalized RGB, HSL, HIS, HSV, YIQ, YUV, or YCrCb.

After the extraction of skin color objects from the original image (denoted as *I*), there may still be some objects with similar skin color in the background, which can make it difficult to locate the face ROI accurately. We can remove small objects with similar skin color by using open operations (denoted as S_O), close operations (denoted as S_C), and morphology operations until the largest skin area (i.e., the face) is found. The open and close operations are defined as

$$S_O = I \ominus E_s \oplus D_s, \tag{6}$$

and

$$S_C = I \oplus D_s \ominus E_s, \tag{7}$$

where $\ominus$ and $\oplus$ denote erosion and dilation, respectively. E_s and D_s are the structure element matrices for erosion and dilation, which can be defined as

$$E_s = D_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \tag{8}$$

Given the experimental design, since the subject's face is about 60–70 cm away from the LCD screen, the largest skin area in the image should be the subject's face.

(d) Convex hull for face ROI

Although we can obtain the location of the face ROI (denoted as *A*) from the original image, a face ROI with many holes and disconnected particles is still not a complete solid block. Therefore, further morphological convex hull operations are required to fill the face area with the same binary value for each pixel. The convex hull operation is expressed as

$$C(A) = \bigcup_{i=1}^{4} P_k^i, \tag{9}$$

where $i = 1$–4 and P_k^i is defined as

$$P_k^i = (P_{k-1}^i \otimes S^i) \cup A. \tag{10}$$

S^i is the ith structuring element array and defined as

$$S^1 = \begin{bmatrix} 1 & \times & \times \\ 1 & c & \times \\ 1 & \times & \times \end{bmatrix}, \ S^2 = \begin{bmatrix} 1 & 1 & 1 \\ \times & c & \times \\ \times & \times & \times \end{bmatrix}, \ S^3 = \begin{bmatrix} \times & \times & 1 \\ \times & c & 1 \\ \times & \times & 1 \end{bmatrix}, \ S^4 = \begin{bmatrix} \times & \times & \times \\ \times & c & \times \\ 1 & 1 & 1 \end{bmatrix}, \quad (11)$$

where c is the corresponding point and $\times$ is an arbitrary value. $\boldsymbol{P}_k^i$ is the convergence result of the convex hull for the ith structuring element array at the kth iteration and $\boldsymbol{P}_0^i = \boldsymbol{A}$. $\otimes$ is the morphological hit-or-miss transform,[15] which can be used to look for particular patterns of foreground and background pixels in an image. An example of a convex hull operation is shown in Fig. 3.

(e) Match with original RGB image

The complete face ROI boundary is defined clearly using a convex hull operation called the binary face ROI. To accomplish a lip pattern match for lip recognition, we extract the original RGB image at the corresponding location of the binary face ROI to locate the face RGB image. The original face RGB image and the lip image pattern can be transferred into gray level images from RGB images for normalized gray cross-correlation analysis as shown in Fig. 4.
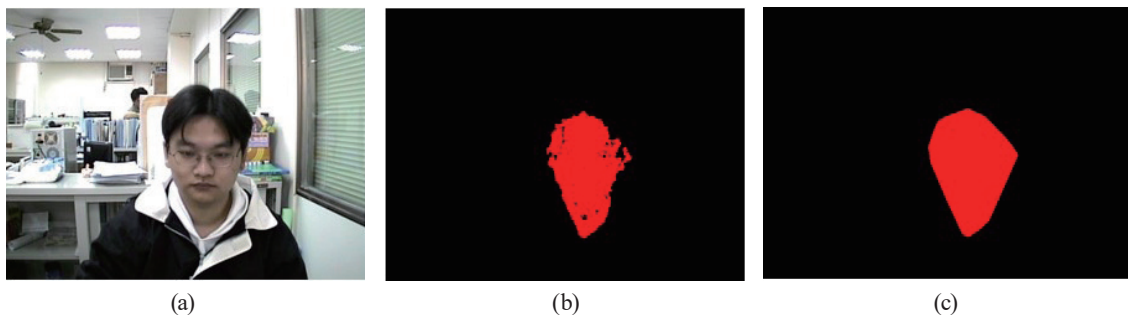

| (a) | (b) | (c) |

Fig. 3.   (Color online) Convex hull operation of face ROI: (a) original image, (b) binarization, and (c) convex hull processing.

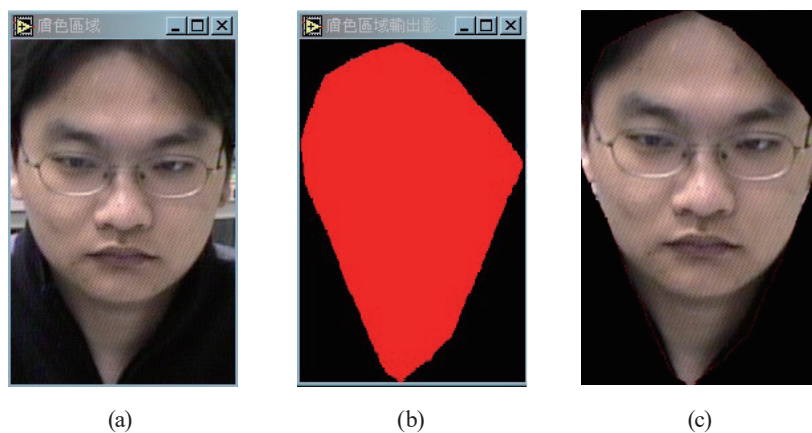| (a) | (b) | (c) |

Fig. 4.   (Color online) Face ROI block matches with the skin color area of the original RGB image: (a) original image, (b) convex hull for the skin color, and (c) superimposition of (a) and (b).

2.2 Lip image recognition and translation into Morse codes

(a) Creation of lip image pattern

As shown in Fig. 2, prior to system operation, the RGB lip image of the subject is clipped manually by a USB camera for use as the pattern for subsequent lip extraction analysis.

(b) Lip pattern matching and locating process

In general, brightness easily affects RGB images. We employ gray cross-correlation analysis to reduce the impact of brightness changes during the process of lip pattern matching. The size of the captured image always depends on the distance between the camera and the subject. To maintain consistency in the cross-correlation calculation, the normalized gray cross-correlation analysis for the RGB image is adopted. To find the correct location of the lips, the normalized gray cross-correlation algorithm is used to compare the lip image pattern with the original facial RGB image. The gray cross-correlation coefficient, R, can be defined as

$$R(i,j) = \frac{\sum_{u=0}^{L-1}\sum_{v=0}^{K-1}\left(w(u,v)-\overline{w}\right)\left(f(i+u,j+v)-\overline{f}(i,j)\right)}{\left(\sum_{u=0}^{L-1}\sum_{v=0}^{K-1}\left(w(u,v)-\overline{w}\right)^2\right)^{\frac{1}{2}}\left(\sum_{u=0}^{L-1}\sum_{v=0}^{K-1}\left(f(i+u,j+v)-\overline{f}(i,j)\right)^2\right)^{\frac{1}{2}}}, \tag{12}$$

where $w(u, v)$ is the position of one pixel of the lip pattern image. L and K are the width and height of the lip pattern image, respectively. $f(i, j)$ is the gray intensity at position (i, j). $\overline{w}$ and $\overline{f}$ are the average gray intensity of the lip pattern image and target image, respectively.

A schematic diagram of gray cross-correlation computation is shown in Fig. 5. Following the lip pattern matching procedure mentioned above, we searched for and found the most similar lip area with the maximum gray cross-correlation value in the original face image, i.e., block (a) in Fig. 5.

(c) Lip profile calculation

As a result of the lip pattern matching procedure,[16] we can locate the position of the lips. Next, we start the recognition procedure of the opened/closed status of the lips, as shown in Fig.
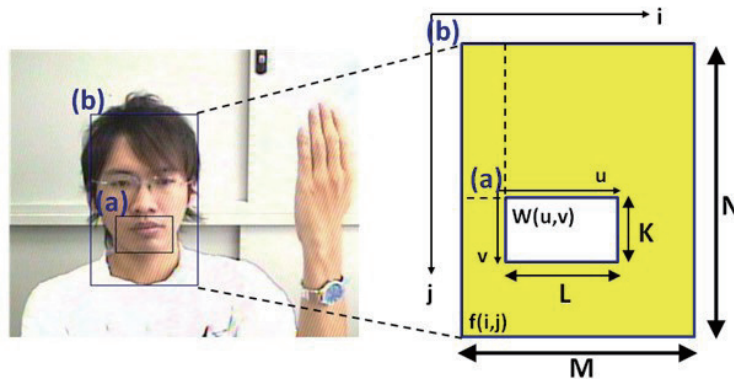

Fig. 5.   (Color online) Gray cross-correlation computation for locating the position of the lips.

6. By observing the pixels of lips, we find that the color of lips ranges from dark red to purple under normal light conditions (360 lux). We need to extract the lips from facial skin of any color to account for people from different races.[17] According to this observation for different faces, the color distribution of the lips and the surrounding skin should be distinguishable. It shows that the lip colors are distributed in the lower part of the crescent area defined by the skin colors on the r-g plane. We also define another quadratic polynomial discriminant function of lip pixels for the more rapid extraction of lip colors.

From the distributions of lip colors, we find the quadratic polynomial of the lower boundary $f_{lower}(r)$. However, the color of the dark area between the upper lip and lower lip when the lips are opened should be darker than the lower limitation of the lip color as shown in Fig. 7. Therefore, we adopt the lower limitation of lip colors $f_{lower}(r)$ as the upper limitation of the color of the area between the upper and lower lips as shown in Eq. (13).

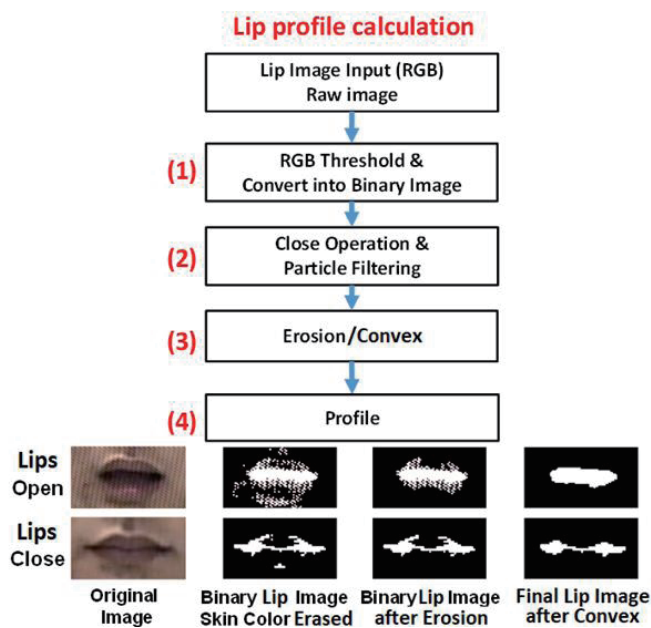$$f_{lower}(r) = -0.776r^2 + 0.5601r + 0.1766 \tag{13}$$

Fig. 6.   (Color online) Recognition procedure of lip status.

(a)   (b)   (c)
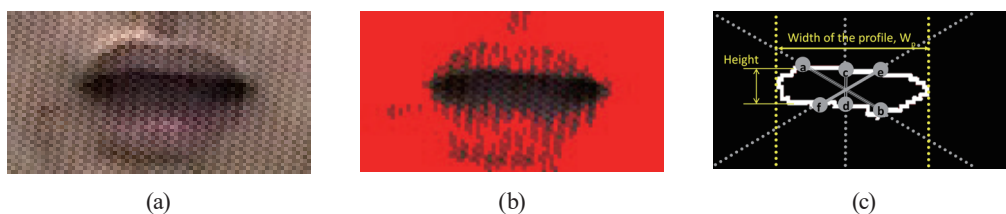
Fig. 7.   (Color online) Comparison of the color discrimination among the upper lip, the lower lip, and the area between the upper and lower lips: (a) original RGB lip image, (b) binary lip image, and (c) height and width of the binary lip block profile.

The key point in distinguishing the opened/closed status of lips is to find the dark pixels in the dark area between the upper and lower lips when the subject opens his/her mouth. Thus, the color of the area between the upper and lower lips is applied to detect the pixels inside the dark area between the upper and the lower lips. The detection rule (L_c) is

$$L_c = \begin{cases} 0, \text{if } 0 \le g \le f_{lower}, R \le 20, G \le 20, \text{and } B \le 20, \\ 1, \text{otherwise,} \end{cases} \tag{14}$$

where R, G, and B are the intensity values in the red, green, and blue channels, respectively. $L_c = 0$ and $L_c = 1$ mean that the pixel is inside and outside the mouth, respectively.

There are many particles in the detected dark area between the upper and lower lips. Morphological operations including erosion, dilation, open, close, and convex hull operations are applied to delete the small particles until the largest particle or block is left. Finally, the profile of the largest block is used to identify the opened/closed status of lips.

As mentioned above, the profile of the opened/closed lip block is extracted. We utilize the boundary of this profile to determine whether the lips are opened or closed. First, we define the threshold value of the vertical height in the center of the binary lip block image in Fig. 7(c) according to the subject's condition. Next, the opened/closed status of the lips is recognized on the basis of the threshold value of the binary lip block image. Finally, the opened/closed status of the lips is transferred into the tone and silence state of Morse code. The input Morse code, L_s, is defined as

$$L_s = \begin{cases} 0, & \text{when mouth is opened,} \\ 1, & \text{when mouth is closed.} \end{cases} \tag{15}$$

The subject controls the duration of opening or closing the lips to input a dot tone or dash tone (similarly, dot silence or dash silence) of Morse code.

2.3 Morse code fuzzy recognition algorithm

A Morse code sequence includes tones and silences. A tone is defined as the pressing time of a switch and a silence is defined as the releasing time of the switch. Tones can be divided into long tone (dash, "-") and short tone (dot, "*") elements according to the duration of the processing time of the status switch, as is the case for silences. Different long/short tone Morse code combinations can represent different characters,[18] and part of the characters used in this study are shown in Table 1. The ratio of the long to short elements (tones or silences) by definition is always 3:1. However, the Morse code input rate may not be consistent even for users familiar with inputting Morse code. Thus, it is very difficult for disabled people to maintain a stable 3:1 ratio for long and short elements. When the long-to-short ratio becomes highly irregular, adaptive recognition fails frequently, especially for disabled persons. Therefore, the adjustable FCIMCI system with a fuzzy recognition algorithm can solve this problem.

Table 1
Table of Morse codes corresponding to characters "A"–"Z".

character	Morse code	character	Morse code	character	Morse code
A	*‒	K	‒*‒	U	**‒
B	‒***	L	*‒**	V	***‒
C	‒*‒*	M	‒‒	W	*‒‒
D	‒**	N	‒*	X	‒**‒
E	*	O	‒‒‒	Y	‒*‒‒
F	**‒*	P	*‒‒*	Z	‒‒**
G	‒‒*	Q	‒‒*‒		
H	****	R	*‒*		
I	**	S	***		
J	*‒‒‒	T	‒	* dot ‒ dash	

To improve the recognition rate, a long-short separation fuzzy recognition algorithm was developed to trace the variation of a Morse code sequence. This algorithm, used for tracing the variation of long (l_k) and short (s_k) elements by employing two predicting loops to recognize long and short elements separately, is shown in Fig. 8.

The recognition procedure is described as follows:

Step 1. Obtain a long (l_k) or short (s_k) element from an input I_k using the function f_T in Eq. (16).

$$f_T \begin{cases} s_k = I_k, & \text{if } I_k < T_k \left(\text{i.e., } I_k \text{ go to } s_k \text{ loop} \right) \\ l_k = I_k, & \text{if } I_k > T_k \left(\text{i.e., } I_k \text{ go to } l_k \text{ loop} \right) \end{cases} \tag{16}$$

Step 2. Calculate the prediction error e_k as

$$e_k = \begin{cases} l_k - ly_{k-1}, & \text{in } l_k \text{ loop,} \\ s_k - sy_{k-1}, & \text{in } s_k \text{ loop.} \end{cases} \tag{17}$$
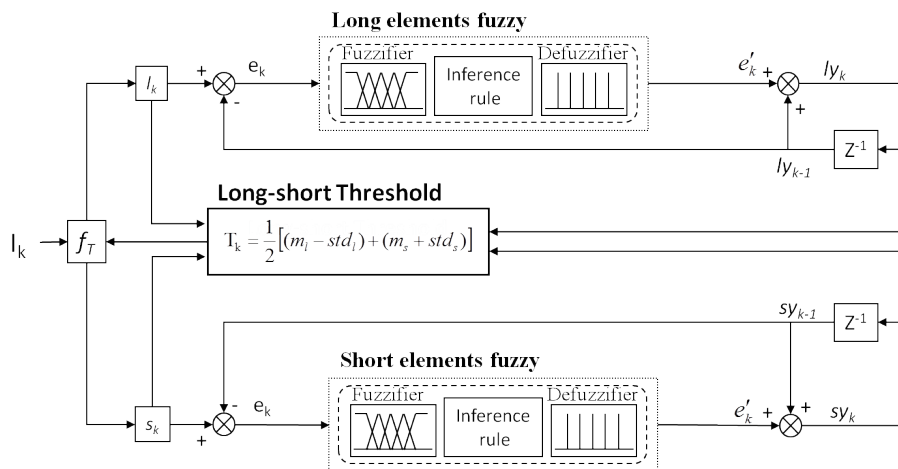

Fig. 8.   Fuzzy recognition algorithm for long-short Morse code separation.

In the fuzzy algorithm, a linguistic fuzzy rule is utilized to calculate the modified error e'_k. Five linguistic rules are given as follows:

Fuzzy rule 1: if e_k is LN, then e'_k is LN (highest speed)

Fuzzy rule 2: If e_k is SN, then e'_k is SN (high speed)

Fuzzy rule 3: If e_k is ZE, then e'_k is ZE (normal speed)

Fuzzy rule 4: If e_k is SP, then e'_k is SP (low speed)

Fuzzy rule 5: If e_k is LP, then e'_k is LP (lowest speed)

LN, SN, ZE, SP, and LP are negative large, negative small, zero, positive small, and positive large, respectively.

Step 3. Update the predicted output value at each loop by Eqs. (18) and (19).

$$ly_k = ly_{k-1} + e'_k, \text{ in } l_k \text{ loop} \tag{18}$$

$$sy_k = sy_{k-1} + e'_k, \text{ in } s_k \text{ loop} \tag{19}$$

Step 4. Update the threshold value, the mean values (m_l, m_s), and the standard deviations (std_l, std_s) in each loop from inputs I_{k-3}, I_{k-2}, I_{k-1}, I_k and the predicted output ly_k or sy_k. For example, m_1 and std_l in the l_k loop are given as follows:

$$m_l = \frac{1}{5}\left(\sum_{i=0}^{3} l_{k-i} + ly_k\right), \tag{20}$$

$$std_l = \frac{\sqrt{\sum_{i=0}^{3}\left(l_{k-i} - m_l\right)^2 + \left(ly_k - m_l\right)^2}}{2}. \tag{21}$$

In the same derivation for a short element, m_s and std_s can be obtained by replacing l_{k-i} and ly_k in Eqs. (20) and (21) with s_{k-i} and sy_k, respectively. Threshold T_k is obtained by calculating the middle value between the lowest border of long elements ($m_1 - std_l$) and the highest border of short elements ($m_s + std_s$), i.e.,

$$T_k = \frac{1}{2}\left[(m_l - std_l) + (m_s - std_s)\right]. \tag{22}$$

By repeating steps 1–4, the system adjusts the threshold value adaptively in response to the typing speed variation and the varying ratio of long-element duration to short-element duration.

Once both the opened/closed lip status recognition and the fuzzy Morse code recognition are completed, the text/character can be inputted in accordance with a Morse code table (Table 1). An example of text input by a subject is shown in Fig. 9. In the Windows operating system, LabVIEW software is used to call up keyboard and mouse application programming interface (API) functions, which assist applications in achieving the actions of opening windows, depicting graphics, and using peripheral devices, to achieve text processing ability.
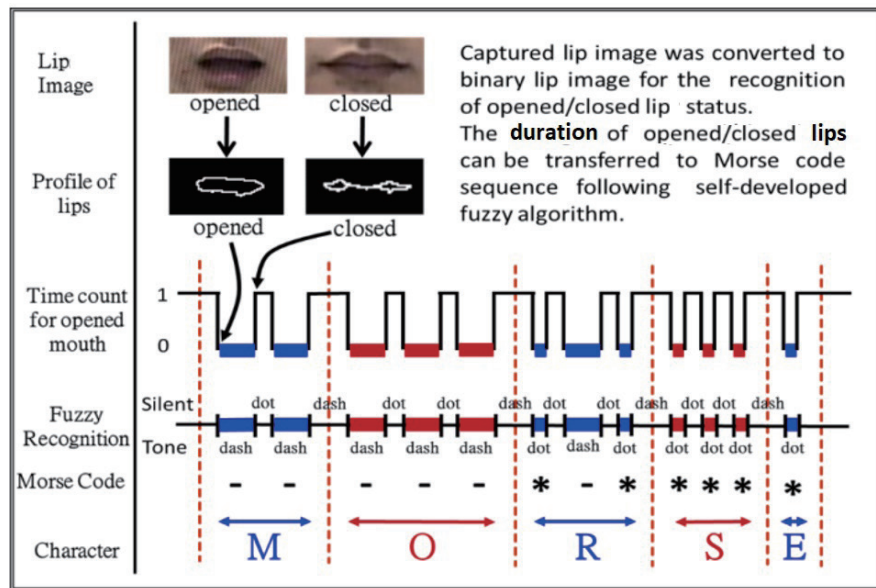
Fig. 9.   (Color online) Morse code generation depending on the opened/closed lip status of the subject.

3.   Results

The performance of the system is demonstrated successfully from several aspects as described in the following sections.

3.1   Face and lip tracking

As shown in Fig. 10, the face tracking results successfully verified the effectiveness of the face recognition algorithm, and the recognition results of the opened/closed status of the mouth with/without hand interference in the background were accurately identified by the lip recognition algorithm regardless of whether it was daytime or nighttime. The image processing strengthens color characteristics in the image and reduces the impact of light. The face tracking algorithm not only tracks but also extracts human faces effectively under variation in the light and environment. The overall average accuracy rate was 97.87% in the lab (360 lux) and 94.44% in a home environment (280 lux). Hence, the image recognition algorithm is not sensitive to light interference in different environments.

3.2   Accuracy analysis of lip recognition

After the face recognition and tracking process, the algorithm starts the process of lip image recognition. The steps of the opened/closed lip image recognition procedure are labeled as (A) to (E) on the left of Fig. 11, and the corresponding lip image recognition results are shown on the right of Fig. 11. Twenty healthy subjects were subjected to 50 time cycles of opened/closed lip image testing. The average recognition accuracy and standard deviation of the test results for the 20 subjects reached up to $98.3 \pm 2.27\%$.
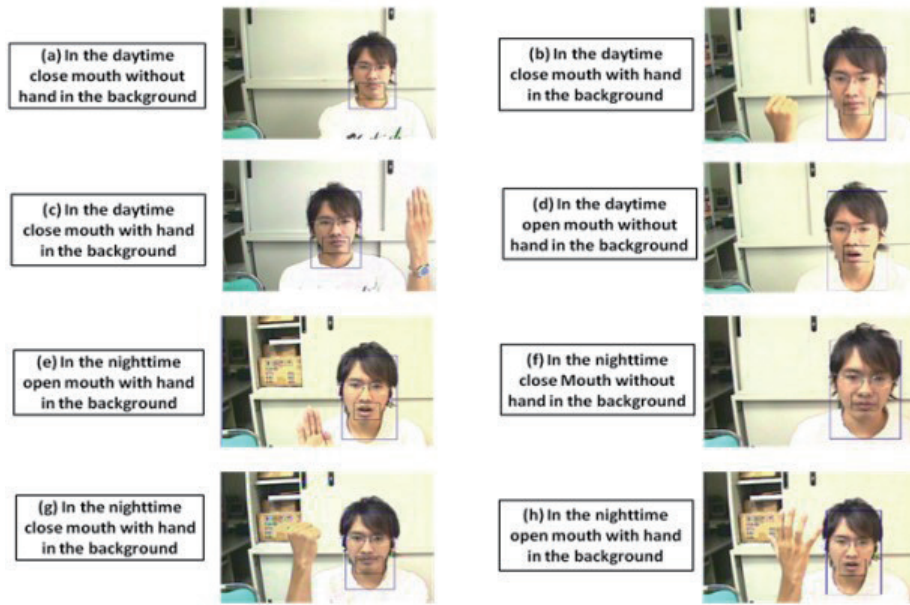
Fig. 10. (Color online) Results for tracking the face and lips of the subject and recognizing the opened/closed status of lips (mouth) with/without a hand in the background.
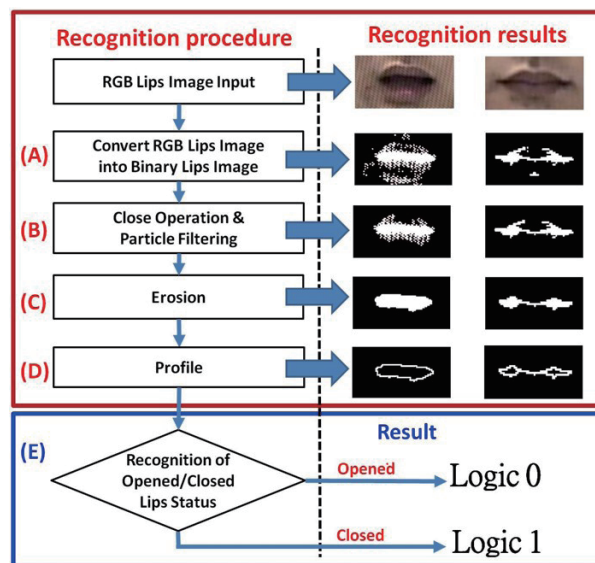

Fig. 11. (Color online) Procedure of lip image recognition and the corresponding logic output for opened/closed lip status.

3.3 FCIMCI system performance test

To identify the FCIMCI system performance, we completed experiments with the FCIMCI system operated by three types of subject: an expert who is familiar with Morse codes, a person with SCI, and a person with CP. These three subjects were asked to enter Morse codes from "A" to "Z" ten times. Figure 12 shows the long and short elements of the Morse code tone sequence ("A" to "Z") inputted by the subjects. The best test result is in Fig. 12(a), in which the threshold
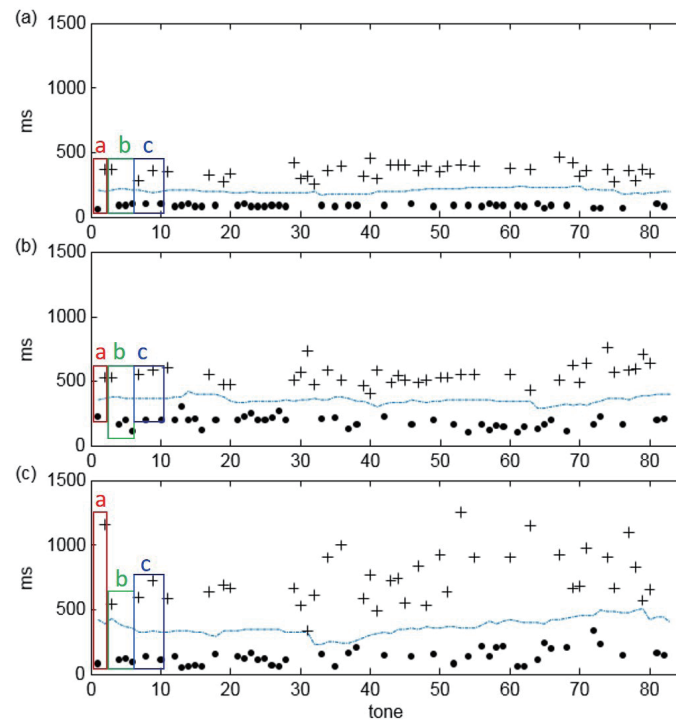
Fig. 12. (Color online) Fuzzy recognition results of tone Morse code for inputting "A" to "Z" by (a) expert, (b) person with SCI, and (c) person with CP. The blue dashed line (−·) represents the fuzzy-controlled threshold values for distinguishing the dot (●) and dash (+) in Morse code.

value line is very smooth, showing high stability and a high input speed, and the worst test result is in Fig. 12(c), in which the threshold value line fluctuates considerably, indicating low stability, and the input speed is lowest. Although Fig. 12 shows different stabilities and input speeds for different subjects, the FCIMCI system still achieves 100% recognition accuracy in some system performance experiments.

The system performance was evaluated on the basis of the data typed by the different subjects. Thirty datasets were given to the subjects, with ten datasets typed by an expert, ten datasets typed by a user with SCI, and the other ten datasets typed by a teenager with CP. Figure 13 shows the recognition results for the data typed by the expert, the data typed by the person with SCI, and the data typed by the person with CP. As shown in Fig. 13, the average recognition rates for the data typed by the expert, the person with SCI, and the person with CP are 99.67, 98.35, and 99.15%, respectively. In a scenario with an unstable typing pattern, the fuzzy recognition algorithm for image Morse code maintains high recognition accuracy. These results demonstrate that the proposed fuzzy recognition algorithm is suitable for recognizing unstable Morse code patterns.

3.4   Case study

To further evaluate the overall FCIMCI system performance, we collaborated with a 28-year-old subject with SCI who had Morse code training experience and conducted an experiment over
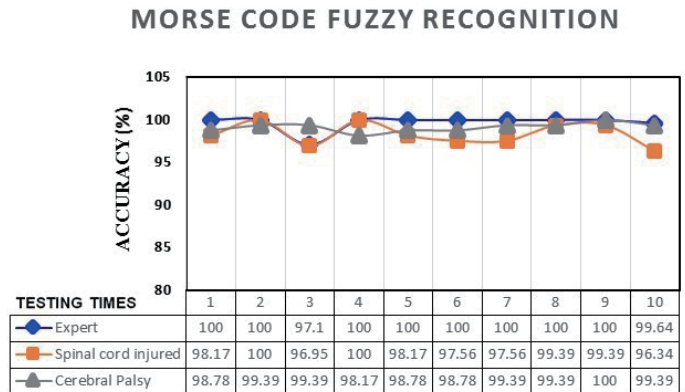
MORSE CODE FUZZY RECOGNITION

TESTING TIMES	1	2	3	4	5	6	7	8	9	10
Expert	100	100	97.1	100	100	100	100	100	100	99.64
Spinal cord injured	98.17	100	96.95	100	98.17	97.56	97.56	99.39	99.39	96.34
Cerebral Palsy	98.78	99.39	99.39	98.17	98.78	98.78	99.39	99.39	100	99.39

Fig. 13.   (Color online) Recognition results of FCIMCI system operated by three subjects: an expert, a person with SCI, and a person with CP.
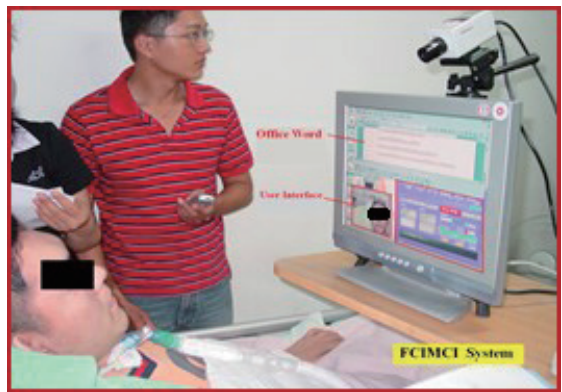

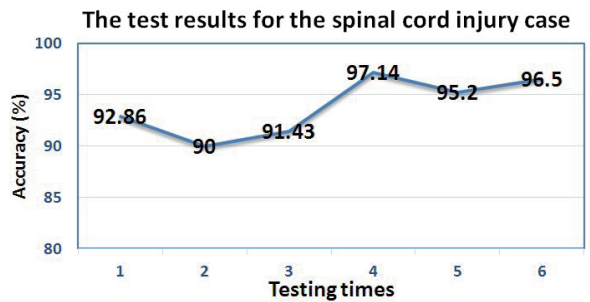Fig. 14.   (Color online) FCIMCI system operated by a subject with SCI.

Fig. 15.   (Color online) Accuracy of FCIMCI system performance test for subject with SCI.

3 months, during which we trained and tested the subject every other week. The SCI subject was required to type the letters "A" to "Z" ten times in each test to assess the performance of the FCIMCI system as shown in Fig. 14. We conducted six tests in one complete experiment. The accuracies for the six tests were between 90 and 97.14% as shown in Fig. 15. As shown by the positive trend of the curve, the system accuracy increased with additional training of the subject.

4.   Discussion and Conclusions

To help severely disabled people input text by mouth action, we designed an FCIMCI system based on contactless mouth image recognition. From the experimental results, the overall performance of the FCIMCI system was demonstrated to be successful in executing digital image processing techniques and in achieving facial tracking, lip image location, extraction, processing, and recognition. The FCIMCI system successfully combined opened/closed lip image recognition techniques with fuzzy control Morse code techniques to provide a communication service for severely disabled people as a convenient communication assistive device in daily life. The image recognition algorithm can accurately detect the human face and lip status with average accuracy rates of 94.44% in a home environment and 97.87% in a lab and is not sensitive to light interference. An image Morse code based on the lip image recognition algorithm can be correctly inputted into the FCIMCI system in real time. High accuracy rates were achieved for fuzzy Morse code recognition for experts, subjects with SCI, and subjects with CP, with average recognition rates of 99.67% for data typed by an expert, 98.35% for data typed by a person with SCI, and 99.15% for data typed by a person with CP. The FCIMCI system was shown to have accuracies between 90 and 97.14%, with an average accuracy of about 93.85% in six tests conducted on a subject with SCI. The recognition speed of the FCIMCI system reached 119.89 ± 39.32 ms/per recognition cycle under the employed image resolution (480×360), providing strong evidence for the real-time function of the FCIMCI system. Moreover, the FCIMCI system is inexpensive and can be easily implemented as a communication assistive device for severely disabled people. It comprises simple hardware (only one PC and one camera) without requiring any appurtenance attached to the subject, eliminating subject discomfort.

In conclusion, we accomplished a practical communication assistive device for severely disabled people by integrating image and Morse code fuzzy recognition algorithms. Despite the advantages of the FCIMCI system, such as cost-effectiveness, simple hardware, real-time response, and high accuracy, further improvements are possible in the future. Since most functions of the system are implemented in software, for example, the image recognition and Morse code fuzzy recognition algorithms, the kernel of the FCIMCI system can be upgraded by implanting improved new software algorithms and hardware (such as an embedded system with firmware and a monitor). Such improvements will result in a better quality of life for people with severe disabilities.

References

1. A. Caranica, H. Cucu, C. Burileanu, F. Portet, and M. Vacher: 2017 Int. Conf. Speech Technology and Human-Computer Dialogue. (SpeD, 2017) 1–8.
2. F. Kong, M. N. Sahadat, M. Ghovanloo, and G. D. Durgin: IEEE Trans. Biomed. Circuits Syst. **13** (2019) 848. https://doi.org/10.1109/TBCAS.2019.2926755
3. F. Koochaki and L. Najafizadeh: IEEE Trans. Neural Syst. Rehabil. Eng. **29** (2021) 974. https://doi.org/10.1109/TNSRE.2021.3083815
4. E. F. LoPresti, D. M. Brienza, and J. Angelo: Interact. Comput. **14** (2002) 359. https://doi.org/10.1016/S0953-5438(01)00058-3
5. D. G. Evans, R. Drew, and P. Blenkhorn: IEEE Trans. Rehabil. Eng. **8** (2000) 107. https://doi.org/10.1109/86.830955
6. R. Palaniappan, R. Paramesran, S. Nishida, and N. Saiwaki: IEEE Trans. Neural Syst. Rehabil. Eng. **10** (2002) 140. https://doi.org/10.1109/TNSRE.2002.802854
7. A. Çakir: Behav. Inf. Technol. **32** (2013) 625. https://doi.org/10.1080/0144929X.2013.796625
8. M. C. Hsieh, C. H. Luo, and C. W. Mao: IEEE Trans. Rehabil. Eng. **8** (2000) 405. https://doi.org/10.1109/86.867882
9. C. H. Shih and C. H. Luo: Int. J. Med. Inf. **44** (1997) 193. https://doi.org/10.1016/S1386-5056(97)00020-8
10. D. T. Fuh and C. H. Luo: J. Med. Eng. Technol. **25** (2001) 118. https://doi.org/10.1080/03091900110052441
11. C. M. Wu and C. H. Luo: J. Med. Eng. Technol. **26** (2002) 202. https://doi.org/10.1080/03091900210156904
12. C. H. Yang, L. Y. Chuang, C. H. Yang, and C. H. Luo: IEICE Trans. Fundamentals Electron. Commun. Comput. Sci. E89-A(7) (IEICE, 2002) 1995.
13. C. H. Yang: IEICE Trans. Fundamentals Electron. Commun. Comput. Sci. E84-A(1) (IEICE, 2001) 356.
14. O. Ikeda: Proc. 2003 Int. Conf. Image Processing. (Cat. No. 03CH37429, 2003) III 913.
15. O. M. Elrajubi, I. El-Feghi, and M. A. B. Saghayer: Int. J. of Comput. Inf. Eng. **8** (2014) 1. https://doi.org/10.5281/zenodo.1093325
16. S. Bakshi, R. Raman, and P. K. Sa: 2011 Annual IEEE India Conf. (IEEE, 2011) 1–4.
17. C. C. Chiang, W. K. Tai, M. T. Yang, Y. T. Huang, and C. J. Huang: Real Time Imaging **9** (2003) 227. https://doi.org/10.1016/j.rti.2003.08.003
18. C. M. Wu, C. Y. Chuang, M. C. Hsieh, and S. H. Chang: Biomed. Eng.: Appl., Basis Commun. **25** (2013) 1350006. https://doi.org/10.4015/S1016237213500063

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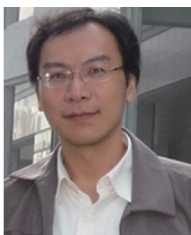
Chung-Min Wu received his B.S. degree in automatic control engineering from Feng Chia University, Taichung, Taiwan, his M.S. degree in biomedical engineering from National Cheng Kung University, Tainan, Taiwan, and his Ph.D. degree in electrical engineering from National Cheng Kung University in 1994,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

Cheng-Fa Yen received his B.S. degree from the Department of Physics, National Central University, Taoyuan, Taiwan, and his M.S. degree from the Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan, in 2000 and 2010, respectively, and is currently pursuing his Ph.D. degree in electrical engineering at National Cheng Kung University, Tainan, Taiwan. He is currently involved in research on biomedical integrated systems. He is also working on non-destructive testing and semiconductor device reliability simulation by numerical analysis. His research interests include emerging nano-biomedical technologies. (fredyentw@hotmail.com)

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)

Cheng-Chi Tai was born in Tainan, Taiwan, R.O.C., on Nov. 10, 1962. He received his B.S. degree in electronic engineering from Chung Yuan Christian University, Chung Li, Taiwan, in 1986, his M.S. degree in electrical engineering from National Cheng Kung University (NCKU), Tainan, Taiwan, in 1988, and his Ph.D. degree in electrical engineering from Iowa State University, Ames, Iowa, USA, in 1997. He is now a professor with the Department of Electrical Engineering, NCKU. His research interests include bio-electronic instrumentation, electromagnetic thermotherapy, medical signal and image processing, and partial discharge nondestructive evaluation using eddy current, ultrasound, and acoustic emission techniques.
 (ctai@mail.ncku.edu.tw)