

# Recognition of Motor Imagery Electroencephalography Using Independent Component Analysis and Machine Classifiers

Chih-I Hung<sup>1,4</sup>, Po-Lei Lee<sup>4</sup>, Yu-Te Wu<sup>1,4,\*</sup>, Hui-Yun Chen<sup>1,4</sup>, Li-Fen Chen<sup>3,4</sup>  
Tzu-Chen Yeh<sup>2,3,4</sup>, Jen-Chuen Hsieh<sup>2,3,4</sup>

<sup>1</sup>Institute of Radiological Sciences, <sup>2</sup>Institute of Neuroscience, <sup>3</sup>Center for Neuroscience  
National Yang-Ming University  
No.155, Sec. 2, Linong St., Beitou District,  
112, Taipei, Taiwan,

<sup>4</sup>Integrated Brain Research Laboratory, Dept. of Medical Research and Education,  
Taipei Veterans General Hospital,  
No.201, Sec. 2, Shihpai Rd., Beitou District  
112, Taipei, Taiwan

e-mail : [runtothewater@pie.com.tw](mailto:runtothewater@pie.com.tw); [pllee2@vghtpe.gov.tw](mailto:pllee2@vghtpe.gov.tw); [ytwu@ym.edu.tw](mailto:ytwu@ym.edu.tw);  
[airrb@pchome.com.tw](mailto:airrb@pchome.com.tw); [lfchen3@vghtpe.gov.tw](mailto:lfchen3@vghtpe.gov.tw); [tcyeh@vghtpe.gov.tw](mailto:tcyeh@vghtpe.gov.tw); [jchsieh@vghtpe.gov.tw](mailto:jchsieh@vghtpe.gov.tw)

## ABSTRACT

Motor imagery electroencephalography (EEG), which embodies cortical potentials during mental simulation of left or right finger lifting tasks, can be used as neural input signals to activate brain computer interface (BCI). The effectiveness of such an EEG-based BCI system relies on two indispensable features: distinguishable patterns of brain signals and accurate classifiers. This work aims to extract a reliable neural feature, termed as beta rebound map, out of motor imagery EEG by means of independent component analysis, and employ four classifiers to investigate the efficacy of beta rebound map. Results demonstrated that, with the use of ICA, the recognition rates of four classifiers, linear discriminant analysis (LDA), back-propagation neural network (BP-NN), radial-basis function neural network (RBF-NN), and support vector machine (SVM) improved significantly from 54%, 54%, 57.3% and 55% to 69.8%, 75.5%, 76.5% and 77.3%, respectively. In addition, the areas under the ROC curve, which assess the quality of classification over a wide range of misclassification costs, also improved greatly from .65, .60, .62, and .64 to .78, .73, .77 and .75, respectively.

## Keywords

Electroencephalography (EEG), Independent component analysis (ICA), brain computer interface (BCI), beta rebound, linear discriminant analysis (LDA), back-propagation neural network (BP-NN), radial-basis function neural network (RBF-NN), support vector machine (SVM)

## 1. INTRODUCTION

In recent years, great progress in neuroscience has inspired studies in developing brain computer interface (BCI) [Mul99a] [Pfu98a] [Pfu00a] [Pol98a], a novel technique in assisting people to communicate with external environments or trigger surrounding devices by means of their brain signals. These

systems are particularly useful for ones who suffer from amyotrophic lateral sclerosis or locked-in syndrome and are unable to produce any motor activity. Their cognition or sensor functions, however, may be intact so that they can be trained to perform mental tasks, for example, in simulating right or left hand or foot movements without any overt motor output. The success of BCI systems relies on two integral parts: distinguishable neural patterns and effective classifiers. This work aims to extract a reliably distinguishable feature from the motor imagery EEG recording by means of independent component analysis and employ machine classifiers to investigate the efficacy of extracted pattern.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.  
Copyright UNION Agency – Science Press

It has been pointed out that imagination of hand movement elicits rhythmic EEG patterns in the primary sensorimotor areas similar to that from a real hand movement [Pfu96a]. When a specific movement or imagined movement is performed, it composes of three phases: planning, execution and recovery. The planning and execution results in localized alpha and lower beta bands amplitude attenuation or event-related desynchronization (ERD) which can be viewed as an EEG correlate of an activated cortical motor network, while the recovery phase produces focal mu and beta amplitude enhancement or event-related synchronization (ERS) which may reflect deactivation/inhibition in the underlying cortical network.

Several BCI systems have been proposed based on the induced ERD when subjects performed imagery hand or foot movements [Pfu98a] [Pfu00a]. Pfurtscheller *et al.* used a learning vector quantization to classified ERD signals on-line in a subject specific band which was determined by distinctive sensitive learning vector quantization. They also adopted adaptive autoregressive model to analyzed ERD signal off-line and applied linear discrimination analysis to improve the detection of imagined left and right hand movements. The reported error rates varied 5.8 and 32.8%. Muller-Gerking *et al.* applied common spatial filter to detect real (not imagined) left, right hand or right foot movements in single trial and reported 84%, 90% and 94% accuracies for three subjects, respectively [Mul99a].

Although the ERD elicited by imagined movement has been extensively used as a feature pattern in BCI systems, we have observed that not every subject can produce discernible ERD during the imagery movement, whereas the beta ERS was persistently appeared for each subject. This motivated us to adopt ERS, rather than the ERD, as the feature pattern. The peaked ERS of imaged left or right hand movement, referred to as beta rebound, exhibits on bilateral sensorimotor areas but with distinct patterns. When the imagination of right hand movement is executed, the beta rebound over left hemisphere produces stronger amplitude than that on the right hemisphere, and the vice versa.

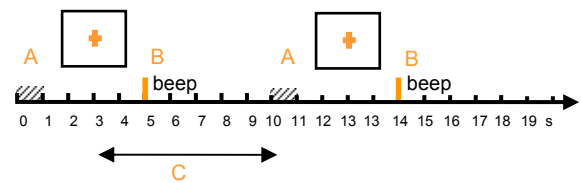
The recorded EEG signals were inevitably contaminated by system noise, artifacts, spontaneous EEG, etc. Following our previous works for MEG/EEG de-noise [Lee03a], we employed the Independent Component Analysis (ICA) technique to decompose each pre-processed epoch into a set of temporally independent components along with corresponding spatial maps, and selected the task-related components by matching designed spatial

templates with the decomposed spatial maps. As a result, the signal-to-noise ratio of each EEG single-trial was improved, which lead to the promotion of classifiers' performance.

This paper is organized as follows. Section 2 reports our experimental paradigm for motor imagery task and EEG recording configuration. Section 3 presents the extracted features, with and without applying ICA, based on peaked beat ERS and termed as beta rebound maps. Section 4 reviews four classifiers in this study. Section 5 summarizes the classification results and Section 6 concludes this study.

## 2. EXPERIMENTAL PARADIGM FOR MOTOR IMAGERY

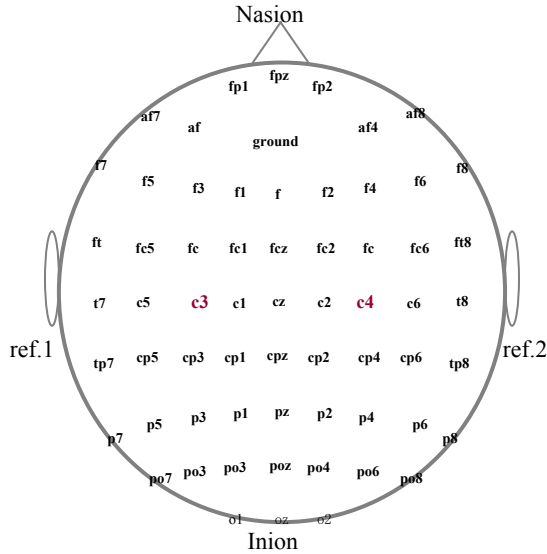
Four right-handed healthy subjects (two males and two females), aged between 20 and 28, participated in this study. Each subject was naive to the experiment and trained only twenty minutes prior to the first session. During each session, the subject was asked to perform 100 trials of imagery right index finger lifting, followed by another 100 trials of imagery left index finger lifting. The length of each trial was ten seconds. Each trial began with one-second presentation of random noise during which subjects were allowed to blink his/her eyes (A in Fig. 1). The subject was then instructed to stare at the fixation cross in the center of the monitor from 2s and started to image right or left index finger lifting right after he/she heard an acoustic cue "beep" (with frequency 1k Hz and 10ms duration) at 5s (B in Fig. 1.). The inter-stimulus interval was 10 second.



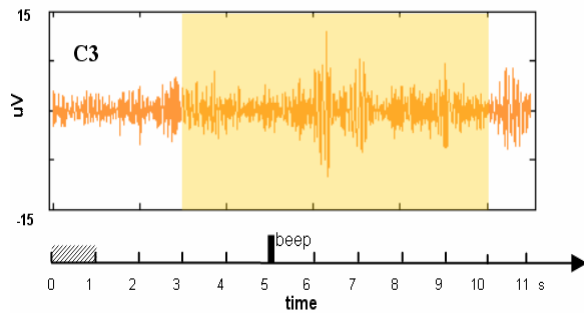
**Figure 1.** Timing of two consecutive trials of the motor imagery task.

A 64-channels electroencephalography (EEG) 10-20 system (with an electro-cap) was used to record the cortical potentials. The configuration of standard 1-20 system is shown in Fig. 2. The vertical and horizontal electro-oculograms (VEOG and HEOG) were applied to reject bad epochs induced by eye blinking during the recording. The data were digitized at 250 Hz. Since we focused on beta-activities, the signals were further bandpass-filtered with 6-50 Hz to remove the dc drifts and 60 Hz noise.

Throughout the recordings, the surface electromyogram (EMG) was monitored from the m. extensor digitorum communis (digitized at 2 KHz) for the detection of motion status. Data of four sessions were collected for each subject. Signals from 3s to 10s (C in Fig. 1.) in each trial (excluding bad epochs) were extracted for further classifiers training and testing. Figure 3 exhibits such a pre-processed epoch from sensorimotor area (channel C3 in 10-20 system).



**Figure 2.** The configure of standard 10-20 system with 64 channels.



**Figure 3.** A pre-processed epoch recorded at C3.

### 3. FEATURE EXTRACTION WITH AND WITHOUT ICA

Extraction of reliable feature from measured data is vital in facilitating the subsequent classification procedure. Since the measured signals were inevitably contaminated by system noise, artifacts, spontaneous EEG, etc., we employed the ICA

technique to decompose each pre-processed epoch into a set of temporally independent components along with corresponding spatial maps, and selected the task-related components by matching designed spatial templates with the decomposed spatial maps. Two types of feature, one using ICA to extract task-related components and the other without using ICA, were created from pre-processed data for the purpose of comparison with their efficacies. The detailed steps for feature extraction with ICA were described in the following:

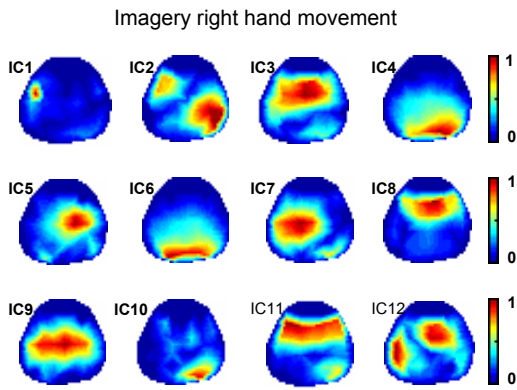
**Step 1: Signal decomposition by using ICA.** We first arranged each pre-processed epoch across  $m$  channels ( $m=62$ ) and  $n$  sampled points ( $n=1750$ ) into an  $m \times n$  matrix  $X$ . The  $i_{th}$  row contains the observed signal from  $i_{th}$  EEG channel, and the  $j_{th}$  column vector contains the observed samples at the  $j_{th}$  time point across all channels. In the present study, all calculations were performed using the FastICA algorithm [Cov65a] [Cov98a]. The FastICA technique first removed means of the row vectors in the  $X$  matrix followed by a whitening procedure to transform the covariance matrix of the zero-mean data into an identity matrix. The whitening process was implemented using the Principal Component Analysis. Only the first  $N$  most significant eigenvectors ( $N=15$  in our analysis) were preserved in the subsequent ICA calculation. In the next step, FastICA searched a matrix to further separate the whitened data into a set of components which were as mutually independent as possible. Combining with previous whitening process, the matrix  $X$  can be transformed into a matrix  $S$  via an un-mixing matrix  $W$ , i.e.,

$$S = WX \quad (1)$$

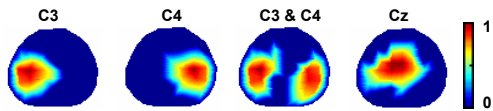
in which the rows of  $S$  were mutually independent. Each column of  $W^{-1}$ , i.e. mixing matrix, represents a spatial map describing the relative projection weights of the corresponding temporal components at each of the EEG channels. They will be referred to as IC spatial maps henceforth. Figure 4 shows 12 IC spatial maps of 12 independent components (not shown) decomposed from a single-trial imagery right hand movement. The maps IC3, IC5, IC7 and IC9 were highly related to motor imagery task and categorized as task-related components, while the IC4 and IC6 maps were associated with the occipital alpha rhythm, and IC1 map was the noise emanated from a bad channel.

**Step 2: Correlating the IC spatial maps with pre-defined spatial templates to select task-related components.** Since the motor imagery task elicits

bilateral activation in the vicinity of sensorimotor areas, four spatial patterns encompassing C3, C4, Cz and both C3 and C4 areas, respectively, were considered as spatial templates (see Fig. 5) in selecting the task-related spatial maps. Please note that four spatial templates rather than single template covering C3 and C4 were taken into account because the task-related activities can be separated by ICA and exhibited in multiple IC spatial maps. Each template was correlated with 12 IC spatial maps of single trial and the best two matches were selected. For example, the spatial maps IC3, IC5, IC7 and IC9 in Figure 4 were selected automatically due to their high similarity. The task-related IC spatial maps as well as the corresponding temporal components were used to reconstruct the signal  $X$  by means of equation (1).



**Figure 4.** The normalized IC spatial maps of a single-trial imagery right hand movement.



**Figure 5.** Spatial templates used to select task-related IC spatial maps.

**Step 3: Computing the envelopes of beta reactivity from reconstructed signals using the Amplitude Modulation method.** The optimal beta frequency band encompassing the prominent and relevant brain activities may vary across subjects and subjects. To tackle this problem, we divided the beta band, into five sub-frequency bands, 8~12, 12~16, 16~20, 20~24, and 24~28 Hz, and used them with additional beta band 8~30 Hz to band-pass filter the reconstructed signals. The Amplitude Modulation (AM) method based on the Hilbert transform was applied to detect the envelope of the filtered EEG signals and quantify the event-related oscillatory

activities [Clo96a]. Each envelope, referred to as AM waveform, was computed by (see Figure 6 (a))

$$m(t) = \sqrt{M_{BP}(t)^2 + H(M_{BP}(t))^2} \quad (2)$$

where  $M_{BP}(t)$  is the single-trial band-passed EEG signal, and  $H(M_{BP}(t))$  is its Hilbert transform. Contrary to the classical measurement of ERS reactivity and the original AM approach in which a relative percentage as indexed to the initial baseline was used [Clo96a], we computed the beta ERS reactivity (termed as beta rebound) using the amplitude difference between the maximum values of beta ERD and beta ERS of the AM envelope.

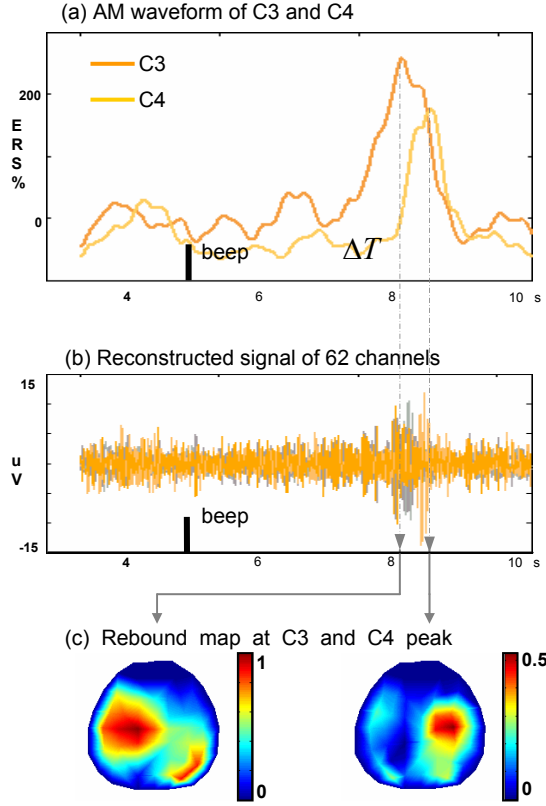
**Step 4: Extracting the beta rebound maps.** The imagery finger lifting task, similar to real finger movement, induced larger beta rebound in the contralateral sensorimotor area than that in the ipsilateral one. In addition, the contralateral beta rebound appeared earlier than the ipsilateral one. The co-existence of prominent beta rebounds at C3 and C4 and the constrained time lag between them suggested that the topographical maps with maximum rebounds at C3 and C4 were reliable features. Specifically, we looked for two time points at which both the AM waveforms of C3 and C4 have maximum peaks but with time lag ( $\Delta T$  in Figure 6(a)) less than 0.5 second. The topographical maps at these two time points, referred to as beta rebound maps (Figure 6. (c)), were concatenated into a  $124 \times 1$  column vector and used as a feature vector.

Using the same time points of the peaked beta rebound resulted from steps 1 ~ 4, we processed the data using step 3 only, i.e. without using ICA. Figure 7 depicts the extracted beta rebound maps and appears to be contaminated due to noise compared with those in Fig. 6.

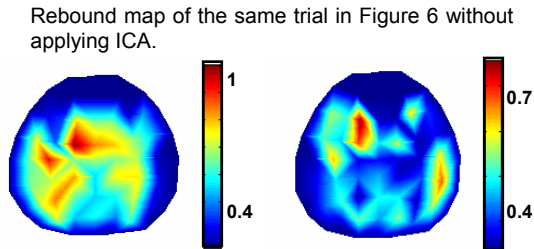
## 4. TWO-CLASS SUPERVISED CLASSIFICATION

In this section, four two-category classifiers used in our study are briefly reviewed. They were linear discriminant analysis (LDA), back-propagation neural network (BP-NN), radial basis function network (RBF-NN) and support vector machine (SVM). The beta rebound maps, denoted by  $\vec{x}_i$ , of imagery right and left hand movement, each of them is a  $124 \times 1$  column vector and, were divided into two data sets, one for training and the other for testing the classifiers. The numbers of beta rebound maps used in the training and testing phases for each subject at each session were 60 and 30. These beta rebound

maps were randomized before being used. For the sake of simplicity, we use the notation  $R$  and  $L$  to denote the category of imagery right and left hand movement, respectively, in the following discussion.



**Figure 6.** Computation of the beta rebound maps. (a) The AM waveform of C3 and C4.  $\Delta T$  was the time lag between prominent beta rebounds at C3 and C4. (b) Reconstructed signals of 62 channels (excluded HEOG and VEOG) which were used to calculate the AM waveforms in (a). (c) The beta rebound maps created from reconstructed signals on 62 channels indexed to the time points of peaked beta rebounds at C3 and C4.



**Figure 7.** The computed beta rebound maps only using steps 3 without applying ICA.

## Classifiers

### 4.1.1 LDA

The idea of LDA is to seek a vector  $\vec{w}$  so that two projected clusters of  $R$  and  $L$  feature vectors  $\vec{x}_i$ 's on  $\vec{w}$  can be well separated from each other while keeping small variance of each cluster. This can be done by maximizing the so-called Fisher's criterion

$$J(w) = \frac{w' S_b w}{w S_w w}$$

with respect to  $\vec{w}$ , where  $S_b$  is the between-class scatter matrix:

$$S_b = (m_R - m_L)(m_R - m_L)'$$

and  $S_w$  is the within-class scatter matrix:

$$S_w = \sum_{x \in R} (\vec{x} - m_R)(\vec{x} - m_R)' + \sum_{x \in L} (\vec{x} - m_L)(\vec{x} - m_L)'$$

in which two summations run over all the training samples of classes  $R$  and  $L$ , respectively, and  $m_R$  and  $m_L$  represent the group mean of classes  $R$  and  $L$ , respectively. The optimal  $\vec{w}$  is the eigenvector corresponding to the largest eigenvalue of  $S_w^{-1} S_b$ . After  $\vec{w}$  is obtained by means of the training data, we projected the test samples on it, and then classified the projected points by the k-nearest-neighbor decision rule.

### 4.1.2 BP-NN

The BP-NN was trained in a supervised manner based on the error-correction learning rule. The hierarchy of a BPNN in our implementation is depicted in Figure 8, which consists of one input layer, one hidden layer, and one output layer. The training phase was accomplished by iterating two passes: the forward and backward passes. In the forward pass of the back-propagation learning, as show in the Figure 8, the output of the BP-NN at iteration  $n$  was computed by

$$y(n) = \varphi(v(n))$$

where  $\varphi(\cdot)$  was the activation function and  $v(n)$  was the induced local field of output neuron

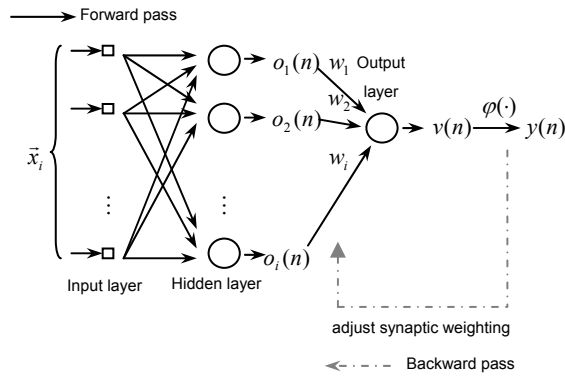
$$v(n) = \sum_{i=1}^m w_i(n) o_i(n)$$

in which  $m$  was the total number of the inputs applied to output neuron,  $w_i$  was the weight connecting neuron  $i$  to the output neuron, and  $o_i(n)$  was the output signal of neuron  $i$ . The error signal,  $e(n)$ , between  $y(n)$  and the desired

output,  $d(n)$ , was computed at each iteration. If the error met the stopping criterion, the training procedure was terminated. Otherwise, it was minimized in the subsequent backward pass to update the synaptic weighting  $w_i(n)$

$$w_i(n+1) = w_i(n) + \alpha[w_i(n-1)] + \eta\delta(n)o_i(n)$$

where  $\alpha$  was the momentum constant, and  $\delta(n)$  is the local gradients of the output layer in the network, given by  $\delta(n) = e(n)\varphi'(v(n))$ . In the testing phase, input feature vectors,  $\bar{x}$ 's, can be linearly classified according to the value of  $y(n)$  in the output layer.



**Figure 8.** The hierarchy of BP neural network.

#### 4.1.3 RBF-NN

The RBF neural network [Hay94a] uses a nonlinear function to map the input data into high-dimension space so that they are more likely to be linearly separable than in the low-dimension space [Cov65a] [Cov91a] [Cov88a]. The hierarchy of (regularization) RBF neural network is depicted in Figure9, which consists of one input layer, one hidden layer, and one output layer.

Each RBF network is designed to have a nonlinear transformation from the input layer to the hidden layer, followed by a linear mapping from the hidden layer to the output layer. The mapping between the input and output space is expressed by:

$$F(\bar{x}) = \sum_{i=1}^N w_i \varphi(\|\bar{x} - \bar{x}_i\|) \quad (4.1)$$

where  $\varphi(\|\bar{x} - \bar{x}_i\|) = e^{-\|\bar{x} - \bar{x}_i\|^2}$  and  $w_i$  represents the weighting from the  $i_{th}$  hidden neuron to output neuron, and  $\bar{x}_i$  represents the  $i_{th}$  known feature vector with dimension  $m$ ,  $i = 1, 2, \dots, N$ . The distance between input vector,  $\bar{x}$ , and center,  $\bar{x}_i$ , is mapped into high-dimension space by means of a Gaussian

function ( $\varphi(\|\bar{x} - \bar{x}_i\|)$ ) in this study. In the phase of supervised learning, training feature vectors  $\bar{x}_i$ ,  $i = 1, 2, \dots, N$ , and output desired output  $F(\bar{x}_i) = d_i$  which is either 1 or -1 in our design, are given. For the sake of simplicity, the training feature vectors are used as centers. With the known  $N$  input feature vectors and the corresponding designed outputs, the weighting  $w_i$  can be computed from the input-output relationship in equation (4.1):

$$Gw = d \quad (4.2)$$

where

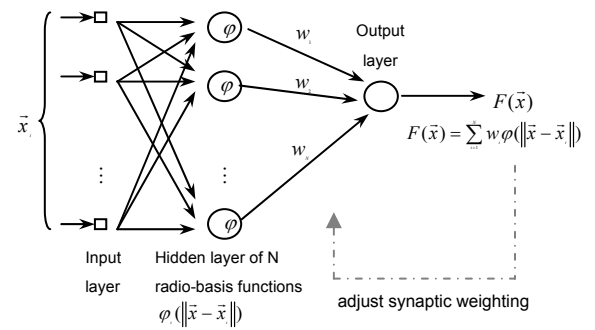
$$G = \begin{bmatrix} \varphi(\|\bar{x}_1 - \bar{x}_1\|) & \varphi(\|\bar{x}_1 - \bar{x}_2\|) & \dots & \varphi(\|\bar{x}_1 - \bar{x}_N\|) \\ \varphi(\|\bar{x}_2 - \bar{x}_1\|) & \varphi(\|\bar{x}_2 - \bar{x}_2\|) & \dots & \varphi(\|\bar{x}_2 - \bar{x}_N\|) \\ \dots & \dots & \ddots & \dots \\ \varphi(\|\bar{x}_N - \bar{x}_1\|) & \varphi(\|\bar{x}_N - \bar{x}_2\|) & \dots & \varphi(\|\bar{x}_N - \bar{x}_N\|) \end{bmatrix},$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_N \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_N \end{bmatrix}$$

By solving the linear system (4.2), the resultant weighting  $w$  vector is

$$w = G^+ d \quad (4.3)$$

where  $G^+ = (G^T G)^{-1} G^T$  is the pseudoinverse matrix of  $G$ . Compared with other neural network which uses gradient-based optimization process to estimate the weightings, for example, the back-propagation recurrent neural network, the RBF neural network solve for a set of linear equations to avoid trapping in a local minimum and greatly reduce the training time. In the testing phase, input feature vectors,  $\bar{x}$ 's, can be linearly classified based on the values of  $F(\bar{x})$ 's.



**Figure 9.** The hierarchy of RBF neural network.

#### 4.1.4 SVM

The basic idea of support vector machine hinges on two mathematical operations: (1) With an appropriate nonlinear mapping  $\varphi(\cdot)$  of an input vector into a high-dimensional feature space, data

from two categories can be linearly separated by a hyperplane [Cov65a], (2) Construction of an optimal hyperplane for separating the features in (1). Let  $\bar{x}$  denote a vector drawn from the input space, assumed to be of dimension  $m_0$  and let  $\{\varphi_j(\bar{x})\}_{j=1}^{m_1}$  denote a set of nonlinear transformations from the input space to the feature space:  $m_1$  is the dimension of the feature space. Given such a set of nonlinear transformations, we may define a hyperplane acting as the decision surface as follows:

$$\sum_{j=0}^{m_1} w_j \varphi_j(\bar{x}) = 0 \quad (4.4)$$

where  $w = \{w_0, w_1, \dots, w_{m_1}\}$  denotes a set of linear weights connecting the feature space to the output space. And it is assumed that  $\varphi_0(\bar{x}) = 1$  for all  $\bar{x}$ , so that  $w_0$  denotes the bias. Equation (4.4) defines the decision surface computed in the feature space in terms of the linear weights of the machine. Define the vector  $\varphi(\bar{x}) = [\varphi_0(\bar{x}), \varphi_1(\bar{x}), \dots, \varphi_{m_1}(\bar{x})]^T$ , and  $w = [w_0, w_1, \dots, w_{m_1}]^T$  we rewrite the decision surface in the compact form:

$$w^T \varphi(\bar{x}) = 0 \quad (4.5)$$

Given the training feature samples  $\varphi(\bar{x}_i)$  corresponds to the input pattern  $\bar{x}_i$ , and the corresponding desired response  $d_i, i = 1, \dots, N$ , which is either 1 or -1 in our design, it has been shown that [Hay94a] the optimal weight vector  $w$  can be expressed as

$$w = \sum_{i=1}^N \alpha_i d_i \varphi(\bar{x}_i) \quad (4.6)$$

where  $\{\alpha_i\}_{i=1}^N$  is the optimal Lagrange multipliers resulted from maximizing the subject function

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j \varphi^T(\bar{x}_i) \varphi(\bar{x}_j) \quad (4.7)$$

subject to the constraints (1)  $\sum_{i=1}^N \alpha_i d_i = 0$ , and (2)

$0 \leq \alpha_i \leq C$ , where  $C$  is a user-specified constant.

Substituting equation (4.6) into (4.5), we obtain the optimal hyperplane

$$\sum_{i=1}^N \alpha_i d_i \varphi^T(\bar{x}_i) \varphi(\bar{x}) = 0 \quad (4.8)$$

which will be used for linearly separating the testing data, i.e. for any testing sample  $x$ , if

$$\sum_{i=1}^N \alpha_i d_i \varphi^T(\bar{x}_i) \varphi(\bar{x}) \geq 0$$

then  $x$  is classified into the subset having the training response  $d_i = 1$ , otherwise it is classified into the other subset with  $d_i = -1$ . In our implementation, we chose the radial basis function in defining the inner-product kernel  $\varphi^T(\bar{x}_i) \varphi(\bar{x})$  as follows:

$$K(\bar{x}, \bar{x}_i) \equiv \varphi^T(\bar{x}) \varphi(\bar{x}_i) = \exp(0.0005 \|\bar{x} - \bar{x}_i\|^2) .$$

According to equation (4.8), once the number of nonzero Lagrange multipliers,  $\alpha_i$ , is determined, the number of radial-basis functions and their centers are determined automatically. This differs from the design of the conventional neural network, for example, the back-propagation neural network or radial-basis function network [Hay94a], where the numbers of hidden layers or of hidden neuron are usually determined heuristically.

## 5. RESULTS

Table 1 summarizes the averaged recognition results for detecting the right and left imagined finger lifting in four subjects (denoted by s1 ~ s4). With the use of ICA in the extraction of the beta rebound maps, each classifier has superior performance regardless of subjects and the overall averaged recognition score improved significantly from 55.0% to 74.8%. In addition, the SVM outperformed other classifiers.

Classifier	ICA	s1	s2	s3	s4	mean
LDA	without	58	55	57	51	54
	<b>with</b>	<b>63</b>	<b>79</b>	<b>74</b>	<b>63</b>	<b>69.8</b>
BP-NN	without	63	52	50	51	54
	<b>with</b>	<b>72</b>	<b>84</b>	<b>79</b>	<b>67</b>	<b>75.5</b>
RBF-NN	without	66	59	54	50	57.3
	<b>with</b>	<b>75</b>	<b>86</b>	<b>79</b>	<b>66</b>	<b>76.5</b>
SVM	without	66	53	50	51	55
	<b>with</b>	<b>72</b>	<b>87</b>	<b>77</b>	<b>73</b>	<b>77.3</b>

Table 1. Averaged recognition rates (in percentages) over four sessions resulted from different classifiers with and without using ICA for feature extraction.

The receiver operating characteristics (ROC) curve, a plot of true-positive rate versus false-positive rate, provides another way to evaluate the performance of binary detection classifiers. The area under the ROC

curve, which can be interpreted as the probability of a random sample being assigned to positive class than that to negative class, assesses the quality of classification over a range of misclassification costs. Table 2 reports that the use of ICA improved the performance of each classifier and the overall averaged ROC area increased from 0.63 to 0.75.

Classifier	ICA	s1	s2	s3	s4	mean
LDA	without	.71	.64	.58	.67	.65
	with	<b>.75</b>	<b>.86</b>	<b>.74</b>	<b>.68</b>	<b>.78</b>
BP	without	.65	.56	.61	.58	.60
	with	<b>.68</b>	<b>.78</b>	<b>.74</b>	<b>.71</b>	<b>.73</b>
RBF	without	.73	.60	.54	.62	.62
	with	<b>.65</b>	<b>.91</b>	<b>.77</b>	<b>.74</b>	<b>.77</b>
SVM	without	.64	.61	.66	.63	.64
	with	<b>.69</b>	<b>.87</b>	<b>.77</b>	<b>.65</b>	<b>.75</b>

Table 2. Averaged ROC areas over four sessions resulted from different classifiers with and without using ICA for feature extraction. The numbers of beta rebound maps used for training and testing for each subject at each session were 60 and 30.

## 6. CONCLUSIONS

We have presented a novel method using ICA in extracting a reliable feature, the beta rebound map, from the peaked ERS of motor imagery EEG. With a minimum training for each subject (20 minutes only), satisfactory classification rates from four classifiers have been achieved. This demonstrated the suitability of beta rebound map as neural input signals in the application of BCI systems.

## 7. ACKNOWLEDGMENTS

The study was funded by the Taipei Veterans General Hospital, Taiwan 91380, the Ministry of Education of Taiwan (89BFA221401), and the National Science of Council, Taiwan (NSC-92-2218-E-010-016).

## 8. REFERENCES

- [Clo96a] Clochon, P., Fontbonne, J. M., Etevenon, P. A new method for quantifying EEG event-related desynchronization: amplitude envelope analysis, *Electroencephalography & Clinical Neurophysiology*, 98: 126-129, 1996.
- [Cov65a] Cover, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, *IEEE transactions on electronic computers*, EC-14: 326-334, 1965.
- [Cov91a] Cover, T. M., Thomas, J. A. Elements of Information Theory. New York: Wiley, 1991.
- [Cov88a] Cover, T. M., Capacity problems for linear machines. Washington, DC: Thompson Book, Pattern Recognition: 293-289, 1988.
- [Hay94a] Haykin, S. Neural Network: A Comprehensive Foundation. New York: Macmillan College Publishing Company, 1994.
- [Lee03a] Lee, P.L., Wu, Y.T., Chen, L.F., Chen, Y.S., Cheng, C.M., Yeh, T.C., Ho, L.T., Chang, M.S., Hsieh, J.C. ICA-based spatiotemporal approach for single-trial analysis of post-movement MEG beta synchronization, *NeuroImage*, in press, 2003.
- [Mul99a] Muller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H. Designing optimal spatial filters for single-trial EEG classification in a movement task, *Clinical neurophysiology*, 110:787-798, 1999.
- [Pfu96a] Pfurtscheller, G., Stancak Jr, A., Neuper, C. Post-movement beta synchronization. A correlate of an idling motor area?, *Electroencephalography & Clinical Neurophysiology*, 98:281-293, 1996.
- [Pfu98a] Pfurtscheller, G., Neuper, C., A., Schlogl, Lugger, K. Separability of EEG Signals Recorded During Right and Left Motor Imagery Using Adaptive Autoregressive Parameters, *IEEE transactions on Rehabilitation Engineering*, Vol. 6, No. 3: 316-325, 1998.
- [Pfu00a] Pfurtscheller, G., Guger, C., Muller, G. Krausz, G., Neuper, C. Brain oscillations control hand orthosis in a tetraplegic, *Neuroscience letters*, 292: 211-214, 2000.