# An Approach for Pattern Recognition of EEG Applied in Prosthetic Hand Drive

**Xiao-Dong ZHANG, Yun-Xia WANG, Yao-Nan LI, Jin-Jin ZHANG**
**School of Mechanical Engineering, Xi'an Jiaotong University**
**Xi'an, Shaanxi Province 710049 China**

**and**

**Xiao-Dong ZHANG**
**State Key Laboratory of Robotics, Shenyang Institute of automation Chinese Academy of Sciences**
**Shenyang, Liaoning Province 110016 China**

## ABSTRACT

For controlling the prosthetic hand by only electroencephalogram (EEG), it has become the hot spot in robotics research to set up a direct communication and control channel between human brain and prosthetic hand. In this paper, the EEG signal is analyzed based on multi-complicated hand activities. And then, two methods of EEG pattern recognition are investigated, a neural prosthesis hand system driven by BCI is set up, which can complete four kinds of actions (arm's free state, arm movement, hand crawl, hand open). Through several times of off-line and on-line experiments, the result shows that the neural prosthesis hand system driven by BCI is reasonable and feasible, the C-support vector classifiers-based method is better than BP neural network on the EEG pattern recognition for multi-complicated hand activities.

**Keywords:** Pattern Recognition, EEG, Brain-Computer Interface, Prosthesis hand**.**

## 1. INTRODUCTION

As a control source of prosthetic hand, EEG (Electroencephalography) is the emerging new and hot research direction in the world. Neurophysiologic studies have shown that EEG information as an information source controlling prosthetic can not only maximize the patients' residual function of brain, but also can transmit directly the control information of the human brain. In addition, the mode of EEG domestic and abroad have made some achievements on the research of BCI-driven prosthetic hand, however, due to the complexity of the EEG signal, the research of neutral prosthetic hand of BCI-driven is still in the laboratory phrase of the study[1]. In addition, the researchers are committed to raising the recognition speed and accuracy by the way of improving the algorithm feature recognition. There is not yet one accepted approach that can achieve the good results of information processing. This paper is based on the characteristics of prosthetic hand's control signals, study in-depth the algorithm of pattern recognition based on that and finally reached the research purpose of raising pattern recognition.

## 2. CONSTRUCTION OF NEURAL PROSTHESIS HAND SYSTEM DRIVEN BY BCI

**Principles of System Construction**

According to structure of the BCI system, the control system of neural prosthesis hand driven by BCI is mainly consist of EEG Signal Detection System, Feature Extracting and Pattern Recognition system and Prosthetic Hand Driven System. The whole model of the system is shown as the Fig. 1.


Fig.1 The neural prosthesis hand system driven by BCI

The Prosthetic Hand is the external control object of the BCI system in this article. Because of the limit of the patterns which can be recognized by the BCI, the Prosthetic Hand which has three degrees of freedom is applied in this BCI control system. And to meet the need of system, EMG testing system is removed and the EMG controlling Circuit is changed to EEG. Three functions that hand opening and closing, wrist of 360-degree rotation, Elbow flexion are achieved in this article. And at last the Prosthetic Hand is fixed on the Mannequins for easy during the experiment.

**Feature Extraction for EEG Driving Neural Prosthesis Hand System**

EEG contains a wealth of information of brain activity, so in order to recognize different hand movement in the BCI driven system, effective feature of the EEG should be extracted.

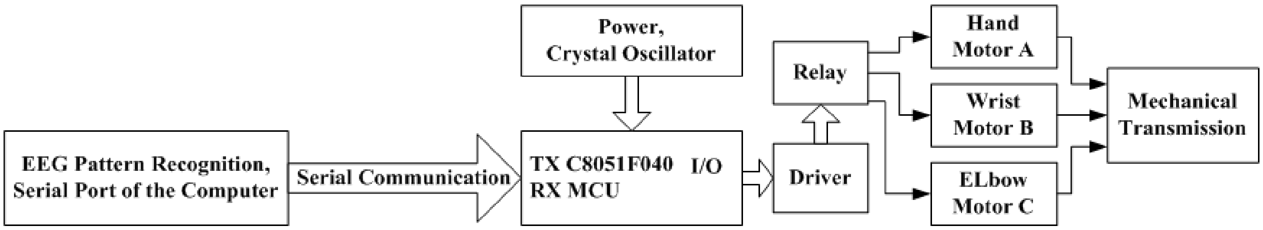


Fig.2 The structure of the neural prosthesis hand system

1) Basic characteristics of EEG signal: EEG signal is a dynamic, random, non-linear bio-signal, its amplitude is low and it has little energy. According to the International 10-20 system electrode, the amplitude of the EEG got from the scalp is between 0.1~200μV, and the frequency is between 0.5~50Hz. And it mainly includes five components in frequency domain:δ：0.5~3.5 Hz, θ: 4~7 Hz, α：8~13 Hz，β：14~30 Hz,γ：31Hz. According to related research, the Stimulus of different areas of cortex is different when people have different actions or imagine different things, so different EEG signal is produced. For P3 and P4 is the best measuring point for hand actions, the signal of P3 and P4 is applied as the object of data analysis and feature extraction[2-5].
2) Feature extraction of EEG signals: In this paper, the EEG signal is decomposed to five levels on the basis of Daubechies4 wavelet, and at last five useful frequency levels are obtained. They are 0-4Hz, 4-8Hz, 8-16Hz, 16-32Hz and 32-64Hz; Which are exactly similar to the δ, θ, α，βand γ.So an Eigenvector matrix which has five lines and 1024 columns can be easily got from the data of the five EEG frequency band, and then singular value decomposition can be applied to this matrix. And a 5-dimensional feature vectors can be obtained from the data, which can be used as the feature for the Pattern Recognition.

**Construction of Neural Prosthesis Hand System Driven by BCI**

Prosthetic hand drive control system is shown as Fig.2, the whole control system is mainly consist of prosthetic hand controller, peripheral device, power driver IC, electrical relays and so on.
Through a serial communications PC send the results of recognize to the prosthetic hand drive controller, and the controller processes the corresponding three motor-driver's rotation, stop and turn. And the prosthetic hand movement is achieved through the installation of mechanical.

## 3. PATTERN RECOGNITION METHODS

**EEG pattern recognition based on Neural network**

Artificial neural network (ANN) have self-learning, self-organization, better fault-tolerance and good nonlinear approximation ability and about 90% of them is used BP network. In this paper a BP neural network is designed for EEG pattern recognition on hand movement.
In this study, the input layer is set by 12 nodes, corresponding to elements in the normalized wavelet characteristic matrix Tk. You know, four basement hand movement states, namely free state, arm movement, hand crawl and hand open should be concerned if a robot arm will complete a whole grasp operation. So two neurons (y1 and y2) are set in the output layer, while the output results on the corresponding hand movement patterns as shown in Table 1.

TABLE I. OUTPUT AND ITS HAND PATTERN

| Action Pattern | Output | |
|---|---|---|
| | $y_1$ | $y_2$ |
| free state | 0 | 0 |
| arm movement | 0 | 1 |
| hand crawl | 1 | 0 |
| hand open | 1 | 1 |

In this paper, the networks is small or medium structure, and therefore a hidden layer is preferred. The number of hidden layer nodes, in theory, there is not a clearly defined, so this paper we uses cut-and-trial method to determine the optimal number of hidden nodes. Cut-and-trial method is to use the same sample set with different hidden layer nodes in the network for training, until the weight does not change, then the network stabilized. By this method, we found that 22 hidden layer neurons is appropriate. Therefore, the finalization of the BPNN structure is 12-22-2. That is, the BP neural network input layer has 12 neurons for the 12 characteristic of EEG signal and the output layer has two neurons expressing four different actions of hands; network has only one hidden layer neuron with 22 nodes.
After the feature vectors of EEG signals obtained, in accordance with the method described in this section, we can recognize the pattern by recognizing feature vectors of hand movements.

**Multi-Classification of EEG Signals based on Support Vector Machine**

Using support vector machine classifier as a multi-class model, its basic principles and classification are as follows.

1) Structural Risk Minimization principle: In 1971, Vapnik have pointed out relationship between empirical risk and expected risk[6] in the machine learning early.

$$R(f) \leq R\_emp\ (f) + \Phi(n/h) \tag{1}$$

Where R(f) is the expected risk, Remp (f) is the empirical risk, Φ(n/h) is the fiducially range, f is a learning machine functions, n is the number of training samples, h is the VC dimension of function reflecting the learning machine complexity and Φ is greater while h is greater, leading the difference between expected risk and experience risk is the greater risk. So in order to make the expected risk minimization, we need make the empirical risk and fiducially range minimization at the same time ;thus function f have a small VC dimension. This is to make f have a smaller VC dimension, which is the principle Structural Risk Minimization, and the support vector machine is a learning algorithm based on this principle.

2) C-Support Vector Classification Machine (C-SVC): Setting the training samples set is: {(xi, yj), i = 1,2, ..., l}, {xi, yj}, 1, ..., l, xi is a d-dimensional vector, yi∈{-1, +1}. The original problem of classification is that the process of solving Optimal Hyper plane and that is to maximize the distance. As shown in Fig.3 whose mathematical expression is a constrained optimization problem.

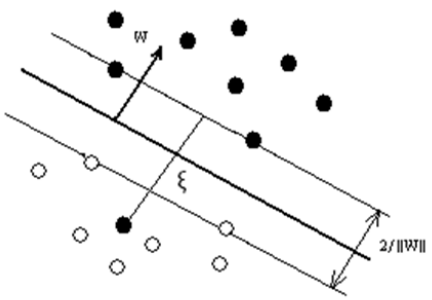


Fig.3 Optimal hyperplane

$$\begin{aligned} &\min_{\xi,\omega,b} \quad \langle\omega\cdot\omega\rangle + C\sum_{i=1}^{l}\xi_{i,} \\ &s.t. \quad y_i(\langle\omega\cdot\Phi(x_i)\rangle + b) \geq 1-\xi_i, i = 1,\dots,l, \\ &\qquad \xi_i \geq 0,\ i = 1,\dots,l. \end{aligned} \tag{2}$$

In the meantime, the Lagrange dual expression is as.

$$\begin{aligned} &\max\ W(\alpha) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i, x_j), \\ &s.t. \quad \sum_{i=1}^{l} y_i\alpha_i = 0, \\ &\qquad C \geq \alpha_i \geq 0, i = 1,\dots,l. \end{aligned} \tag{3}$$

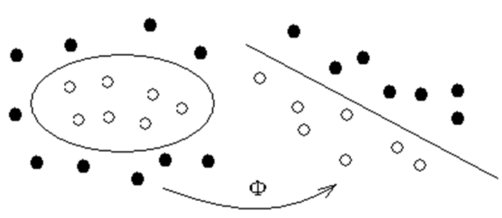


Fig.4 Nonlinear mapping

The ξi is slack variables, which allow the border is disrupted and C is the corresponding coefficient of error of punishment; and therefore such a support vector classification machine is also known as C-SVC. Φ(xi) is a non-linear mapping (Fig.4), it make sample xi mapped to a high-dimensional linear separable feature space; K(xi,yi) satisfy Mercer's theorem of symmetric functions, and also known as the kernel for completing the sample's inner product operations in high-dimensional space, the introduction of <Φ(xi), Φ(xj)> sovle the "curse of dimensionality" in high-dimensional feature space. Commonly the kernel function has the following three kinds.

Polynomial kernel：$K(x,z) = \left((x\cdot z)+t\right)^d$ (4)

Gaussian RBF kernel：$K(x,z) = exp(-\|x-z\|^2/\sigma^2)$ (5)

Sigmoid nucleus：$K(x,z) = tanh(k_1(x\cdot z)+k_2)$ (6)

It can be seen that equation (3) is a quadratic optimization problem, whose solution is the global optimal solution, and the $\alpha_i$ which dose not correspond to 0 is corresponding to the sample called support vector. The discriminated function as follows.

$$f(x) = sgn\left(\sum_{i=1}^{l}\alpha_i y_i K(x_i,x)+b\right) \tag{7}$$

According to KKT conditions, we can see $b = y_j - \sum_{i=1}^{l} y_i\alpha_i K(x_i, x_j)$, and it can be calculated by any 0<αi <C which is not unique and often preferable to the mean.

1) Parameter Optimization: As we can see, there are many parameters in the support vector machine, the choices of which determine the complexity, generalization capacity and robustness of the classifier at a great extend, so parameter optimization is significant. Many optimization algorithms has been applied in the support vector machine parameter optimization problem, such as grid-point method, k-fold cross-validation method, gradient algorithm, genetic algorithms, particle swarm optimization, ant colony algorithm and simulated annealing algorithm, and so on.

2) Multi-classification problem in support vector machine: Like most of the classifiers, support vector machine is proposed for the two classification problems, we need a certain way to convert it to a multiple classifiers when it was applied in the Multi-classification problem. At first, we give multi-classification problem the following description, assuming that there are k categories, the training sample ((xi, yj), i = 1,2, ..., l), here. Commonly used support vector machine multi-classification methods are the following kinds.

One VS all(OVA)[7],regarding the one class of K classes and the other classes as a two sub-problems, and classes them by constructing k sub-problems vectors, but there will be "rejection" phenomenon, which reduces the recognition rate.

One VS one (OVA) [7], the k class is divided to k (k-1) / 2 two sub-problems by two in a class, and constructing some classifier to achieve multi-classification task, and finally give the classification by voting decision-making. The method often requires a number of classifiers, and the "rejection" phenomenon exist the same.

Decision tree method (DT) [8],it combine the binary trees and support vector classification machines to construct K classifiers to achieve multi-classification. In the method, the classification of each node parameters may be different, which will help improve accuracy; but once the error in classification extended to the next node, the following classification will make no sense. As there are many possibilities when we choose a class and others to classify at a node, it increased the complexity.

Error-correcting output coding method (ECOC) [9], it uses pass phrase encode the classifiers, and classes them by measuring the Hamming distance between the DUT code and someone code. It has good error correction capacity to error data;

Determining the many types of target function method[10], The idea is constructing a multi-class classifier by changing the objective function; it optimizes all the variables in a optimization objective. Due to too many variables to optimize, optimize results are often unsatisfactory, which make it difficult for use in large-scale problems.

Multi-class Classification based on one classification method[11],The method evolves from a classification problem, and determining the respective categories by solving the center of more than one category and then calculating the sample to the distance from the center of each type.

Principle of the above methods varies form each other, each with advantages and disadvantages, there are different manifestations in different occasions, specific problem should choose which methods not yet have a better selection standard. For the ease of use, the one-to-all method is used widely relatively. When choosing the best method of multiple classifiers, we still need to test the various methods to select.

## 4. EXPERIMENT

### System Building

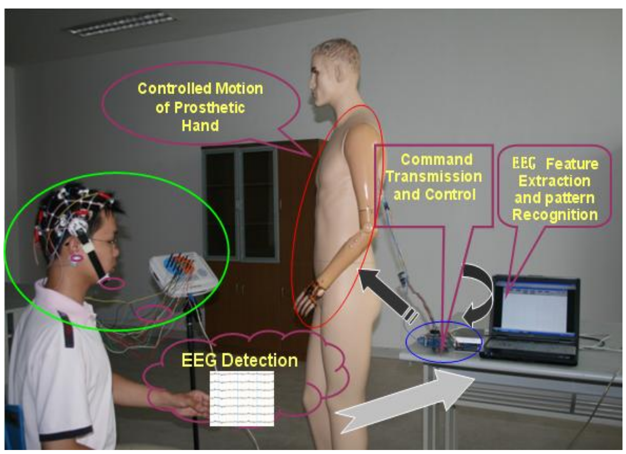


Fig.5 The neural prosthesis hand system driven by BCI

According to the system model, a neural prosthesis hand system driven by BCI is constructed as shown in Fig.5,by integrating the EEG signal detector system, EEG signal analyzing and recognition system and prosthesis hand control system, which have been developed in previous work.

This experiment conduct the pattern recognition by using BP neural network (BPNN) and multi-classification algorithms based on C-support vector machine respectively. The outcome of the EEG recognition realized the corresponding movement of the prosthesis hand by prosthesis hand control driver unit. Fig.6 demonstrate the 4 different experimental statuses: arm free, arm fluctuating, hand close and hand open.

During the experiment, the experimenter must concentrate on the movement imagination as well as doing the corresponding movements to strong the EEG signal generated from the experimenter under each pattern. The nonstandard movements of the arm or hand should be avoiding to the best of the experimenter's ability to minimize the error of the EEG acquisition. The time that the experimenter starts the movement must be synchronous with the time start the online pattern recognition. Each movement of each experimenter is repeated 20 times and totally 80 times. The highest recognition rate is statistical base on the correct recognition times of each pattern.

### Pattern Recognition by Neural Network

Conduct the online EEG pattern recognition experiment and prosthesis hand drive experiment of each hand movement using the neural network structure parameter of each experimenter's EEG data in the highest offline recognition rate condition.

Table.2 shows that the highest recognition rate, using by BPNN under 4 different movement-arm free, arm fluctuating, hand close and hand open, is about 60%.The rate is of little different due to the different experimenter and the experimental environment.

TABLE II. THE BEST ONLINE RECOGNITION RATE

| Experimenter | Experimenter 1 | Experimenter 2 | Experimenter 3 | Average |
|---|---|---|---|---|
| Recognition Rate | 55% | 65% | 62.5% | 60.8% |

Fig.6 Experiment of the neural prosthesis hand driven by BCI

TABLE III. THREE METHODS AND THEIR PARAMETER

| Multi-classification Method | Kernel Function | Kernel Parameter | C | Support Vector Number | Recognition Rate |
|---|---|---|---|---|---|
| DT | Poly | $\gamma$ =10, d=10, t=1 | 1 | 37 | 90% |
| DT | Rbf | $\sigma^2 = 10$ | 100 | 50 | 80% |
| MOC | Rbf | $\sigma^2 = 2$ | 1 | 60 | 65% |

**Multi-Classification Recognition Based on C-Support Vector Machine**

Use the experimenter 2's experimental data to validate the recognition precision. Select the top 15 item of the 4 group EEG signal data corresponding to each hand movement as the training samples and rest of them as the testing samples.

Using the MATLAB7.01 as the simulation platform and the LSSVM [12], OSUSVM 3.00 [13] as the tool of the Multi-classification based on support vector machine. Both of them support the multi-classification task. The classification methods of the LSSVM are MOC, OVA, ECOC and include linear function, Rbf, Poly function and Sigmoid function as well as cross validation function. The classification method of the latter is decision tree, which has 1 sigmoid kernel function less than the former. This experiment conduct the comparative test by combining the different multi-classification methods and different kernel functions mentioned above and confirm an optimized result. It is worthy point out that the form of the kernel function-Poly function in OSUSVM toolbox is different with expression (4), it is as follow.

$$K(x,z) = (\gamma(x \cdot z) + t)^d \tag{8}$$

Combining the different multi-classification methods and kernel function, and selecting the suitable parameter of each group to reach the highest recognition rate respectively. Because of the much combination and the huge recognition result difference between each group, the highest rate is 90% and the lowest is only 25%.Only the top 3 methods and parameters are given as in Table 3.

In the poly kernel function, the coefficient $\gamma$ is of the significant effect. The bigger $\gamma$ means the higher recognition rate of the classifier. This conclusion is accordance with the deduce of the reference [14]. It can be concluded from the Tab.3 that the resolution is improved obviously and the amount of the support vectors are reduces by using the poly kernel function form of expression (8) and enlarge the value of $\gamma$.

## 5. RESULT ANALYSIS AND CONCLUSIONS

It can be seen that in the pattern recognition of prosthetic-driven EEG the identification method based C-support vector machine multi-classification is more efficient than the BP neutral network. This may be mainly due to poor BP network fault tolerance and the incompleteness of its algorithm. BP algorithm goal is to minimize the "error", therefore in the iterative process of training the network, the shock should not be too big and should get smaller and smaller. That is, the final iteration results with only a small difference between the target values. Therefore BP network fault tolerance is inevitably poor. At the same time, because BP algorithm is based on the gradient method which only has local search ability, so it would be equally difficult to avoid trapping in local minimum. So the algorithm is incomplete. While the algorithm which can enhance the BP network's global search ability such as genetic one and simulated annealing once appeared, but it can not overcome the increased difficulties in their calculations. What's more, random selection of initial weight values increase the uncertainty and complexity of BP algorithm search the global minimum point. And when they are used as samples one by one "forgotten" phenomenon may appear. What's corresponding with the BP network's mentioned shortcomings, support vector machine based on the role of maximizing the interface interval introduce slack variable and corresponding penalty coefficient C, achieves soft margin classification and improves fault tolerance. Its optimization equation is quadratic one. Its solution is global optional solution. Because of these advantages, in this experiment with C-SVS as classification Support Vector Machine Identification, the recognition rate is significantly higher than BP neutral network, and shows good classification ability.

So, some important conclusions can be obtain as follows.

1) The organized brain-computer interface driven neutral prosthetic hand system is reasonable and feasible.
2) To the pattern recognition of the complex EEG, the identification method based on C-support vector machine multiple-classification is more efficient than the BP neutral network.
3) The support vector machine based on the principle of structural risk minimization are ideally suited for small sample learning. The application of its nuclear function is a good solution to the "curse of dimensionality" problem, and the generalization ability of classifiers are more strong.
4) Multi-classification support vector machine based on decision tree selected in this paper can effectively raise the recognition rate of multi-hand operation of the complex EEG signal classification, for the application of BCI system it has great significance.

However, on how to choose multi-classification method of support vector machine, there is still no fast-track approach which is systematic and has guiding significance. This paper is to finish the choice through multiple ways of testing. Therefore, in the viewpoint of supporting vector machine applied to the issue of pattern recognition there is still much research space. So multi-class pattern recognition evaluation system should be established from BCT system as a whole, considering the choosing methods of the samples, the judgment of the classifier is reliability.

In addition, it has great significance to study EEG's relation with different characteristics of identities, to find the general categorization of EEG among people and study the EEG general processing model which is more appropriate to specific individual under some specific mode.

## ACKNOWLEDGMENT

The authors are very grateful for the supports from the State Key Laboratory of Robotics Open Project (RLO200809) and the Natural Science foundation of Jiangsu Province(BK2007059).

## REFERENCES

[1] W.H.Dai , X.D.Zhang, “Design on the System of Brain-Computer Interface Driving Neural Prosthesis Hand”, **Key Engineering Materials,** Vols. 392-394, 2009, pp 1012-1018.

[2] X.D.Zhang, T.H.Kang, H.R.Choi, “An Approach for Pattern Recognition of Hand Activities Based on EEG and Fuzzy Neural Network”, **Journal of Mechanical Science and Technology,** Vol.19, No.1, 2005, pp.87-96.

[3] X.D.Zhang, H.R.Choi, “Xiaodong Zhang, Hyoukryeol Choi, Pattern Recognition of Human Grasping Operations Based on EEG”, **International Journal of Control, Automation, and Systems**, vol. 4, no. 5, 2006, pp.592-600.

[4] P.Zhuang, “Electroencephlography event-related desynchronization and synchronization activities and movement related exercises”, **Chinese Journal Of Clinnical Rehabilitation**, Vol.8, No.1, 2004, pp.152-154.

[5] A.Grossman., J.Morlet, “Decimposition of hardy function into square integrable wavelets of construct shape”, **SIAM, J.Math. Anal.**, Vol.15, 1984, pp.723-736.

[6] V. Vapnik. **The Nature of Statistical Learning Theory**. **Springer-Verlarg** [M]. New York, 1995.

[7] H.C.Wei, L.C.Jen, “A comparison of methods for multi-class support vector machines”, **IEEE Transaetions on Neural Networks,** Vol.13, No.2, 2002, pp.415-425.

[8] K.P.Bennett, J.A.Blue, “A support vector machine approach to decision trees”, **Proceeding of IJCNN’98**[C]. Anchorage Alaska: IEEE Press, 1998: 2396-2401.

[9] T.G. Dietterieh, G.Bakiri, “Solving multiclass learning problems via error-correcting output codes”, **Journal of Artificial Intelligence Research**, Vol.2, 1995, pp.263-286.

[10] J.Weston, C.Watkins, “Support vector machines for multi-class pattern recognition”, **In Proceedings of ESANN'99**. Brussels, 1999, pp.219-224.

[11] D.S.Sun; J.P.Wu, “Vehicle Recognition Based on 1—-SVM”, **Communication and Transportation Systems Engineering and Information**, Vol.3, No.4, 2003, pp.34-37.

[12] http://www.esat.kuleuven.ac.be/sista/lssvmlab.

[13] http://www.eleceng.ohio-state.edu/~maj/osusvm.

[14] T.T.Luo; T.H.Fan, “Survey of SVM Kernel Function and Optimum”, **Ordnance Industry Automation**, Vol.26, No.10, 2007, pp.34-48.