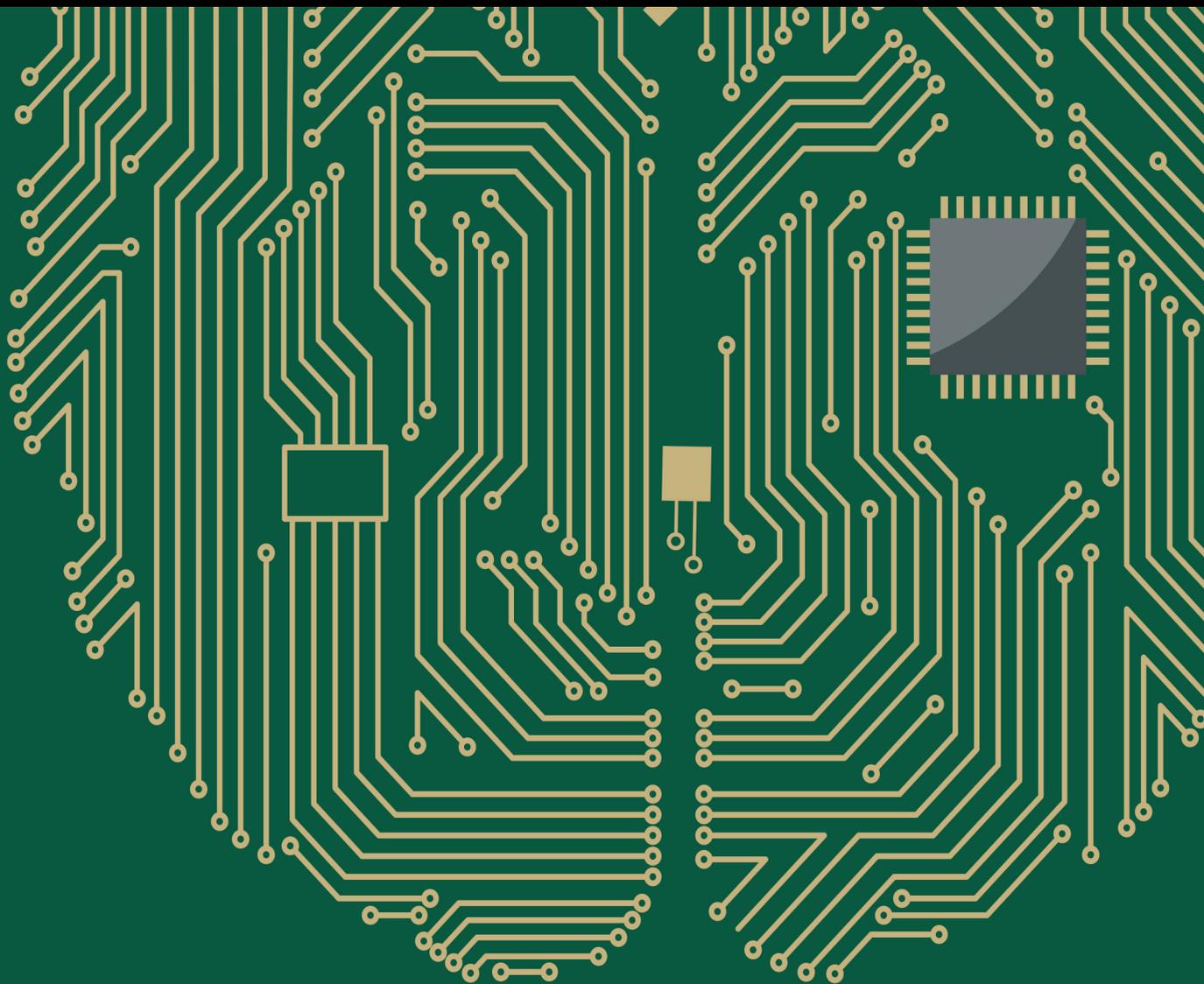


Robot and Neuroscience Technology: Computational and Engineering Approaches in Medicine

Guest Editors: Hiroki Tamura, Shangce Gao, Qixin Cao, Chuntao Leng, Hui Yu,
and Harold Szu





**Robot and Neuroscience Technology:
Computational and Engineering Approaches
in Medicine**

**Robot and Neuroscience Technology:
Computational and Engineering Approaches
in Medicine**

Guest Editors: Hiroki Tamura, Shangce Gao, Qixin Cao,
Chuntao Leng, Hui Yu, and Harold Szu



Copyright © 2016 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “Computational Intelligence and Neuroscience.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Ricardo Aler, Spain
Pietro Aricò, Italy
Hasan Ayaz, USA
Sylvain Baillet, Canada
Theodore W. Berger, USA
Steven L. Bressler, USA
Vince D. Calhoun, USA
Francesco Camastra, Italy
Ke Chen, UK
Michela Chiappalone, Italy
Andrzej Cichocki, Japan
Jens Christian Claussen, Germany
Silvia Conforto, Italy
Justin Dauwels, Singapore
Artur S. d'Avila Garcez, UK
Christian W. Dawson, UK
Paolo Del Giudice, Italy
Thomas DeMarse, USA
Piotr Franaszczuk, USA
Leonardo Franco, Spain
Doron Friedman, Israel
Samanwoy Ghosh-Dastidar, USA
J. Manuel Gorriz Saez, Spain
Manuel Graña, USA

Rodolfo H. Guerra, Spain
Christoph Guger, Austria
Stefan Haufe, Germany
Dominic Heger, Germany
Stephen Helms Tillery, USA
J. A. Hernández-Pérez, Mexico
Luis Javier Herrera, Spain
Etienne Hugues, USA
Paul C. Kainen, USA
Pasi A. Karjalainen, Finland
Dean J. Krusienski, USA
Mikhail A. Lebedev, USA
Yuanqing Li, China
Cheng-Jian Lin, Taiwan
E. López-Rubio, Spain
Reinoud Maex, France
Hong Man, USA
Kezhi Mao, Singapore
J. D. Martín-Guerrero, Spain
Sergio Martinoia, Italy
Elio Masciari, Italy
Michele Migliore, Italy
Haruhiko Nishimura, Japan
Klaus Obermayer, Germany

Karim G. Oweiss, USA
Massimo Panella, Italy
Fivos Panetsos, Spain
Jagdish Patra, Australia
S. Samarasinghe, New Zealand
Saeid Sanei, UK
Michael Schmuker, UK
Sergio Solinas, Italy
Stefano Squartini, Italy
Hiroshige Takeichi, Japan
Toshihisa Tanaka, Japan
Jussi Tohka, Spain
C. M. Travieso-González, Spain
Lefteri Tsoukalas, USA
Marc Van Hulle, Belgium
Pablo Varona, Spain
Meel Velliste, USA
F. B. Vialatte, France
Ricardo Vigario, Finland
Thomas Villmann, Germany
Michal Zochowski, USA
Rodolfo Zunino, Italy

Contents

Robot and Neuroscience Technology: Computational and Engineering Approaches in Medicine

Hiroki Tamura, Shangce Gao, Qixin Cao, Chuntao Leng, Hui Yu, and Harold Szu
Volume 2016, Article ID 5204510, 1 page

Ubiquitous Robotic Technology for Smart Manufacturing System

Wenshan Wang, Xiaoxiao Zhu, Liyu Wang, Qiang Qiu, and Qixin Cao
Volume 2016, Article ID 6018686, 14 pages

EOG-sEMG Human Interface for Communication

Hiroki Tamura, Mingmin Yan, Keiko Sakurai, and Koichi Tanno
Volume 2016, Article ID 7354082, 10 pages

A Fast Framework for Abrupt Change Detection Based on Binary Search Trees and Kolmogorov Statistic

Jin-Peng Qi, Jie Qi, and Qing Zhang
Volume 2016, Article ID 8343187, 16 pages

Adaptive Shape Kernel-Based Mean Shift Tracker in Robot Vision System

Chunmei Liu, Yirui Wang, and Shangce Gao
Volume 2016, Article ID 6040232, 8 pages

An Interactive Astronaut-Robot System with Gesture Control

Jinguo Liu, Yifan Luo, and Zhaojie Ju
Volume 2016, Article ID 7845102, 11 pages

Rhythmic Oscillations of Excitatory Bursting Hodgkin-Huxley Neuronal Network with Synaptic Learning

Qi Shi, Fang Han, Zhijie Wang, and Caiyun Li
Volume 2016, Article ID 6023547, 8 pages

Editorial

Robot and Neuroscience Technology: Computational and Engineering Approaches in Medicine

Hiroki Tamura,¹ Shangce Gao,² Qixin Cao,³ Chuntao Leng,⁴ Hui Yu,⁵ and Harold Szu⁶

¹*Department of Environmental Robotics, Faculty of Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai-nishi, Miyazaki-shi 889-2192, Japan*

²*Department of Intellectual Information Engineering, Faculty of Engineering, University of Toyama, No. 5506, Information Technology Building (G9), Gofuku 3190, Toyama 930-8555, Japan*

³*Research Institute of Robotics, Shanghai Jiao Tong University, No. 800 Dong Chuan Road, Shanghai 200240, China*

⁴*Engineering Training Center, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Minhang District, Shanghai 200240, China*

⁵*School of Creative Technologies, University of Portsmouth, Winston Churchill Avenue, Portsmouth PO1 2UP, UK*

⁶*Biomedical Engineering, Catholic University of America, Washington, DC 20064, USA*

Correspondence should be addressed to Hiroki Tamura; htamura@cc.miyazaki-u.ac.jp

Received 5 June 2016; Accepted 6 June 2016

Copyright © 2016 Hiroki Tamura et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Robot and neuroscience technology based on computational and engineering approaches, which has been successfully applied to a wide variety of fields such as medicine, provides giant opportunities for the advancement of human-computer interface's application to promote the medical research, improve the quality of life, and enhance the patient's safety. Academia generally takes it for granted that the findings in robot and neuroscience have greatly promoted the development of artificial intelligence and consequently developed robotic technology. The research which focuses on interdisciplinary fields includes various areas like artificial intelligence, models, and computational theories of human cognition, perception, and motivation and brain models, artificial neural nets, and neural computing. Thus, more and more attention has been paid to the research progress on computational intelligence and neuroscience and their potential applications in robotics.

The scope of this issue has been restricted to items that are relevant to building theoretical and practical systems including contributions in the area of applicable neural networks theory, supervised and unsupervised learning methods, algorithms, architectures, performance measures, applied statistics, software simulations, hardware implementations, benchmarks, system engineering and integration, and innovative applications. There were a total of 15 manuscripts received, six of which were accepted for publication. The adopted articles embrace a scope that is the representative of computational

intelligence and neuroscience technologies, namely, rhythmic oscillations with synaptic learning, human-computer interface for communication, interactive astronaut-robot system, novel abrupt change detection framework, ubiquitous robotic technology's application, and object tracking in robot vision system.

The six papers offer many ideas in robot and neuroscience technology.

This special issue sheds light on bridging the gap between neuroscience, artificial intelligence, and engineering. Editors truly hope this special issue celebrates the communal efforts in this field and becomes the academic citation reference for the researchers of robot and neuroscience technology and of the beneficiaries.

Acknowledgments

We thank all authors for submitting their papers to this special issue. We also thank all reviewers for providing useful comments.

*Hiroki Tamura
Shangce Gao
Qixin Cao
Chuntao Leng
Hui Yu
Harold Szu*

Research Article

Ubiquitous Robotic Technology for Smart Manufacturing System

Wenshan Wang, Xiaoxiao Zhu, Liyu Wang, Qiang Qiu, and Qixin Cao

Research Institute of Robotics, Shanghai Jiao Tong University, Shanghai 200240, China

Correspondence should be addressed to Wenshan Wang; amigo@sjtu.edu.cn

Received 29 November 2015; Revised 26 April 2016; Accepted 9 May 2016

Academic Editor: Hiroki Tamura

Copyright © 2016 Wenshan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the manufacturing tasks become more individualized and more flexible, the machines in smart factory are required to do variable tasks collaboratively without reprogramming. This paper for the first time discusses the similarity between smart manufacturing systems and the ubiquitous robotic systems and makes an effort on deploying ubiquitous robotic technology to the smart factory. Specifically, a component based framework is proposed in order to enable the communication and cooperation of the heterogeneous robotic devices. Further, compared to the service robotic domain, the smart manufacturing systems are often in larger size. So a hierarchical planning method was implemented to improve the planning efficiency. A test bed of smart factory is developed. It demonstrates that the proposed framework is suitable for industrial domain, and the hierarchical planning method is able to solve large problems intractable with flat methods.

1. Introduction

As the era of Industry 4.0 comes, industrial robots are no longer the preprogrammed robots setting separately doing their repeating jobs [1, 2]. As the manufacturing tasks become more individualized and more flexible, it shows great prospect to develop smart manufacturing systems, where machines are not likely to be preconfigured by traditional teaching methods, but doing variable tasks and coping with a wide variety of unexpected environmental and operational changes. The future manufacturing industry also requires that the system could dynamically schedule the tasks for these machines according to their work loads and the received tasks.

This feature of doing various tasks utilizing collaboration of distributed devices shares common ideas with the ubiquitous robotic technology, which is mainly applied in service robots domain [3]. In this perspective, the novel industrial manufacturing system could take advantages of the ubiquitous robotic technology.

In a typical ubiquitous robotic system, robotic devices are developed into modules [4–6]. These modules are connected through network, enabling data sharing and functionality calling. This modularized framework, which brings painless

modification, expansion, and deletion, could also be applied to the smart manufacturing domain. We propose in this paper a framework of smart factory that takes advantage of a component based method, which abstracts each machinery process as a module with standardized communication ports. So different machines are able to communicate and cooperate with each other upon these ports.

Another important issue of ubiquitous robotic systems is the development of a task level learning and planning module that handles various tasks and dynamic environment without recoding the robots [7, 8]. This is also critical for smart factories, where there may be a variety of orders and different situations for each order. For example, in a future smartphone assembly factory, customers could make highly customized orders, such as individualized color, button shape, and cover material. The manufacturing process could be varied from order to order. In addition, it should take processing failures, human interferences, order changes, and other uncertainties into consideration. As a result, the task planning module for large-scale problems with uncertainty shows great importance.

Compared to the ubiquitous robotic systems, the task planning in industrial domain is even more challenging due

to its larger planning space. For example, even in a small and medium factory, there could be dozens of machinery process and the planning space grows exponentially. Notice that human solves tasks in a hierarchical way, and fortunately most tasks in industrial domain have hierarchical structures. As a result, the large task can be divided into a task tree consisting of small subtasks, which can be solved more efficiently. Furthermore, some subtasks are reusable among up level tasks. In this study, a hierarchical task planning method is proposed to improve the planning efficiency. A study case of the smart assembly line is implemented as a demonstration platform for our methods.

2. Related Works

Many existing studies on smart factory focus on how to integrate RFID into the manufacturing system to collecting more data [9–11]. The manufacturing is smarter by tracking the processing information. We argue that it would achieve higher flexibility and intelligence if connecting not only the production but all the machinery processes. So different robotic devices could collaborate into different groups according to different tasks.

The ubiquitous robotic technology is widely studied these years. A number of frameworks have been proposed [12–14]. Recently, more efforts have been made on task level planning and learning technologies. The task planning methods for such multiagent systems could be categorized into centralized planning and decentralized planning. Decentralized planning methods are mainly applied to loosely coupled problems such as multi-UAV environmental monitoring [15] and cooperative mapping and localization [16]. As the individual machinery processes are highly coupled in manufacturing tasks, we prefer the centralized planning method.

The most commonly employed centralized techniques are based on automated planning in Artificial Intelligence. Ha et al. used SHOP2 planner to decompose services based on semantic knowledge [12]. Erdem et al. presented an application of answer set programming to housekeeping robotics [17]. Niemueller et al. approached the task planning problem by deploying a rule engine [18]. These planning methods cannot deal with dynamic situations with uncertainties as is the case in the real world. In response to this, some researchers have used probabilistic models in task planning problems. For example, Barbosa et al. used Partially Observable Markov Decision Processes (POMDP) to model the tasks with uncertainty [19]. Cirillo et al. implemented RTL plan for probabilistic domains [20]. Planning methods based on probabilistic models such as Markov Decision Process (MDP) model and POMDP model can handle nondeterministic problems but at significant cost. They suffer dimension explosion, which limits the size of the state space to impractical applications.

The researches of MDP planning methods for large problems mainly consist of two kinds, the state approximation and hierarchical planning. The former has considerable difficulty in applying to general purpose planner discussed in this study. So we focus on the hierarchical planning methods.

The efforts of achieving the hierarchical planning of MDP problems are divided into two parts: first, how to automatically generate the hierarchical structures [21, 22]; second, how to develop planning algorithms to solve subproblems introduced by the hierarchical structure [23, 24]. Sutton et al. [25] used options to temporally abstract knowledge based on Semi-Markov Decision Process (SMDP) theory. Parr [26] developed an approach to hierarchically structuring MDP policies called Hierarchies of Abstract Machines. Dietterich [23] developed another approach called the MAXQ Value Function Decomposition. These methods assume the hierarchy is predefined by human experts. For the automatic task decomposition problem, Hengst [27] proposed the HEXQ approach for the construction of a hierarchy of abstractions based on the change frequency of state variables. Jonsson [28] proposed the VISA approach for decomposing factored MDPs based on causal relations between variables. Kheradmandian and Rahmati [21] incorporated and represented the ability of data mining techniques in automatic discovering of structures and patterns. Most of these methods are based on statistic methods that try to learn the critical states as the subgoals. This learning process is time consuming and does not have any optimality guarantees. We followed the research of Hengst and Jonsson, who generate the task hierarchy depending on state variables. We improve Jonsson's work by abstracting hierarchical options instead of searching for exits. Consequently the optimality is improved from recursively optimal to hierarchically optimal.

3. System Architecture

In contrast to traditional manufacturing processes, the smart manufacturing offers the advantage of distributed networked machines to complete different tasks through collaboration. The framework for smart factory is designed as in Figure 1.

In the low level, the robotic devices are developed into components that they can “plug and play” in the system and be reused and reconfigured according to different manufacturing process. These components are the foundation of the system. As mentioned, robotic components are highly heterogeneous with respect to platforms such as operating system, programming language, and communication media. Middleware is thus employed to generalize the components into a uniform abstraction which enables dynamic communication and coordination between any two of the modules [29]. This also brings benefits to the modification of existing devices and the expansion of new ones.

In the middle level, a number of functionalities are developed in the internal cloud, such as the human-system interface, storage management, task planning, virtual manufacturing, and big data collection. The customer orders products through a human-system interface. The order includes customized requests, for instance, the favorite color and shape of the parts and whether the parts are being polished and so forth. These orders are sent to the task planning module, which also utilizing the information from the storage management module. The planner is the key part of the system's agility and intelligence. It turns customers' orders into subtask sequences, which can be directly carried out

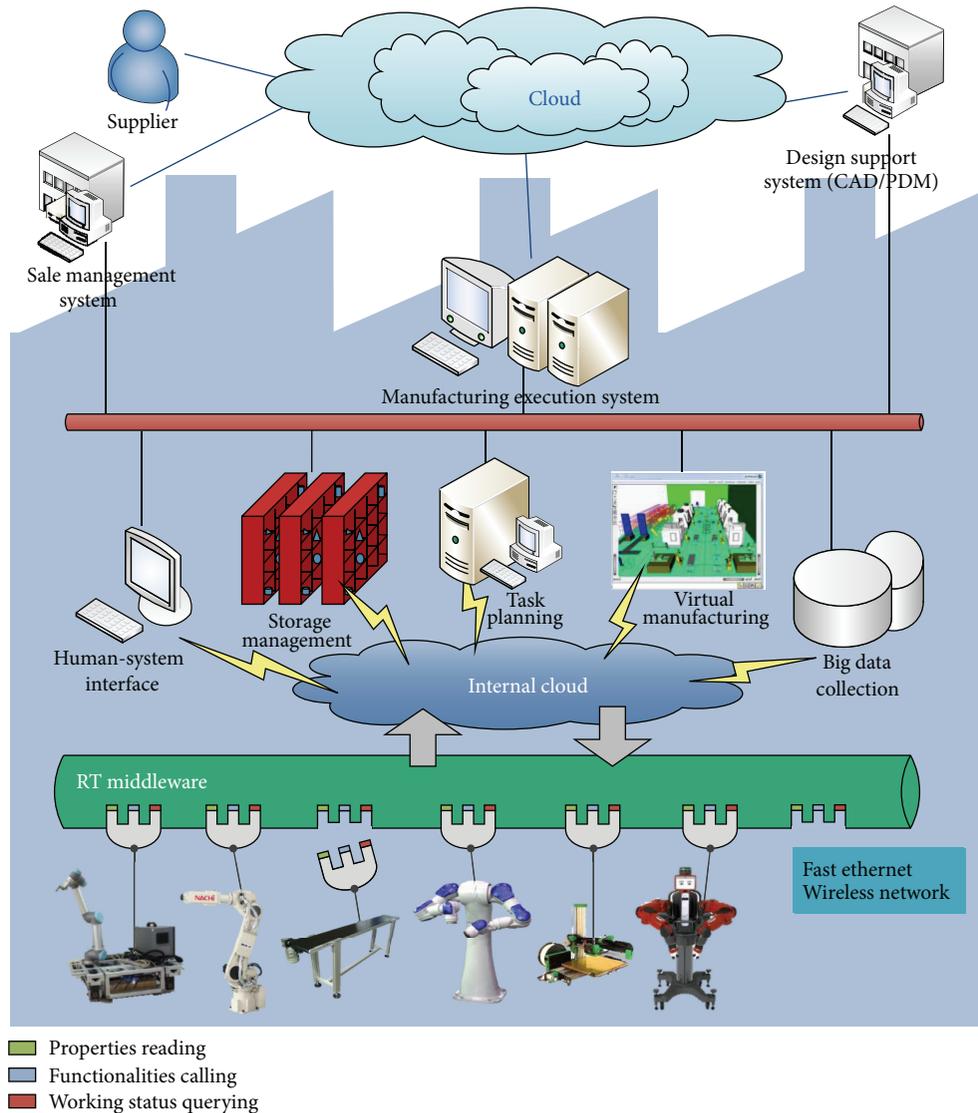


FIGURE 1: System architecture of the smart factory.

by corresponding robotic components. It is a general purpose planner based on Reduced Markov Decision Process (RMDP) model, which will be detailed later.

In the upper level, there are manufacturing execution system, sale management system, and design support system. These are all critical part of the industrial production process. This paper will not get into details of these big systems but mainly focus on the task planning module and the component based technology.

4. Component Based Machinery Process

Components use ports to communicate with each other and with high level controller. The ports are categorized into data ports and service ports [30]. The data port is responsible for the continuous exchange of data. Each component can have any number of data in-ports and out-ports. A data out-port sends the data to a corresponding in-port which

receives the data. The service port provides the command based communication. The component with a service port, offering a set of services, listens for requests for those services via a connector.

Each component has three service ports, namely, FuncGet, FuncSet, and ExeStatusGet. The service port is responsible for the interaction with the upper layer. FuncGet port reports to the service layer about the components' state. For example, the polishing robot reports the available polishing configuration; the Autonomous Intelligent Mobile Manipulator (AIMM) reports its states including its coordinates, whether the manipulator is empty, and battery level. FuncSet port provides the functionality invoking, such as setting the target position for the AIMM, starting polishing with certain configuration, and so forth. ExeStatusGet port returns the execution status, for example, whether or not the AIMM has reached its destination, or whether the polishing robot succeeds or fails in doing the task.

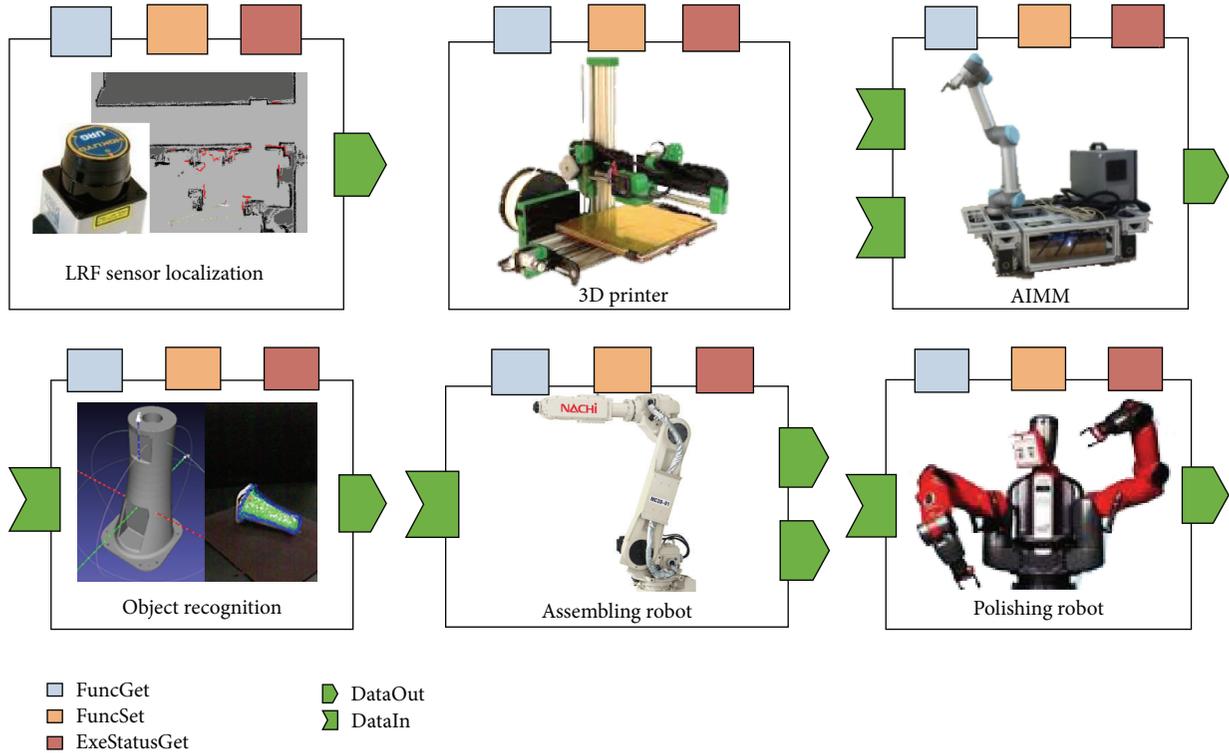


FIGURE 2: Robotic components in our system.

Each component may have any number of data ports for continuous data exchange between components. For instance, the localization information is transferred from the data out-port of laser component to the data in-port of the path planning component. Once two data ports are connected, those two components are able to perform real-time communication to accomplish the task collaboratively.

The individual robotic functions are also critical to the system's intelligence. The traditional industrial robots are like blind and deaf muscles repeating some predefined motions. In the smart factory, robotic components are capable of sensing the environment and making decisions in optimization of resources and time. Some of the robotic components in our system are shown in Figure 2.

There are five 3D printers with materials in different colors, one dual-arm robot for polishing, one assembling robot, and one AIMM. The AIMM is equipped with laser sensor for localization and obstacles avoiding. Other software components such as localization, path planning, and object recognition are also implemented.

4.1. Polishing Component with Auto Path Generation. Traditionally, the polishing path is taught by the expert engineers. This teaching process could be complex and tedious. In our smart factory, the polishing path is automatically generated from the CAD data (Figures 3(b) and 3(c)). Then, the robot follows this path by a motion planning algorithm with collision avoidance (Figure 3(d)). Besides, the polishing area is easy to specify with a user-friendly GUI as in Figure 3(a).

4.2. AIMM Component. AIMM is responsible for the transportation task that transports parts and work pieces between workstations and storages (Figure 4). Such transportation tasks contain physical separation larger than the workspace of the robot manipulator. This requires a lot of technologies such as object recognition, grasp point generating, motion planning, localization, and path planning. It uses RGB-D camera to do the object recognition and obstacle avoidance and uses laser sensor to do the localization.

4.3. Assembling Component. The assembling robot also has the sensing capability (Figure 5). It grasps the working parts by online detecting the location and orientation. The visual detection is based on template matching method and is able to recognize complex shape with localization error below 1 mm. We also employ a motion planning and motion controlling module for assembling and obstacle avoiding.

5. Hierarchical Task Planning

The challenges of task planning for smart factory domains are introduced by their large problem size and uncertainty. This study follows the techniques of automated planning derived from the AI field. Firstly, a task model called RMDP model is proposed. This model is designed for describing problems with large size and limited uncertainties such as smart factory. Secondly, the relations of variables are analyzed based on this model. The relations are depicted by the causal graph. Thirdly, according to the causal graph, the original

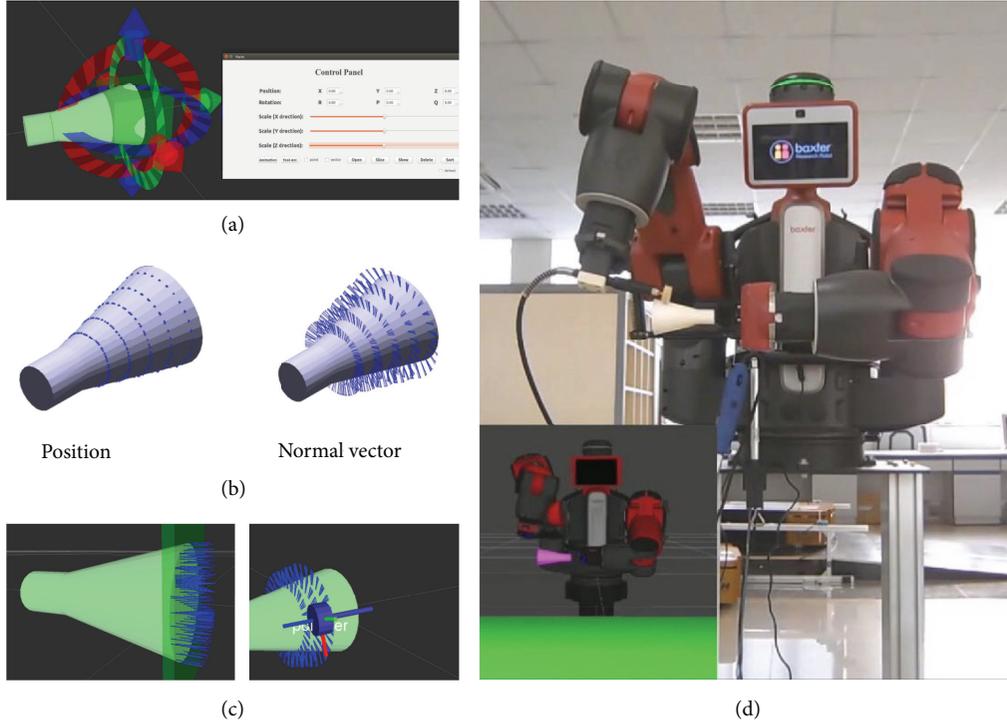


FIGURE 3: (a) Configuring the polishing area, (b) auto-generating the polishing path, (c) path generation and tool simulation, and (d) motion planning and polishing with dual-arm robot.

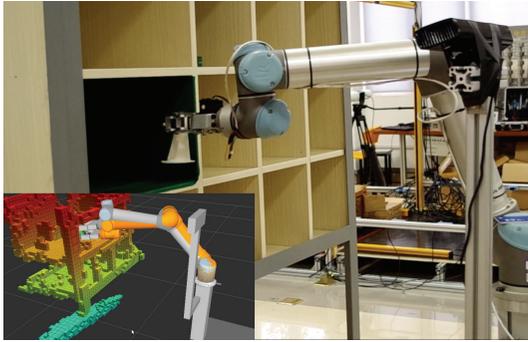


FIGURE 4: AIMM is picking up a working part from the warehouse.

actions are hierarchically abstracted into options, which induce smaller subproblems. At last, the subproblems and the original problem are solved based on Semi-MDP theory.

5.1. Task Modeling Based on Multivalued State Variables.

The task planning problem is modeled as a state transition system. Depending on different assumptions, various models are proposed. The two most commonly used models are the classical planning model and the MDP model. However, the classic model cannot deal with dynamic situations with uncertainties as is the case in the real world. The MDP model supports nondeterministic actions and dynamic situations, but it scales poorly to large problems.

We propose the RMDP model by making the following assumption in line with the manufacturing domains. It is

assumed that, after actions are executed by robotic components, the outcome could be among a few predictable states, which are the successful state and a few failed states. This assumption simplifies the MDP model by decreasing the branching factor of the state space. In addition, the model is designed based on multivalued state variables, which is more compact and natural compared to the propositional based models. This is important for the following variable analysis.

Definition 1. RMDP model is defined as a five-tuple $\Pi = (V, D, A, I, G)$:

- (i) $V = \{v_1, v_2, \dots, v_n\}$ is a finite set of state variables;
- (ii) $D = \{d_1, d_2, \dots, d_n\}$ is a finite set of variable domains, each $v_i \in V$ with a finite domain $d_i \in D$. V and D define the planning space S , where state $s \in S$ is represented as a vector $[x_1, x_2, \dots, x_n]$, where $x_i \in d_i$ is the value of variable v_i ;
- (iii) $A = \{a_1, a_2, \dots, a_m\}$ is a finite set of actions; each $a_i \in A$ is a triple (pc, ef, c) referring to the action's preconditions, effects, and cost, respectively. The preconditions of action a are defined with a list $[pc_1, pc_2, \dots, pc_j]$, where $pc_i = (v, x)$ denotes that the value of variable v should be x to satisfy the precondition. The effects of action a are defined with an effect list $[e_1, e_2, \dots, e_k]$, where $e_i = (p, v, x)$ denotes that the variable v will change its value to x with probability p , after the action's execution. Each action has a cost c , which acts like a reward function in MDP model;

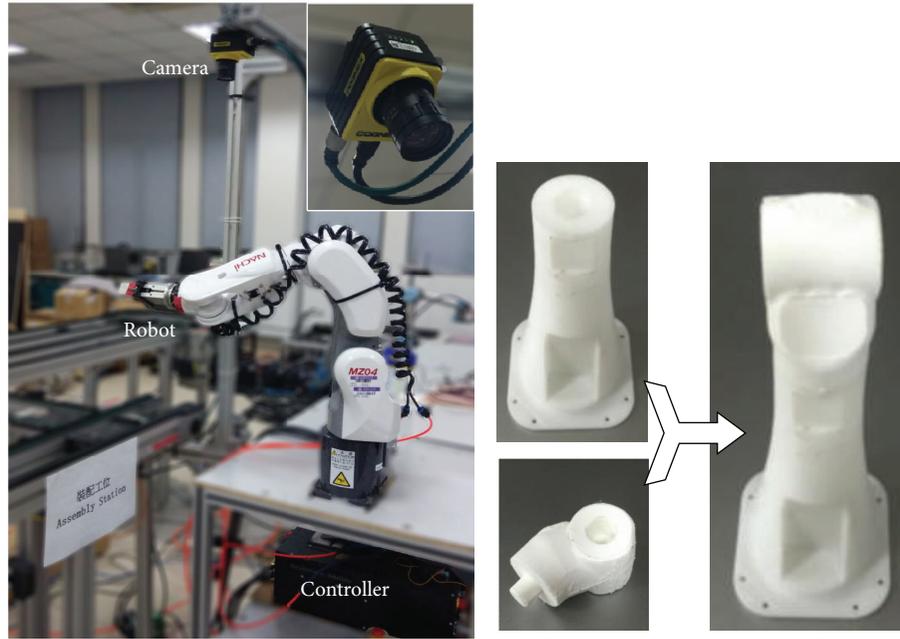


FIGURE 5: Assembling robot with visual detection.

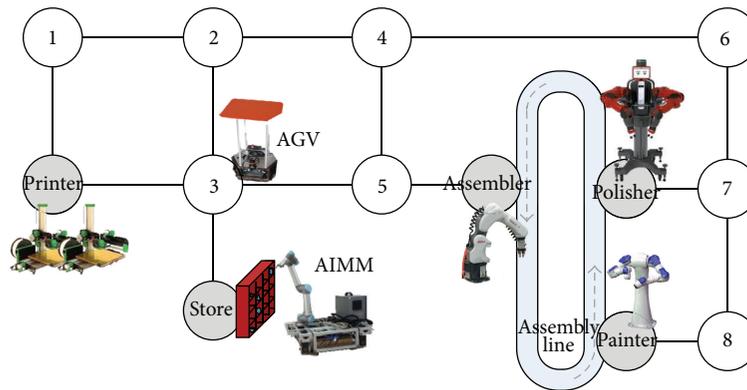


FIGURE 6: The smart factory planning domain.

- (iv) $I \in S$ denotes the initial state;
- (v) $G \subseteq S$ denotes the set of goal states.

The demonstrating task in this study is shown in Figure 6. As described in Section 4, the smart factory in our study case includes 3D printers, polishing robot, assembling robot, and AIMM. This task is designed according to the physical system in our laboratory, which will be detailed in the next section. One workstation of painting robot and one Automatic Guided Vehicle (AGV) are added to increase the complexity. Detailed task description is shown in Tables 1 and 2.

These actions in Table 2 are grounding actions. There are too many grounding actions to be defined by hand. In practice, the actions are defined in lifted manner. For example, there are 28 “Move” actions in total, 14 for AIMM and 14 for AGV. These 28 actions are presented by one lifted

“Move” action: $\text{Move}(\text{Robot}, \text{Location}, \text{Location})$, where the action is parameterized with variable types “Robot” and “Location”. The lifted actions are compiled to the grounding actions in a preprocessing stage.

5.2. Variable Dependency Analysis. Notice that there are dependencies between different variables. For example, according to the “pickup” action, the change of value of “part1_loc” is dependent on the value of “AIMM_loc”. According to “polish” action, the change of value of “polished_part1” is dependent on the value of “part1_loc”. We depict these dependencies by a causal graph, following the work of Helmert [31] and Jonsson [28].

Definition 2. The *causal graph* of Π is a directed graph $\text{CG}(\Pi)$ with vertices V_{cg} and an arc (u, v) whenever there exists an action $a \in A$ so that either (i) there exists $a \in A$

TABLE 1: Variable V and domain D in the smart factory task.

	Variable	Domain
1	AIMM_loc	loc1, loc2, . . . , loc8, printer_spot, store_spot, . . . , painter_spot
2	AGV_loc	loc1, loc2, . . . , loc8, printer_spot, store_spot, . . . , painter_spot
3	AIMM_hand_empty	true, false
4	part1_loc	AIMM, AGV, printer_spot,
5	part2_loc	store_spot, assembler_spot,
6	part3_loc	polisher_spot, painter_spot
7	part1_color	
8	part2_color	red, blue, black, white
9	part3_color	
10	part1_polished	
11	part2_polished	true, false
12	part3_polished	
13	assembled_p1_p2	true, false
14	assembled_p1_p3	true, false

so that $u \in a(\text{pc})$ and $v \in a(\text{ef})$ are both defined, or (ii) there exists $a \in A$ so that $u \in a(\text{ef})$ and $v \in a(\text{ef})$ are both defined.

The causal graph is independent of the initial state and goal state. As a result, it can be calculated offline. The causal graph of the example task is shown in Figure 7(a), where each circle represents a variable associated with Table 1. If we add one more AGV and three more parts to the factory (the variables are listed in Table 3) and the causal graph is shown in Figure 7(b), the associated algorithm is as Algorithm 1 shows.

The causal graph reflects the structure of the planning problem. The overall task is decomposed according to the causal relations of variables.

If the causal graph is acyclic, the decomposition is very intuitive. The task can be decomposed into the same structure as the causal graph. Then the task could be solved hierarchically. However, most tasks have a cyclic causal graph, such as in Figure 7. In these cases, we find out all the strongly connected components (SCCs) in the causal graph and combine the variables in each SCC. As a result, the task of Figure 7(a) is decomposed as in Figure 8. Because the v_7-v_{14} in the high layer all have very small domain size, we combine them in one layer. Because these SCCs do not have interdependencies, the combination will not change the number of subtasks, but only for structural simplicity.

Given the task structure, the task is able to be solved hierarchically from low layer to the high layer. This process is divided into two phases called iteratively. These two phases are abstracting options and solving Semi-MDP, which will be detailed in the following two sections.

5.3. Hierarchical Option Causal Abstraction. Given the hierarchical structure, one of the key problems is to find the

```

(1) CG = (Vcg, Ecg), where Vcg = V, Ecg = ∅
(2) for each a ∈ A:
(3)   for each u ∈ a(pc):
(4)     for each v ∈ a(ef):
(5)       if u ≠ v then add e = (u, v) to Ecg
(6)     endfor
(7)   endfor
(8)   for each u ∈ a(ef):
(9)     for each v ∈ a(ef):
(10)      if u ≠ v then add e = (u, v) to Ecg
(11)    endfor
(12)  endfor
(13) endfor

```

ALGORITHM 1: Calculate causal graph.

reusable subtasks. So the original problem could be decomposed into the combination of these subtasks. We propose an algorithm called Hierarchical Option Causal Abstraction (HOCA). In general, the actions are abstracted into hierarchical options based on causal relations. Each option induces a subtask that could be solved offline. The planning efficiency is remarkably improved using these options instead of the primitive actions. Options are used for the generalization of temporally extended primitive actions by Sutton et al. [25]. In their work, options are designed by human expert. This term is modified in this study in order to enable the automatic abstraction.

Definition 3. An *option* is a four-tuple $o = (\text{fa}, \text{pc}, \beta, \pi)$, where

- (i) fa denotes the father option of o ;
- (ii) pc is the preconditions of option o , similar with the action's definition;
- (iii) β is the set of goals of option o , each "var-value" pair $(v, x) \in \beta$ requiring that the goal value of variable v is x ;
- (iv) $\pi : s \mapsto o$ is the policy for this option, which is calculated by the method detailed in the next section.

Definition 4. The *option hierarchy* derived from action a is represented as $H^a = \{o_0^a, o_1^a, \dots, o_m^a\}$, where $o_0^a = ([], a(\text{pc}), a(\text{ef}), [])$ is directly converted from action a , and $\forall 1 < k \leq m$, there is $o_k^a(\text{fa}) = o_{k-1}^a$.

To automatically generate options, we define the preconditions of the action as a subgoal based on the causal relations of the variables. This subgoal is solved within an abstracted state space, which is much smaller than the original space. Through this process, a primitive action is abstracted into an option, which could be further abstracted into higher level options. These options derived from action a form an option hierarchy as Definition 4.

Practically, in each layer k , we further define the abstract option (ABO) and active option (ACO). Each ABO derives a

TABLE 2: Some of the actions in the smart factory task.

Action	Precondition	Effect	Probability
Move_AIMM_loc1_loc2	AIMM_loc = loc1	{AIMM_loc = loc2}	0.85
		{}	0.1
		{AIMM_loc = printer_spot}	0.05
Move_AGV_loc3_loc5	AIMM_loc = loc3	{AIMM_loc = loc5}	0.85
		{}	0.05
		{AIMM_loc = loc2}	0.05
		{AIMM_loc = store_spot}	0.05
Pickup_AIMM_store_part1	AIMM_loc = store_spot part1_loc = store_spot AIMM_hand_empty = true	{part1_loc = AIMM, AIMM_hand_empty = false}	0.9
		{}	0.1
Putdown_AIMMPainter_part3	AIMM_loc = painter_spot part3_loc = AIMM AIMM_hand_empty = false	{part3_loc = painter_spot, AIMM_hand_empty = true}	1.0
Putdown_AIMM_printer_part2_AGV	AIMM_loc = printer_spot AGV_loc = printer_spot part2_loc = AIMM AIMM_hand_empty = false	{part2_loc = AGV, AIMM_hand_empty = true}	0.9
		{part2_loc = printer_spot, AIMM_hand_empty = true}	0.1
Polish_part1	part1_loc = polish_spot part1_polished = false	{part1_polished = true, part1_color = white}	0.8
		{}	0.2
Paint_part1_red	part1_loc = paint_spot	{part1_color = red}	0.9
Assemble_p1_p2	part1_loc = assemble_spot part2_loc = assemble_spot assembled_p1_p2 = false	{assembled_p1_p3 = true}	0.95
		{}	0.05

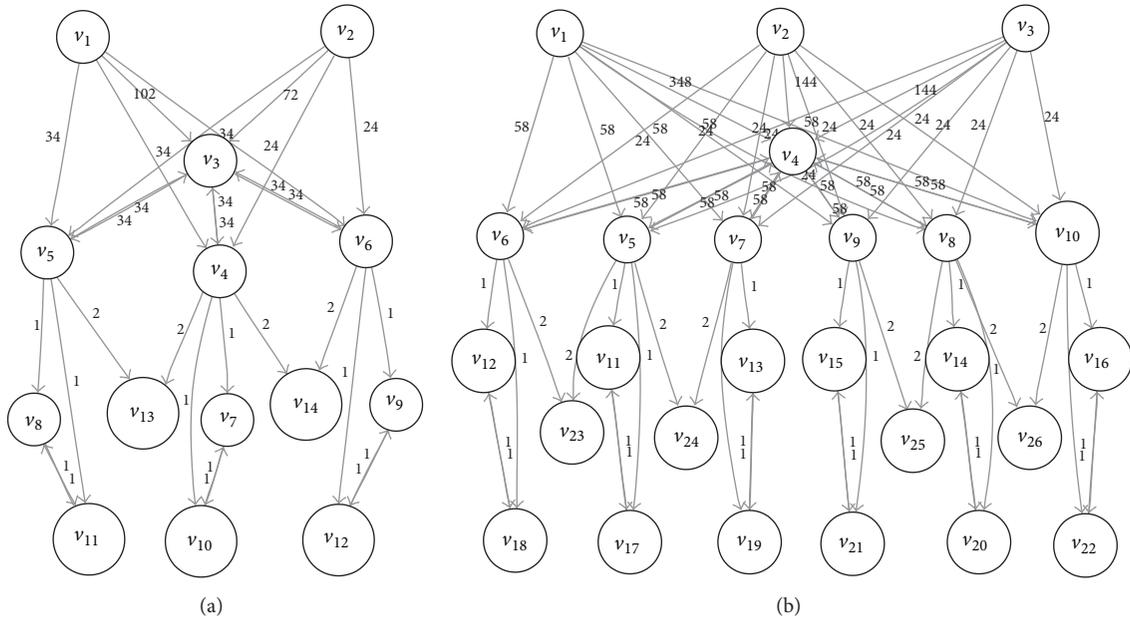


FIGURE 7: Causal graphs of smart factory tasks.

hierarchical option in layer k , while ACO is used for solving the Semi-MDP, which will be detailed in the next section. Assume the hierarchical structure is $Z = \{V^0, V^1, \dots, V^L\}$; V^k represents the set of variables in the k th layer. Define

$V^{k-} = \bigcup_{i=1}^k V^i$ as the union of the variables in layer lower than or equal to k . Define $V^{k+} = \bigcup_{i=k+1}^L V^i$ as the union of the variables in layer higher than k . Define V^a as the set of variables appearing in action a 's preconditions and effects.

TABLE 3: Variables of the smart factory task in Figure 7(b).

Tag	Variable
v_1	AIMM_loc
v_2	AGV1_loc
v_3	AGV2_loc
v_4	AIMM_hand_empty
v_5-v_{10}	part1_loc - part6_loc
$v_{11}-v_{16}$	part1_color - part6_color
$v_{17}-v_{22}$	part1_polished - part6_polished
v_{23}	assembled_p1_p2
v_{24}	assembled_p1_p3
v_{25}	assembled_p4_p5
v_{26}	assembled_p4_p6

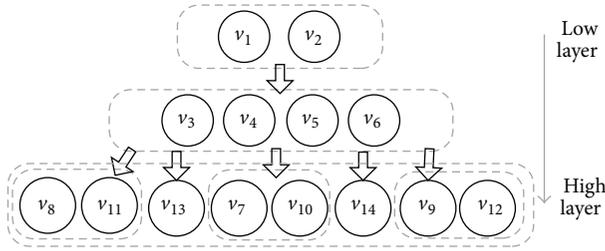


FIGURE 8: Combine the variables in strongly connected components.

Definition 5. The hierarchical option o^a at layer k is an *active option* (ACO) if and only if $V^a \subseteq V^{k-}$. The set of ACOs in layer k is denoted by \tilde{O}^k .

Definition 6. The hierarchical option o^a at layer k is an *abstract option* (ABO) if and only if $o^a(\text{pc}) \cap V^k \neq \emptyset$ and $V^a \cap V^{k+} \neq \emptyset$. The set of ABOs in layer k is denoted by \tilde{O}^k .

In layer k , the ACOs are options that all the associated variables are within the k th and above layer. So the ACO in layer k is fully abstracted. It can be used for solving the Semi-MDPs. The ABOs of layer k are options that satisfy two conditions: firstly existing variable both in the option's precondition and in layer k ; secondly existing variable in higher layer than k . Following the definitions, the hierarchical option in the k th layer is derived as Algorithm 2.

For example, in the low layer of the smart factory task, all the “move” actions are ACOs. The “pickup” and “putdown” actions satisfy the ABO conditions. The action “Pickup_AIMM_store_part1” is abstracted to option, which induces “AIMM_loc = store_spot” as a subgoal. Similarly in layer two, all the “pickup” and “putdown” options become ACOs. The actions “polish”, “paint”, and “assemble” are ABOs abstracted into options.

The options of one task are also independent of the initial and goal states, but they only depend on the task definitions of variable, domain, and actions. As a result, they can be calculated offline.

```

(1) for each  $o_i^a \in O^k$ 
(2)   if  $V^a \subseteq V^{k-}$  // this is an ACO
(3)     Add  $o_i^a$  to  $\tilde{O}^k$ 
(4)   else if  $o_i^a(\text{pc}) \cap V^k \neq \emptyset$  // this is an ABO
(5)     Add  $o_i^a$  to  $\tilde{O}^k$ 
(6)      $V_i^{a,k} = o_i^a(\text{pc}) \cap V^k$ 
(7)      $o_{i+1}^a = (o_i^a, o_i^a(\text{pc}) - V_i^{a,k}, \beta_i^{a,k}, \pi)$ 
(8)     Add  $o_{i+1}^a$  to  $O^{k+1}$ 
(9)   else // neither ACO nor ABO
(10)    Add  $o_i^a$  to  $O^{k+1}$ 
(11)  endif
(12) end for

```

ALGORITHM 2: Abstract options in the k th layer.

5.4. Solve the Hierarchical Semi-MDPs. A Semi-Markov Decision Process (SMDP) is a MDP model with temporally extended actions [32]. Efforts have been done to extend planning algorithms from MDP to SMDP problems [25]. As described above, the options abstracted in this study are also temporally extended. The subtasks relying on these options are consequently SMDPs.

Definition 7. The SMDP problem is defined as four-tuple $\Sigma = (V, D, O, \beta)$ where V is the variable set; D is the domain set; O is the option set; β is the goal for this problem.

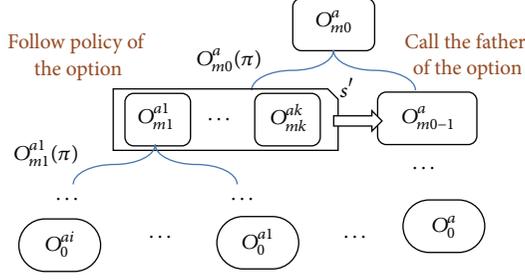
The solution of a SMDP problem is a policy π , mapping from states to options. To calculate the policy there are a bunch of algorithms extended from MDP problems, such as Dynamic Programming based on Bellman equation. For any state $s \in S$, the value function of policy π is

$$\begin{aligned}
VF^\pi(s) &= E \{r_t + r_{t+1} + \dots + r_{t+k} + VF^\pi(s_{t+k}) \mid \varepsilon(\pi(s), s, t)\} \\
&= c^{\pi(s)}(s) + \sum_{s' \in S} p^{\pi(s)}(s' \mid s) VF^\pi(s'),
\end{aligned} \tag{1}$$

where $\varepsilon(\pi(s), s, t)$ denotes the event of executing $\pi(s)$ in state s at time t and $t+k$ is the random time at which $\pi(s)$ terminates. $c^o(s)$ and $p^o(s' \mid s)$ denote the cost and transition probability of option o . They compose the option's model.

The optimal value function is the one with maximum value

$$\begin{aligned}
VF^*(s) &= \max_{\pi} VF^\pi(s) = \max_{\pi} E \{r_t + r_{t+1} + \dots + r_{t+k} \\
&+ VF^\pi(s_{t+k}) \mid \varepsilon(\pi(s), s, t)\} = \max_{\pi} \left[c^{\pi(s)}(s) \right. \\
&+ \left. \sum_{s' \in S} p^{\pi(s)}(s' \mid s) VF^\pi(s') \right].
\end{aligned} \tag{2}$$

FIGURE 9: Execution of option o forms an execution tree.

The optimal policy is the one that maximizes the value function

$$\begin{aligned} \pi^* &= \arg \max_{\pi} VF^{\pi}(s) \\ &= \arg \max_{\pi} \left[c^{\pi(s)}(s) + \sum_{s' \in S} p^{\pi(s)}(s' | s) VF^{\pi}(s') \right]. \end{aligned} \quad (3)$$

To calculate the policy, one important issue is how to get the models of the options in the option set. Since option o is abstracted from a hierarchy, the outcome state s' and the option cost $c^o(s)$ are random variables. According to Sutton et al. [25], they proposed a multitime model

$$p^o(s' | s) = \sum_{k=1}^{\infty} p(s', k) \gamma^k, \quad (4)$$

where $p(s', k)$ is the probability that the option terminates in s' after k steps and γ is a discount factor. On the other hand, the cost of o is a function of the state s :

$$c^o(s) = E \{ c_{t+1} + c_{t+2} + \dots + c_{t+k} | \varepsilon(o, s, t) \}, \quad (5)$$

where $\varepsilon(o, s, t)$ denotes the event of o being executed in state s at time t and $t+k$ is the random time at which o terminates.

The subtask induced by abstract option o in layer k is denoted by $\Sigma_o^k = (V^{k-}, D^{k-}, \widehat{O}^k, o(\beta))$, where \widehat{O}^k is the ACO set in this layer and $o(\beta)$ is the subgoal of option o . The variables and domains are all subset of the original problem. This makes the state space smaller.

According to the option hierarchy in Definition 4, the execution of o forms an execution tree. It is a recursive process as Figure 9 shows. The execution consists of two phases. It firstly follows the policy of the option and secondly calls the father option. This process is done recursively until all options reach down to the primitive actions, which lie on the leaf nodes of the execution tree. For the leaf nodes, the costs are equal to the cost of the primitive actions. In practice, the trees usually are not very deep, two or three layers in our example.

As a result, let o' be the father of o ; the cost of option o is represented by the following recursive formula:

$$\begin{aligned} c^o(s) &= VF^{o(\pi)}(s) + E \{ c^{o'(\pi)}(s') \} \\ &= VF^{o(\pi)}(s) + \sum_{s'} p^{o'(\pi)}(s' | s) c^{o'(\pi)}(s'), \end{aligned} \quad (6)$$

where s' is a random variable that denotes the state at which it terminates when satisfying $o(\beta)$; and $p^{o'(\pi)}(s' | s)$ is probability that the state terminates at s' following the policy $o'(\pi)$.

However in practice, the above model is difficult to compute. One solution is to employ model-free algorithms such as Temporal Difference (TD) and Q-learning. Instead of calculating, it just observes the outcome state and cost and updates the value function with small steps. The TD form of the updating rule is

$$VF^{\pi}(s) = (1 - \alpha) VF^{\pi}(s) + \alpha [c^{\pi(s)}(s) + VF^{\pi}(s')], \quad (7)$$

where

$$c^o(s) = VF^{o(\pi)}(s) + c^{o'(\pi)}(s''), \quad (8)$$

where α is the step size, $o' = o(\text{fa})$, s' is the outcome state after the hierarchical tree of o has completed, and s'' is the outcome state after policy $o(\pi)$ has terminated.

Similarly, the Q-learning version of the updating rule is

$$VF^{\pi}(s) = (1 - \alpha) VF^{\pi}(s) + \alpha \max_{o'} [c^o(s) + VF^{\pi}(s')]. \quad (9)$$

The Q-learning algorithm for SMDP problem is as in Algorithm 3.

After the SMDP for option o has been calculated, this option has been abstracted in the current layer. It may become an ACO of the higher layer or be abstracted again in higher layer. In summary, this abstracting process and calculating SMDP process are called iteratively from low layer to high layer as Algorithm 4 shows.

Although the planning efficiency is remarkably improved, the policy achieved by HOCA algorithm is not a global optimal policy. As many hierarchical planners do [26, 27], HOCA achieves a hierarchical optimal policy. This means that the solution is optimal given the constraints of the hierarchy. It is often a tradeoff. If we want a policy that is closer to the global optimal one, we should use simpler hierarchy with options of lower abstraction level. But if the speed is more wanted, it needs more abstract options.

6. Experiments and Results

A smart factory was implemented based on the ubiquitous robotic technology. A demonstration video could be found in the Supplementary Material available online at <http://dx.doi.org/10.1155/2016/6018686> as well as in this link: <https://youtu.be/MVO4yGF0GwY>. It took in customers' individualized order and arranged the producing process accordingly. Figure 10 shows one execution of the smart factory task. First, the customer made an order through the user interface. The order was then sent to the task planning module, which calculated the action sequence hierarchically. 3D printers started to make parts with specific color and shape as Figure 10(b). Meanwhile, the AIMM transported the part from the storage to the polishing station as shown in Figures 10(c)–10(e). After that, the dual-arm polishing robot polished the part according to customer's configuration as

```

(1) problem  $\Sigma = (V, D, O, \beta)$ , initialize Value Function VF to zero
(2) while error > threshold
(3)    $s = \text{random state}$ 
(4)   while  $s$  not satisfy  $\beta$ 
(5)     for each  $o \in O$  that is applicable to  $s$ 
(6)       apply  $o$  to  $s$ , observe the outcome state  $s'$  and cost  $c$ 
(7)     endfor
(8)      $\text{VF}(s) = (1 - \alpha)\text{VF}^\pi(s) + \alpha \max_{o \in O} [c^o(s) + \text{VF}(s')]$ 
(9)      $\text{error} = \max(\text{error}, \Delta \text{VF}(s))$ 
(10)     $s = s'$ 
(11)   end while
(12) end while
(13)  $\pi(s) = \arg \max_{o \in O} [c^o(s) + \text{VF}^\pi(s')]$ 

```

ALGORITHM 3: Q-learning for SMDP.

```

(1) RMDP  $\Pi = (V, D, A, I, G)$ , compute the causal graph CG using Algorithm 1
(2) combine the SCC of CG, obtain the task hierarchy  $Z = \{V^0, V^1, \dots, V^L\}$ 
(3) convert the primitive actions to options, add them to  $O^0$ 
(4) for each layer  $0 \leq k \leq L$ 
(5)   calculate ACO  $\tilde{O}^k$  set and ABO set  $\tilde{O}^k$  using Algorithm 2
(6)   for each  $o \in \tilde{O}^k$ 
(7)     induce a sub-problem  $\Sigma_o^k = (V^{k-}, D^{k-}, \tilde{O}^k, o(\beta))$ 
(8)     solve  $\Sigma_o^k$  using Algorithm 3, obtain the policy  $o(\pi)$ 
(9)   endfor
(10) endfor
(11) divide  $G$  from high layer to low layer  $G = \{g_1, g_2, \dots, g_l\}$ 
(12) for each  $g_i \in G$ , solve  $g_i$  using Algorithm 3, obtain the policy  $o_i(\pi)$ 
(13) the policy is a sequence  $o_1(\pi), o_2(\pi), \dots, o_l(\pi)$ 

```

ALGORITHM 4: Hierarchical option causal abstraction.

in Figure 10(f). At last, the parts were transported to the assembling spot after which the product was successfully processed as in Figures 10(g) and 10(h).

With the component based framework, every two of the robotic devices are ready to cooperate with each other. For instance, the continuous localization data is transferred from the laser sensor to the AIMM's path planning module through data port. And the AIMM can pass the working part directly onto the polishing robot, after calling on its service port. Further, this modular framework also facilitates the easy expansion of new devices and painless modification of the existing devices.

The hierarchical task planning method decomposes the original big problem into a hierarchy of small problems. The problem of smart factory in Figure 6 has $2.4e + 8$ states. It takes about 1700 episodes for the flat Q-learning algorithm to converge to the optimal value. HOCA algorithm is firstly run offline to compute the causal graph and hierarchical structure. Then according to the three-layer structure as Figure 8 shows, two layers of option abstraction are applied. Using HOCA algorithm described previously with one layer of abstraction, namely, just using the abstracted options

“pickup” and “putdown”, it converges in less than 1200 episodes. In this problem, it still converges to the optimal value using one layer of abstraction (Figure 11). When HOCA is run with two layers abstraction, in which case all the primitive actions including “polish”, “paint”, and “assemble” are abstracted, the convergence time is significantly reduced. This is because the domain sizes of the high layer variables are relatively small in this example. Using those abstracted options, it only takes 4 or 5 steps to reach the goal (Figure 11). However, it does not reach global optimal policy using two layer abstractions. In the optimal solution, three working parts are picked and placed on the AGV in sequence and transferred together. This strategy is unavailable when using highly abstracted options, in which the parts are transferred one by one. We plan to study this problem in the next step. It may achieve the global optimum if we flat the policy and refine it with low layer options.

The improvements are even larger on more complex tasks. In the smart factory task in Figure 7(b), the size of the state space is about $4.8e + 15$. It is generally unsolvable for flat algorithms. But for HOCA, it will not be much harder than the previous task. The additional AGV and three working

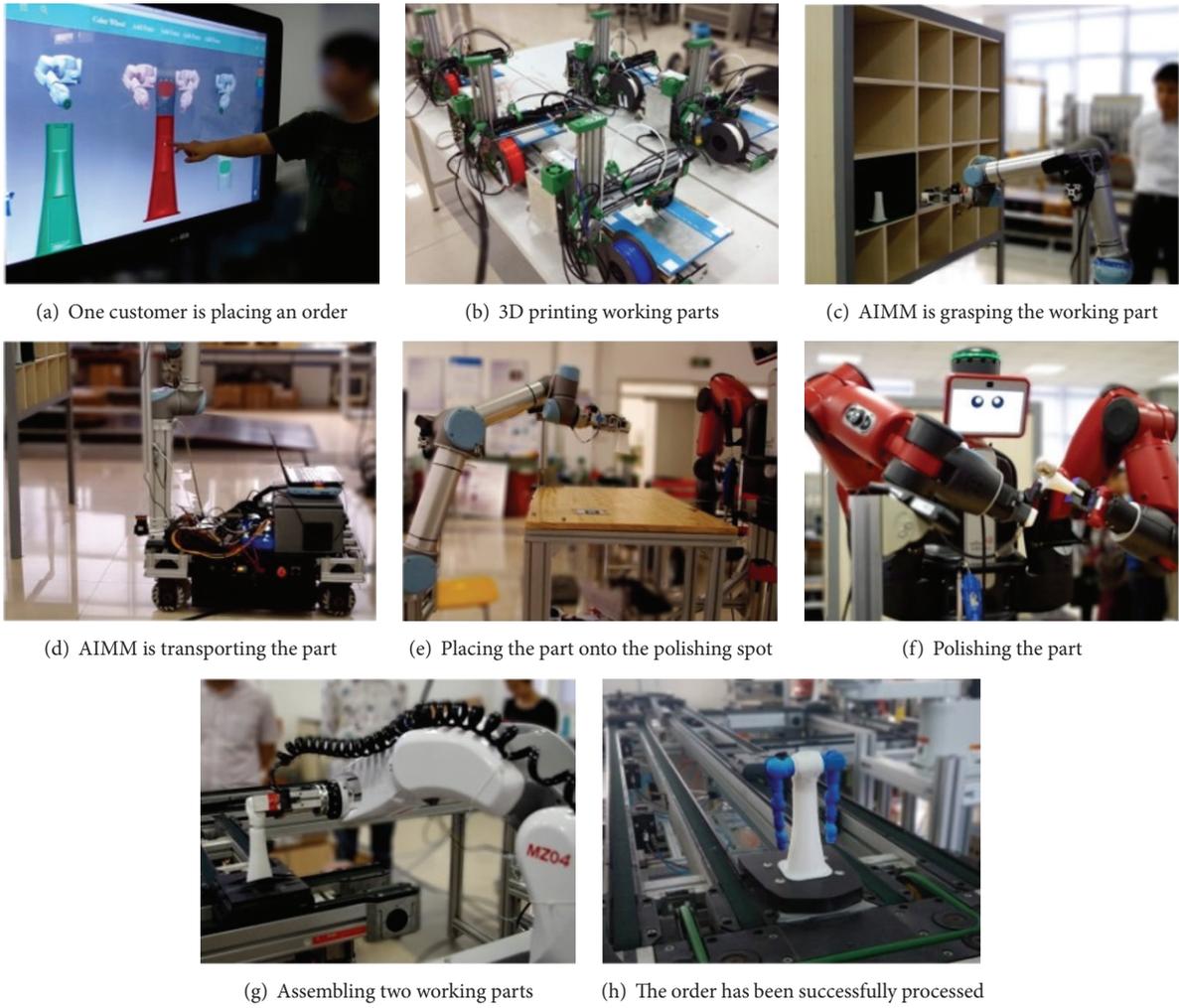


FIGURE 10: Manufacturing process in smart factory.

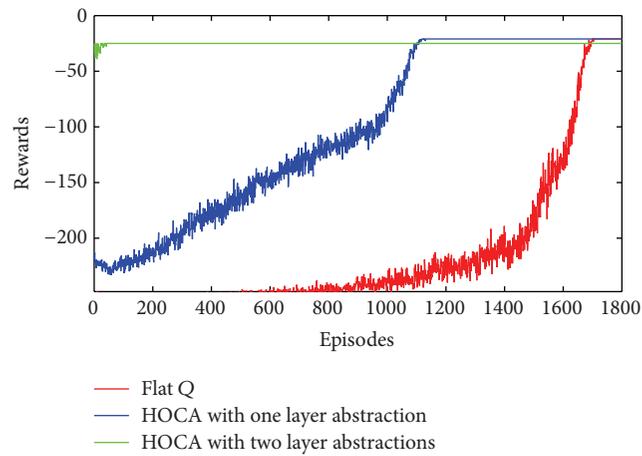


FIGURE 11: Performance of HOCA under different abstractions against a flat Q-learning.

parts in this task have same domain and subtasks with the ones in the previous task. Those subtasks are reused and bring no extra computational loads for offline option abstractions. The online iteration is still sustainable.

7. Conclusions

Given the increasing popularity of smart manufacturing as a solution offering better autonomy, this paper discussed the similarity of the smart manufacturing with the ubiquitous robotic system. A component based framework has been proposed and proved to be applicable for industrial domain. Further, since the manufacturing problems are often in large size with uncertainties, a hierarchical task planning method called HOCA based on RMDP model has been developed. This method decomposes the original big problem into a hierarchy of small problems by automatically abstracting primitive actions to a hierarchy of options according to variable dependencies.

A smart factory was implemented as the testing bed of our framework and algorithms. The individualized orders were processed by the system that arranged the producing process accordingly. The results showed that the framework facilitates the communication and cooperation between the robotic components. Further the hierarchical planning method has remarkably reduced the problem size and makes large problem tractable. However, the planner can only obtain hierarchical optimal results. This is the obvious defects and should be improved in the future.

It is our view that the results obtained from this work represent a substantial improvement. This method is not restricted to the particular domain discussed in this paper. As such, these results could also be beneficial to the researchers attempting to design smart manufacturing systems for other complex tasks in large-scale environment.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research has been supported in part by the National High Technology Research and Development Program of China under Grant 61273331. The authors gratefully acknowledge YASKAWA Electric Corporation for supporting the collaborative research funds on the project “Research and Development of Key Technologies for Smart Factory.”

References

- [1] J. Lee, H.-A. Kao, and S. Yang, “Service innovation and smart analytics for industry 4.0 and big data environment,” *Procedia CIRP*, vol. 16, pp. 3–8, 2014.
- [2] J. Davis, T. Edgar, J. Porter, J. Bernaden, and M. Sarli, “Smart manufacturing, manufacturing intelligence and demand-dynamic performance,” *Computers & Chemical Engineering*, vol. 47, pp. 145–156, 2012.
- [3] A. Chibani, Y. Amirat, S. Mohammed, E. Matson, N. Hagita, and M. Barreto, “Ubiquitous robotics: recent challenges and future trends,” *Robotics and Autonomous Systems*, vol. 61, no. 11, pp. 1162–1172, 2013.
- [4] J. Rashid, M. Broxvall, and A. Saffiotti, “A middleware to integrate robots, simple devices and everyday objects into an ambient ecology,” *Pervasive and Mobile Computing*, vol. 8, no. 4, pp. 522–541, 2012.
- [5] R. B. Rusu, B. Gerkey, and M. Beetz, “Robots in the kitchen: exploiting ubiquitous sensing and actuation,” *Robotics and Autonomous Systems*, vol. 56, no. 10, pp. 844–856, 2008.
- [6] W. Wang, Q. Cao, X. Zhu, and S. Liang, “A framework for intelligent service environments based on middleware and general purpose task planner,” in *Proceedings of the 11th International Conference on Intelligent Environments (IE '15)*, pp. 184–187, Prague, Czech Republic, July 2015.
- [7] R. Lundh, L. Karlsson, and A. Saffiotti, “Autonomous functional configuration of a network robot system,” *Robotics and Autonomous Systems*, vol. 56, no. 10, pp. 819–830, 2008.
- [8] G. Amato, M. Broxvall, S. Chessa et al., “Robotic ubiquitous cognitive network,” in *Ambient Intelligence—Software and Applications*, pp. 191–195, Springer, Berlin, Germany, 2012.
- [9] G. Q. Huang, Y. F. Zhang, and P. Y. Jiang, “RFID-based wireless manufacturing for real-time management of job shop WIP inventories,” *International Journal of Advanced Manufacturing Technology*, vol. 36, no. 7-8, pp. 752–764, 2008.
- [10] E. W. T. Ngai, D. C. K. Chau, J. K. L. Poon, A. Y. M. Chan, B. C. M. Chan, and W. W. S. Wu, “Implementing an RFID-based manufacturing process management system: lessons learned and success factors,” *Journal of Engineering and Technology Management*, vol. 29, no. 1, pp. 112–130, 2012.
- [11] R. Y. Zhong, Q. Y. Dai, T. Qu, G. J. Hu, and G. Q. Huang, “RFID-enabled real-time manufacturing execution system for mass-customization production,” *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 2, pp. 283–292, 2013.
- [12] Y.-G. Ha, J.-C. Sohn, Y.-J. Cho, and H. Yoon, “A robotic service framework supporting automated integration of ubiquitous sensors and devices,” *Information Sciences*, vol. 177, no. 3, pp. 657–679, 2007.
- [13] A. Sanfeliu and J. Andrade-Cetto, “Ubiquitous networking robotics in urban settings,” in *Proceedings of the Workshop on Network Robot Systems. Toward Intelligent Robotic Systems Integrated with Environments*, pp. 10–13, 2006.
- [14] A. Saffiotti and M. Broxvall, “PEIS ecologies: ambient intelligence meets autonomous robotics,” in *Proceedings of the Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies (sOcEUSAI '05)*, pp. 277–281, October 2005.
- [15] J. Capitan, M. T. J. Spaan, L. Merino, and A. Ollero, “Decentralized multi-robot cooperation with auctioned POMDPs,” *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 650–671, 2013.
- [16] B. K. Kim, N. Tomokuni, K. Ohara, T. Tanikawa, K. Ohba, and S. Hirai, “Ubiquitous localization and mapping for robots with ambient intelligence,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06)*, pp. 4809–4814, IEEE, Beijing, China, October 2006.
- [17] E. Erdem, E. Aker, and V. Patoglu, “Answer set programming for collaborative housekeeping robotics: representation, reasoning, and execution,” *Intelligent Service Robotics*, vol. 5, no. 4, pp. 275–291, 2012.

- [18] T. Niemueller, G. Lakemeyer, and A. Ferrein, "Incremental task-level reasoning in a competitive factory automation scenario," in *Proceedings of the AAAI Spring Symposium: Designing Intelligent Robots*, pp. 43–48, March 2013.
- [19] M. Barbosa, A. Bernardino, D. Figueira et al., "ISROBOTNET: a testbed for sensor and robot network systems," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '09)*, pp. 2827–2833, IEEE, St. Louis, Mo, USA, October 2009.
- [20] M. Cirillo, L. Karlsson, and A. Saffiotti, "A human-aware robot task planner," in *Proceedings of the 19th International Conference on Automated Planning and Scheduling*, pp. 58–65, AAAI Press, Thessaloniki, Greece, September 2009.
- [21] G. Kheradmandian and M. Rahmati, "Automatic abstraction in reinforcement learning using data mining techniques," *Robotics and Autonomous Systems*, vol. 57, no. 11, pp. 1119–1128, 2009.
- [22] D. Xiao, Y.-T. Li, and C. Shi, "Autonomic discovery of subgoals in hierarchical reinforcement learning," *The Journal of China Universities of Posts and Telecommunications*, vol. 21, no. 5, pp. 94–104, 2014.
- [23] T. G. Dietterich, "An overview of MAXQ hierarchical reinforcement learning," in *Abstraction, Reformulation, and Approximation*, T. Walsh and B. Y. Choueiry, Eds., vol. 1864 of *Lecture Notes in Computer Science*, pp. 26–44, Springer, New York, NY, USA, 2000.
- [24] T. G. Dietterich, "Hierarchical reinforcement learning with the MAXQ value function decomposition," *Journal of Artificial Intelligence Research*, vol. 13, pp. 227–303, 2000.
- [25] R. S. Sutton, D. Precup, and S. Singh, "Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning," *Artificial Intelligence*, vol. 112, no. 1-2, pp. 181–211, 1999.
- [26] R. Parr, *Hierarchical control and learning for markov decision processes [Ph.D. thesis]*, University of California, Berkeley, Calif, USA, 1998.
- [27] B. Hengst, *Discovering Hierarchy in Reinforcement Learning with HEXQ*, vol. 2, ICML, 2002.
- [28] A. Jonsson, *A Causal Approach to Hierarchical Decomposition in Reinforcement Learning*, University of Massachusetts Amherst, 2006.
- [29] N. Ando, T. Suehiro, K. Kitagaki, T. Kotoku, and W.-K. Yoon, "RT-middleware: distributed component middleware for RT (robot technology)," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3933–3938, Edmonton, Canada, August 2005.
- [30] W. Wang, Q. Cao, X. Zhu, and M. Adachi, "An automatic switching approach of robotic components for improving robot localization reliability in complicated environment," *Industrial Robot*, vol. 41, no. 2, pp. 135–144, 2014.
- [31] M. Helmert, "The fast downward planning system," *Journal of Artificial Intelligence Research*, vol. 26, pp. 191–246, 2006.
- [32] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, New York, NY, USA, 2014.

Research Article

EOG-sEMG Human Interface for Communication

Hiroki Tamura,¹ Mingmin Yan,² Keiko Sakurai,³ and Koichi Tanno⁴

¹Department of Environmental Robotics, University of Miyazaki, Miyazaki 889-2192, Japan

²Organization for Promotion of “Center of Community” Program, University of Miyazaki, Miyazaki 889-2192, Japan

³Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki, Miyazaki 889-2192, Japan

⁴Department of Electrical and Systems Engineering, University of Miyazaki, Miyazaki 889-2192, Japan

Correspondence should be addressed to Hiroki Tamura; htamura@cc.miyazaki-u.ac.jp

Received 26 November 2015; Revised 26 April 2016; Accepted 17 May 2016

Academic Editor: Hasan Ayaz

Copyright © 2016 Hiroki Tamura et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this study is to present electrooculogram (EOG) and surface electromyogram (sEMG) signals that can be used as a human-computer interface. Establishing an efficient alternative channel for communication without overt speech and hand movements is important for increasing the quality of life for patients suffering from amyotrophic lateral sclerosis, muscular dystrophy, or other illnesses. In this paper, we propose an EOG-sEMG human-computer interface system for communication using both cross-channels and parallel lines channels on the face with the same electrodes. This system could record EOG and sEMG signals as “dual-modality” for pattern recognition simultaneously. Although as much as 4 patterns could be recognized, dealing with the state of the patients, we only choose two classes (left and right motion) of EOG and two classes (left blink and right blink) of sEMG which are easily to be realized for simulation and monitoring task. From the simulation results, our system achieved four-pattern classification with an accuracy of 95.1%.

1. Introduction

During eye movements, a potential exists across the cornea and the retina, and it is the basis for the electrooculogram (EOG). The EOG can be modeled by a dipole [1] and used in medical systems. Several EOG-based human-computer interface studies are found in the literature. For example, a wheelchair controlled with eye movements is being developed for disabled and elderly people. The eye movement signals and the sensor signals are combined, and both direction and acceleration are controlled [2]. Surface electromyogram signals (sEMGs) are detected over the skin surface and are generated by the electrical activity of muscle fibers during contraction [3]. Moving muscles can be detected by analyzing sEMGs. One of the important applications of sEMGs is controlling artificial legs. Although head movement, which is a natural gesture, can be used to indicate a certain direction [4], seriously disabled people cannot move their neck and head. However, many of these people can employ facial muscle movement. sEMG is a way of studying facial muscle activities by recording the action potentials

from contracting fibers. sEMG can be detected with surface electrodes, which are easy to apply and noninvasive and pose no health and safety risks to the users. Computer systems can also be controlled by using face sEMG signals [5, 6]. These computer systems can recognize facial motion (left blink, right blink, and bite) by using sEMG sensors. Furthermore, EOG, electroencephalogram (EEG), and EMG signals can be classified in real time and can control movable robots by using an artificial neural network classifier [7].

Investigating the possibility of employing EOGs for a human-computer interface, the relation between the sight angle and an EOG is determined. In-depth studies have found that the slowly changing baseline drift makes it difficult to estimate continuous EOG signals, and this drift only appears in direct current (DC) signals in the circuit. To overcome this issue, our system previously proposed the use of alternating current (AC) of EOGs to reduce the baseline drift by segmentation of the signal [8–10]. In these papers [8–10], we introduced the mouse cursor control system using our EOGs technique. The initial thresholds of the eyes movement class (right, left, and voluntary blink) are

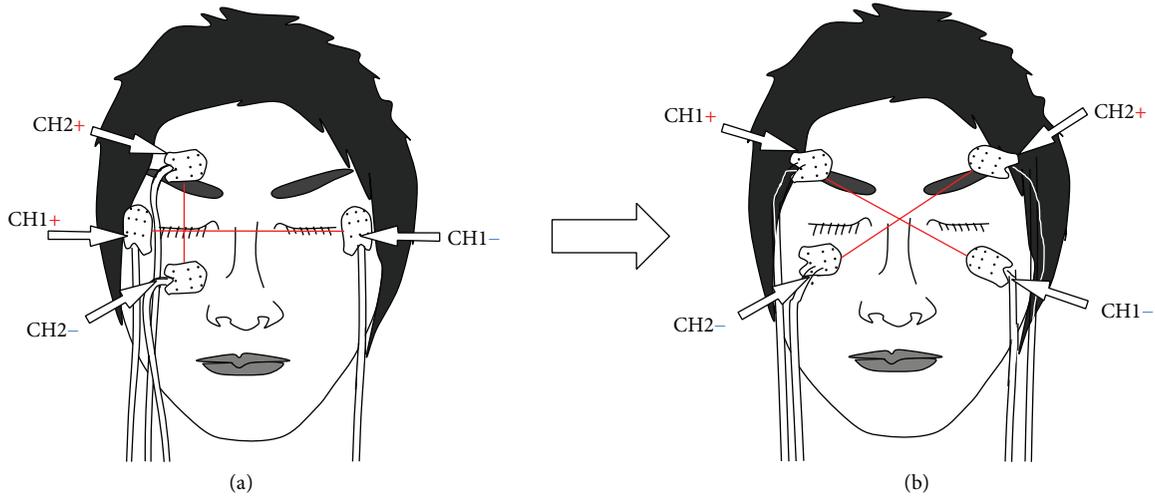


FIGURE 1: Conventional method (a) and the cross-channel method of the EOG (b) [8–10].

empirically determined individually for each user. Two eyes movement classes (right and left) are command of the same movement of the mouse cursor. In these papers [8–10], the eyes movement of diagonally lower right direction is a command of click processing. In addition, the voluntary blink is a command of click processing, too. However, these systems have a problem in which the face sEMG signals become noisy for EOG signals.

In this study, we are developing an EOG-sEMG human interface system for communication. Our proposed EOG device does not have the problem of artifacts from eye blinking. We apply an algorithm that uses the both dynamic variation of the DC element and the pattern classification of the AC element. This segmentation of the signal reduces baseline drift. Although there was a 3-electrode method which could measure vertical and horizontal components in EOG which was considered as the noise during the EEG measuring [11], our EOG system still uses cross-channels which use 4 electrodes to both improve the accuracy of EOG measuring and realize the EOG-sEMG “dual-modality” process simultaneously. Furthermore, the cross-channel EOG signals have a similar performance to the plus-channel method which is widely used in EOG measuring. In addition, the electrode position of cross-channel EOG signals method has a good feature that sEMG can also be effectively measured at the same time. In comparison with other papers [12], this is a useful merit.

Compared with other related researches on human-computer interfaces for helping people with disabilities such as eye tracking systems [13], which use image processing, the EOG-sEMG based on our proposal human-computer interface has strong anti-interference ability from the environment such as lights and noises. The patients could use this type of interface even without sight. In particular, the ALS patients could still control their eyeballs and muscles around the eyes even in terminal stage, in which they could use the EOG-sEMG based on our proposal human-computer interface also. Although image processing devices are widely used because of being intuitive and more predictable,

the EOG-sEMG based on our proposal human-computer interface still provides good choice for the severely handicapped persons.

In this paper, we propose a technique that can perform face pattern recognition by recording the EOGs and sEMGs. In order to prove the performance of our proposed method, we tried 3 related experiments step by step. At first, although as much as 7 classes could be recognized, dealing with the real state of the patients and the initial accuracy tests, we only choose two classes (left and right motion) of EOG and two classes (left blink and right blink) of sEMG which are easily to be realized for simulation and monitoring task. Next, we carried out experiments of character inputs by using our proposed system with those 4 classes. The experimental participants were healthy men in their twenties who agree to join in our experiments with no coercion. From these experiments, we showed that the four-pattern recognition of our proposed system is easy to learn and to use. Further, we applied our proposed simplified method (software that can do character input with one click (one pattern recognition)) for the severely handicapped persons (muscular dystrophy patients).

2. The Measurement System

2.1. Cross-Channels. As shown in Figure 1(a), the plus-channel method is widely used as the most conventional way for recording EOG signals in the previous research: the horizontal signals were recorded by Channel 1 (CH1), and the vertical signals were recorded by Channel 2 (CH2) [8–10]. In this paper, to improve the accuracy of the EOG signals, a new cross-channel method is proposed as shown in Figure 1(b). The horizontal and the vertical signals could be recorded by both channels at the same time. It is much easier to analyze data using double signals.

2.2. The EOG-sEMG Measurement System. In this subsection, the design of the EOG-EMG measurement system

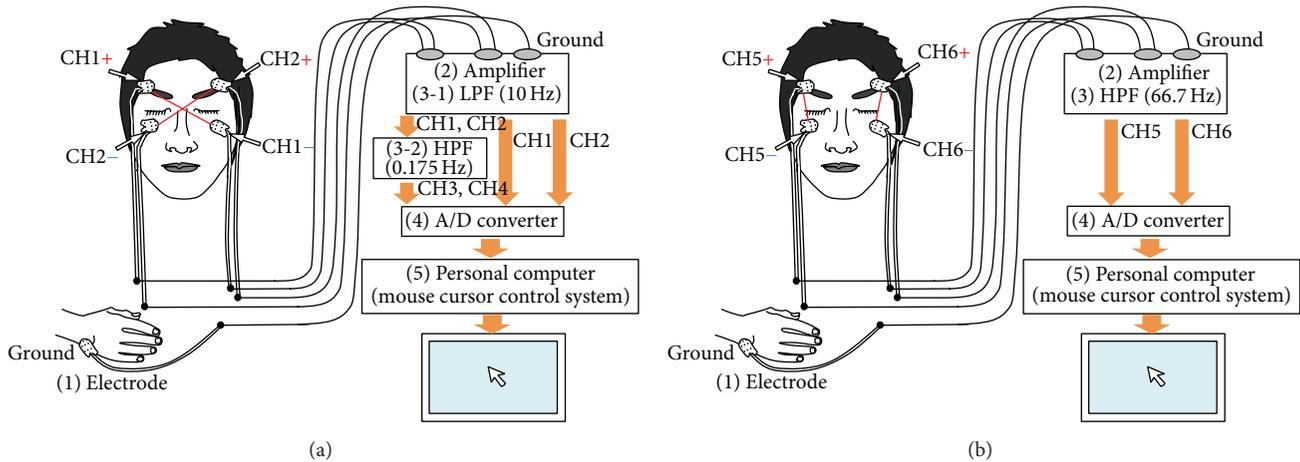


FIGURE 2: Proposed EOG-sEMG human interface system: (a) EOG flow and (b) sEMG flow. The EOG flow and the sEMG flow use the same five electrodes.

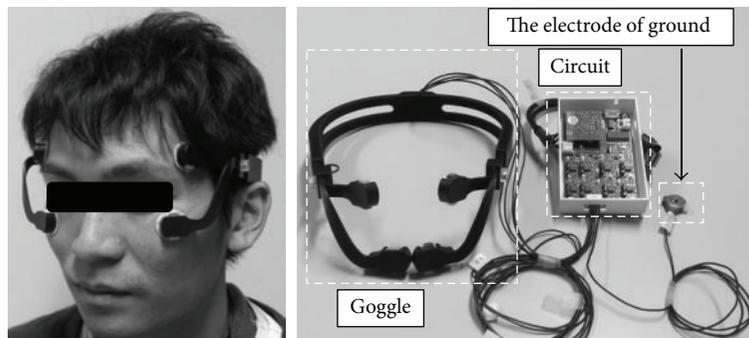


FIGURE 3: Interface device to fix the electrodes on the face.

is proposed. Figure 2 shows the formal scheme for acquisition and analysis of the EOG and sEMG signals for control and flow of information through the system. Our proposed system has four components: (1) amplifier, (2) filter, (3) A/D converter, and (4) mouse cursor control system. Specifically, this means the system consists of five electrodes, an A/D converter, a personal computer, and a monitor (shown in Figure 2). Four electrodes for the two channels of sEMG signals are pasted on the face, and one electrode is pasted on the right or left wrist as the ground.

2.3. The EOG-sEMG Sensing Device. In this subsection, we show the interface device used in this study. Figure 3 shows a person wearing the interface device. This interface device resembles goggles or glasses, fixing the electrodes. The signals measured by the device are sent to the PC using Bluetooth. The sent data are shown in Figure 4. The measured signals are DC currents for two EOG channels (2 ch), which are CH1 and CH2, and AC currents for two EOG channels (CH3 and CH4). In the same time, 2 ch signals of sEMG are also measured, which are CH5 and CH6. We will show the details of the measured data in the following section.

3. The EOG Recording System

In this section, we will introduce the EOG recording system. Our proposed EOG system is based on two flows: (1) the amplifier and the low-pass filter (DC element) and (2) the amplifier, low-pass filter, and the high-pass filter (AC element). After recording the signal amplitude (1000 times) and noise reduction measures for the biopotential data acquisition system, four kinds of eye movements (right, left, up, and down) are recognized accurately and the electronic noise reduction is also successful. It should be noted that the horizontal EOG signals are stronger than the vertical EOG signals. That is because vertical saccades are slower than horizontal saccades, and downward saccades are the slowest [14]. Figures 5–8 show that the four eye movements (right, left, up, and down) are clearly different. Moreover, CH1 and CH2 are DC signals, which can be used for recording the continuous movements of the eyes. CH3 and CH4 are AC signals of the EOG. Therefore, CH3 and CH4 strongly react to eye movements only.

Because the EOG signals change for all four eye movements, we imported CH1 + CH2, which is used for the vertical signals and CH1 – CH2 for the horizontal signals. In the experiments, we asked each participant to move his

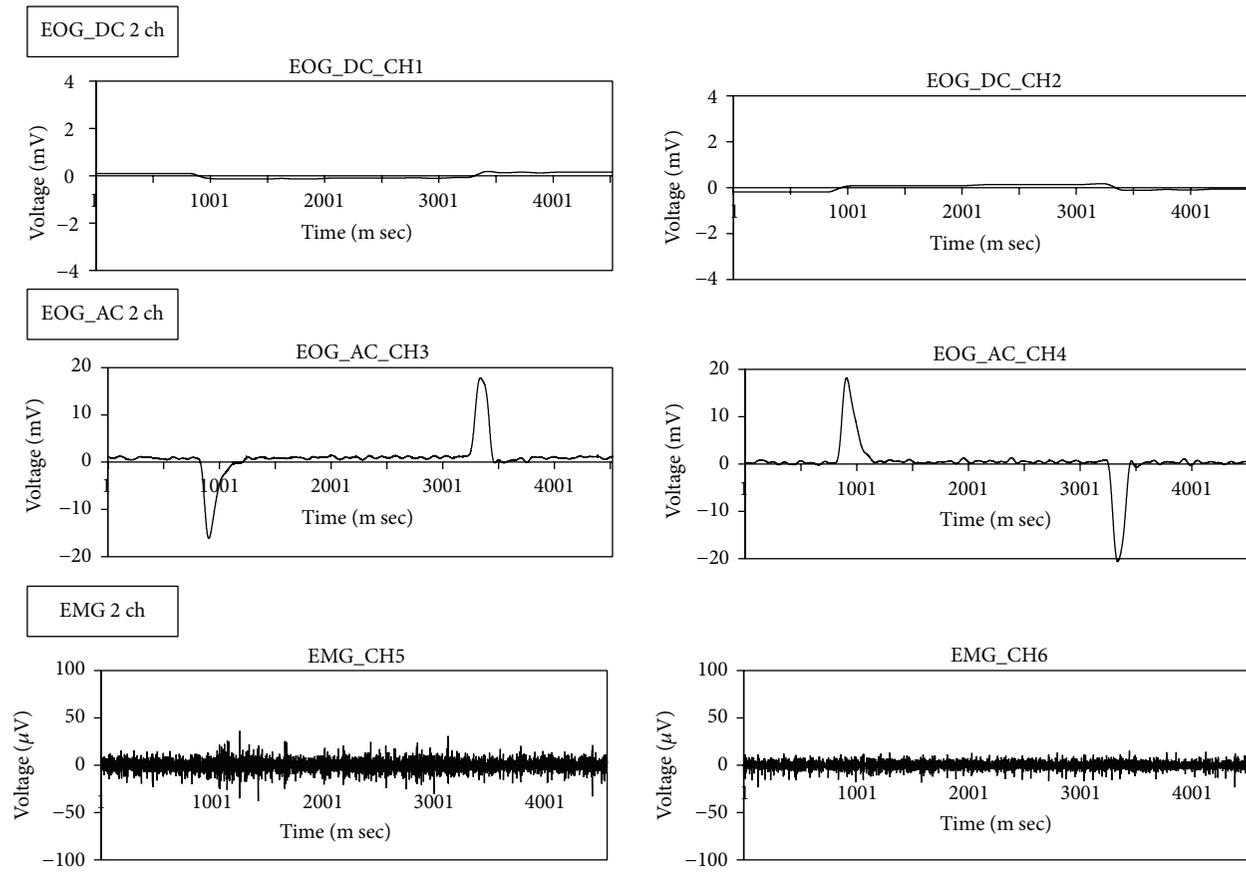


FIGURE 4: Data sent from the interface device. The signals are sEMG 2 ch and EOG 2 ch. The EOGs measure the AC and DC signals.

eyeballs to follow the next sequence: center-right-center-left-center-up-center-down-center. The results of these two new procedures are shown in Figure 9.

4. sEMG Signal Recording

Figure 2 shows the formal scheme for the acquisition and analysis of the sEMG signals for control and flow of information through the system. Our system consists of these four components: (1) surface electrodes, (2) amplifier, (3) high-pass filter, and (4) personal computer for sEMG signal classification. The sEMG signals detected by the surface electrodes are amplified and filtered prior to data acquisition to reduce noise artifacts and enhance spectral components that contain the information for data analysis. Two channels of sEMG signals can be used to recognize facial movement. To remove the DC level and the noise of the 60 Hz power line, the high-pass filter is used. The cutoff frequency of the high-pass filter is 66.7 Hz.

The recordings in Figure 10 show the applied noise reduction measures in our system. As a result, the data of the three sEMG classes (right blink, left blink, and bite) are clearly different. After filtering and amplifying (approximately 1000 times), the sEMG signals are digitized and then transferred to the personal computer. The sampling frequency of the measurement data is 1 KHz on a band from 0 Hz to 500 Hz.

The sEMG signals are processed by the moving average processing. The moving average processing calculates the rectified and unweighted average mean of the previous n ($n = 50$) data points. Then, the value after the moving average processing is determined as “active” or “inactive” based on the threshold. The thresholds (CH5 and CH6) are set according to the users. This method is necessary to set the threshold value of each user. Moreover, this system does not react to the usual blink. Figure 11 shows a diagram of this process.

5. Pattern Recognition Algorithm

In this section, we will introduce our proposed EOG-sEMG pattern recognition algorithm. Figure 12 shows the overall flow of our proposed system. This process consists of repeating steps. The pattern recognition consists of five classes: two classes (left and right movement) for the EOG and the three classes (left blink, right blink, and bite) for the sEMG.

If the sEMG data after signal processing exceeds a threshold, the pattern of the sEMG data is determined by CH1 only (right blink-like motion) or CH2 only (left blink-like motion) or both (the bite or strong blink). Our proposed algorithm initializes the DC elements of the EOG after the sEMG activity is completed. Furthermore, when the AC element of the EOG has not changed (eyes are not moving) and the sEMG is not active, our proposed algorithm initializes

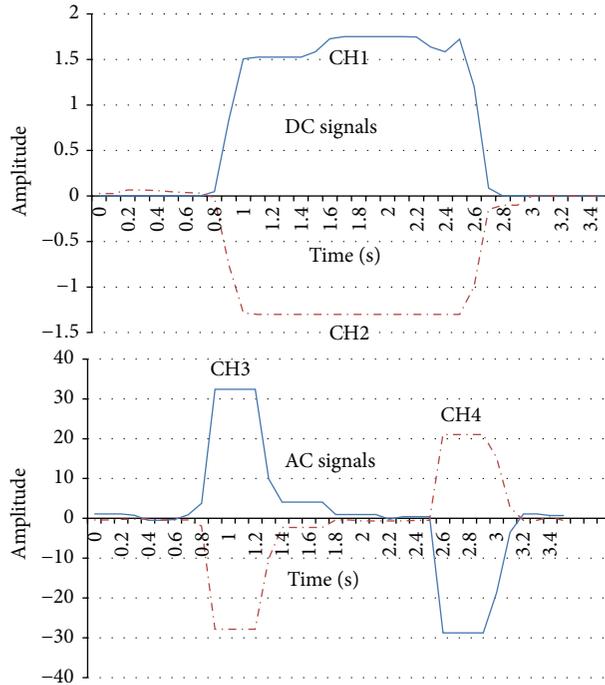


FIGURE 5: The EOG signals recording samples in CH1, CH2, CH3, and CH4 (right).

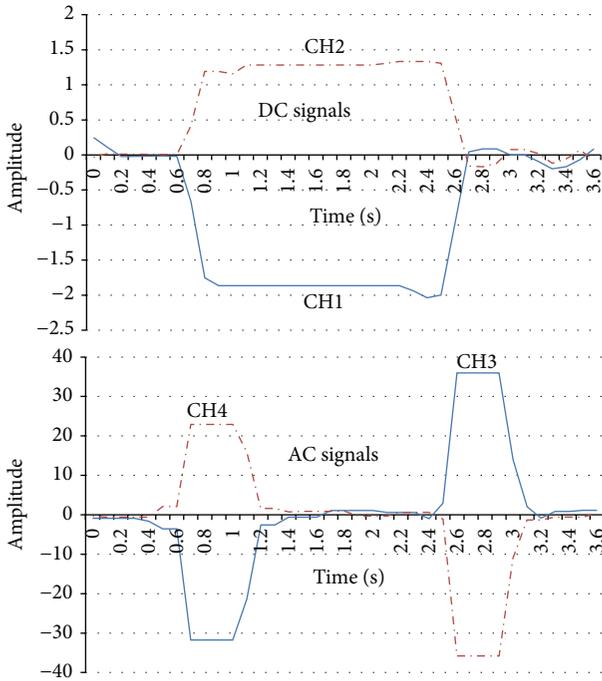


FIGURE 6: EOG signal recording samples in CH1, CH2, CH3, and CH4 (left).

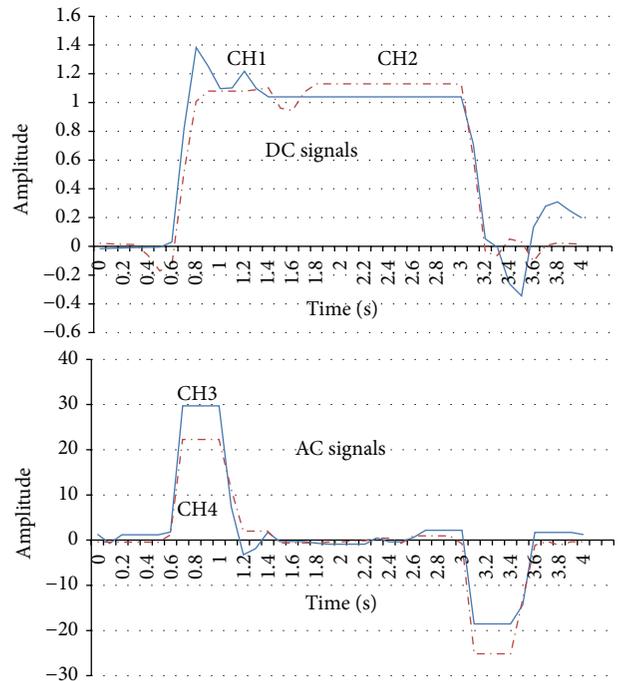


FIGURE 7: EOG signal recording samples in CH1, CH2, CH3, and CH4 (up).

the reference value which eyes are looking as the front (“EOG renewal” in Figure 12). In other words, our system determines that eyes are looking at the front, because AC element is not changed. At that time, our system updates as the reference value of the DC elements. After that, our

system uses the amount of change in DC from this reference value.

From the experimental rule, judging threshold between active and inactive is set. Next, when the eyes move, our algorithm compares the changing range of CH1 + CH2 and

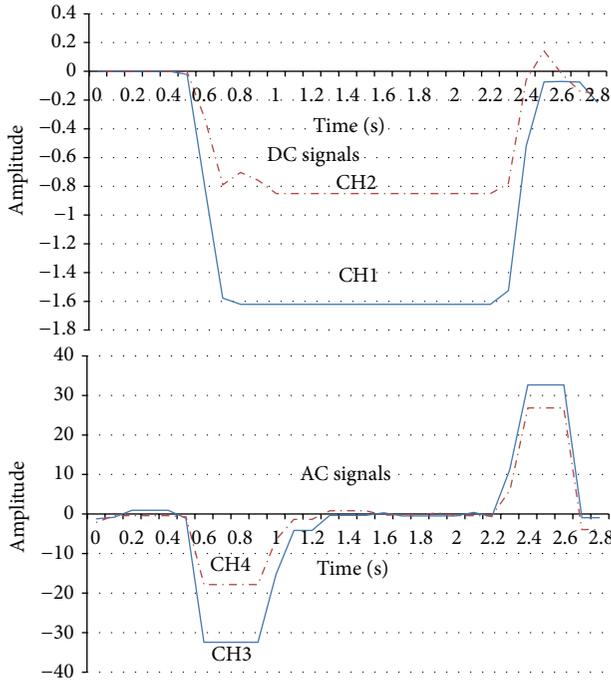


FIGURE 8: EOG signal recording samples in CH1, CH2, CH3, and CH4 (down).

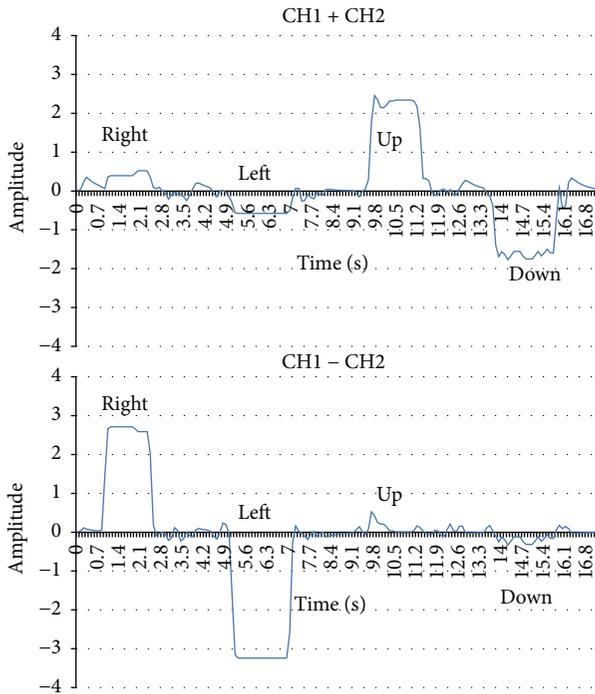


FIGURE 9: EOG signal recording samples in CH1 + CH2 (upside) and CH1 - CH2 (downside).

CH1 - CH2. When CH1 - CH2 is larger, our algorithm performs the determination process of the EOG. When CH1 + CH2 is larger, our algorithm determines that the eyes moved in the vertical direction, because the EOG data for the vertical direction of the eyes is similar to the EOG data for the blink pattern. In addition, our experimental results indicate

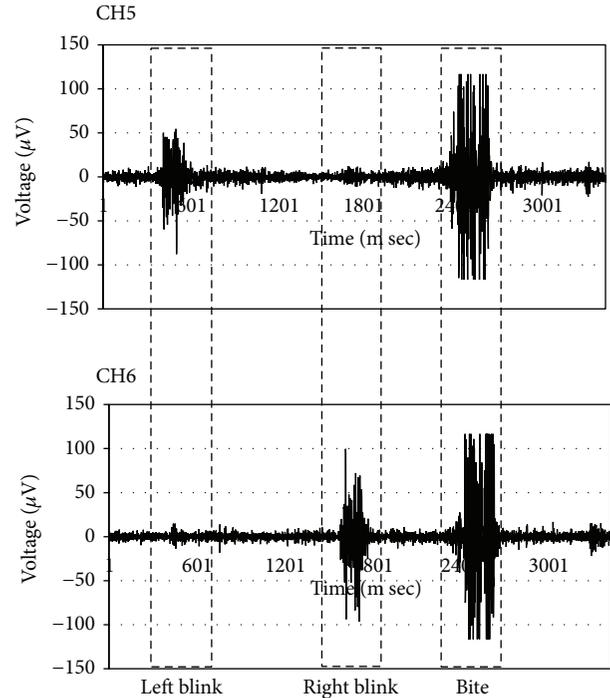


FIGURE 10: Activation levels of two-channel sEMG.

that many people cannot easily control the up direction of the eye.

Next, we introduce the EOG pattern recognition algorithm, an example of EOG pattern recognition processing shown in Figure 13. Although Figure 9 shows that all of eyes motions could be checked by values for CH1 + CH2 and CH1 - CH2, pattern marching algorithm is still necessary to solve the baseline shift caused by drift problem. The algorithm steps of Figure 13 are as follows:

- (1) The DC and AC elements of the EOG exceed the threshold for the right direction. It is determined that this is the right direction of the eyes.
- (2) The DC element continues to exceed the threshold for the right direction and the AC element returns to the baseline. The direction of the eyes continues to be to the right.
- (3) The AC element greatly changes in the negative direction when the eye direction returns to the center position.
- (4) The AC element and the DC element return to each element's baseline. Then, the baseline is updated.

Our system performs the determination of the right and left movement of the eyes by using this algorithm. This algorithm makes it possible to determine direction. This is a difficult process when using only the AC element.

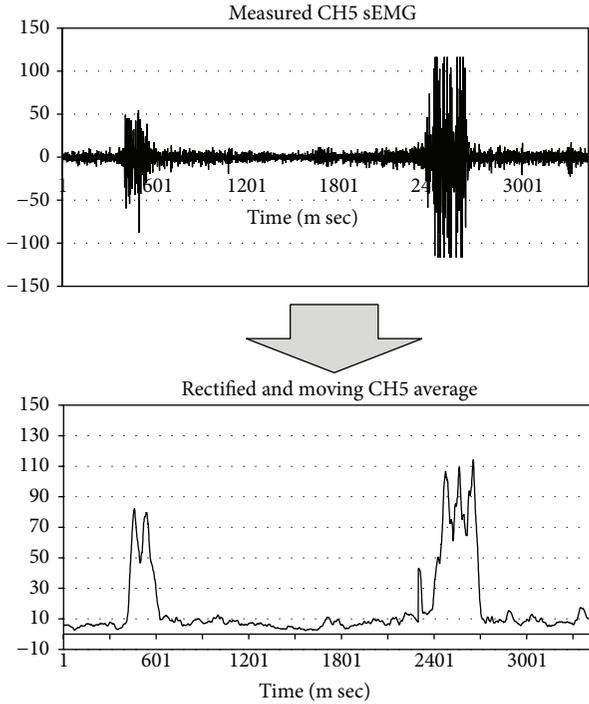


FIGURE 11: Diagram of process. SEMG signals are computed to obtain the rectified and moving averages. The signal is determined to be active or inactive based on the threshold.

6. Experiments and Results

To test the effectiveness of our proposed system, we conducted two experiments: one is a pattern recognition test and the other is a character input test.

6.1. Pattern Recognition Experiment. First, we conducted pattern classification experiments for the two eye movement classes (right and left) and the three facial sEMG classes (right blink, left blink, and bite). The interval of each action is 3 seconds (shown in Figure 14). Each participant performed each eye movement nine times. In addition, we tested whether our proposed system could reject a normal blink. We carried out the normal blink rejection test 30 times. The participants of the experiments were eight healthy males who are all the numbers of our lab in their twenties who gave their consent to participate in this experiment.

6.2. Pattern Recognition Experimental Results. The pattern recognition experimental results are shown in Table 1 and the last line shows the performance evaluation (PE) of each pattern under 5 points rule (correct for 5 points, reject for 2 points, and miss for 0 points). Reject means that input pattern did not react. Miss means that input pattern was wrong response. From the experiments, the EOG pattern recognition (right and left) of our proposed system is reliable. From these experimental results, our proposed system shows good performance in the recognition of the four classes (right, left, right blink, and left blink). The average recognition rate was

TABLE 1: Pattern recognition experimental results.

	Right	Left	Right blink	Left blink	Bite
K.G	8/9 (reject: 1)	9/9	9/9	9/9	4/9 (reject: 5)
T.N	9/9	9/9	9/9	7/9	8/9 (reject: 1)
M.Y	9/9	8/9	9/9	8/9	4/9 (reject: 5)
K.N	9/9	9/9	9/9	9/9	8/9 (reject: 1)
M.F	9/9	9/9	9/9	9/9	3/9 (reject: 6)
R.K	8/9	9/9	7/9	7/9	9/9
K.M	7/9 (reject: 2)	8/9 (reject: 1)	8/9	9/9	5/9 (reject: 4)
T.T	9/9	9/9	9/9	9/9	7/9 (reject: 2)
Ave.	94% (reject: 4%)	97% (reject: 1%)	96%	94%	67% (reject: 33%)
P.E 5	4.78	4.87	4.80	4.70	4.01

95.1%, the average reject rate was 1.4%, and the average miss rate was only 3.5%, and the average PE is 4.79 of 5.

However, the sEMG bite pattern recognition is not as good as expected. One reason is that some people have troubles of occlusal irregularity that cause divergence between their cheeks. This kind of bite actions was usually recognized as right or left blink. Another reason is that we used a threshold tuning method for distinguishing normal blink, and the normal blink rejection test had a 97% success rate (232/240). Under this method, the bite actions which could not pass the threshold will be rejected the same as the normal blink.

6.3. Character Input Experiment. Next, we conducted the character input experiment using the four classes (right (EOG), left (EOG), right blink (sEMG), and left blink (sEMG)). The alphabet sentence input software, Hearty Ladder [6], is shown in Figure 15. We used the four-division selection method. We assigned the four patterns as follows:

Right (EOG) is the command to select lower right area.

Left (EOG) is the command to select lower left area.

Right blink (sEMG) is the command to select higher right area.

Left blink (sEMG) is the command to select higher left area.

The user repeated the filtering of four patterns until the character of the last pattern was selected. The control interval was 0.1 s. We conducted experiments of the “miyazaki” input task (8 characters: 4 operations for 1 character) which is the name of our university. The participants of the experiments were

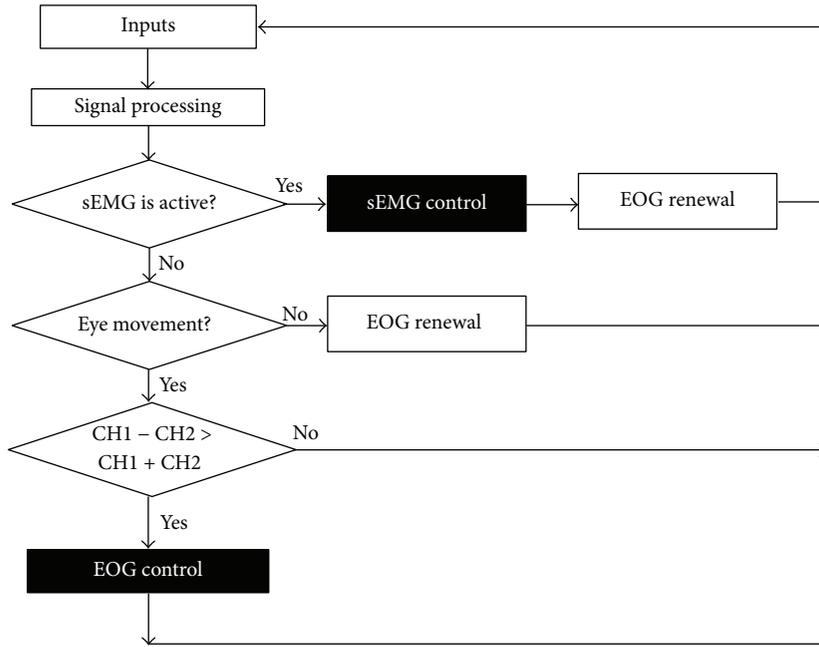


FIGURE 12: Flow of EOG-sEMG pattern recognition algorithm.

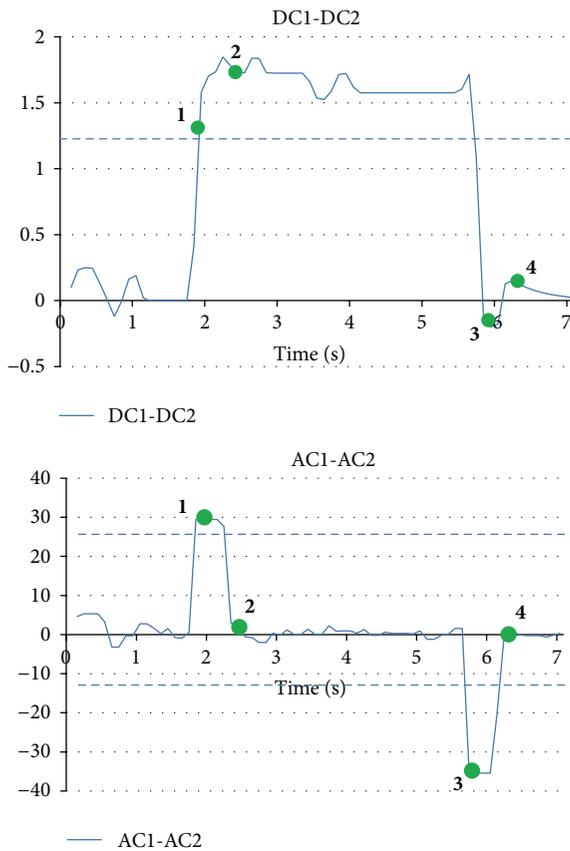


FIGURE 13: Example of EOG pattern recognition processing. The bold numbers represent the algorithm steps.

nine healthy males in their twenties who gave their consent to participate in this experiment. The experimental participants included six people experienced with this system.

TABLE 2: Character input experimental results.

	Average time (sec)	SD	Average miss rate (%)
K.I	42.0	3.1	0.0
K.M	35.8	5.0	0.0
R.K	60.2	12.4	3.5
T.T	53.2	13.0	2.5
K.G	61.0	15.5	5.0
T.N	43.2	2.7	0.0
M.F (inexperienced)	68.6	5.5	0.0
M.S (inexperienced)	36.6	5.2	2.0
R.K (inexperienced)	49.4	19.8	0.0
Ave.	50.0	—	1.45

6.4. Character Input Experimental Results. We tried five trials of the “miyazaki” input task. We counted the number of misses and the time duration until the end of input. The experimental results are shown in Table 2. From these experiments, the average time of one trial was 50.0 s. The average input time of one character of the alphabet was approximately 6.3 s (50 s/8 characters) in our proposed system. The time of the fastest participant, K.M, was 4.5 s (35.8 s/8 characters). In addition, if we considered the participants M.F, M.S, and R.K who had been using our proposed system over 100 h as the experienced and the other 6 participants who did not use our proposed system by any chance as inexperienced, the average time of the experienced participants was 49.2 s and that of the inexperienced participants was 51.5 s. The miss rate of the inexperienced participants was also low. From these results,

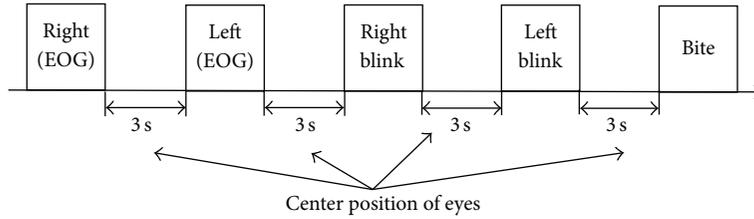


FIGURE 14: Basic tasks in pattern recognition experiment.

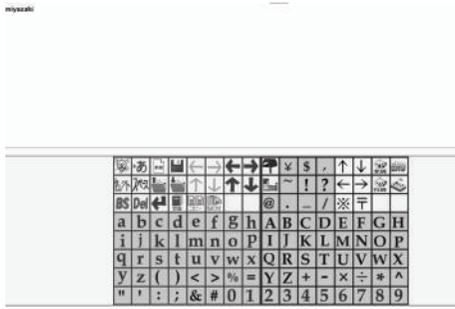


FIGURE 15: The alphabet sentence input software for the “miyazaki” task.

we consider that one advantage of our proposed system is that users do not need a lot of training.

We compared the performance of our proposed method with the EOG method and the sEMG method. The EOG method used right and left eye motions only as the 2-division selection method. The sEMG method used right blinks and left blinks only as the 2-division selection method. We tried the same experiments of five trials of the “miyazaki” input task with the same participants. These experimental results are shown in Table 3. It can be seen that our proposed method has the best performance of all the three methods. We calculated the time required for one operation. The EOG method was 1.93 s, the sEMG method was 1.65 s, and our proposed method was 1.56 s. In addition, the miss rate of the sEMG method was less than that of our proposed method. One of the most conclusive reasons is that the number of operations (7 operations for 1 character) is bigger in the 2-division selection method than the 4-division selection method.

As we mentioned in this section, 2 or 4 patterns could be recognized by our proposed system. All these patterns could be assigned as separated functions and, dealing with the state of users, all these patterns could be employed by any combination of them. In the next section, we will discuss the application of our proposed system for the severely handicapped persons using only 1 pattern.

7. Experimental Results of the Severely Handicapped Persons

In this paper, we applied our proposed method for the severely handicapped persons (muscular dystrophy patients).

TABLE 3: Comparison of experimental results of our proposed method with the EOG method and sEMG method.

	Average time (sec)	SD	Average miss rate (%)
EOG method (2-division)	77.0	8.7	2.06
sEMG method (2-division)	66.0	14.1	0.07
Our method (4-division)	50.0	11.6	1.45

We have gotten the approval of the ethics committee of the University of Miyazaki. The experiment of Section 6 with complicated operation will cause great fatigue to the people with disabilities, so experiments were carried out in the way that one character will be entered with a single click using EOG or sEMG. Only one pattern was selected from EOG or sEMG patterns that could be easily used. In this character input method, character is automatically selected by software; user can be selected by clicking when the target character has been chosen by the software. In this experiment, clicking processing is carried out by the recognition of EOG or sEMG activity.

This experiment had the two subjects. In order to minimize the influences of physical condition of subjects and environments, we tried the same experiments 10 times on different occasions in different days. The continuous 30 minutes in one experiment was used for testing the Japanese character input. Table 4 shows the experimental results. From Table 4, recognition rate of sEMG method is 99.2% and recognition rate of EOG method is 98.0%. The input time was slow about 1 to 2 seconds compared to healthy subjects. Therefore, the handicapped persons have the same result as healthy ones using a single click. From these results, we can say that our approach is effective in severely handicapped persons.

8. Conclusions

In this study, we introduced an EOG-sEMG human-computer interface device designed for patients suffering from amyotrophic lateral sclerosis or other illnesses. We use both cross-channels and parallel lines channels on the face with the same 4 electrodes. This system could record EOG and sEMG signals as dual-threading for pattern recognition simultaneously. Furthermore, our proposed method using

TABLE 4: The experimental results of the two handicapped persons.

	Average time of one character input (sec) (healthy person)	Average miss rate (%)
EOG method (1-click)	16.1 (13.5)	2.0
sEMG method (1-click)	14.6 (13.5)	0.8

the combination of AC and DC elements of EOG reduces the corresponding drift and enables continuous operation in the recording of eye movements. The experimental results showed that our proposed method is effective for four-pattern recognition (right (EOG), left (EOG), right blink (sEMG), and left blink (sEMG)). In particular, our proposed method demonstrated good performance for the character input experiments. The miss rate was only 1.4%. In addition, the results of inexperienced and experienced participants showed very little difference. From these results, we think that one advantage of our proposed system is that users do not need a lot of training.

Furthermore, we compared our proposed method with the EOG method and the sEMG method. Our proposed method required the shortest time for character input. Our proposed human-computer interface can be applied in the EOG system, the sEMG system, and the EOG-sEMG system. It is possible to use our system for patients who can only control their eye motions. It is also possible to use only the sEMG signals with a high success rate until the user is comfortable using the device. Our proposed human-computer interface has the advantage that it can be used depending on the situation.

We hope to be able to communicate with patients suffering from amyotrophic lateral sclerosis or other illnesses by using our system. In our future work, we plan to test many subjects and more severely disabled people.

Competing Interests

This research was funded by SYOWA Co. Ltd.

Acknowledgments

This research was supported by Miyazaki Prefectural Industrial Support Foundation (R&D) and a Grant-Aid for Young Scientists (B) (23700668) KAKEN. The authors are grateful to Mr. Kazuhiko Inami and Mr. Kazuya Gondou. They gratefully acknowledge the work of past and present members of their laboratory.

References

- [1] S. Venkataramanan, P. Prabhat, S. R. Choudhury, H. B. Nemade, and J. S. Sahambi, "Biomedical instrumentation based on Electrooculogram (EOG) signal processing and application to a hospital alarm system," in *Proceedings of the Proceedings of International Conference on Intelligent Sensing and Information Processing (ICISIP '05)*, pp. 535–540, IEEE, Chennai, India, January 2005.
- [2] R. Barea, L. Boquete, M. Mazo, and E. López, "System for assisted mobility using eye movements based on electrooculography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 4, pp. 209–218, 2002.
- [3] K. Tomohiro, M. Tadashi, K. Tohru, and S. Tsugutake, *Biomechanism Library Practical Usage of Surface Electromyogram*, Tokyo Denki University Press, Tokyo, Japan, 2006.
- [4] A. Assareh, S. Konjkav, A. Fallah, and S. M. P. Firoozabadi, "A new approach for navigating automatic wheelchairs using EMG signals feature extraction and classification with an adaptive controller," in *Proceedings of the 12th International Conference on Biomedical Engineering*, Singapore, December 2005.
- [5] H. Tamura, T. Murata, Y. Yamashita, K. Tanno, and Y. Fuse, "Development of the electric wheelchair hands-free semi-automatic control system using the surface-electromyogram of facial muscles," *Artificial Life and Robotics*, vol. 17, no. 2, pp. 300–305, 2012.
- [6] H. Tamura, T. Manabe, T. Goto, Y. Yamashita, and K. Tanno, "A study of the electric wheelchair hands-free safety control system using the surface-electromyogram of facial muscles," in *Intelligent Robotics and Applications: Third International Conference, ICIRA 2010, Shanghai, China, November 10–12, 2010. Proceedings, Part II*, vol. 6425 of *Lecture Notes in Computer Science*, pp. 97–104, Springer, Berlin, Germany, 2010.
- [7] C. K. Young and M. Sasaki, "Mobile robot control by neural network EOG gesture recognition," in *Proceedings of the 8th International Conference on Neural Information Processing*, vol. 1, pp. 322–328, 2001.
- [8] H. Tamura, M. Miyashita, K. Tanno, and Y. Fuse, "Mouse cursor control system using electrooculogram signals," in *Proceedings of the World Automation Congress, IFMIP 239*, pp. 1–6, September 2010.
- [9] H. Tamura, M. Yan, M. Miyashita, T. Manabe, K. Tanno, and Y. Fuse, "Development of mouse cursor control system using DC and AC elements of electrooculogram signals and its applications," *International Journal of Intelligent Computing in Medical Sciences and Image Processing*, vol. 5, no. 1, pp. 3–15, 2013.
- [10] M. Yan, H. Tamura, and K. Tanno, "Gaze estimation using electrooculogram signals and its mathematical modeling," in *Proceedings of the CDROM, IEEE 43rd International Symposium on Multiple-Valued Logic*, pp. 18–22, Toyama, Japan, May 2013.
- [11] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical Neurophysiology*, vol. 118, no. 1, pp. 98–104, 2007.
- [12] H. S. Dhillon, R. Singla, N. S. Rekhi, and R. Jha, "EOD and EMG based virtual keyboard: a brain-computer interface," in *Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology*, pp. 259–262, 2009.
- [13] Tobii, <http://www.tobii.com>.
- [14] J. G. Thomas, "The dynamics of small saccadic eye movements," *Journal of Physiology*, vol. 200, no. 1, pp. 109–127, 1969.

Research Article

A Fast Framework for Abrupt Change Detection Based on Binary Search Trees and Kolmogorov Statistic

Jin-Peng Qi,¹ Jie Qi,¹ and Qing Zhang²

¹College of Information Science & Technology, Donghua University, Shanghai 201620, China

²Australia e-Health Research Centre, CSIRO Computation Informatics, Brisbane, QLD 4060, Australia

Correspondence should be addressed to Jin-Peng Qi; qipengkai@126.com

Received 28 September 2015; Accepted 28 April 2016

Academic Editor: Hiroki Tamura

Copyright © 2016 Jin-Peng Qi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Change-Point (CP) detection has attracted considerable attention in the fields of data mining and statistics; it is very meaningful to discuss how to quickly and efficiently detect abrupt change from large-scale bioelectric signals. Currently, most of the existing methods, like Kolmogorov-Smirnov (KS) statistic and so forth, are time-consuming, especially for large-scale datasets. In this paper, we propose a fast framework for abrupt change detection based on binary search trees (BSTs) and a modified KS statistic, named BSTKS (binary search trees and Kolmogorov statistic). In this method, first, two binary search trees, termed as BSTcA and BSTcD, are constructed by multilevel Haar Wavelet Transform (HWT); second, three search criteria are introduced in terms of the statistic and variance fluctuations in the diagnosed time series; last, an optimal search path is detected from the root to leaf nodes of two BSTs. The studies on both the synthetic time series samples and the real electroencephalograph (EEG) recordings indicate that the proposed BSTKS can detect abrupt change more quickly and efficiently than KS, t -statistic (t), and Singular-Spectrum Analyses (SSA) methods, with the shortest computation time, the highest hit rate, the smallest error, and the highest accuracy out of four methods. This study suggests that the proposed BSTKS is very helpful for useful information inspection on all kinds of bioelectric time series signals.

1. Introduction

Abrupt change detection is to identify abrupt changes in the statistical properties of a signal series, which occur at unknown instants [1–3]. These changes are interesting because they are indicative of qualitative transitions in the data generation mechanism (DGM) underlying the signals. Currently, CP detection has attracted considerable attention in the fields of data mining and statistics, and it has been widely studied in many real-world problems, such as atmospheric and financial analyses [1], fault detection in engineering system [4, 5], climate change detection [6], genetic time series analyses [7], signal segmentation [8, 9], and intrusion detection in computer network [4].

In community of statistics, some nonparametric approaches for CP detection have been widely explored. For example, KS statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution or

between the empirical distribution function of two samples [10, 11]. Also, KS statistic and its modified versions are broadly investigated on many application fields, for example, testing hypotheses regarding activation in blood oxygenation level-dependent functional MRI data [12], modeling the cumulative distribution function of rub-induced AE signals, quantifying the goodness of fit to offer a suitable signal feature for diagnosis [13], as well as abrupt change detecting on EEG signals [14], and gene expression time series [15]. Meanwhile, as for the model-related statistic approaches, some modified cumulative sum (CUSUM) methods provide the asymptotic distributions of test statistics and the consistency of procedures and behave better in finite samples and have a higher stability with respect to the time of change than ordinary CUSUM procedures [16]. The CUSUM method and its revised versions have been widely applied to detect the structural breaks in the parameters of stochastic models, as well as the abrupt changes in the regression parameters of multiple time series regression models, such as multiple

CP detection in biological sequences [17], abrupt change detection in the regression parameters of a set of capital asset pricing data related to the Fama-French extension of the CAPM [16], and abrupt change detection in a shape-restricted regression model [18].

On the other hand, SSA is a powerful technique for time series analyses. SSA is nonparametric and requires no prior knowledge on the properties of time series signal [19]. The main idea of SSA is applied in the principal component analyses on the trajectory matrix with subsequent reconstruction of the original time series. SSA has been proved to be very successful and has already become a standard tool in the analyses of climatic [10], meteorological, and geophysical time series [11, 19]. Currently, SSA has been successfully applied in the real time series recordings, for example, abrupt change analyses on EMG-onset detection [12] and CP detection in time series [13]. Although SSA is a model-free method, it is not scalable to large-scale datasets, because it is time-consuming and sometimes invalid for time series analyses with less significant data fluctuation.

In addition, Wavelet Transform (WT) is another important tool for time series analyses [14, 15, 20–23]. WT has been widely applied in anomaly detection, time series prediction, image processing, and noise reduction [15, 23–25]. WT can represent general function at different scales and positions in a versatile and sophisticated manner, so that the data distribution features can be easily extracted from different time or space scales [25, 26]. As a simple WT, Haar Wavelet (HW) owns some attractive features including fast implementation and ability to analyze the local features. HW is very useful to find abrupt changes of discontinuity and high frequency in time series, so it is a potential candidate in modern electrical and computer engineering applications, such as signal and image compression, eye detection [27], abnormality detection on time series [28, 29], and abrupt change detection on autoregressive conditional heteroscedastic processes [30].

However, all of these methods above are time-consuming and sometime invalid for abrupt change detection near the left or the right boundary, especially for insignificant data fluctuation in large-scale time series. To resolve these problems, we propose a fast framework for CP detection based on binary search trees and a modified KS statistic, termed BSTKS for short. In this novel method, first, two BSTs are derived from a diagnosed time series. Second, three search criteria are introduced in terms of the statistic and variance fluctuations between two adjacent time series segments, and then an optimal search path is detected from the root to leaf nodes of two BSTs. Last, the proposed BSTKS and other KS, t , and SSA methods are tested on both the synthetic time series and real EEG recordings and evaluated in terms of computation time, hit rate, error, accuracy, and area under curve (AUC) of Receiver Operating Characteristic (ROC) curve analyses.

In general, for a certain bioelectric signal, an abrupt change means an important transition of biological functions or health states before and after a strong attack or an acute perturbation from internal or external environment. Therefore, it is very necessary to not only discern abrupt

change from all kinds of physiological and psychological time series signals, but also inspect the significant fluctuation between adjacent time series segments with different scales. The following sections focused on not only presenting the framework of the proposed BSTKS method through theoretical foundation, simulation, and evaluation, but also discussing how it can more quickly and efficiently detect abrupt change on both synthetic and real bioelectric EEG signals than other existing KS, t , and SSA methods. The rest of this paper is organized as follows. Section 2 gives the preliminary of abrupt change by introducing the statistic and variance fluctuations between two adjacent time series segments. Section 3 implements the integrated framework of the BSTKS method in terms of three search criteria in detail. Section 4 provides some representative experiments by using the synthetic time series and real EEG recordings and then analyzes the performance of BSTKS by comparing with other KS, t , and SSA methods. Section 5 gives summary and conclusion from previous sections.

2. Preliminary

2.1. Statistic Fluctuation. KS statistic is sensitive to differences in both location and shape of the cumulative distribution functions (c.d.f) of two samples. The null distribution of KS statistic is calculated under the null hypothesis that the two samples are drawn from the same distribution or one sample is drawn from the reference distribution. To detect an abrupt change from a diagnosed time series Z , we define the statistic fluctuation between two adjacent segments within Z by means of KS statistic as follows [1, 4, 19].

Definition 1. Supposing a time series sample, $Z = \{z_1, \dots, z_N\}$, one observes

$$Z = f\left(\frac{i}{n}\right) + X, \quad i = 1, \dots, N, \quad (1)$$

where $X = \{x_i\}_{i=1, \dots, N}$ is a set of the discrete and centred i.i.d random variables and f is a noisy mean signal with unknown distribution. The statistic fluctuation between two adjacent segments $Z_L = \{z_a, \dots, z_c\}$ and $Z_R = \{z_{c+1}, \dots, z_b\}$ is defined as

$$S_{mn}(x) = \left(\frac{mn}{m+n}\right)^{1/2} \sup_{x \in R} |F_m(x) - G_n(x)|, \quad (2)$$

in which $F_m(x)$ and $G_n(x)$ are the c.d.f of Z_L and Z_R , respectively; $m = c - a$, $n = b - c - 1$, and $m + n \leq N$. Supposing the hypothesized $F_m(x)$ and $G_n(x)$ in (2) are not available, we can derive the empirical cumulative distribution functions (e.c.d.f) of $F_m(x)$ and $G_n(x)$ from Z_L and Z_R . Then, $F_m(x)$ and $G_n(x)$ can be redefined as

$$\begin{aligned} F_m(x) &= P_m(Z_L \leq x) = \frac{1}{m} \sum_{i=a}^c I(z_i \leq x), \\ G_n(x) &= P_n(Z_R \leq x) = \frac{1}{n} \sum_{j=c+1}^N I(z_j \leq x), \end{aligned} \quad (3)$$

where $F_m(x)$ and $G_n(x)$ count the proportion of the sample points below level x .

Hypothesis 1. In order to discern an abrupt change on Z in terms of statistic fluctuation defined above, we introduce KS test for two adjacent segments Z_L and Z_R in Z as

(H_0) if $S_{mn}(z_c) \leq \delta$, no abrupt change occurs in Z ;

(H_1) if $S_{mn}(z_c) > \delta$, abrupt change occurs in Z ,

in which $\delta \in R$ is a threshold of the statistic fluctuation within Z belonging to an identical distribution. Then, we test (H_0) against (H_1) from observations. If an abrupt change c occurs in Z , there exists a value c satisfying $S_{mn}(z_c) > \delta$, $z_c \in [z_1, z_N]$, and $\delta \in R$. In this hypothesis, we assume that the number, the location, and the size of the function f in (1) are unknown, and the upper bound of the statistic fluctuation δ is supposed to be known.

2.2. Variance Fluctuation. Provided the statistic fluctuation defined in (2) is insignificant enough, it is difficult to detect abrupt change near the left or the right boundary within Z , especially when sample size N gets smaller. Therefore, we need to introduce another variable to calculate the variance fluctuation between two adjacent parts within a time series sample.

Definition 2. Supposing two adjacent segments $Z_L = \{z_a, \dots, z_c\}$ and $Z_R = \{z_{c+1}, \dots, z_b\}$ in $Z = \{z_1, \dots, z_N\}$, the variance fluctuation between Z_L and Z_R is defined as

$$D_{mn}(c) = \sup_{1 \leq L, R \leq N} \left| \frac{1}{m} \sum_{L=a}^c z_L - \frac{1}{n} \sum_{R=c+1}^b z_R \right|, \quad (4)$$

in which $m = c - a$, $n = b - c - 1$, and $m + n \leq N$.

Hypothesis 2. (H_0) If $D_{mn}(c) \leq \beta$, no abrupt change occurs at c in Z ; (H_1) if $D_{mn}(c) > \beta$, abrupt change occurs at c in Z .

Here, $\beta \in R$ is a variance threshold of time series Z which obeys an identical distribution. If there exists a value c satisfying $D_{mn}(c) > \beta$, $z_c \in [z_1, z_N]$, then an abrupt change occurs at c in Z .

3. Method

3.1. Two BSTs' Construction. In the first part of the proposed BSTKS method, two BSTs, that is, BSTcA and BSTcD, are constructed from a time series sample Z , by using multilevel HWT. Generally, as shown in Figure 1, a discrete time series signal $Z = \{z_1, z_2, \dots, z_N\}$ can be decomposed into the k th-level trend cA^k and k -level fluctuations, that is, cD^1, cD^2, \dots, cD^k , $k = 1, 2, \dots, \log_2 N$. The k -level HWT is the mapping H_k defined as [13]

$$Z \xrightarrow{H_k} (cA^k \mid cD^k \mid cD^{k-1} \mid \dots \mid cD^2 \mid cD^1), \quad (5)$$

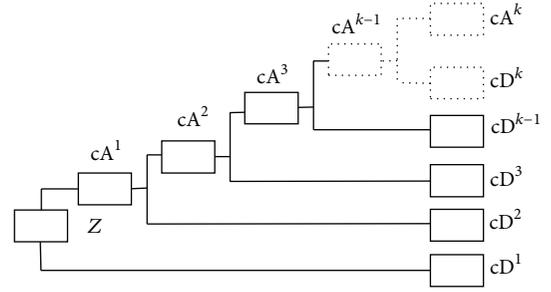


FIGURE 1: The diagram of a discrete time series Z decomposition by k -level HWT, which is composed of k -level cA and cD , that is, the average and difference coefficient vectors.

and then, the mapping H_k can be represented by the approximation and detail coefficient matrices, termed McA and McD as follows:

$$McA = \begin{bmatrix} cA_{1,1} & \dots & cA_{1,N} \\ \vdots & cA_{k,j} & \vdots \\ cA_{M,1} & 0 & 0 \end{bmatrix}, \quad (6)$$

$$McD = \begin{bmatrix} cD_{1,1} & \dots & cD_{1,N} \\ \vdots & cD_{k,j} & \vdots \\ cD_{M,1} & 0 & 0 \end{bmatrix},$$

where $0 \leq k \leq M = \log_2 N$ and $1 \leq j \leq N/2^k$.

Supposing the size of a diagnosed Z is divisible k times by 2, the j th element $cA_{k,j}$ in cA^k and the j th element $cD_{k,j}$ in cD^k can be denoted as

$$cA_{k,j} = \frac{1}{(\sqrt{2})^{\wedge k}} \left(\sum_{i=a}^b z_i \right), \quad (7)$$

$$cD_{k,j} = \frac{1}{(\sqrt{2})^{\wedge k}} \left(\sum_{L=a}^c z_L - \sum_{R=c+1}^b z_R \right),$$

where $1 \leq k \leq \log_2 N$ and $2^k(j-1) + 1 \leq i \leq j * 2^k$; $a = 2^k(j-1) + 1$, $c = 2^k(j-1) + 2^{(k-1)}$, and $b = 2^k * j$.

During two BSTs' construction, as shown in Figure 2, the non-leaf nodes in BSTcA and BSTcD are assembled by the k -level coefficient vectors of McA and McD , respectively; and then the leaf nodes are derived directly from the original time series Z . Therefore, the features of abrupt change in Z can be reflected and distributed into the different non-leaf nodes of BSTcA and BSTcD, in accordance with the k level coefficient vectors in McA and McD .

3.2. CP Detection Based on Three Search Criteria. To find an optimal path towards the potential CP within a given time series Z quickly and efficiently, some search criteria need to be introduced, and then the data exceptions can be detected from the root to leaf nodes of two BSTs. As for the statistic

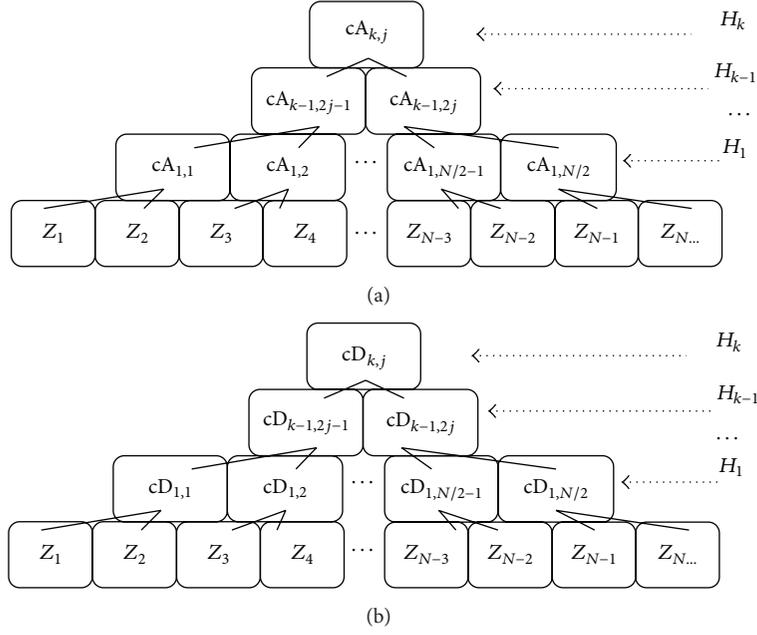


FIGURE 2: The diagrams of two binary trees, BSTcA and BSTcD, which are constructed by McA and McD, as well as the original time series Z .

fluctuation within BSTcA, first, a new variable $z_{k,j}$ is defined according to a current non-leaf node $cA_{k,j}$ in BSTcA,

$$z_{k,j} = \frac{1}{2^k} \left(\sum_{i=a}^b z_i \right) = \frac{(\sqrt{2})^{\wedge k}}{2^k} cA_{k,j}, \quad (8)$$

where $1 \leq k \leq \log_2 N$, $1 \leq j \leq N/2^k$; $a = 2^k(j-1) + 1$, $b = 2^k * j$, and $a \leq i \leq b$. Then, the statistic fluctuation between two adjacent segments $Z_L = \{z_a, \dots, z_c\}$ and $Z_R = \{z_{c+1}, \dots, z_b\}$ can be defined by a modified KS statistic as

$$S'_{mn}(k, j) = \left(\frac{nm}{n+m} \right)^{1/2} \cdot \left| \left\{ \frac{1}{n} \sum_{iL=a}^c I(z_{iL} \leq z_{k,j}) - \frac{1}{m} \sum_{iR=c+1}^b I(z_{iR} \leq z_{k,j}) \right\} \right|, \quad (9)$$

in which $z_{k,j}$ is a new element defined in (8); m and n stand for the sizes of Z_L and Z_R , respectively; $1 \leq k \leq \log_2 N$, $1 \leq j \leq N/2^k$; $a = 2^k(i-1) + 1$, $b = 2^k j$, and $c = 2^k(j-1) + 2^{(k-1)}$. $S'_{mn}(k, j)$ measures the e.c.d.f difference between Z_L and Z_R , and the larger $S'_{mn}(k, j)$ means the more significant statistic fluctuation between Z_L and Z_R . Therefore, a potential abrupt change might occur at c in Z with more probability.

Definition 3. For a current non-leaf node $cA_{k,j}$ in BSTcA, with its left and right-child nodes $cA_{k-1,2j-1}$ and $cA_{k-1,2j}$, the distance of e.c.d.f, $S_{k,jL}$, and $S_{k,jR}$ can be defined as

$$S_{k,jL} = S'_{mn}(k, j; k-1, 2j-1) = \left(\frac{nm}{n+m} \right)^{1/2} \cdot \left| \frac{1}{n} \left(\sum_{i=a}^b I(z_i \leq z_{k,j}) \right) - \frac{1}{m} \sum_{iL=a}^c I(z_{iL} \leq z_{k,j}) \right|$$

$$= \left(\frac{nm}{n+m} \right)^{1/2} W \left| \frac{1}{n} \left(\sum_{i=a}^b I(z_i \leq cA_{k,j}) \right) - \frac{1}{m} \sum_{iL=a}^c I(z_{iL} \leq cA_{k,j}) \right|,$$

$$S_{k,jR} = S'_{mn}(k, j; k-1, 2j) = \left(\frac{nm}{n+m} \right)^{1/2}$$

$$\cdot \left| \frac{1}{n} \left(\sum_{i=a}^b I(z_i \leq z_{k,j}) \right) - \frac{1}{m} \sum_{iR=c+1}^b I(z_{iR} \leq z_{k,j}) \right|$$

$$= \left(\frac{nm}{n+m} \right)^{1/2} W \left| \frac{1}{n} \left(\sum_{i=a}^b I(z_i \leq cA_{k,j}) \right) \right.$$

$$\left. - \frac{1}{m} \sum_{iR=c+1}^b I(z_{iR} \leq cA_{k,j}) \right|,$$

(10)

where $2 \leq k \leq \log_2 N$, $1 \leq j \leq N/2^k$; $a = 2^k(j-1) + 1$, $b = 2^k j$, $c = 2^k(j-1) + 2^{(k-1)}$; $n = 2^k$, $m = 2^{k-1}$; and $W = (\sqrt{2})^{\wedge k} / 2^k$. To estimate an optimal path towards the potential change position within Z , without loss of generality, the first search criterion is introduced based on the statistic fluctuations $S_{k,jL}$ and $S_{k,jR}$.

Criterion 1. Given two statistic fluctuation variables $S_{k,jL}$ and $S_{k,jR}$ in accordance with two non-leaf child nodes $cA_{k-1,2j-1}$ and $cA_{k-1,2j}$ of the current selected node $cA_{k,j}$ in BSTcA, and $2 \leq k \leq \log_2 N$,

- (a) if $(S_{k,j;L} > S_{k,j;R}) \wedge (S_{k,j;L} > C(\alpha))$ holds true, then the left-child node $cA_{k-1,2j-1}$ is selected and involved into the current search path; meanwhile, the right-child $cA_{k-1,2j}$ is discarded;
- (b) if $(S_{k,j;R} > S_{k,j;L}) \wedge (S_{k,j;R} > C(\alpha))$ holds true, then the right-child node $cA_{k-1,2j}$ is selected and involved into the current search path; meanwhile, the left-child $cA_{k-1,2j-1}$ is discarded.

Proof. For a selected non-leaf node $cA_{k,j}$ in BSTcA, as shown in Figure 3, the original time series Z is divided equally into two adjacent segments Z_L and Z_R , which are covered by two non-leaf child nodes $cA_{k-1,2j-1}$ and $cA_{k-1,2j}$, respectively. According to the definitions of $S_{k,j;L}$ and $S_{k,j;R}$ in (10), the satisfied $S_{k,j;L} > S_{k,j;R}$ indicates that the statistic fluctuation within Z_L is more significant than that one within Z_R ; that is, a potential abrupt change might be contained in Z_L with more probability than in Z_R , and vice versa. Furthermore, if $S_{k,j;L} > C(\alpha)$ holds true, then (H_1) of Hypothesis 1 is satisfied; that is, abrupt change occurs in Z_L , and vice versa, where $C(\alpha)$ is the critical value predefined in an identical distribution and α is the significance level. Therefore, one of the two child nodes $cA_{k-1,2j-1}$ and $cA_{k-1,2j}$ is selected and involved into the current search path; meanwhile, the remaining one is discarded. Once the statistic fluctuation is significant enough, an optimal search path can be detected by Criterion 1 from the top to the last non-leaf level in BSTcA. However, the search procedure is probably forced to cease because the statistic fluctuation is so insignificant that Criterion 1 is invalid for detecting it, especially for the left or the right boundary when sample Z is with smaller size N . Therefore, it is necessary to introduce another search criterion based on the variance fluctuations within BSTcD. \square

Definition 4. For a current non-leaf node $cD_{k,j}$ in BSTcD, with its left and right-child nodes $cD_{k-1,2j-1}$ and $cD_{k-1,2j}$, respectively, the variance fluctuations $D_{k,j;L}$ and $D_{k,j;R}$ are defined in terms of (4) as

$$D_{k,j;L} = D'_{mn}(k, j; k-1, 2j-1) = \left(\frac{nm}{n+m}\right)^{1/2} \cdot \left\| \frac{1}{n} \left\{ \left(\sum_{iL=a}^c z_{iL} \right) - \left(\sum_{iR=c+1}^b z_{iR} \right) \right\} \right\| - \left\| \frac{1}{m} \left\{ \left(\sum_{La=a}^{lc} z_{La} \right) - \left(\sum_{Lb=lc+1}^c z_{Lb} \right) \right\} \right\|$$

$$= \left(\frac{nm}{n+m}\right)^{1/2} \left\| |N'(cD_{k,j})| - |M'(cD_{k-1,2j-1})| \right\|,$$

$$D_{k,j;R} = D'_{mn}(k, j; k-1, 2j) = \left(\frac{nm}{n+m}\right)^{1/2} \cdot \left\| \frac{1}{n} \left\{ \left(\sum_{iL=a}^c z_{iL} \right) - \left(\sum_{iR=c+1}^b z_{iR} \right) \right\} \right\|$$

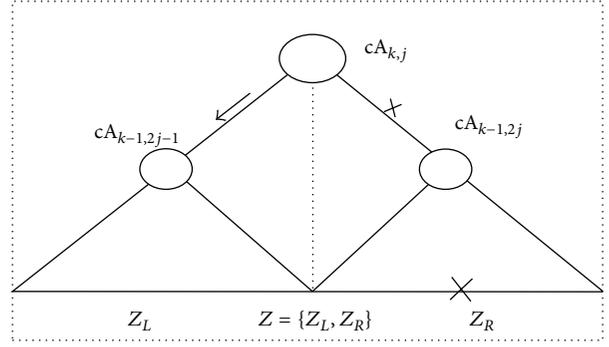


FIGURE 3: The scheme of Criterion 1 based on the statistic fluctuations within BSTcA. In terms of this criterion, the left or right-child node, that is, $cA_{k-1,2j-1}$ or $cA_{k-1,2j}$, might be selected to be involved in the current search path; meanwhile the remaining one is discarded. Thereafter, an optimal path towards the potential abrupt change in Z is expected to be obtained from BSTcA, after $\log_2 N$ binary search steps.

$$- \left\| \frac{1}{m} \left\{ \left(\sum_{Ra=c+1}^{rc} z_{Ra} \right) - \left(\sum_{Rb=rc+1}^b z_{Rb} \right) \right\} \right\|$$

$$= \left(\frac{nm}{n+m}\right)^{1/2} \left\| |N'(cD_{k,j})| - |M'(cD_{k-1,2j})| \right\|, \quad (11)$$

where $2 \leq k \leq \log_2 N$, $1 \leq j \leq N/2^k$; $a = 2^k(j-1)+1$, $b = 2^k j$, $c = 2^k(j-1)+2^{(k-1)}$; $lc = 2^k(j-1)+2^{(k-2)}$, $rc = c+2^{(k-2)}$; $n = 2^k$, $m = 2^{k-1}$; and $N' = (\sqrt{2}) \wedge k/2^k$, $M' = (\sqrt{2}) \wedge (k-1)/2^{(k-1)}$.

Suppose Criterion 1 is invalid as $(S_{k,j;L} = S_{k,j;R}) \parallel (\max(S_{k,j;L}, S_{k,j;R}) \leq C(\alpha))$ holds true; the second search criterion needs to be introduced in terms of the two variance fluctuation variables $D_{k,j;L}$ and $D_{k,j;R}$ as follows.

Criterion 2. Given two variance fluctuation variables $D_{k,j;L}$ and $D_{k,j;R}$ according to the two non-leaf child nodes $cD_{k-1,2j-1}$ and $cD_{k-1,2j}$ of the selected node $cD_{k,j}$ in BSTcD, and $2 \leq k \leq \log_2 N$,

- (a) if $(D_{k,j;L} > D_{k,j;R}) \wedge (D_{k,j;L} > C(\beta))$ holds true, then the left-child node $cA_{k-1,2j-1}$ in BSTcA is accordingly selected and involved into the current search path; meanwhile the right one is ignored;
- (b) if $(D_{k,j;L} < D_{k,j;R}) \wedge (D_{k,j;R} > C(\beta))$ holds true, then the right-child node $cA_{k-1,2j}$ in BSTcA is accordingly selected and involved into the current search path; meanwhile the left one is ignored.

Proof. Similarly, as illustrated in Figure 4, the satisfied $(D_{k,j;L} > D_{k,j;R})$ in Criterion 2 means that the variance fluctuations within Z_L are stronger than that one within Z_R , in terms of the definitions of $D_{k,j;L}$ and $D_{k,j;R}$ in (11). That is, a potential abrupt change might exist in Z_L with more probability than in Z_R , and vice versa. Meanwhile, if $D_{k,j;L} > C(\beta)$ holds true, then (H_1) in Hypothesis 2 is satisfied; that is, abrupt change occurs in Z_L , and vice versa, where $C(\beta)$ is the critical value predefined in an identical

between two adjacent segments Z_L and Z_R . Meanwhile, the search procedure is implemented from the root to non-leaf nodes in the last second level of BSTcA, in terms of Criteria 1 and 2. Then, the estimated CP can be obtained from the leaf nodes in BSTcA, by using Criterion 3. Thereafter, an optimal path towards a potential CP within Z is detected from BSTcA, after about $\log_2 N$ binary search steps.

3.3. Methods Compared with BSTKS. There are many methods proposed for abrupt change detection in time series, and the following are some typical methods, to evaluate the proposed BSTKS framework.

KS Statistic (see [31]). In this method, a diagnosed time series Z is divided into two adjacent segments $Z_L = \{z_1, z_2, \dots, z_m\}$ and $Z_R = \{z_{m+1}, z_{m+2}, \dots, z_N\}$, and then KS statistic is applied to calculate the statistic distance between Z_L and Z_R as

$$\begin{aligned} D_{mn}(x) &= \left(\frac{mn}{N}\right)^{1/2} \sup_{x \in R} |F_n(x) - G_m(x)| \\ &= \left(\frac{mn}{N}\right)^{1/2} \sup_{x \in R} \left| \sum_{R=m+1}^n I(z_R < x) - \sum_{L=1}^m I(z_L < x) \right|, \end{aligned} \quad (15)$$

where $F_n(x)$ and $G_m(x)$ stand for the e.c.d.f of Z_L , and Z_R , respectively; $N = m + n$, N is the total length of Z , and m refers to a current test position within Z .

t -Statistic (see [32]). t also known as Welch's t -test is used only when the two population variances are assumed different (the two sample sizes may or may not be equal) and hence must be estimated separately. Suppose a diagnosed Z is divided into $Z_L = \{z_1, z_2, \dots, z_m\}$ and $Z_R = \{z_{m+1}, z_{m+2}, \dots, z_N\}$. Then, t -statistic is calculated as

$$t = \frac{\bar{Z}_L - \bar{Z}_R}{S_{\bar{Z}_L - \bar{Z}_R}}, \quad S_{\bar{Z}_L - \bar{Z}_R} = \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}, \quad (16)$$

where \bar{Z}_L and \bar{Z}_R are the sample means of Z_L and Z_R , respectively; S is an unbiased estimator of the standard deviation, $N = m + n$, and m and n are the sizes of two segments Z_L and Z_R , respectively.

SSA (see [12, 13, 33]). In SSA method, a windowed portion is chosen within a time series $Z = \{z_1, z_2, \dots, z_N\}$, where N is large enough and a window width m and the lag parameter M are set such that $M = m/2$, $K = m - M + 1$. For each $n = 0, 1, \dots, N - m - M$, this method takes an interval of the time series $[n + 1, n + m]$ and then defines the $M \times K$ trajectory matrix X_n and describes the structure of the windowed portion as an L -dimensional subspace. If the structure changes further, it will not be well described by the computed subspace. Then, the distance between this subspace and the new trajectory vectors will increase; therefore, this increase will signal that an abrupt change occurs in Z .

4. Results and Discussion

In this section, the proposed BSTKS is evaluated on the synthetic time series and real EEG recordings with different

TABLE 1: The averaged results on four methods with datasets G_1 to G_7 .

	Time	Hit rate	Error	Accuracy	AUC
BSTKS	.0063	.4797	38.5268	.9018	.8922
KS	.3537	.0841	38.7321	.8804	.8984
t	1.0068	.0168	56.3036	.8878	.7960
SSA	1.5218	.0583	41.9464	.8762	.9941

size N . By comparing with existing KS, t , and SSA methods, the efficiency, sensitivity, and performance are analyzed in terms of the computation time, error and accuracy, hit rate, and AUC of ROC analyses. Furthermore, the novelty of our algorithm and necessity for real application are discussed in the following paragraphs.

4.1. CP Detection on Synthetic Time Series. In our simulations, some typical time series samples were derived from the normally distributed datasets (mean, $\mu = 0$, and standard deviation, $\text{sd} = 1$). Each diagnosed sample of size N is composed of a normal segment of size k and an adjacent segment of size $N - k$, in which the abnormal part is simulated by adding a constant variation v into the random numbers of size $N - k$. The proposed BSTKS and other three methods, namely, KS, t , and SSA, were tested, respectively, on 200 samples which were derived from each time series group G_i with $N_i = 2 \wedge (4 + i)$, $i = 1, 2, \dots, 7$, and $v_i = d(1 + \log_2(k - 4))$, where $k = \log_2(N_i)$ and $d = 1.0$. For each sample in G_i , a series of test positions were arranged by $CPK_j = j * (2 \wedge (k - 4))$, $k \geq 5$, and $j = 1, 2, \dots, 15$.

First, simulations were carried out according to different value of sample size N_i and test position CPK_j . The average analyses on four methods were listed in Table 1, and the results of simulations on datasets G_1 – G_7 were illustrated in Figure 5. In general, our BSTKS is the most promising with the shortest computation time, the highest hit rate, the smallest error, and the highest accuracy out of all four methods. Particularly, as sample size N increases from N_1 to N_7 , all four methods take longer time for bigger N , and BSTKS is always the fastest one. Meanwhile, BSTKS owns the highest level of hit rate against the low tracks of other three methods; and BSTKS is much more efficient with the smallest error and the highest accuracy, though all four methods tend to be better with N increasing. However, BSTKS has smaller AUC of ROC analyses, that is, bigger search space, than SSA and KS.

Second, simulations were carried out based on the datasets G_1 , G_4 , and G_7 . The proposed BSTKS and other three methods were tested according to the different value of variance $v = d(1 + \log_2(k - 4))$, $k = 5, 8, 11$, and $d = 0.5, 1.0, 2.0, 3.0$, respectively. The average results of four methods on G_1 , G_4 , and G_7 were summarized in Table 2, and the typical simulations were selected on G_4 and represented in Figure 6. Generally, when v gets larger, all four methods get better hit rate, accuracy, and AUC of ROC analysis, except for longer computation time for bigger size N . Compared with other three methods, the proposed BSTKS is more encouraging because of the shortest computation time, especially when N gets bigger, as well as the highest hit rate and accuracy,

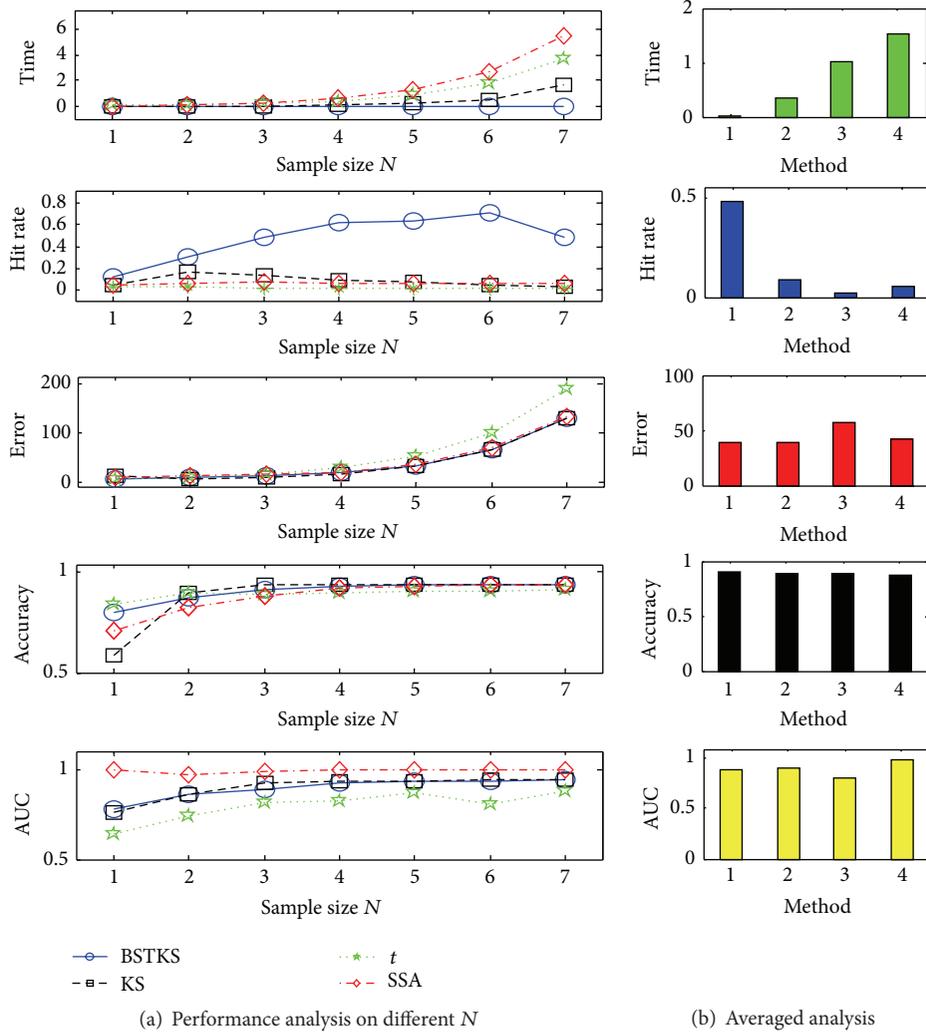


FIGURE 5: The simulations on G_1 to G_7 with size N from 2^5 to 2^{11} . (a) The results in terms of computation time, hit rate, error and accuracy, and AUC of ROC analyses. (b) The average analyses on BSTKS and other three methods. In the histograms, “1,” “2,” “3,” and “4” stand for BSTKS, KS, t , and SSA, respectively.

especially when N gets smaller. Moreover, the simulations on G_4 with different variance ν (Figure 6) explicitly illustrate that BSTKS has the best performance when ν gets larger, in terms of the shortest time and the biggest increase of the hit rate out of four methods. For the accuracy and AUC, both BSTKS and KS keep higher sensitivity than t and SSA, as ν increases from 0.5 to 3.0. Moreover, the simulations on G_1 and G_7 were omitted, because similar results can be obtained like G_4 above.

Third, simulations were implemented based on different CP test positions within G_1 and G_4 . The proposed BSTKS and other three methods were analyzed according to the different value of test position CPK and variance ν . The results of simulations on G_1 and G_4 were illustrated in Figure 7, and the results near the left and right boundaries in G_1 and G_4 were summarized in Table 3. In general, all four methods tend to be better, when N increases under a fixed ν , or when ν increases under a fixed N . Meanwhile, for test position CPK near the left and right boundaries, the

proposed BSTKS produces better performance than other three methods, because of the highest hit rate, the smallest error in all four methods, and higher accuracy and AUC than t and SSA. Moreover, the simulations on G_1 and G_4 near the left and right boundaries were illustrated in Figure 8 in detail. In terms of the distribution of estimated CP (e-CP), PDF of e-CP, and AUC of ROC analysis, these simulations indicate that BSTKS is more sensitive for both left and right boundaries than other three methods, especially when sample size N and variance ν get smaller.

Therefore, all simulation results above suggest that our proposed BSTKS is an encouraging and efficient method for abrupt change detection from the synthetic time series datasets, because of the shortest computation time, the highest hit rate, and accuracy out of four methods, especially for less significant statistic fluctuation when N gets smaller, as well as for less significant variance fluctuation when N gets bigger, and ν gets smaller.

TABLE 2: The summary of simulations according to different variances in G_1 , G_4 , and G_7 .

Items	Methods												
	$N = 2^5$				$N = 2^8$				$N = 2^{11}$				
	<i>BSTKS</i>	KS	<i>t</i>	SSA	<i>BSTKS</i>	KS	<i>t</i>	SSA	<i>BSTKS</i>	KS	<i>t</i>	SSA	
$d = 0.5$	Time	.018	.035	.227	.116	.029	.335	1.85	2.81	.060	6.96	172	24.6
	Hit rate	.046	.005	.010	.038	.093	.093	.005	.025	.106	.056	.006	.034
	Accuracy	.792	.515	.792	.748	.984	.995	.905	.899	.999	.999	.944	.884
	AUC	.694	.694	.644	.997	.954	.951	.978	.971	.983	1.00	.999	1.00
$d = 1.0$	Time	.018	.034	.223	.113	.031	.356	1.97	2.98	.061	7.09	18.1	26.5
	Hit rate	.086	.013	.007	.035	.041	.001	.099	.142	.135	.045	.000	.035
	Accuracy	.846	.552	.839	.756	.998	.998	.939	.974	.999	.999	.940	.986
	AUC	.695	.695	.851	.997	.993	.992	.997	.991	.992	1.00	.998	1.00
$d = 2.0$	Time	.018	.035	.229	.115	.031	.345	1.91	2.88	.065	7.60	19.7	29.1
	Hit rate	.165	.061	.007	.049	.181	.090	.000	.049	.167	.028	.000	.053
	Accuracy	.927	.737	.958	.765	.998	.997	.971	.983	.999	.999	.984	.997
	AUC	.754	.754	.908	.997	.998	.999	.996	1.00	.995	1.00	.999	1.00
$d = 3.0$	Time	.019	.037	.245	.125	.034	.382	2.08	3.28	.067	8.08	20.5	31.4
	Hit rate	.225	.086	.002	.037	.189	.084	0.00	.046	.169	.035	0.00	.045
	Accuracy	.942	.818	.952	.773	.997	.996	.986	.983	.999	.999	.991	.998
	AUC	.857	.938	.655	.997	1.00	.996	.728	.100	1.00	.999	.467	1.00

TABLE 3: The summary of simulations on G_1 and G_4 near the left and right boundaries according to different variance d .

Items	Methods																
	$N = 2^5, CPK = 8$				$N = 2^5, CPK = 24$				$N = 2^8, CPK = 16$				$N = 2^8, CPK = 240$				
	<i>BSTKS</i>	KS	<i>t</i>	SSA	<i>BSTKS</i>	KS	<i>t</i>	SSA	<i>BSTKS</i>	KS	<i>t</i>	SSA	<i>BSTKS</i>	KS	<i>t</i>	SSA	
$d = 0.5$	Hit rate	.200	.055	.036	.160	.215	.060	.010	0.0	.230	.160	0.0	0.0	.265	.165	0.0	.065
	Error	2	20	8	3	3	7	8	12	24	1	134	49	28	21	102	33
	Accuracy	.937	.375	.750	.906	.906	.781	.750	.625	.960	.996	.476	.808	.891	.918	.601	.871
	AUC	.635	.963	.641	.978	.656	.987	.599	.978	.599	.946	.780	.797	.750	.884	.564	.797
$d = 1.0$	Hit rate	.515	.160	.040	.295	.485	.190	.005	0.0	.490	.195	0.0	0.0	.535	.175	.025	.100
	Error	0	10	4	2	0	3	8	12	1	0	112	10	0	1	90	1
	Accuracy	1.0	.687	.875	.937	1.0	.906	.750	.625	.996	1.0	.562	.960	1.00	.996	.648	.996
	AUC	.831	.883	.641	.978	.922	.927	.599	.978	.864	.986	.780	.829	.999	.988	.657	.911
$d = 2.0$	Hit rate	.655	.240	0.0	.220	.725	.175	0.0	0.0	.510	.160	0.0	0.0	.530	.150	.005	.065
	Error	0	6	1	2	0	1	4	12	0	0	107	5	0	1	36	4
	Accuracy	1.00	.812	.968	.937	1.0	.968	.875	.625	1.0	1.0	.582	.980	1.0	.996	.859	.984
	AUC	.976	.979	.770	.978	.938	.875	.808	.978	.978	.985	.780	.999	1.0	.990	.752	.995
$d = 3.0$	Hit rate	.715	.210	0.0	.265	.730	.215	0.0	0.0	.530	.195	0.0	0.0	.545	.145	0.0	.060
	Error	0	6	1	2	0	1	1	13	0	0	119	5	0	2	11	4
	Accuracy	1.0	.812	.968	.937	1.0	.968	.968	.593	1.0	1.0	.535	.980	1.0	.992	.957	.984
	AUC	.999	.960	.770	.978	.996	.822	.808	.978	.999	.940	.780	.999	1.0	.990	.752	.998

4.2. *Abrupt Change Analyses on EEG Recordings.* To verify the proposed method further, we take some representative samples from the CHBMIT Scalp EEG Database. In the PhysioBank platform, the CHBMIT Scalp EEG Database (CHBMIT) was collected at the Children's Hospital Boston; it consists of EEG recordings from pediatric subjects with

intractable seizures [34, 35]. In this CHBMIT EEG database, some subjects were monitored up to several days after withdrawal of antiseizure medication in order to characterize their seizures and assess their candidacy for surgical intervention. Based on these EEG recordings in the CHBMIT EEG database, as well as some existing experiments in [36–39],

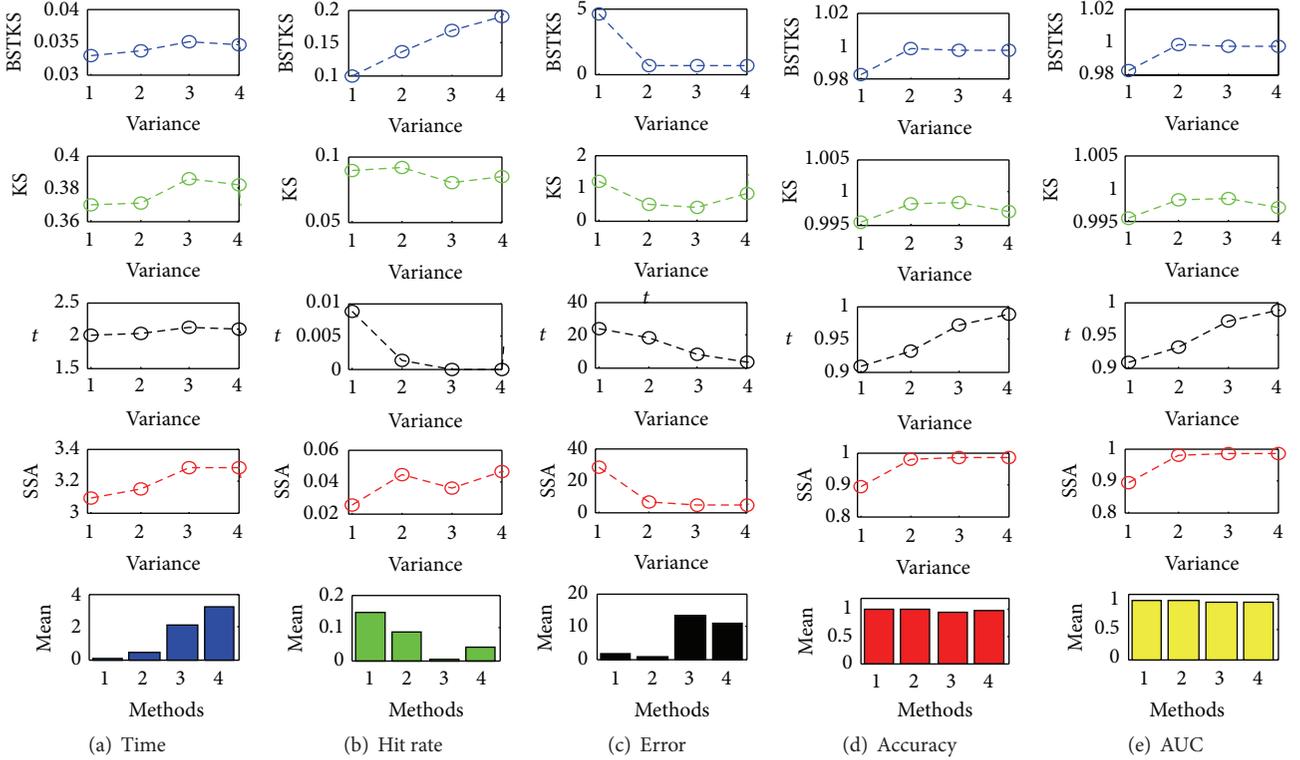


FIGURE 6: The simulations on 200 samples in G_4 with different variances. Under different variances ν from 0.5 to 3.0, (a) the computation time, (b) the hit rate, (c) the error, (d) the accuracy, and (e) the AUC of ROC analysis, for BSTKS, KS, t , and SSA, respectively. In all “mean” histograms, “1,” “2,” “3,” and “4” in x -axis stand for BSTKS, KS, t , and SSA methods, respectively.

the proposed BSTKS and other three methods were tested according to different value of test position CPK and sample size N .

First, a diagnosed EEG sample $Z = [Z_L, Z_R]$ was assembled from two significantly different segments, in which $Z_L = \{z_1, \dots, z_{CPK}\}$ and $Z_R = \{z_{CPK+1}, \dots, z_N\}$ were derived from `chb01_04_edfm` and `chb01_05_edfm`, respectively. Then, BSTKS and other three methods were tested on the assembled EEG recordings Z_1-Z_8 , respectively, according to the different value of assigned test position CPK and sample size N . The results of abrupt change detection on these assembled EEG samples were illustrated in Figure 8 and summarized in Table 4. Generally speaking, all four methods can roughly estimate the assigned test position from each assembled EEG recording and then divide it into two adjacent segments Z_L and Z_R . It is worth stressing that the proposed BSTKS can discern the different EEG segments accurately with the smallest error and the highest accuracy out of four methods. Also, BSTKS is the most efficient and encouraging with the shortest time in all four methods.

Moreover, for CPK near the left and right boundaries in Z_1-Z_8 , BSTKS has much better sensitivity than other KS, t , and SSA methods because of the smallest error and the highest accuracy, especially for less statistic fluctuation when N gets smaller, as well as less significant variance fluctuation when N gets bigger. Supposing the assembled EEG sample indicates that a sharp transition of one’s mental situation

occurs before and after a sudden attack or acute stimulation, it is meaningful to estimate the location of the abrupt change and the maximal difference of data distribution exists between two adjacent EEG segments. These experiments above suggest that the proposed BSTKS can successfully detect the change position where a sudden change occurs under a potential mental shock, more quickly and efficiently than KS, t , and SSA methods.

Second, the original EEG samples Z_1-Z_6 were selected directly from different recordings in the `chb01_05_edfm`; then the proposed BSTKS and other three methods were tested according to different sample size N . Because the distance of e.c.d.f (V.e.c.d.f) can partly reflect the data fluctuation between two adjacent EEG segments, we use this V.e.c.d.f variable to distinguish different performance of BSTKS and other three methods. The results of abrupt change analyses on these original EEG recordings were shown in Figure 9 and summarized in Table 5. For all methods above, they can estimate an abrupt change from each of these original EEG samples Z_1-Z_6 and then divide it into two adjacent EEG segments. Compared with other three methods, the proposed BSTKS is encouraging for the shortest time out of four methods. Moreover, BSTKS has bigger V.e.c.d.f than t and SSA, which means that it can more reasonably distinguish two adjacent EEG segments with different state of mental health. Although KS has the biggest V.e.c.d.f in all four methods, it takes much more search time than BSTKS, especially when

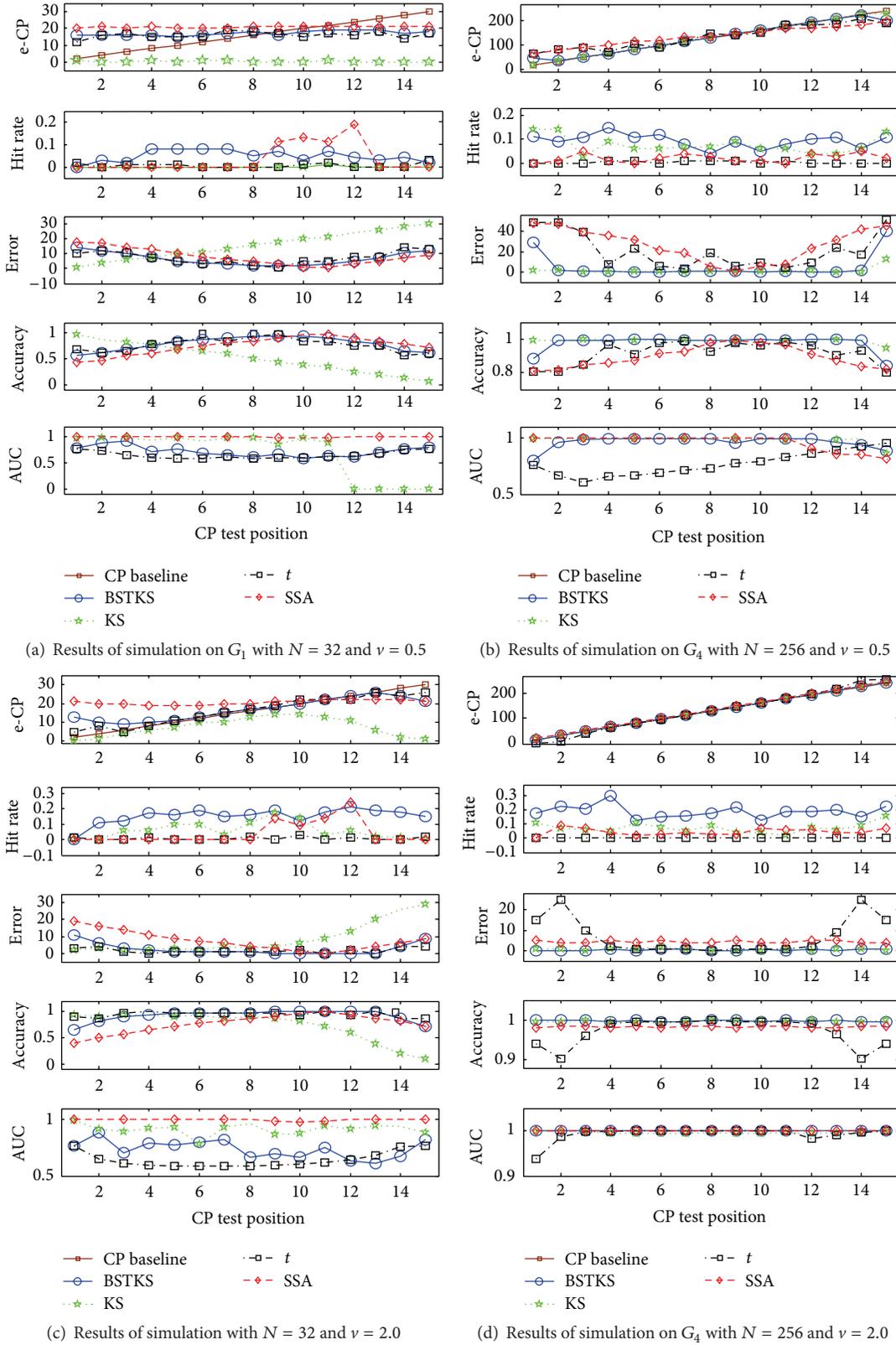


FIGURE 7: The simulations on G_1 and G_4 according to the different variance ν and test position CPK . The results were shown in (a) G_1 with $N = 32$ and $\nu = 0.5$, (b) G_4 with $N = 256$ and $\nu = 0.5$, (c) G_1 with $N = 32$ and $\nu = 2.0$, and (d) G_4 with $N = 256$ and $\nu = 2.0$, in terms of e-CP, hit rate, error, accuracy, and AUC, respectively.

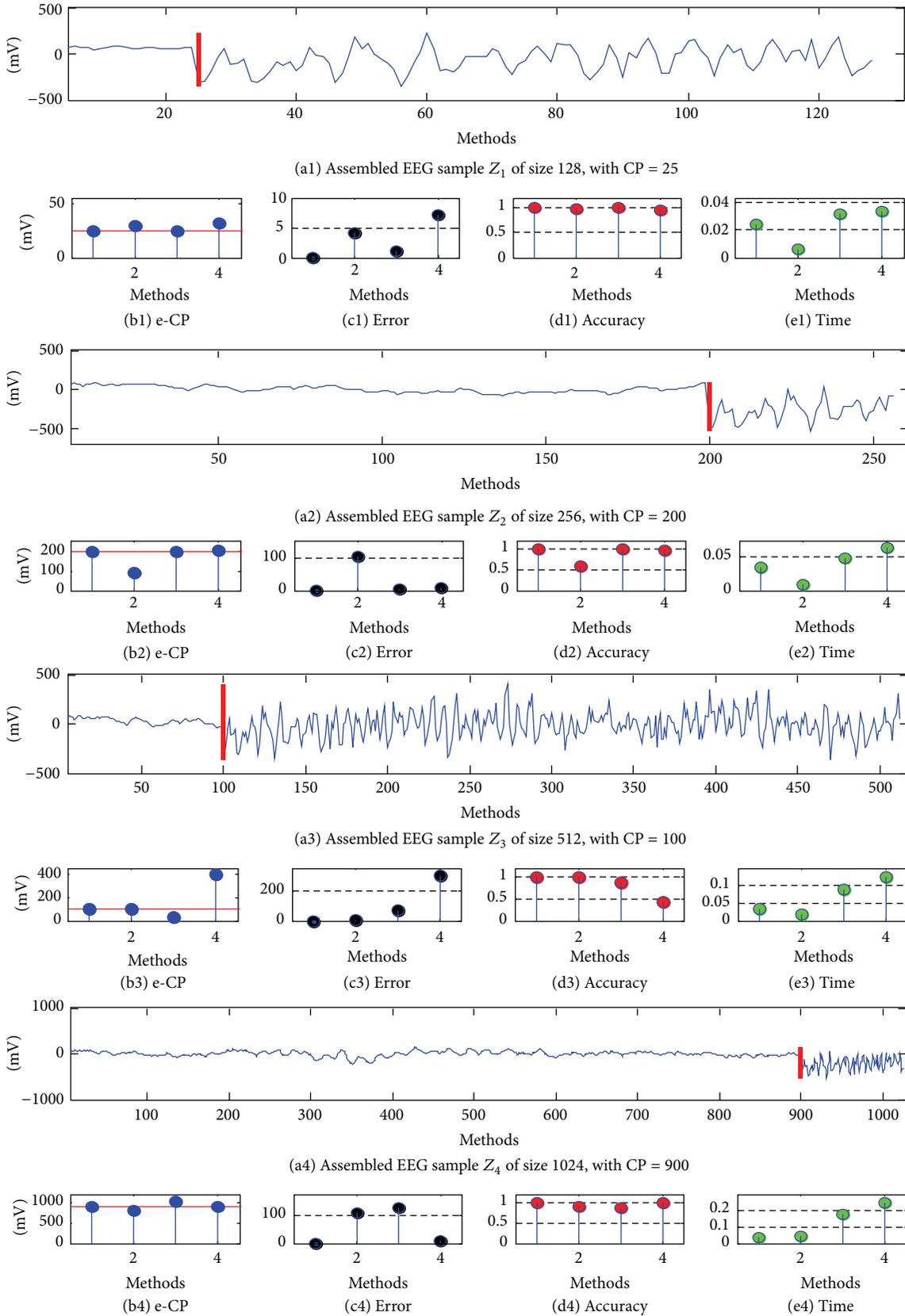


FIGURE 8: The results of CP detection on the assembled EEG samples Z_1 - Z_4 with different value of sample size N and test position CPK . For Z_1 - Z_4 with different N from 2^7 to 2^{10} , (a1-a4) the assembled EEG samples Z_1 - Z_4 with the assigned test position CPK , (b1-b4) the e-CP, (c1-c4) the error of e-CP, (d1-d4) the accuracy of e-CP, and (e1-e4) the computation time. In the x-axis of (b-e), the methods "1," "2," "3," and "4" stand for BSTKS, KS, t , and SSA, respectively.

TABLE 4: The summary of abrupt change detection on Z_1-Z_8 .

M	N	Z								Mean
		2^7	2^8	2^9	2^{10}	2^7	2^8	2^9	2^{10}	
	CPK	25	50	100	200	100	200	400	900	
	BSTKS	25	50	100	276	100	200	376	900	NA
e-CP	KS	29	36	95	206	34	92	301	795	NA
	<i>t</i>	24	255	31	1023	31	199	33	1023	NA
	SSA	32	55	398	1007	106	208	500	907	NA
	BSTKS	0	0	0	76	0	0	24	0	12.5
Err	KS	4	14	5	6	66	108	99	105	50.9
	<i>t</i>	1	205	69	823	69	1	367	123	207.3
	SSA	7	5	298	807	6	8	100	7	154.8
	BSTKS	1.0	1.0	1.0	.93	1.0	1.0	.95	1.0	.98
Acc	KS	.97	.94	.99	.99	.48	.57	.81	.89	.83
	<i>t</i>	.99	.20	.86	.20	.46	.99	.28	.88	.61
	SSA	.94	.97	.42	.21	.94	.97	.80	.99	.78
	BSTKS	.023	.031	0.034	.036	.028	.033	.035	.038	.032
Time	KS	.019	.021	.038	.049	.020	.029	.039	.052	.033
	<i>t</i>	.03	.063	.088	.170	.031	.050	.081	.174	.086
	SSA	.037	.071	.126	.239	.035	.065	.118	.245	.117
	BSTKS	.023	.031	0.034	.036	.028	.033	.035	.038	.032

TABLE 5: The summary of CP detection from the original EEG samples Z_1-Z_6 .

M		N						Mean
		2^9	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	
e-CP	BSTKS	328	316	1286	2633	4352	6224	NA
	KS	348	317	1342	2252	4673	5947	NA
	<i>t</i>	511	314	17	4095	10	16383	NA
	SSA	426	854	90	2634	408	11271	NA
V.e.c.d.f	BSTKS	.0649	.2608	.2822	.0997	.1318	.0388	.1464
	KS	.4603	.3829	.4407	.3050	.3325	.2234	.3574
	<i>t</i>	0	.1257	.5384	0	0	0	.1106
	SSA	.1368	.0850	.1260	.0745	.0212	.0012	.0741
Time	BSTKS	.020	.020	.024	.030	.019	.0320	.0241
	KS	.016	.041	.112	.466	1.461	5.638	1.289
	<i>t</i>	.072	.137	.281	.913	1.726	4.709	1.306
	SSA	.107	.209	.415	1.103	1.769	3.548	1.192

is detected from the root to leaf nodes of two BSTs in terms of three search criteria. The novelty of the proposed method is addressed by comparing with other KS, *t*, and SSA methods, and simulations on the synthetic time series indicate that the proposed BSTKS is more efficient due to the shortest time, the highest hit rate, and the smallest error and highest accuracy out of four methods. Meanwhile, BSTKS has better sensitivity than KS near the left and right boundaries, because of shorter search time, higher hit rate, and bigger AUC, especially when sample size N gets smaller with less significant statistic fluctuation. In addition, the necessity of the proposed method in the real domain is analyzed on real EEG recordings, and the results indicate

that the proposed method can successfully discern an abrupt change and then obviously distinguish two adjacent EEG segments from the real EEG recordings. Through inspecting the significant fluctuation between adjacent segments signals, it is encouraging further for useful information inspection on all kinds of physiological and psychological time series signals. In a word, our BSTKS is a novel, efficient, and promising method for abrupt change analysis, and it is very helpful for useful information inspection on all kinds of real time series with different scales.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors would like to thank Professor Mohan Karunanithi in the Australia e-Health Research Centre, CSIRO Computation Informatics, for his assistance, support, and advice for this paper. This paper is supported by National Natural Science Foundation of China (no. 13K10414 and no. 61104154) and Specialized Research Fund for Natural Science Foundation of Shanghai (nos. 16ZR1401300 and 16ZR1401200).

References

- [1] R. J. Bolton and D. J. Hand, "Statistical fraud detection: a review," *Statistical Science*, vol. 17, no. 3, pp. 235–249, 2002.
- [2] B. E. Brodsky and B. S. Darkhovsky, *Nonparametric Methods in Change-point Problems*, vol. 243 of *Mathematics and its Applications*, Springer, Dordrecht, The Netherlands, 1993.

- [3] B. E. Brodsky and B. S. Darkhovsky, *Non-Parametric Statistical Diagnosis: Problems and Methods*, vol. 509 of *Mathematics and Its Applications*, Springer, Dordrecht, The Netherlands, 2000.
- [4] K. Yamanishi, J.-I. Takeuchi, G. Williams, and P. Milne, "Online unsupervised outlier detection using finite mixtures with discounting learning algorithms," *Data Mining and Knowledge Discovery*, vol. 8, no. 3, pp. 275–300, 2004.
- [5] U. Murad and G. Pinkas, "Unsupervised profiling for identifying superimposed fraud," in *Principles of Data Mining and Knowledge Discovery*, pp. 251–261, Springer, 1999.
- [6] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Lu, "A review and comparison of changepoint detection techniques for climate data," *Journal of Applied Meteorology and Climatology*, vol. 46, no. 6, pp. 900–915, 2007.
- [7] Y. Wang, C. Wu, Z. Ji, B. Wang, and Y. Liang, "Non-parametric change-point method for differential gene expression detection," *PLoS ONE*, vol. 6, no. 5, Article ID e20060, 2011.
- [8] M. E. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, New York, NY, USA, 1993.
- [9] J.-P. Qi, Q. Zhang, Y. Zhu, and J. Qi, "A novel method for fast change-point detection on simulated time series and electrocardiogram data," *PLoS ONE*, vol. 9, no. 4, Article ID e93365, 2014.
- [10] P. Yiou, E. Baert, and M.-F. Loutre, "Spectral analysis of climate data," *Surveys in Geophysics*, vol. 17, no. 6, pp. 619–663, 1996.
- [11] R. Vautard and M. Ghil, "Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series," *Physica D: Nonlinear Phenomena*, vol. 35, no. 3, pp. 395–424, 1989.
- [12] L. Vaisman, J. Zariffa, and M. R. Popovic, "Application of singular spectrum-based change-point analysis to EMG-onset detection," *Journal of Electromyography and Kinesiology*, vol. 20, no. 4, pp. 750–760, 2010.
- [13] V. Moskvina and A. Zhigljavsky, "An algorithm based on singular spectrum analysis for change-point detection," *Communications in Statistics—Simulation and Computation*, vol. 32, no. 2, pp. 319–352, 2003.
- [14] F. Gustafsson, *Adaptive Filtering and Change Detection*, vol. 1, John Wiley & Sons, New York, NY, USA, 2000.
- [15] V. Alarcon-Aquino and J. A. Barria, "Anomaly detection in communication networks using wavelets," *IEE Proceedings: Communications*, vol. 148, no. 6, pp. 355–362, 2001.
- [16] S. Fremdt, "Page's sequential procedure for change-point detection in time series regression," *Statistics*, vol. 49, no. 1, pp. 128–155, 2015.
- [17] M. Priyadarshana, T. Polushina, and G. Sofronov, "Hybrid algorithms for multiple change-point detection in biological sequences," *Advances in Experimental Medicine and Biology*, vol. 823, pp. 41–61, 2015.
- [18] K. Nosek and Z. Szkutnik, "Change-point detection in a shape-restricted regression model," *Statistics*, vol. 48, no. 3, pp. 641–656, 2014.
- [19] R. Vautard, P. Yiou, and M. Ghil, "Singular-spectrum analysis: a toolkit for short, noisy chaotic signals," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1–4, pp. 95–126, 1992.
- [20] M. Khalil and J. Duchêne, "Detection and classification of multiple events in piecewise stationary signals: comparison between autoregressive and multiscale approaches," *Signal Processing*, vol. 75, no. 3, pp. 239–251, 1999.
- [21] M. Kobayashi, "Wavelets and their applications in industry," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 47, no. 3, pp. 1749–1760, 2001.
- [22] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, vol. 4 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge, UK, 2006.
- [23] M. Salam and D. Mohamad, "Segmentation of Malay Syllables in connected digit speech using statistical approach," *International Journal of Computer Science and Security*, vol. 2, pp. 23–33, 2008.
- [24] J. Qi, Y. Ding, Y. Zhu, and Y. Wu, "Kinetic theory approach to modeling of cellular repair mechanisms under genome stress," *PLoS ONE*, vol. 6, no. 8, Article ID e22228, 2011.
- [25] V. S. Tseng, C.-H. Chen, C.-H. Chen, and T.-P. Hong, "Segmentation of time series by the clustering and genetic algorithms," in *Proceedings of the 6th IEEE International Conference on Data Mining Workshops (ICDM '06)*, pp. 443–447, Hong Kong, December 2006.
- [26] V. Alarcon-Aquino and J. Barria, "Change detection in time series using the maximal overlap discrete wavelet transform," *Latin American Applied Research*, vol. 39, pp. 145–152, 2009.
- [27] S. Chen and C. Liu, "Eye detection using discriminatory Haar features and a new efficient SVM," *Image and Vision Computing*, vol. 33, pp. 68–77, 2015.
- [28] B. S. Darkhovski, *Nonparametric Methods in Change-Point Problems: A General Approach and some Concrete Algorithms*, vol. 23 of *Lecture Notes—Monograph Series*, 1994.
- [29] J. S. Walker, *A Primer on Wavelets and Their Scientific Applications*, CRC Press, New York, NY, USA, 2002.
- [30] P. Fryzlewicz and S. S. Rao, "Multiple-change-point detection for auto-regressive conditional heteroscedastic processes," *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, vol. 76, no. 5, pp. 903–924, 2014.
- [31] R. Simard and P. L'Ecuyer, "Computing the two-sided Kolmogorov-Smirnov distribution," *Journal of Statistical Software*, vol. 39, no. 11, pp. 1–18, 2011.
- [32] T. Sørli, R. Tibshirani, J. Parker et al., "Repeated observation of breast tumor subtypes in independent gene expression data sets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 14, pp. 8418–8423, 2003.
- [33] H. Hassani and A. Zhigljavsky, "Singular spectrum analysis: methodology and application to economics data," *Journal of Systems Science and Complexity*, vol. 22, no. 3, pp. 372–394, 2009.
- [34] A. L. Goldberger, L. A. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [35] A. H. Shueb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*, Massachusetts Institute of Technology, Cambridge, Mass, USA, 2009.
- [36] Z. H. Xia, X. H. Wang, X. M. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 340–352, 2016.
- [37] S. D. Xie and Y. X. Wang, "Construction of tree network with limited delivery latency in homogeneous wireless sensor networks," *Wireless Personal Communications*, vol. 78, no. 1, pp. 231–246, 2014.

- [38] J. Li, X. Li, B. Yang, and X. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 507–518, 2015.
- [39] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 7, pp. 1403–1416, 2015.

Research Article

Adaptive Shape Kernel-Based Mean Shift Tracker in Robot Vision System

Chunmei Liu,¹ Yirui Wang,² and Shangce Gao³

¹Department of Computer Science and Technology, Tongji University, Shanghai 201804, China

²The College of Information Sciences and Technology, Donghua University, Shanghai 201620, China

³Faculty of Engineering, University of Toyama, Toyama 930-8555, Japan

Correspondence should be addressed to Chunmei Liu; chunmei.liu@tongji.edu.cn

Received 27 November 2015; Revised 25 April 2016; Accepted 10 May 2016

Academic Editor: Michele Migliore

Copyright © 2016 Chunmei Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes an adaptive shape kernel-based mean shift tracker using a single static camera for the robot vision system. The question that we address in this paper is how to construct such a kernel shape that is adaptive to the object shape. We perform nonlinear manifold learning technique to obtain the low-dimensional shape space which is trained by training data with the same view as the tracking video. The proposed kernel searches the shape in the low-dimensional shape space obtained by nonlinear manifold learning technique and constructs the adaptive kernel shape in the high-dimensional shape space. It can improve mean shift tracker performance to track object position and object contour and avoid the background clutter. In the experimental part, we take the walking human as example to validate that our method is accurate and robust to track human position and describe human contour.

1. Introduction

Object tracking is one major component in many robot domains as it directly affects the whole processing. Although significant amount of work has been done on object tracking in the literature, some difficulties still exist in objects tracking, such as nonrigid object structures, object occlusion, multiple connected objects, low contrast to the background, object scale variation, and complex object motion.

A large number of approaches have focused on object tracking. These works can be divided into three groups: point tracking, silhouette tracking, and kernel tracking. In [1], a point-based tracking approach is proposed by corresponding detected object points across frames. For tracking small objects, it is effective as it can use a single point to represent small objects, but for large objects it must use multiple points to represent large objects, which unavoidably leads to misdetection and occlusion. In comparison with the point-based tracking, a silhouette-based method focuses on an object shape description for tracking [2], which can flexibly handle a variety of object shapes. The kernel-based tracking approaches utilize a model region to represent the object in order to estimate object motion.

The kernel-based tracking approaches have many kinds of trackers. An early kernel-based visual tracker is the CAMSHIFT [3], which tracks the human faces by assigning each pixel a positive weight to find the location of a rectangular window in which the total weight of the pixels in the window is maximal. A kernel-based tracker [4] is proposed by minimizing a Bhattacharyya coefficient-based distance between the reference color distribution and the target's color distribution. An extended Kalman filter is suggested by updating the reference color histogram in [5]. The tracker in [6] was used by a mixture-of-Gaussians color model of the target. The affine object tracker is also performed in [7]. A kind of affine kernel-based trackers is presented by combining color-related kernel and boundary-related kernel to improve the tracking accuracy [8]. A kernel-based tracker is proposed by calculating the Gaussian pyramids of the images and applying mean shift algorithm at each pyramid level for the target [9]. Considering the computational complexity, it is proposed to use segmentation technique instead of using mean shift tracking algorithm [10]. Bolme et al. [11] exploit correlation filters based trackers to track visual objects.

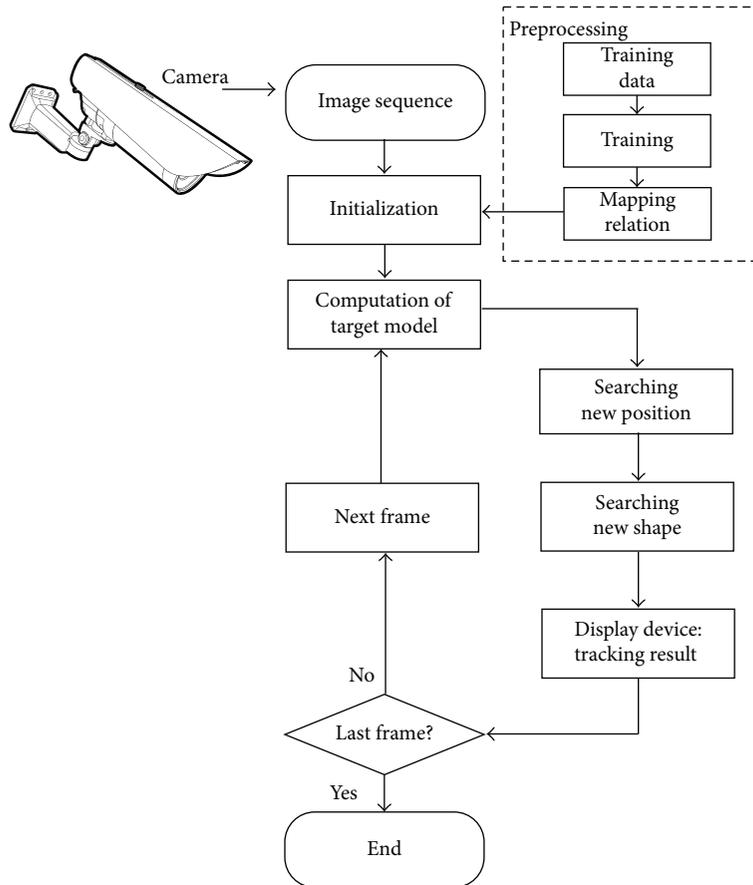


FIGURE 1: The flow chart of the proposed approach.

Mean shift algorithm is an efficient kernel-based tracking approach which is a kind of nonparametric methods for seeking the nearest mode of a point sample distribution based on kernel density estimation [4, 12]. It is popular and widely applied in object tracking as it has many merits of low computation, easy implementation, real time response, and robust tracking performance.

It is challenging to have an ideal kernel with the shape adaptive to the object where no background points reside, especially the arbitrary shape of the nonrigid object. For the nonrigid object, it is very difficult to make the kernel shape the same as the object shape as the nonrigid object shape always varies. Some kernel-based methods are proposed to adapt to arbitrarily shaped object and overcome the background disturbance [13–17]. An asymmetric kernel mean shift algorithm is proposed to estimate object location, orientation, and scale by Yilmaz [14]. It introduces an implicit level set of functions to reduce the estimation bias and improve the density estimation process. The detected object mask is presented to construct a kernel by Yi et al. [16]. It is robust to background clutter and tracks the object very accurately if the object is detected accurately. So it depends heavily on the detection results. A GMM-SAMT algorithm is applied to achieve an asymmetric shape adapted kernel by Quast and Kaup [13]. An asymmetric kernel-based visual tracker is proposed by Leichter et al. [17]. It uses the target's color PDF to enhance the tracker's robustness.

The work in this paper presents an adaptive shape kernel-based mean shift tracker to improve the position estimation and track the object contour under the environment captured by a static camera. The first contribution is the proposal of the adaptive shape kernel whose shape is reconstructed from the low-dimensional shape space to the high-dimensional shape space. Figure 1 shows the whole algorithm architecture. In the preprocessing stage, we perform nonlinear manifold learning technique to obtain the mapping relation from the high-dimensional shape space to the low-dimensional shape space, which is trained by training dataset with the same view as the tracking video sequence. The second contribution of our paper is to work on a combination of the adaptive shape with color feature to describe the object's appearance. In contrast with the symmetric constant kernel used in the traditional tracker, it can better adapt to the object shape change to reduce the estimation error and improve the density estimation process. The whole processing is performed to find the right shape in the shape space and find the right position in position space for each frame in the video. Experiments demonstrate that this kind of tracker can outperform the traditional tracker significantly. Our method is accurate and robust to track object position and describe object contour especially when target shape deformation and background clutter occur.

The remainder of this paper is organized as follows. The traditional mean shift algorithm is reviewed in Section 2.

Section 3 explains how to construct an adaptive shape kernel from the embedding low-dimensional shape space to the high-dimensional shape space. Section 4 introduces the process of the adaptive kernel shaped mean shift algorithm. In Section 5, the experimental studies are presented to prove the advantages of the proposed algorithm. Finally, Section 6 summarizes the main contributions of the paper together with discussions on some opening issues.

2. Mean Shift Tracking

Mean shift is a robust statistical algorithm, which applies a kernel density function in the new image based on the color histogram of the object in the previous image, and uses mean shift to find the maxima of a kernel density near the object's old position iteratively. It works with a search window that is positioned over a kernel density distribution. Within this search window, we compute the mean shift vector Δx to evaluate the displacement of the object centroid. So the local maxima of the kernel density can be obtained by moving the searching window around the original position.

Firstly, we move the initial object position \hat{x}_{old} to a new position \hat{x}_{new} and repeat the average computation iteratively until the local maximum can be found. And the new position is updated: $\hat{x}_{new} = \hat{x}_{old} + \Delta x$. The mean shift vector is computed as follows:

$$\Delta x = \frac{\sum K(x - \hat{x}_{old}) w(x) (x - \hat{x}_{old})}{\sum K(x - \hat{x}_{old}) w(x)}, \quad (1)$$

$$w(x) = \sqrt{\frac{h_t(I(x))}{h_c(I(x))}},$$

where h_t and h_c are the color distribution functions generated from the target model and the candidate object region, respectively.

3. Adaptive Kernel Shape

Kernel shape is an important parameter of the mean shift algorithm, which decides which points participate in the computation. As all points in the kernel contribute to finding the local extrema, kernel shape plays an important part in mean shifting. The traditional mean shift method applies a symmetric kernel, such as a circle or an ellipse. This kind of kernel shape cannot match up with the object shape. The kernel window unavoidably covers many background points as well as the foreground object points. These background points inside the kernel window work as a part of the object. Therefore, the target tracking easily shifts to the wrong position as background clutter.

We consider making the kernel shape consistent with the object shape. However, it is not easy to describe the object shape accurately, especially for nonrigid object. During nonrigid object tracking, nonrigid object shape always varies. It becomes very crucial to keep the consistency between the kernel shape and the nonrigid object shape. In this paper, a kind of the adaptive kernel shape is proposed to describe the

nonrigid object shape, which can avoid the disturbance of background points in the searching kernel window.

We use Γ to define the kernel shape parameter set. The transformed kernel parameters can be described as

$$P(x, y, \sigma, \theta, \Gamma), \quad (2)$$

where (x, y) are the center coordinates of the tracking window position. σ is the scale dimension, θ is the orientation variation parameter, and Γ is the kernel shape parameter set. Here, we focus on the importance of the kernel shape parameters Γ . So the scale parameter σ and the orientation variation parameter θ are not considered in this paper.

In the following discussion, we provide the details on how to learn the adaptive kernel shape parameter set Γ . It is performed by three steps. Firstly, an object shape is represented by sampling points on the object contour. Secondly, we apply a nonlinear dimensionality reduction to transform the high-dimensional shape space into the low-dimensional shape space. This step includes two stages: training and tracking. In training stage, we use the training data which are required to have the similar view and the same object movement with the tracking object to obtain the low-dimensional shape space. Finally, in the tracking stage, the kernel shape of mean shift is reconstructed from the low-dimensional shape space to original shape space. In the following discussion, we take the walking human as example to illustrate the whole processing.

3.1. Shape Representation. It is necessary to represent precisely the object shape when tracking object. We can use some sampling points on the object contour to represent the object shape. The object has some key positions. These key position points are convex. These convex points play a significant role in the shape representation, which can be obtained by the projection of the object silhouette in some directions. These convex points can separate the shape contour into some parts. Each part has few singular points. It is flat. So we can sample uniformly the points on each part and use these sampling points to represent the object shape.

We take the walking human as example to illustrate how to handle the whole processing. For the walking human, human shape varies in a gait cycle in a camera's viewpoint. We represent each shape instance by sampling the points on the human contour. There are three key positions of human body: head point, left foot point, and right foot point to describe human shape. These three key positions are taken as reference points to separate human shape into three parts: head-right foot, right foot-left foot, and left foot-head.

In order to find the head position, we projected the object silhouette onto the vertical axis (Figure 2). Head point is the top position of this projection. In the same way, the right part and the left part of the object silhouette are, respectively, projected onto the vertical axis. Left foot point and right foot point are, respectively, the bottom positions of projection.

After we obtain these three key points, we can separate the shape boundary into three parts: head-right foot, right foot-left foot, and left foot-head. For each part, 20 consecutive

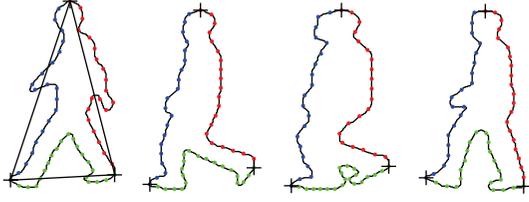


FIGURE 2: The shape points of silhouettes.

points are sampled. Thus, 60 points are employed to represent an instance shape.

3.2. Nonlinear Embedding. After sampling, the object shape can be represented in 60-dimensional space. However, it is difficult to search the right kernel shape in 60-dimensional shape space for the current instance. We provide a global geometric transformation for the shape instance between the low-dimensional shape space and the high-dimensional shape space. In this low-dimensional shape space, mean shift algorithm can search the right kernel shape window for the current instance.

As all points in the kernel contribute to finding the local extrema, kernel shape plays an important part in mean shifting. The traditional mean shift method applies a symmetric kernel, such as a circle or an ellipse. This kind of kernel shape cannot match up with the object shape. The kernel window unavoidably covers many background points as well as the foreground object points. These background points inside the kernel window work as a part of the object. Therefore, the target tracking easily shifts to the wrong position as background clutter.

For human walking, human shape has one cycle of the gait variation and suffers to deformation and self-occlusion which lead to the shape points lying on a nonlinear, twisted manifold. Isomap is a nonlinear dimensionality reduction method. It is one of the low-dimensional embedding methods, which provides a simple method for estimating the intrinsic geometry of a data manifold. It is highly efficient and generally applicable to a dimensionalities reduction. Here, we employ Isomap to embed 60 shape points in a 2-dimensional space.

Given a set of N_F shape observations, $S = \{s_1 \cdots s_{N_F}\}$, their corresponding embedding shapes are $T = \{t_1 \cdots t_{N_F}\}$. Thus, each shape s_i can be represented as t_i in low-dimensional space. Figure 3 shows an example of embedding walking cycles on a fixed view. We use a two-dimensional embedding space to describe the walking shape sequence in one cycle. As can be noticed, several embedding shape frames can represent a walking cycle.

3.3. Shape Reconstruction. When we obtain the low-dimensional shape space, it is possible to search the right kernel shape in the shape space. The searching of the kernel shape can be implemented in the low-dimensional shape space. However, if the right kernel shape was found, it is necessary to map the kernel shape from the low-dimensional space to the original high-dimensional space.

Here, we can finish searching the kernel shape in the low-dimensional shape space. Similar to the mapping from the low-dimensional shape space into the original shape space, RBF is applied to reconstruct the shape in original shape space.

4. Adaptive Kernel-Based Mean Shift

As it is difficult to finish searching the right shape in high-dimensional shape space for tracking, kernel shape parameter is generally ignored. The points in the kernel participate in the computation to find the local extrema, so the kernel shape plays an important role in mean shift algorithm. An ideal kernel should have the same shape with tracking object without background clutter. However, as the dimensionality of shape space is very high, it is difficult to finish the shape searching in such high-dimensional space. Particularly, the shape of nonrigid object is a kind of arbitrary shape. It increases the difficulties of the kernel shape searching. In this paper, an adaptive shape based kernel is proposed for mean shift tracker. Firstly, the embedding low-dimensional shape space is found after training samples. In this embedding low-dimensional shape space, it is applied to finish searching the kernel shape for mean shift tracker. When the searching shape in low-dimensional space is identified, it is necessary to reconstruct it from low-dimensional space to original shape space. As mentioned above, RBF is applied to reconstruct the shape in original shape space.

We use Γ to define the kernel shape parameter set. Γ_h ($S = \{s_1 \cdots s_{N_F}\}$) is the kernel shape parameter set in original shape space. In this paper, Γ_l ($T = \{t_1 \cdots t_{N_F}\}$) is the shape parameter set in the embedding low-dimensional shape space. In the following discussion, we provide details on how to design adaptive kernel mean shift tracker, which combines adaptive kernel shape and color features to describe the object's appearance.

4.1. Color Feature. The proposed mean shift combines the reconstructed shape kernel and color features to find the most probable position of the target object through iteration. The detailed tracking process is implemented as follows.

The color histogram features can provide a better discriminating ability than gray histogram features. Here, m bin histograms in color space are selected to represent the objects' color probability density functions. So the color weighting histogram for the target model is obtained as follows:

$$\hat{q}_u = C \sum_{i=1}^N K(\bullet) \delta [b(x_i) - u], \quad (3)$$

where δ is the delta function, u is the quantitative level of histograms, $u = 1, \dots, m$, C is the normalized constant, and N is the pixel number in the color-related kernel acting region. $\{x_i^*\}_{i=1, \dots, n}$ are the normalized pixel locations. $b(x_i^*)$ is the index of its bin in the quantized color feature space, $b(x_i) = 1, \dots, m$. The function $K(\bullet)$ is the proposed kernel function which is based on the reconstructed kernel shape.

In the tracking window based on kernel shape, central pixels can provide more reliable information than boundary

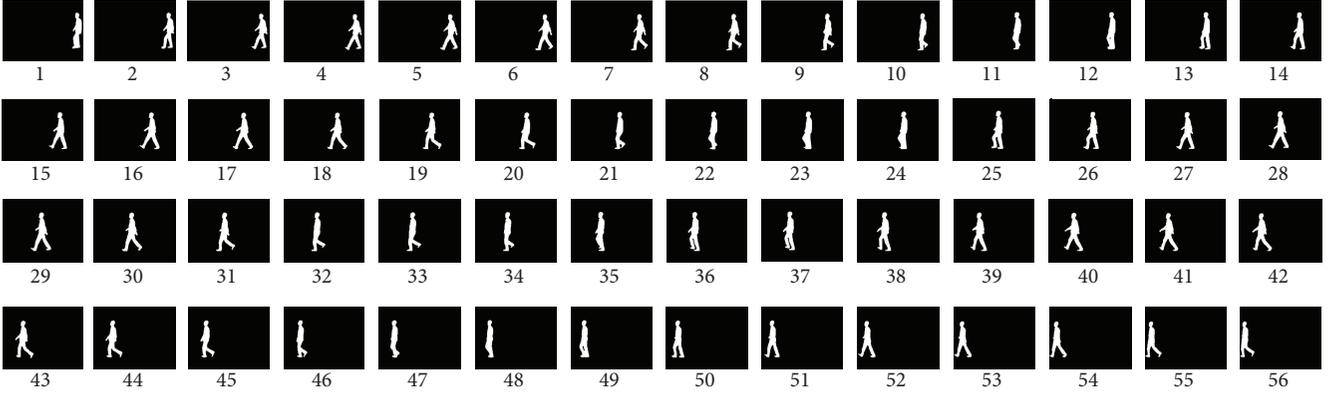


FIGURE 3: The silhouette sequence for training.

pixels for tracking as it suffers from less interference by background or occlusion. The distance from each pixel to the center of kernel shape is applied to compute its weight which decides the action rate for tracking matching.

Based on the reconstructed kernel shape, the adaptive kernel is computed by the normalization of each pixel distance $d(x_i)$ to the shape boundary:

$$K(x_i) = \frac{d(x_i)}{d_{\max}}. \quad (4)$$

Here, d_{\max} is the maximum distance from the boundary to the center.

4.2. Implementation Process. The proposed mean shift combines the reconstructed shape kernel and color features to find the most probable position of the target object through iteration. The detailed tracking process is implemented as follows.

- (1) Initialize the location of object y_0 and the kernel shape Γ_0 to compute color kernel K_{y_0} to obtain the target model q_u .
- (2) For a location of the tracking window y_k , compute the model p_k by using previous color $K_{y_{k-1}}$ based on shape Γ_{k-1} and estimate the new position \hat{y}_k and update the weight $w(y_k)$:

$$\hat{y}_k = \frac{\sum_{i=1}^{n_h} [x_i w_i g(\|(y_k - x_i)/h\|^2)]}{\sum_{i=1}^{n_h} [w_i g(\|(y_k - x_i)/h\|^2)]}, \quad (5)$$

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u}} \delta[b(x_i) - \mu].$$

- (3) For a kernel shape of the tracking window y_k , compute the model p_k by using previous color K_{y_k} based on shape Γ_k and estimate the new shape $\hat{\Gamma}_k$ by mapping between embedding shape space and original shape space.
- (4) Track the object repeating steps (2)-(3).

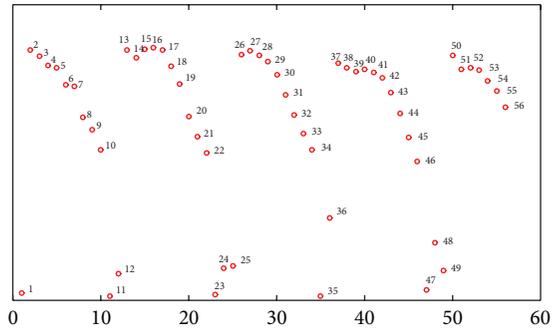


FIGURE 4: The embedded shape in 2-dimensional shape space.

5. Experiments

To demonstrate the performance of the proposed adaptive mean shift tracker, we have experimented with various sequences [18]. In all of the experiments with the proposed tracker, in the hsv color space the color-related component is used with equally spaced values in each color band, namely, with $m = 2 \times 2 \times 16$ bins, and the weight of the color feature kernel is set to 0.9.

5.1. Shape Reconstruction. In all tested sequences, the initialization was manually performed in the first frame. The tracked shape is constructed from a training silhouette sequence in 0-degree direction (CASIA Gait Database [19]). As shown in Figure 3, the training silhouette sequence includes 56 frames, which corresponds to points in 2-dimensional shape space. Each point corresponds to one shape in original shape space. Furthermore, it demonstrates 12 shape points form one walking gait circle. The shape point on one walking circle corresponds to one shape in original shape, as shown in Figure 4. So it is executable to reconstruct any shape from 2-dimensional shape space to original shape. Figure 5 shows the results. Figures 5(a) and 5(c) are original training shapes of frames 19 and 20. Figure 5(b) is the reconstructed shape between training shapes of frames 19 and 20. So we can apply the proposed method to construct the shape of the arbitrary shape point in 2-dimensional shape

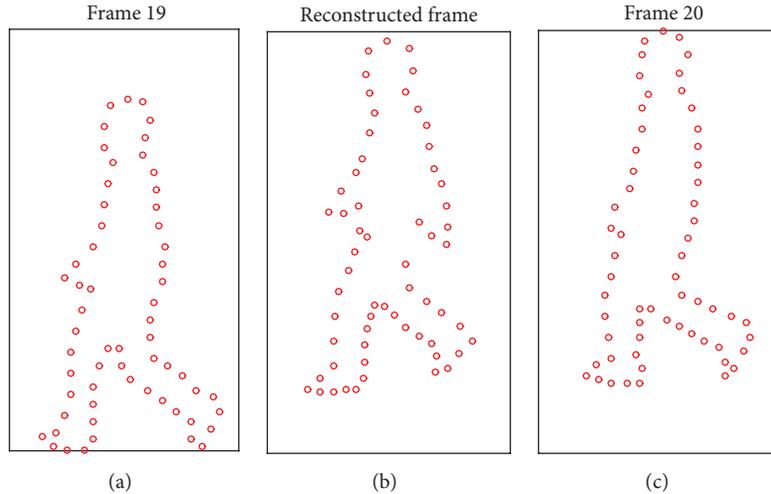


FIGURE 5: The reconstructed shape from 2-dimensional shape space to original shape space.

space. It helps in searching shape parameter for kernel of mean shift tracker.

5.2. Bhattacharyya Coefficients Analysis. In order to demonstrate the robustness of the adaptive shape kernel-based mean shift to the background clutter, we implement a Bhattacharyya coefficients analysis comparing with the traditional mean shift algorithm based on a rectangular kernel shape. We make analysis of the Bhattacharyya coefficients which are used to find the similarity between the current tracking instance and the target model.

The Bhattacharyya coefficients are shown corresponding to the tracking windows centered in a 50×50 neighborhood around the object center [12] in Figure 6. In Figure 6(a), the green rectangle block is the object center area, and the yellow rectangle blocks are the tracking windows centered in the object center. Bhattacharyya coefficient 51×51 area is constructed by computing Bhattacharyya coefficients of the tracking windows as shown in Figures 6(b)-6(c). It is distributed convexly and monotonically. Generally the maximum corresponds to the object center.

Figure 6(b) shows that the Bhattacharyya coefficient surface by the ellipse shape is flat as there is less difference of Bhattacharyya coefficients between the object center and the object center's neighborhood. The reason is that the ellipse shape of kernel is inevitable to include some background points. These background points participate in the computation of Bhattacharyya coefficient. Thus this leads to the flat Bhattacharyya coefficient surface. In Figure 6(c), the Bhattacharyya coefficient surface by the proposed algorithm is steep because less background points participate in the computation of Bhattacharyya coefficient in the adaptive kernel of object center than the neighborhoods' of the object center. The kernel shape is adaptive to the object shape. There is less background influence in the adaptive kernel. So it is effective and robust when tracking the object.

5.3. Adaptive Shape Kernel-Based Mean Shift. The proposed algorithm works with the assumption that the video sequence

TABLE 1: Overall tracking accuracy.

	Precision	Recall
ν_1	0.822	0.882
ν_2	0.818	0.8418
ν_3	0.801	0.8633

Precision = $\{\text{true shape}\} \cap \{\text{tracked shape}\} / \{\text{tracked shape}\}$.

Recall = $\{\text{true shape}\} \cap \{\text{tracked shape}\} / \{\text{true shape}\}$.

to be processed is captured by a static camera. So in order to illustrate the performance of the proposed method, we tested three sequences from UIUC database [18] which have a similar view to training silhouette dataset. The video has been converted into sequences of image frames with resolution of 720×480 pixels.

We first present results for estimating the location and shape of a walking human. Figure 7 shows four frames of one walking human sequence. As the proposed adaptive kernel can describe the object shape well and there is less background in the shape of an adaptive kernel, it is effective for human tracking. Compared with traditional mean shift tracker, it can not only track human position but also describe human contour. In Figure 7, the human contour tracking is not accurate enough as we performed position tracking and shape tracking only once for every frame in order to save computation time. If we performed the position searching once for each iterative shape during shape tracking, the results would be improved. But it would increase lots of computation. For the compromise considered, we apply the proposed iteration.

Here, we present the video sequences with the ground truth annotated by hand. We use precision and recall to evaluate the overall tracking performance for the proposed mean shift algorithm. In Table 1, our proposed method obtains overall tracking accuracy of 0.8~0.9. It demonstrates that the proposed algorithm can track not only the object position but also the object shape.

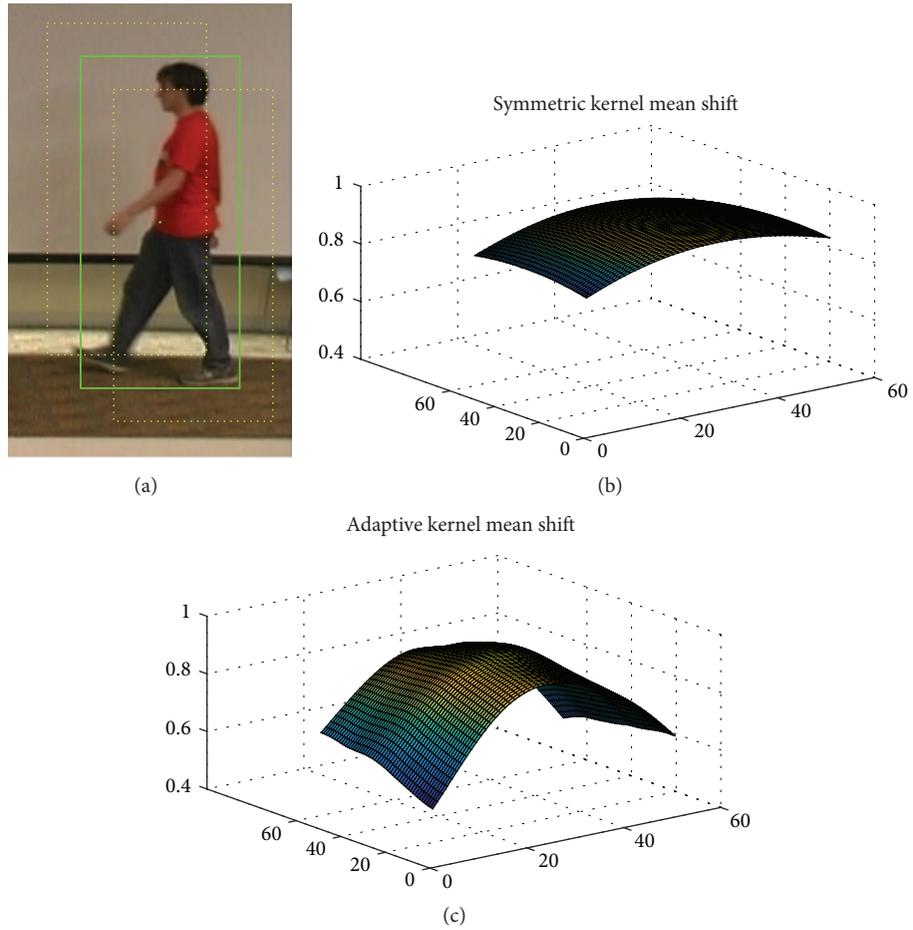


FIGURE 6: Bhattacharyya coefficients. (a) Tracking object; (b) Bhattacharyya coefficients result by symmetric kernel; and (c) Bhattacharyya coefficients result by adaptive kernel.

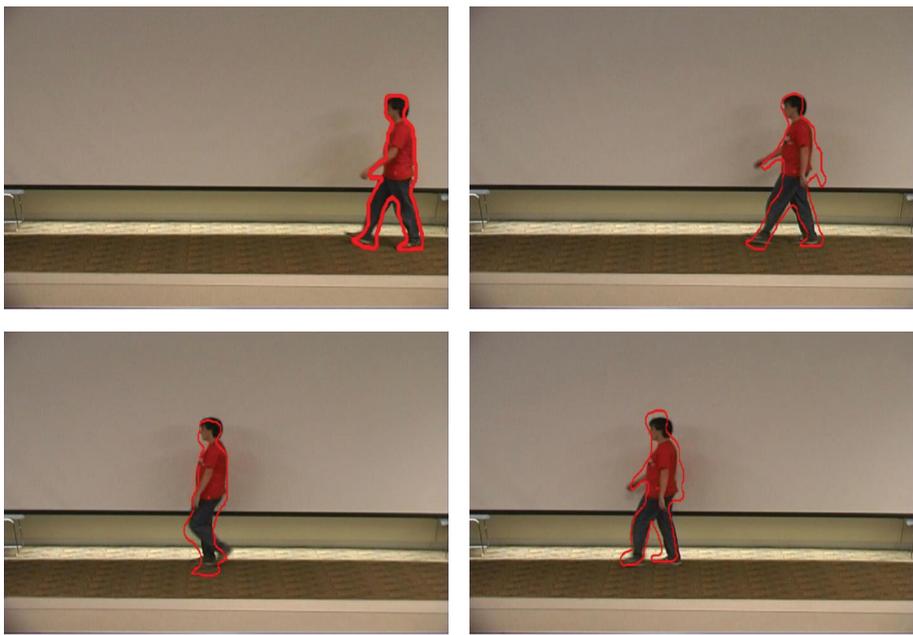


FIGURE 7: The results by adaptive kernel mean shift tracker.

6. Conclusion

In this paper, we have proposed a novel adaptive kernel-based mean shift tracker, which integrates color feature kernel based on adaptive shape to improve object tracking performance. Experiments have validated that our method is accurate and robust to track human position and describe human contour. We believe that improvements are due to the increased accuracy of the kernel shape construction integrating color feature. We have noticed that the proposed method only adapts to the fixed camera view. In the future, we will try some other adaptive shape models for more views. On the other hand, recently considering performance and computation, the correlation filter-based trackers behave very well. A well-defined model shape helps in object tracking. We will further integrate the proposed adaptive shape with the correlation filter-based trackers to improve the tracking performance for robot vision system.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (nos. 61003102 and 61472284).

References

- [1] C. J. Veenman, M. J. T. Reinders, and E. Backer, "Resolving motion correspondence for densely moving points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 1, pp. 54–72, 2001.
- [2] A. Yilmaz, X. Li, and M. Shah, "Contour-based object tracking with occlusion handling in video acquired using mobile cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1531–1536, 2004.
- [3] G. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology Journal*, vol. 2, p. 1, 1998.
- [4] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 564–577, 2003.
- [5] N. S. Peng, J. Yang, and Z. Liu, "Mean shift blob tracking with kernel histogram filtering and hypothesis testing," *Pattern Recognition Letters*, vol. 26, no. 5, pp. 605–614, 2005.
- [6] S. J. McKenna, Y. Raja, and S. Gong, "Tracking colour objects using adaptive mixture models," *Image and Vision Computing*, vol. 17, no. 3-4, pp. 225–231, 1999.
- [7] H. Zhang, W. Huang, Z. Huang, and L. Li, "Affine object tracking with kernel-based spatial-color representation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 293–300, June 2005.
- [8] I. Leichter, M. Lindenbaum, and E. Rivlin, "Visual tracking by affine kernel fitting using color and object boundary," in *Proceedings of the IEEE 11th International Conference on Computer Vision (ICCV '07)*, pp. 1–6, Rio de Janeiro, Brazil, October 2007.
- [9] Z. Ali, S. Hussain, and I. A. Taj, "Kernel based robust object tracking using model updates and Gaussian pyramids," in *Proceedings of the IEEE International Conference on Emerging Technologies (ICET '05)*, pp. 144–150, September 2005.
- [10] D. Chakraborty and D. Patra, "Real time object tracking based on segmentation and kernel based method," in *Proceedings of the 5th International Conference on Industrial and Information Systems (ICIIS '10)*, pp. 426–429, IEEE, Mangalore, India, August 2010.
- [11] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2544–2550, San Francisco, Calif, USA, June 2010.
- [12] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '00)*, pp. 142–149, June 2000.
- [13] K. Quast and A. Kaup, "Shape adaptive mean shift object tracking using Gaussian mixture models," in *Proceedings of the 11th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '10)*, pp. 1–4, Desenzano del Garda, Italy, April 2010.
- [14] A. Yilmaz, "Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–6, Minneapolis, Minn, USA, June 2007.
- [15] A. Yilmaz, "Kernel-based object tracking using asymmetric kernels with adaptive scale and orientation selection," *Machine Vision and Applications*, vol. 22, no. 2, pp. 255–268, 2011.
- [16] K. M. Yi, H. S. Ahn, and J. Y. Choi, "Orientation and scale invariant mean shift using object mask-based kernel," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, Tampa, Fla, USA, December 2008.
- [17] I. Leichter, M. Lindenbaum, and E. Rivlin, "Tracking by affine kernel transformations using color and boundary cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 164–171, 2009.
- [18] N. İkizler and D. A. Forsyth, "Searching for complex human activities with no visual examples," *International Journal of Computer Vision*, vol. 80, no. 3, pp. 337–357, 2008.
- [19] Database:CASIA, <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>.

Research Article

An Interactive Astronaut-Robot System with Gesture Control

Jinguo Liu,¹ Yifan Luo,^{1,2} and Zhaojie Ju³

¹State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

²University of Chinese Academy of Sciences, Beijing 100864, China

³School of Computing, University of Portsmouth, Portsmouth, Hampshire PO1 3HE, UK

Correspondence should be addressed to Jinguo Liu; liujinguo@sia.cn and Zhaojie Ju; zhaojie.ju@port.ac.uk

Received 23 November 2015; Accepted 2 March 2016

Academic Editor: Hiroki Tamura

Copyright © 2016 Jinguo Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human-robot interaction (HRI) plays an important role in future planetary exploration mission, where astronauts with extravehicular activities (EVA) have to communicate with robot assistants by speech-type or gesture-type user interfaces embedded in their space suits. This paper presents an interactive astronaut-robot system integrating a data-glove with a space suit for the astronaut to use hand gestures to control a snake-like robot. Support vector machine (SVM) is employed to recognize hand gestures and particle swarm optimization (PSO) algorithm is used to optimize the parameters of SVM to further improve its recognition accuracy. Various hand gestures from American Sign Language (ASL) have been selected and used to test and validate the performance of the proposed system.

1. Introduction

When astronauts conduct EVA missions on the surface of other planets, they generally need to collaborate with some agents or some systems to complete the missions smoothly and efficiently. Reducing the crew workload is a primary concern, particularly during EVA. The robot's autonomy can make the robot finish some tasks independently and allow the robot to complete certain tasks with little crew's attention. The robot used in the space exploration always has a high level of autonomy (LOA). However, in current real operations, a human operator has a better insight in the task completion than the robot system. Autonomous systems are not yet as efficient as humans in modeling the richness of interactions and balancing the trade-off between the various crewmembers and their mission requests. Therefore, astronauts must interact with the robot at various levels, from high level goal commands to detailed activity sequences and then to direct teleoperation, to cope with the full spectrum of situations expected. This creates significant challenges with regard to communication, human-robot interface, and human-understandable state representation.

As for the HRI problem, considerable effort has been made to the development of intelligent and natural interfaces

between users and computer systems, and HRI has been developed by leaps and bounds [1–6]. Now there are many mature ways of HRI; among those ways, voice recognition and gesture recognition are two major developing directions. Speech recognition system now is developing towards two important directions: one is the large vocabulary continuous speech recognition system and the other is the application of miniaturization, portable audio products. The large vocabulary and continuous speech recognition system is now generally based on one or more PCs. The portable processing chip for recognition usually has limitations in computing speed and storage capacity. In planetary exploration missions, these limitations indicate that there is still a long way to go to apply speech recognition in this area. Hand gestures, which have been addressed in the sign language for the deaf people for many years, can represent rich language and have also attracted a lot of attention. Gesture recognition is a technology often used in HRI applications, and there are lots of methods for hand gesture recognition, such as the methods based on image recognition, curvature, and surface electromyography (EMG) signal.

This paper proposes a way of using hand gestures of astronauts to intervene in the autonomy of the agent. An example of astronauts cooperating with agents to complete a

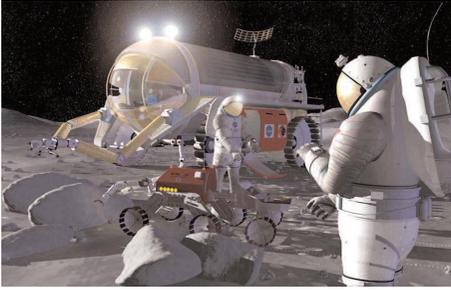


FIGURE 1: Schematic diagram of astronauts collaborating with agent [3].

mission is shown in Figure 1. Though recent image processing techniques have achieved a fascinating development [2], they are not suitable for the space applications, because the clumsy suit may bring some of the most difficult problems in the field of machine vision [7]. For surface EMG signals, there is a large gap in the space suits and the atmospheric pressure inside spacesuit is only 40 percent of the standard atmosphere, so whether the EMG signals in this case change or not is unknown.

Increasing numbers of industrial and service robots [8, 9] have focused on designing the HRI technology in order to increase robot efficiency and effectiveness. HRI refers to a process of conveying operators' intentions and interpreting the sequence of robot motions and working requirements in task descriptions. The complement of HRI through the application of suitable interaction methods and interfaces has been an essential factor as well as a challenge in the robot industry. Recent development of robotics has introduced haptic interaction, through which the users can feel both virtual and real environments, such as in teleoperations and telesurgeries [10]. There have been many works providing technical and theoretical support for HRI to be more efficient and suitable. Now commonly used methods include multimodal interaction, teaching model, virtual reality, and augmented reality.

Nowadays, the space activity is still in the early stage, and the technology needs further improvement. In the near future, with the development of aerospace technology, the astronauts will not be limited to the technical personnel; other people, such as engineers, physicists, biologists, surgeons, and even philosophers, also have the opportunities to become astronauts in the space exploration and carry out relevant scientific experiments. Therefore, the individual agent or multiagent system, which collaborates with astronauts, requires a higher LOA and friendly HRI. Making HRI more effective, efficient, and natural is crucial to the success of sustained space exploration. In particular, we assume that humans and robots must be able to (1) communicate clearly about their goals, abilities, plans, and achievements; (2) collaborate to solve problems, especially when situations exceed autonomous capabilities; and (3) interact via multiple modalities (dialogue, gestures, etc.), both locally and remotely. To achieve these goals, a number of HRI challenges must be addressed.

TABLE 1: Main parameters of the device.

Items	Properties
Bend sensors	Temperature range: $-35^{\circ}\text{C}\sim+80^{\circ}\text{C}$; resistance tolerance: $\pm 30\%$
Stm32 controller	Cores: Cortex-M3 32-bit RISC, 512 K Flash, 64 K RAM; operating frequency: 72 MHz, 1.25 DMIPS/MHz
NRF24L01	Transmission distance: 150 m; digital interface (SPI) speed: 0~10 Mbps; on the air data rate 1 or 2 Mbps
Servos	Power supply range: 7~10 V; operating temperature: $-5^{\circ}\text{C}\sim+85^{\circ}\text{C}$; communication speed: 7343 bps~1 Mbps

Using gestures to convey information has become an important part of human computer interaction [4–7]. Hand gesture recognition is widely used in many applications, such as computer games, machinery control (e.g., crane), and household electrical appliance remote control. Hand gesture analysis can be divided into three main approaches, namely, glove-based methods, vision-based methods, and methods for drawing gestures [5]. For approaches based on the data-glove, the relative position of a finger is captured by an additional sensor, which is normally a magnetic or acoustic sensor attached to a glove. A lookup table software toolkit is usually provided for hand gesture recognition [7]. The second way is based on the image processing, which is stricter with the image background, and thus it is not suitable for applications in a complex working environment [6]. The third method involves the analysis of gesture drawing [5], using a stylus as an input device. This method is often used for identifying written words, which has problems of reliability, accuracy, and electromagnetic interference noise.

The paper is organized as follows. In Section 2, the interactive astronaut-robot system is introduced in detail, including the system devices, the overall plan and the main functions, and the snake-like robot. In Section 3, we introduce the application of SVM and PSO for the hand gesture recognition. In Section 4, we designed two experiments to verify the reliability and robustness of the proposed system. Conclusions and future work are discussed in Section 5.

2. Interactive Astronaut-Robot System

The system integrates bending sensors in a glove to capture the bending angles of all the fingers. Then the finger angles are classified through the model trained by the SVM, and corresponding instructions generated control the snake-like robot, so that the snake-like robot can assist astronauts to complete the mission. The main components include bending sensor system, STM32 controller, wireless communication module, and the modular snake robot composed with servos. The main parameters of each device are shown in Table 1.

2.1. The Control System. The main function of this control system is designed to achieve the modular robot moving with the planned movement according to the instructions from the gesture recognition system. Detailed implementation is

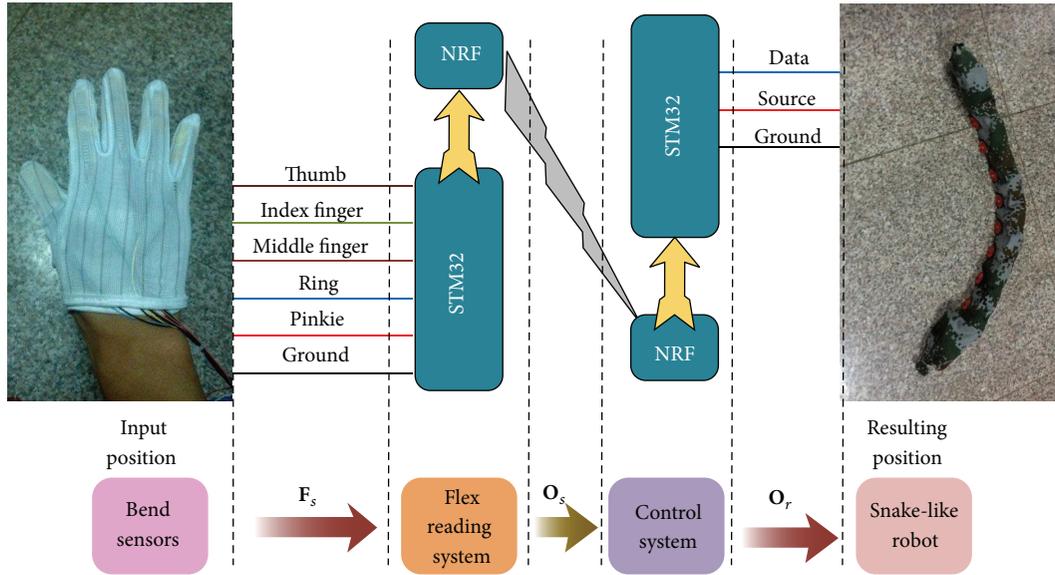


FIGURE 2: Outline of the control system.

shown in Figure 2. After the controller gets the signal F_s from the bend sensors mounted on the glove, the signal goes through a filter and a normalization preprocessing stage, and O_s is sent to the controller mounted in the snake-like robot through wireless module. This controller processes O_s by SVM and gets the predicting label. Then corresponding operation instructions are sent to the snake-like robot. Finally the snake-like robot executes the corresponding movement.

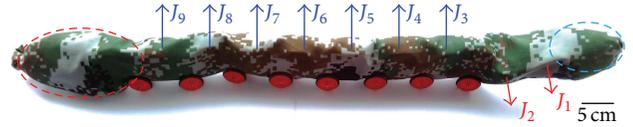


FIGURE 3: Structure of the snake-like robot.

2.2. Snake-Like Robot. Snakes could do very well in the rough terrain like Mars, by going over and through broken ground and sand, and squeeze through tight spaces. Thus, great interest in the snake-like robot research has been generated. The European Space Agency is developing snake-like robots aiming at providing robot with more mobility during space exploratory activities. The snake-like robot applied in the mission of lunar exploration and Mars exploration will be helpful for the rover to travel over the complex rugged surface and narrow gaps on the ground.

During some missions where a wheeled rover collaborates with a snake-like robot, the wheeled rover can be used to travel long distances, while the snake robot could detach and reach places where the rover cannot reach. And if the rover gets stuck, the snake robot could conceivably be used to help pull it away.

Hirose has proposed the serpentine curve early in 1993 [11]. The curvature of the serpenoid curve is given by

$$\rho = -\alpha b \sin(bs), \quad (1)$$

where α is amplitude angle (rad); b is constant of proportionality (rad/m); s is length of serpentine curve (m).

The snake-like robot is composed of modular units, which are connected by active revolute joints, and the change of position between relative modules results in the movement of the robot. The flexible architecture of snake-like robot makes

it hard to make a turning movement like other legged robots. To ensure the snake-like robot can achieve high efficiency in turning movement, Ye et al. proposed several methods for the turning motion of snake-like robot [12]. The snake-like robot used in this paper is shown in Figure 3 and made up of ten serial joints and each joint has one degree of freedom. A camera (the one encircled by the blue circle) is arranged on the head and a control module (the one encircled by the red circle) is fixed at the tail. Its physical connection is shown in Figure 4.

In the design of the communication system in a snake-like robot, a half-duplex asynchronous serial communication (8 bits, 1 stop, no parity) is utilized. Transmission speed is up to 1 Mbps. Link (physical) is TTL level multidrop (daisy chain type connector) considering minimizing physical cable.

The protocol of each modular unit communicating with the main controller is shown in Figure 5. Two 0XFF are the start code, ID is the number for the corresponding actuator, LENGTH is the length of the instruction, instruction is the instruction for the actuator to perform, PARAMETER is additional information needed to be sent other than the instruction, and the checksum is used to verify the signal. Distributed feedback compensation control is used as the control method. The specific control block diagram is shown in Figure 6.

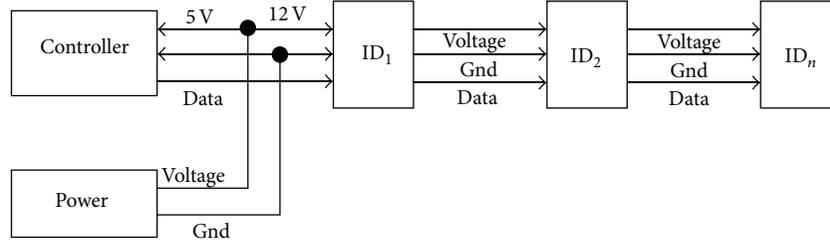


FIGURE 4: Physical connection.

0xFF	0xFF	ID	LENGTH	INSTRUCTION	...
PARAMETER ₁	...	PARAMETER _N	CHECKSUM		

FIGURE 5: Snake-like robot communication protocol.

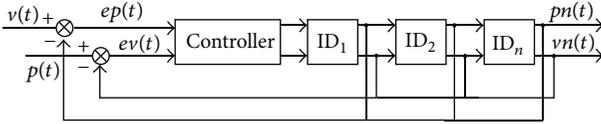


FIGURE 6: Control block structure.

3. Motion Recognition and Parameter Optimization

Machine learning based on data is an important aspect of modern intelligence technology. Statistics study begins with the observation of data to conclude a model, which is the base of the forecast for future data or the data cannot be observed. Traditional statistics study the asymptotic theory when the number of samples tends to infinity. Existing learning methods are mostly based on this assumption. But, in practical problems, the number of samples is often limited, so they usually have an unsatisfactory performance. Compared with the traditional statistics, Statistical Learning Theory (SLT) is a specialized theory, which systematically studies the relationship between experiences risk and actual risk for various types of sets of functions, namely, the generalization bounds [5]. Vapnik and Kotz began to dedicate themselves to researching this theory from the 1960s [13]. In the mid-90s, because of the development of Vapnik's theory and the lack of substantive progress in the theory of neural network learning methods, SLT began to receive more appreciation. SLT was based on a solid theory and provided a unified framework for solving the learning problem with the small samples. It incorporates many of the existing methods, expected to help solve many difficult problems, for example, the selection of neural network architecture and the local minima problem. Based on this theory, there is a new universal learning method; support vector machine (SVM), using geometry classification method to find the optimal hyperplane and get the maximum margin classifier, has shown a lot of superiority compared to the existing method [14, 15].

SVM is a more practical part of statistical theory, which was originally proposed by Vapnik et al. in 1992 to 1995 [14,

16–18]. It is currently still in the development stage. SVM is a structure of risk minimization strategies, which compromise the empirical risk and confidence interval to obtain the actual minimum risk [19]. A SVM approaches problems by searching for the Maximum Marginal Hyperplane (MMH) where a hyperplane has an equal distance from the hyperplane to both sides of its margin to ensure the hyperplane is more accurate at classifying future data tuples [20]. Compared with the new algorithms like Extreme Learning Machine (ELM) [21], SVM is committed to using less parameters to express a complex model; it still has its advantage in methodology and is more plausible.

SVM classifies linear data directly. When the data is linearly inseparable, it transforms the original data into a higher dimensional space by using a nonlinear mapping, and then searches for a linear separating hyperplane in the new space. Nonlinear data processing steps are shown in Figure 7.

There are several modes of SVM, which can be used for data classification, regression, and distribution estimation [22]. This paper uses the C-Support Vector Classification (C-SVC) [17, 23] to classify the data.

The distinguished hyperplane of the sample set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_1, \dots, \mathbf{x}_N\}$ can be shown by the formula

$$\mathbf{W}^T \mathbf{x} + w_{d+1} = 0, \quad (2)$$

where \mathbf{W} is the weight vector and the direction of hyperplane. d is the dimension of the feature space. w_{d+1} is the offset of the hyperplane. During the course of looking for the best \mathbf{W}^* to maximize the interval between the hyperplane and the closest sample, Lagrange multiplier method can be used to solve the problem of inequality constraint. The corresponding Lagrange function is

$$L(\mathbf{W}, w_{d+1}, \lambda) = \frac{1}{2} \mathbf{W}^T \mathbf{W} - \sum_{k=1}^N \lambda_k [y_k (\mathbf{W}^T \mathbf{x}_k + w_{d+1}) - 1], \quad (3)$$

where $\lambda_k \geq 0$ and $k = 1, 2, \dots, N$ is the Lagrange coefficients to be determined.

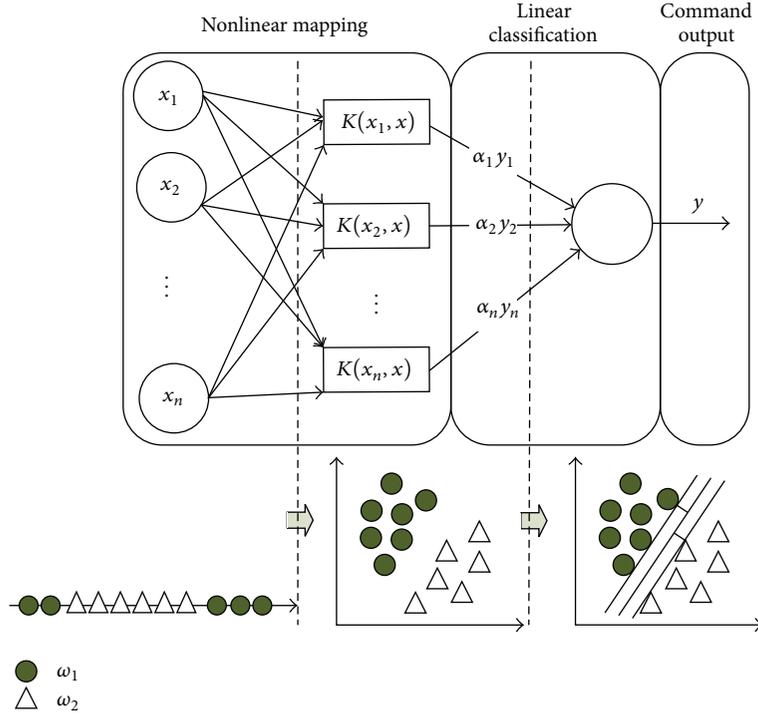


FIGURE 7: SVM nonlinear data processing principle.

To obtain a necessary condition for the extreme value in Lagrange function, the course of seeking the partial derivatives equaling zero of \mathbf{W} and w_{d+1} is shown below:

$$\frac{\partial}{\partial \mathbf{W}} L(\mathbf{W}, w_{n+1}, \lambda) \Big|_{\mathbf{W}=\mathbf{W}^*} = \mathbf{W}^* - \sum_{k=1}^N \lambda_k y_k \mathbf{x}_k = 0, \quad (4)$$

$$\frac{\partial}{\partial \mathbf{W}} L(\mathbf{W}, w_{n+1}, \lambda) \Big|_{w_{d+1}=w_{d+1}^*} = -\sum_{k=1}^N \lambda_k y_k = 0.$$

Namely,

$$\mathbf{W}^* = \sum_{k=1}^N \lambda_k y_k \mathbf{x}_k, \quad (5)$$

$$\sum_{k=1}^N \lambda_k y_k = 0.$$

Convert it to the dual form:

$$L_D(\lambda) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j + \sum_{k=1}^N \lambda_k. \quad (6)$$

To ensure distinguished hyperplane has the smallest risk of classification,

Maximise $L_D(\lambda)$

$$= \text{Maximise} \left\{ \sum_{k=1}^N \lambda_k - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \right\}$$

Subject to $\lambda_k \geq 0, k = 1, 2, \dots, N,$

$$\sum_{k=1}^N \lambda_k y_k = 0.$$

(7)

The function showed above is the simple quadratic programming problem, which has standard solving algorithm. Once the problem is solved under the condition of $\lambda_k \geq 0, k = 1, 2, \dots, N,$ the optimal weight vector \mathbf{W}^* will be got based on the formula shown in (5). Solutions meeting the requirements are called support vector.

When it comes to nonlinear classification, the data is usually mapped to a high-dimensional linear space by the kernel function in Figure 7. In this way the linearly inseparable data can be converted into linear separable data in a high-dimensional space. Three kinds of kernel functions are commonly used, namely, polynomial kernel of degree $h,$ Gaussian radial basis function kernel, and Sigmoid kernel. Three kernel functions are as follows.

Polynomial kernel of degree h is

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^h. \quad (8)$$

Gaussian radial basis function kernel is

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\|\mathbf{x}_i - \mathbf{x}_j - 1\|^2 / 2\sigma^2}. \quad (9)$$

Sigmoid kernel is

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\kappa \mathbf{x}_i \cdot \mathbf{x}_j - \delta). \quad (10)$$

There are no golden rules for determining which admissible kernel will result in the most accurate result in SVM. In

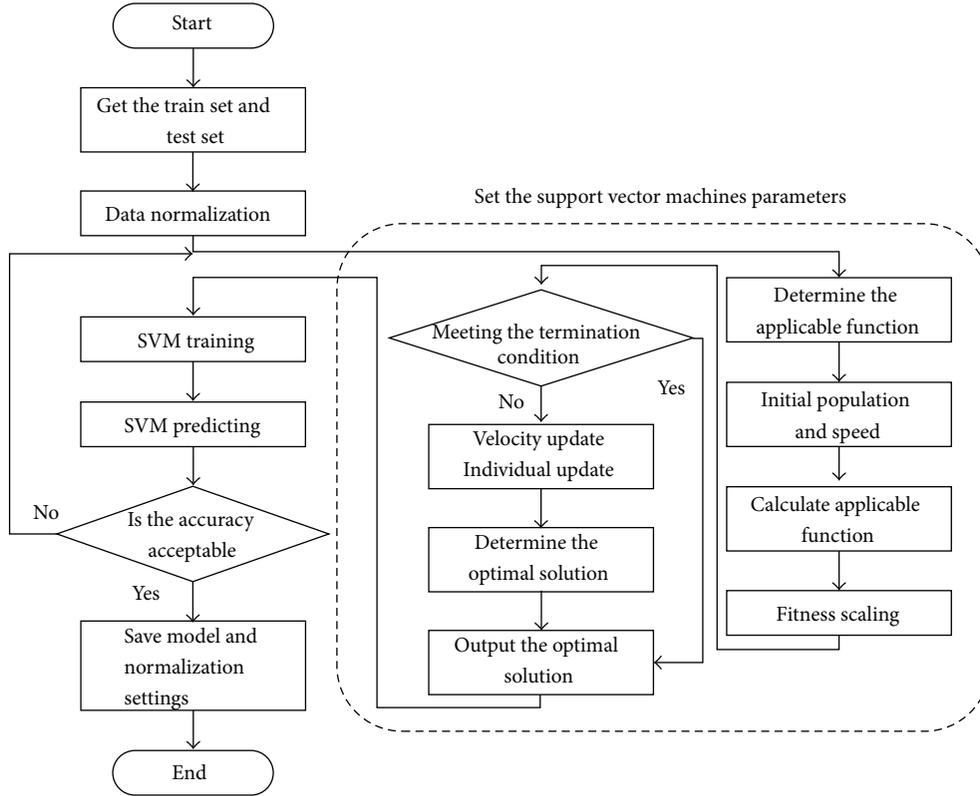


FIGURE 8: Flow diagram of the overall program.

practice, the kernel chosen does not generally make a large difference in the resulting accuracy. SVM training always finds a global solution, unlike neural networks, such as backpropagation, where many local minima usually exist.

For the using of SVM, although the choosing of kernel generally does not make a large difference in result accuracy, when a kernel is chosen, there are still a number of parameters that should be optimized. In this paper, after selecting the Gaussian radial basis function, there are two parameters c and g that need to be optimized, where c is the penalty coefficient that means error tolerance; the higher the value is, the smaller the error can be tolerated. Parameter g determines the distribution of data after mapping to the new feature space.

There is no best way to select the SVM parameters. The most common way is to let c and g be within a certain range. In this paper, cross-validation method based on grid-search was used for the parameter optimization. Cross-validation is one of the more classic solutions [22]. The algorithm is conducted according to a basic idea that in the inner loop of cross-validation, once the recognition rate for the first time appears to be a local maximum, the parameter values are recorded and the inner loop ends. Finally, estimate the optimal parameters by calculating the arithmetic mean of the entire local maximum.

PSO is a new Evolutionary Algorithm (EA) developed in recent years [24]. The particle swarm is more than just a collection of particles. A particle by itself has almost no power to solve any problem. Progress occurs only when the particles

interact. Particle swarm follows the optimal particle to search the solution space; each particle obtains a search direction and speed in next loop by comparing with the individual optimum value and global optimum value respectively with random perturbations distributed uniformly in a certain range. Compared with other EAs, the advantages of PSO are being simple, being easy to achieve, and few parameters to be adjusted. Using PSO with appropriate parameters can significantly improve the accuracy of SVM [25–28]. The formulas to update the primitive velocity and location are shown as follows:

$$\begin{aligned}
 \vec{v}_i &\leftarrow \omega \vec{v}_i + \vec{U}(0, \phi_1) \otimes (\vec{p}_i - \vec{x}_i) + \vec{U}(0, \phi_2) \\
 &\quad \otimes (\vec{p}_g - \vec{x}_i), \\
 \vec{x}_i &\leftarrow \vec{x}_i + \vec{v}_i,
 \end{aligned} \tag{11}$$

where \vec{x}_i is the current location; \vec{p}_i is the previous personal best position; \vec{p}_g is the previous global best position; \vec{v}_i is velocity and ω is inertia weight; $\vec{U}(0, \phi_i)$ represents a vector of random numbers uniformly distributed in $[0, \phi_i]$ which is randomly generated at each iteration and for each particle; \otimes is componentwise multiplication

In the original version of PSO, velocity of each particle is limited to $[-V_{\max}, +V_{\max}]$.

The program flow using the SVM, whose parameters were chosen by the PSO to obtain a classification model, is shown in Figure 8.

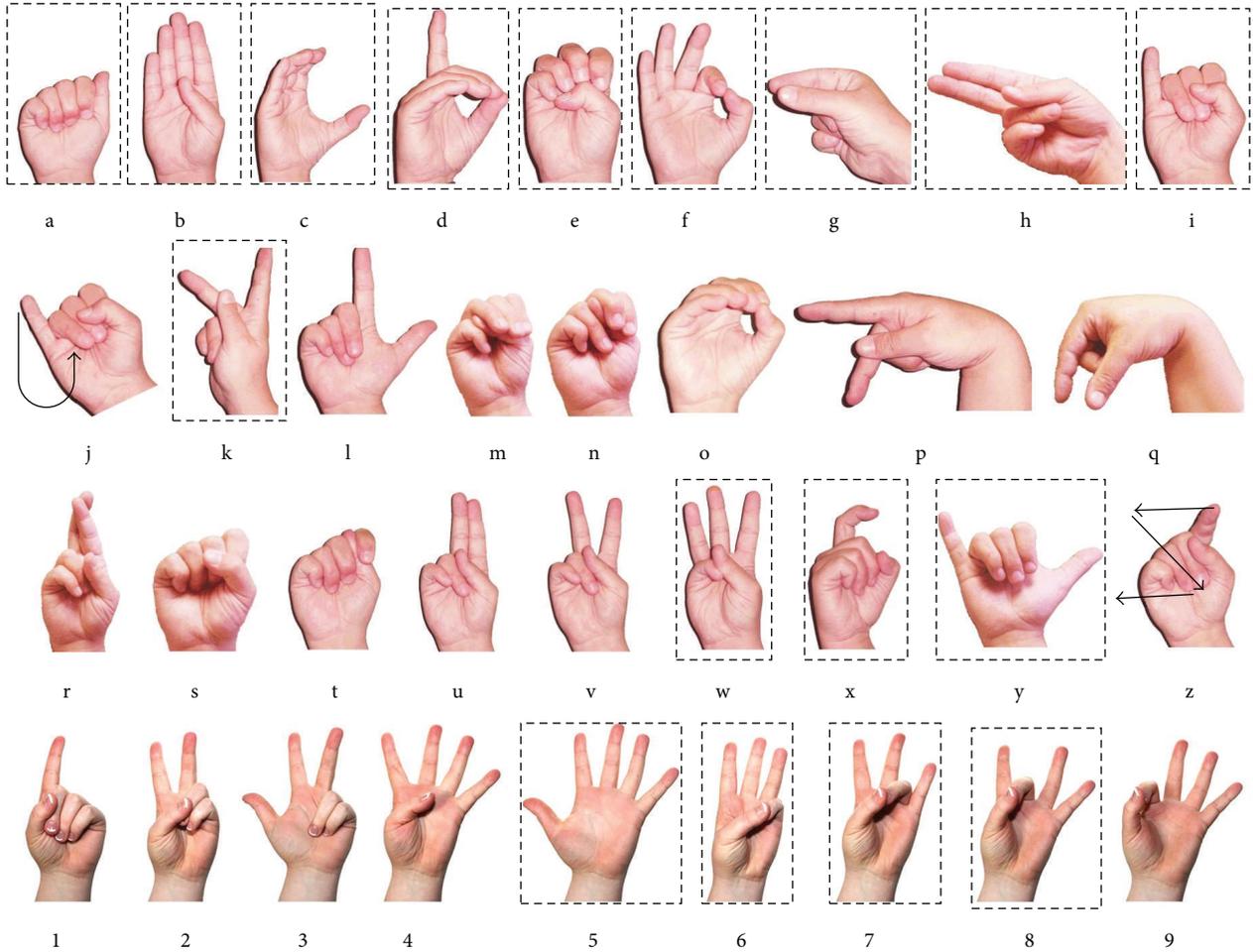


FIGURE 9: 16 kinds of gestures in ASL.

4. Experiment

In this paper, we focus on the planet surface EVA, where the autonomous robots need assistance on path planning, mission guidance, and so forth. In the process of classification of HRI instructions, the above learning method, SVM, with a small number of learning samples is used to classify the instructions. For the SVM parameter optimization, PSO algorithm was used to optimize SVM parameters c and g by the way of cross-validation. The found optimal parameters will be used to find the best SVM model. The software package LIBSVM we used was developed in [22].

In order to verify the accuracy and robustness of the proposed method, two experiments are conducted. First, the proposed methods are evaluated on 16 hand gestures selected from 36 hand gestures in the ASL. Second, the hand recognition algorithm has been integrated into a snake-like robot, and validation is then made with a space suit.

4.1. Hand Gesture Recognition. ASL has 36 hand gestures, 26 letters, and 6000 words. Although most of the ASL alphabet letters depend on finger bending, some of them also depend on hand orientation and two of them are dynamic. There are some similarities between g and q , h and u , and k and

p . These couples have basically the same hand shape, but their hand orientation differs from the others. There are hand shape similarities between i and j and x and z , but j and z are dynamic characters.

In this paper, we selected 16 in 36 of ASL shown in Figure 9 for the classification and identification experiment; corresponding gestures in the experiment are shown in Figure 10. For each gesture, we collected 15 sets of data, from which we use 10 for training and other 5 for the testing. The test data is normalized before testing the accuracy. In addition, we collected 5 new sets of hand shapes for each gesture to test the trained model.

For a more detailed analysis on the effect of using PSO for SVM cross-validation, we calculated the average Cross-Validation Accuracy (CVA) of the SVM cross-validation for different parameters of PSO, as shown in Tables 2 and 3 (each group has 6 experiments). At first, we fixed the maximum generation as 5 and adjusted the size of PSO population shown in Table 2. The parameter of PSO is the maximum generation or the size of PSO population. Obviously, when the size of PSO population was equal to 5, the CVA reached its maximum. With the population as 5 and the maximum generation, the results shown in Table 3 indicated that when

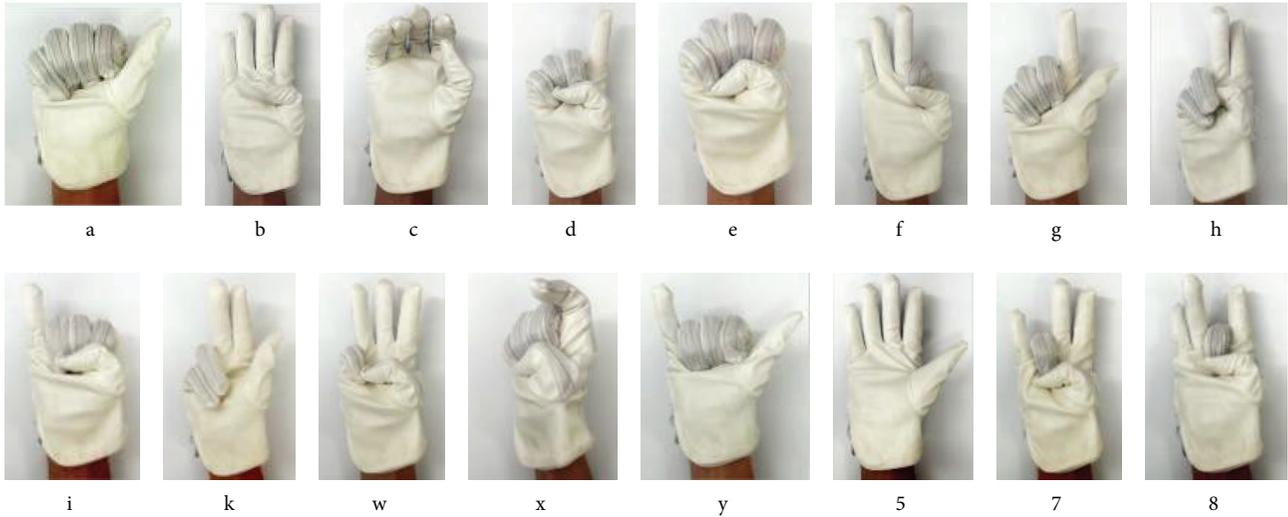


FIGURE 10: Corresponding gestures in experiment.

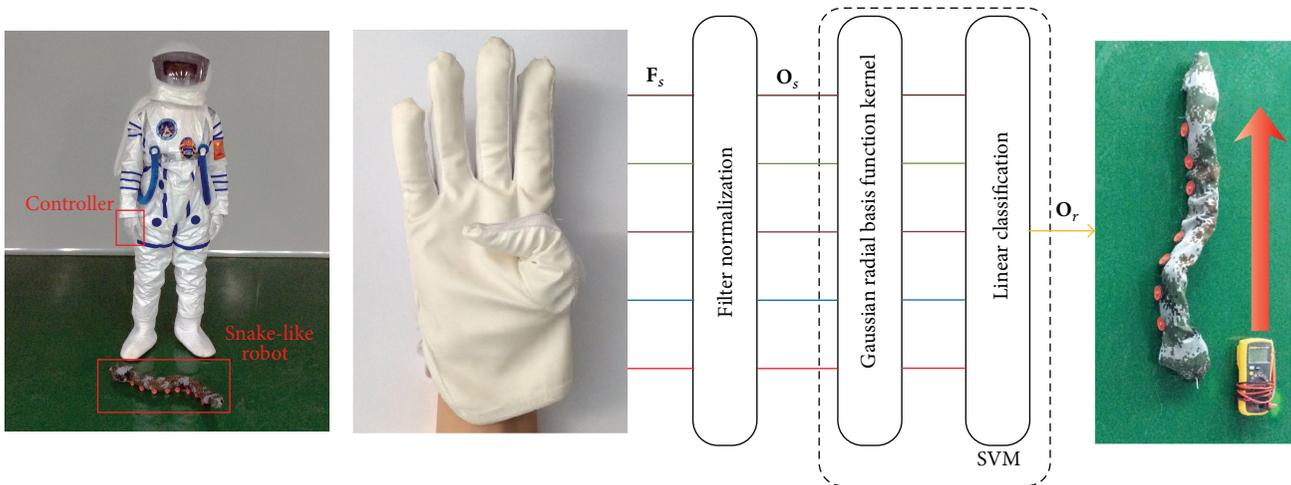


FIGURE 11: Experimental system.

the maximum generation is equal to 5, the CVA reached its maximum.

From Tables 2 and 3, we can also know that using PSO with appropriate parameters can significantly improve the accuracy of SVM in the process of cross-validation. Compared with the results of gesture recognition using ELM in [29], few parameters were used in this paper, and SVM had more stable results than the Extreme Learning Machine.

As we can see, the classification accuracy can always reach 100% except the size of PSO population that is too small. It demonstrates that the method in this paper has a high accuracy and strong robustness.

4.2. Snake-Like Robot Remote Control with Hand Gestures. A snake-like robot plays a powerful role in space exploratory activities. In this paper, a snake-like robot motion control was employed of testing the accuracy, stability, and robustness of the proposed approach. We modeled the environment of

TABLE 2: Average CVA when the maximum generation is fixed.

Parameter of PSO	CVA	Classification accuracy
5\2	55.1968%	64.79%
5\3	77.1160%	100%
5\4	77.2569%	100%
5\5	81.3294%	100%
5\6	78.9453%	100%
5\7	77.5563%	100%

astronauts on other planets, embedded the controller in the glove, and controlled the movement of the snake-like robot. Overall structure of the experiment is shown in the left of Figure 11 and a schematic diagram of control signal flow shown in the right.

The hand gestures have been integrated into the glove-robot control system. Various motions have been identified



FIGURE 12: Snake-like robot gesture control experiments.

TABLE 3: Average CVA when the size of PSO population is fixed.

Parameter of PSO	CVA	Classification accuracy
2\5	74.1435%	100%
3\5	76.9485%	100%
4\5	76.7045%	100%
5\5	81.3294%	100%
6\5	80.3364%	100%
7\5	80.2778%	100%

for the snake-like robot, such as turning left/right and moving forward/backward. It demonstrates that the proposed system including the hardware and software is effective and robust. It is a good prototype for the HRI used in the space exploration. The corresponding hand gestures, robot movements, and simulated motion tracks are shown in Figure 12, respectively.

5. Conclusion and Future Works

This paper proposed a gesture-type user control system for the space exploration based on the actual application

environment for the purpose of utility and stability. In this study, bending sensors were integrated with a space suit to control a snake-like robot, which was designed for the space exploration. SVM was used as the gesture signal pattern recognizer, and PSO algorithm was used for optimizing the parameters of SVM. The system classified the action sequence and ensured the accuracy and real-time performance of the control process. The experimental results showed that this system was effective with a high accuracy, reliability, and robustness.

In the future, the system will be improved with a series of command functions so that astronauts can interrupt robot's operations whenever necessary to provide guidance and assistance for the mission. Simultaneously, the collaboration between the astronaut and the robot will be strengthened and the interactions will be more precise and concise with advanced nonlinear methods [30, 31]. Finally, the HRI system will be further improved with a natural and friendly interface so that nontechnical astronauts can also have a barrier-free communication with robots.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research is supported by Research Fund of China Manned Space Engineering (050102), the Key Research Program of the Chinese Academy of Sciences (Y4A3210301), the National Science Foundation of China (51175494, 61128008, and 51575412), and the State Key Laboratory of Robotics Foundation.

References

- [1] D. Rempel, M. J. Camilleri, and D. L. Lee, "The design of hand gestures for human-computer interaction: lessons from sign language interpreters," *International Journal of Human Computer Studies*, vol. 72, no. 10-11, pp. 728-735, 2014.
- [2] H. Hasan and S. Abdul-Kareem, "Human-computer interaction using vision-based hand gesture recognition systems: a survey," *Neural Computing and Applications*, vol. 25, no. 2, pp. 251-261, 2014.
- [3] T. Fong and I. Nourbakhsh, "Interaction challenges in human-robot space exploration," *Interactions*, vol. 12, no. 2, pp. 42-45, 2005.
- [4] M. Kaushik and R. Jain, "Gesture based interaction NUI: an overview," *International Journal of Engineering Trends and Technology*, vol. 9, no. 12, pp. 633-636, 2014.
- [5] B. Ionescu, D. Coquin, P. Lambert, and V. Buzuloiu, "Dynamic hand gesture recognition using the skeleton of the hand," *Eurasip Journal on Applied Signal Processing*, vol. 2005, no. 13, pp. 2101-2109, 2005.
- [6] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1-54, 2015.
- [7] G. R. S. Murthy and R. S. Jadon, "Hand gesture recognition using neural networks," in *Proceedings of the IEEE 2nd International Advance Computing Conference (IACC '10)*, pp. 134-138, IEEE, Patiala, India, February 2010.
- [8] S. Lechevalier, J. Nishimura, and C. Storz, "Diversity in patterns of industry evolution: how an intrapreneurial regime contributed to the emergence of the service robot industry," *Research Policy*, vol. 43, no. 10, pp. 1716-1729, 2014.
- [9] Y. Gao and J. Liu, "China's robotics successes abound," *Science*, vol. 345, no. 6196, p. 523, 2014.
- [10] R. A. Castillo Cruces and J. Wahrburg, "Improving robot arm control for safe and robust haptic cooperation in orthopaedic procedures," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 3, no. 4, pp. 316-322, 2007.
- [11] S. Hirose and M. Mori, "Biologically inspired snake-like robots," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO '04)*, pp. 1-7, IEEE, August 2004.
- [12] C. Ye, S. Ma, B. Li, and Y. Wang, "Turning and side motion of snake-like robot," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '04)*, pp. 5075-5080, IEEE, May 2004.
- [13] V. Vapnik and S. Kotz, *Estimation of Dependences Based on Empirical Data*, Springer, 2006.
- [14] V. Cherkassky and F. M. Mulier, *Learning From Data: Concepts, Theory, and Methods*, John Wiley & Sons, New York, NY, USA, 2007.
- [15] H. Zhong, C. Miao, Z. Shen, and Y. Feng, "Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings," *Neurocomputing*, vol. 128, pp. 285-295, 2014.
- [16] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144-152, ACM, July 1992.
- [17] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [18] P. B. Schiilkop, C. Burgest, and V. Vapnik, "Extracting support data for a given task," in *Proceedings of the 1st International Conference on Knowledge Discovery & Data Mining*, pp. 252-257, August 1995.
- [19] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, Mass, USA, 2013.
- [20] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, Southeast Asia edition, 2006.
- [21] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
- [22] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [23] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "Training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144-152, ACM, July 1992.
- [24] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*, pp. 760-766, Springer, New York, NY, USA, 2010.
- [25] C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Systems with Applications*, vol. 31, no. 2, pp. 231-240, 2006.
- [26] C.-L. Huang and J.-F. Dun, "A distributed PSO-SVM hybrid system with feature selection and parameter optimization," *Applied Soft Computing Journal*, vol. 8, no. 4, pp. 1381-1391, 2008.

- [27] M. J. Abdi and D. Giveki, "Automatic detection of erythemato-squamous diseases using PSO-SVM based on association rules," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 603–608, 2013.
- [28] P. J. García Nieto, E. García-Gonzalo, F. Sánchez Lasheras, and F. J. De Cos Juez, "Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability," *Reliability Engineering and System Safety*, vol. 138, pp. 219–231, 2015.
- [29] J. Tang, C. Deng, and G. B. Huang, "Extreme learning machine for multilayer perceptron," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 27, no. 4, pp. 809–821, 2016.
- [30] Z. Ju and H. Liu, "A unified fuzzy framework for human-hand motion recognition," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 5, pp. 901–913, 2011.
- [31] Z. Ju and H. Liu, "Human hand motion analysis with multisensory information," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 2, pp. 456–466, 2014.

Research Article

Rhythmic Oscillations of Excitatory Bursting Hodgkin-Huxley Neuronal Network with Synaptic Learning

Qi Shi, Fang Han, Zhijie Wang, and Caiyun Li

Engineering Research Center of Digitized Textile & Apparel Technology, College of Information Science and Technology, Donghua University, Shanghai 201620, China

Correspondence should be addressed to Fang Han; yadiahhan@163.com

Received 27 November 2015; Accepted 24 February 2016

Academic Editor: Hiroki Tamura

Copyright © 2016 Qi Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rhythmic oscillations of neuronal network are actually kind of synchronous behaviors, which play an important role in neural systems. In this paper, the properties of excitement degree and oscillation frequency of excitatory bursting Hodgkin-Huxley neuronal network which incorporates a synaptic learning rule are studied. The effects of coupling strength, synaptic learning rate, and other parameters of chemical synapses, such as synaptic delay and decay time constant, are explored, respectively. It is found that the increase of the coupling strength can weaken the extent of excitement, whereas increasing the synaptic learning rate makes the network more excited in a certain range; along with the increasing of the delay time and the decay time constant, the excitement degree increases at the beginning, then decreases, and keeps stable. It is also found that, along with the increase of the synaptic learning rate, the coupling strength, the delay time, and the decay time constant, the oscillation frequency of the network decreases monotonically.

1. Introduction

Neural systems could exhibit rhythmic oscillations, which are a type of synchronous state. Synchronization in neural systems is thought to be important for processing of sensory information and motor function [1], but the occurrence of synchronization in some specific areas of the brain may also be associated with some diseases, such as the epilepsy and Parkinson's disease [2]. Because of the importance of synchronization in neural systems, it has been studied for a long time from many aspects in neuroscience research [3–7].

There are up to 10^{11} neurons in human brain, and each neuron is connected to approximately 10^4 other neurons. Neurons are coupled with each other by electrical or chemical synapses, and chemical synapses are dominant in quantity. The chemical synapse is related to the exchange of neurotransmitters between neurons and can be inhibitory or excitatory. Neurons coupled by different types of synapses constitute different networks, in which the dynamical behaviors can be very different [8]. Han et al. [9] found that the synchronization for inhibitory neuronal systems is more robust and stable than that for excitatory neuronal systems,

so they investigated robust synchronization for a globally coupled inhibitory neuronal network. However, they did not study dynamical behaviors for excitatory neuronal systems, which should also be explored to reveal the underlying mechanisms of rhythmic oscillations in neural systems.

Synaptic plasticity is a prevalent feature of biological neural systems and considered to be critical for memory and learning functions of brains. Synaptic efficacy could be regulated by the plasticity at a variety of time scales, like from milliseconds to minutes. To study how synaptic plasticity works in neural systems, many synaptic learning rules, such as Hebbian learning rule and STDP rule, are proposed. And accordingly, the dynamics of neuronal systems under the influence of synaptic plasticity has been explored [10–13]. For example, Han et al. [13] investigated the dynamical properties of Newman-Watts (NW) small-world neuronal networks with a short-term synaptic plasticity named Oja rule and got some interesting findings for electrically and chemically coupled neuronal networks, respectively. But the study is insufficient, because synapses have more detailed structure which influences the dynamics of the neural systems a lot, especially for chemical ones.

This paper aims to find out how the synaptic learning and chemical synaptic parameters influence rhythmic oscillations in excitatory neuronal network. It is organized as follows. The model of a globally coupled excitatory bursting Hodgkin-Huxley (HH) neuronal network is presented in Section 2. The results of simulations on the excitatory neuronal network, including the effects of the coupling strength and the synaptic learning rate and the effects of chemical synaptic parameters, are presented in Section 3. The conclusions are given in Section 4.

2. Model and Dynamics

The traditional Hodgkin-Huxley (HH) model neuron only emits spikes. By incorporating a slow calcium ionic channel into the HH model, a modified model neuron which could emit bursts can be obtained [9]. In this paper, the modified HH model neuron is used to construct neuronal network.

2.1. Model Neuron. The equations and parameters of a single modified HH model neuron can be described as follows (the membrane potential is measured in mV and time in ms):

$$C_m \frac{dV}{dt} = -g_{Na} m_{\infty}^3 h (V - E_{Na}) - g_K n^4 (V - E_K) - g_{Ca} M^2 H (V - E_{Ca}) - g_L (V - E_L) + I^{app}, \quad (1)$$

where the parameter V is the membrane potential of the modified model neuron. In (1), the activation or inactivation variables h, n, M , and H are associated with Na^+ , K^+ , and Ca^{2+} voltage-dependent ion currents, respectively, which changes quickly and satisfies the following steady-state functions:

$$\frac{dh}{dt} = \theta (\alpha_h (1 - h) - \beta_h h), \quad (2a)$$

$$\frac{dn}{dt} = \theta (\alpha_n (1 - n) - \beta_n n), \quad (2b)$$

$$\frac{dM}{dt} = \tau_M^{-1} (-M + M_{\infty}), \quad (2c)$$

$$\frac{dH}{dt} = \tau_H^{-1} (-H + H_{\infty}). \quad (2d)$$

In (1), the activation variable m_{∞} is related to transient sodium current, which changes very fast and satisfies the steady-state function:

$$m_{\infty} = \frac{\alpha_m}{(\alpha_m + \beta_m)}, \quad (3a)$$

$$\alpha_m = \frac{-0.1 (V_i + 35)}{(\exp(-0.1 (V_i + 35)) - 1)}, \quad (3b)$$

$$\beta_m = 4 \exp\left(\frac{-(V_i + 60)}{18}\right). \quad (3c)$$

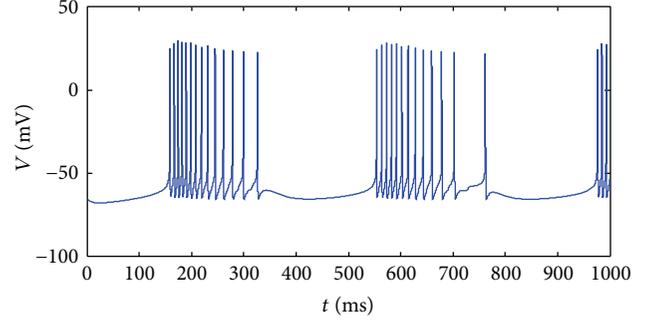


FIGURE 1: The firing pattern of a single neuron.

In (2a)–(2d), the related variables satisfy the following functions:

$$\alpha_h = 0.07 \exp\left(\frac{-(V_i + 58)}{20}\right), \quad (3d)$$

$$\beta_h = \frac{1}{(\exp(-0.1 (V_i + 28)) + 1)}. \quad (3e)$$

$$\alpha_n = \frac{-0.01 (V_i + 34)}{(\exp(-0.1 (V_i + 34)) - 1)}, \quad (3f)$$

$$\beta_n = 0.125 * \exp\left(\frac{-(V_i + 44)}{80}\right). \quad (3g)$$

$$M_{\infty} = \frac{1}{(1 + \exp(-(V_i + 57)/6.2))}, \quad (3h)$$

$$H_{\infty} = \frac{1}{(1 + \exp((V_i + 81)/4))}, \quad (3i)$$

$$\tau_M = 0.612 + \frac{1}{(\exp(-(V_i + 132)/16.7) + \exp((V_i + 16.8)/18.2))}, \quad (3j)$$

$$\tau_H = \begin{cases} \exp\left(\frac{(V_i + 467)}{66.6}\right), & V_i < -80 \text{ mV}, \\ 28 + \exp\left(\frac{-(V_i + 22)}{10.5}\right), & V_i \geq -80 \text{ mV}. \end{cases} \quad (3k)$$

The values of parameters in (1), (2a), (2b), (2c), and (2d) are set as follows: $C_m = 1 \mu\text{F}/\text{cm}^2$, $g_{Na} = 35 \text{ ms}/\text{cm}^2$, $E_{Na} = 55 \text{ mV}$, $g_K = 9 \text{ ms}/\text{cm}^2$, $E_K = -90 \text{ mV}$, $g_{Ca} = 3 \text{ ms}/\text{cm}^2$, $E_{Ca} = 120 \text{ mV}$, $g_L = 0.1 \text{ ms}/\text{cm}^2$, and $E_L = -65 \text{ mV}$, $\theta = 5$.

In (1), I^{app} is injected current (in $\mu\text{A}/\text{cm}$). Neurons could exhibit different firing patterns for different values of I^{app} . If I^{app} belongs to $[-0.95, -0.25]$, the modified HH neuron exhibits bursting behaviors (see Figure 1). The value of the injected current is set as -0.5 in this figure.

2.2. Neuronal Network Model. By using the above model neuron, we can set up a globally coupled neuronal network with 50 neurons. Considering that most of synapses are chemical ones and the weights of synapses are always changing, we apply a synaptic learning rule, Oja learning rule [13], to

the chemical synapses in the neuronal network model. The globally coupled neuronal network is described as follows:

$$C_m \frac{dV_i}{dt} = -g_{Na} m_{\infty}^3 h(V_i - E_{Na}) - g_K n^4 (V_i - E_K) - g_{Ca} M^2 H(V_i - E_{Ca}) - g_L (V_i - E_L) + I_i^{app} - \sigma \sum_{j=1}^{50} W_{ij} a_{ij} s(t) (V_i - E_{syn}), \quad (4)$$

where the parameter V_i is the membrane potential of the i th neuron, $i = 1, 2, \dots, 50$, σ is the maximal synaptic conductance, a_{ij} is the element of the adjacency matrix and

equals 0 or 1 depending on whether there is a synapse between neurons i and j , and E_{syn} is the reversal potential.

In (4), W_{ij} is the value of the weight between neurons i and j , which changes according to the following equation:

$$\dot{W}_{ij} = L * \arctan [V_i (V_j - V_i W_{ij})], \quad (5)$$

where L denotes the learning rate of the chemical synapse and V_i and V_j denote the membrane potentials of neuron i and neuron j , respectively.

In (4), $s(t)$ is the level of openness of ion channels, $s(t) = \sum_m s_m(t)$, where $s_m(t)$ is associated with the m th spike of the presynaptic neuron, which is characterized by the following α -function:

$$s_m(t) = \begin{cases} 0, & t < t_m^f + d \\ \frac{(\exp(-(t - t_m^f - d)/\tau_d) - \exp(-(t - t_m^f - d)/\tau_r))}{(\tau_d - \tau_r)}, & t \geq t_m^f + d, \end{cases} \quad (6)$$

where t_m^f is the m th firing time of the presynaptic neuron and τ_d , τ_r , and d are the decay time constant, rise time constant, and synaptic delay of the synapse, respectively.

The chemical synapse can be excitatory or inhibitory. If E_{syn} is high enough, the synapse is excitatory; if E_{syn} is low enough, the synapse is inhibitory. In this paper, E_{syn} is set as 0 mV to make all the synapses excitatory.

3. Simulation Results

In this section, the dynamical properties of the excitatory neuronal network are studied. I^{app} for each neuron is chosen randomly in $[-0.5, -0.4]$ so that all the neurons can show different bursting behaviors. In the next simulations, the time step is 0.01 ms and the total simulation time is 1000 ms.

We first explore the variations of synaptic weights under the influence of the synaptic leaning rule. Figures 2(a) and 2(b) show the spike trains of two arbitrarily chosen connected neurons in the network and Figure 2(c) shows their synaptic weights. The learning rate and the coupling strength are set as $L = 2$, $\sigma = 0.05$.

In order to see the variations in Figure 2 clearly, we extract some points in Figure 2 and list their values in Table 1. In Table 1, the values of the membrane potentials of the two neurons and the weights between them at 550 ms, 593.47 ms, 644.80 ms, 668.00 ms, 698.34 ms, and 750 ms are shown, respectively. It can be seen that, at 550 ms and 750 ms, both of the two neurons are at rest and the weights between them almost do not change; at 593.47 ms and 668.00 ms, neuron 1 is excited and neuron 2 is at rest, and obviously the weight from neuron 1 to neuron 2 (W_{12}) increases and the weight from neuron 2 to neuron 1 (W_{21}) decreases; at 644.80 ms and 698.34 ms, neuron 2 is excited and neuron 1 is at rest, and obviously the weight from neuron 2 to neuron 1 (W_{21}) increases and the weight from neuron 1 to neuron 2 (W_{12}) decreases.

So from Figure 2 and Table 1, we can come to the conclusion that the synaptic weights between two neurons do not change when both of them are in the state of rest, and if one neuron is excited and the other one is at rest, the weight from the excited neuron to the rest one is strengthened, and the weight from the rest one to the excited one is weakened. It means that, when ion currents are conducted between two neurons by the chemical synapses, the synaptic efficacy would be strengthened from the active neuron to the inactive one, and on the contrary, the synaptic efficacy would be weakened from the inactive one to the active one. This is plausible in real neural systems.

3.1. Effects of Coupling Strength and Synaptic Learning Rate.

In order to see the dynamical properties of the whole network, some spatiotemporal patterns of the network with different values of coupling strength σ and learning rate L are plotted in Figure 3.

Figure 3 shows the effects of different parameters on the overall dynamics of the network, which can be divided into two groups. The first group of (a), (b), and (c) shows the effect of the coupling strength and it can be seen that the network with larger coupling strength is less excited. The second group of (d), (e), and (f) demonstrates the effect of the learning rate and it can be found that the neuronal network with $L = 2.5$ is much more excited than that with $L = 0.25$ and $L = 10$. However, it seems difficult to see the exact variation tendency of the influence of the parameters, so we need to use some measurements to carry out quantitative analysis.

In the next simulations, we use two measurements, excitement degree D_{exc} and oscillation frequency F_{mean} , to describe the properties of the neuronal networks. All the results are obtained by average of 20-time repetition.

The measurement excitement degree is defined as $D_{exc} = \langle n_{exc}(t) \rangle$, where $n_{exc}(t)$ indicates the proportion

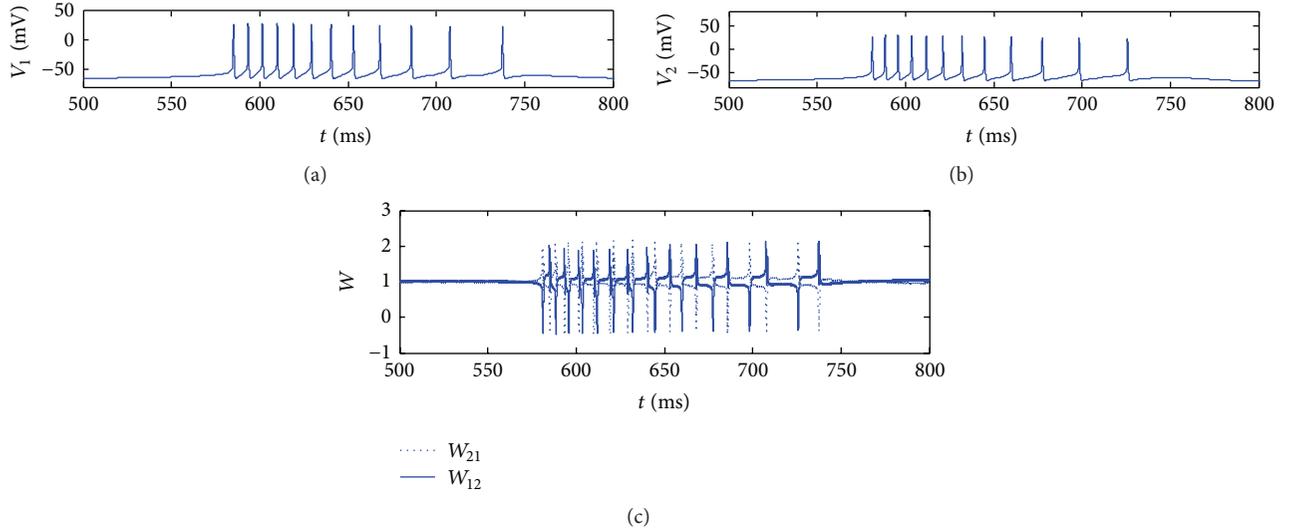


FIGURE 2: Spike trains of two arbitrarily chosen connected neurons in the network ((a), (b)) and synaptic weights between them in two directions, respectively (c).

TABLE 1: The values of some extracted points as examples in Figure 2.

Time (ms)	550.00	593.47	644.80	668.00	698.34	750
V_1 (mV)	-62.6378	28.6678	-62.4997	24.6842	-59.6050	-60.6416
V_2 (mV)	-62.8390	-55.9874	26.0039	-60.5042	22.8384	-58.9853
W_{12}	1.0189	1.3345	-0.4245	1.4635	-0.2923	0.9867
W_{21}	0.9811	-0.3117	1.4746	-0.3372	1.5400	1.0451

of the number of excited neurons in the neuronal network at the moment and $\langle \cdot \rangle$ indicates averaging over time. This measurement reflects the extent of excitement of the neuronal network. Here, a neuron is thought to be excited once its spike exceeds a threshold, like -50 mV, or else it is thought to be inhibited. Larger value of the quantity D_{exc} means the neuronal network is more excited.

As the network activity is very rhythmic, we can also check the oscillation frequency of the neuronal network. The oscillation frequency F_{mean} can be defined as $F_{\text{mean}} = (1/N) \sum_{i=1}^N f_i$, $f_i = (\phi_i^{t_2} - \phi_i^{t_1}) / 2\pi(t_2 - t_1)$, where f_i is the bursting frequency of neuron i and $\phi_i^{t_2}$ and $\phi_i^{t_1}$ are the burst-phases at times t_2 and t_1 ($t_2 > t_1$), respectively.

By using the above two measurements, the dynamical properties of the excitatory neuronal network are studied. Firstly, we study the influence of the coupling strength and the learning rate on excitement degree and oscillation frequency. The parameters of delay time d , decay time τ_d , and rise time τ_r are set as 8 ms, 12 ms, and 0.3 ms, respectively.

Figures 4(a) and 4(b) display the effect of coupling strength on excitement degree and oscillation frequency, respectively. It can be seen that both excitement degree and oscillation frequency decrease monotonically along with the increase of the coupling strength. We may draw a conclusion that high coupling strength could decrease the excitement in a certain range.

Figures 5(a) and 5(b) show the effect of the learning rate on excitement degree and oscillation frequency for

the network. It can be seen that excitement degree increases at first and then decreases along with the increasing learning rate. It achieves the maximum with $L = 2.5$. When it comes to the oscillation frequency (see Figure 5(b)), it increases at the very beginning and achieves the maximum 18 Hz and then decreases monotonously to 9 Hz at the end. It implies that the learning rate can regulate the excitement and oscillation periods of the network at a certain range.

3.2. Effects of the Chemical Synaptic Parameters. Then, we explore the influence of the synaptic parameters of chemical synapses. The values of the coupling strength, the learning rate, and the rise time are set as 0.25, 0.25, and 0.3 ms, respectively.

Figure 6 also can be divided into two groups. The first group of (a), (b), and (c) shows the effects of the delay on the overall dynamics of the network and it can be found that the network with delay time $d = 3$ ms is more active. The second group of (d), (e), and (f) demonstrates the effects of the decay time constant. Still, we need to quantify these phenomena.

To study the effect of the delay time and the decay time constant, we plot Figures 7(a), 7(b), 8(a), and 8(b).

Figures 7(a) and 7(b) show the effect of the delay time on excitement degree and oscillation frequency for the network. The parameter of the decay time is set as $\tau_d = 12$ ms. It can be seen from Figure 7(a) that excitement degree increases at first, then decreases a little, and then keeps stable along with the increasing synaptic delay. It achieves the maximum

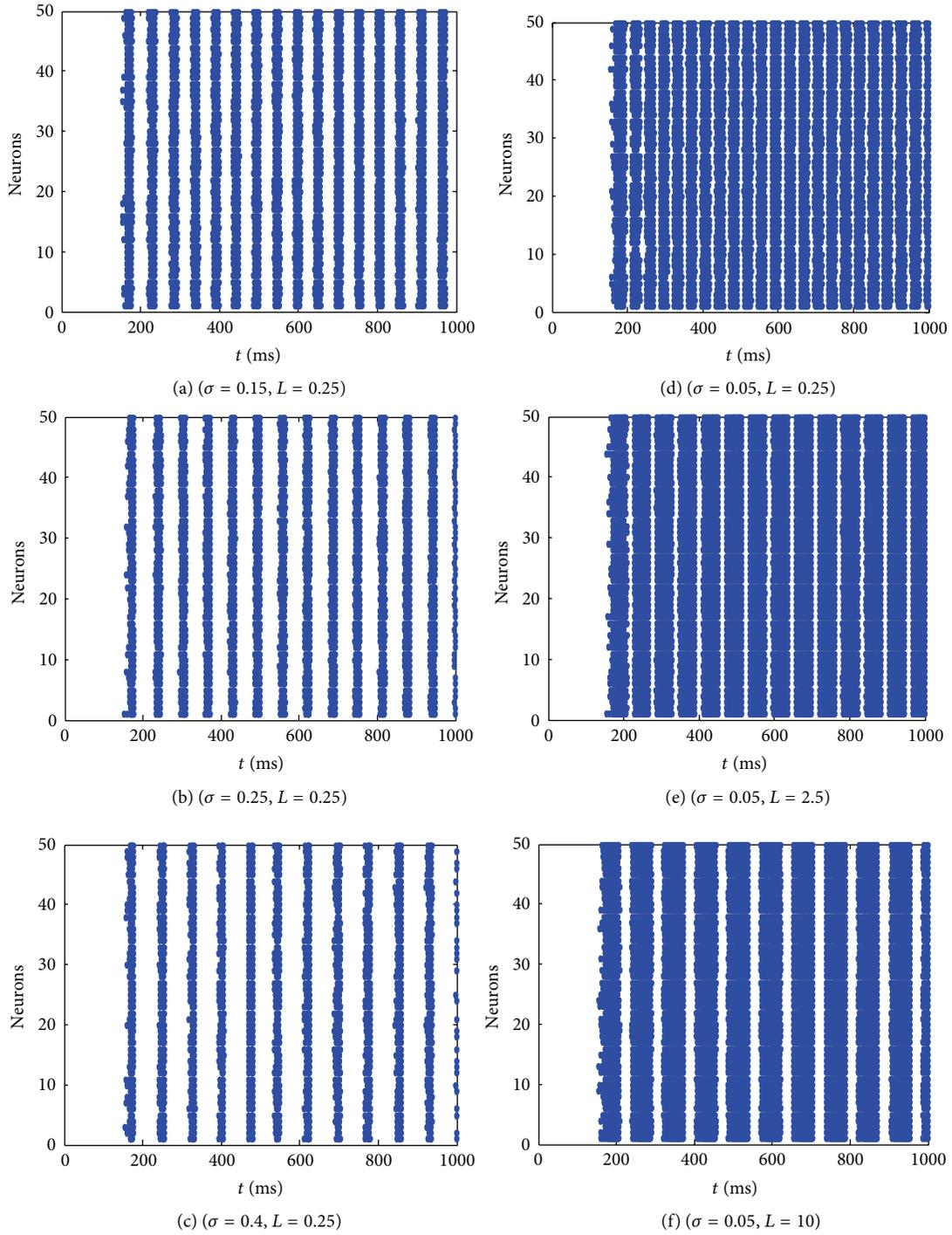


FIGURE 3: Some spatiotemporal patterns of the coupled excitatory neuronal network with different values of coupling strength and learning rate.

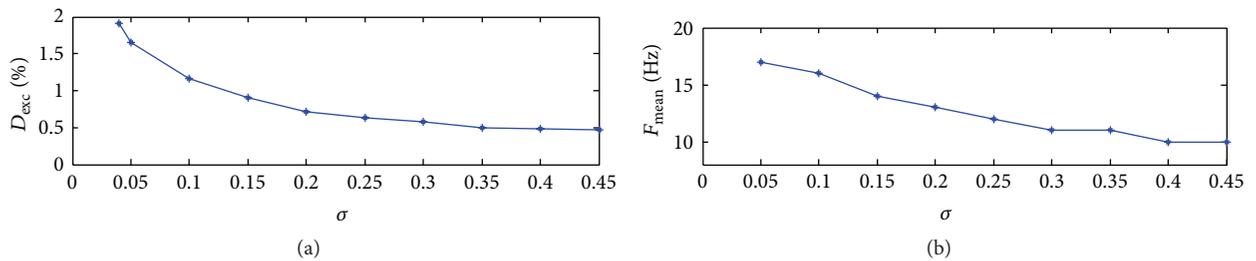


FIGURE 4: Effect of the coupling strength on excitement degree (a) and oscillation frequency (b) with $L = 0.25$.

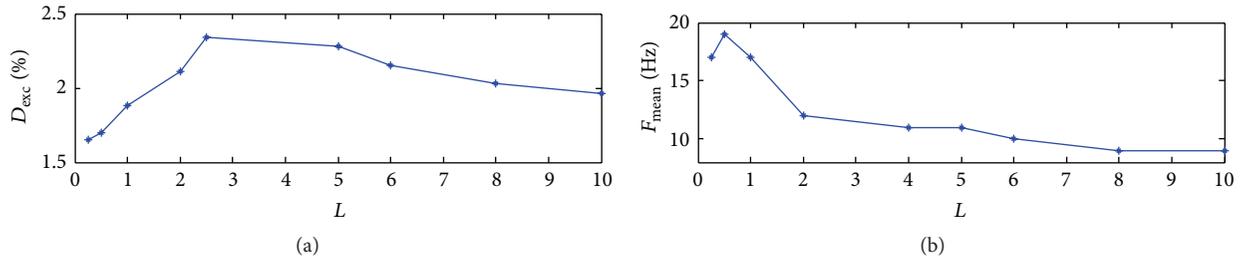


FIGURE 5: Effect of the learning rate on excitement degree (a) and oscillation frequency (b) with $\sigma = 0.05$.

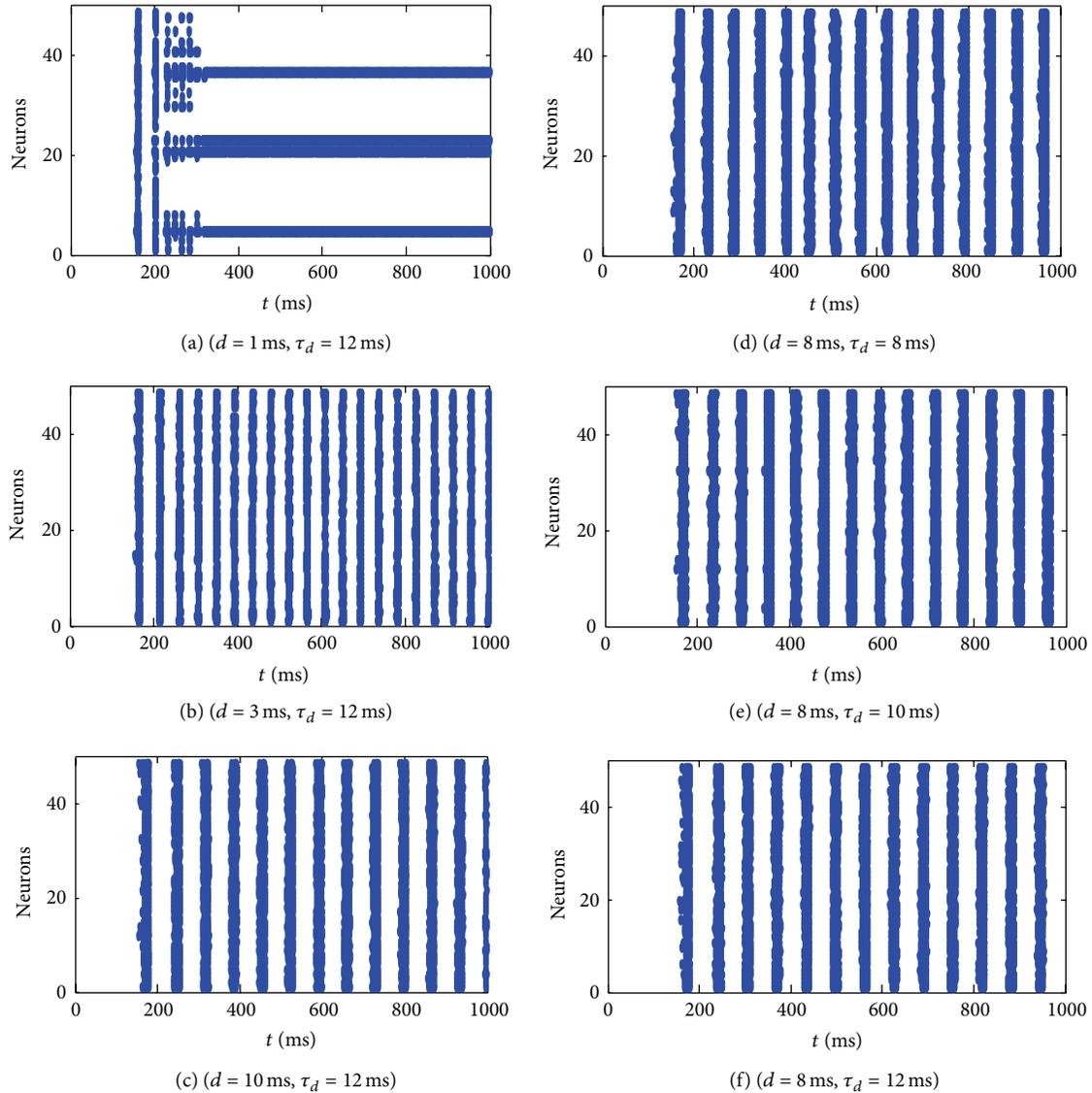


FIGURE 6: Some spatiotemporal patterns of the excitatory neuronal network under the different values of synaptic delay and decay time constant.

with the delay $d = 3$ ms. From Figure 7(b) it can be seen that the oscillation frequency decreases monotonically with the increasing synaptic delay. It implies that synaptic delay time can improve the excitement of the network but decrease periods of the rhythmic behaviors for neuronal networks.

The effects of the decay time constant on excitement degree and oscillation frequency for the network are shown in Figures 8(a) and 8(b). The value of the synaptic delay is set as $d = 8$ ms. It can be seen that the excitement degree decreases in general except for the very beginning and the oscillation

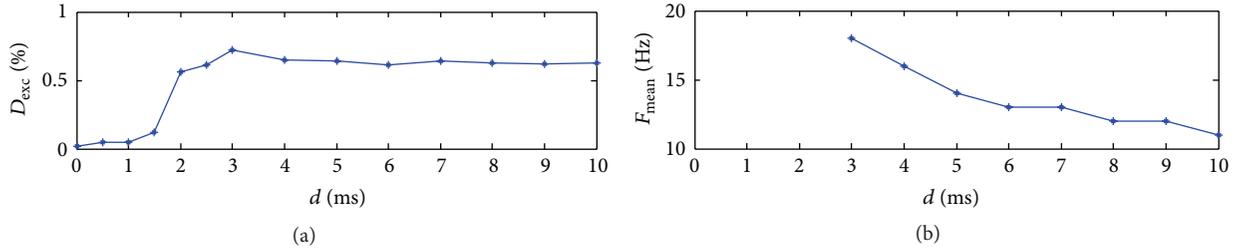


FIGURE 7: Effect of the synaptic delay on excitement degree (a) and oscillation frequency (b) with $\tau_d = 12$ ms.

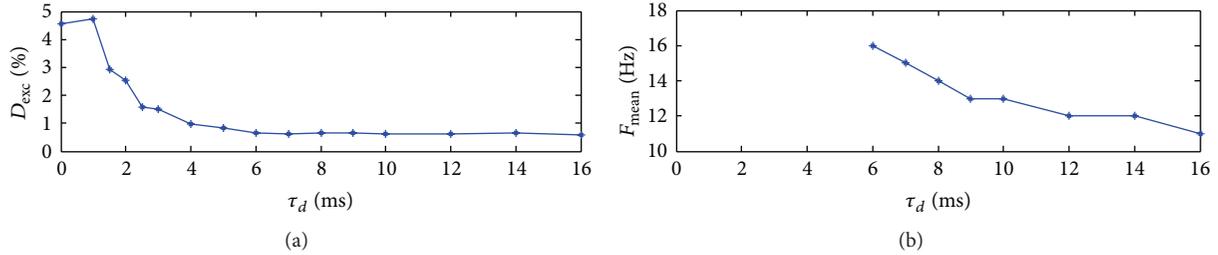


FIGURE 8: Effect of the decay time constant on excitement degree (a) and oscillation frequency (b) with $d = 8$ ms.

frequency also decreases monotonically when the decay time constant is increasing. We may draw a conclusion that the decay time constant can weaken the excitement of the network and decrease periods of the rhythmic oscillations.

4. Conclusions and Discussions

The dynamics of an excitatory neuronal network with a synaptic learning rule are explored in this paper. It is found that both the coupling strength and the synaptic learning rate have great influence on the excitement degree and oscillation frequency of the neuronal network. The increasing of the coupling strength can weaken the extent of excitement, whereas increasing the synaptic learning rate makes the network more excited in a certain range. What is more, the increasing of both the coupling strength and the learning rate could decrease the oscillation frequency of the network. These results can be demonstrated well by the spatiotemporal patterns of the neuronal networks.

Besides, it is found that the synaptic parameters of delay time and decay time constant also have great influence on the excitement degree and oscillation frequency of the network. Along with the increasing of the delay time and the decay time constant, the oscillation frequency of the network decreases monotonically, and the excitement degree increases at the beginning, then decreases, and keeps stable.

In addition, we also check the situations for inhibitory neuronal networks. We may draw a conclusion that the coupling strength and the synaptic learning rate have little influence on the excitement of the network. Along with the increasing of the coupling strength and learning rate, the excitement degree and oscillation frequency change hardly. But large decay time and delay time can decrease the oscillation frequency of the network.

The results we obtained in this paper may be instructive for understanding the information processing mechanisms of neural systems and for the study of controlling synchronization in neural systems. It is known that human brain waves are sometimes rhythmic, which relate to human's certain behavior or brain diseases. So, the study of rhythmic oscillations can help us know about our brains and cure some diseases. By regulating some parameters of synapses, the rhythmic behaviors of neural systems could be altered.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants nos. 11572084, 11472061, and 71371046), the Fundamental Research Funds for the Central Universities, and DHU Distinguished Young Professor Program.

References

- [1] P. J. Uhlhaas and W. Singer, "Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology," *Neuron*, vol. 52, no. 1, pp. 155–168, 2006.
- [2] K. Abuhassan, D. Coyle, and L. Maguire, "Compensating for thalamocortical synaptic loss in Alzheimer's disease," *Frontiers in Computational Neuroscience*, vol. 8, article 65, 2014.
- [3] M. S. Baptista, F. M. Moukam Kakmeni, and C. Grebogi, "Combined effect of chemical and electrical synapses in Hindmarsh-Rose neural networks on synchronization and the rate of

- information,” *Physical Review E*, vol. 82, no. 3, Article ID 036203, 12 pages, 2010.
- [4] H. T. Yu, J. Wang, C. Liu, B. Deng, and X. Wei, “Delay-induced synchronization transitions in small-world neuronal networks with hybrid electrical and chemical synapses,” *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 21, pp. 5473–5480, 2013.
- [5] H. T. Yu, X. M. Guo, J. Wang, B. Deng, and X. Wei, “Spike coherence and synchronization on Newman-Watts small-world neuronal networks modulated by spike-timing-dependent plasticity,” *Physica A: Statistical Mechanics and its Applications*, vol. 419, pp. 307–317, 2015.
- [6] E. L. Lameu, C. A. Batista, A. M. Batista et al., “Suppression of bursting synchronization in clustered scale-free (rich-club) neuronal networks,” *Chaos*, vol. 22, Article ID 043149, 2012.
- [7] H. Yu, J. Wang, C. Liu, B. Deng, and X. Wei, “Delay-induced synchronization transitions in modular scale-free neuronal networks with hybrid electrical and chemical synapses,” *Physica A: Statistical Mechanics and its Applications*, vol. 405, pp. 25–34, 2014.
- [8] C. A. S. Batista, R. L. Viana, F. A. S. Ferrari, and S. R. Lopes, “Control of bursting synchronization in networks of Hodgkin-Huxley-type neurons with chemical synapses,” *Physical Review E*, vol. 87, Article ID 042713, 2013.
- [9] F. Han, Z. J. Wang, Y. Du, X. J. Sun, and B. Zhang, “Robust synchronization of bursting Hodgkin-Huxley neuronal systems coupled by delayed chemical synapses,” *International Journal of Non-Linear Mechanics*, vol. 70, pp. 105–111, 2015.
- [10] E. M. Izhikevich and N. S. Desai, “Relating STDP to BCM,” *Neural Computation*, vol. 15, no. 7, pp. 1511–1523, 2003.
- [11] R. R. Borges, F. S. Borges, E. L. Lameu et al., “Effects of the spike timing-dependent plasticity on the synchronisation in a random Hodgkin-Huxley neuronal network,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 34, pp. 12–22, 2016.
- [12] H. Yu, X. Guo, J. Wang, B. Deng, and X. Wei, “Spike coherence and synchronization on Newman-Watts small-world neuronal networks modulated by spike-timing-dependent plasticity,” *Physica A: Statistical Mechanics and its Applications*, vol. 419, pp. 307–317, 2015.
- [13] F. Han, M. Wiercigroch, J. A. Fang, and Z. J. Wang, “Excitement and synchronization of small-world neuronal networks with short-term synaptic plasticity,” *International Journal of Neural Systems*, vol. 21, no. 05, pp. 415–425, 2011.