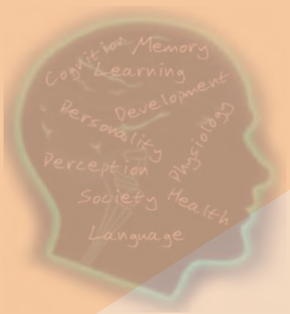
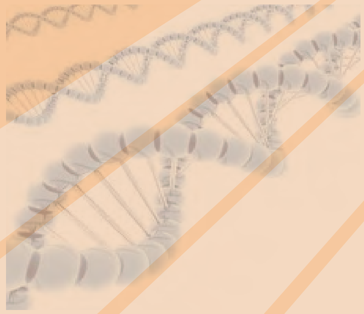
BMI Brain-Mind Institute

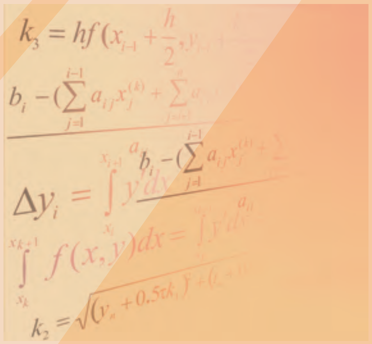
FOR FUTURE LEADERS OF BRAIN-MIND RESEARCH

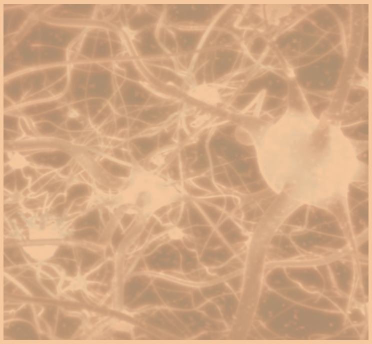




Proceedings of 2014 BMI the Third International Conference on Brain-Mind

July 12-13
Beijing, China





BMI Press ISBN 978-0-9858757-0-1 US$35

Contents

Messages From the Chairs

We have entered the third year of the Brain-Mind Institute (BMI) and the International Conference on Brain-Mind (ICBM).

Several U.S. government-funding agencies have started their funding programs as part of the President Barack Obama's BRAIN Initiative. The European Union's Human Brain Project is going on. After a series of workshops, China seems to be still preparing its own brain project. Understanding how the brain works is one of the last frontiers of the human race. The progress on this subject will lead to not only new technology but also human improved ways to develop human societies.

Therefore, BMI has been an earlier platform that treats every human activity as a part of science, including, but not limited to, biology, neuroscience, psychology, computer science, electrical engineering, mathematics, intelligence, life, laws, policies, societies, politics, and philosophy. BMI plans to further span its service to the scientific community and public by promoting science in human activities.

This year BMI offered BMI 831 Cognitive Science for Brain-Mind Research and BMI 871 Computational Brain-Mind. We would like to thank the Institute of Automation of the Chinese Academy of Sciences (CASIA) for hosting the BMI 831 and BMI 871 classes as well as ICBM 2014. The Brain-Mind Institute and the Brainnetome Center of CASIA co-sponsored and co-organized BMI 2014. As BMI planned to host BMI courses and ICBM at an international location, this year, it is Beijing. BMI 2014 co-locates with the World Congress on Computational Intelligence 2014.

As a multi-disciplinary communication platform, ICBM is an integrated part of the BMI program. ICBM 2014 includes invited talks, talks from submitted papers, and talks from submitted abstracts. As last year, ICBM talks will be video recorded and available publicly through the Internet.

As before, the BMI Program Committee tries to be open-minded in its review of submissions. This open-mindedness is necessary for all subjects of science, not just brain-mind subjects.

Welcome to Beijing!

Tianzi Jiang, General Co-Chair and Program Co-Chair
Juyang Weng, General Chair and Program Chair

Committees

Advisory Committee

- Dr. Stephen Grossberg, Boston University, USA
- Dr. James McClelland, Stanford University, USA
- Dr. James Olds, George Mason University, USA
- Dr. Linda Smith, Indiana University, USA
- Dr. Mriganka Sur, Massachusetts Institute of Technology (MIT), USA

Program Committee

- Dr. Yiannis Aloimonos, University of Maryland, USA
- Dr. Minoru Asada, Osaka University, Japan
- Dr. Luis M. A. Betterncourt, Los Alamos National Laboratory, USA
- Dr. Angelo Cangelosi, University of Plymouth, UK
- Dr. Shantanu Chakrabartty, Michigan State University, USA
- Dr. Yoonsuck Choe, Texas A&M University, USA
- Dr. Wlodzislaw Duch, Nicolaus Copernicus University, Poland
- Dr. Zhengping Ji, Los Alamos National Laboratory, USA
- Dr. Tianzi Jiang, Institute of Automation of Chinese Academy of Sciences, China
- Dr. Yaochu Jin, University of Surrey, UK
- Dr. Pinaki Mazumder, University of Michigan, USA
- Dr. Yan Meng, Stevens Institute of Technology, USA
- Dr. Ali A. Minai, University of Cincinnati, USA
- Dr. Jun Miao, Institute of Computing Technology, Chinese Academy of Sciences, China
- Dr. Thomas R. Shultz, McGill University, Canada
- Dr. Linda Smith, Chancellor's Professor, Indiana University, USA
- Dr. Juyang Weng, Michigan State University, USA
- Dr. Xiaofeng Wu, Fudan University, China
- Dr. Ming Xie, Nanyang Technological University, Singapore
- Dr. Xiangyang Xue, Fudan University, China
- Dr. Chen Yu, Indiana University, USA
- Dr. Cha Zhang, Microsoft Research, Redmond WA, USA
- Dr. Shuqing Zeng, GM Research and Development, General Motors, USA

Conference Committee

- Dr. James H. Dulebohn, Michigan State Universit, USA
- Dr. Jianda Han, Shengyang Institute of Automation, China
- Dr. Kazuhiko Kawamura, Vanderbilt University, USA
- Dr. Minho Lee, Kyungpook National University, Korea
- Dr. Gonzalo Munevar, Lawrence Technological University, USA
- Dr. Danil Prokhorov, Toyota Research Institute of America, USA
- Dr. Robert Reynolds, Wayne State University, USA
- Dr. Katharina J. Rohlfing, Bielefeld University, Germany
- Dr. Matthew Schlesinger, Southern Illinois University, USA
- Dr. Jochen Triesch, Frankfurt Institute for Advanced Studies, Germany
- Dr. Hiroaki Wagatsuma, Kyushu Institute of Technology, Japan

MSU Steering Committee

- Dr. Andrea Bozoki, Michigan State University
- Dr. Jay P. Choi, Michigan State University
- Dr. Lynwood G. Clemens, Michigan State University
- Dr. Kathy Steece-Collier, Michigan State University
- Dr. Steve W. J. Kozlowski, Michigan State University
- Dr. Jiaguo Qi, Michigan State University
- Dr. Frank S. Ravitch, Michigan State University
- Dr. Fathi Salem, Michigan State University
- Dr. George Stockman, Michigan State University
- Dr. Yang Wang, Michigan State University
- Dr. Juyang Weng, Michigan State University

Sponsors

International Neural Network Society

IEEE Computational Intelligence Society

Toyota

Invited Talks

Brainnetome: A New Avenue to Understand the Brain and its Disorders

Tianzi Jiang *(Brainnetome Center, Institute of Automation)*

Abstract

The Brainnetome (Brain-net-ome) is a new "-ome" in which the brain network is its basic research unit. It includes at least the following essential components: network topological structure (connectome), performance, dynamics, manifestation of functions and malfunctions of brain on different scales, genetic basis of brain networks, and simulating and modeling brain networks on supercomputing facilities. Here we will review progress on some aspects of the Brannetome, including Brainnetome atlas, Brainnetome-wise Association Studies (BWAS) of neurological and psychiatric diseases, such as schizophrenia and Alzheimer's disease, and how the Brainnetome meets genome, and so on. It envisions that the Brainnetome will become an emerging co-frontier of brain imaging, information technology, neurology and psychiatry. Some long-standing issues in neuropsychiatry may be solved by combining the Brainnetome with genome.

Short Biography

Tianzi Jiang is Professor of Brain Imaging and Cognitive Disorders, Institute Automation of Chinese Academy of Sciences (CASIA), and Professor of Queensland Brain Institute, University of Queensland. He is the Director of Brainnetome Center of CASIA and the Chinese Director of the Sino-French Laboratory in Computer Science, Automation and Applied Mathematics (LIAMA), one National Center for International Research. His research interests include neuroimaging, Brainnetome, imaging genetics, and their clinical applications in brain disorders and development. He is the author or co-author of over 200 reviewed journal papers in these fields and the co-editor of six issues of the Lecture Notes in Computer Sciences. He is Associate Editor of IEEE Transactions on Medical Imaging, IEEE Transactions on Autonomous Mental Development, Neuroscience Bulletin and an Academic Editor of PLoS One.

Visual Perceptual Learning and Its Brain Mechanisms: A New Perspective

Cong Yu *(Peking University)*

Abstract

Visual perceptual learning is regarded as a powerful tool to understand brain plasticity at the behavioral level. Learning is known to be specific to the trained retinal location and orientation, which places important constraints on perceptual learning theories, many of which assume that perceptual learning occurs in the early visual areas that are retinotopic and orientation selective.

However, we created new experimental paradigms to demonstrate that location and orientation specificities can be eliminated from perceptual learning. In a "double training" paradigm, location specific learning can transfer completely to a new retinal location following additional training at the new location with an irrelevant task. Similarly, in a training-plus-exposure (TPE) paradigm, orientation/direction-specific learning can transfer completely to an untrained new orientation/direction if an observer is also passively exposed to the new orientation/direction through an irrelevant task.

These results suggest that perceptual learning is more likely a high-level process that occurs beyond the retinotopic and orientation/direction selective visual cortex. What is being actually learned in perceptual learning? I will present evidence that perceptual learning may be a form of concept learning, in that the brain may learn a highly abstract "concept" of orientation/direction. On the other hand, why high-level perceptual learning shows specificity in the first place? I will present evidence that learning specificity may result from high-level learning not being able to functionally connect to the untrained visual inputs that are under-activated due to insufficient stimulation or suppression, as well as unattended, during training. It is the double training and TPE paradigms that bottom-up and top-down reactivate untrained inputs to establish functional connections and enable learning transfer.

Short Biography

Cong Yu received his Ph.D. in experimental psychology from University of Louisville in 1995. After postdoc trainings in basic and clinical vision sciences at University of Houston and UC Berkeley, he joined in the Inst. of Neuroscience, Chinese Academy of Sciences in 2003, the Inst. of Cognitive Neuroscience and Learning, Beijing Normal University in 2006, and the Department of Psychology, Peking University in 2012. He is currently a professor with the Department of Psychology and an Investigator with the Peking-Tsinghua Center for Life Sciences at Peking University.

Scale-Free Music of the Brain

Dezhong Yao *(University of Electronic Science and Technology of China)*

Abstract

To listen to brain activity as a piece of music, we proposed scale-free brainwave music (SFBM) technology, which translated the electroencephalogram (EEG) into musical notes according to the power law of both the EEG and music. In this talk, first shown is the general scale-free phenomena in the world, then a few versions of brain music are presented, they are music from one EEG channel, Quartet from multichannel EEGs, and music from both EEG and fMRI. Finally, potential application for mental health is discussed.

Short Biography

Dezhong Yao, PhD (1991, 2005), Professor (1995-), ChangJiang Scholar Professor (2006-); Dean, School of Life Science and Technology (2001-), University of Electronic Science and Technology of China, Chengdu 610054, China; Director, Key Laboratory for NeuroInformation, Ministry of Education of China (2010-); Head, Domestic Team of the 111 Project on NeuroInformation, Ministry of Education of China (2011-). His main research interests are Brain-X interaction including brain-computer interface, brainwave music interface, and Neuro-Imaging on music, epilepsy, cerebral palsy and plasticity. Since 1990, he has published 100 more peer-reviewed international journal papers, with 1000 more citations. He won the Outstanding Youth Resaerch Fund of NSFC (2005), and the first class Natural Science Reward of the Ministry of Education (2010). Websites: http://www.neuro.uestc.edu.cn/bci/member/yao/yao.asp, and http://scholar.google.com/citations?user=ClUoWqsAAAAJ

The Brain Works like Bridge-Islands with Modulation

Juyang Weng *(Fudan University and Michigan State University)*

Abstract

On one hand neuroscience is rich in data and poor in theory. On the other hand, many computer scientists are busy with engineering inspired methods, not motivated by brain inspired methods. However, in this talk, I argue that it is no longer true that "we do not know how the brain works". The knowledge of computer science is also necessary to understand how the brain works. Supported by a series of experimental studies known as Where What Networks (WWN-1 through WWN-8), I present an overarching but intuitive analogical model called bridge-islands. Each island is either a sensor (e.g., an eye or an ear) or an effector (an arm, or a gland). The brain is a multi-exchange bridge that connects to all the islands in bidirectionally. It is not productive to model the brain statically as a connected set of Brodmann areas, because in the born blind, the visual areas are automatically assigned to audition and touch. Therefore, the bridge-island model describes how various brain areas emerge from pre-natal and post-natal activities based on largely statistics. In other words, the brain wires itself. We also discuss how the self-wired basic circuits become motivated through four additional neural transmitters beyond glutamate and GABA --- serotonin, dopamine, acetylcholine, and norepinephrine.

Short Biography

Juyang (John) Weng is a professor at the Dept. of Computer Science and Engineering, the Cognitive Science Program, and the Neuroscience Program, Michigan State University, East Lansing, Michigan, USA, and a Changjiang visiting professor a Fudan University, Shanghai, China. He received his BS degree from Fudan University in 1982, his MS and PhD degrees from University of Illinois at Urbana-Champaign, 1985 and 1989, respectively, all in Computer Science. From August 2006 to May 2007, he was also a visiting professor at the Department of Brain and Cognitive Science of MIT. His research interests include computational biology, computational neuroscience, computational developmental psychology, biologically inspired systems, computer vision, audition, touch, behaviors, and intelligent robots. He is the author or coauthor of over two hundred fifty research articles, including a book Natural and Artificial Intelligence: Introduction to Computational Brain-Mind. He is an editor-in-chief of International Journal of Humanoid Robotics and an associate editor of the IEEE Trans. on Autonomous Mental Development, and the editor-in-chief of the Brain-Mind Magazine. He is instrumental in the establishment and operation of the Brain-Mind Institute, a nonprofit for cross-disciplinary education and research. He was an associate editor of IEEE Trans. on Pattern Recognition and Machine Intelligence, an associate editor of IEEE Trans. on Image Processing. He is a Fellow of IEEE.

The New Memory Technology to Support Brain-Like Computer

Luping Shi, Jing Pei, and Lei Deng Beata Jarosiewicz *(Tsinghua University)*

Abstract

The memory is the one of the main components of computer system as well as the key component for hand phone and cloud computing. For the last century, scaling has been being the main driving force for all of the current memory technologies in order to increase the density and reduce the cost. There are several kinds of memory to form a data storage hierarchy, such as DRAM,SRAM, Flash, optical disks, and hard disk drive. Although people have put a lot of effort to break scaling limitation, it could be estimated that all of above technologies might reach their scaling limits in about 10 to 15 years. On the other hand CPU also faces the same problem. Thus it is the time to find a new way to develop memory and CPU and to further develop computer. Brain like computer is one of the best approaches to solve the above problem. In this talk, the current statuses of memory and brain-like computer are briefly introduced. The requirement for the new memory technology to support brain-like computer is analyzed. The new memory should be capable of emulating some of brain functions. It should have the unique integrated function of storage and processing. The main problem and the possible approaches will be discussed.

Short Biography

Prof.Luping Shi received his Doctor of Science from University of Cologne, Germany in 1992. In 1993, he worked as a Post-doctoral fellow in Fraunhofer Institute of Applied Optics and Precision Instrument, Jena, Germany. From 1994 to 1996, he worked as a research fellow in Department of Electronic Engineering, City University of Hong Kong. From 1996 to 2013 he worked in data storage institute, Singapore as a senior scientist and division manager and led nonvolatile solid-state memory (NVM) and artificial cognitive memory (ACM) and optical storage researches. He joined Tsinghua university, China, as a national distinguish professor and director of optical memory national engineering research center in Mar 2013. His main research areas include NVM, ACM, optical data storage, integrated opto-electronics, and nanoscience. He has published more than 150 papers in prestigious journals including Science, Nature Photonics, Advanced Materials, Physical Review Letters, filed and granted more than 10 patents and conducted more than 60 keynote speech or invited talks at many important conferences during last 10 years. He is the recipient of the National Technology Award 2004 Singapore. He served as general co-chair of The 9th Asia-Pacific Conference on Near-field Optics2013, IEEE NVMTS 2011- 2014,East-West Summit on Nanophotonics and Metal Materials 2009 and ODS'2009.

Brain-Inspired Multi-Anything Algorithms for Medical Image Analysis and Computer Vision

Bart M. Romeny *(Eindhoven University of Technology, Netherlands and Northeastern University)*

Abstract

Electrophysiological, optical, opto-genetic, fMRI-, diffusion MRI and other brain imaging techniques have revealed an astoundingly well organized visual front-end. However, real understanding and generic modeling of the complex representations in the huge filter banks still offers many challenges. The multi-scale structure has inspired to (now in computer vision widely used) robust differential shift- and rotation invariant operators and keypoint detectors, and to hierarchical segmentation approaches and recognition techniques. The multi-orientation structure, recognized in the cortical pinwheels and their interconnections, has inspired to robust contextual tracking and adaptive enhancement operations.

We will discuss an innovative Lie-group based model for simultaneous analysis in the multi-scale, multi-orientation, multi-velocity, multi-disparity and multi-color domain. Applications will be presented for contextual, crossing preserving enhancement of elongated structures, such as 2D and 3D brain vasculature (e.g. quantitative retinal and extra-orbital vessel analysis exploited in a large-scale program for screening for early diabetes), and complex 3D brain dwMRI tractography, and perceptual grouping. The results are highly promising, and regularly outperform classical approaches, but need substantial processing, which today can be directed to, also brain-inspired, massively parallel GPU processing.

Short Biography

Bart M. ter Haar Romeny received the MSc degree in Applied Physics from Delft University of Technology in 1978, Ph.D. from Utrecht University in 1983 in biophysics. He became principal physicist of the Utrecht University Hospital Radiology Department. He was co-founder and associate professor at the Image Sciences Institute (ISI) of Utrecht University (1989-2001). From 2001, ter Haar Romeny holds the chair of Biomedical Image Analysis at the Department of Biomedical Engineering of Eindhoven University of Technology and Maastricht University in the Netherlands, and since 2011 is appointed distinguished professor at Northeastern University, Shenyang, China. His research interests include quantitative medical image analysis, its physical foundations and clinical applications. His interests are in particular the mathematical modeling of the visual brain and applying this knowledge in operational computer-aided diagnosis systems. He authored an interactive tutorial book on multi-scale computer vision techniques, edited a book on non-linear diffusion theory in computer vision.

He is author of over 200 refereed journal and conference papers, 12 books and book chapters, and holds 2 patents. He supervised many PhD students, of which 4 graduated cum laude. He is senior member of IEEE, and chairman of the Dutch Society for Pattern Recognition and Image Processing.

Information Processing in the Visual Pathway

Zhongzhi Shi *(Institute of Computing Technology)*

Abstract

Intelligence Science is an interdisciplinary subject that dedicates to joint research on basic theory and technology of intelligence by brain science, cognitive science, artificial intelligence and other disciplines. We have proposed a mind model CAM which is a general framework for brain-like machines. This talk will focus on the information processing in the visual pathway. Information processing in the visual pathway can be separated into objective processing and spatial processing. The Conditional Random fields based Feature Binding (CRFB) computational model is applied to visual objective processing. Feature integration theory is widely approved on the principles of the binding problem, which supplies the roadmap for our computational model. We construct the learning procedure to acquire necessary pre-knowledge for the recognition network on reasonable hypothesis–maximum entropy. With the recognition network, we bind the low-level image features with the high-level knowledge. For visual spatial processing, we explore three important kinds of relationship between objects that can be queried: topology, distance, and direction.

Short Biography

Zhongzhi Shi is a professor at the Institute of Computing Technology, Chinese Academy of Sciences, leading the Intelligence Science Laboratory. His research interests include intelligence science, machine learning, multi-agent systems and image processing. Professor Shi has published 14 monographs, 15 books and more than 450 research papers in journals and conferences. He has won a 2nd-Grade National Award at Science and Technology Progress of China in 2002, two 2nd-Grade Awards at Science and Technology Progress of the Chinese

Academy of Sciences in 1998 and 2001, respectively. He is a fellow of CCF and CAAI, senior member of IEEE, member of AAAI and ACM, Chair for the WG 12.2 of IFIP. He serves as Editor-in-Chief for International Journal of Intelligence Science and Series on Intelligence Science. He has served as Vice President for Chinese Association of Artificial Intelligence and General Secretary of China Computer Federation.

Control for the Autonomy of Mobile Robots

Jianda Han *(State Key Laboratory of Robotics, Shenyang Institute of Automation)*

Abstract

With the great development of robotics in recent years, many field robots have been expected to carry out tasks in outdoor surroundings, where the robots may suffer from complex terrains, dynamic obstacles/dangerous, bad weather conditions, and so on. Thus, one of the challenging topics is: how a field robot can survive the environment while handling the assigned tasks in an optimal/intelligent approach. The autonomy, which enables robots working on those complicated circumstances with reduced human intervention, has been becoming one of the main goals of mobile robotics. In this talk, I will introduce a feasible control scheme that has been implemented on and experimentally tested on ground mobile and flying robots. The scheme includes four aspects: 1) modeling and understanding the behavior environment; 2) behavior optimization; 3) autonomous learning; and 4) cooperation and coordination of multiple robots. By this scheme, we have realized the autonomous flight of the 100kg-level flying robots and the long-distance autonomous navigation of polar robots. Some of the experimental tests and the applications will be also demonstrated in this talk.

Short Biography

Han Jianda received his PhD degree in Electrical Engineering from the Harbin Institute of Technology in 1998. Currently he is a professor and deputy director of the State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences. His research interests include nonlinear estimation and robust control, control for the autonomy of robots, and robotic system integrations and applications such as medical and assistant robots, ground mobile robots, as well as flying robots. His team developed the first polar robot of China, which was tested in Antarctica in 2008 and 2011; and also the 100kg flying robot, which has realized its applications such as rescue, precision agriculture, power cable construction, etc. Dr. Han currently also serves as a member of the 5-person Expert Panel of the Intelligent Robot Division, the National High Technology Research and Development (863) Program of China.

Brainnetome Studies of Alzheimer's Disease with Neuroimaging

Yong Liu *(Brainnetome Center, Institute of Automation)*

Abstract

The human brain has been described as a large, sparse, complex network. Some of previous neuroimaging studies have provided consistent evidence of dysfunctional connectivity among the brain regions in the AD; however, little is known about whether or not this altered functional connectivity causes disruption of the efficient of information transfer of brain functional networks in the AD. We will introduce the altered functional connectivity pattern of the AD from region of interest analysis, to local network analysis and to whole brain network analysis. And there has been a considerable amount of work recently on the characterization of brain structure and function in the context of networks. This includes identifying correlated changes, defining various network properties (such as long-distance connectivity, rich club behavior, or more general information theoretic measures) and evaluating changes in these properties in the AD groups. Detection and estimation of these alterations could be helpful for understanding the functional alteration of the AD.

Short Biography

Dr. Yong Liu is an associate professor in Brainnetome Center, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. He received his PhD degree from CASIA in 2008 and obtained his MSc degree from Beijing University of Technology in 2005. Since June 2008, he joined CASIA as an assistant/associate research professor. He is a visiting scholar from April 2011 to March 2012 in Brain Mapping Unit in

University of Cambridge, where he worked with Professor Ed Bullmore. His main interests are the analysis of brain networks using multi-modal approaches as well as the application in cognitive disorders, especially in Alzheimer disease and mild cognitive impairment. To date, he has authored about 52 peer-reviewed journal articles and has an h-index of 19. He is a member of Youth Innovation Promotion Association of CAS (2013). He has published in several international peer-reviewed journals including Brain, Cerebral Cortex, NeuroImage and Human Brain Mapping. He is a member of academic editor of PloS One and the Organization for Human Brain Mapping.

Neuromorphic Motivational Systems for Sequential Tasks in Dynamic Environment

Dongshu Wang, Yihai Duan, and Juyang Weng, *Fellow, IEEE*

***Abstract*— Although reinforcement learning has been extensively studied, few agent models that incorporate values use neuromorphic architecture. By neuromorphic, we mean that the motivational system is a biologically plausible neural network. This work proposes a neuromorphic motivational system, which includes two subsystems — the serotonin system and the dopamine system for a sequential task in dynamic environment. To testify the effects of serotonin and dopamine, we experiment with this motivational system for robot navigation in dynamic environment under different settings. The three experiments all illustrated that the reinforcement learning via the serotonin and the dopamine systems is beneficial for developing desirable behaviors in this set of sequential tasks — staying statistically close to its friend and away from its enemy. It was indeed the punishment and reward from serotonin and dopamine that cause the agent's behaviors. The agent can decide how to react in given situations by itself rather than being explicitly taught where to go.**

I. Introduction

A Major function of human brain is to develop circuits for processing sensory signals and generating motor actions. The signals in the brain are largely transmitted through neurotransmitters, endogenous chemicals that are sent from a neuron to a target cell across a synapse [5]. This modulatory system of the brain is often called motivational system or value system in neuroscience and psychology. It is about how neurons in the brain use a few particular types of neural transmitter to indicate certain properties of signals. A modulatory system goes beyond information processing and sensorimotor behaviors. It provides mechanisms to developmental system so that it develops dislikes and likes [21].

A. Neural Modulatory Transmitters

In the subject of neural modulation, a small group of neurons synthesize and release a particular type of neural modulatory transmitters which diffuse through large areas of the nervous system, producing an effect on many neurons. Such processing is characterized by direct synaptic transmission — the pre-synaptic neuron directly influences the post-synaptic neuron through different types of neurotransmitters. For example, glutamate is a kind of neurotransmitter. Nerve impulses trigger release of glutamate from the pre-synaptic cell [5]. A sufficient number of bindings by neurotransmitters results in the firing of the post-synaptic neuron [3]. Each cell typically has many receptors, of many different kinds.

While glutamate and GABA are neurotransmitters whose values are largely neutral in terms of reference at the time of birth, some other neurotransmitters appear to have been used by the brain to represent certain signals with intrinsic values [21], [3], [4]. For example, serotonin (5-HT) seems to be involved with pain, punishment, stress and threats; while dopamine (DA) appears to be related to pleasure, wanting, anticipation and reward [4]. Therefore, 5-HT and DA, along with many other neurotransmitters that have inherent values, seem to be useful for modeling the intrinsic value system of the central nervous system in artificial neural networks.

B. Motivational System

Psychological studies have provided rich behavioral evidence about the existence of the motivational system [21], [1], [10], [11], [13], [2], [15], [16], [22], [23]. It is known that the motivational system is important to the autonomous learning of the brain. However, although there is a very rich literature about models of neuromodulatory systems, such models are limited in terms of computational functions due to a few well known limitations in prior neural network models. Weng argued that we have overcome such limitations [20].

The motivational systems are often referred to as diffuse systems in the sense that each modulatory neuron in such a system uses a particular neurotransmitter (e.g., serotonin or dopamine) and makes diffuse connections, via such neurotransmitters, with many other neurons [21]. Both serotonin and dopamine come in many different forms, since the brain may conveniently use a particular neurotransmitter for different purposes in different parts of the body. There are other forms of serotonin and dopamine that have different effects that are not widely understood [11], [14]. We will focus on the forms of dopamine and serotonin that affect the brain as rewards and punishment, respectively.

Dopamine is released in the brain's Ventral Tegmental Area (VTA) and particularly the nucleas accumbens act as a general facilitative agonist for pleasure. Dopamine is associated with reward prediction [11]. If an agent gets a reward, then dopamine is released in the brain. If an agent is expecting a reward, dopamine is also released. For example, if a dog is trained that it always gets food when it hears a bell, then dopamine will be released into the brain at the time that it normally would receive the food. If the dog does not get food at the expected time, then the dopamine levels will drop below normal for a short period of time.

Dongshu Wang and Yihai Duan are with School of Electrical Engineering, Zhengzhou University, Zhengzhou, Henan, 450001, China. (email: wangdongshu@zzu.edu.cn, duanyihai88@163.com). Juyang Weng is with Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48823, USA. He is also with the MSU Cognitive Science Program and the MSU Neuroscience Program. (e-mail: weng@cse.msu.edu).

This work was supported by the Natural Sciences Foundation of China (Grant No. 61174085).

TABLE I

DIFFERENCE BETWEEN OUR PAPER AND MAIN REFERENCES

Type	Motor (Z area) neurons	Internal (Y area) neurons
Episodic tasks	Palaski [17]	Zheng [25], Weng [22]
Sequential tasks	Daly [5]	This work

Serotonin often appears to be dopamines counterpart [6]. Dopamine excites the neurons while serotonin inhibits them. One specific type of serotonin with this effect is 5-HT. Serotonin leads to behavior inhibition and aversion to punishment. There are two parts of the brain that release serotonin, the dorsal raphe and the median raphe. The dorsal raphe connects serotonin to all of the areas that have dopamine connections [12]. Serotonin from the dorsal raphe interacts with dopamine to cause the agent to avoid behaviors that the dopamine encourages.

C. Novelty and Importance

Our improvement on the motivated development network lies in the following three aspects:

Firstly, though Zheng [25] and Weng [22] also considered the effect of serotonin and dopamine on the neurons in Y_u, their experiment object is face recognition which is a pattern recognition problem. The recognition result is determined only by the current one decision. But robot navigation is a typical sequential problem because the current decision can affect the following one, and the final result is determined by all the former decisions instead of the last one.

Secondly, previous work [5] and [22] studied the effect of serotonin and dopamine on the learning rate of Z qualitatively. Our research studied their effect on the learning rate of Z and Y_u quantitatively. In other words, we calculate the linear punishment (or reward) value according to the distance between the agent and the friend (or enemy) and the threshold, instead of the qualitative way.

Finally, we studied the effects of different environments (such as different teachers) on the final navigation performance. To the strict “teacher”, even though you do something very well, he still consider you have not done enough. On the contrary, to the tolerant “teacher”, though you do not do well, he may think you have done quite well. To the same situation, different environments will produce different effects.

The difference between our paper and the main references is illustrated in table 1.

In sequential tasks, the behavior of the agent is determined by its past experience, and it can imitate and adapt the social environment. In other words, if the agent has many friends and little enemies, thus it always receives reward, and it will become braver. Conversely, if the agent has many enemies and little friends, thus it always receives punishment, and it will become very more careful. It is very important and essential to study the effect of serotonin and dopamine on the sequential tasks.

The remainder of the paper is organized as follows. Section II introduces the theory for value systems of sequential tasks. Section III describes the theory of developmental network behind our model. Simulation experiments are presented and analyzed in section IV, while the conclusions are given in the last section.

II. MOTIVATIONAL SYSTEM OF SEQUENTIAL TASKS

In terms of context dependence, there are two types of tasks, episodic and sequential.

A. Episodic and Sequential Tasks

In an episodic task environment, agent’s experience is divided into atomic episodes. Each episode consists of the agent perceiving and then performing a single action. Crucially, the next episode does not depend on the actions taken in previous episodes. In episodic environments, the choice of action in each episode depends only on the episode itself. Many classification tasks are episodic. For example, an agent that has to spot defective parts on an assembly line bases each decision on the current part, regardless of previous decisions; moreover, the current decision does not affect whether the next part is defective. In sequential environments, on the other hand, the current decision could affect all future decisions. Chess, taxi driving are sequential. In these cases, short-term actions can have long-term consequences [18].

Robot navigation is a typical sequential task because the behavior of the agent also depends on its past actions. In Fig. 2, there are four trajectories: T_1 (denotes strong reward, weak punishment), T_2 (denotes strong reward, strong punishment), T_3 (denotes weak reward, strong punishment) and T_4 (denotes weak reward, weak punishment). Though the four trajectories are different, they use the same learning rule. If there is no previous punishments from the enemy, the agent will move along the direction of the blue arrow (it depends on how much reward the agent receives in the past). Similarly, if there was no previous rewards from the friend, the agent will still move along the direction of the red arrow (it depends on how much punishment the agent receives in the past). If the agent receives both the reward and the serotonin, it will move along the direction of black arrow(it depends on how much punishment and reward the agent receives in the past), so its moving is determined by the past experience and actions. The memories of past trajectories are in the network weights and the network predicted amounts of serotonin and dopamine, but not exactly the trajectories themselves.

B. Reinforcement Learning in Sequential Tasks

Reinforcement learning is the problem faced by an agent that must learn behavior through trial-and-error interactions with a dynamic environment. Without some feedback about what is good and what is bad, the agent will have no grounds for deciding which move to make [18]. The agent needs to know that something good has happened and that something bad has happened. This kind of feedback is called a reward or punishment. In the natural world, signals from a pain sensor is associated with bad and those from a sweet sensor is associate with good. There are many additional reward and

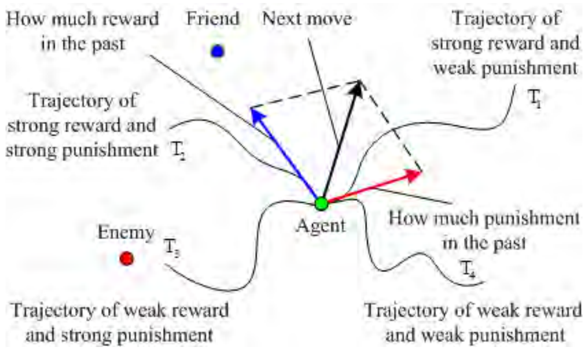


Fig. 1. How much rewards and how much punishments that the agent predicts depend on how the agent moved in the past. The closer it moved toward the enemy, the more punishment it would receive and the enemy would affect the action of the agent more. The further it moved away from the friend, the more reward it would receive and the friend would affect the action of the agent more.

punishment sensors. However, the brain mechanisms of such an association is largely unknown.

Reinforcement learning has been carefully studied by animal psychologists for many years [7], [9], [8], [19]. However, those models are *symbolic* in the sense that each mode has specific, handcrafted meanings for a specific task.

Using brain-like *emergent* representations, each neuron does not have a specific task meaning. Patterns of neuronal firing emerge from interactions of the physical external world. In particular, a motor neuron (or a firing patten of neurons) does not represent a bad action until it consistently fires with the presence of serotonin. Weng et al. [22] modeled that serotonin and dopamine are associated with pain sensors and sweet sensors, or punishment and reward, respectively, in general. Serotonin inhibits the firing of the current motor neurons and dopamine excites the firing of the current motor neurons. Hopefully, since the diffusions of such neural transmitters are relatively slow, statistically, the level of such neurotransmitters correlates with the responsible motor actions reasonably well.

The above is about motor neurons. In a sequential task, however, each reinforcer (punishment or reward) is a consequence of a sequence of past state trajectories indicated by sensory-state pairs in terms of $(\mathbf{x}(t), \mathbf{z}(t))$:

$$(\mathbf{x}(t-m), \mathbf{z}(t-m)), ..., (\mathbf{x}(t-1), \mathbf{z}(t-1)), (\mathbf{x}(t), \mathbf{z}(t)) \quad (1)$$

where $\mathbf{x}(t)$ and $\mathbf{z}(t)$ are the sensory and state vectors, respectively. In our Developmental Network (DN) framework, state and action are the same since reporting state is an action. E.g., say bad words is an action punishable.

The Y area, represented by a large number of Y neurons, represents a multi-exchange *bridge* that form feature clusters in the two islands X area and Z area (as discussed in the next section):

$$(\mathbf{x}(t-1), \mathbf{z}(t-1)) \rightarrow \mathbf{y} \rightarrow (\mathbf{x}(t), \mathbf{z}(t))$$

where $\rightarrow$ means "predicts". If the following $(\mathbf{x}(t), \mathbf{z}(t))$ is independent with y, the task is episodic. In a sequential task, the following $(\mathbf{x}(t), \mathbf{z}(t))$ depends on y.

In our prior work [25], we modeled that serotonin and dopamine both increase the learning rate of the firing Y neurons to memorize the important event in episodic tasks. In this paper, we study how serotonin and dopamine increase the learning rate of firing Y neurons for sequential tasks. As we can see from the above analysis, changing the rate using serotonin and dopamine transmitter should improve the performance of learning sequential tasks.

III. Theory of Developmental Network

A. Developmental Network

Developmental network is the basis of a series of Where-What networks, whose 7th version, namely, the latest version, appeared in [24]. The simplest version of a Development Network (DN) has three areas, the sensory area X, the internal area Y, and the motor area Z, with an example in Fig. 2. The internal area Y as a "bridge" to connect its two "banks"- the sensory area X and the motor area Z. The DN

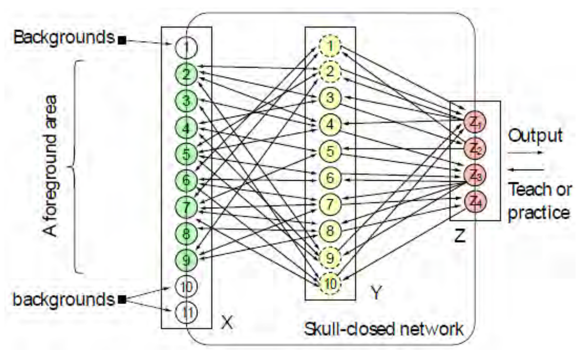


Fig. 2. The architecture of DA. It contains top-down connections from Z to Y for context represented by the motor area. It contains top-down connections from Y to X for sensory prediction. Pink areas are human taught. Yellow areas are autonomously generated (emergent and developed). Although the foreground and background areas are static here, they change dynamically.

algorithm is depicted as follows:

1) At time $t = 0$, for each area A in $\{X, Y, Z\}$, initialize its adaptive part $N = (V, G)$ and the response vector $\mathbf{r}$, where V contains all the synaptic weight vectors and G stores all the neuronal ages.

2) At time $t = 1, 2, ...$, for each area A in $\{X, Y, Z\}$, do the following two steps repeatedly forever:

a) Every area A computes using area function f.

$$(\mathbf{r}',\ N') = f(\mathbf{b}, \mathbf{t}, N) \quad (2)$$

where f is the unified area function described in the following equation (3), $\mathbf{b}$ and $\mathbf{t}$ are areas bottom-up and top-down inputs from current network response $\mathbf{r}$, respectively; and $\mathbf{r}'$ is its new response vector.

b) For each area A in $\{X, Y, Z\}$, A replaces: $N \leftarrow N'$ and $\mathbf{r} \leftarrow \mathbf{r}'$.

If X is a sensory area, $x \in X$ is always supervised and then it does not need any synaptic vector. The $z \in Z$ is supervised only when the teacher chooses to. Otherwise, z

gives (predicts) motor output. Next, we describe the area function f.

Each neuron in area A has a weight vector $\mathbf{v} = (\mathbf{v}_b, \mathbf{v}_t)$, corresponding to the area input $(\mathbf{b}, \mathbf{t})$, if both bottom-up part and top-down part are applicable to the area. Otherwise, the missing part of the two should be dropped from the notation. Its pre-action energy is the sum of two normalized inner product:

$$\mathbf{r}(\mathbf{v}_b, \mathbf{b}, \mathbf{v}_t, \mathbf{t}) = \frac{\mathbf{v}_b}{||\mathbf{v}_b||} \cdot \frac{\mathbf{b}}{||\mathbf{b}||} + \frac{\mathbf{v}_t}{||\mathbf{v}_t||} \cdot \frac{\mathbf{t}}{||\mathbf{t}||} = \dot{\mathbf{v}} \cdot \dot{\mathbf{p}} \quad (3)$$

where $\dot{\mathbf{v}}$ is the unit vector of the normalized synaptic vector $\mathbf{v} = (\dot{\mathbf{v}}_b, \dot{\mathbf{v}}_t)$, and $\dot{\mathbf{p}}$ is the unit vector of the normalized synaptic vector $\mathbf{p} = (\dot{\mathbf{b}}, \dot{\mathbf{t}})$.

To simulate lateral inhibition (winner takes all) within each area A, only top-k winners fire and update. Considering $k = 1$, the winner neuron j is identified by:

$$j = \arg\max_{1 \le i \le c} \mathbf{r}(\mathbf{v}_{bi}, \mathbf{b}, \mathbf{v}_{ti}, \mathbf{t}) \quad (4)$$

where c is the neuron number in the area A.

The area dynamically scale top-k winners so that the top-k responses with values in [0,1]. For $k = 1$, only the single winner fires with responses value $y_j = 1$ and all other neurons in A do not fire.

All the connections in a DN are learned incrementally based on Hebbian learning—co-firing of the pre-synaptic activity $\dot{\mathbf{p}}$ and the post-synaptic activity y of the firing neuron. Consider area Y, as other area learn in a similar way. If the pre-synaptic end and the post-synaptic end fire together, the synaptic vector of the neuron has a synapse gain $y\dot{\mathbf{p}}$. Other non-firing neurons do not modify their memory. When a neuron j fires, its weight is updated by a Hebbian-like mechanism:

$$\mathbf{v}_j \leftarrow \omega_1(n_j)\mathbf{v}_j + \omega_2(n_j) y_j \dot{\mathbf{p}} \quad (5)$$

where $\omega_2(n_j)$ is the learning rate depending on the firing age n_j of the neuron j and $\omega_1(n_j)$ is the retention rate with $\omega_1(n_j) + \omega_2(n_j) \equiv 1$. The simplest version of $\omega_2(n_j)$ is $1/n_j$, which gives the recursive computation of the sample mean of input $\dot{\mathbf{p}}$:

$$\mathbf{v}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \dot{\mathbf{p}}(t_i) \quad (6)$$

where t_i is the firing time of the neuron. The age of the winner neuron j is incremented $n_j \leftarrow n_j + 1$. A component in the gain vector $y_j \dot{\mathbf{p}}$ is zero if the corresponding component in $\dot{\mathbf{p}}$ is zero. Each component in $\mathbf{v}_j$ so incrementally computed is the estimated probability for the pre-synaptic neuron to fire under the condition that the post-synaptic neuron fires. A more complicated version of $\omega_2(n_j)$ is presented in the next section when we discuss the architecture of our motivated system.

B. Motivational Developmental Network (MDN)

According to literature [6], serotonin and dopamine receptors are also found in brain neurons except the motor neurons. It means that the release of serotonin and dopamine, which occurs in RN and VTA areas, should also have effect on neurons in Y_u neurons.

Fig. 3 presents the architecture of the MDN. It links all pain receptors with RN located in the brain stem–represented as an area, which has the same number of neurons as the number of pain sensors. Every neuron in RN releases serotonin. Similarly, it also links all sweet receptor with VTA–represented as an area, which has the same number of neurons as the number of sweet sensors. Every neuron in VTA releases dopamine.

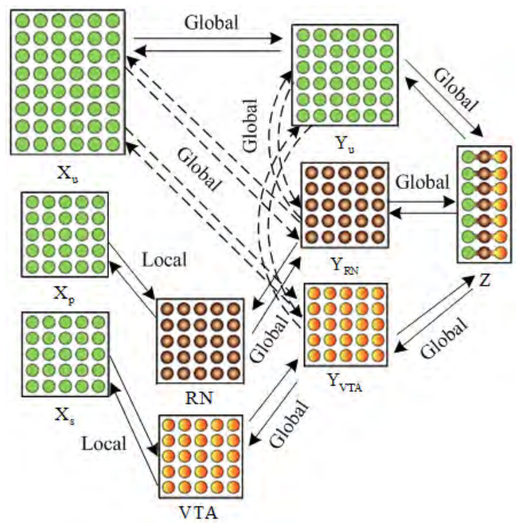


Fig. 3. A motivated DN with serotonin and dopamine modulatory subsystems. It has 9 areas. RN has serotonergic neurons. Neurons in Y_{RN} and Z have serotoninergic synapses. VTA has dopaminergic neurons. Neurons in Y_{VTA} and Z have dopaminergic synapses. The areas Y_u, Y_{RN} and Y_{VTA} should reside in the same cortical areas, each represented by a different type of neurons, with different neuronal densities.

Serotonin and dopamine are synthesized by several brain areas. For simplicity, we use only RN to denote the area that synthesizes serotonin and only the VTA to denote the area that synthesizes dopamine, although other areas in the brain also involved in the synthesis of these neurotransmitters.

Therefore the sensory area $X = (X_u, X_p, X_s)$ consisting of an unbiased array X_u, a pain array X_p and a sweet array X_s. $Y = (Y_u, Y_{RN}, Y_{VTA})$ connects with $X = (X_u, X_p, X_s)$, RN and VTA as bottom-up inputs and Z as top-down input.

Within such a motivational developmental network, the motor area is denoted as a sequence of neurons $Z = (z_1, z_2, \cdots, z_m)$, where m is the number of motor neurons whose axons innervate muscles or glands. Each z_i has three neurons $z_i = (z_{iu}, z_{ip}, z_{is})$, where z_{iu}, z_{ip} and z_{is} $(i = 1, 2, \cdots, m)$ are unbiased, pain and sweet, respectively. And these indicate the effects of glutamatergic synapses, serotoninergic synapses and dopaminergic synapses, respectively.

Whether the action i is released depends on not only the response of z_{iu} but also on those of z_{ip} and z_{is}. z_{ip} and z_{is} report how much negative value and positive value are associated with the i-th action, according to past experience. They form a triplet for the pre-action energy value of each motor neuron, glutamate, serotonin and dopamine.

Modeling the cell's internal interactions of the three different types of neurotransmitter, the composite pre-action value of a motor neuron is determined by

$$z_i = \; z_{iu}\gamma(1 - \alpha z_{ip} + \beta z_{is}) \tag{7}$$

with positive constants α, β and γ. In other words, z_{ip} inhibits the action but z_{is} excites it. α is a relatively larger constant than β since punishment typically produces a change in behavior much more significantly and rapidly than other forms of reinforcers.

Then the j-th motor neuron fires and action is released where

$$j = \arg \max_{1 \leq i \leq m} \{z_i\} \tag{8}$$

That is, the primitive action released at this time frame is the one that has the highest value after inhibitory modulation through serotonin and excitatory modulation through dopamine, respectively, by its bottom-up synaptic weights. Other Z neurons do not fire.

IV. Experiments

This section, we will describe the experiment procedure designed to test the above theory and algorithm.

A. Experiment Design

In our experiment, we uses three robots to test our algorithm. One of the robots is the agent which can think and act, the other two are its friend and enemy, respectively. If the agent approaches the friend robot, it is rewarded with dopamine. If it approaches the enemy, it is punished with serotonin. In this way, the agent will learn to close its friend and avoid its enemy. But it must learn this behavior through its own trial and error experience.

The agent's brain is the MDN with three areas, X, Y and Z, as depicted in Fig. 3. Where X is the sensor area which has three sub-areas, X_u is the unbiased area, X_p is the pain area and X_s is the sweet area. The X_u vector is created directly from the sensors' input. The X_p vector identifies in which ways the robot is punished and represents the release of the serotonin in RN. The X_s vector identifies in which ways the robot is rewarded and represents the release of the dopamine in VTA.

All of the Y_{RN}, Y_{VTA} and Z sub-areas compute their response vectors in the same way. At the end of each time step, the neurons in Y_{RN}, Y_{VTA} and Z areas that fired update themselves. Their weights are updated according to the former equations (5) and (6). Their ages are updated as follows: a_i = a_i+1.

According to the work of [25], serotonin and dopamine levels are released at different levels rather than binary values. The release gives specific neurons in the Y_{RN} and Y_{VTA} areas a non-zero response, which will have effect on the learning rate of neurons in Y_u. Moreover, the roles of a motor neuron and an inter neuron are different. The former roughly corresponds to the action that is responsible for the corresponding punishment and reward; the latter corresponds to the memory of the corresponding event. Therefore, serotonin and dopamine should increase the efficiency of learning in Y_u, instead of directly discouraging and encouraging the firing. One way to reach such an effect is to increase the learning rate depicted as follows:

$$\omega_2(n_j) = \min((1 + \alpha_{RN} + \alpha_{VTA})\frac{1}{n_j}, 1) \tag{9}$$

where α_{RN} and α_{VTA} are constants related with RN and VTA respectively. This expression shows that reward and punishment change the learning rate in Y_u neurons. If neurons in X_p and X_s do not fire, responses in RN and VTA are zero, thus the learning rate (9) will turn into its original form.

B. Experiment Parameters Setting

The Y and Z_u areas are initialized to contain small random data in their state vectors. The Z_p and Z_s areas are initialized to zero since the agent has no idea which actions will cause pleasure or pain. The ages of all neurons are initialized to 1. The number of neurons in Y layer and the fire number k can be selected based on the resources available. The size of Z area is equal to the number of actions that the agent can perform. At any time, the agent can perform one of nine possible actions, it can move in each of the cardinal or intercardinal directions or it can maintain its current position. So the neurons in Z areas has 9 rows and 3 columns. 3 columns denote the Z_u, Z_p and Z_s, respectively.

The size of each vector in the X area is determined by the transformation function through which the robot can sense the locations of its friend and enemy. If we define the following entities, a (agent), f (friend), e (enemy) , we can draw a sketch of the location relation among the three robots as shown in Fig. 4, and get the following expressions:

$$\theta_f = \arctan(a_x - f_x, a_y - f_y), \tag{10}$$

$$d_f = \sqrt{(a_x - f_x)^2 + (a_y - f_y)^2}, \tag{11}$$

$$\theta_e = \arctan(a_x - e_x, a_y - e_y), \tag{12}$$

$$d_e = \sqrt{(a_x - e_x)^2 + (a_y - e_y)^2}, \tag{13}$$

$$x_u = \{\cos\theta_f, \sin\theta_f, \cos\theta_e, \sin\theta_e, \frac{d_f}{d_f + d_e}, \frac{d_e}{d_f + d_e}\}, \tag{14}$$

where θ_f and θ_e are the angle between the heading of the agent and the direction of the friend robot and enemy robot, respectively; d_f and d_e are the distance between the agent and the friend and the enemy, respectively.

The pain sensor and the sweet sensor has just one value to denote the fear and desire. Consulting the reference [5], the fear threshold is set 125, namely, if $d_e > 125$, there is no punishment. If $30 < d_e \leq 125$, punishment value is set

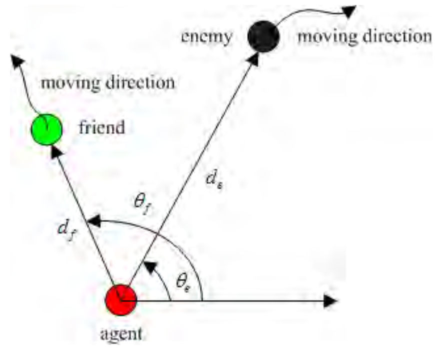


Fig. 4. The setting of the wandering plan which includes the agent, the friend robot and enemy robot. The size of the square space used is 500×500.

4. Otherwise, the punishment value is calculated through the fear threshold divided by the actual distance d_e.

Similarly, the desire threshold is set 50, namely, if $d_f < 50$, there is no reward. If $50 < d_f \leq 150$, the reward value is calculated through the actual distance d_f divided by the desire threshold. Otherwise, the reward value is set 3.

C. Experiment Setup

We designed a simulation environment to illustrate how such a motivated agent would response in the presence of its friend and enemy. The motivated robot (agent) is controlled by the motivated "brain" which is actually a MDN. The "brain" releases serotonin and dopamine for enemy and friend based on specific circumstances. Through the simulation, the agent will learn by reinforcement, deciding which one to avoid and which one to go after, based on the release of serotonin or dopamine.

At each time step, the horizontal and vertical coordinates are collected for each entity. With these data, we can calculate the distance between the agent and its friend (or enemy). Through observing the distances of the agent to its friend (or enemy), we can measure the learning procedure of the agent.

The agent starts with a behavior pattern determined by its initial neural network configuration. The unbiased regions are initialized with small random data while the biased regions are initialized to zero. This gives the initial appearance of a random behavior pattern. Eventually, it performs an action that causes it to be rewarded or punished, causing it to be either favor or avoid that action when placed in similar situations in the future.

D. Results and Analysis

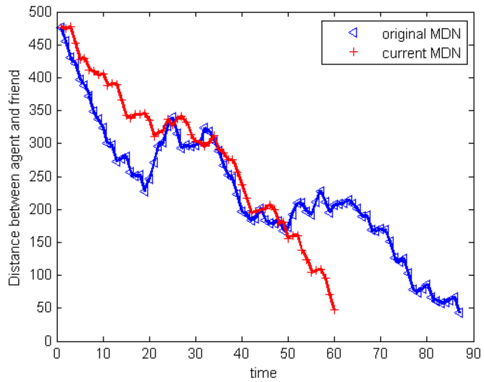
In the first experiment, in order to test the effect of serotonin and dopamine on the algorithm performance of current MDN, we compare the distance between the agent and its friend (or enemy) under the original MDN (based on Daly's work, reference 2, it only considered the effects of serotonin and dopamine on the motor area qualitatively) and the current MDN (we designed, it not only considered the effects of serotonin and dopamine on the motor area, but also considered their effects on the Y_u area quantitatively).

The experiment is carried out in dynamic environment, namely, the friend and the enemy move randomly in the environment. This setting is without the static objects, such as house and trees. Initial location of the agent is [470, 400], the friend's initial location is [100, 100], and the enemy's initial location [250, 250]. Their distance situations change with the time are shown in Fig. 5. From Fig. 5 , we can see that at the initial time t=0, d_e is about 270, which is bigger than the punishment threshold (125), so the agent does not receive the punishment. The d_f is about 480, which is much bigger than the desire threshold (50), so the agent receives the reward. In other words, at initial time t=0, the agent only receives the reward, thus it moves towards the friend along the fastest direction. About between the time 18 and 20, the agent begins to receive both the reward and the punishment, and its moving direction is determined by three elements (i.e., punishment, reward and the guidance of Y_u to Z_u, as shown in Eq. (7)). Though it receives the punishment, the punishment is not very big, and it also receives the reward, but the reward is very big, so the hybrid effect of the three elements is that it still moves towards the friend. At time 22, the punishment value becomes very big and the reward is not very big, and punishment becomes the decisive factor, so from this moment on, the agent moves far away from its enemy instead of moving towards the friend. After the time 30, the agent only receives the reward so it moves towards the friend.

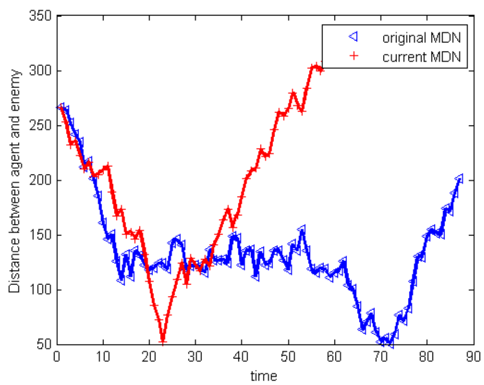
From Fig. 5, we can also see that in the experiment, under the same initial condition, the current MDN can track the friend more quickly and get a smaller d_f, and avoid the enemy more quickly and get bigger d_e than the original MDN. The reason is the effects of the serotonin and dopamine on the learning of Y_u. Learning rate of Y_u in current MDN introduces the effect of serotonin and dopamine as denoted in Eq. (10), but the original MDN do not. So the learning rate of current MDN is bigger than that of the original MDN, and its learning speed is faster than that of the original MDN. Graphically, the current MDN can achieve a smaller d_f and bigger d_e than the original MDN.

Moreover, the neurons in Y_u area can memorize the corresponding events (e.g., punishment or reward), and these events will affect the agent's psychology if it often receives punishments or rewards. Consequently, it will generate like or dislike to certain environment, which will affect the corresponding behaviors (depicted in Eq. (7)) of the neurons in Z area, thus will determine the agent's following behaviors. For example, if a neuron receives reward, it will strengthen the corresponding behavior to repeat the similar behaviors in the following actions. Graphically, it represents the nearer distance d_f. Conversely, if a neuron receives punishment, it will also strengthen the corresponding behavior to avoid the similar behaviors in the following actions. Graphically, it represents the further distance d_e.

In addition, in Eq. (7), the constant α is relatively larger than β, and the effect of serotonin should be much more significant and rapid than that of the dopamine. Fig.5 testify



(a)



(b)

Fig. 5. Distance comparison between agent and friend (a) and enemy (b) under original MDN and current MDN.

this inference. From time 40, it is very clear that the agent avoids its enemy much faster than it approaches the friend.

In the second experiment, we test the effects of different environmental parameters (different punishment and reward thresholds) on the sequential task quantitatively. The results are illustrated in Fig. 6, six groups of numbers along the horizontal axis denote the different distances between the agent and the friend (or enemy), and these distances are adopted as the reward (or punishment) thresholds in the corresponding experiments, and the vertical axis denotes the average distances we measured.

Based on Fig. 6, we can see that in these cases, the distances between the agent and the friend are much smaller than that to the enemy. All average distances to the enemy are greater than corresponding punishment thresholds, which indicate the effect of serotonin on the agent's behavior. Similarly, all average distances to the friend are smaller than corresponding reward thresholds, which indicate the effect of dopamine on the agent's behavior. These different thresholds reflect the environmental parameters, indicating that the serotonin and dopamine systems appear to work as expected.

In the third experiment, we study the effect of same environment and different punishment and reward values on the robot navigation. Initially, the three robots are set to

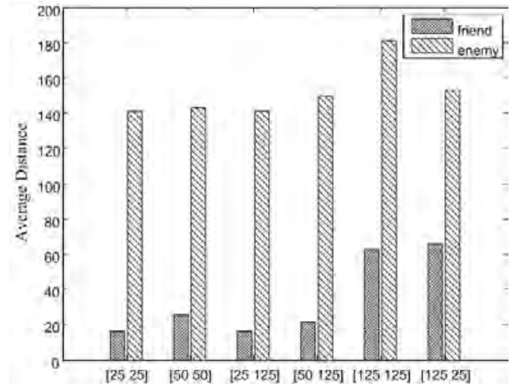


Fig. 6. Average distance between agent and its friend and enemy in different situations.

constant locations in different cases. But the punishment and/or reward values are variable. Then we analyze the effects of these different punishment and reward values on the agent's behavior improvement. The effects are shown in Fig. 7. From Fig. 7, we can see that:

(1) Case of small reward or big punishment. In this case, the maximum reward is set 1, and the minimum punishment is set 4. From Fig. 7 (b), we can see that because the punishment value is relatively big, and the reward is relatively small, so the agent always keeps a big distance with the enemy. Thus when the agent receives the punishment, it moves out of the punishment scope immediately. From Fig. 7, we can see that between the time step 50 and 60, the agent does not receive punishment ($d_e > 125$) and only receive the reward. But the agent moves away from the friend instead of moving towards it (graphically, the distance d_f increases instead of decreasing). This phenomenon shows that under the condition of relatively small reward, when the agent is far away from the enemy, the effect of weights is more obvious than that of the reward.

(2) Case of small punishment or big reward. In this case, the max of punishment is set 1 and reward is set according to section B. From Fig. 7 (b), we can see that the nearest distance between the agent and the enemy is 20. At the former one time step, the agent is far from its friend. The agent receives the punishment and reward at the same time, but it moves toward the enemy instead of moving away from it and moves towards the friend. So we can conclude that when the agent, friend and enemy are located on the same line, and if the reward value is bigger than the punishment value, the punishment can be omitted. This phenomenon also means that in this case, the reward is set too big and the punishment is relatively too small. This implicates that minimum punishment should be bigger than the maximum reward. This is one of the requirements in designing punishment and reward value, and it is consistent with the rules in choosing the α and β values in Eq. (12).

(3) Case of teacher-reinforcement. In this case, the punishment is set 4 and reward is set 3. From Fig. 7, we can see that in this case, the agent can not only keep certain distance

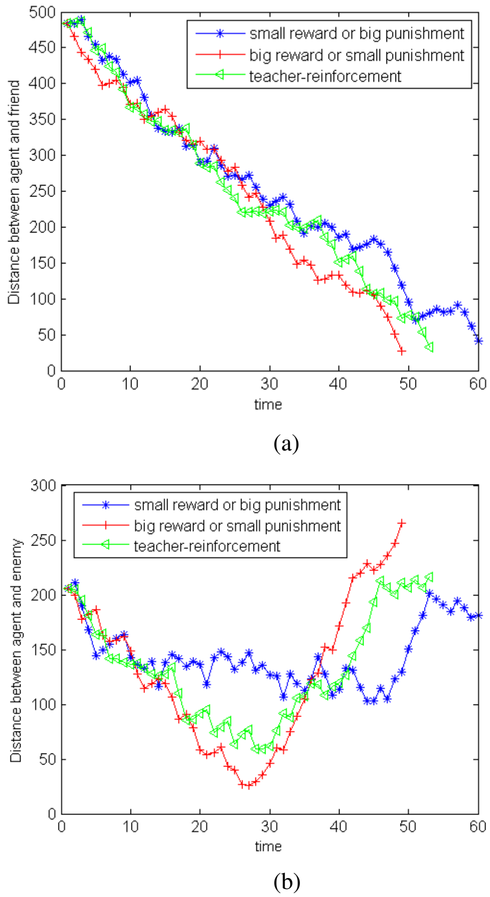


Fig. 7. Distance comparisons among different punishments and rewards.

from the enemy, but also moves towards the friend in short time. These mean that the punishment and reward values we set are suitable.

V. Conclusions

We analyzed the effect of serotonin and dopamine systems on Y neurons for sequential tasks and conducted three simulation experiments to test the effects. Experiment results illustrated that it was indeed the punishment and reward from serotonin and dopamine that cause the agent's behaviors. The agent can decide how to react in given situations by itself rather than being explicitly taught where to go. In future work, the agent will be placed in a more complicated situation with multiple friends and enemies to further study the effects of serotonin and dopamine.

References

[1] G. Baldassarre, F. Mannella, V. G. Fiore, P. Redgrave, K. Gurney, and M. Mirolli. Intrinsically motivated action-outcome learning and goal-based action recall: A system-level bio-constrained computational model. *Neural Networks*, 41:168–187, 2013.

[2] S. Cavallaro. Genomic analysis of serotonin receptors in learning and memory. *Behavioral Brain Research*, 195:2–6, 2008.

[3] B. R. Cox and J. L. Krichmar. Neuromodulation as a robot controller, a brain-inspired strategy for controlling autonomous robots. *IEEE Robotics and Automation Magazine*, pages 72–80, 2009.

[4] M. J. Crockett, A. Apergis-Schoute, B. Herrmann, M. Lieberman, U. Muller, T. W. Robbins, and L. Clark. Serotonin modulates striatal responses to fairness and retaliation in humans. *The Journal of Neuroscience*, 33:3505–3513, 2013.

[5] J. Daly, J. Brown, and J. Weng. Neuromorphic motivated systems. In *Proceedings of International Joint Conference on Neural Networks, (IJCNN 2011)*, pages 2917–2924, San Jose, California, USA, July 31-August 5 2011.

[6] N. D. Daw, S. Kakada, and P. Dayan. Opponent interactions between serotonin and dopamine. *Neural Networks*, 15:603–616, 2002.

[7] P. Dayan and B.W. Balleine. Reward, motivation and reinforcement learning. *Neuron*, 36:285–298, 2002.

[8] X. Huang and J. Weng. Novelty and reinforcement learning in the value system of developmental robots. In *Proceedings of the Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, pages 1–9, Edinburgh, Scotland, August 10-11 2002.

[9] X. Huang and J. Weng. Inherent value systems for autonomous mental development. *International Journal of Humanoid Robotics*, 4:407–433, 2007.

[10] Z. Ji, M. D. Luciw, and J. Weng. A biologically-motivated developmental system towards perceptual awareness in vehicle-based robots. In *Proceedings of the Seventh International Conference on Epigenetic Robotics: Modeling, Cognitive Development in Robotic Systems*, pages 1335–1342, Piscataway, NJ, USA, November 5-7 2007.

[11] S. Kakade and P. Dayan. Dopamine: generalization and bonuses. *Neural Networks*, 15:549–559, 2002.

[12] J. L. Krichmar. The neuromodulatory system: a framework for survival and adaptive behavior in a challenging world. *Adaptive Behavior*, 16:385–399, 2008.

[13] M. Kubisch, M. Hild, and S. Hofer. Proposal of an intrinsically motivated system for exploration of sensorimotor state spaces. In *Proceedings of the Tenth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, pages 51–56, Orenas, Glumslov, Sweden, November 3 2010.

[14] K. E. Merrick. A comparative study of value systems for self-motivated exploration and learning by robots. *IEEE Transactions on Autonomous Mental Development*, 2:119–131, 2010.

[15] S. Niekum, A. G. Barto, and L. Spector. Genetic programming for reward function search. *IEEE Transactions on Autonomous Mental Development*, 2:83–90, 2010.

[16] P. Oudeyer, F. Kaplan, and V. V. Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11:265–286, 2007.

[17] S. Palaski, C. Vandam, and J. Weng. Modeling dopamine and serotonin systems in a visual recognition network. In *Proceedings of International Joint Conference on Neural Networks, (ICJNN 2011)*, pages 3016–3023, San Jose, California, USA, July 31-August 5 2011.

[18] S. J. Russel and P. Norvig. *Artificial Intelligence: A Modern Approach, Third Edition*. Prentice hall, Upper Saddle River, New Jersey, 2010.

[19] S. Singh, R. L. Lewis, A. G. Barto, and J. Sorg. Intrinsically motivated reinforcement learning: an evolutionary perspective. *IEEE Transactions on Autonomous Mental Development*, 2:70–82, 2010.

[20] J. Weng. Why have we passed neural networks no not abstract well. *Natural Intelligence: the INNS Magazine*, 1:13–22, 2011.

[21] J. Weng. *Natural and Artificial Intelligence: Introduction to computational brain-mind*. BMI press, Okemos, Michigan, USA, 2012.

[22] J. Weng, S. Paslaski, J. Daly, C. Vandam, and J. Brown. Modulation for emergent networks: serotonin and dopamine. *Neural Networks*, 41:225–239, 2013.

[23] H. Wersing, S. Kirstein, M. Gotting, H. Brandl, M. Dunn, I. Mikhailova, C. Goerick, J.Steil, H. Ritter, and E. Korner. A biologically motivated system for unconstrained online learning of visual objects. In *Proceedings of the 16th International Conference on Artificial Neural Networks, (ICANN 2006), Lecture Notes in Computer Science*, pages 508–517, Athens, Greece, Sepetember 10-14 2006.

[24] X. Wu, Q. Guo, and J. Weng. Skull-closed autonomous development: Wwn-7 dealing with scales. In *Proceedings of the International Conference on Brain-Mind,*, pages 1–9, East Lansing, Michigan, USA, July 27-28 2013.

[25] Z. Zheng, K. Qian, J. Weng, and Z. Zhang. Modeling the effects of neuromodulation on internal brain areas: serotonin and dopamine. In *Proceedings of International Joint Conference on Neural Networks, (IJCNN 2013)*, pages 1404–1411, Dallas, Texas, USA, August 4-9 2013.

Comparison between WWN and Some Prior Networks

Zejia Zheng and Juyang Weng

***Abstract*—Although supervised learning is effective, the cost demand on teachers is often too high to be practically applied. How can an agent learn effectively from both motor supervision as well as environmental reinforcement without handcrafting task-specific rules? In this paper we present the Developmental Network framework for emergent development. We compare Deep Learning Network and convolutional neural network with our Where-What-Network (WWN) in two dimensions: whether the network uses emergent representation, and whether the reinforcement system is non-symbolic. We discuss the drawbacks of error-backpropagation for deep learning networks and convolutional filters for convolutional neural networks. In terms of reinforcement learning, Deep Convolutional Network has been recently combined with symbolic Q-learning. In contrast, WWN combines hebbian learning with top-k competition to preserve long-term memory, extracting most relevant information directly from cluttered background. Neuromodulatory systems (i.e. non-symbolic serotonin and dopamine pathways) are integrated into the supervised learning pathways in WWN for reinforcement learning. Experimentally, the motivated DN and WWN have demonstrated capability to tackle traditional reinforcement learning tasks (i.e. simple path finding) as well as multi-concept reinforcement learning tasks (i.e. image classification and location recognition).**

I. Introduction

Modern machines have already out-performed human beings in many domains. But despite the computing powers, we have seen a paradox in the field of artificial intelligence: on the one hand we have those machines out-performing humans in tasks that are generally considered difficult (by humans), but on the other hand those task-specific machines perform poorly in areas that are commonly considered easy (by humans), such as vision, audition, and natural language understanding. What is the proper architecture for general intelligence? Can we generate intelligence by pure logic base manipulation?

Newell and Simon wrote in [18] that a physical symbol system has the necessary and sufficient means for general intelligent action. They define a physical symbol system as a number of symbols related in some physical way (such as one token being next to each other). The hypothesis implies that computers, when we provide them with the appropriate symbol-processing programs, will be capable of intelligent action [19]. Their assumption is the bases for the research from the school of Symbolists. Symbolic approaches have generated success in well-defined problems, but it has been steadily receiving criticisms from various sources.

Zejia Zheng and Juyang Weng are with Michigan State University, East Lansing, MI, USA (email zhengzej, weng@cse.msu.edu). Juyang Weng is also with the MSU Cognitive Science Program and the MSU Neuroscience Program.

The most serious of all criticisms is the grounding problem of all symbolic machines. The symbol grounding problem, which refers to how symbols relate to the real world, was first discussed by Steven Harnard in [10]. The problem describes the scenario where the learner tries to learn Chinese by consulting a Chinese-Chinese dictionary. The endeavour would bound to be unsuccessful because the learner cannot ground the symbol's meaning to the environment. In the case of autonomous agents, we have to take into account that the system needs to interact with the environment on its own. Thus, the meaning of the symbols must be grounded in the system's interaction with the real world.

Researchers have established that a valid solution of the symbol grounding problem will need to combine the bottom-up sensor input with a top-down feedback approach to ground the symbols. The agent should also have developmental ability to develop its internal representation without explicit supervision [26]. Many connectionist networks seems to satisfy those conditions.

The best representatives of the recent development in the field of connectionist models are Deep Learning models (by deep-learning, we mean a cascade of areas like those that have been published in [11] and [3]) and Convolutional Neural Networks (e.g. HMAX and LeCunn Network). A Convolutional neural network is comprised of one or more convolutional layers (often with a subsampling step) and then followed by one or more fully connected layers as in a standard multilayer neural network. Deep learning network is a family of multi-layer neural networks with hierarchies of representation. Both convolutional network and deep belief networks use back-propagation as its training algorithm [11] [14] [3].

A detailed discussion of properties of convolutional networks and deep architectures is presented in the following section, after which we discuss our approach of autonomous learning based on Where-What-Network. Generalization of the network in the direction of reinforcement learning and concept scaffolding is discussed in Section IV. By concept scaffolding, we mean that the learning procedure of complex tasks is facilitated by the already learned simpler tasks, as is observed in stages of early language acquisition where children concatenate the already learned simple words to express complex meanings [1].

II. DN framework and its embodiment

A. Developmental Network framework

Here we define the framework of Developmental Network (DN). DN is introduced in [33] to explain basic principles about an autonomous agent.

A DN has three areas X, Y, and Z. X serves as the input area, Y is called the hidden area in traditional neural network literature, and Z is the motor area.

To generate a corresponding action $\mathbf{z}$ for the current input $\mathbf{x}$, there should be an update function to update the states in each area:

$$(\mathbf{x}', \mathbf{y}', \mathbf{z}', N') = f_{DN}(\mathbf{x}, \mathbf{y}, \mathbf{z}, N) \tag{1}$$

N' stands for the other adaptive part of the network, such as strength of the connection between neurons.

The flow of information can be characterized in the following two steps:

1) $X(t-1) \rightarrow Y(t-0.5) \leftarrow Z(t-1)$ Current input and output pair, current input image and motor supervision, for example, would trigger specific neuron firing in the Y area. The firing neuron in Y would learn the input pair based on the learning algorithm. In WWN, the learning algorithm is LCA. In DBF, the learning algorithm can be incremental version of error back propagation.
2) $X(t) \leftarrow Y(t-0.5) \rightarrow Z(t)$ The internal states of the agent would generate corresponding action in Z area and prediction in X area in the next time stamp.

A DN following the above paradigm does not require the task to be known to the programmer. Without given any task, a human designs (or nature evolves) the general-purpose DP which resides in the DN as a functional equivalence of the genome that regulates the development. A DN can incrementally learn any complex finite qutomata but the DN is grounded to the real world.

One successful embodiment of DN is the Where-What-Network. Where What Networks [15] are a visuomotor version of the Developmental Network, modeling the dorsal (where) stream and the ventral (what) stream of visual and behavioral processing. Where-What Networks (WWN) have been successfully trained to perform a number of tasks such as visual attention and recognition from complex backgrounds [30], stereo vision without explicit feature matching to generate disparity outputs [25], and early language acquisition and language-based generalization [16].

Many other connectionist models seem to be possible candidates for the implementation of DN framework. Recent Deep Learning Networks and Convolutional Networks claims to be inspired by at least part of human brain functions (visual cortex). We look into the learning algorithms of these two architectures in the following section.

B. Deep Architecture and Convolutional Networks

This is a perfect place to examine the details of existing vision-attention models. The first model we choose to compare with here is the convolution networks. Convolutional networks are widely used for its minimum requirement on weight storage space and the ability to find transformation invariants. A convolutional network for visual recognition would be the HMAX model introduced in [24]. HMAX is a hierarchical model for visual cortex. C1 Feature vector is extracted by pooling local maximum responses from S1 neurons, which are essentially layers of gabor filters with specific orientation preferences. C1 Feature vector is then fed into a dictionary of learned feature vectors(S2) to extract the final feature vector. Learning in HMAX is somehow rudimentary as the low-level feature detector is handcrafted and non-adaptive.

TABLE I
COMPARISON BETWEEN WWN AND CONVOLUTIONAL NETWORK

	Conv Net	WWN
Feature Detector	Shared	Locally Optimal
Top-down attention	No	Yes
Multi-concept	No	Yes
Sparse Response	No	Yes

Our model, WWN, is better suited for autonomous development compared to convolutional models in the following aspects:

1) Filter optimality. The low-level features (receptive fields of S1 units) in convolutional networks like HMAX are handcrafted instead of learned. We have to stack layers of gabor filters at the same location before we can have better performance. In WWN, spatial and temporal optimality in LCA learning algorithm guarantees efficiency of the learned feature detectors.
2) Top-down attention. Convolutional neural networks models like HMAX, LeNet [14] or NeoCognitron [8] are strictly feedfoward ventral pathway models without top-down feedbacks. Inspired by the bidirectional connections prevalent in the visual cortex, we modeled the top-down effect of attention as the top-down connection from the motor layer to the hidden layers. This allows guided feature extraction from the bottom-up inputs and gives better representation results. It is worth to notice that other models have top-down attention mechanisms as well. For example, Walther and Koch [28] modeled object-specific attention in HMAX using feedback connections from the object layer back to the complex features CF in order to infer suitable complex feature maps. Their model is similar to the top-down connection from the TM in WWN to the hidden layers. WWN has one more set of top down connection from LM to the hidden layers, allowing the network to pay special attention to certain locations in the image. The combination of space based attention and object based attention in one single network is an important novelty in WWN.
3) Multi-concept learning. As far as we know, there is no single network that can report multiple concepts at the same time. Existing visual models, especially those with convolution filters, would first determine whether the object is present in the given image, and then convolve the image with a template to find the location. However, ventral and dorsal pathways work in parallel instead of sequential. The recognition of 'where' should independent of the recognition of 'what' [9]. In WWN,

TABLE II
COMPARISON BETWEEN WWN AND DEEP LEARNING NETWORKS

	WWN	Deep Nets
Major Assumption	Optimal Input Action Abstraction	Optimal Input Reconstruction
Long Term Memory	Yes	No
Emergent Hierarchy	Yes	No

the two pathways share the same feature maps (e.g. the short-time activation in Y area). Firing and learning in one motor (e.g. LM), however, can be affected by firing in the other motor (e.g. TM) via the top-down connection from motor to Y (e.g. TM$\rightarrow Y$), thus WWN is more biologically plausible compared to sequential localization architectures.

4) Sparse response. WWN supports the sparse coding idea of Olshaushen and Field [20]: each cortical area has few neurons to fire. But our model is different in a number of important aspects. Instead of starting from a hypothetical objective function, our model starts from the cortical connection patterns experimentally observed in many physiological studies, as surveyed and summarized by Felleman and Van Essen [7]. HMAX, as is proposed in [24], does not have sparse response. Hu and co-workers attempted to combine HMAX with sparse coding [12], but the sparsity lies only in the S1 layer instead of the entire network.

Another important class of visual recognition network is the Deep Learning Networks. Much of this work appears to have been motivated by the hierarchical organization of the cortex, and the output of the networks are often compared with the simple receptive fields found in V1 and V2 [23] [21].

Here we argue that WWN is more biologically plausible compared to Deep Architectures in the following aspects:

1) Major assumption. The major assumption of deep architectures for visual recognition is to model the distribution of training data using limited neurons. In other words, deep architectures like Deep Belief Networks [11] and Stacked Autoencoders [27] aim to minimize the representational error of the training dataset. This assumption however, is highly unlikely in the course of learning, as learning is closely connected with the immediate action. Therefore, A better option is to minimize the error in action, or motor skills, rather than just to restore the input image. By implementing the bidirectional connection from the motor layer to the hidden layer, WWN links its internal representation directly with actions, minimizing the error in motor skills via hebbian learning.
2) Learning algorithm. Most deep architectures use error back-propagation as its learning algorithm [11] [27] [23]. Error back-propagation finds the gradient of the error function with respect to the current parameter , which is the entire set of weights in the network (or the current layer if layer wise training is used). In this sense, error back propagation would adjust all the weights, however minutely, as it is doing gradient descent. There is no competition in hidden neurons when training with back-prop. This corrupts the long-term memory in the network because neurons irrelevant to the current input would be forced to adjust their weight when error back propagation is finding the gradient. LCA however, only adjust the weights of the neurons with the highest response values, preserving the long-term memory in other neurons.
3) Emergent hierarchy. As is discuss in previous sections, hierarchical architecture would emerge in WWN via synaptic maintenance [29]. Learning is not constrained in specific layers within WWN. Deep architecture, however, hand-craft the layers and train them by greedy layer-wise training [11]. Biologically speaking, it is unlikely that V1 neurons would develop prior to V2 neurons. It is even more unlikely that V1 neurons would freeze their development while V2 neurons are fine-tuning their synaptic weight. WWN offers a developmental hierarchical model for visual cortex, which is more biologically plausible.

C. *Generalization: Reinforcement Learning*

In terms of reinforcement learning, techniques used in traditional reinforcement learning include: the dynamic-programming approach which uses basic Bayesian reasoning to solve for an optimal strategy [4], Markov decision processes (MDPs) models which can handle delayed reinforcement [5], adaptive heuristic critic algorithm in which the value-function computation is implemented by a reinforcement-learning component modified to deal with multiple states and non-stationary rewards [2], and Watkins Q-learning algorithm which is able to compare the expected utility of the available actions without requiring a model of the environment [32]. These are all symbolic methods which suffers from the ground problem as is explained at the beginning of the section.

Reinforcement learning in Deep Learning architectures or convolutional neural networks, however, is a relatively new topic and has not been studied until recently. The proposed method in [17] combines deep learning with Q-learning method to train an agent to learn how to play a simulated shooting game. A similar approach is proposed in [6] which combines Dayan style planning (based on Markov states) with Deep Belief Networks for image observation. The reinforcement learning modules in these networks are all symbolic modules which are non-developmental and contradicts the purpose of autonomous development.

WWN, on the other hand, deals with reward and punishment by simulating the dopamine serotonin pathways in human brain. We implemented neuro-modulation in WWN by integrating reward and punishment learning pathways with the supervised learning system, allowing the network to explore without supervision. Concept scaffolding in location motor enabled the agent to learn finer location concepts with the help from previously learned skills. Our generalization uses

TABLE III
TWO DIMENSIONAL COMPARISON AMONG LEARNING METHODS

	Symbolic	Emergent
Not Motivated	Finite Automata	Deep Learning Conv Nets
Motivated	Q-learning	WWN

consistent neural network based modulatory systems compared to the traditional symbolic methods (e.g. Markov state based methods). The generalization is more biologically plausible and more compatible with our Developmental Network framework for autonomous development compared to existing reinforcement learning modules for deep architectures and convolutional networks.

III. NETWORK ARCHITECTURE

A. Where What Network

Here we introduce how learning takes place in Where-What-Network.

It is known that our visual system has two major pathways: ventral (what) for object identification and dorsal (where) that deals more with visuomotor aspects (i.e., where to reach for an object), which presumably codes an objects location. These pathways separate from early visual areas and converge at prefrontal cortex, which is known to be active in top-down attention. Prefrontal cortex connects to motor areas. WWN was built inspired by the idea of these two separating and converging pathways, as is illustrated in Fig.1. Meaningful foregrounds in the scene will compete for selection in the ventral stream, and locations in the scene will compete for processing in the dorsal stream.

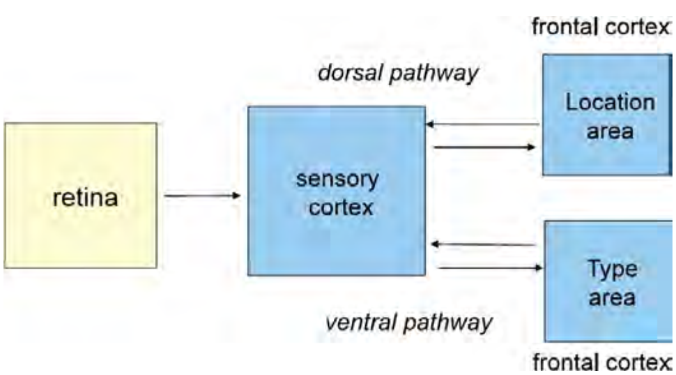


Fig. 1. A schematic of the Where-What Networks (WWN). It consists of a sensory cortex which is connected to the What area in the ventral pathway and to the Where area in the in the dorsal pathway.

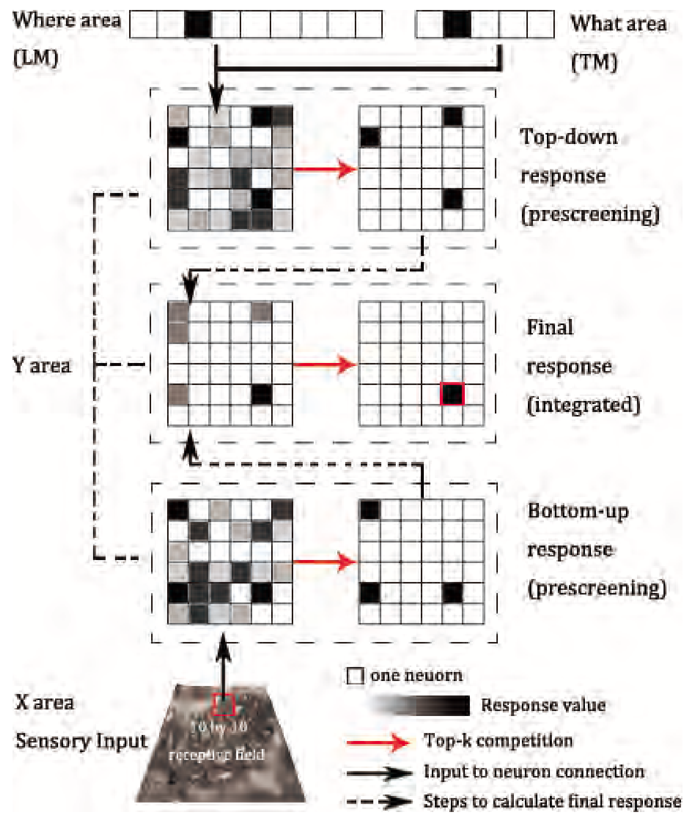


Fig. 2. An example of activation patterns and neuronal changes during learning process in the network. The network selects winning neuron based on top-k competition over the top-down responses and bottom-up responses. The pattern on the left shows the preresponse of the neuron, and the grid on the right shows the final response of the neuron. The winning neuron is marked in red at the final response layer.

B. The learning algorithm

The learning algorithm of WWN is described in detail in previous papers [25] [15] [30].

1) Pre-response of the neurons: Each brain area uses the same area function f, which allows the internal area to develop representations of the training data. Generally speaking, each neuron in internal area stores two sets of weight vector $(\mathbf{v}_b.\mathbf{v}_t)$, representing bottom-up weight and top-down weight separately. Similarly, neurons in motor area only have bottom-up weight, while neurons in sensors only have top-down weight. The top-down weight in sensors is useful when we need to predict future images based on current internal responses. In the current program we do not need that set of weight.

The pre-response value for each neuron is calculated as:

$$r(\dot{\mathbf{v}}_b, \dot{\mathbf{b}}, \dot{\mathbf{v}}_t, \dot{\mathbf{t}}) = (\dot{\mathbf{v}}_b \cdot \dot{\mathbf{b}} + \dot{\mathbf{v}}_t \cdot \dot{\mathbf{t}})/2 \tag{2}$$

where $\mathbf{b}$ and $\mathbf{t}$ are bottom up input and top down input respectively. Each vector in equation (2) is normalized before calculation:

$$\dot{\mathbf{v}} = \mathbf{v}/||\mathbf{v}||$$

Each neuron in the Y area extracts local input from the input image. The local window is called *receptive field* of that neuron, depicted in Fig. 2 as the red box in the input image.

Neurons in the Z area accepts the global response values of all the neurons in the Y area as bottom up input. The response values are calculated based on top-k competition explained in the following subsection.

2) Top-k competition: The final neuron response in Y area is given by top-k competition. The k neurons with the highest pre-response value would fire with the adjusted responses while other neurons would be suppressed. To adjust the response values based on their ranking:

$$r' = \begin{cases} r \cdot (r - r_{k+1})/(r_1 - r_{k+1}) & r_{k+1} \leq r \leq r_1 \\ 0 & otherwise \end{cases}$$

where r_1 is the highest response value; r_{k+1} is the $k+1$th highest response value.

3) Hebbian-like learning: If a neuron wins in the multistep lateral competitions described above, its bottom up weight and top down weight would update using the following Hebbian learning rule:

$$\mathbf{w}_{u,i} \leftarrow \beta_1 \mathbf{w}_{u,i} + \beta_2 r_i \mathbf{x}_t$$

where $\beta 1$ and $\beta 2$ determine retention and learning rate of the neuron, respectively:

$$\beta_1 = \frac{m_i - 1 - \mu(m_i)}{m_i}, \beta_2 = \frac{1 + \mu(m_i)}{m_i} \tag{3}$$

with $\beta_1 + \beta_2 \equiv 1$, m_i is the neuron's firing age, i.e. $m_i = 1$ in the beginning of training, and increments by one every time the neuron wins lateral competition.

μ is a monotonically increasing function of m_i that prevents the learning rate β_2 from converging to zero as m_i increases.

$$\mu(m_i) = \begin{cases} 0, & \text{if } m_i < t_1 \\ c(m_i - t_1)/(t_2 - t_1), & \text{if } t_1 < m_i < t_2 \\ c + (t - t_2)/\gamma, & m_i > t_2 \end{cases}$$

Typical value for those parameters in the experiments reported afterwards would be : $t1 = 10, t2 = 10^3, c = 1 and \gamma = 10^4$.

The same Hebbian learning rule updates the top-down weights of neuorns using similar equation:

$$\mathbf{w}_{d,i} \leftarrow \beta_1 \mathbf{w}_{d,i} + \beta_2 r_i \mathbf{r}_t$$

The firing Z neuron accepts Y area firing patterns as bottom up input and updates using the same Hebbian learning rule.

IV. Generalization

To achieve full autonomous development, the agent must be able to learn without direct supervision. That means that the agent must be able to refine its existing motor skills based on previous experience. Inspired by the instructional scaffolding process proposed by Applebee in [13] when he was observing how young children learn to write, we implement a coarse to fine location concept learning scheme to the network which allows the network to learn finer location concepts using previously learned locations. To learn from exploration also requires the agent to have some sort of reinforcement learning mechanism. We integrated the reward and punishment learning pathways into WWN network, modeling the neuromodulatory systems found in human brains.

A. Motor neuron splitting: use old knowledge to learn new skills

When a young child learns to recognize the object at certain location, he learns the spatial concepts gradually. He also uses the already learned knowledge to help him acquire new skills. [31] shows how children develop projective/Euclidean understanding before they could comprehend *in front of* and *behind*.

Our network currently models this procedure in the location motor (LM), but similar procedures can be applied to type motor (TM) easily. Coarse to fine learning is achieved by splitting each of the location motor neuron into four child neurons, each representing a finer spatial concept, as is illustrated in Fig.3.

The following steps would take place when LM splitting occurs:

1) Child location motor neuron copies $\mathbf{v}_b$ from parent neuron.
2) Firing age of child location motor neuron set to 1 (or a very low number).
3) Y area neurons copy connection to the parent neuron $\mathbf{v}_t$ to the child location motor neuron.

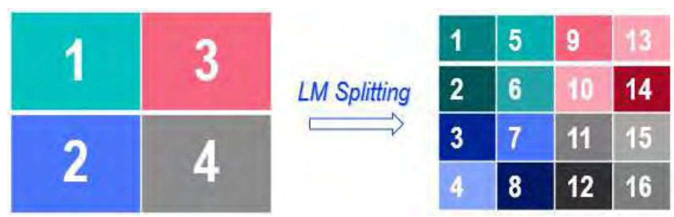


Fig. 3. LM splitting mechanism. Each location motor neuron splits into four child neurons to learn finer spatial concepts.

B. Reward and punishment pathways: Modulated Developmental Network

The architecture of the two pathways are introduced in detail in the previous papers [22] [34]. Here we only give a brief summary due to limited space.

Previous work modeled two neurotransmitters in human brain: serotonin and dopamine. These two neurotransmitters are released separately in rewarded or punished events. The model simplifies the role of those two neurotransmitters, i.e. dopamine is released when the agent is rewarded, and serotonin is released when the agent is punished. The papers demonstrated that the reward and punishment system built based on Developmental Network enables the agent to learn according to the sweetness(reward) and pain(punishment) it receives when making educated guesses rather than specific instructions of correct movement(supervision).

Although Fig.4 has 11 areas on the plot, in the program we simplify that architecture into three pathways: $X_u \rightleftharpoons Y_u \rightleftharpoons Z_u$ is the unbiased pathway, $\{X_u, X_p\} \rightleftharpoons Y_p \rightleftharpoons Z_p$ is the punishment pathway, and $\{X_u, X_p\} \rightleftharpoons Y_s \rightleftharpoons Z_s$ is the reward pathway.

VTA and RN are treated as conceptual area that trigger firing when reward or punishment is present, corresponding to two *if* clauses in the program. Y_{VTA} and Y_{RN} are the same areas as Y_p and Y_s, using different neuromodulators (i.e. GABA for inhibitory connections, and glutamate for excitatory connections.) This means that these two regions are still active even when no punishment and reward is present, allowing the network to recall punished and rewarded instances with pure input from X.

The network calculates three responses: $\mathbf{r}_u$ for response value in the unbiased pathway, $\mathbf{r}_p$ for response value in the punishment pathway and $\mathbf{r}_s$ for response value in the reward pathway.

The final response value is given by:

$$r_i \leftarrow r_{iu}(1 + r_{is} - \gamma r_{ip}) \tag{4}$$

γ is usually larger than 1, indicating that inhibition from the pain pathway is much more effective compared to excitation from the reward pathway.

Another effect of neuromodulators is that they would increase the learning rate in the corresponding areas. This would change equation 3:

$$\beta_1 = 1 - \beta_2, \beta_2 = \alpha \cdot \frac{1 + \mu(m_i)}{m_i} \tag{5}$$

In the experiment, $\alpha = 2$.

C. Maze exploration: an example

To illustrate how reinforcement learning to take place in DN (or WWN), we discuss in detail how DN can solve traditional reinforcement learning tasks in this section. The maze and agent track is shown in Fig.5.

The task for the agent is to find the food at the lower right corner of the maze by exploration. These are the constraints imposed on the agent to make the task non-trivial:

1) Limited sight. The agent can only see the tile immediately adjacent to its current position.
2) Walls. For this demonstration, we built walls at 3 places on the map (dark blue tiles in the figure). When the agent bumps into walls, it would receive a pain signal of level 5.
3) Sense of smell. The food gives out some sort of oder. The further the agent is away from the food, the less intensive the smell is. Thus the agent can compare the current scent with the scent of the previous states. If the scent is intensified by the action (say step to the right), the agent would receive a mild reward of level 1. If the scent is the same, the agent receives no reward or punishment. If the scent is weakened, the agent would perceive this as a punishment of level 1.
4) Ultimate reward. Finding the food would give the agent a reward of level 10.

As is shown in the figure, the agent starts wondering around for a little bit. But it soon learns that stepping back is bad and hitting walls is even worse. After 40 epochs of training, it finds the best route to the food.

There is no objective function involved in the teaching scheme. The agent learns that food is 'good' from the input of its reward sensor. Similarly, the concept to 'avoid hitting walls' is learned by avoiding punishment during exploration. As is proposed in the introduction, the training for this agent is task-non-specific. The programmer does not need to know what is in the environment or what the task is.

This example is not intended for a full scale implementation of maze exploration. The agent in the example is short sighted and can not plan ahead, thus would perform poorly in a natural environment. To find a path in a complicated environment, the agent must be able to develop the idea of localization. This means that we need one more Z area to allow the agent to develop the concept of 'states'. The bidirectional connection between Y area and Z area would allow the agent to handle

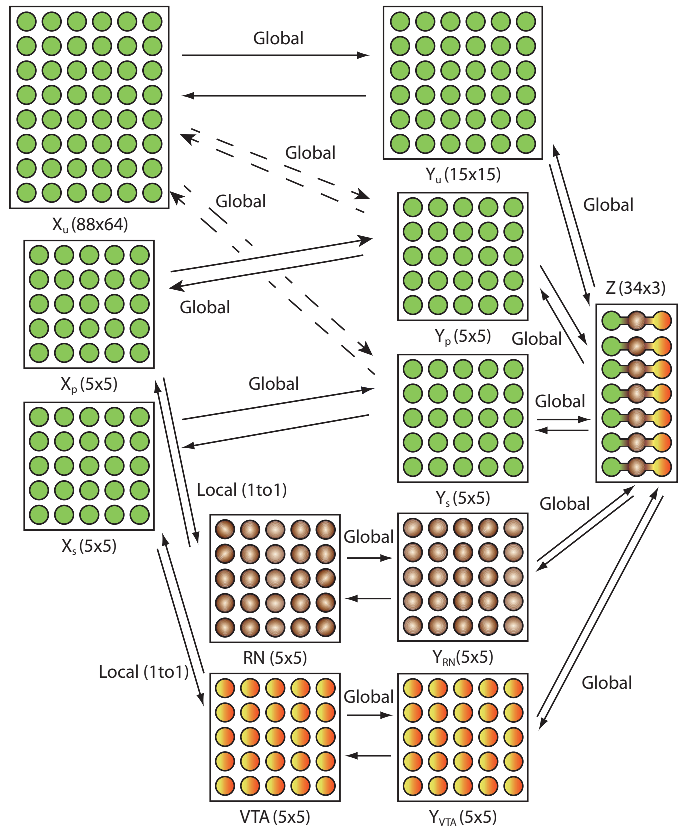


Fig. 4. A DN with 5-HT and DA modulatory subsystems. It has 11 areas. RN has serotonergic neurons. Neurons in Y_{RN} or Z have serotoninergic synapses. VTA has dopaminergic neurons. Neurons in Y_{VTA} or Z have dopaminergic synapses. The areas $Y_u, Y_p, Y_s, Y_{RN}, Y_{VTA}$ should reside in the same cortical areas, each represented by a different type of neurons, with different neuronal densities. Within-area connections are simulated by top-k competition.The number of neurons in the figure does not necessarily corresponds to the number of neurons we use in the experiment. The figure does not show all the neurons in the subareas due to limited space. The Z neurons compute unbiased response(the first column, using glutamate), inhibition response (second column, using serotonin), and excitation response (third column, using dopamine). "Global" indicates that the neurons accept input from all neurons in the previous area. RN (raphe nuclei) is the subarea inside human brain that secretes serotonin. VTA (Ventral tegmental area) is the subarea that secretes dopamine. Basically Xp and Xs does not produce those neurotransmitters but only sends input signals to those two brain areas. The function of these two layers is for biological correctness. Thus each neuron in X_p (or X_s) corresponds to one neuron in RN (or VTA). The link is one to one, thus "Local".

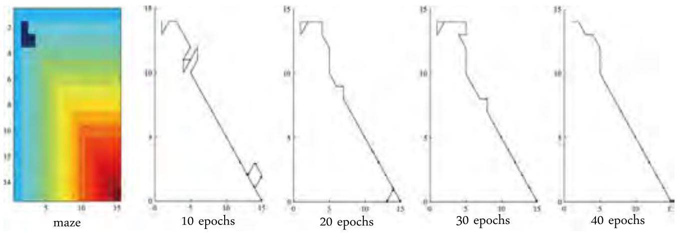


Fig. 5. Maze exploration example. The agent, starting from the top left corner of the maze, aims to find the food located at the lower left corner of the maze.

delayed reward and form strategic plans.

D. Other experiments

The Modulated Developmental Network has been tested in the following tasks:

1) Face recognition task with reinforcement learning teaching schedule. We collected 913 training images and 171 testing images of 33 people, with each person representing a different class. The images are fed to the agent sequentially. Given a image, the network would make an educated guess based on its current network weights, and then the teacher would issue reward or punishment based on the educated guess. The sample input is presented in Fig.6. The result is presented in Fig.7. The result is originally reported in [34]. We showed that Modulated DN with dynamically changing learning rate and top-k competition would give better learning result compared to the original DN.
2) Integrated learning in WWN with concept scaffolding. We trained the agent to recognize location and type information of the foreground object. The complexity of the task resulted in too many wild guesses which impaired learning in the network. Concept scaffolding in LM was implemented in the network to solve this issue. This allows the network to learn new finer location based on its already learned motor skills. This minimizes the number of educated guesses and thus gives better learning in WWN. The experiment details are reported in [35].
3) Realtime training and testing. We developed a graphic user interface, shown in Fig.8, to train and test the agent with real-time video stream. The interface grabs image from ip camera realtime and trains the network according to the set up parameter. The agent learns directly from cluttered background at a speed of 4 frames per second. The network is currently trained under supervised learning mode. Future interface improvement involves integration of reinforcement learning and concept scaffolding with the current training scheme.

E. Discussion

The reward and punishment pathway enabled our agent to develop likes and dislikes. The system is different from traditional symbolic reinforcement learning in the sense that the system does not need a symbolic representation. The

Fig. 6. Sample training images for face recognition task. Each picture contains a different expression. The data set contains 913 different images of 33 people.

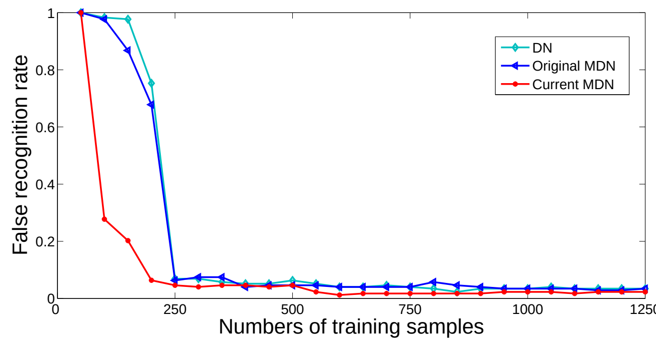


Fig. 7. Speed of learning in DN and MDN for face recognition task. Red line: Modulated Developmental Network with dynamically changing learning rate and top-k competition. Blue Line: Original Modulated Developmental Network. Green line: Developmental Network. X axis: Number of Training Samples. Y axis: Recognition error rate. Training samples are randomly selected from the 913 face image dataset. Our version of MDN converges to the best recognition much faster than DN and original MDN. Original MDN and DN shares the same recognition rate, while our version of MDN gives a better recognition rate than the former methods.

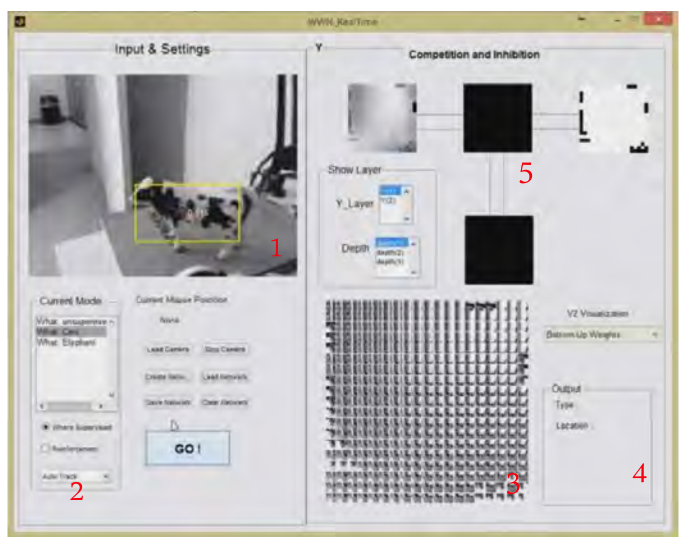


Fig. 8. GUI layout for realtime reinforcement learning WWN. 1.Viewer. Shows current input video. When teaching also highlights the supervised location. 2.3 modes of supervision: auto tracking, which calls the built in tracking algorithm to feed the agent with the location information; mouse tracking, which tracks the movement of mouse on the screen; and mouse click, which only updates with the current frame and the location where the teacher clicks mouse. 3. Bottom-up weight visualization of current layer. 4.Output. When testing shows detected type and location. 5.Firing in the Y area. A short demonstration is available at https://www.youtube.com/watch?v=id9j7Pb1bOA.

learning mode can be switched back to supervised learning on-line.

For complicated tasks, the agent would learn to solve the puzzle based on its previous experience via concept scaffolding. The neuron splitting scheme can be easily generalized to other motor areas (TM, for example).

Our network is the first integrated learning system that combines reinforcement learning and supervised learning algorithms. The research presented here is important to achieve full autonomous development. Future works involve the development of planning mechanism in the brain area. We would also test the network with large data sets to check its representational capability.

V. Conclusion

In this paper we present the DN framework for autonomous development and WWN as its embodiment. By comparing WWN with the existing networks we argue that WWN overcomes the underlying flaws in those networks and would thus be an ideal candidate for modelling development with biologically plausible learning algorithms. We present our research on the generalization of the network in the direction of reinforcement learning and concept scaffolding. Our result shows that by adding the two biologically inspired pathways the network can learn based on environmental feedbacks without direct supervision. The work presented in the paper is an important stepping stone to achieve full autonomous development.

References

[1] Arthur N Applebee and Judith A Langer. Instructional scaffolding: Reading and writing as natural language activities. *Language arts*, pages 168–175, 1983.

[2] Andrew G Barto, Richard S Sutton, and Charles W Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *Systems, Man and Cybernetics, IEEE Transactions on*, (5):834–846, 1983.

[3] Yoshua Bengio. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, 2(1):1–127, 2009.

[4] Donald A Berry and Bert Fristedt. *Bandit Problems: Sequential Allocation of Experiments (Monographs on Statistics and Applied Probability)*. Springer, 1985.

[5] Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena Scientific Belmont, MA, 1995.

[6] Ryan Faulkner and Doina Precup. Dyna planning using a feature based generative model. [Online; accessed 2-April-2014].

[7] Daniel J Felleman and David C Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, 1(1):1–47, 1991.

[8] Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202, 1980.

[9] Melvyn A Goodale and A David Milner. Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25, 1992.

[10] Stevan Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1):335–346, 1990.

[11] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.

[12] Xiaolin Hu, Jianwei Zhang, Jianmin Li, and Bo Zhang. Sparsity-regularized hmax for visual recognition. *PloS one*, 9(1):e81813, 2014.

[13] Judith A Langer and Arthur N Applebee. Reading and writing instruction: Toward a theory of teaching and learning. *Review of research in education*, 13(1):171–194, 1986.

[14] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[15] Matthew Luciw and Juyang Weng. Where what network 3: Developmental top-down attention with multiple meaningful foregrounds. In *International Joint Conference on Neural Networks*, pages 4233–4240, 2010.

[16] Kajal Miyan and Juyang Weng. Wwn-text: Cortex-like language acquisition with what and where. In *Development and Learning (ICDL), 2010 IEEE 9th International Conference on*, pages 280–285. IEEE, 2010.

[17] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[18] Allen Newell and Herbert A Simon. Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126, 1976.

[19] Nils J Nilsson. The physical symbol system hypothesis: status and prospects. In *50 years of artificial intelligence*, pages 9–17. Springer, 2007.

[20] Bruno A Olshausen, Charles H Anderson, and David C Van Essen. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience*, 13(11):4700–4719, 1993.

[21] Simon Osindero, Max Welling, and Geoffrey E Hinton. Topographic product models applied to natural scene statistics. *Neural Computation*, 18(2):381–414, 2006.

[22] Stephen Paslaski, Courtland VanDam, and Juyang Weng. Modeling dopamine and serotonin systems in a visual recognition network. In *Neural Networks (IJCNN), The 2011 International Joint Conference on*, pages 3016–3023. IEEE, 2011.

[23] Christopher Poultney, Sumit Chopra, Yann L Cun, et al. Efficient learning of sparse representations with an energy-based model. In *Advances in neural information processing systems*, pages 1137–1144, 2006.

[24] Thomas Serre, Lior Wolf, Stanley Bileschi, Maximilian Riesenhuber, and Tomaso Poggio. Robust object recognition with cortex-like mechanisms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(3):411–426, 2007.

[25] Mojtaba Solgi and Juyang Weng. Developmental stereo: Emergence of disparity preference in models of the visual cortex. *Autonomous Mental Development, IEEE Transactions on*, 1(4):238–252, 2009.

[26] Mariarosaria Taddeo and Luciano Floridi. Solving the symbol grounding problem: a critical review of fifteen years of research. *Journal of Experimental & Theoretical Artificial Intelligence*, 17(4):419–445, 2005.

[27] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 9999:3371–3408, 2010.

[28] Dirk B Walther and Christof Koch. Attention in hierarchical models of object recognition. *Progress in brain research*, 165:57–78, 2007.

[29] Yuekai Wang, Xiaofeng Wu, and Juyang Weng. Synapse maintenance in the where-what networks. In *Neural Networks (IJCNN), The 2011 International Joint Conference on*, pages 2822–2829. IEEE, 2011.

[30] Yuekai Wang, Xiaofeng Wu, and Juyang Weng. Skull-closed autonomous development: Wwn-6 using natural video. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pages 1–8. IEEE, 2012.

[31] Susan Wanska. The relationship of spatial concept development to the acquisition of locative understanding. *The Journal of genetic psychology*, 145(1):11–21, 1984.

[32] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

[33] Juyang Weng. *Natural and Artificial Intelligence: Introduction to Computational Brain-Mind*. BMI Press, East Lansing, MI, 2012.

[34] Zejia Zheng, Qian Kui, Juyang Weng, and Zhengyou Zhang. Modeling the effects of neuromodulation on internal brain areas: Serotonin and dopamine. In *Neural Networks (IJCNN), The 2013 International Joint Conference on*. IEEE, 2013.

[35] Zejia Zheng, Juyang Weng, and Zhengyou Zhang. Wwn: Integration with coarse-to-fine, supervised and reinforcement learning. In *Neural Networks (IJCNN), The 2013 International Joint Conference on*. IEEE, 2014.

Salient Region Detection by Combining Background Contrast and Connectivity

Jiamei Shuai, Zhiguo Ma, Laiyun Qing, and Jun Miao, *Member, IEEE*

***Abstract*—Saliency detection is a challenging and important problem in computer vision. Early works [1] mostly focused on predicting the fixation of human eyes, while recent works [2] mainly work on predicting salient objects in natural images. In this work we focus on the task of object level salient region detection.**

Previous approaches are mainly based on the contrast prior of the foreground object, usually utilizing the appearance contrast with image neighborhood of a certain scope. Although various computational models have been proposed, the problem remains challenging due to the large variation of the objects, complex textures in background, etc.

As a first exploration of the background prior in saliency detection, Wei [2] *et al.* make use of the observation that most image boundary areas belong to background and the background is usually large and connected, they propose a salient region detection algorithm by measuring the saliency of an image patch using the length of its shortest path to the image boundary. Some follow-up works along this direction are proposed recently and achieve quite impressive results on various salient region detection datasets.

In this work, we propose a novel salient region detection algorithm based on background prior with two saliency measurements proposed. Instead of measuring the contrast between objects and its neighborhoods, we evaluate the saliency of an image region by measuring its distance with the estimated background in both feature domain and connectivity domain.

The input image is first preprocessed to get an appropriate superpixel representation. In order to suppress small-scale textures of images that are supposed to be not sensitive for human vision system, we employ a structure extraction algorithm [3] to smooth out the texture areas of the input image. Superpixel segmentation [4] is then applied to the texture-suppressed image to reduce the number of processing units. Superpixels based on the texture-suppressed image provide an efficient image representation with homogeneous elements. Then we estimate the background using the superpixels near the image boundary by utilizing the location prior of background.

Motivated by the observation that the variations of image background can be mostly described in the estimated background, we make a statistics of the distance between the foreground/background superpixels with the estimated background. For each background, we select several most similar superpixels in the estimated background and the mean Lab average color distances between the background superpixel and each of these selected superpixels are calculated. Then the average summation these distances in each image is plotted in a histogram. Same procedure is applied to get the histogram of the distance between the salient object and the estimated background. The statistics show that the distances between the image background with the estimated background are relatively small and tend to be zero, while the distances between the salient region with the estimated background are relatively large and can be highlighted using this measurement.

Based on the observation stated above, we propose a saliency measurement called background contrast. The background contrast of a superpixel is defined as the summation of its *k* minimum color distances to the superpixels of estimated background.

The surround cue of the object is also incorporated by the measurement of background connectivity. As the Gestalt psychological studies suggest that several factors are likely to influence figure-ground segregation, e.g. size, surround and symmetry, we explore the surround cue for saliency detection. Regarding the superpixels as nodes in a weighted graph and the weight between neighboring nodes is the color distance, we define the background connectivity of a superpixel as the shortest geodesic path among each superpixel to its *k* most similar superpixels of the estimated background. The feature utilized is mean *Lab* color of each superpixel and the distance is measured in L_2.

Then we integrate the two measurements by linear combination and finally a post-processing involving spatial and color adjacency is employed to generate a per-pixel saliency map.

Experiments on three publicly available salient region detection datasets show that our algorithm performs favorably against the state-of-the-art algorithms. The incorporation of texture suppression is also proved to be effective for improving the detection accuracy.

This research is partially sponsored by Natural Science Foundation of China (Nos.61070116, 61175115, and 61272320).

J.-M. Shuai is with University of Chinese Academy of Sciences, Beijing 100049, China.

Z.-G. Ma is with Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China.

L.-Y. Qing is with University of Chinese Academy of Sciences, Beijing 100049, China (corresponding author, e-mail: lyqing@ucas.ac.cn).

J. Miao is with Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing100190, China.

References

[1] Laurent Itti, Christof Koch, and Ernst Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, 1998.

[2] Y. Wei, F. Wen, W. Zhu, J. Sun, "Geodesic saliency using background priors," in Proc. European Conference on Computer Vision (3), 2012, pp. 29-42.

[3] Li Xu, Qiong Yan, Yang Xia, and Jiaya Jia, “Structure extraction from texture via relative total variation,” ACM Trans. Graph., vol. 31, no. 6, pp. 139, 2012.

[4] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Au- re´lien Lucchi, Pascal Fua, and Sabine Su sstrunk, “Slic superpixels compared to state-of-the-art superpixel methods,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, 2012.

An Experimental Study on Simulating Human Eye Searching Paths Using an Eye-Movement Control Model Based on Neural Coding of Visual Context

Lijuan Duan, Zhiguo Ma, Jili Gu, Jun Miao

Abstract—**Searching tasks by humans are performed in a more intuitive and efficient manner by selecting only a few regions to focus on, while observers never form a complete and detailed representation of their surroundings.**

Usually two kinds of top-down cues are used for gaze movement control in target searching: the cues about targets such as shape, color, scale and the cues about the visual context that contains the target and the relevant objects or environmental features with their spatial relationship. Torralba used global context to predict a horizontally long narrow region where the target is more likely to appear. Ehinger and Paletta used object detectors to search the target in such predicted region for accurate localization. Kruppa and Santana used an extended object template containing local context to detect extended targets and infer the location of the target via the ratio between the size of the target and the size of the extend template. Different from the above methods which adopted global context or local context cues separately, Miao and et al. proposed a serial of eye-movement control models based on neural coding of both global context and local context cues to control a virtual gaze movement for target searching [1-4].

In this study, we simulate human searching paths by Miao et al's model [4] in a facial feature locating task on a face dataset. To test the performance of the introduced model, we collect human eye-movement data of 27 subjects by a SMI iVIEW X Hi-Speed eye tracker when the subjects are carrying out the facial feature locating task on the face dataset. We compare the searching paths generated by the introduced model against human eye-movement data. Experimental results demonstrate that the model achieves a good prediction accuracy on both dynamic scan-paths and static fixation locations.

Twenty-seven college students from Beijing University of Technology participated in this study, with 15 females and 12 males. A set of 30 face pictures are prepared as stimuli. Among this set, 15 pictures are female-face, and the rest are male-face. The size of each picture is 1024×768 pixels. Pictures are presented on a color computer monitor at a resolution of 1024×768 pixels. The monitor size was 41 cm by 33.8 cm, and the participants were sited in a chair about 76 cm in front of the screen. The independent variables of the experiment are search time and fixation numbers. The dependent variable measures are the gender that the participants belonged to, the sex gender of the face pictures being shown, the initial points from four different quadrants and the facial feature to be located: the left eye or the right eye. Following a nine-point calibration procedure, participants were shown written instructions asking them in preparation for a target search test in the following face pictures. For each trial, a black indicator was first appeared in the middle of the white screen for 1000 ms. Then a "+" was presented at one position of four starting points in a random order to make sure searching from the pre-defined positions. After a face picture presented in the middle of the screen for 2000 ms, the participants were asked to find the target eye as accurately and quickly as possible. Participants were demanded not to look at anything else in the pictures after they had found the target.

We compare the searching paths generated by the model with the scan-paths recorded from the subjects. The two paths are all divided into pieces each consisting of two adjacent fixations. The Hausdorff distance is used to evaluate the predicting performance of the searching paths by the model with the scan-paths of all subjects recorded by the eye tracker. The scan-paths between different subjects are also evaluated. We also compute the precision for target locating via the searching paths from four different quadrants to left eyes and right eyes.

Experimental results show that the simulated scan-paths are similar to human saccades: the average of the Hausdorff distances between the searching paths generated by the model and the subjects on all the corresponding pictures is 29.18, which is similar to the average Hausdorff distance of 26.36 between the scan-paths generated by every two subjects of the total 27 subjects. Experimental results also show that the model achieves an average search precision that is above 96%.

This research is partially sponsored by Natural Science Foundation of China (Nos.61070116, 61175115, and 61272320).

L. Duan is with the College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China.

Z. Ma is with the Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China.

J. Gu is with the College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China.

J. Miao is with the Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China. (Corresponding author, e-mail: jmiao@ict.ac.cn).

References

[1] J. Miao, X. Chen, W. Gao, and Y. Chen, A visual perceiving and eye ball-motion controlling neural network for object searching and locating, in Proc. Int. Joint. Conf. Neural Netw., Vancouver, BC, Canada, 2006, pp. 4395–4400.

[2] J. Miao, B. Zou, L. Qing, L. Duan and Y. Fu, Learning internal representation of visual context in a neural coding network, in Proc. Int. Conf. on Artificial Neural Networks, Thessaloniki, Greece, 2010, vol. 6352, pp. 174–183.

[3] J. Miao, L. Qing, B. Zou, L. Duan and W. Gao, Top-down gaze movement control in target search using population cell coding of visual context, IEEE Trans. Autonomous Mental Development, vol. 2, no. 3, September 2010.

[4] J. Miao, L. Duan, L. Qing and Y. Qiao, An Improved Neural Architecture for Gaze Movement Control in Target Searching, Proc. of IEEE International Joint Conference on Neural Networks'2011, San Jose, USA, Jul. 31-Aug. 5, 2011.

BMI Press ISBN 978-0-9858757-0-1 US$ 35