

คอสต์ฟังก์ชันชนิดใหม่ที่ขึ้นกับพฤติกรรมสำหรับการสร้างแมนิโฟลด์

นางสาวกชกร ณ นครพนม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A NEW BEHAVIOR-BASED COST FUNCTION FOR MANIFOLD
CONSTRUCTION

Miss Kodchakorn Na Nakornphanom

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Computer Science

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2007

Copyright of Chulalongkorn University

501900

กชกร ณ นครพนม: คอสต์ฟังก์ชันชนิดใหม่ที่ขึ้นกับพฤติกรรมสำหรับการสร้างแมนิโฟลด์. (A NEW BEHAVIOR-BASED COST FUNCTION FOR MANIFOLD CONSTRUCTION) อ.ที่ปรึกษา: ศ. ดร. ชิดชนก เหลือสินทรัพย์, อ.ที่ปรึกษาร่วม: รศ. ดร.จักร์ อัครวานันท์, 68 หน้า.

ในการเรียนรู้ของอัลกอริทึมแบบแพร่ย้อนกลับ ค่าคลาดเคลื่อนกำลังสองถูกนำมาใช้เป็นคอสต์ฟังก์ชันสำหรับการปรับค่าน้ำหนักเพื่อคำนวณหาค่าความผิดพลาด คอสต์ฟังก์ชันแบบนี้เพียงวัดค่าผลต่างกำลังสองระหว่างค่าเป้าหมายและค่าผลลัพธ์ ซึ่งผลต่างแท้จริงที่ได้อาจให้ค่าที่เป็นบวกหรือลบขึ้นอยู่กับค่าผลลัพธ์ว่าอยู่สูงหรือต่ำกว่าค่าเป้าหมาย ดังนั้นแมนิโฟลด์ที่ถูกสร้างขึ้นจากคอสต์ฟังก์ชันแบบนี้จึงไม่สามารถรักษาลักษณะที่แท้จริงจากแมนิโฟลด์ของค่าเป้าหมายสำหรับทุกๆบริเวณของข้อมูลได้ จากข้อเสียที่กล่าวมาข้างต้นทำให้อัตราเร็วในการเรียนรู้ต่ำหรือทำให้โครงข่ายประสาทเทียมไม่สามารถเรียนรู้ข้อมูลได้ตามค่าความผิดพลาดที่ยอมรับได้ ซึ่งจะทำให้ผลลัพธ์ที่ได้จากการทำนายขาดประสิทธิภาพ ดังนั้นเป้าหมายของงานวิจัยนี้จึงเสนอคอสต์ฟังก์ชันตัวใหม่ที่สามารถเรียนรู้พฤติกรรมของค่าเป้าหมายและค่าผลลัพธ์ที่ได้จากโครงข่ายให้ไปในทิศทางเดียวกัน โดยใช้หลักการของผลคูณจุดระหว่างค่าเป้าหมายและผลลัพธ์ และจะหยุดการฝึกสอนเมื่อผลคูณจุดเข้าใกล้ศูนย์ ในงานวิจัยนี้ได้ทำการทดสอบคอสต์ฟังก์ชันที่สร้างขึ้นกับชุดข้อมูลมาตรฐาน การวัดประสิทธิภาพของคอสต์ฟังก์ชันนี้ทำโดยนำไปหาค่าโคไซน์ระหว่างเวกเตอร์ของค่าเป้าหมายและเวกเตอร์ของค่าผลลัพธ์ ซึ่งให้ค่าความถูกต้องมากกว่า 90 % และยังช่วยลดจำนวนรอบที่ใช้ในการเรียนรู้อย่างน้อยสองเท่าอีกด้วย นอกจากนี้ยังทำการเปรียบเทียบผลลัพธ์ของวิธีการที่นำเสนอกับวิธีพื้นฐานด้วยการวัดค่าที่เทสต์

ภาควิชา คณิตศาสตร์
สาขาวิชา วิทยาการคอมพิวเตอร์
ปีการศึกษา 2550

ลายมือชื่อนิสิต..... *K. N. N. N.*
ลายมือชื่ออาจารย์ที่ปรึกษา..... *C. L.*
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม..... *Jack Asavant*

4673801023 : MAJOR COMPUTER SCIENCE

KEY WORD : ARTIFICIAL NEURAL NETWORK / COST FUNCTION / FUNCTION APPROXIMATION

KODCHAKORN NA NAKORNPHANOM: A NEW BEHAVIOR-BASED COST FUNCTION FOR MANIFOLD CONSTRUCTION. THESIS ADVISOR: PROF. CHIDCHANOK LURSINSAP, Ph.D., THESIS COADVISOR: ASSOC. PROF. JACK ASAVANANT, Ph.D., 68 pp.

In backpropagation (BP) neural network learning algorithm, the mean squared error (MSE) has been used as a cost function. The cost function measures the squared difference between the target and the output. The real difference may be either positive or negative depending on whether the target is above or below the actual output. Obviously, the generated manifold obtained by the cost function does not preserve the real topology of the target manifold at any location. These disadvantages cause the slow convergent rate. Therefore, the results show an ineffective forecasting. The objective of this research is to propose a new behavior-based cost function to capture the parallel behavior between the target and the output. The method is based on the dot product between the target and the output vectors. If the dot product is sufficiently close to zero, then the training can be terminated. Therefore, the number of steps needed to train the network can be minimized. The proposed cost function is tested with the benchmarked test sets. The performance is evaluated by measuring the values of cosine between the actual target and the output. The accuracy of the proposed methodology is over 90%, and it can reduce the amount of the learning process at least two times. The difference of the proposed method and the standard methodology is also checked by the t-test value.

Department : Mathematics
Field of Study : Computer Science
Academic Year : 2007

Student's Signature : *K. N. Nakornphanom*
Advisor's Signature : *C. L.*
Co-advisor's Signature : *Jack Asavanant*

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Dr. Chidchanok Lursinsap at The Advanced Virtual and Intelligent Computing (AVIC) Center, for all his great support and patience that tremendously helped me to accomplish this thesis, and also my co-advisor, Associate Professor Dr. Jack Asavanant for guiding me through finishing this dissertation with untiring help. Their patience and support helped me overcome many crisis situations and finish the research.

I am deeply indebted to Professor Dr. Frank C. Lin for his abundantly helpful, and has assisted me in numerous ways.

I would also like to thank Associate Professor Suchada Siripant for taking good care of me and help in difficult times. And, I express my special thanks to the committee, Dr. Chularat Tanprasert for correcting the manuscripts and sharing the comments and suggestions, which helped to make further improvements.

I wish to thank my best friend in primary school (Angkhanang Atichati), my best friend in high school (Nuntamon Werasopon), my best friend in a graduate school (Supaporn Sidhisoradej), and my loving brother (Jetrawin Chitsumran) for helping me get through the difficult times, and for all the emotional support, camaraderie, entertainment and caring they provided. I would also like to thank the members of AVIC, especially, Orawan Tinnungwattana, Prem Junsawang, Piyamas Kanuan and most importantly my Ph.D friends (Sittisak Sai-ngern, Maytee Bamrungrajhirun, Kamol Keatruangkamala, and Yuttana Lila) for great support, helping, patience, and warm environment in which to learn and grow.

I appreciate the financial support from the National Electronics and Computer Technology Center (NECTEC) for the great support in my Ph.D. program, the National Research Council of Thailand under the project "Tsunami Detection and Early Warning by Remote Sensing", and also Dr. Mansour Ioualalen for The Franco-Thai Cooperation Program in Higher Education and Research.

May I devote this work to all the people as I mentioned above. Without them, this work will never be done. Through out this thesis, I had encountered many problems, but they were so little, compared to all the supports given by these people. The encouragement from them was so great, and those impressions will always be in my memory.

Lastly, and most importantly, I wish to thank my family, Mr.Paiboon and Ms. Pongsiri Na Nakornphanom, my parents, and Miss Pornsajee Na Nakornphanom, my sister, for providing a loving environment for me. They raised me, supported me, taught me, and love me always. To them I dedicate this thesis.

Contents

	Page
Abstract (Thai)	iv
Abstract (English)	v
Acknowledgements	vi
Contents	vii
List of Tables	ix
List of Figures	x
CHAPTER	
1 INTRODUCTION	1
1.1 Statement of Problems	2
1.2 Literature Review of Literature Related to Cost Functions	3
2 THEORETICAL BACKGROUND	6
2.1 Feedforward Multilayer Perceptron	7
2.2 Learning Algorithm of neural network	9
2.2.1 Backpropagation Neural Network	9
2.2.2 Levenberg-Marquardt Training Method	11
2.3 Activation Function	13
2.4 Cost Function	16
2.4.1 Mean Squared Error	16
2.4.2 Relative Entropy	16
2.5 Neural Cellular Automata	17
3 PROPOSED SOLUTIONS	18

CHAPTER	Page
3.1 Objectives	18
3.2 Scope of Work and Limitations	18
3.3 Network Architecture	18
3.4 A New Behavior-Based Cost Function	19
3.4.1 Learning Procedure	23
3.5 Shifting the result	28
3.6 Using Neural Cellular Automata To Simulate The Color Diffusion	30
4 EXPERIMENTAL RESULTS	34
4.1 Result of Proposing A New Behavior-Based Cost Function	34
4.1.1 Approximation problem	34
4.1.2 Classification problem	41
4.2 Shifting the Result	44
4.3 Result of Using Neural Cellular Automata To Simulate The Color Diffusion	47
4.4 Result of Using Neural Cellular Automata To Simulate The Tsunami Behavior	50
5 CONCLUSION	52
5.1 Result discussion	52
5.2 Conclusion	54
5.3 Further work	54
References	55
Vitae	58

List of Tables

Table		Page
4.1	The properties of the experimental data.	35
4.2	The comparison of the results in the approximation problem (number of epochs and cosine values) obtained from our paper and standard LM method.	36
4.3	The comparison of the results of the classification problem (number of hidden units, number of epochs and cosine value) obtained from our method and from normal LM method.	44
4.4	The comparison of the results in testset 1 (number of epochs and cosine values) obtained from our method and standard LM method.	49

List of Figures

Figure	Page
2.1 Two interconnected biological cells of neural network.	6
2.2 Feedforward three-layer perceptron architecture.	8
2.3 Three different types of activation Functions (a)Threshold function, (b)Piecewise-Linear Function and (c)Logistic Activation Functions.	14
2.4 The block-scheme of a general neural cellular automata iteration.	17
3.1 The structure of the neural network.	19
3.2 The acceptable error criterion (a) the cost function used in the backpropagation neural network.	20
3.3 This figure shows the idea of creating a new behavior-based cost function.	21
3.4 The concept of selecting the properly activation function.	23
3.5 The comparison between standard hyperbolic tangent function and the hyperbolic tangent function with $\alpha = 2.5$	24
3.6 The figure (a) shows the idea of setting the α . (b) The first criterion.	24
3.7 The figures show the idea of setting the α in (a) the second criterion. (b) The third criterion.	25
3.8 This figures show the idea of setting the α in (a) the fourth criterion. (b) The fifth criterion.	26
3.9 This figure shows the idea of setting the α in the general case. . .	27
3.10 This figure shows the idea of shifting the result.	28
3.11 This figures show the idea of shifting the result. (a) Result from the training patterns (b) Result after shift the training patterns (c) Result from the testing patterns (d) Result after shift the testing patterns.	29

Figure	Page
3.12 The structure of our proposed 2D tightly coupled cellular neural network.	32
4.1 This figure shows the results from the <i>sinc</i> testing set.	36
4.2 The comparison between target and output from the proposed cost function of testing set 2, two dimensional sinc: (a) The target image (b)The output image.	37
4.3 The comparison between target and output of testing set 2, two dimensional sinc. The top graph represents the target and bottom graph represent the output.	38
4.4 Result from MG_{17} testing set.	38
4.5 Result from MG_{30} testing set.	39
4.6 Result from lorenz testing set.	39
4.7 Result from titanium testing set.	40
4.8 Result from sunspot testing set.	40
4.9 This figure shows the concept of input pattern sequence in <i>xor</i> testing set.	41
4.10 This figure shows the concept of input pattern sequence in the general case.	42
4.11 Result from xor testing set	43
4.12 Result from iris testing set.	43
4.13 Shifting the result of Lorenz testing set with the <i>shift</i> value = 0.4827.	44
4.14 Shifting the result of Titanium testing set with the <i>shift</i> value = 1.347.	45
4.15 Shifting the result of MG17 testing set with the <i>shift</i> value = 0.9997.	45
4.16 Shifting the result of MG30 testing set with the <i>shift</i> value = -0.4891.	46

Figure	Page
4.17 Shifting the result of Sunspot testing set with the <i>shift</i> value = -0.1168.	46
4.18 The actual snapshot images taken at different times from the experiments. The area inside the rectangle is the training area and the area outside is the testing area. (a) At time 2. (b) At time 10. (c) At time 20. (d) At time 30.	47
4.19 The predicted images 3×3 compared with the actual image in Figure 4.18. (a) At time 2. (b) At time 10. (c) At time 20. (d) At time 30.	48
4.20 The predicted images of 5×5 compared with the actual image in Figure 4.18. (a) At time 2. (b) At time 10. (c) At time 20. (d) At time 30.	48
4.21 The value of <i>cosine</i> obtained from the neural network by the period of predictable time.	48
4.22 The actual snapshot images taken at different times from the test set 2. The area inside the rectangle is the training area and the area outside is the testing area. (a) At time 50. (b) At time 100. (c) At time 150. (d) At time 200.	49
4.23 The predicted images compared with the actual image in Figure 5. (a) At time 50. (b) At time 100. (c) At time 150. (d) At time 200.	49
4.24 The actual snapshot images taken at different times from the experiments (a) At time 241. (b) At time 341. (c) At time 441. (d) At time 600.	50
4.25 The predicted images 3×3 compared with the actual image in Figure (a) At time 241. (b) At time 341. (c) At time 441. (d) At time 600.	51
5.1 The weight adjustment comparison. (a) the standard LM training step and (b) the proposed training step.	53