一個具有加速因子的快速贏者取走類神經網路
A Fast Winner-take-all Neural Network with an Accelerated Factor

陳奇銘
Chen Chi-Ming

淡水工商管理學院資訊管理技術系
Department of Information Management Technology
Tamsui Oxford University College
ming0310@ms13.hinet.net

摘要

本論文，提出增加一個加速因子的觀念來改進以平均值的贏者取走類神經網路[16]，利用此方法可以提高收斂的速度。改進的贏者取走類神經網路主要將抑制的門檻由平均值提昇到更高的階層。

關鍵字: 加速因子，贏者取走類神經網路

ABSTRACT

In this paper, we extend the concept of the general mean-based winner-take-all network (GEMNET) proposed in [16] by introducing an accelerated factor to further improve its convergence speed. The improvement of the WTA network is based on the dynamic increase of the threshold which is attained by a higher order statistics than the mean in mutual inhibition.

keyword: accelerated factor, winner-take-all neural network

1. INTRODUCTION

In the *W*inner-*T*ake-*A*ll (WTA) process, the neuron initially most activated will gradually dominate and become maximally activated while the other competitive neurons die out. In many well-known neural networks [1-7], the WTA process is required to select the neuron which has the maximum activation or best correspondence to the input during learning or processing. Many continuous networks [7-10] have been introduced with asynchronous competitive behaviors originally. As to synchronous iterative WTA networks [11-17], one-layer feedback networks with mutual inhibitions can be effectively achieved. The MAXNET proposed in [11] is a famous WTA network which adopts heavy lateral inhibition. Recently, the GEMNET [16] with dynamic mutual inhibitions has been developed under the concept of the statistical mean to achieve fast convergence. The improved GEMNET [17] suggested a method to enhance the convergence speed of the GEMNET under the assumption of known-distribution of inputs. However, the convergence of the improved GEMNET is not assured if the inputs are not fallen in the class of the designated distribution.

In this paper, we propose a higher order statistics neural network (HOSNET) to further improve the convergence speed of mutual inhibited WTA networks. In Section 2, the WTA behavior of the GEMNET is briefly described. The HOSNET with an accelerated factor to improve the convergence of the GEMNET is addressed in detail in Section 3. The optimal accelerated factor for achieving the fastest convergence is also suggested in this section. In Section 4, simulation results of the HOSNET compared to those of the GEMNET are discussed.

2. MEAN-BASED WTA NEURAL NETWORKS

Consider M competitors, whose initial activations are X_1, X_2, ..., and X_M, assumed in the range of $[X_{min}, X_{max}]$ where X_{min} and X_{max} represent the minimum and the maximum bounds of all possible inputs, respectively. If the inputs are arranged in ascending order of magnitude, which satisfy

$$X_{<1>} \leq X_{<2>} \leq \ldots \leq X_{<M>}, \qquad (1)$$

where $X_{<m>}$ is called the m^{th} activation and <m> carries with the original index of the corresponding input. Thus, the maximum activation is denoted by $X_{<M>}$, and the second maximum activation is denoted by $X_{<M-1>}$ where <M> and <M-1> represent the original indices of the maximum and the second maximum, respectively. The MAXNET [11] and the GEMNET [16] share the same one-layer competitive architecture to resolve the WTA problem. However, the GEMNET with built-in dynamics outperforms the MAXNET, which only uses fixed mutual inhibitions.

The general mean-based WTA neural network (GEMNET) shown in Reference [16] is a feedback one-layer neural network developed under the concept that the maximum is always greater than the mean of any subset of activations. The connection weight between Node i and Node j in the GEMNET is given by

$$W_{ij}(k) = \begin{cases} \gamma & i = j \\ \dfrac{\gamma}{M_G(k) - 1} & i \neq j \quad 1 \leq i,j \leq M, \end{cases} \qquad (2)$$

where γ is the compensation factor to approximately maintain the maximum activation constant [16]. In summary, the GEMNET physically performs the thresholding to the mean of active activations for achieving the mutual inhibition WTA process.

3. HIGHER ORDER STATISTICS WTA PROCESS

In the GEMNET, the WTA process inhibits the neurons, whose activations are less than the mean of active activations. Theoretically, the mean-thresholding approach, which inhibits about one half of current active neurons, is too conservative to achieve the fast WTA process for a large number of competitors. If we dynamically increase the level of inhibition, the convergence speed of the GEMNET will be further improved.

By using an accelerated factor, the GEMNET can be extended to a higher order statistics neural network (HOSNET) for achieving a faster WTA process. The connection weight between Node i and Node j in the HOSNET is suggested as

$$W_{ij}(k) = \begin{cases} \gamma & i = j \\ \dfrac{\gamma\beta(k)}{M_H(k) - \beta(k)}, & i \neq j \quad 1 \leq i,j \leq M \end{cases} \qquad (3)$$

where $M_H(k)$ denotes the number of active neurons in the HOSNET and $\beta(k) \geq 1$ is the proposed accelerated factor. If $\beta(k) = 1$, the HOSNET is identical to the GEMNET. From the connection weights depicted in (2) and (3), we know that the HOSNET and GEMNET share the similar structure. The GEMNET performs the mean-thresholding mutual inhibition controlled by $M_G(k)$ [16]. However, the HOSNET uses $M_H(k)$, the effective number of active neurons instead of $M_G(k)$. Conceptually, $\mu_H(k)$ can be treated as an estimator of the higher order statistics from the obtained mean estimator, $\mu(k)$. Thus, the convergence behavior of the HOSNET should be faster than that of the GEMNET. In order to avoid the over-inhibition, the accelerated factor should be limited in certain bounds. How to determine an optimal accelerated factor and its upper and lower bounds becomes an important issue in this paper.

Now, we should find the optimal accelerated factor, which not only can prevent from the over-inhibition but also can complete the fastest WTA process. For the fastest convergence, the optimal thresholding for the WTA process should be given in

$$X_{<M-1>} \leq \mu_H(k) < X_{<M>}, \qquad (4)$$

which can inhibit all the neurons except the maximum one and avoid the over-inhibition in the HOSNET. For achieving the fastest convergence and avoiding the possible over-estimation, therefore, $\mu_H(k) = X_{<M-1>}$ is suggested in this paper. Actually, the optimal thresholding, $\mu_H(k)$, can be treated as an adaptive estimator for estimating the second maximum activation in each iteration. Thus, the optimal accelerated factor should be chosen to satisfy

$$\beta_{opt}(k) = \frac{X_{<M-1>}}{\mu(k)}. \qquad (5)$$

Since the threshold $\mu_H(k)$ should be less than $X_{<M>}$ to avoid the over-inhibition, the bounds of the accelerated factor is limited in

$$1 \leq \beta_{bound}(k) < \frac{X_{<M>}}{\mu(k)}. \qquad (6)$$

Unfortunately, the exact values of $X_{<M>}$ and $X_{<M-1>}$ are not available for real applications. The accelerated factor in the HOSNET should be designed in the average sense.

From order statistics [18], the dynamic optimal accelerated factor for M uniformly-distributed activations can be expressed by

$$\beta_{opt}(k) = \frac{2M(k) - 2}{M(k) + 1} \quad \text{for } M(k) > 2 \qquad (7a)$$

and

$$\beta_{opt}(k) = 1 \quad \text{for } M(k) \leq 2. \qquad (7b)$$

If the activations are uniformly distributed in any fixed range, we can also find the same conclusion. With the above selected sequence of accelerated factors, each HOSNET process intends to complete the WTA process in one iteration. If the accelerated factor is fixed as $\beta(k) = \beta_o$ for simplifying the implementation, β_o should be less than $\dfrac{2M}{M+1}$.

4. Convergence Analysis

In Reference [16], we know that the GEMNET requires

$$K = \mathrm{Log}_2(M) \qquad (8)$$

iterations to achieve the convergence for uniformly distributed inputs. If the initial activations are assumed to be uniform distribution in [0,1], we can also analyze the convergence performance of the HOSNET.

If the accelerated factor in the HOSNET is optimally selected as (7), we can analyze the convergence speed of the HOSNET easily. The performance analysis is based on the statistical average of the first HOSNET iteration by the assumption that the HOSNET achieves the worst case in the second HOSNET iteration.

At the first iteration, the optimal selection of accelerated factor implies that the thresholding of the inhibition is exactly at $\frac{M-1}{M+1}$. The probability of the neuron being inhibited is $p=\frac{M-1}{M+1}$. On the contrary, each neuron only has probability of $q = \frac{2}{M+1}$ to become active. For M neurons, the number of active neurons after the first HOSNET process possesses M+1 cases which are $M_H(1) = i$ for $i = 0, 1, 2, .., M-1, M$. Based on Bernoulli's theorem, the probability that i neurons remain active, is equal to

$$P_i = C_i^M (p)^{M-i}(q)^i = \frac{M!}{i!(M-i)!}\left(\frac{M-1}{M+1}\right)^{M-i}\left(\frac{2}{M+1}\right)^i . \quad (9)$$

Now, we can discuss that the averaged number of iterations required for the HOSNET to achieve the complete WTA process for M-competitors .

For $M_H(1) = 0$ case, the HOSNET will be classified as the over-inhibition status. The HOSNET is automatically switched back to the GEMNET, which requires further $Log_2(M)$ iteration to achieve the complete WTA process. Therefore, the HOSNET will require $Log_2(M)+1$ iterations to achieve the convergence in this case.

For $M_H(1) = 1$ case, the HOSNET completes the WTA process. No more iteration is required. It is obvious that the HOSNET totally needs only one iteration to attain the convergence in this case.

For $M_H(1) = 2$ and 3 cases, it is easy to check from (16) and (17) that the optimal accelerated factor in the second iteration $\beta(2)$ will be one. The performance of the second HOSNET process is the same as that of the GEMNET process. Therefore, the convergence speed of the HOSNET totally needs $1+Log_2(2)$ or $1+Log_2(3)$ iterations for $M_H(1)=2$ and $M_H(1)=3$ cases, respectively.

For $M_H(1)= i \geq 4$ cases, we can assume that the second HOSNET process becomes the over inhibition condition, which is the worst case. The HOSNET requires further $1+Log_2(i)$ iterations for achieving the convergence. Thus, the convergence speed of the HOSNET totally needs $2+Log_2(i)$ iteration to complete the WTA process in these cases.

The upper bound of the total number of iterations for the HOSNET to achieve the convergence can be found as

$$K < 0.1353Log_2(M) + 2.1884 = K_{bound}. \quad (10)$$

In other words, if the number of iterations is limited in K_{bound}, the HOSNET can complete the WTA process.

Comparing (8) and (10), if the number of competitors, M is greater than 6, the HOSNET shows better convergence behavior than the GEMNET. Table 1 shows the theoretical averaged numbers of iterations required for the GEMNET and the HOSNET to complete a WTA process. We learnt that the HOSNET with the optimally accelerated factor is faster than the GEMNET in the case of uniformly distributed inputs.

Various numbers of inputs with uniform distribution in [0,1] are randomly generated as the competitors to evaluate the WTA behaviors of the HOSNET. By using the dynamic accelerated factor depicted in (7) and the constant accelerated factor fixed to 2(M-1)/(M+1), Table 2 shows the averaged number of iterations for 1000 WTA cases performed by the HOSNET in simulations. Comparing Tables 1 and 2, we found that the simulation results are close to the theoretical analyses stated in (10). Since the most neurons are inhibited in the first iteration, the dynamic optimal accelerated factor is close to but slightly better than the fixed one. As to the normal distribution in N(0,1), Table 3 also shows that the HOSNET achieves better convergence performance than the GEMNET in the case of a large number of competitors.

5. CONCLUSIONS

The HOSNET WTA neural network with an accelerated factor has been proposed to improve the convergence performance of the mean-thresholding WTA process. The optimally accelerated factor and convergence performance of the HOSNET are also suggested and discussed. Both theoretical analyses and simulation results show that the convergence speed of the HOSNET is higher than that of the GEMNET if the number of competitors is very large.

REFERENCES

[1] S. Grossberg, "Adaptive pattern classification and universal recording: I. Parallel development and coding of neural feature detectors", *Biological Cybernetics*, 23, p.121-134, 1976.

[2] G. Carpenter and S. Grossberg, "Adaptive resonance theory: Stable self-organization of neural recognition codes in response to arbitrary lists of input patterns", *Eighth Annual Conference of the Cognitive Science Society*, pp.45-62, Hillsdale, NJ: Lawrence Erlbaum Associates, 1986.

[3] B. Kosko, "Fuzzy associative memories", In A. Kandel (Ed.), *Fuzzy Expert System*, Reading, MA,

Addison-Wesley, 1987.

[4] T. Kohonen, "Automatic formation of topological maps in a self-organizing system", In E. Oja and O. Simula (Eds.) *Proceedings of the 2nd Scandinavian Conference on Image Analysis*, p214-220, 1981.

[5] T. Kohonen, *Self-organization and associative memory*, Springer-Verlag, New York, USA, 1989, 3rd Ed.

[6] S. Grossberg, "Nonlinear neural networks: Principles, mechanisms, and architectures", *Neural Networks*, Vol. 1, pp.17-61, 1988.

[7] S. Grossberg, "Contour enhancement, short term memory, and constancies in reverberating neural networks", *Studies in Applied Mathematics*. LII. 213-257, 1973.

[8] A. L. Yuille, N. M. Grzywacz, "A winner-take-all mechanism based on presynaptic inhibition feedback", *Neural Computation* 1, pp. 334-337, 1989.

[9] R. Coultrip, R. Granger, and G. Lynch, "A cortical model of winner-take-all competition via lateral inhibition", *Neural Networks*, Vol. 5, pp. 47-54, 1992.

[10] J. A. Feldman and D. H. Ballard, "Connectionist models and their properties", *Cognitive Science*, 6, pp.205-254, 1982.

[11] R. P. Lippmann, "An introduction to computing with neural nets", *IEEE ASSP Magazine*, April 1987.

[12] H. K. Kwan, "One-layer feedforward neural network fast maximum/minimum determination", Electronics Letters, pp. 1583-1585 1992.

[13] J. C. Yen and S. Chang, "Improved winner-take-all neural network", Electronics Letters pp662-664 March 1992.

[14] B. W. Suter and M. Kabrsisky, "On a magnitude preserving iterative Maxnet algorithm", *Neural Computation 4*, pp. 224-233, 1992.

[15] J. C. Yen, F. J. Chang, and S. Chang, "A New Winners-tak-all Architecture in Artificial Neural Networks", IEEE Trans. on Neural Networks, Vol.5, No.5, pp. 838-843 Sep. 1994.

[16] J. F. Yang, C. M. Chen, W. C. Wang, and J. Y. Lee, "A General mean based iteration winner-take-all neural network", *IEEE Tran. on Neural Networks*, Vol 6, No.1, pp. 14-24, Jan. 1995.

[17] J. F. Yang and C. M. Chen, "An improved General Mean Based Iterative Winner-Take-All Neural Network", 1994 International Symposium on Artificial Neural Network, pp. 429-434.

[18] H. A. David, *Order Statistics*, Second Edition, A Wiley Publication, 1980.

Number of Competitors	10	50	100	500	1000	2000	3000	4000	5000
GEMNET	3.322	5.418	6.421	8.744	9.966	10.97	11.55	11.97	12.29
HOSNET	2.638	2.952	3.106	3.401	3.537	3.672	3.751	3.808	3.851

Table 1. The theoretically averaged number of iterations required in the GEMNET and the HOSNET for completion of a WTA process

Number of Competitors	10	50	100	500	1000	2000	3000	4000	5000
Simulated Results (constant β)	2.973	3.179	3.302	3.511	3.673	3.802	3.885	3.945	3.958
Simulated Results (dynamic β(k))	2.484	2.802	2.953	3.362	3.416	3.511	3.523	3.657	3.676

Table 2. The averaged number of iterations required for the completion of WTA process by the HOSNET with the constant and dynamic optimal accelerated factors

Number of Competitors	10	50	100	500	1000	2000	3000	4000	5000
GEMNET	2.481	4.336	5.048	6.762	6.901	7.605	8.043	8.428	9.194
HOSNET (β=3)	1.304	3.864	4.092	3.692	4.407	4.836	4.915	4.825	4.694
HOSNET (β=7)	3.294	4.758	4.818	4.002	3.326	3.796	4.270	4.654	4.930

Table 3. The averaged number of iterations required in the GEMNET and the HOSNET with various cases of constant accelerated factors for completion of a WTA process in N(0,1).