2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

DT-CNN Annealing with Additive Noise Generated by Class 3 CA: A Comparison with Chaos Annealing

Tomohiro FUJITA† and Takeshi OGURA†

†Ritsumeikan University 1-1-1 Noji-higashi, Kusatsu, Shiga, Japan
Email: {tfujita, togura}@se.ritsumei.ac.jp

Abstract—An annealing for global optimization on discrete time cellular neural network (DT-CNN) with additive noise is investigated. In our method the additive noise is generated using chaotic behavior of class 3 cellular automata (CA). The optimization with our method shows higher performance than with uniform random noise. In this paper we investigate the relation between our method and chaos annealing, which is another superior optimization method with neural network. The auto-correlations of the CA noise and logistic map, which is often used in chaos annealing, are evaluated, and we can see the similarity of both systems. The annealing performances are also evaluated, and the results with the CA noise and logistic map show both methods exhibit higher performances under the same condition. These results may suggest that the CA noise and logistic map have common basis and it makes the ability of global search high.

1. Introduction

Combinatorial optimization problems arise in a lot of scientific and technological fields. In this paper we investigate a global optimization with discrete time cellular neural network (DT-CNN).

Many researches apply the optimization on DT-CNN to various problems to date. One of examples is solving minimization problem of spin glass energy by DT-CNN [2]. In statistical physics, with the Ising model [3] the minimization problem of spin glass energy is formulated as a quadratic assignment problem which is an NP-hard problem. This formulation has the same structure of Lyapunov function of DT-CNN, so this can be solved by DT-CNN. In recent years the Ising model is applied in probabilistic information processing [4], and in some applications this method achieves great success. Therefore solving this class of problems is paid much attention from the image processing field. Another example is an image coding and decoding [5]. The image coding and decoding are formulated as an optimization problem and are solved by DT-CNN.

The drawback of the optimization with DT-CNN is that in many cases the state of the network is trapped at a local minimum, and thus a global solution can not be found. To overcome this difficulty we proposed an annealing method on a DT-CNN that realizes global optimization [6, 7]. In this scheme, noise is induced into network dynamics then gradually reduced. In this process, the state of the network is initially random but eventually becomes convergent. Due to the randomness of the noise, the network escapes from local minima.

In previous work [8] we proposed a hardware-oriented method of noise generation. The noise is generated using the chaotic behavior of class 3 cellular automata (CA) [1] on Cellular AutoMata on Content Addressable Memory (CAM^2). CAM^2 is a dedicated hardware for CA and CNN [9, 10], so that the noise can be generated easily. We also showed the annealing performance with the noise generated by chaotic behavior of CA (CA noise) is superior to that with uniform random noise [7].

Preceding studies reported that the Hopfield Neural Network with the chaotic noise as the additive noise increases the optimization ability [11, 12]. It was pointed out that this ability relates to the auto-correlation of the chaotic noise [13, 14]. Hasegawa and Umeno showed the noise with auto-correlation which has negative auto-correlation at first data lag and gradually decays has high solving ability of minimization [14]. Our system also uses the chaos of class 3 CA, and so the CA noise may have the common property which the chaos noise has.

In this paper we take up logistic map as a representation of chaos, and investigate whether both our method and logistic map have common statistical property. The auto-correlation of the CA noise is compared with that of logistic map. The performance of annealing is also compared CA noise and logistic map.

2. Noise Induced DT-CNN Model

DT-CNN is a temporally discretized CNN. The DT-CNN consists of an $M \times N$ rectangular array of cells $C(i, j)$, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$. These cells have three variables $u_{i,j}$, $x_{i,j}$ and $y_{i,j}$, denoting input, state and output, respectively. The dynamics of the DT-CNN takes the following form:

$$x_{i,j}(t+1) = \sum_{C(k,l)\in N_r(i,j)} A(i, j; k, l)\, y_{i,j}(t)$$

$$+ \sum_{C(k,l)\in N_r(i,j)} B(i,j;k,l)\, u_{i,j} + I \quad (1)$$

$$y_{i,j}(t) = \frac{1}{2}\left(|x_{i,j}(t)+1| - |x_{i,j}(t)-1|\right) \quad (2)$$

where $N_r(i,j)$ is a set of neighborhood cells and $A(i,j;k,l)$ and $B(i,j;k,l)$ are parameters called templates. I is also a parameter called the threshold value.

We can define the Lyapunov function $E(t)$ of the DT-CNN as follows [15].

$$\begin{aligned} E(t) = & \frac{1}{2}\sum_{(i,j)} y_{i,j}^2(t) \\ & -\frac{1}{2}\sum_{(i,j)}\sum_{(k.l)} A(i,j;k,l)\, y_{i,j}(t)\, y_{k,l}(t) \\ & -\sum_{(i,j)}\sum_{(k.l)} B(i,j;k,l)\, y_{i,j}(t)\, u_{k,l} \\ & -\sum_{(i,j)} I\, y_{i,j}(t) \end{aligned} \quad (3)$$

State $x_{i,j}$ changes as the Lyapunov function decreases and converges at one of the local minima. In order to obtain the global minimum, hardware annealing on the DT-CNN is performed using a DT-CNN with noise. The dynamics of the DT-CNN with noise is defined as the following form:

$$\begin{aligned} x_{i,j}(t+1) = & \sum_{C(k,l)\in N_r(i,j)} A(i,j;k,l)\, y_{i,j}(t) \\ & + \sum_{C(k,l)\in N_r(i,j)} B(i,j;k,l)\, u_{i,j} + I \\ & + a(t)\, n_{i,j}(t) \end{aligned} \quad (4)$$

$$a(t+1) = (1-\delta)\, a(t) \quad (0 < \delta < 1) \quad (5)$$

$n_{i,j}$ is the noise in cell $C(i,j)$. The range of $n_{i,j}$ is $[-1, 1)$. The amplitude of noise is controlled by $a(t)$, which decreases exponentially in accordance with Eq. (5). δ controls the speed of damping.

Noise is added to the dynamics of the DT-CNN to search for a global minimum. Since state $x_{i,j}$ fluctuates randomly in the presence of noise, escaping from local minima becomes possible. As the noise becomes smaller, the state of the DT-CNN becomes stable at an optimal minimum or suboptimal minima.

3. Noise Generator

To implement the DT-CNN with noise on hardware, a generating system of noise $n_{i,j}$ is required. We propose the use of two-dimensional CA as a noise generator [8]. CA are computational models proposed by Neumann [1] and consist of lattice-shaped cells. Since the architecture of CA is similar to that of the DT-CNN, we can implement CA on a universal CNN machine, CAM[2]. Figure 1 shows the concept of the DT-CNN with noise, which has two layers: a DT-CNN layer and a CA layer.

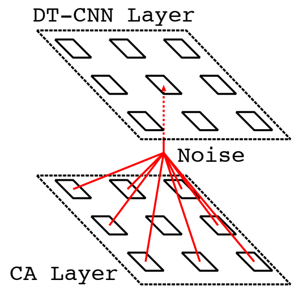

Figure 1: Hardware annealing on DT-CNN

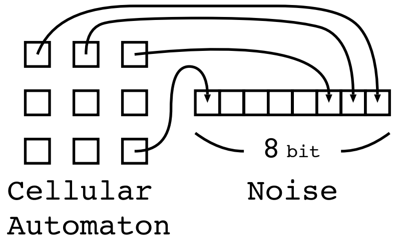

Figure 2: Noise generator

The states of the cells vary on the basis of the state transition function F, which is determined by the rule number R as follows:

$$R = \sum_{n=0}^{N}\sum_{v_{i,j}=0}^{1} F(v_{i,j}, n) \times 2^{v_{i,j}+2n} \quad (6)$$

where $v_{i,j}$ is the state of cell (i,j) and takes binary values, n is the number of neighborhood cells with state $v_{i,j}$ equal to 1 and N is the total numbers of neighborhood cells. Wolfram [1] sorted state transition functions into four classes. The functions in class 3 have chaotic behavior. Since we require disordered noise, we use CA in class 3 as the noise generator. In this paper, $R = 143954$.

We generate noise $n_{i,j}$ from the CA as follows.

$$\begin{aligned} n_{i,j}(t) = & \left(\frac{v_{i-1,j-1}}{2^8} + \frac{v_{i-1,j}}{2^7} + \frac{v_{i-1,j+1}}{2^6} + \frac{v_{i,j-1}}{2^5}\right. \\ & + \left.\frac{v_{i,j+1}}{2^4} + \frac{v_{i+1,j-1}}{2^3} + \frac{v_{i+1,j}}{2^2} + \frac{v_{i+1,j+1}}{2^1}\right) \times 2 \\ & - 1 \end{aligned} \quad (7)$$

The range of $n_{i,j}$ is $[-1, 1)$. Figure 2 shows the process by which the noise is generated.

4. Auto-correlation of CA noise

Hasegawa et al. [13] analyzed which characteristic of the chaos affects on the annealing performance using surrogates. From their experimental results, the noise which

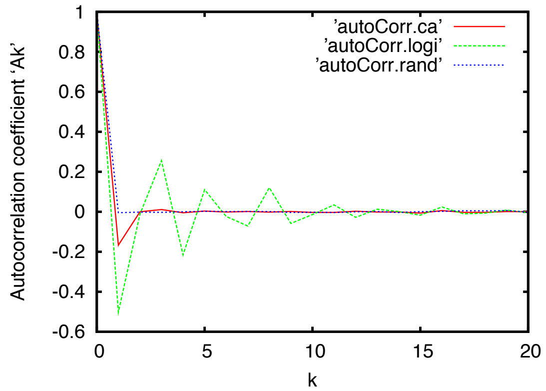

Figure 3: Autocorrelation

preserves the auto-correlation of chaotic noise exhibits high solving ability. This shows that temporal structure of the chaos affects on the annealing performance. We analyze the auto-correlation of CA noise and investigate whether CA noise has the same temporal structure as logistic map.

The auto-correlation of time-series $\{x_0, x_1, \cdots\}$ is defined as $A_k = E[(x_0-\mu)(x_k-\mu)]$, where k is a time-shift from original data. Figure 3 shows the auto-correlation of CA noise. In Fig. 3 the auto-correlations of logistic map and uniform random number are also depicted. The auto-correlation of uniform random number is 0 at $k \neq 0$, because there is no correlation among the time series of the random number. On the other hand, in the logistic map, the correlation coefficients at $k = 1$ is negative and it gradually decays. Hasegawa and Umeno pointed out that this characteristic shape of the auto-correlation makes the solving ability of optimization high [14].

The auto-correlation of CA noise also takes negative auto-correlation coefficient at $k = 1$. That feature is similar with logistic map, but the auto-correlation decays so fast that the coefficients is nearly equal to zero at $k \geq 2$.

5. Annealing with CA noise and logistic map

We use three types of noise sources, namely, CA noise, logistic map and uniform random number, as the noise in Eq. (4). In this experiment we use the following templates.

$$A = \begin{bmatrix} 0.0 & 1.0 & 0.0 \\ 1.0 & 3.0 & 1.0 \\ 0.0 & 1.0 & 0.0 \end{bmatrix}, B = \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 \end{bmatrix}, I = 0 \tag{8}$$

The input variable $\boldsymbol{u}$ is randomly generated within $[-U, U]$ to create various problems, where U is the range of random numbers. We apply our model (Eq.(4)) on DT-CNN with cell size of 100×100.

Figure 4 is the result of the 100 trials with various initial conditions, and in this case U is equal to 1. The x axis is the serial number for the trials and the y axis is the convergent energy of the network. In Fig. 4 the green, red and blue lines show the result of the annealing with the CA noise, logistic map and uniform random number, respectively. This results show in CA noise and logistic map the 100 trials converge to one solution and these convergent energies with two methods are exactly same. We also tried to solve the same problem in GA, but we could not find better solution than this. Comparing with the uniform random number, CA noise and logistic map find the better solution in all 100 trials.

We change the parameter U into 4, and the experimental results by this parameter show in Fig. 5. In Fig. 5, the results using the CA noise and logistic map do not show the convergency into one solution. In all 100 trials, the convergent energies with the CA noise and logistic map are worse than that with the random noise.

Randomly generating the input variable $\boldsymbol{u}$ within the range of $[-U, U]$, we also try the same evaluation on the 100 various problems. When U is equal to 1, in the all problems, the annealing results with the CA noise and logistic map show the convergence to one solution, and they obtain better solution than the random number. On the other hand, when $U = 4$, in all problems, the convergence to one solution does not reveal, and the random number obtains better solutions than the CA noise and logistic map.

We summarize the above results. When U is equal to 1, the annealing with both the CA noise and logistic map shows higher performances than with the uniform random number. In the case with the CA noise and logistic map, the network converges to exact one solution in all trials from 100 initial conditions. As U changes into 4, the annealing performances with the CA noise and logistic map degrade. Their performances are worse than the random number.

The results with the CA noise and logistic map show the same tendency in annealing performance. That may suggest that they have the common basis and it makes solving ability higher.

6. Conclusion

From our analyses the CA noise has some common natures with logistic map. Firstly, the auto-correlation coefficient of the CA noise takes negative value at the first time lag. Logistic map also has this shape of the auto-correlation and the noise which takes this shape of auto-correlation has high ability of global search. Secondly, both the CA noise and logistic map reveal high performances of DT-CNN annealing under the same condition. This may suggest the common principle of both systems makes the ability of global search high.

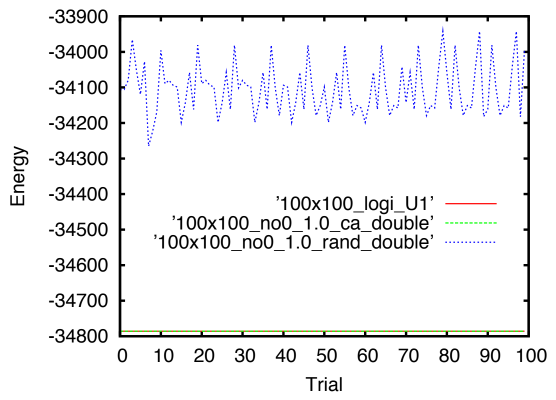

Figure 4: Convergent energy in the case of $U = 1$ (Cell size $= 100 \times 100$)

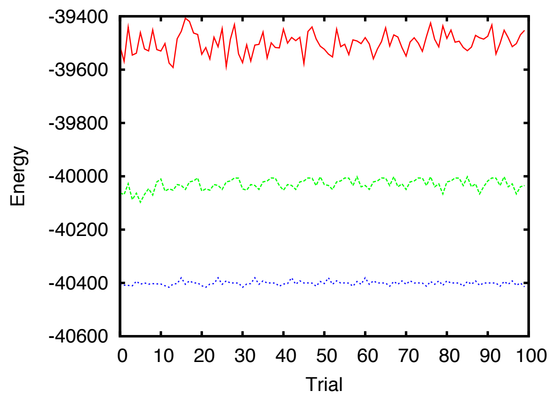

Figure 5: Convergent energy in the case of $U = 4$ (Cell size $= 100 \times 100$)

References

[1] S. Wolfram, *Cellular automata and complexity : collected papers*. Westview Press, 1994.

[2] M. Ercsey-Ravasz, T. Roska, and Z. Neda, "Cellular neural networks for NP-hard optimization," in *11th International Workshop on Cellular Neural Networks and Their Applications*, 2008, pp. 52–56.

[3] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford University Press, 2001.

[4] K. Tanaka, "Statistical-mechanical approach to image processing (topical review)," *J. Phys. A: Math. & Gen.*, vol. 35, no. 37, pp. R81–R150, 2002.

[5] M. Tanaka, K. R. Crounse, and T. Roska, "Parallel analog image coding and decoding by using cellular neural networks," *IEICE Trans. Fundamentals*, vol. E77-A, no. 8, pp. 1387–1395, 1994.

[6] T. Fujita, K. Sakomizu, and T. Ogura, "Hardware annealing on DT-CNN using CAM^2," in *Cellular Nanoscale Networks and Their Applications (CNNA), 2010 12th International Workshop on*, Feb. 2010.

[7] T. Fujita and T. Ogura, "Performance evaluation of DT-CNN annealing with CA noise," in *Proceedings of the 2010 International Workshop on Nonlinear Circuits, Communication and Signal Processing*, Mar. 2010, pp. 592–595.

[8] K. Sakomizu, T. Fujita, and T. Ogura, "Hardware annealing on discrete-time cellular neural network for object extraction," *J. Signal Processing*, vol. 13, no. 6, pp. 463–468, 2009.

[9] T. Ikenaga and T. Ogura, "CAM^2: A highly-parallel two-dimensional cellular automaton architecture," *IEEE Trans. Comput.*, vol. 47, no. 7, pp. 788–801, 1998.

[10] T. Fujita, T. Okamura, M. Nakanishi, and T. Ogura, "CAM^2-universal machine: A DTCNN implementation for real-time image processing," in *Cellular Neural Networks and Their Applications, 2008. CNNA 2008. 11th International Workshop on*, July 2008, pp. 219–223.

[11] Y. He, "Chaotic simulated annealing with decaying chaotic noise," *IEEE Trans. Neural Networks*, vol. 13, no. 6, pp. 1526–1531, Nov. 2002.

[12] L. Wang and K. Smith, "On chaotic simulated annealing," *IEEE Trans. Neural Networks*, vol. 9, no. 4, pp. 716–718, 1998.

[13] M. Hasegawa, T. Ikeguchi, T. Matozaki, and K. Aihara, "An analysis on additive effects of nonlinear dynamics for combinatorial optimization," *IEICE Trans. Fundamentals*, vol. E80-A, no. 1, pp. 206–213, 1997.

[14] M. Hasegawa and K. Umeno, "Solvable performances of optimization neural networks with chaotic noise and stochastic noise with negative autocorrelation," in *ICONIP (1)*, ser. Lecture Notes in Computer Science, M. Ishikawa, K. Doya, H. Miyamoto, and T. Yamakawa, Eds., vol. 4984. Springer, 2007, pp. 693–702.

[15] L. Chua and L. Yang, "Cellular neural networks: theory," *IEEE Trans. Circuits Syst.*, vol. 37, no. 12, pp. 1257–1272, Oct. 1988.