Journal of Information Hiding and Multimedia Signal Processing ©2014 ISSN 2073-4212
Ubiquitous International Volume **5**, Number **2**, April **2014**

Analysis and Elimination of Digital Chaotic Key Sequence's Autocorrelation

Baoxiang Du, Qun Ding and Xiaoli Geng

Key Laboratory of Electronics Engineering
Heilongjiang University,Key Laboratory of Electronics Engineering
74 Xuefu Road, Harbin, Heilongjiang, China
dubaoxiang@sina.com, qunding@aliyun.com, 361195315@qq.com

Received September, 2013; revised March, 2014

Abstract. *The special properties of the chaos make the chaotic encryption technology to be an important research field of information science and technology. However, because digital chaos is affected by finite precision of computer, chaos system properties present degradation—short periodicity. This paper firstly analyses reason that causes strong autocorrelation of digital chaotic sequence, then uses statistical method to deal with correlation of the sequence, and last chooses K-L transform method for eliminating correlation through comparing different methods. Simulation results prove that this method can eliminate the autocorrelation of logistic sequence, and its autocorrelation figure is similar to white Gaussian noise autocorrelation's. Meanwhile, it also can increase the period of logistic sequence, which makes up short periodicity of digital chaotic sequence and makes digital chaotic key sequences can be applied safely to encryption system.*
Keywords: digital chaos, K-L transform, phase space reconstruction

1. **Introduction.** Chaos phenomena was first found by meteorologist Lorenz as a random phenomena in the 1960s[1], then digital chaotic sequence that was generated by nonlinear difference equation was applied in physics, cryptography, information security and so on . Meanwhile, a large number of scholars started studying the utilizations of chaos. Chaos system is a complex nonlinear dynamic system. Its good Pseudo random property, orbital unpredictability, extremely sensitive characteristics for initial state and the control parameters make chaotic secure communication become an important research subject in the field of secure communication in recent years. With the understanding of chaotic system, more and more researchers began to design digital encryption algorithm based on chaotic system[2]-[5]. The initial stage of the digital chaotic encryption system is in 1989 to 1993. There are a lot of papers about digital chaotic encryption system during this period[6]-[10]. However, through analyzing discrete chaotic system, people found that security flaws of the digital chaotic encryption system, that is the degradation of the digital chaotic encryption system, some progress of digital chaotic encryption field was not obvious during those years. Although the ideal chaotic systems are not repetitive infinitely, random, sensitive and so on, chaotic system in digital encryption algorithm may not be true definition of chaos mapping on continuous domain. Because of all algorithms are operated in limited precision, some chaotic sequence that is generated by iteration is not chaotic system orbit theoretically, and incompletely has chaotic characteristics. Moreover based on the result of approximation, the chaotic sequence will be evolved into a sequence with cycle, which makes encryption algorithm be cracked based on chaos theory design.

But the initial value sensitivity and the huge initial key space of digital chaotic system still can't be substituted by traditional digital encryption algorithm. There are still many original basic characteristics of chaotic system. If we can improve influence of limited accuracy in digital chaotic system, digital chaotic encryption domain still has prospects of research and development. So people have been making a breakthrough, studying, and improving for the degradation of the digital chaotic system characteristics from 1997 to now, the papers about digital chaotic sequence encryption emerge in endlessly[11]-[19]. This paper first analyzes autocorrelation of chaotic time sequence, then researches the method of eliminating correlation for chaotic sequence with stronger autocorrelation, last proposes K-L transform. The experiment result shows that K-L transform method is better in eliminating correlation than the traditional singular value decomposition (SVD), discrete cosine transform (DCT) and so on. This method increases the period and complexity of the digital chaotic sequence, reduces the correlation between chaos sequences, and makes the chaos sequence present better randomness, so it can be applied greatly in the secure communication system.

2. **Analysis of sequence autocorrelation.** Mathematical expression of logistic mapping is followed as

$$x_{n+1} = \mu x_n(1 - x_n),\ \mu \in (0, 4),\ x_n \in (0, 1). \tag{1}$$

Due to the limitation of operation precision, when computer conducts iterative operation every time, mathematical expression needs to add a random error term. The random error term added every time is totally different. Assuming the i_{th} random error term added is μ_i, mathematical model is written as

$$x_{n+1} = \mu x_n(1 - x_n) + \mu_i,\ \mu \in (0, 4),\ x_n \in (0, 1),\ i = 1, 2, \cdots, n. \tag{2}$$

We can assume that random error term is mutual independent, and then it is expressed as

$$Cov(\mu_i, \mu_j) = 0,\ i \neq j,\ i, j = 1, 2, \cdots, n. \tag{3}$$

If

$$Cov(\mu_i, \mu_j) \neq 0,\ i \neq j,\ i, j = 1, 2, \cdots, n. \tag{4}$$

for the different sample points, the random error terms are totally no longer independent, which have some kinds of correlation. We can think that this sequence is correlative. Since random error terms, of which mean is 0, are normally distributed, the correlation of sequence is expressed as

$$E(\mu_i, \mu_j) = 0,\ i \neq j,\ i, j = 1, 2, \cdots, n.$$

If

$$E(\mu_i, \mu_j) \neq 0,\ i \neq j,\ i, j = 1, 2, \cdots, n,$$

we believe where the sequence is first-order correlative, or autocorrelative, so the correlation of Logistic sequence is autocorrelation.

Why does correlation of sequence appear in discrete chaotic sequence? From observing (2) above, due to the rounding-off reasons, we may find that there is a random error in discrete chaotic sequence. This error is caused by operation precision of computer. It can not be guaranteed that each of random error is dependent, so there is the correlation between sequences. If the autocorrelation of sequence is strong, its period is very short. This kind of sequence is not suitable to be applied in the chaotic encryption system.

3. **Selection of eliminating sequence's correlation methods.** Firstly, we need to check the correlation of sequence. There are a variety of testing methods for the correlation of sequence such as Test Procedure of Mr Norman, Regression Test, D.W.Test, ect. The common idea of these testing methods is that: firstly we use the common Least Squares Estimate Model to get the Random error's "approximate estimator", then through analyzing the correlation among these "approximate estimator" to judge that the random error term is whether correlative or not. If the model is tested to be correlative, then we need develop new method to estimate the model. The most common methods are generalized by least square method and finite difference method. Generalized least square method, just as its name implies, is the most universal significance least square method. Ordinary least squares and Weighted least squares are its special cases. Finite difference method is a kind of method to overcome correlation of sequence effectively, so it can be widely used. Finite difference method is the transformation for difference model, which can be divided into a order difference method and general finite difference method. Applying the general finite difference method, we must know the random error term's correlation coefficient in different sample points. As a matter of fact, people do not know their specific values, then we must firstly estimate these values, so many estimation methods, such as Iteration method, Doberman two step method and so on, are developed. The basic idea of these methods is to use the common Least squares estimate model for getting a random error of the approximate estimated value, and then use the approximate estimated value to obtain the estimator of random error's correlation coefficient. Different methods are used to make these estimators closer to the reality. In recent years, the common used decorrelation methods are Principal component analysisPCA, Singular value decompositionSVD, Discrete cosine transformDCTand K-L transform method in signal processing field. Among these methods, discrete cosine transform is the easiest way to realize and K-L transform method is the best way to eliminate correlation.

We give K-L transform to eliminate the correlation of discrete chaotic sequence below.

4. **K-L transform theory.** Karhunen-Loeve transform is a transformation built on the basis of statistical characteristics, and some documents also call Hotelling Transform, because Hotelling first gave the method which turns discrete signals into a string of non-correlation coefficient[20]. Outstanding advantage of K-L transform is good correlation. Simply speaking, a simplified analysis method of the complex relationship between each interrelated variables, is the best transformation in the sense of the mean square error (MSE, Mean Square Error). It occupies an important position in the data compression and image rotation technology.

K-L transform is a special case of orthogonal linear transformation.

K-L transformation formula is

$$Y = AX. \tag{5}$$

Inverse transformation formula is

$$X = A^T Y. \tag{6}$$

A must meet the following requirement: A is orthogonal matrix. If A is orthogonal matrix, we have $A^T A = I, A^T Y = A^T AX = X$, So we can get X. A is a transposed matrix of eigenvectors matrix of covariance matrix of X matrix. According to the calculation principle of covariance matrix, covariance matrix of X is a real symmetrical matrix. Real symmetrical matrix with n order must have n linear independent feature vector eigenvectors corresponding to different eigenvalues in real symmetric matrix are orthogonal, according to the theory of linear algebra[21]. So A that we get is a orthogonal matrix

and its lines or columns are linear independent. So it reduces correlation of Y by K-L transform.

5. **Generation of Logistic binary sequence and K-L transform.** Mathematical expression of logistic mapping is followed as

$$x_{n+1} = \mu x_n(1 - x_n),\ \mu \in (0, 4),\ x_n \in (0, 1). \tag{7}$$

When μ belongs to $[3.5699456, 4]$, logistic mapping turns into chaotic state and presents complex dynamic characteristics. We design digital circuit model of logistic chaos mapping (as shown in Fig. 1) based simulink according to (7), where initial key is $x \in (0, 1)$, and output sequence of key sequence generator is 0/1 sequence[22].

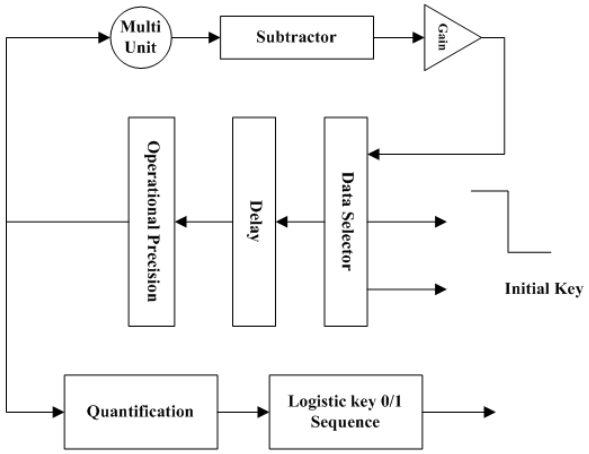


FIGURE 1. Circuit of Logistic Chaotic 0/1 Sequence Based on simulink

For logistic binary sequence $(x_1, x_2, \cdots, x_{N_m}$, because K-L transform is a kind of orthogonal transformation, and what the circuit diagram produce is 0/1 sequence, we choose the embedded dimension m, delay time τ (we choose numbers of 0/1 sequence shifted right as delay time τ according to discrete character of digital chaos) to restructure phase space, so matrix which the time sequence structures is $X_{m\times N_m}$ [23].

$$X_{m\times N_m} = \begin{pmatrix} x_1 & x_2 & \cdots & x_{N_m} \\ x_{1+\tau} & x_{2+\tau} & \cdots & x_{N_m+\tau} \\ \cdots & \cdots & \ddots & \cdots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \cdots & x_{N_m+(m-1)\tau} \end{pmatrix} \tag{8}$$

Calculate covariance matrix of $X_{m\times N_m}$: define each column as a set of vector, so matrix $X_{m\times N_m}$ is expressed as $X_{m\times N_m} = [X_1, X_2, \cdots, X_{N_m}$, and each element X_i in matrix $X_{m\times N_m}$ has m samples respectively. Its average vector is defined as follow

$$M_{X_i} = E\{X_i\} \cong \frac{1}{m}\sum_{j=1}^{m} X(j, i). \tag{9}$$

Its covariance matrix is a $N_m \times N_m$ matrix, we assume that the covariance matrix is

$$Z = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1N_m} \\ c_{21} & c_{22} & \cdots & c_{2N_m} \\ \cdots & \cdots & \ddots & \cdots \\ c_{N_m1} & c_{N_m2} & \cdots & c_{N_mN_m} \end{pmatrix} \tag{10}$$

In covariance matrix Z, every element c_{ij} is covariance of i_{th} column and j_{th} column. Set $\lambda \in [\lambda_1, \lambda_2, \cdots, \lambda_{N_m}$ as eigenvalue of matrix Z, and $F \in [F_1, F_2, \cdots, F_{N_m}]$ as its corresponding eigenvector, then their relation is showed in (11) and (12)

$$| Z - \lambda I | = 0, \tag{11}$$

$$ZF = \lambda F. \tag{12}$$

Eigenvector F is m dimension, so we could solve m feature value $\lambda_1, \lambda_2, \cdots, \lambda_m$. When the m feature values are used respectively in (12) above, we can obtain m eigenvectors corresponding to eigenvalues:

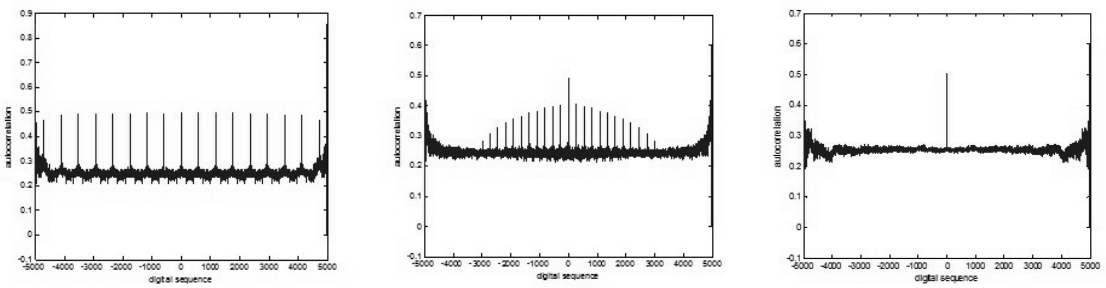
$$F_i = [f_{i1}, f_{i2}, \cdots, f_{iN_m}], \ (i = 1, 2, \cdots, N_m). \tag{13}$$

Matrix that all transposed eigenvectors structure is transformation matrix.

$$\Phi = [F_1^T, F_2^T, \cdots, F_{N_m}^T] = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1N_m} \\ f_{21} & f_{22} & \cdots & f_{2N_m} \\ \cdots & \cdots & \ddots & \cdots \\ f_{N_m1} & f_{N_m2} & \cdots & f_{N_mN_m} \end{pmatrix}^T \tag{14}$$

namely $Y = \Phi X$,is K-L transform, which reduces correlation of X (covariance matrix of Y is diagonal matrix), so correlation of Logistic binary sequence $(x_1, x_2, \cdots, x_{N_m})$ also decreases.

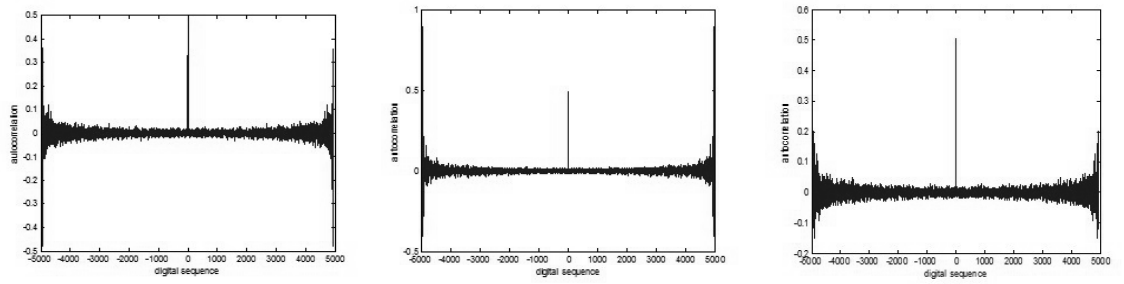
6. **Autocorrelation Analysis.** Choose $\mu = 4$ and initial value $X(0) = 0.2$ of logistic equation as the conditions of initial key in Fig. 1 under Matlab environment. When we choose different precisions, their autocorrelation tests are shown in Fig. 2.



(a) Precision is 18 (b) Precision is 24 (c) Precision is 36

FIGURE 2. Autocorrelation Test Charts

Therefore, operation precision is the main reason leading to chaotic dynamical characteristics degradation in the digital chaotic system. The higher the operation precision is, the better the digital chaotic dynamical characteristic is, however, operation precision in system is extremely limited, so we cannot arbitrarily choose bigger operation precision. Short period digital chaotic sequence that is generated by specific operation precision for encrypting the information will generate short period ciphertext sequence, which leads to eventual insecurity of encrypted system. When we realize chaotic iterative operation in digital system, time domain and value domain are dispersed, that is to say, each iteration calculation brings quantization error; the higher precision realized is, the smaller average quantization error is, conversely, the average quantization error is quite big[24], So this paper proposes K-L transform method for logistic based on phase space reconstruction that using delay coordinate state method. The first line sequence of the matrix is key sequence. Autocorrelation tests are shown in Fig. 3:



(a) Precision is 18 (b) Precision is 24 (c) Precision is 36

FIGURE 3. K-L Transform Autocorrelation Test

Using statement $X = \text{wgn}(1, 4551, 0)$ generates Gaussian white noise sequence and analyzes autocorrelation as shown in Fig. 4:

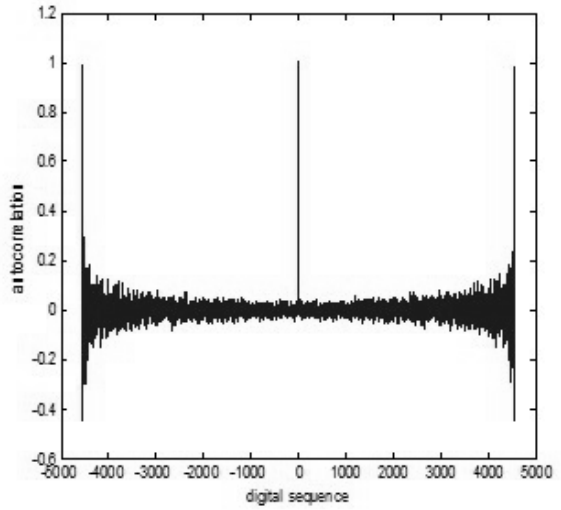


FIGURE 4. Gaussian White Noise Sequence Autocorrelation

It shows that K-L transform can reduce the correlation between the logistic sequence, its autocorrelation chart is similar to Gaussian white noise's, and could increase the periodicity of logistic sequence, moreover, makes up the short periodicity phenomenon, which makes the digital chaotic key sequence applied to the encryption system safely.

7. **Analysis of period and complexity.** Using neural network approach[25], the paper analyzes period and complexity of logistic 0/1 sequence produced by figure 1 and normalized 0/1 sequence after K-L transform. This paper selects 100000/1 sequences, and then simulates them under Matlab environment. Period and complexity measured are shown in table. 1: We can see through the table 1: 0/1 sequence period after K-L transform changes from original 588 to no-period. Minimum complexity, maximum complexity and average complexity has increased greatly. We change the sequence's length, and then conduct simulation experiments. The results show that normalized 0/1 sequences after K-L transform basically have no periods and the complexities have increased significantly. It shows that the new method is significantly improved in period and complexity.

TABLE 1. Sequence Period and Complexity Comparison of before and after K-L Transform

Sequence name	Logistic 0/1 sequence produced by figure 1	Normalized 0/1 sequence of phase space reconstruction K-L transform
Period	588	No period
Minimum complexity	0.117505	0.123653
Maximum complexity	0.127293	0.131246
Average complexity	0.123136	0.127419

8. **Conclusion.** Due to limitation of system accuracy, chaotic sequences produced by digital chaotic systems present chaotic degeneration phenomenon, which strengthens autocorrelation of sequence and causes periodic secret key sequence, finally leads to insecurity of encryption system. As to short periodicity of the digital chaotic system and strong correlation of digital chaotic sequence, this paper uses delay coordinate state space method for logistic sequence to restructure the phase space, and then makes K-L transform for reconstruction phase space. So this method increases the period of digital logistic sequence, eliminates the burrs, reduces the correlation of digital chaotic sequence, and increases complexity of key sequence, which makes digital chaotic sequences be safely applied to key encryption system.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (No.61072072) and the Science and Technology Project Supported by Electronic Engineering Key Laboratory of Heilongjiang Provincial Department of Education (DZZD20100010). This paper was supported by Innovated Team Project of 'Modern Sensing Technology' in colleges and universities of Heilongjiang Province (No. 2012TD007) and Institutions of Higher Learning by the Specialized Research Fund for the Doctoral Degree(No. 20132301110004).

REFERENCES

[1] E. N. Lorenz, Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130-141, 1964.
[2] T. Y. Li, and J. A. Yorke, Period three implies chaos, *The American Mathematical Monthly*, vol. 82, no. 10, pp. 985-992, 1975.
[3] R. M. May, Simple mathematical models with very complicated dynamics, *Nature*, vol. 261, pp. 459-467, 1976.
[4] R. Matthews, On the derivation of a chaotic encryption algorithm, *Cryptologia*, vol. 8, no. 1, pp. 29-41, 1984.
[5] L. O. Chua, and T. Lin, Chaos in Digital Filters, *IEEE Trans. Circuits and Systems*, vol. 35, no. 6, pp. 648-658, 1988.
[6] L. M. Pecora, and T. L. Carroll, Synchronization in chaotic systems, *Physical Review Letters*, vol. 64, no. 8, pp. 821-824, 1990.
[7] D. R. Frey, Chaotic digital encoding: an approach to secure communication, *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 10, pp. 660-666, 1993.
[8] M. Tatebayashi, N. Matsuzaki, and D. B. Newman, A Cryptosystem Using Digital Signal Processors for Mobile Communication, *Proc. of IEEE International Conference on World Prosperity Through Communications*, pp. 1145-1148, 1989.

[9] J. Palmore, and C. Herring, Computer arithmetic, chaos and fractal, *Physica D*, vol. 42, no. 1-3, pp. 99-110, 1990.
[10] P. H. Borcherds, The digital tent map and the trapezoidal map, *chaos, Soliton & Fractals*, vol. 3, no. 4, pp. 451-466, 1993.
[11] M. J. WERTER, An improved chaotic digital encoder, *IEEE trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no, 2, pp. 227-229, 1998.
[12] T. Kohda, and A. Tsuneda, Statistics of chaotic binary sequences, *IEEE Trans. Information Theory*, vol. 43, no. 1, pp. 104-112, 1997.
[13] D. Utami, H. Suwastio, and B. Svmadjudin, FPGA implementation of digital chaotic cryptography, *Proc. of the 1st EurAsian Conference on Information and Communication Technology*, LNCS 2510, pp. 239-247, 2002.
[14] L. Kocarev, and G. Jakimoski, Pseudorandom bits generated by chaotic maps, *IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, no. 1, pp. 123-126, 2003.
[15] J. Gernak, Digital generators of chaos, *physical letters A*, vol. 241, no. 3-4, pp. 151-160, 1996.
[16] C. Jin, Y. Yang, and C. Qi, Relevant key attack of chaos sequence cipher, *Journal of electronics and information technology*, vol. 28, no. 3, pp. 410-414, 2006.
[17] Shujun Li, Analysis and design of digital chaotic cipher, Ph. D. Thesis, Xi'an Jiaotong University, Xi'an, China, 2003.
[18] Q. Ding, J. Pan, L. Wang, and G. Chen, The cipher code parameter selection and impact on output cycles, *Proc. of the 2009 International Workshop on Chaos-Fractals Theories and Applications*, pp. 143-147, 2009.
[19] Q. Ding, X. Peng, and Z. Yang, Combined sequence password chip based on neural network algorithm, *Journal of electronics and information technology*, vol. 34, no. 3, pp. 409-412, 2006.
[20] Guangshu Hu, *Digital signal processing*, Tsinghua University Press, 2003.
[21] Sujuan Li, *Linear algebra*, Harbin Press,2003.
[22] Q. Ding, J. Pang, J. Fang, and X. Peng, Designing of chaotic system output sequence circuit based on FPGA and its possible applications in network encryption card, *International Journal of Innovative Computing, Information and Control(IJICIC)*, vol. 2, no. 3, pp. 449-456, 2007.
[23] Yan Wang, Phase space restruction, bifurcation and economic system attractor analysis, Ph. D. Thesis, Northwestern Polytechnical University, Xi'an, China, pp. 43-50, 2006.
[24] Q. Ding, Y. Zu, F. Zang, and X. Peng, Discrete chaotic circuit and the property analysis of output sequence, —textitIEEE International Symposium on Communications and Information Technologies, vol.2, pp. 1043-1046, 2005.
[25] F. Chen, G. Chen G. He, X. Xu, and Q. He, Universal perceptron and DNA-like learning algorithm for binary neural networks: LSBF and PBF implementations, *IEEE trans. Neural Networks*, vol. 20, no, 10, pp. 1645-1658, 2009.