# ON TRAINING NEURAL NETWORK ALGORITHMS FOR ODOR IDENTIFICATION FOR FUTURE MULTIMEDIA COMMUNICATION SYSTEMS

*Ki-Hyeon KWON[1], Namyong KIM[1], Hyung-Gi BYUN[1], and Krishna C. PERSAUD[2]*

[1] Dept. Of Information & Communication Eng. Kangwon National University, Korea
[2] School of Chemical Eng. Analytical Science, University of Manchester, Manchester, U.K.

## ABSTRACT

Future multimedia communication system can be developed to identify, transmit and provide odors besides voice and image. In this paper, an improved odor identification method is introduced. We present an analysis of center-gradient and a new method of using convergence parameters in training RBFN-SVD-SG (Radial Basis Function Network using Singular Value Decomposition combined with Stochastic Gradient) algorithm for odor identification. Through mathematical analysis, it was found that the steady-state weight fluctuation and large values of convergence parameter can lead to an increase of variance of center-gradient, which induces ill-behaving convergence. The proposed method of using raised-cosine functions for time-decaying convergence parameter shows faster convergence and better recognition performance.

## 1. INTRODUCTION

Multimedia systems have been developed to process odor information as well as voice, image and video information. A multimedia PC providing odors has been introduced in [1]. The odors are related to the text information, pictures, audio and video clips on the PC. In [2], methods on the expression of the odor on virtual space have been studied. Future multimedia communication systems like mobile phones equipped with odor sensors can be developed to collect, identify, transmit and provide extra sensory information of odor besides voice and image. To odor identification problems, Radial basis function networks (RBFN) and other neural networks have recently been applied [3][4]. The centers and widths of the RBFN are calculated by fuzzy c-means algorithm and the distribution of input patterns. The weights of RBFN are calculated by SVD (Singular Value Decomposition) method in a single shot process. This RBFN-SVD is considered superior to the other learning algorithms, particularly in odor analysis. But the performance of RBFN as a classifier is highly dependent on the choice of centers and widths in basis function. So, the fine-tuning of centers and widths is needed and the

stochastic gradient (SG) method [5] is successfully applied for it. The adaptive RBFN-SVD-SG has shown good identification performance for complex and mixture chemical patterns and confirmed by experimental trials [6]. In this paper we present a theoretical analysis of the factors contributing to center-gradient fluctuation of RBFN-SVD-SG algorithm and propose a method to acquire faster convergence and better recognition performance.

## 2. RBF-SG ALGORITHM

Let input vector $x_n$ from a conducting polymer sensor array $x_n = [x(n), x(n-1), ..., x(n-M+1)]^T$, and $d_n$ denotes desired output. Let $y_n$ denote the RBF output for input $x(n)$ and let the error be denoted by $e_n = d_n - y_n$, all at the training time $n$. The output of RBFN with M Gaussian basis functions is

$$output(x_n) = \sum_{j=1}^{j=M} w_j \exp(-\frac{\left\| x_n - c_j^{(n)} \right\|^2}{(\sigma_j^{(n)})^2}) \tag{1}$$

where, $c_j^{(n)}$ is j-th element of center vector $C^{(n)}$, $w_j$ is j-th element of weight and $\sigma_j^{(n)}$ is j-th element of width at time n. The RBF-SG algorithm adapts the network parameters according to the following equations (2) and (3) [7]. In (2) and (3), $\mu_s$ and $\mu_c$ are adaptation parameters. The weights are tuned by SVD in the same iteration. [6].

$$c_j^{(n+1)} = c_j^{(n)} + \mu_c e_n w_j^{(n)} \exp(\frac{-\left\| x_n - c_j^{(n)} \right\|^2}{(\sigma_j^{(n)})^2}) \frac{[x_n - c_j^{(n)}]}{(\sigma_j^{(n)})^2} \tag{2}$$

$$\sigma_j^{(n+1)} = \sigma_j^{(n)} + \mu_s e_n w_j^{(n)} \exp(\frac{-\left\| x_n - c_j^{(n)} \right\|^2}{(\sigma_j^{(n)})^2}) \frac{\left\| x_n - c_j^{(n)} \right\|^2}{(\sigma_j^{(n)})^3} \tag{3}$$

## 3. MATHEMATICAL ANALYSIS OF CENTER-GRADIENT

The performance of the stochastic gradient method is dependent on how accurate the estimation of MSE gradient is [8]. In RBF-SG algorithm, the center-gradient estimate is

$$\nabla_c^{\wedge}(n) = \frac{\partial e^{(n)2}}{\partial C^{(n)}} \qquad (4)$$

The center-gradient estimate for j-th center becomes

$$\nabla_{cj}^{\wedge}(n) = \frac{\partial e^{(n)2}}{\partial c_j^{(n)}} \qquad (5)$$

can be expressed as

$$\nabla_{cj}^{\wedge}(n) = -4e^{(n)}w_j^{(n)}\exp(\frac{-\left\|x^{(n)} - c_j^{(n)}\right\|^2}{\sigma_j^{(n)2}})\frac{[x^{(n)} - c_j^{(n)}]}{\sigma_j^{(n)2}} \qquad (6)$$

Defining

$$O_j^{(n)} = \exp(\frac{-\left\|x^{(n)} - c_j^{(n)}\right\|^2}{\sigma_j^{(n)2}}) \qquad (7)$$

Equation (2) and (4) become (8) and (9), respectively.

$$c_j^{(n+1)} = c_j^{(n)} + \mu_c e^{(n)}w_j^{(n)}O_j^{(n)}\frac{x^{(n)} - c_j^{(n)}}{\sigma_j^{(n)2}} \qquad (8)$$

$$\nabla_{cj}^{\wedge}(n) = -4e^{(n)}w_j^{(n)}O_j^{(n)}\frac{[x^{(n)} - c_j^{(n)}]}{\sigma_j^{(n)2}} \qquad (9)$$

Assuming centers in the steady-state, (i.e., as $n \to \infty$), are located close to the optimum centers and the Euclidian distance between the input patterns and the cluster's center is less than the width, $\left\|x_n - c_j^{(n)}\right\|^2 << \sigma_j^{(n)2}$, $O_j^{(n)}$ in (7) closely approximates to 1. Then the variance of $\nabla_{cj}^{\wedge}(n)$, $Var[\nabla_{cj}^{\wedge}(n)]$, is

$$Var[\nabla_{cj}^{\wedge}(n)] = \frac{16}{(\sigma_j^{(n)})^4}E[e^{(n)2}]E[w_j^{(n)2}]$$
$$\cdot E[(x^{(n)} - c_j^{(n)})\cdot(x^{(n)} - c_j^{(n)})] \qquad (10)$$

Defining the difference between input pattern $x^{(n)}$ and optimum center $c_j^o$ as $v_j^{(n)} = x^{(n)} - c_j^o$ and center deviation $f_j^{(n)}$ as $f_j^{(n)} = c_j^{(n)} - c_j^o$, $x^{(n)} - c_j^{(n)}$ becomes

$$x^{(n)} - c_j^{(n)} = v_j^{(n)} - f_j^{(n)} \qquad (11)$$

Substituting (11) into (10) yields

$$Var[\nabla_{cj}^{\wedge}(n)] = \frac{16}{\sigma_j^{(n)4}}E[e^{(n)2}]E[w_j^{(n)2}]$$
$$\cdot E[(v_j^{(n)} - f_j^{(n)})\cdot(v_j^{(n)} - f_j^{(n)})]$$

$$= \frac{16}{\sigma_j^{(n)4}}E[e^{(n)2}]E[w_j^{(n)2}]$$
$$\cdot(E[v_j^{(n)}\cdot v_j^{(n)}] - 2E[f_j^{(n)}\cdot v_j^{(n)}] + E[f_j^{(n)}\cdot f_j^{(n)}]) \qquad (12)$$

It can be assumed that $v_j(n)$ and $f_j(n)$ are uncorrelated and $E[f_j^{(n)}\cdot v_j^{(n)}] = 0$. By defining the variance of input patterns $E[v_j^{(n)}\cdot v_j^{(n)}]$ as $V_j$,

$$Var[\nabla_{cj}^{\wedge}(n)] = \frac{16}{\sigma_j^{(n)4}}E[e^{(n)2}]E[w_j^{(n)2}]$$
$$\cdot(V_j + E[f_j^{(n)}\cdot f_j^{(n)}]) \qquad (13)$$

Using $f_j^{(n)} = c_j^{(n)} - c_j^o$, (8) can be expressed as

$$f_j^{(n+1)} = f_j^{(n)}[1 - \frac{\mu_c e^{(n)}w_j^{(n)}}{\sigma_j^{(n)2}}] + \mu_c e^{(n)}w_j^{(n)}\frac{v_j^{(n)}}{\sigma_j^{(n)2}} \qquad (14)$$

For convenience's sake, we define

$$B = \mu_c e^{(n)}w_j^{(n)} / \sigma_j^{(n)2} \qquad (15)$$

Taking the Z transform of (14), the transfer function for the variance of j-th center-gradient is the system $A(j,z)$ with input $v_j^{(n)}$ and output $f_j^{(n)}$ as follows

$$A(j,z) = \frac{Bz^{-1}}{1 - (1-B)z^{-1}} \qquad (16)$$

Since in the steady-state, (i.e., as $n \to \infty$), variance of $f_j^{(n)}$, $E[f_j^{(n)}\cdot f_j^{(n)}]$, equals the sum of the squares of the impulse response of the $A(j,z)$, times the variance of the input, $E[f_j^{(n)}\cdot f_j^{(n)}]$ is given by

$$E[f_j^{(n)}\cdot f_j^{(n)}] = V_j\Phi_j \qquad (17)$$

where

$$\Phi_j = \sum_{n=1}^{\infty}[(1-B)^{n-1}B]^2 = \frac{B^2}{1 - (1-B)^2}$$
$$= \mu_c e^{(n)}w_j^{(n)} / 2\sigma_j^{(n)2} \qquad (18)$$

for $\mu_c e^{(n)}w_j^{(n)} / 2\sigma_j^{(n)2} << 1$

Substituting (18) into (13) and using $E[e^{(n)2}] = MSE_{min}$ in the steady-state, lead to

$$Var[\nabla_{cj}^{\wedge}(n)]$$
$$= \frac{16MSE_{min}}{\sigma_j^{(n)4}}E[w_j^{(n)2}]\cdot(V_j + \mu_c e^{(n)}w_j^{(n)} / 2\sigma_j^{(n)2}) \qquad (19)$$
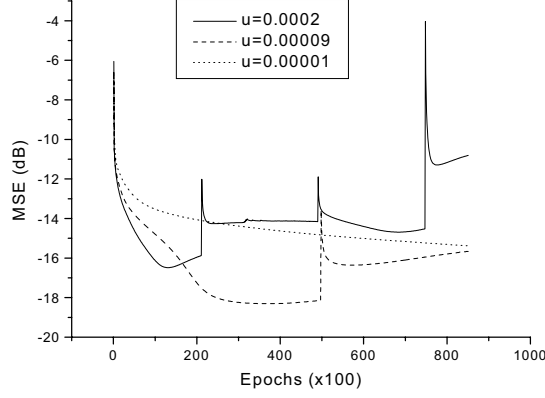
Fig. 1. MSE curves of RBF-SVD-SG (u= $\mu_c$ )



Fig. 2. MSE Performance of the proposed method ( $\mu_c$ =0.0001, $\alpha$ =0.00001, and Threshold = 0.00001)

In (19), we can find that steady-state weight fluctuation and large values of $\mu_c$ can lead to large fluctuation of center-gradient estimate. This large center-gradient variance can cause ill-behaving MSE convergence. We observed experimentally that MSE learning curves of RBF-SVD-SG algorithm bounce if not sufficiently small value of $\mu_c$ is used as presented in Fig. 1.

## 4. THE PROPOSED TIME-DECREASING CONVERGENCE PARAMETER

A method of avoiding the MSE increase can be to use very small value of $\mu_c$ . But small $\mu_c$ makes the convergence speed of the RBF-SG algorithm slow. It is desired to use methods of acquiring fast learning speed and low MSE performance. The use of time-decreasing convergence parameter [7] can be a solution to the problem of RBF-SG algorithm. We propose a modified version of raised-cosine function with roll-off factor = 1 [9] to use as a time-decreasing function. In (20), $\alpha$ is the smallest convergence parameter for large $n$ and $\mu(t)$ is a decreasing function from $\mu_c$ to $\alpha$ . The proposed time-decreasing convergence parameter in (20) is applied when the algorithm converges

$$\mu(t) = \begin{cases} \alpha + 1/2(\mu_c - \alpha)[1 + \cos(\pi t)]; 0 \le t < 1 \\ \alpha; t \ge 1 \end{cases} \quad (20)$$

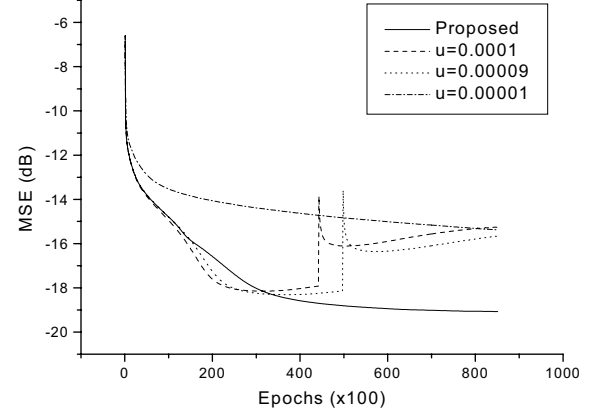to an acceptable level of error performance by observing the averaged center-gradient estimate in (6), i.e., when

$\frac{1}{M}\sum_{j=1}^{M} \nabla_{cj}^{\wedge}(n) \le Threshold$ . For gradually decreasing convergence parameter, $t$ can be set as $t = count * \Delta$ , where $\Delta$ is a small number.

$$if(\frac{1}{M}\sum_{j=1}^{M} \nabla_{cj}^{\wedge}(n) \le Threshold)$$

$$\{t = count * \Delta$$

$$\mu_c = \mu(t); \quad count = count + 1;\}$$

## 5. RESULTS AND DISCUSSION

We used a odor sensing system which has an array of conducting polymer sensors mounted on a ceramic substrate together with associate electronics developed by Prof. K.C. Persuad at University of Manchester, U.K. Measurement of chemicals were repeated to collect patterns from solvent vapors over periods of 4 weeks. After having trained the RBF network with 8 centers using weeks 1 and 2 data sets obtained from 1% acetonitrile (ac1), 10% acetonitrile (ac10), 1% acetone (ae), 1% butanone (bu), 10% methanol (me), 1% propanol (pr1), 10% propanol(pr10), and water (wa) , the adaptive RBFN-SVD-SG was applied to the previously unseen data from weeks 3 and 4 to evaluate odor prediction. The ill-behaving MSE learning could be avoided by using very small value of $\mu_c$ but it made the convergence speed very slow as in Fig. 2. The MSE curves of RBFN-SVD-SG with constant convergence parameters reach the minimum
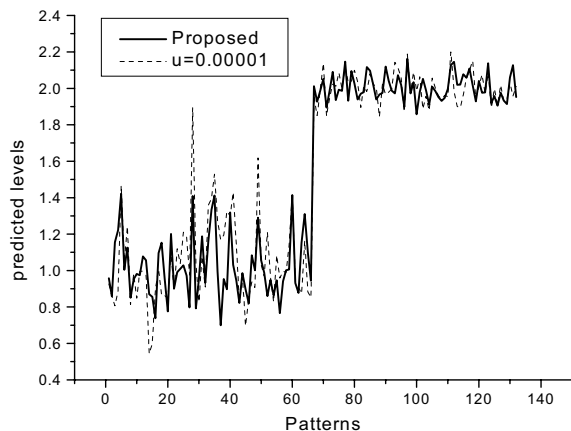
Fig. 3. Identification results for ac1 and ac10 (ac1: level 1, ac10: level 2).

MSE but they increase again and have abrupt bounces at about 45000 or 50000 epochs. The proposed method using a time-decreasing convergence parameter starts at about epoch number 10000 to lower the value of $\mu_c$ according to the raised-cosine function $\mu(t)$ when center-gradient estimate is judged to be less than the threshold. The proposed method for RBFN-SVD-SG converges to –19 dB, but the RBFN-SVD-SG with constant convergence parameters shows bouncing and increasing MSE curves after having reached to –18 dB of MSE. For observations in more detail, the identification performance at the epoch 80000 for level 1 and level 2 is shown in Fig. 3, where level 1 indicates ac1, level 2 ac10, respectively. When the level 1.5 is chosen as a decision threshold to discriminate ac1 and ac10, many decision errors for level 1 are detected in the algorithm using constant convergence parameter. On the other hand, all patterns for level 1 of the proposed method are within the decision region.

## 6. CONCLUSIONS

This paper presents an improved odor identification method for multimedia systems that can collect and provide extra sensory information, odor. Firstly an analysis of center-gradient variance of RBFN-SVD-SG algorithm was carried out. The center-deviation can have a negative effect on the RBF hidden node output. We found theoretically that the steady-state weight fluctuation and large values of $\mu_c$ can lead to an increase of variance of center-gradient estimate. The MSE learning curves of RBFN-SVD-SG algorithm can show bounces in accordance with the unstable behavior of center-gradient. Using a smaller value of convergence parameter, the problem can be avoided but it causes a slower convergence. The proposed method of using raised-cosine functions for time-decreasing convergence parameters shows stable, fast learning and improvement in minimum MSE. The test odors are identified with much better accuracy when the proposed method is used than constant convergence parameter is used.

## 7. REFERENCES

[1] Tan EC, Wahab A, Goh GH and Wong SH, "PC-Controlled Scent System", IEEE Transactions on Consumer Electronics, V.44 N.1, Feb. 1998.

[2] Tominaga, K., Honda, S., Ohsawa, T., Shigeno, H., Okada, K., Matsushita, Y., "Friend Park-expression of the wind and the scent on virtual space", International Conference on Virtual Systems and Multimedia, Berkeley, USA, pp. 507-515, Oct. 2001.

[3] K. C. Persaud and H. Byun, "Industrial applications of neural networks" World Scientific, pp. 85-90

[4] D. S. Lee, J. S. Huh, H. Byun and D. D. Lee, "An Electronic Nose for Recognizing Combustible Gases using Thick Film Sensor Array and Neural Network", Proceedings of ISOEN'2000, pp. 115-120, 2000.

[5] I. Cha and S. A. Kassam, "Interference cancellation using radial basis function networks", Signal processing, vol. 47. pp. 247-268, 1995.

[6] H. Byun, N. Y. Kim, K. C. Persaud, J. S. Huh, and D. D. Lee, "Application of Adaptive RBF Networks to Odour Classification using Conducting Polymer Sensor Array", Proceedings of ISOEN'2000, pp. 121-126, 2000.

[7] I. Cha and S. A. Kassam, "Channel equalization using adaptive complex radial basis function networks," IEEE Journal on Selected Areas in Communications, vol. 13., pp. 122-131, Jan. 1995.

[8] B. Widrow and S. D. Stearns, *Adaptive signal processing*, Prentice-Hall, 1985.

[9] J. G. Proakis, *Digital Communications*. 2nd ed, McGraw-Hill, 1989.