Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Development of the Color Constancy Vision Algorithms using Bio-inspired Information Processing

Yasunori Takemura, Kazuo Ishii / Kyushu Institute of Technology
Dept. of Brain Science and System Engineering, KIT
2-4 Hibikino Wakamatsu Kitakyushu, Fukuoka, Japan
{takemura-yasunori@edu.}{ishii@}brain.kyutech.ac.jp

***Abstract*—One of important subjects for mobile robots is the vision based decision-making system with environmental recognition. In order to extract features from obtained images, how to realize color constancy by adjusting color property is the important technical issue. We have been working on color constancy vision algorithms using bio-inspired information processing methods as Self-Organizing Map (SOM), modular network SOM (mnSOM) and Neural Gas (NG). In this paper, we introduce the color constancy algorithms to a mobile robot platform for RoboCup middle size league and evaluate the performances through a vision based self-localization problem in various light conditions.**

Keywords— color constancy, robot vision, bio-inspired information processing

I. INTRODUCTION

One of important subjects for most of mobile robots is the vision based decision-making system, where the perceptual constancy is big problem for robots which use vision to recognize environments. There are several types of perceptual constancies: shape constancy, size constancy, color constancy, lightness constancy, distance constancy, and location constancy. The color constancy is a feature of the human color perception system which ensures that the perceived color of objects remains relatively constant under varying illumination conditions. As the color constancy is not realized, the working space of robots is limited and robots often miss to detect target objects.

In order to realize robust robot system which works outdoor environments and in various and varying lighting conditions, a system to keep color constancy is needed by adapting vision system to various color profiles of working environments. As the target robots system, we use RoboCup robots which move around using omni-vision camera indoor, because the lighting condition varies from hour to hour, weather also affects to color recognition, and a certain level of repeatability of lighting condition can be expected.

RoboCup is an international joint project to promote Artificial Intelligent (AI), robotics, and related fields. It is an attempt to foster AI and Intelligent robotics research by providing standard problems where a wide variety of technologies can be integrated and examined. In RoboCup, soccer game is selected as a main topic of research, aiming at innovations to be applied for socially significant problems and industries in the future [1]. “Hibikino-Musashi” is a joint RoboCup middle-size league (MSL) soccer team in Kitakyushu Science and Research Park [2,3]. “Musashi” robot is developed based on the concept of the “omni-directional” and “modularity” concept (Fig.1) [4].

Fig. 1 Omni-directional mobile soccer robot “Musashi”

In this paper, we present the color recognition system based on color constancy algorithm using bio-inspired information technology. As the bio-inspired technology, we evaluate Multi-Layer Perceptron (MLP), Self-Organizing Map (SOM), modular network SOM (mnSOM), k-means, Neural Gas (NG). For examples, Self-Organizing Map (SOM) is an unsupervised learning algorithm that performs topology-preserving transformation from higher-dimensional vector data spaces to low map spaces. The SOM has become a powerful tool in many areas such as data mining, data analysis, data classification, and data visualization [5, 6, 7, 8].

In the RoboCup, robots must detect a ball colored by orange, field by green and lines by white. Our current robot vision system uses two color models in YUV and HSV formats to detect target objects, and obtain different color images from both color models. Then, the images are banarized to detect target colors using a certain thresholds and added logically [9].

The important question is how the thresholds should be decided. In this paper, we introduce bio-inspired processing algorithms such as Self-Organizing Map (SOM), modular network SOM (mnSOM) and Neural Gas (NG) into the optimization of threshold parameters. The experimental results in various light conditions are discussed and evaluated.

II. VISION SYSTEM OF "MUSASHI" ROBOT

The vision system of the "Musashi" robot consists of an omni-directional mirror and an IEEE 1394 digital camera. Many RoboCup robots recently have strong kicking devices for high loop-shoots, so that the vision system should estimate the position of the ball in the air from single vision. The obtained image from the omni-vision camera (Fig.2a) is in the YUV format and also converted into HSV format (Fig.2c). The upper and lower thresholds to extract target colors are decided empirically in both YUV and HSV formats, and the resulted image is given as the AND operation of both images. In the example of Fig.2, the blue regions are extracted in YUV color space (Fig.2b) and HSV color space (Fig.2c) and the obtained image is in Fig.2d. It is shown that the blue goal is detected clearly using this method. In RoboCup, the robot should recognize three kinds of objects: the orange ball, the green field and the white lines. The orange ball is extracted as follows (1)~(3):

$$O = (V^o \cap H^o) \tag{1}$$

$$V^o \in [V^o_{min}, V^o_{max}] \tag{2}$$

$$H^o \in [H^o_{min}, H^o_{max}] \tag{3}$$

Let *O*, V^o and H^o be the extracted orange region by AND operation, the orange region estimated from *V* value in YUV format, and the orange region from H value in HSV format, respectively. YUV format values and HSV format values are expressed by 8 bits numbers [0, 255]. Let *G* and *W* be the green region and the white line region, respectively. The following describes the process to extract the field region and line region.

$$G = (U^g \cap V^g) \tag{4}$$

$$U^g \in [U^g_{min}, U^g_{max}] \tag{5}$$

$$V^g \in [V^g_{min}, V^g_{max}] \tag{6}$$

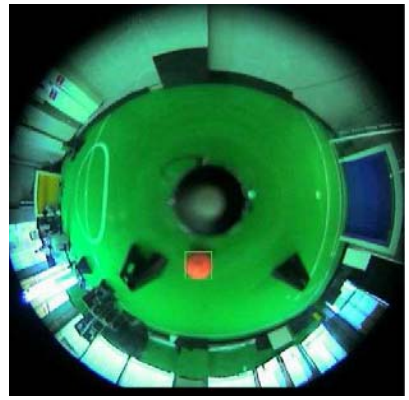
$$W = Y^w \tag{7}$$

$$Y^w \in [Y^w_{min}, Y^w_{max}] \tag{8}$$

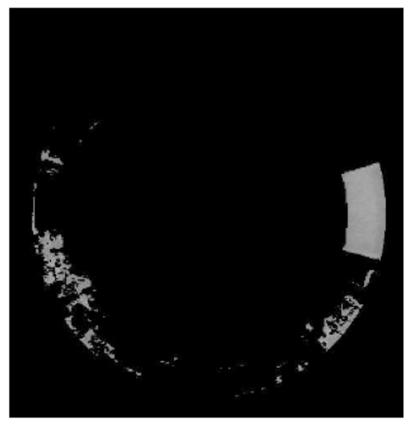
Y and *U* indicate the Y and U values in YUV, respectively. *B* and *Ye* indicate the each goal color, so that Goals are also extracted using this method (Fig.2d).

$$B = (U^b \cap H^b) \tag{9}$$

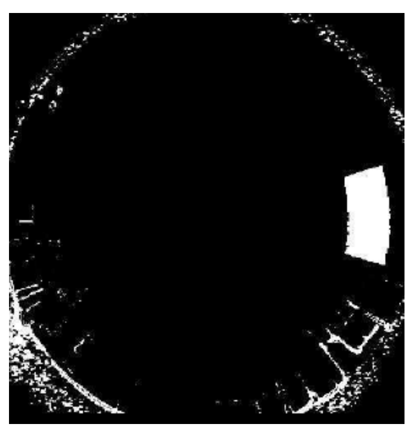
$$U^b \in [U^b_{min}, U^b_{max}] \tag{10}$$



(a) Image from the IEEE 1394 camera.

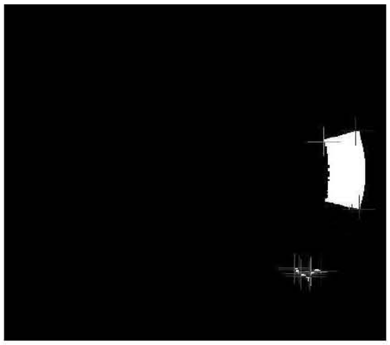


(b) Extracted blue objects of using the YUV color map.



(c) Extracted blue objects by using the HSV color map.

Fig.2 Blue color extraction result of "Musashi" robot vision system



(d) Result image of blue sampling.

Fig.2 Blue color extraction result of "Musashi" robot vision system

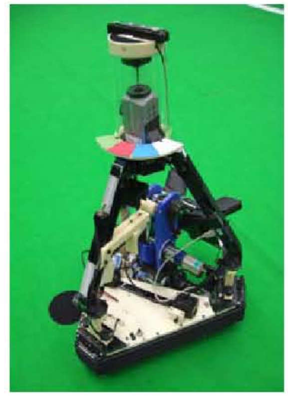


Fig.3 Template of the basic color bar around the omni-directional camera

$$H^{b} \in [H^{b}_{min}, H^{b}_{max}] \tag{11}$$

$$Ye = (V^{ye} \cap H^{ye}) \tag{12}$$

$$V^{ye} \in [V^{ye}_{min}, V^{ye}_{max}] \tag{13}$$

$$H^{ye} \in [H^{ye}_{min}, V^{ye}_{max}] \tag{14}$$

The upper suffix means the name of color, and the bottom means minimum (min) or maximum (max) values. This method improves the color extraction robustness and accuracy in variable lighting condition. We set the thresholds manually in the current system.

II. COLOR CONSTANCY ALGORITHM OF USING SELF-ORGANIZING MAPS

A Basic Input data and Output data

In this paper, Bio-Inspired Algorithms are used for the recognition to the light environment condition. At first, we explain a basic input data and output data.

The basic four kinds of colors are prepared around the omi-directional camera (Fig.3). In this time, the input data $\boldsymbol{x}$ describes as follows:

$$x = (Y_g, U_g, V_g, Y_r, U_r, V_r, Y_b, U_b, V_b, Y_w, U_w, V_w) \tag{15}$$

The variable Y_*, U_* and V_* correspond to luminance information (Y) and color information (U and V) and take values between 0 – 255 originally and are normalized between -1.0 and 1.0. The bottom index describes green (*g*), red (*r*), blue (*b*) and white (*w*). This input data set is made from camera law data signal.

In this time, we need extract the colors. Output data needs the three kinds of the color thresholds: There are orange (ball), white (line) and green (fields), because the RoboCup Middle Size League (MSL) changes the rule which changes color goal (blue and yellow) to no color goal in this year. Orange is recognized with equation (1) ~ (3). $H_{\min}$ are always constant value. Therefore, $H_{\min}$ except the output data. Green is recognized with equation (4) ~ (6). $U_{\min}$ are always constant value. Therefore, $U_{\min}$ except the output data. White is recognized with equation (7) ~ (8). $Y_{\max}$ are also always constant value. Therefore $Y_{\max}$ also except the output data. In this time, output data set *y* is described as follows:

$$y = (H^{o}_{\max}, V^{o}_{\min}, V^{o}_{\max}, V^{g}_{\max}, V^{g}_{\max}, Y^{w}_{\min}) \tag{16}$$

Upper index means orange (o), green (g), and white (w). All parameters are normalized -1.0 to 1.0 vectors. The proposed color constancy algorithms learn the mapping functions between the teaching vectors in (15) and (16). We suppose that the *x* in (15) expresses the fundamental environmental information and is used to decide the light condition. The *y* corresponds to the lighting condition.

B Color constancy algorithm using Self Organizing Map

In a Self-organizing map (SOM), the neurons are placed at the nodes of a lattice that is usually two-dimensional. Higher-dimensional maps are also possible but not as common. The neurons become selectively tuned to various input patterns or classes of input patterns in the course of a competitive learning process. The locations of the neurons so tuned (i.e., the winning neurons), neurons become ordered with respect to each other in such a way that a meaningful coordinate system for different input features is created over the lattice [10].

A SOM is therefore characterized by the formation of a topographic map of the input patterns in which the spatial locations (i.e., coordinates) of the neurons in the lattice are indicative of intrinsic statistical features contained in the input patterns, hence the name "Self-Organizing Map". In this time, the SOM algorithm is used on color detection algorithm.

SOM algorithm has the four processes: there are *Evaluative process*, *Competitive process*, *Cooperative process* and *Adaptive process*. Table I shows the variables used explanation of the color constancy algorithm of using batch type SOM. The learning data is defined by eq. (17).

TABLE I

VARIABLE USED EXPLANATION OF ALGORITHM

Symbol	Quantity
θ	Learning data
x	Input vector
y	Output vector
i	Index expressing classes (i = 1, ... , I)
w	Reference vector
k	Index expressing unit (k = 1,......, K)
E	Distance between input vector and reference vector
k^*	Best Matching Unit (BMU)
ϕ	Neigubor Function
$d(a,b)$	Euclidean distance between a and b
ψ	Learning late
σ	Neighbor radius
τ	Time constant

$$\theta_i = [x_i \quad y_i] \tag{17}$$

(a) *Evaluative Process*

In evaluative process, all learning data sets are calculated the distance between each reference vector.

$$E_i^k = \left\| w^k - x_i \right\|^2 \tag{18}$$

(b) *Competitive process*

In competitive process, to find the best matching of the input vector x with the reference vector w, the best matching unit (BMU) (k*) is defined that the minimum distance of eq. (18).

$$k_i^* = \arg_k \min E_i^k \tag{19}$$

(c) *Cooperative process*

In cooperative process, the winning units (best matching units) locate the center of a topological neighborhood of cooperating units. Let $d(k,k^*)$ denote the Euclidean distance between *k-th* module and k^*-th best matching unit. Then, we may assume that the topological neighborhood ϕ_i^k is an unimodal function of the distance $d(k, k^*)$, such that it satisfies two distinct requirement:

(i) The topological neighborhood ϕ_i^k is symmetric about the maximum point defined by $d(k, k^*) = 0$

(ii) The amplitude of the topological neighborhood ϕ_i^k decreases monotonically with increasing distance $d(k, k^*)$, decaying to zero for $d(k,k^*) \to \infty$; this is a necessary condition for convergence.

A typical choice of ϕ_i^k that satisfies these requirements are the Gaussian function (eq. (20)).

$$\phi_i^k = \exp\left(d(k,k_i^*)^2 \big/ 2\sigma^2 \right) \tag{20}$$

The parameter σ is the "effective width" of the topological neighborhood. It is called neighbor radius. Its use also makes the SOM algorithm converge more quickly than a rectangular topological neighborhood world [11, 12, 13]. Another unique feature of the SOM algorithm is that the size of the topological neighborhood shrinks with time. This requirement is satisfied by making the neighbor radius σ (eq. (21)) of the topological neighborhood function ϕ_i^k decrease with time. A popular choice for the dependence of σ on discrete time t is the exponential decay described by [14, 15].

$$\sigma = \sigma_{\min} + (\sigma_{\max} - \sigma_{\min}) \exp\left(-t \big/ \tau \right) \tag{21}$$

In batch type SOM, each unite learning rate is defined by ψ_i^k which is normalized by summation of the ϕ_i^k (eq. (22))

$$\psi_i^k = \phi_i^k \Big/ \sum_{i'} \phi_{i'}^k \tag{22}$$

(d) *Adaptive process*

In the Kohonen's SOM, in adaptive process, all unit vectors are adjusted by using the update formula eq.(23)

$$w^k(t+1) = w^k(t) + \psi_i^k (x_i - w^k) \tag{23}$$

In this color constancy algorithm, learning data θ is described by eq.(17). So, unit vector has also input vectors and output vectors. In other words, reference vectors have also same dimensions of input and output vectors. In this time, we use the batch SOM, then the reference vectors ***w*** are updated by eq. (24).

$$w^k(t+1) = \sum_i \psi_i^k \theta_i \tag{24}$$

In the processing mode, input data form the camera put the evaluation process. And, on the competitive process, the robot finds the BMU, and this BMU output vector y is used for thresholds.

C Color constancy algorithm using Neural Gas

Neural Gas (NG) is network model of using vector quantization. NG is also similar in SOM some respects [16]. About one of the biggest different point, SOM has the relationship of neighborhood unit, on the other hand, NG doesn't have the relationship each unit. Each unit is learning independently, and there can move freely. Therefore, number of the unit is more decrease than SOM. Basically, the algorithm of NG learning process is same as SOM. In the point of competitive process, each unit is decided ranking by using eq. (18).

$$rank(k_i) = rank(E_i^k) \tag{25}$$

TABLE II VARIABLE USED EXPLANATION OF MNSOM ALGORITHM

Symbol	***Quantity***
υ	Reference module vector
x	Input vector
y	Output vector
j	Index expressing classes (j = 1, ... , J)
m	Index of module
Em	Distance between input vector and reference vector
m^*	Best Matching Module (BMM)
ϕ	Neigubor Function
$d(a,b)$	Euclidean distance between a and b
ψ	Learning late
σ	Neighbor radius
τ	Time constant

And, in the cooperative process, learning rate is changed to eq. (26)

.

$$\phi_i^k = \exp\left(rank(k_i) \Big/ 2\sigma^2 \right) \tag{26}$$

Equation of (26) substitute eq. (22), and each units are learned by eq. (24). On the processing mode, the algorithm of processing mode is same as SOM.

D Color constancy algorithm using modular network SOM

Modular network SOM (mnSOM) is extended to conventional SOM. The idea of mnSOM is very simple; each vector unit of a conventional SOM which is arrayed on a 2-dimentional lattice, is replaced by a function module of a neural network (e.g. Multi-layer perceptron (MLP), SOM, etc.) [17]. mnSOM inherits all other properties of the conventional SOM. The mnSOM strategy has several advantages. First, the application targets are widely expanded from fields involving just vectorized data to those dealing with more general classes of datasets relevant to functions, systems, time series and so on [18].

The learning process of mnSOM consists of 4 processes. Table II shows the functions and parameters of using mnSOM. In our algorithm, the basic mnSOM does not have reference vectors. However, our algorithm has reference vectors same dimensions as input vector x, and neural network is used MLP which is learning under the Back-Propagation algorithm. In this time input is used input vector x, and teaching data is used output vector y.

(a) Evaluative process.

At first, input vector x is compare to the module's vector u (eq.(26)).

$$Em_j^m = \left\| u^m - x_j \right\|^2 \tag{26}$$

(b) Competitive process

The minimum of distance eq. (26) is decided as Best Matching Module (BMM).

$$m_j^* = \arg_m \min Em_j^k \tag{27}$$

(c) Cooperative process

By using each BMM, each neighbor hood module are given to the learning rate which is based on Gaussian function (eq.(28)).

$$\psi_j^m(t) = \phi(d(m,m_j^*)/\sum_{j'=1}^{J}\phi(d(m,m_{j'}^*) \tag{28}$$

$$\phi(d(m,m_j^*);t) = \exp\left[-d(m,m_j^*)^2/2\sigma^2(t)\right] \tag{29}$$

(d) Adaptive process

By using the each learning rate, the MLP and reference vector are learned.

$$\Delta w^m = -\eta\sum_{j=1}^{J}\psi_j^m(t)\frac{\partial E_j^m}{\partial w^m} = -\eta\frac{\partial E^m}{\partial w^m} \tag{30}$$

$$u^j(t+1) = \sum_j \psi_j^m x_j \tag{31}$$

The processing mode, on the evaluative process, input from the robot evaluates the reference vector, and competitive process decides the BMU. The output is calculated by using BMU's MLP.

III. EXPERIMENT

In order to evaluate the performances of color constancy algorithms, 17 sets of teaching data in various lighting conditions are prepared for learning. Table III shows the lighting conditions where the data are sampled and their luminance. The data are measured in fluorescent light (FL), fluorescent light sunshine (FL and SUN), white mercury lamp (WM), orange mercury lamp (OM) and under sunshine (SUN) and their luminance change from 10 to 18640 [lx]. The data sets in table IV are the test data for performance evaluation,

TABLE III ENVIRONMENT OF THE LEARNING DATA

Light environment (Location)	***Illuminance [lx]***
FL (RoboCup room)	10
FL (RoboCup room)	53
FL (RoboCup room)	110
FL (RoboCup room)	453
FL and SUN (RoboCup room)	81
FL and SUN (RoboCup room)	90
FL and SUN (RoboCup room)	450
FL and SUN (Entrance)	25
FL and SUN (Entrance)	45
FL and SUN (Entrance)	125
FL and SUN (Entrance)	252
SUN (Outside)	228
SUN (Outside)	18640
SUN (Gym)	35
WM and SUN (Gym)	113
OM and SUN (Gym)	142
WM, OM and SUN (Gym)	216

SUN: Sunshine, FL: fluorescent light,
WM: white mercury lamp, OM: orange mercury lamp)

TABLE IV ENVIRONMENT OF THE NON LEARNING DATA

Light environment (Location)	*Illuminance [lx]*
FL (RoboCup room)	310
FL and SUN (RoboCup room)	69
SUN (RoboCup room)	2
FL and SUN (Entrance)	337
FL and SUN (Entrance)	168
SUN (Entrance)	30
SUN (Outside)	58
SUN (Outside)	980

SUN: Sunshine, FL: fluorescent light,
WM: white mercury lamp, OM: orange mercury lamp)

TABLE V RESULT OF EXPERIMENT

	SOM	*NG*	*mnSOM*
Learn speed [sec]	27.9	17.1	81.5
Error	0.0007	0.0000	0.0367
Execution speed [fps]	16.3	17.1	8.7
Recognize Rate[%]	83.3	85.4	85.4

not used in the learning process. The learning times for each algorithms are 30000.

Table V shows the comparison of the learning results. With 30000 times learning, the errors of SOM and Neural Gas based algorithms converged to almost zero, but in mnSOM algorithm still error remains. The execution speed means how many frames can be processed online in the RoboCup robot. The SOM and NG based algorithms can calculate 2 times more faster than mnSOM based algorithm. Regards to the recognition rate, NG and mnSOM based algorithms show the best performance.

In this problem, the NG based algorithm shows the best performance. mnSOM, an expansion of conventional SOM shows also good performance, however, this color constancy problem is not so complex as the mnSOM supposes.

IV. RESULT

In this paper, we propose and develop the color recognition system of using the bio-inspired algorithm. Considering about mounting the autonomous mobile soccer robot, NG algorithm is the best of color detection algorithm.

In this algorithm, the algorithm detect the thresholds under the non-linear light environment, however, in the future work, the algorithms of bio-inspired can also adjust the camera parameters (Iris, white balance, gamma parameter and so on.)

ACKNOWLEDGMENT

Many people and graduated students have helped us along the way. Their guidance, good humor, advice, and inspiration sustained us through the years of work. And, this work was supported by the 21st Century COE program, Ooba sangyou Co., Ltd., Nippon telephone center corporation, Neural image Co., Ltd., and TOSHIBA.

REFERENCES

[1] RoboCup Homepae, http://www.robocup.org
[2] H.Matsubara, M.Asada, et. al., The RoboCup: The Robot World Cup Initiative, In Proceedings of IJCAI-95
[3] Y.Takemura, A.A.F. Nassiraei, et. al., “HIBIKINO-MUSASHI”, RoboCup 2007 Atlanta, CD-ROM Proc. of RoboCup 2007
[4] A.A.F.Nassiraei, Y.Takemura, A.Sanada, et. al., “Concept of Mechatronics Modular Design for an Autonomous Mobile Soccer Robot, CIRA 2007, Jacksonville, pp.178-183
[5] T.Kohonen, “Self-organized formation of topologically correct feature maps”, Biol. Cybernetics, vol.43, pp.59-69, 1982
[6] Saalbach G.Heidemann and H.Ritter, “Parametrized SOMs for Object Recognition and Pose Estimation”, Proc. of ICANN 2002, pp.902-907, 2002
[7] Barreto, G.A., Araujo, A.F..R. Ducker. C. and Ritter. H., “A distribution of a simulator for robots and its application to position estimation by self-organization”, Technical Report of IEICE, Vol.100, No.687, NC2000-147,pp.149-156
[8] K.Ishii, S.Nishida, T.Ura, “A Self-Organizing Map Based navigation System for an Underwater Vehicle”, Proc. of ICRA’04, pp.4466-4471, 2004
[9] K.Azeura, I.Godler, “Color Sampling with Omni-Directional Camera in Robot Soccer”, Proc. of RSJ 2006, Okayama, 1B25.pdf, 2006, In Japanese
[10] T.Kohonen, “The self^organizaing map,” Proceedings of the Instituite of Electrical and Electronics Engineers, vol.78, pp.1460-1480
[11] Lo,Z.-P., M.Fujita and B.Bavarian, “Analysis of neighborhood interaction in Kohonen neural networks”, 6th International Parallel Processing Symposium Proceedings, pp.247-249, Los Alamitos, CA, 1991
[12] Lo, Z.-P., Y.Yu and B.Bavarian, “Analysis of the convergence properties of topology preserving neural networks”, IEEE Transactions on Neural Networks, vol.4, pp.207-220, 1993
[13] Erwin, E., K.Obermayer and K.Schulten, “I:Self-organizing maps: Stationary states, metastability and convergence rate”, Biological Cbernetics, vol.67, pp.47-55
[14] Ritter, H., T.Martinets and K.Schulten, Neural Computation and Self-Organizing Maps: An Introduction, Reading, MA: Addison-Wesley
[15] Obermayer, K., H.Ritter and K.Schulten, “Development and spatial structure of cortical feature maps: A model study”, Advances in Neural Information Processing Systems, vol.3, pp.11-17, San Mateo, CA: Morgan Kaufmann
[16] Thomas M. Martinets, Stanislav G. Berkovich and Klaus J. Schulten, “”Neural Gas” Network for Vector quanitization and its Application to Time-Series Pred iction”, IEEE Transactions on neural networks, Vol.4, No.4, July, 1993
[17] T.Tokunaga, T.Furukawa and S. Yasui, “Modular network SOM:Extention of SOM to the realm of function space”, Proc. Of Workshop on Self-Organizing Map 2003 (WSOM2003), Kitakyushu, Japan, 2003, pp.173-178
[18] T.Furukawa, T.Tokunata, K.Morishita, S.Yasui, “Modular network SOM (mnSOM): From vector space to functio space”, Lecture Notes in ComputerScience, Vol.3696, pp.391-396, 2005