ASYMPTOTIC STATE OF ONE-DIMENSIONAL SOM AT RAYLEIGH POINT DENSITY INPUT

Ivo R. Draganov, Antoaneta A. Popova

Radiocommunications and Videotechnologies Dept.
Technical University - Sofia
8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria
Tel.: +359 2 965 2274; E-mail.: {idraganov, antoaneta.popova}@tu-sofia.bg

Abstract

In this paper an analysis is presented concerning the asymptotic state of the one-dimensional self-organizing map (SOM) with finite grid in the case of Rayleigh point distribution input. The main goal is to find the diversion of the neurons' location after a certain number of epochs. The SOM distortion measure is analyzed with its value found approximately using Taylor series. The stationary values of the statistical expectations covered by the neurons are found solving a set of non-linear equations. Also the objective function of the SOM is found along with its gradient and using gradient-descent approach the minimum of the distortion measure is calculated. Based on the values obtained useful tips for proper initialization of the SOM in this case are given. The results are considered useful enough in wide variety of practical cases in telemedicine, image processing, optical communications and other areas.

1. INTRODUCTION

It is well known fact that the area allocated for storing the most important feature set inside a self-organizing map (SOM) is proportional to the frequency of occurrence of that very same feature in the observations [1].

So far an investigation of the point density for the linear map is led in the presence of a very large number of codebook vectors over a finite area for linear, linear-quadratic and quadratic distributions [2], [3]. It is revealed that the asymptotic point density is proportional to the probability of a certain feature vector occurring raised to some exponent depending on the number of neighbors including the winning neuron and some scalar factor. Similar research on the change of this power is done in [4] when the neighbor function is Gaussian kernel and its normalized second moment is independent variable. The resulting range for the power value in this case is from 1/3 to 2/3. Similar results are presented in [5].

Some more recent researches concern the asymptotic state of the SOM at normal [6] and distorted normal distributions when the input passes at first through non-linear channel [7] where the power value range is found to be wider.

Here the influence of the Rayleigh point density of the input over the asymptotic state of a finite one-dimensional SOM is investigated with its distortion measure. In part 2 theoretical analysis is presented and in part 3 some experimental results are given. In part 4 a conclusion is made.

2. SOM ANALYSIS WITH RAYLEIGH POINT DENSITY INPUT

Let one-dimensional feature space of *x* is considered. The number of points must be large enough (e.g. by criteria given in [1]) and they must be stochastic variables so their probability density *p(x)* could be defined. The codebook vectors m_i usually form regular optimal configuration and thus can not be stochastic. Their number is typically low in any cluster as well.

2.1. Asymptotic State of the One-Dimensional Finite-Grid SOM

Let suppose m_i and m_{i+1} are two neighboring points. A way of defining the point density is as $(m_{i+1} - m_i)^{-1}$ but it does not cover the samples around the boundaries of the clusters for which this density does not have meaning. So a better way of defining it is as the inverse of the width of the Voronoi set $[(m_{i+1} - m_i)/2]^{-1}$. The input consists of samples $x(t) \in \Re, t = 0,1,2,...$ while the codebook is $m_i(t) \in \Re, t = 0,1,2,..., i = 1,...,k$. The one-dimensional SOM algorithm with at least one neighbor at each side is given by [1]:

$$\begin{gathered} m_i(t+1) = m_i(t) + \varepsilon(t)[x(t) - m_i(t)], \quad \text{for } \mathrm{i} \in \mathrm{N_c} \\ m_i(t+1) = m_i(t) \quad \text{for} \quad \mathrm{i} \notin \mathrm{N_c} \\ c = \arg\min_{\mathrm{i}} \{|x(t) - m_i(t)|\}, \\ N_c = \{\max(l, c-1), c, \min(k, c+1)\}, \end{gathered} \tag{1}$$

where N_c is the neighbor set around node c and $\varepsilon(l)$ is the learning-rate factor. The Voronoi set V_i around m_i is defined as:

$$V_i = \left[\frac{m_{i-1}+m_i}{2}, \frac{m_i+m_{i+1}}{2}\right], V_1 = \left[0, \frac{m_1+m_2}{2}\right],$$
$$\mathrm{V_k} = \left[\frac{m_{k-1}+m_k}{2}, 1\right], \ \text{for} \ 1 < \mathrm{i} < \mathrm{k}, \tag{2}$$
$$U_i = V_{i-1} \cup V_i \cup V_{i+1}, U_1 = V_1 \cup V_2,$$
$$\mathrm{U_k} = V_{k-1} \cup V_k, \ \text{for} \ 1 < \mathrm{i} < \mathrm{k}$$

In this case U_i is the set of such $x(t)$ which provoke changes in $m_i(t)$ during one learning step. Following (1) and (2) we get to the well known stationary equilibrium for m_i coinciding for the general case [1]:

$$m_i = E\{x \mid x \in U_i\}, \forall \mathrm{i} \ . \tag{3}$$

In other words every m_i becomes centroid of the probability mass for each U_i and then for 2 < i < (k-1) the limits for U_i are:

$$A_i = \frac{1}{2}(m_{i-2} + m_{i-1}),$$
$$B_i = \frac{1}{2}(m_{i+1} + m_{i+2}) \ . \tag{4}$$

For i =1 and i = 2, A_i = 0, and for $i = k - 1$ and $i = k$, B_i = 1.

The case investigated here concerns input data with the following distribution:

$$p(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} . \tag{5}$$

As (5) is too complex to be used in finding the centroids of the probability masses, Taylor series are used instead:

$$p(x) = \sum_{i=0}^{\infty} \frac{p^{(i)}(0)}{i!} x^i . \tag{6}$$

We find the absolute difference between the third and second order approximations:

$$\Delta p_{32}(x) = \sum_{i=0}^{3} \frac{p^{(i)}(0)}{i!} x^i - \sum_{i=0}^{2} \frac{p^{(i)}(0)}{i!} x^i = \frac{x^3 p'''(0)}{6} \ . \tag{7}$$

Since:

$$p'(x) = \frac{e^{-x^2/2\sigma^2}}{\sigma^2}\left(1 - \frac{x^2}{\sigma^2}\right),$$
$$p''(x) = -\frac{x e^{-x^2/2\sigma^2}}{\sigma^4}\left(3 + \frac{x^2}{\sigma^2}\right), \tag{8}$$
$$p'''(x) = \frac{e^{-x^2/2\sigma^2}}{\sigma^4}\left(\frac{x^4}{\sigma^4} - 3\right)$$

then $p(0) = 0$, $p'(0) = 1/\sigma^2$, $p''(0) = 0$, $p'''(0) = -3/\sigma^4$, and $\Delta p_{32}(x) = -x^3/2\sigma^4$.

Now for 5 typical cases of σ the error Δp_{32} is found and the results are presented in Fig. 1. It is visible that only for σ = 0.5 the error between the second and third approximation exceeds considerably 1 by module and this in such a wide range for x from 0 to 10. So it is reasonable to use approximation for the original distribution of second order that is $p(x) \approx x/\sigma^2$.

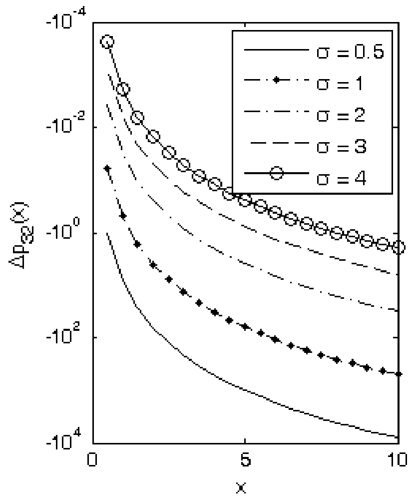

Fig. 1. The absolute error between approximations of second and third order

The stationary values of the m_i are defined by:

$$m_i = E\{x \mid x \in U_i\} = \frac{\int_{A_i}^{B_i} x p(x) dx}{\int_{A_i}^{B_i} p(x) dx} = \frac{2(B_i^3 - A_i^3)}{3(B_i^2 - A_i^2)}, \forall i \ . \tag{9}$$

But they could be expressed in even simpler way - it consists of defining the point density q_i around m_i as the inverse of the length of the Voronoi set – $q_i = [(m_{i+1} - m_{i-1})/2]^{-1}$. As a result of that q_i can be expressed in the form $const.[p(m_i)]^{\alpha}$. Then passing from m_i to m_j it is true:

$$\alpha = \frac{\log(m_{i+1} - m_{i-1}) - \log(m_{j+1} - m_{j-1})}{\log[p(m_j)] - \log[p(m_i)]} . \quad (10)$$

For improved accuracy more values of the m_i are needed as we shall see in the next section.

2.2. Finding the One-Dimensional SOM Distortion Measure with Finite Grids

The objective function of the SOM is given by [1]:

$$E = \sum_i \sum_j \int_{x \in V_i} h_{ij} \left\| x - m_j \right\|^2 p(x) dx , \quad (11)$$

where V_i is the Voronoi set around m_i and h_{ij} is defined as:

$$h_{ij} = \begin{cases} 1, & \text{if } |i-j| < 2 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

and i and j run over all the values defining h_{ij}.

Then (11) becomes:

$$E = \sum_i \sum_j \int_{C_i}^{D_i} (x - m_j)^2 p(x) dx =$$
$$= \sum_i \sum_j \frac{1}{\sigma^2} [\frac{(D_i^4 - C_i^4)}{4} - \quad ,(13)$$
$$- \frac{2m_j(D_i^3 - C_i^3)}{3} + \frac{m_j^2(D_i^2 - C_i^2)}{2}]$$

where N_i is defined in (1) and the borders C_i and D_i of the Voronoi set V_i are:

$$\begin{aligned} C_1 &= 0, \\ C_i &= \frac{m_{i-1} + m_i}{2} \quad \text{for } 2 \le \mathrm{i} \le \mathrm{k}, \\ \mathrm{D_i} &= \frac{m_i + m_{i+1}}{2} \quad \text{for } 1 \le \mathrm{i} \le \mathrm{k-1}, \\ \mathrm{D_k} &= 1. \end{aligned} \quad (14)$$

3. EXPERIMENTAL RESULTS

As a simulation environment we use Matlab® R2009B over MS® Windows® XP® Pro SP3.

First α from (10) is found for different number of grid points. The more m_i are used the more accurate are the results. For $i = 4$ and $j = k - 3$ assuring negligible border effects 10, 25, 50, and 100 grid points are used. The same experiment is done with normally distributed input points in [6], so here a direct comparison can be made. The results are given in Table 1.

Table 1. Experimentally estimated α for two different distributions of the input

Grid points	Exponent α	
	Normal, [6]	Rayleigh
10	0.2989	0.3480
25	0.3330	0.3495
50	0.3331	0.3501
100	0.3330	0.3509

It is clearly seen that the exponent approximation is presented here by higher *α* which is actually expected because of the steeper left slope of the Rayleigh curve in comparison to the symmetric Gaussian one.

Obviously the values obtained for the Rayleigh distribution almost do not depend on the number of grids. Now when we have the real case m_i found it is seen that the exponent of the approximated state of the SOM is close to 1/3. This is actually a case strongly related with the optimal vector quantization [1].

Graphically the results from Table 1 are given in Fig. 2.

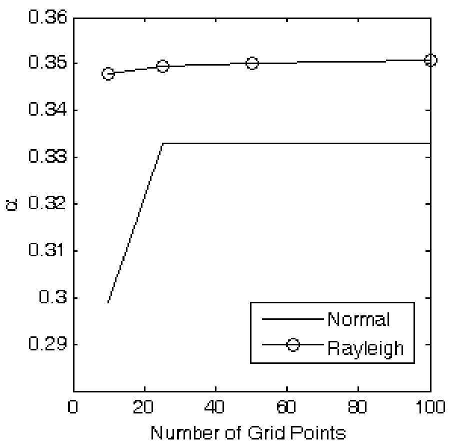

Fig. 2. Experimentally derived *α* as a function of the number of grid points for two different distributions of the input for the SOM

4. CONCLUSION

In this paper an approach for finding the stationary positions of the nodes of one-dimensional SOM has been presented in the case of Rayleigh density point input. The results are precise enough taking the advantage of very fast computation. Furthermore the distortion measure of the SOM using finite grid is calculated in the general case and it is shown that the positions of the nodes could be optimized at the stage of initialization.

The results achieved prove the correctness of the suggested approach which is considered useful in a large number of practical cases where the input data poses Rayleigh point density distribution.

5. ACKNOWLEDGMENT

This work is supported by Research and Development Sector (R&DS) with TU-Sofia under Grant № 102pd177-7 "Creation of models and object segmentation, applied to medicine and education".

References

[1] T. Kohonen, *Self-Organizing Maps*, 3 ed., Springer, New York, 2001.

[2] H. Ritter and K. Schulten, "On the Stationary State of Kohonen's Self-Organizing Sensory Mapping", *Biological Cybernetics*, Vol. 54, Issue 2, June 1986, pp. 99-106.

[3] H. Ritter, "Asymptotic Level Density for a Class of Vector Quantization Processes", *IEEE Transactions on Neural Networks*, No. 2, 1991, pp. 173-175.

[4] D. Dersch and P. Tavan, "Asymptotic Level Density in Topological Feature Maps", *IEEE Transactions on Neural Networks*, Vol. 6, No. 1, 1995, pp. 230-236.

[5] S. Luttrell, "Code Vector Density in Topographic Mappings: Scalar Case", *IEEE Transactions on Neural Networks*, Vol. 2, No.4, 1991, pp. 427-436.

[6] Draganov, "Normal Point Density Input Influence over a Finite One-Dimensional SOM", *Proc. of ICEST'2010*, Ohrid, Macedonia, June 23-26, 2010, pp. SP.II.17-20.

[7] Draganov, A. Popova, and N. Neshov, "Asymptotic State of One-Dimensional SOM at Normal Point Density Input Passed Through Non-Linear Channel", *Proc. 14th WSEAS Conf. on Computers*, Vol. 2, Corfu Island, Greece, July 23-25, 2010, pp. 761-765.