INVARIANCE OF GAUSSIAN-VECTOR MAPPING USING A SELF-ORGANIZING MAP

Igor Farkaš *

Abstract: We study the task of ordered mapping of Gaussian-vector patterns with dot-product version of SOM. Specifically, we investigate the invariance property of such a mapping, in one- and two-dimensional cases, with respect to the standard deviation σ of Gaussians, as well as the variability of σ. For the purpose of evaluation of the mapping performance, the *peak position misplacement* (*PPM*) coefficient is introduced. We provide quantitative results, based on simulations, of the lower limit of σ for both cases, below which ordered mapping is not guaranteed. Further, the mapping is shown to be invariant to variability of σ to a certain degree.

Key words: *self-organizing map, Gaussian patterns, peak position mapping, invariance*

1 Introduction

With regard to representation of input patterns, we may distinguish two basic feature mapping models using a self-organizing map (SOM, see e.g. [1]). The first one, originally proposed by Willshaw and von der Malsburg [2], is biologically motivated. It employs two separate two-dimensional lattices of neurons connected together, with one projecting onto the other. The second model [3] is an artificial one, but computationally more effective due to much lower dimensionality of input representation. Both approaches have been used e.g. in modelling the somatotopic map, i.e. projecting the local touch stimuli from the body surface onto the cortex. In the latter approach [4], each touch stimulus was represented by a couple of coordinates of its "center of gravity", so the dimension of input space in this case is actually only two. In the former approach [5], a stimulus pattern was modelled by a Gaussian activity profile, centered at some location in input sheet. From the geometrical viewpoint, the one-to-one mapping is feasible also when we use these high-dimensional inputs because their *inherent dimensionality* remains unchanged. This must be equal to one, e.g. for the tonotopic map

*Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 842 19 Bratislava, Slovakia

(mapping the frequency of acoustic stimuli) [3], or two, for the somatotopic map [5].

In this paper, we focus on the task of mapping these Gaussian patterns with SOM. By a "Gaussian pattern" we shall thus mean a vector with a single maximum component (peak center) and with the other (N-1) components decreasing in magnitude in such a way that the overall activity shape resembles the sampled Gaussian function. The ordered mapping is empirically known to occur for a relatively free shape of Gaussians. On the other hand, for simple reason we may expect that the ordering effect will fail, if the Gaussians are too sharp (i.e. with a very small overlap): whereas in "ordinary" case the inherent dimensionality of input set may be one, in the latter case, the Gaussians will approximate the set of basis vectors of N-dimensional space, which results in their dimensionality equal to N. [1]

To get some insight into the ordering performance, we investigate by simulation the invariance property of the mapping Gaussian patterns with SOM. First, we study the influence of Gaussian variance σ^2, which parametrizes the Gaussian sharpness (for convenience, we use standard deviation σ as a parameter). Second, we empirically test the mapping invariance with respect to the variability of Gaussian shape. In both subtasks, 1D and 2D cases of the mapping are considered. The topology preservation between input and output spaces is quantified by *peak position misplacement* (*PPM*) coefficient which is introduced.

2 Influence of the Gaussian variance

In the simplest, 1D case the Gaussian-vector input stimuli are modelled by sampling the Gaussian function of the form

$$g_\sigma^{(p)}(r) = A.\exp\{-(r-p)^2/2\sigma^2\} \qquad \text{for } r = 1, 2, ..., N \tag{1}$$

where $p \in \{1, 2, ..., N\}$ is the position of the centre of the stimulus along the chain of N receptors (neurons), $g_\sigma^{(p)}(r)$ represents the magnitude of excitation at position r, A is the peak magnitude. Thus for a fixed parameter, standard deviation σ, we have a set of N-dimensional vectors $\mathbf{g}_\sigma^{(p)} = [g^{(p)}(1), g^{(p)}(2), ..., g^{(p)}(N)]^T$, which differ in the position of their peak (centre of the stimulus). [2] Several half-profiles of such Gaussians with 100 components are plotted in Fig.1 for various values of σ.

[1] It should be mentioned that the observed ordering effect does not relate only to Gaussian patterns. Qualitatively the same situation would occur with other overlapping patterns such as triangular- or box-shaped ones.

[2] For simplicity, we consider the position of an excitation peak to have only discrete values, i.e. every stimulus finds itself in the center of the receptive field of one of the N receptors.

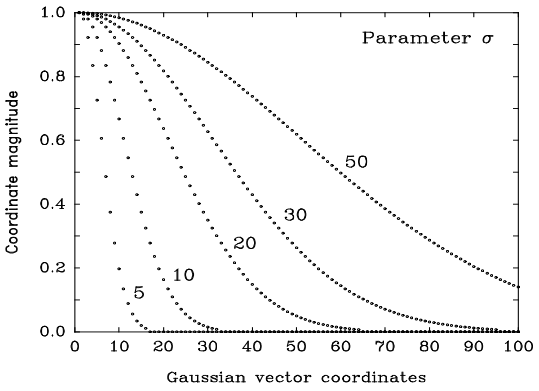

Figure 1: *Several shapes of the Gaussian half-profile plotted for various values of* σ.

For the mapping to succeed, it is necessary that the set of Gaussians (points in N-space) possesses the property of the internal order, with respect to peak positions. Similarity between two vector patterns is commonly measured by their dot product (DP) or the inverse of Euclidean distance (ED) between them. With regard to the either similarity measure it is thus required, that *the more similar the two Gaussian patterns are* (their peaks are closer to each other), *the higher value of the similarity measure should they yield.* Analytical calculation of the above dependence would be rather difficult, therefore we evaluated it using simulations for a set of Gaussians with $N = 100$ components. The dependences for both DP and ED are shown in Fig.2.

In both cases, the x-axis denotes the (discrete) distance between the peaks of the two Gaussians. The y-axis means the DP (Fig.2a) and ED (Fig.2b) of the two Gaussians in N-space. As seen in Fig.2a, for most values of parameter σ, the function of dependence monotonously decreases with increasing $|p - q|$. Consistent, "reverse" behaviour can be observed in Fig.2b. This dependence supports the hypothesis about the existence of an underlying (nonlinear) 1D manifold that could be "passing" through the Gaussians — points in N-space. In addition, *these points appear to be ordered along the manifold with respect to the peak position.*

This hypothesis has been tested by training a 1D SOM on these Gaussians. As known, dimensionality match between input and output spaces is a necessary condition for SOM to be able to topologically map the input space. In this case, the mapped feature of input patterns was expected to correspond to the peak position.

Rather than Euclidean distance we applied the dot-product version of the Kohonen algorithm (DP-SOM) which employs a common input-output

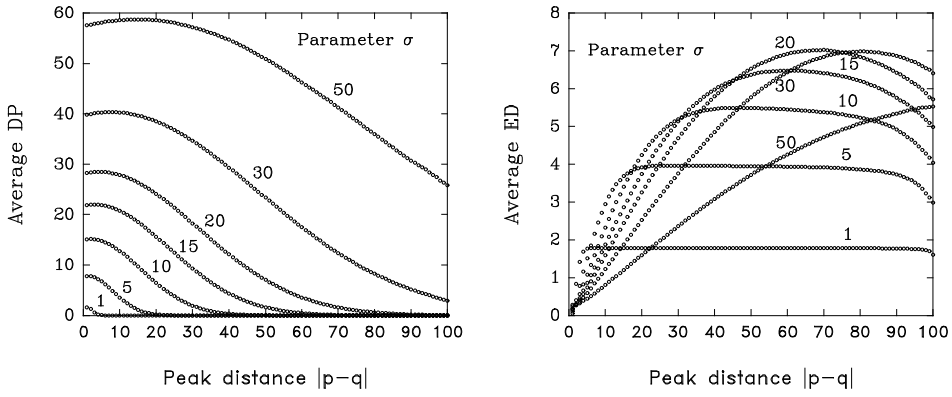

Figure 2: *Average (a) Dot Product of and (b) Euclidean Distance between two Gaussian patterns* $\mathbf{g}^{(p)}$ *and* $\mathbf{g}^{(q)}$ *of dimension* $N = 100$ *(with random amplitudes* $A \in \langle 0.9, 1.0 \rangle$*) as a function of distance of their peaks, for various values of* σ*. For each* σ*, 500 vectors were generated in total (i.e. 5 for each peak position), from which the averages (DP and ED) were computed. Standard deviation of averages is not plotted, as it appeared to be insignificant.*

function of a linear neuron, i.e. the output of the i-th neuron with associated weight vector $\mathbf{w}_i$ is taken as $y_i = \mathbf{w}_i^T.\mathbf{x}$, where $\mathbf{x}$ is a Gaussian pattern $\mathbf{g}_\sigma^{(p)}$. [3] The weight adaptation rule has the form

$$\mathbf{w}_i(t+1) = \frac{\mathbf{w}_i(t) + \alpha(t).h(i^*, i).\mathbf{x}}{\|\mathbf{w}_i(t) + \alpha(t).h(i^*, i).\mathbf{x}\|} \tag{2}$$

with $\alpha(t)$ denoting the common learning rate and $h(i^*, i)$ being the (rectangular) neighbourhood function. At first, the performance of the mapping is studied for various, but constant values of σ. In order to quantitatively evaluate the mapping order obtained after training, we define a coefficient — *peak position misplacement* (*PPM*). It is updated after presentation of each test pattern and eventually, its mean si taken. The evaluation of the coefficient is based on comparing the position $p \in \{1, 2, ..N\}$ of maximum excitation in an input Gaussian pattern $\mathbf{g}_\sigma^{(p)}$ and the corresponding position i^* (index of the winner) of the maximum output response for that pattern in SOM (in the range 1 to n, where n is the number of neurons). Namely, the final value is taken as

$$PPM = \sqrt{E[(q.i^* - p)^2]} \tag{3}$$

[3] It is well known that from the point of view of topological mapping both ED and DP versions are equivalent as long as in the latter case the weight vectors are kept normalized.

with $E[.]$ denoting the mean value. The ratio coefficient $q = N/n$ is included in order to match the lengths of output and input arrays. (Without loss of generality, in all our simulations we considered $N = n$). It may be obvious that *PPM* quantifies the average misplacement of the position of the maximum output response from the optimal one. As a consequence, the closer *PPM* is to zero, the better mapping has resulted.

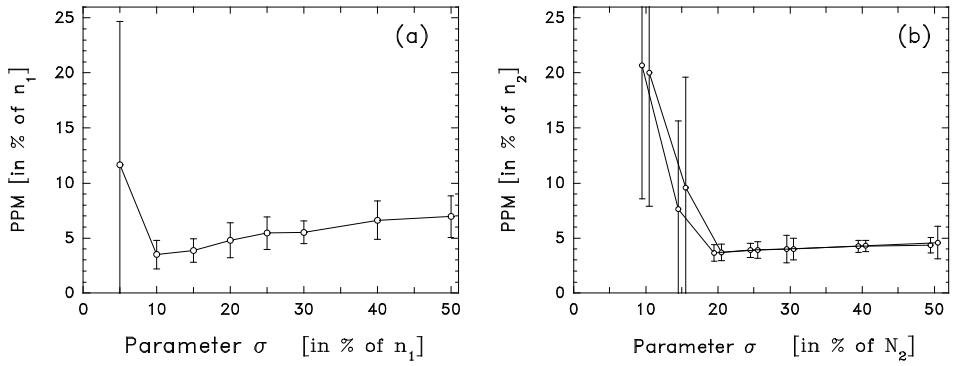

Figure 3: *Average values of PPM in the peak-position-mapping task with DP-SOM. (a) 1D case:* $N_1 = 100$, $n_1 = 100$ *neurons in 1D DP-SOM. SOM is seen to yield the lowest PPM (approx. 3.5% of* n_1*) for* $\sigma = 10\%$ *of* N_1*. (b) 2D case:* $N_2^2 = 400$, $n_2^2 = 400$ *neurons in 2D DP-SOM. Minimum PPM (approx. 3.7% of* n_2*, for both sheet coordinates) in this case occurs for* $\sigma = 20\%$ *of* N_2*, for both coordinates.*

Fig.3a shows the results of 1D simulations. These were run 20 times for each of 8 different values of σ (5, 10,...). Each of the 8 training sets consisted of 500 Gaussian vectors (of dimension $N = N_1 = 100$), whose components were generated according to Eq.1. Parameters of DP-SOM containing $n = n_1 = 100$ neurons were common: $t_{max} = 10000$ iterations, $\alpha_{init} = 0.8$, initial neighborhood radius $= n_1/3$.

A simulation yielding a low value of *PPM* (empirically chosen $PPM < 10$) was considered as an ordered state. In addition, ordering was systematically checked by visual inspection of SOM responses. Ordering was obtained in almost all cases of σ, with the exception of $\sigma = 5$ (SOM failed 7 times of 20). For this reason, standard deviation of *PPM* obtained for $\sigma = 5$ is very large (see Fig.2a). As seen, there exists some limit σ_{min} (according to simulations, somewhere in the interval $\langle 5, 10 \rangle$), below which the ordering can, but may not succeed. This must be due to the fact, that with decreasing σ below σ_{min} the hypothetical 1D manifold transforms to a set of mutually distant clusters positioned close to coordinate axes, resulting in inherent dimension of the set equal to N. Towards the opposite

end, increasing σ only *slightly deteriorates* the ordering property of SOM. It does even occur, if the activity of an individual input pattern spreads all over the input array (cf. Fig.1, the Gaussian half-profile for $\sigma > 20$).

Analogical simulations were performed for the 2D case. Gaussians were generated using the 2D modification of Eq.1, with $N = N_2^2 = 400$ components each. DP-SOM was set to have $n = n_2^2 = 400$ neurons arranged in a 2D sheet. *PPM* was evaluated as in 1D case, now separately for both sheet coordinates (therefore, in Fig.3b there are two connected lines). All SOM parameters were the same as in 1D case, the initial neighborhood radius was set to $n_2/3$. In order to make both 1D and 2D cases comparable, σ's were made relative (i.e. expressed in % of) to input dimension(s): i.e. relative to N_1 in 1D case, to N_2 in 2D case, and *PPM*s were expressed in % of output dimension(s) (SOM size): n_1 in 1D case, n_2 in 2D case. In this way, for 1D the obtained values remain the same, for 2D they become the 5-multiple of the computed ones. The dependences are qualitatively the same as 1D case, for both coordinates.

3 Influence of variability of the Gaussian variance

In order to see how the variability of Gaussian variance affects topological mapping property of DP-SOM, we generated 5 training sets, differing in variability range $\langle \sigma - \Delta\sigma, \sigma + \Delta\sigma \rangle$, each containing 500 patterns. For both cases (1D and 2D), σ was chosen to correspond to the "best" Gaussian case, i.e. the one with σ_{min}. Results are shown in Fig.4a and 4b. In both cases, the mapping performance deteriorates roughly linearly, but still, correct ordering was observed for all considered $\Delta\sigma$'s up to 10% of N_1 and N_2, respectively.

4 Conclusion

The performed simulations confirmed the claim that mapping Gaussian patterns with a DP-SOM is invariant to the width of Gaussians (quantified by σ), provided that their overlap is sufficient ($\sigma \geq \sigma_{min}$). The different values of empirically obtained σ_{min} for 1D (roughly 10% of N_1) and 2D (roughly 20% of N_2) cases may be the consequence of the geometry of corresponding manifolds embedded in N_1- and N_2^2-dimensional spaces. Mapping error for σ_{min}, quantified by *PPM* coefficient, was approximately the same in both cases: 3.5% and 3.7% of SOM size, respectively.

In addition, *PPM* was illustrated to grow roughly linearly with variability of σ (quantified by $\Delta\sigma$), with relatively small slope. As mapping

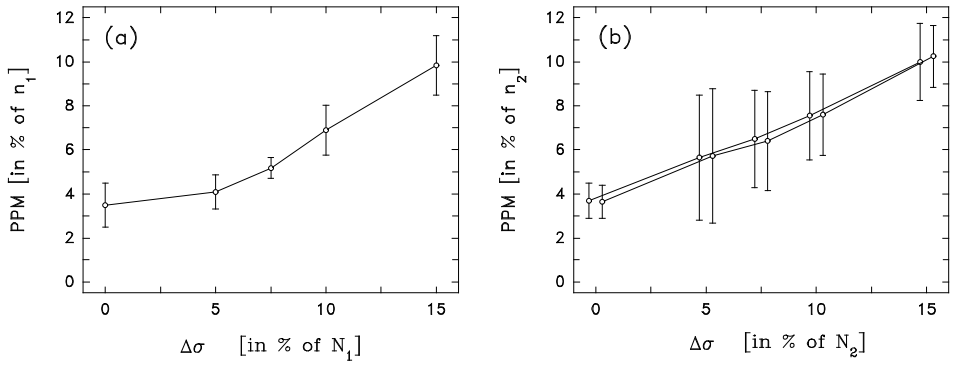

Figure 4: *Average values of PPM (obtained from 20 simulations) for various levels of Gaussian variability. (a) 1D case:* $\sigma = 10\%$ *of* N_1*. (b) 2D case:* $\sigma = 20\%$ *of* N_2*.*

was observed to converge to an ordered state, though with increasing error *PPM*, it can be said to preserve its invariance to $\Delta\sigma$ to a certain degree.

In [6] we suggested that the invariance property of the studied mapping can be exploited in designing a hierarchical feature mapping system consisting of DP-SOM modules. These would generate Gaussian-like patterns as their output representation, optionally tuned by incorporated lateral feedback in order to make the Gaussians fall in invariance range.

Acknowledgement: This work was partially supported by Slovak Grant Agency for Science (grants No. 2/2040/95 and No. 95/5305/468).

References

[1] S. Haykin. *Neural Networks: A Comprehensive Foundation.* Macmillan College Publishing Company, 1994.

[2] D. J. Willshaw and C. von der Malsburg. How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London, Series B*, 194:431–445, 1976.

[3] T. Kohonen. Self-Organizing Maps. *Springer-Verlag*, 1995.

[4] H. Ritter and K. Schulten. On the stationary state of kohonen's self-organizing sensory mapping. *Biological Cybernetics*, 54:99–106, 1986.

[5] K. Obermayer, H. Ritter, and K. Schulten. Large-scale simulations of self-organizing neural networks on parallel computers: Application to biological modelling. *Parallel Computing*, 14:381–404, 1990.

[6] I. Farkaš. *On Vector-Coded Feature Mapping Using Self-Organizing Neural Maps.* PhD thesis, Slovak Technical University in Bratislava, 1995.