

Implementation Issues of Kohonen Self-Organizing Map Realized on FPGA

Rafał Długosz^{1,2}, Marta Kolasa¹, Michał Szulc³,
Witold Pedrycz⁴ and Pierre-André Farine²

1- Faculty of Telecommunication and Electrical Engineering
University of Technology and Life Sciences
ul. Kaliskiego 7, 85-796, Bydgoszcz, Poland

2- Institute of Microtechnology
Swiss Federal Institute of Technology in Lausanne
Rue A.-L. Breguet 2, CH-2000, Neuchâtel, Switzerland

3- Chair of Computer Engineering
Poznań University of Technology
ul. Polanka 3A, 60-965, Poznań, Poland

4- Department of Electrical and Computer Engineering
University of Alberta, Edmonton, AB T6G 2V4, Canada

Abstract. Presented are the investigations showing an impact of the length of data signals in hardware implemented Kohonen Self-Organizing Maps (SOM) on the quality of the learning process. The aim of this work was to determine the allowable reduction of the number of bits in particular signals that does not deteriorate the network behavior. The efficiency of the learning process has been quantified by using the quantization error. The results obtained for the SOM realized on Field Programmable Gate Array (FPGA), as well as by means of the software model of the SOM show that the smallest allowable resolution (expressed in bits) of the weight signals equals seven, while the minimal bit length of the neighborhood signal ranges from 3 to 6 (depending on the map topology). For such values and properly selected values of other parameters the learning process remains undisturbed. Reducing the number of bits has an influence on the number of neurons that can be synthesized on a single FPGA device.

1 Introduction

Artificial Neural Networks (ANNs) have a long history of their hardware implementation, ranging from the Very Large Scale of Integration (VLSI) Application Specific Integrated Circuits (ASICs) with *fully custom* realizations, through Field Programmable Gate Arrays (FPGAs), to numerous purely software based systems. ASIC implementation requires solving specific problems of electronic nature, so it poses a lot of challenge. On the other hand, ANNs implemented in this way offering strong abilities of parallel data processing, can be much faster than their software counterparts, with the power dissipation even four orders of magnitude smaller than the one encountered in PC realizations.

Realization based on FPGA can be viewed as an intermediate solution between the ASIC and the software approach. In this case the ANN is also able to operate in parallel, as it is possible in ASIC designs, but at substantially higher power dissipation that is even two orders of magnitude larger than in ASIC realizations. Nevertheless, there are still numerous applications in which power dissipation is of second importance, while high data rate and short design process are important features. In the comparison with ASIC designs, realizations based on FPGA suffer from some limitations. One of the most important of them is a limited number of neurons that can be synthesized on a single FPGA device. This number depends mostly on the complexity of a single neuron and therefore one of the important optimization directions is to simplify the structure of a single neuron. This can be done, for example, by reducing the number of bits in particular signals used in the ANN. In this paper, we focus on the influence of the number of bits in the neuron weights w and the corresponding input signals x on the quality of the learning process of the Kohonen Self-Organizing maps (SOM). The obtained results find the application in both the FPGA and the ASIC based systems. The number of bits in particular signals creates trade-offs between the energy consumption and the silicon area (in case of ASIC) on the one hand, and the quality of the learning process, on the other one.

The presented results have been obtained for the SOM implemented on the FPGA Virtex XC5VLX110T device. We have realized a complete SOM with full adaptation abilities. This creates an evident advantage in comparison with such FPGA realizations, in which the SOM was realized either with fixed weights, determined e.g. on PC, or with a simplified learning algorithm [1, 2, 3].

2 An influence of particular parameters of the SOM on the quality of the learning process

To determine the influence of the bit lengths of the signals on the quality of the learning process, the series of simulations have been completed using the accurate software model of the SOM and, for selected cases, the FPGA implemented NN. Using the software model was necessary, as not all map sizes can be synthesized on a single FPGA device. The software model enabled determining the optimal values of particular parameters and thus an efficient further hardware realization on the FPGA. The comprehensive simulations have been carried out for three map topologies, namely the rectangular one with 4 and 8 neighbors (Rect4 and Rect8, respectively) and the hexagonal one, for the sizes of the map varying in-between 4x4 and 32x32 neurons, different neighborhood functions and different values of the maximum neighborhood size. The last parameter, R_{\max} , is the value of the neighborhood range at the beginning of the learning process.

The SOM was trained with different data sets. In this paper, we present selected results for 2-D data regularly distributed in the input space. In such case data are divided into P classes (centers) that are located as in a regular rectangular grid. The value of P is equal to the number of neurons in the map. Each data center is represented by an equal number of learning patterns X . Such

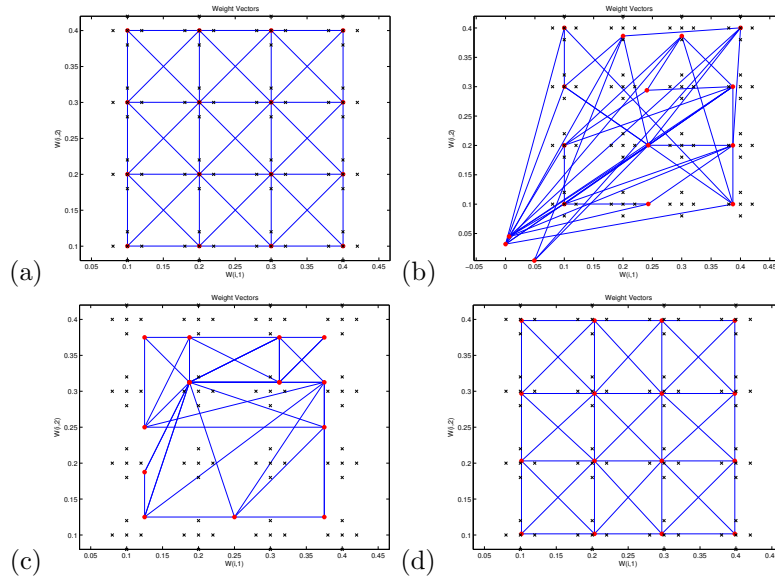


Fig. 1: The quality of the learning process, for the map with 4x4 neurons, reported for the following cases:

- (a) $nb_W = 36 / R_{\max} = 2 / Q_{\text{err}} = 16.2e - 3$,
- (b) $nb_W = 36 / R_{\max} = 1 / Q_{\text{err}} = 30.3e - 3$,
- (c) $nb_W = 4 / R_{\max} = 2 / Q_{\text{err}} = 43.3e - 3$,
- (d) $nb_W = 7 / R_{\max} = 2$ or $R_{\max} = 3 / Q_{\text{err}} = 16.2e - 3$.

data set enables a better illustration and comparison of the results for different network parameters [4]. To achieve comparable results the input space has been fitted to input data i.e. for 8x8 neurons the inputs are in the range of 0 to 1, for 4x4 neurons in the range of 0 to 0.5, while for 16x16 neurons in the range of 0 to 2. As a result, the optimal value of Q_{err} always equals $16.2e - 3$.

The learning process of the SOM has been evaluated by means of the quantization error (Q_{err}), which is a commonly used criterion in such cases. The Q_{err} can be expressed as follows:

$$Q_{\text{err}} = \frac{1}{m} \sum_{j=1}^m \sqrt{\sum_{l=1}^n (x_{j,l} - w_{i,l})^2} \quad (1)$$

where m is the number of the learning patterns, X , in the input data set, n is the number of the NN inputs, while i denotes the winning neuron.

2.1 Results

Selected results for different bit lengths of particular signals for different values of other network parameters, as discussed above, are presented in Figs. 1, 2

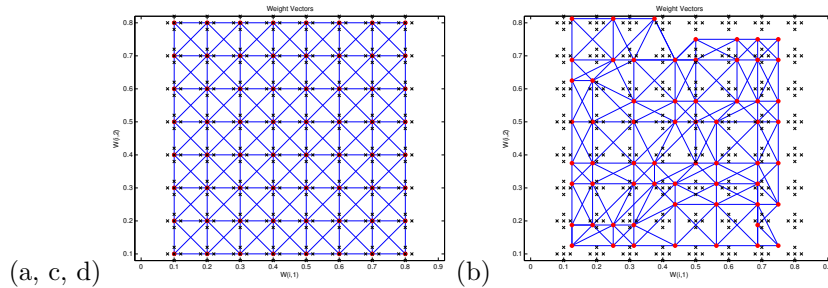


Fig. 2: The quality of the learning process, for the map with 8x8 neurons, reported for the following cases:

- (a) $nb_W = 36 / R_{\max} = 4$ or $R_{\max} = 6 / Q_{\text{err}} = 16.2e - 3$,
- (b) $nb_W = 4 / R_{\max} = 4 / Q_{\text{err}} = 40.3e - 3$,
- (c) $nb_W = 7 / R_{\max} = 6 / Q_{\text{err}} = 18.0e - 3$,
- (d) $nb_W = 7 / R_{\max} = 4 / Q_{\text{err}} = 16.2e - 3$.

and 3, for the ANN with 4x4, 8x8 and 16x16 neurons, respectively. The SOM was trained with the triangular neighborhood function (TNF). The earlier investigations carried out by the authors, presented in [5], show that this function can fully substitute the Gaussian neighborhood function (GNF). A realization of the TNF is much simpler than that of the GNF and therefore this function is in particular suitable for any hardware implementation, both ASIC and FPGA. The other important conclusion discussed in [5] is that even low signal resolution (3 to 6 bits) of the signal at the output of the NF block still allows for a proper performance of the SOM.

In this work, we focus on a new aspect of the complex optimization of the SOM i.e. the influence of the resolution of the neuron weight signals, nb_W , on the learning quality of the SOM. As in this paper, we focus on the FPGA implementation, therefore we are further discussing the influence of this parameter on the number of neurons that can be synthesized on a single FPGA device.

Figures 1–3 present input data and final placement of neurons in the input space for selected cases i.e. for the Rect8 topology, for selected map sizes and different bit lengths of the weight and the input signals. The results of Figs. 1–3 are shown for different initial values, R_{\max} , of the neighborhood range R . The R_{\max} parameter is the neighborhood range used in the first epoch of the learning phase. We have studied the influence of this parameter on the learning process in [5]. It has been demonstrated there that for different input data sets and different other network parameters varying in wide ranges, the optimal values of R_{\max} are usually very small, even for large SOMs with more than thousand neurons. This conclusion is in contrast with a common opinion that R_{\max} at the beginning of training process should be large enough to cover at least half of the map. This conclusion has been confirmed also during the investigations described in this paper for different resolutions of the x and the w signals.

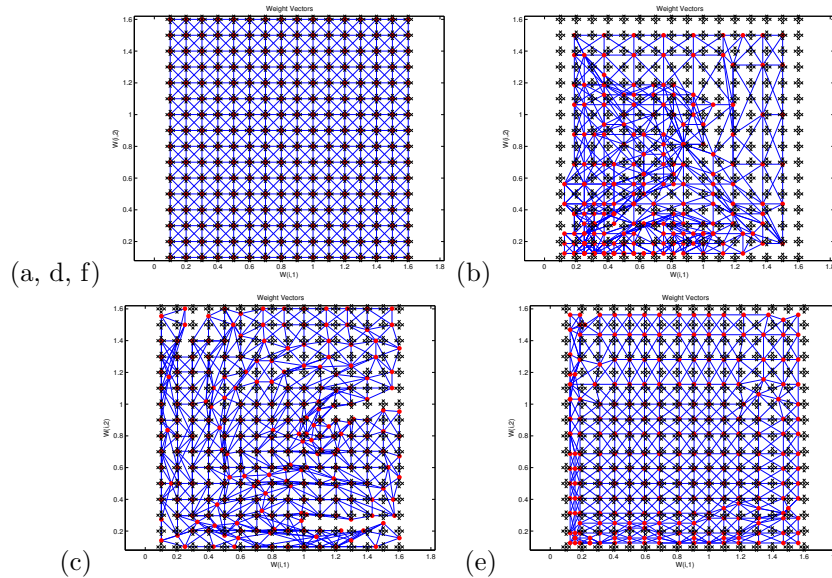


Fig. 3: The quality of the learning process, for the map with 16x16 neurons, reported for the following cases:

- (a) $nb_W = 36 / R_{\max} = 12$ or $R_{\max} = 3 / Q_{\text{err}} = 16.2e - 3$,
- (b) $nb_W = 4 / R_{\max} = 12 / Q_{\text{err}} = 62.3e - 3$,
- (c) $nb_W = 7 / R_{\max} = 12 / Q_{\text{err}} = 28.1e - 3$,
- (d) $nb_W = 15 / R_{\max} = 12 / Q_{\text{err}} = 16.2e - 3$,
- (e) $nb_W = 4 / R_{\max} = 3 / Q_{\text{err}} = 56.3e - 3$,
- (f) $nb_W = 7 / R_{\max} = 3 / Q_{\text{err}} = 16.2e - 3$.

On the basis of the presented results some conclusions can be drawn. It is possible to point out such values of R_{\max} , for which the map becomes properly organized for all presented cases even for 7 bits of the weight resolution (nb_W), as shown in Figure 1 (d), 2 (d) and 3 (f). It can be seen in Figure 2 (c),(d) and Figure 3 (c), (f) that optimal values of R_{\max} , for which the map becomes properly organized, are usually small. For the map with 8x8 neurons it was possible to obtain the optimal value of Q_{err} ($16.2e - 3$) for low resolution of nb_W of 7 bits for $R_{\max} = 4$. For larger values of R_{\max} the error Q_{err} is larger. This shows that if the neighborhood at the beginning of training is “too strong”, the learning process is not optimal. Similar results are observed for the map with 16x16 neurons. In this case the map becomes properly organized (for $nb_W = 7$) for $R_{\max} = 3$, while for $R_{\max} = 12$ the Q_{err} is much larger.

These conclusions are very important from the hardware implementation point of view. Since all neurons in the map are composed of the same blocks, therefore any reduction of the complexity of any block in a single neuron has an effect on the complexity of the overall map.

2.2 Implementation issues of the SOM on the FPGA device

The presented Kohonen SOM together with the adaptation mechanism has been described in the VHDL – hardware description language. This makes the system easy portable between different FPGA devices that is one of the advantages here. The number of bits of the weight w and the input x signals have a direct impact on the size of the map that can be synthesized on a single device. The used FPGA Virtex-5 device contains 17,500 slices. Every slice contains four logic-function generators (or look-up tables), four storage elements, wide-function multiplexers, and carry logic. The comparative study for different signal resolutions are shown in Table 1. The results – the number of slices for particular cases– show the importance of the presented investigations. The ability to shorten the length of some signals to only 7 bits, without disrupting the learning process, increases the number of neurons that can be implemented on a single device even by 240% in the comparison when dealing with the resolution of 16 bits.

Table 1: The number of slices occupied by the SOM for different map sizes

nb_w	one neuron	map (3x3)	map (10x10)
7	72	1253	8180
10	102	2011	13650
16	179	2961	19840 (not synthesizable)

3 Conclusions

Detailed investigations presented in the paper show that even for low resolutions of the weight and the input signals of only 7 bits, the learning abilities of the Kohonen SOM are not affected. For low resolutions of these signals the number of neurons that can be realized on a single FPGA device substantially increases. Even though 7 bits might not be enough in many cases, the optimizing of this parameter is important in any hardware realization.

References

- [1] W. Kurdthongmee, “A novel hardware-oriented Kohonen SOM image compression algorithm and its FPGA implementation”, *Journal of Systems Architecture*, Vol. 54, No. 10, pp. 983–994, October 2008.
- [2] H. Hikawa, “FPGA implementation of self organizing map with digital phase locked loops”, *Neural Networks*, Vol. 18, pp.514–522
- [3] J. Pena, M. Vanegas, A. Valencia, “Digital hardware architectures of Kohonen’s self organizing feature maps with exponential neighboring function”, *IEEE international conference on reconfigurable computing and FPGA’s*, 2006, pp. 1–8
- [4] J. A. Lee, M. Verleysen, “Self-Organizing Maps with Recursive Neighborhood Adaptation”, *Neural Networks*, Vol. 15, Issues 8-9, October-November 2002, pp. 993–1003
- [5] M. Kolasa, R. Długosz, W. Pedrycz and M. Szulc, “Programmable Triangular Neighborhood Function for Kohonen Self-Organizing Map Implemented on Chip”, *Neural Networks*, doi:10.1016/j.neunet.2011.09.002, Vol. 5, pp. 146 – 160, January 2012