Architecture of Memory Blocks for Hardware Implementation of Partially Parallel Fuzzy ART Neural Network

Serhii Shatnyi[1], Pavlo Tymoshchuk[2]

National University of Water and Environmental Engineering[1]
Lviv Polytechnic National University[2]
sha_ser@ukr.net, pavlo.v.tymoshchuk@lpnu.ua

INTRODUCTION

Neural network (NN) systems used for signal and data analysis provide many advantages over traditional techniques. In particular, signal and data processing systems can be designed based on Fuzzy Adaptive Resonance Theory (ART) NNs. An important advantage of the Fuzzy ART NNs is their high performance. Such networks are characterized high accuracy, scalability for large datasets and capability to learn not only in offline but also in online modes. Moreover, Fuzzy ART NNs can correctly process noised data [1].

Fuzzy ART NNs are successfully used in various applications, in particular, for medical data processing, pattern recognition and classification, and clustering [1]. Software implementation of sequential and partially parallel Fuzzy ART NNs can be found, for instance, in [2], [3]. Hardware implementation designs of the networks are presented in [4], [5]. In particular, hardware implementation of clustering system based on partially parallel Fuzzy ART NN on reconfigurable FPGA in LabVIEW environment is described in [5].

In this paper, reconfigurable FPGA hardware architecture of memory blocks is designed for partially parallel Fuzzy ART NN. Results of presented simulations show that parallel memory blocks used is capable to operate much faster than their sequential counterparts.

HARDWARE IMPLEMENTATION ARCHITECTUREOF MEMORY BLOCKS

In order to design hardware implementation architecture of memory blocks, we employ National Instruments LabView environment and Real-Time Single Board RIO sbRIO-9637 as hardware and software tools. The basis of the platform used is reconfigured FPGA integral circuit with 40

MHz operation speed, 48 programmable analog input-output channels, and 60 programmable input-output channels. The software provides a possibility to perform hardware simulations with high accuracy.

The architecture consisting of two independent modules, in particular, the FPGA module and host module is presented in Fig. 1. The FPGA module contains two physically independent memory blocks VI-Memory_1 and VI-Memory_2 and three independent operational channels used for performing reading and writing operations. Auxiliary elements in this module are Tick Count and Iteration Count used for accurate time measurements, analog output interface employed for data communication and transmission, logical if-case operation used for precise limiting of read and write operations. The FPGA module, in which two independent blocks of memory and three channels are implemented, operates in parallel access mode and performs operations of writing and reading.

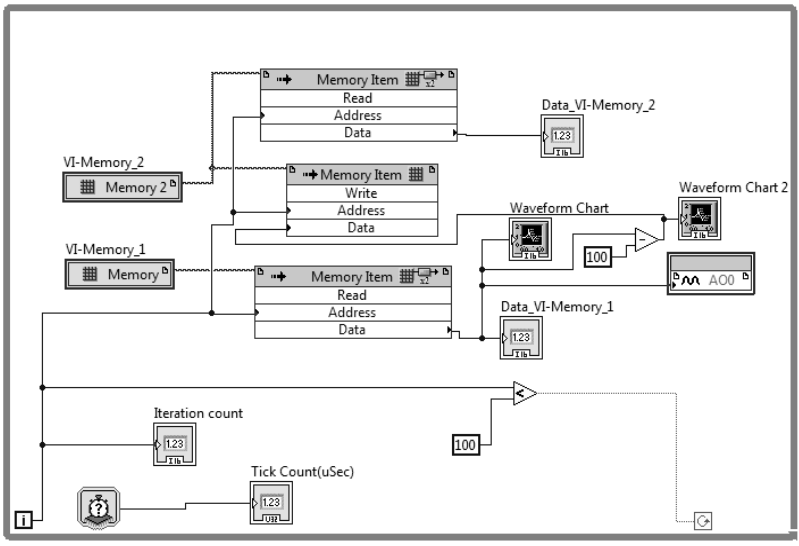

Fig. 1. Block diagram of FPGA module.

The host module, shown in Fig. 2, is used for simulation in the mode of sequential access to memory. As it can be seen in Fig. 2, the host module is presented as one- dimensional block and grid. The host module has peer-to-peer interface used for communication with hardware built based on sbRIO-9637 that is used for time measurements and defining speedup in parallel operation mode.

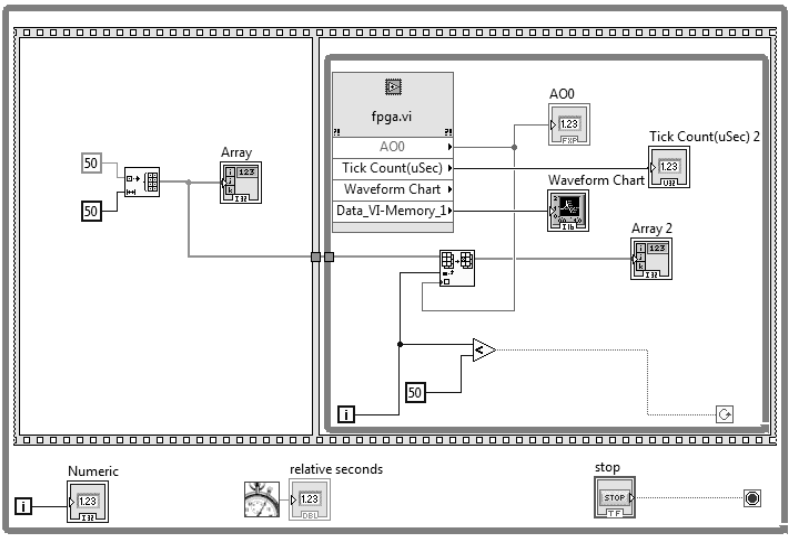

Fig. 2. Block diagram of host module.

Simultaneous operation of the FPGA module and host module provide writing and reading data in partially parallel Fuzzy ART NN.

COMPUTER SIMULATIONS

Consider example of computer simulations of hardware implementation of the FPGA module and host module. For this purpose, corresponding code of high-performance language of graphical computing G of National Instruments and IBM PC based computer system are employed. Timings are set on a dual-core 1.7 GHz Intel Core i5 with 8GB RAM [6]. The main memory block contains sinusoidal signals generated as floating-point data (Fig. 3), initial address of the block 0x00, and final address 0x1024. Each output signal of the first memory block is modified by determining its smallest value, inversion and recording the obtained signal in the second memory block. The reading, processing, and recording of one point of sampled signal are performed in the first computing clock. The results obtained in parallel memory mode are presented in Fig. 4.

Since the architecture described above consists of two independent modules one can determine not only time characteristics but also quality and accuracy of performed operations. The results obtained in memory mode operation are shown in Fig. 5.

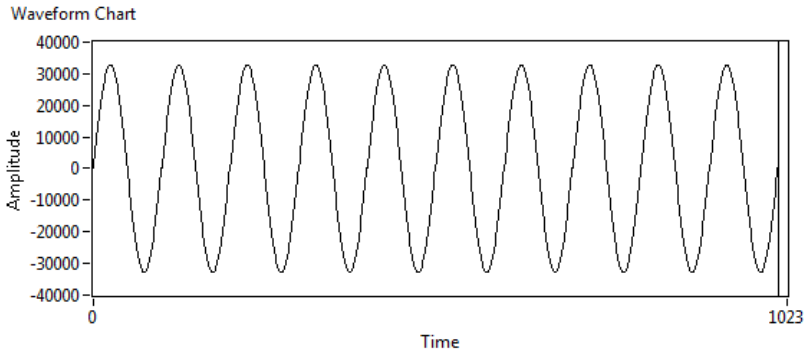

Fig. 3. Data set generated in main memory block.

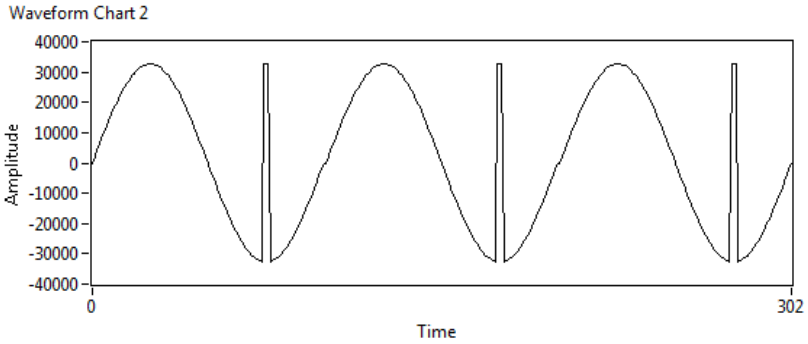

Fig. 4. Updated data set stored in parallel memory block without with defining smallest signal.

The time characteristics obtained by using different sets of input data [2] are presented in Table 1, where C_s is a number of cycles obtained for sequential read\write operations, C_p is a number of cycles obtained for parallel read\write operations, T_s is an operation time in sequential mode (ms), T_p is an operation time in parallel mode (us). As it can be seen in Table 1, the time characteristics are differed depending on a structure of the initialized memory block. On the other hand, the change of time characteristics matches well to one of the Fuzzy ART NN described in [5]. The schematic memory blocks of Fuzzy ART NN designed by specialized software are suitable for testing on both real signals and data sets obtained from available repositories.

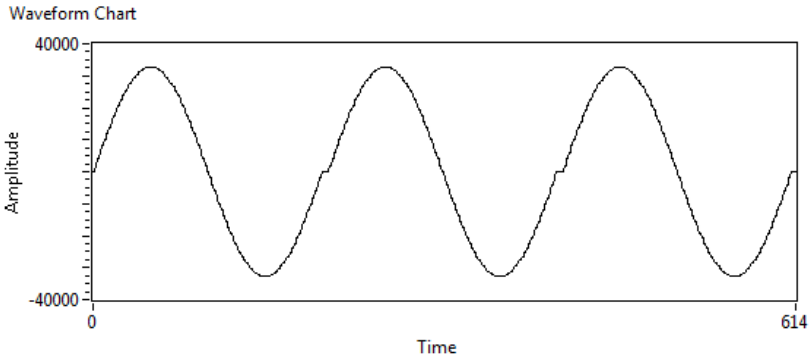

Fig. 5. Updated data set stored in sequential memory block without determining smallest signal.

TABLE 1

RESULTS OF SIMULATIONS OF MEMORY READ\WRITE OPERATIONS IN SEQUENTIAL AND PARALLEL MODES

№	Data Size	C_s	C_p	T_s (ms)	T_p (us)
1	10x1000	7	5	1.704	7.1
2	50x1000	11	7	11.501	18.1
3	100x1000	16	10	38.673	41.651
4	500x200	23	13	41.569	41.1
5	1500x67	31	16	43.137	42.317
6	6500x15	41	20	32.133	21.1
8	10000x10	52	27	55.373	41.1

According to the results of simulations, the parallel memory used for partial parallel Fuzzy ART NN is capable to operate up to 10^3 times faster than its sequential counterpart. Thus, the parallel memory can be recommended to use for hardware implementation of partial parallel Fuzzy ART NN.

REFERENCES

[1] S. Grossberg, "Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world," *Neural Networks*, vol. 37, pp. 1-47, Jan. 2013.

[2] M. Martinez-Zarzuela, F. Pernas, J. Higuera, M. Rodriguez, "Fuzzy ART neural network parallel computing on the GPU," *in Proc. Int. Work-Conf. on Artificial Neural Networks,* Jun. 20, 2007, pp. 463-470.

[3] Serhii Shatnyi, Pavlo Tymoshchuk, "Simulation of sequential and partially parallel Fuzzy ART neural network," *in Proc. 14th Int. Conf. "The Experience of Designing and Application of CAD Systems in Microelectronics"*, Feb. 21-25, 2017, pp. 372 – 374.

[4] T. Serrano-Gotarredona and B. Linares-Barranco, "A modified ART 1 algorithm more suitable for VLSI implementations," *Neural Networks, vol. 9, no. 6, ,* Aug. 1996. pp.1025-1043.

[5] P. V. Tymoshchuk, S. V. Shatnyi, "Hardware implementation design of partially parallel clustering system based on Fuzzy ART using LabVIEW," *in Proc. XXIVth Int. Ukrainian-Polish Conf. "CAD in Machinery Design. Implementations and Educational Issues*, L'viv, Ukraine, Oct. 21-22, 2016, pp. 29-30.

[6] S. Shatnyi and P. Tymoshchuk, "Simulation of electrocardiogram processing and analysis using Fuzzy ART," *Computer Systems of Design. Theory and Practice*, No 882, pp. 70-75, 2017.