2017 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2017, Cancun, Mexico, December 4-7, 2017

Complex-valued neural networks to realize energy-efficient neural networks including reservoir computing

Akira Hirose†, Seiji Takeda‡, Toshiyuki Yamane‡, Daiju Nakano‡,
Shigeru Nakagawa‡, Ryosho Nakane† and Gouhei Tanaka†

†Social Cooperation Program on Energy Efficient Information Processing (EEIP) and
Department of Electrical Engineering and Information Systems,
The University of Tokyo
7–3–1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
‡IBM Research - Tokyo, Kawasaki, Kanagawa 212-0032, Japan
Email: ahirose@ee.t.u-tokyo.ac.jp

Abstract—This paper describes the background and basic ideas of wave-based neural networks including reservoir computing for realizing energy-efficient neural network hardware. We also show the significance of the framework of complex-valued neural networks.

1. Introduction

In these years, neural networks (NNs) are widely and actively used in artificial intelligence (AI) fields. The most basic function in modern AI lies in the extraction of correlations of events in various levels and diverse areas. Finding correlations sometimes leads to explicit expression of causality and/or generation of new information. Storing of correlation is also the most basic function of a neuron microscopically as well as a neuron group macroscopically. From this viewpoint of human beings too, correlation finding and storing should definitely be of great importance more and more.

Present AI systems are realized as neuro-based software on von-Neumann type hardware. Then, a large amount of energy is consumed in a system to deal with large-scale data for learning and processing with deep learning or other methods. Edge nodes in sensor networks exhibit large energy consumption in a total system. Saving the energy is a seriously pressing issue. Hardware innovation has great significance also from this point of view.

2. Wiring and variability issues

Here, let's discuss development and technologies by referring to Fig. 1. Researches on neural hardware hold a long history of several decades in multiple directions. In any direction, however, they encounter a common difficulty of wiring increase at an exponential rate along the growth of the network size, resulting in fatal infabricability. Furthermore, the increase of the total wiring length leads directly to the increase of electric power to charge and discharge the wires. Energy saving definitely requires the solution of wiring explosion.

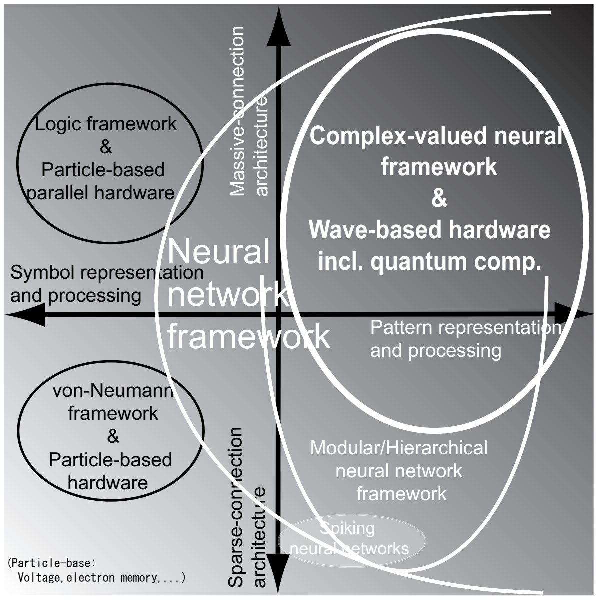

Figure 1: Conceptual diagram showing the promising information-processing frameworks and hardware architectures mapped in the coordinates of the degree of pattern-/symbol-information representations and processing modes as well as the degree of sparsity/massiveness in wiring (modified from Ref. [1]).

The wiring explosion has been a big problem even in von-Neumann computers. In the von-Neumann systems in addition, the downscaling of elementary devices is also a key point. Ultimate downscaling of such a level that an element is composed of only several tens of atoms brings about relatively huge variability in the electronic characteristics of the elementary devices. The variability disables us for precise fabrication of a circuit in detail. This fact suggests strongly the increasing importance of neural adaptability in information processing systems widely in the near future. This issue is also one of the most serious problems though the limitation in Moore's Law refers to it only im-

A part of this work is to be presented at International Conference on Neural Information Processing (ICONIP) 2017 [1].

plicitly.

3. Complex-valued neural networks, a wave-friendly architecture

Complex-valued neural networks (CVNNs) are constructs a framework for dealing with complex amplitude [2–9]. They extend their applicable fields mainly in electronics such as coherent imaging [10–12], channel prediction in multi-path mobile communications [13, 14] to treat complicated electromagnetic field, lightwave information processing systems [15–18], in particular in adaptive processing of phase information [19–21] and lightwave frequency-multiplexed information processing [22, 23], lightwave processing without physical wire interconnections [24–28], as well as quantum computing [29].

The most important advantage of a CVNN lies in the superior generalization ability in application to processing of wave-originating information and wave-based neural hardware [30,31]. The merit is significant also in reservoir computing utilizing lightwave [32, 33] including echo state networks [34, 35].

4. Comparisons of various neural network hardware

Conventional computers represent information by whether one or more static particle (electron) exist or not. In this sense, they use baseband (static) physical representation. This strategy is suitable for realizing a memory using electrons and representing information as digital symbols (bits). In particular, the ease in function updates by rewriting software programs has been worthy to note. This situation applies also to neural network implementation by using field programmable gate arrays (FPGAs) or general purpose computing on graphics processing unit (GPGPU).

Wave implementation of neural connections for adaptive learning and processing is somewhat in contrast. Though it is true that analog use sometimes results in a limited accuracy in each element, this is not a serious problem because the neural adaptability compensates this weakness in the operation as a system. We would rather utilize the great merit of analog use, namely the continuity in response, that realizes diverse and flexible learning ability in the physical level. Contrarily, in digital systems, dynamics such as learning need to be written as software. This is because the metric of bit information is completely separate from the physical metric such as voltage, resulting in the necessity for human beings to assign information and/or meaning of a bit by writing software.

In history, the parametron also dealt with phase information [36]. Its memory employed symbol representation in such a manner that the circuit holds bistable or multistable states in its phase values to choose one of the possible states. Its processing also treated the phase as a symbol. Parametron is a symbol-processing digital system, dealing with bit information, not fully successful in utilizing the merit of employing the wave/phase. This point was revealed as a weak point, as well as the incompatibility with integration, of parametron.

In contrast, the basis of neural networks lies in pattern information representation and pattern information processing. Then, it is possible in principle that a physically natural spatio-temporal gradient determines the neural dynamics. In this sense, neural networks are compatible with physical implementation. They can also include nonlinearity. In addition, they can realize information connections among processing elements without charge and discharge by employing wave-related phenomena without physical lines. Such architecture realizes flexible, massive and energy-efficient networks.

Pulse neural networks, or spiking neural networks, may be located between these two architectures mentioned above. They have a set of merits such as that fact that correlation is obtained simply as a time-domain average of a series of pulses multiplied by another sequence of pulses. However, because of the baseband circuit structure, they have a shortcoming of charge-and-discharge energy consumption just like conventional digital neural networks. In addition, the multiple-pulse information representation requires a frequency bandwidth wider than what is needed intrinsically for the information representation, resulting in larger power consumption.

5. CVNNs based on natural metric of waves

When we deal fully with wave information, we actually work on its amplitude and phase. We have to introduce the natural metric involved in the complex amplitude for use of physical wave nature into the neural dynamics. Nonliearity should also be in harmony with the amplitude and phase so that it works as a meaningful nonlinear function. Such neural framework is the CVNN. When we deal with polarization additionally, we should employ/combine quaternion neural networks, an extended framework of complex-valued neural networks [12, 37–39].

Consequently, it is clear that wave-connection neural networks hold a great advantage in particular in realization of energy-efficient neural-network hardware. Therefore, the framework of CVNNs plays an important role in neural hardware in the next generation.

6. Conclusion

We discussed energy-efficient neural networks and reservoir computing. There the wave-connection neural networks hold a great advantage. The framework of the complex-valued neural networks serves an important role.

References

[1] A. Hirose, S. Takeda, D. Nakano, T. Yamane, S. Nakagawa, R. Nakane, and G. Tanaka, “Complex-valued neural networks for wave-based realization of reservoir computing,” in *International*

Conference on Neural Information Processing (ICONIP) 2017 Guangzhou, to be presented, 2017.

[2] A. Hirose, *Complex-Valued Neural Networks, 2nd Edition*. Heidelberg, Berlin, New York: Springer, 2012.

[3] A. Hirose (Ed.), *Complex-Valued Neural Networks: Advances and Applications*, ser. IEEE Press Series on Computational Intelligence. New Jersey, U.S.A.: IEEE Press and Wiley, April 2013.

[4] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters – Noncircularity, Widely Linear and Neural Models*. Wiley, April 2009.

[5] T. Adali and S. Haykin, *Adaptive Signal Processing: Next Generation Solutions*. Wiley-IEEE Press, 2010.

[6] I. Aizenberg, *Complex-Valued Neural Networks with Multi-Valued Neurons*, ser. Studies in Computational Intelligence. Springer, 2011.

[7] T. Nitta, *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*, T. Nitta, Ed. Pennsylvania: Information Science Reference, 2009.

[8] E. Bayro-Corrochano, *Geometric Computing for Wavelet Transforms, Robot Vision, Learning, Control and Action*. Springer, 2010.

[9] S. Suresh, N. Sundararajan, and R. Savitha, *Supervised Learning with Complex-valued Neural Networks*. Springer, 2013.

[10] T. Hara and A. Hirose, "Adaptive plastic-landmine visualizing radar system: Effects of aperture synthesis and feature-vector dimension reduction," *IEICE Transactions on Electronics*, vol. E88-C, no. 12, pp. 2282–2288, December 2005.

[11] A. B. Suksmono and A. Hirose, "Interferometric sar image restoration using Monte-Carlo metropolis method," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 290–298, 2002.

[12] F. Shang and A. Hirose, "Quaternion neural-network-based PolSAR land classification in Poincare-sphere-parameter space," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5693–5703, September 2014.

[13] A. Hirose and T. Nagashima, "Predictive self-organizing map for vector quantization of migratory signals and its application to mobile communications," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1532–1540, 2003.

[14] T. Ding and A. Hirose, "Fading channel prediction based on combination of complex-valued neural networks and chirp Z-transform," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1686–1695, September 2014.

[15] A. Hirose and R. Eckmiller, "Proposal of frequency-domain multiplexing in optical neural networks," *Neurocomputing*, vol. 10, no. 2, pp. 197–204, 1996.

[16] ——, "Coherent optical neural networks that have optical-frequency-controlled behavior and generalization ability in the frequency domain," *Applied Optics*, vol. 35, no. 5, pp. 836–843, 1996.

[17] ——, "Coherent optical neural networks and the generalization characteristics," *Optical Review*, vol. 3, no. 6A, pp. 418–422, 1996.

[18] S. Kawata and A. Hirose, "Coherent lightwave associative memory system that possesses a carrier-frequency-controlled behavior," *Opt. Eng.*, vol. 42, no. 9, pp. 2670–2675, 2003.

[19] ——, "Coherent optical neural network that learns desirable phase values in frequency domain by using multiple optical-path differences," *Optics Letters*, vol. 28, no. 24, pp. 2524–2526, 2003.

[20] K. Tanizawa and A. Hirose, "Adaptive control of tunable dispersion compensator that minimizes time-domain waveform error by steepest-descent method," *IEEE Photonics Technology Letters*, vol. 18, no. 13, pp. 1466–1468, 2006.

[21] ——, "Performance analysis of steepest-descent-based feedback control of tunable dispersion compensator for adaptive dispersion compensation in all-optical dynamic routing networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, no. 4, pp. 1086–1094, 2007.

[22] S. Kawata and A. Hirose, "Frequency-multiplexed logic circuit based on a coherent optical neural network," *Applied Optics*, vol. 44, no. 19, pp. 4053–4059, 2005.

[23] ——, "Frequency-multiplexing ability of complex-valued Hebbian learning in logic gates," *International Journal of Neural Systems*, vol. 12, no. 1, pp. 43–51, 2008.

[24] A. Hirose and M. Kiuchi, "Coherent optical associative memory system that processes complex-amplitude information," *IEEE Photon. Tech. Lett.*, vol. 12, no. 5, pp. 564–566, 2000.

[25] A. Limmanee, S. Kawata, and A. Hirose, "Phase signal embedment in densely frequency-multiplexed coherent neural networks," in *OSA Topical Meeting on Information Photonics (OSA–IP) 2005 Charlotte*, no. ITuA2, June 2005.

[26] A. Hirose, T. Higo, and K. Tanizawa, "Efficient generation of holographic movies with frame interpolation using a coherent neural network," *IEICE Electronics Express*, vol. 3, no. 19, pp. 417–423, 2006.

[27] C. S. Tay, K. Tanizawa, and A. Hirose, "Error reduction in holographic movies using a hybrid learning method in coherent neural networks," *Applied Optics*, vol. 47, no. 28, pp. 5221–5228, 2008.

[28] M. Takeda, S. Kirihara, Y. Miyamoto, K. Sakoda, and K. Honda, "Localization of electromagnetic waves in three dimensional fractal cavities," *Physical Review Letters*, vol. 92, p. 093902, 2004.

[29] A. Ono, S. Sato, M. Kinjo, and K. Nakajima, "Study on the performance of neuromorphic adiabatic quantum computation algorithms," in *International Joint Conference on Neural Networks (IJCNN) 2008 Hong Kong Nakajima*, June 2008, pp. 2508–2512.

[30] A. Hirose and S. Yoshida, "Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, pp. 541–551, 2012.

[31] A. Hirose, "Nature of complex number and complex-valued neural networks," *Frontiers of Electrical and Electronic Engineering in China*, vol. 6, no. 1, pp. 171–180, 2011.

[32] S. Takeda, D. Nakano, T. Yamane, G. Tanaka, R. Nakane, A. Hirose, and S. Nakagawa, "Photonic reservoir computing based on laser dynamics with external feedback," in *International Conference on Neural Information Processing (ICONIP) 2016 Kyoto*, 2016, pp. 222–230.

[33] T. Yamane, S. Takeda, D. Nakano, G. Tanaka, R. Nakane, S. Nakagawa, and A. Hirose, "Dynamics of reservoir computing at the edge of stability," in *International Conference on Neural Information Processing (ICONIP) 2016 Kyoto*, 2016, pp. 205–212.

[34] G. Tanaka, R. Nakane, T. Yamane, D. Nakano, S. Takeda, S. Nakagawa, and A. Hirose, "Exploiting heterogeneous units for reservoir computing with simple architecture," in *International Conference on Neural Information Processing (ICONIP) 2016 Kyoto*, 2016, pp. 187–194.

[35] R. Mori, G. Tanaka, R. Nakane, A. Hirose, and K. Aihara, "Computational performance of echo state networks with dynamic synapses," in *International Conference on Neural Information Processing (ICONIP) 2016 Kyoto*, 2016, pp. 264–271.

[36] E. Goto, "The parametron – A new circuit element which utilizes non-linear reactors (in Japanese)," *Paper of Technical Group of Electronic Computers and Nonlinear Theory, IECE*, July 1954.

[37] N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura, "Quaternion neural network with geometrical operators," *Journal of Intelligent and Fuzzy Systems*, vol. 15, pp. 149–164, 2004.

[38] N. Usami, A. Muhuri, A. Bhattacharya, and A. Hirose, "PolSAR wet snow mapping with incidence angle information," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 2029–2033, December 2016.

[39] Y. Takizawa, F. Shang, and A. Hirose, "Adaptive land classification and new class generation by unsupervised double-stage learning in poincare sphere space for polarimetric synthetic aperture radars," *Neurocomputing*, vol. 248, pp. 3–10, 2017.