Complex-Valued and Quaternion Neural Networks to Deal with Complex Amplitude and Polarization: Their Strength in Electronics

Akira Hirose* * The University of Tokyo, Tokyo, Japan E-mail: ahirose at ee.t.u-tokyo.ac.jp

Abstract—This paper discusses the advantages in complexvalued and quaternion neural networks (CVNNs and QNNs) by presenting important concepts as well as applications in electronic engineering. Essentially, CVNNs and QNNs have excellent generalization ability in the learning of two- and three-dimensional rotation and scaling, respectively. Its origin is the sparsity in themselves as complex and quaternion numbers. The ability is useful in particular in adaptive processing of waves consisting of complex amplitude (amplitude and phase) and/or polarization. We consider an example in communications where we deal with channel prediction to discuss the superiority in CVNNs and its basics.

I. INTRODUCTION

Neural network research and development have been promoted forward by the cooperation of mathematics, physics, physiology and psychology to elucidate the human brain mechanisms as well as to utilize its findings in engineering so that we can construct systems which provide us with great functionality beyond human beings in specific aspects [2]– [8]. Complex-valued neural networks (CVNNs) and quaternion neural networks (QNNs) are constructed by, and dealing with, complex and quaternion numbers, respectively, to realize such super-human systems.

Most CVNNs and QNNs aim at engineering applications rather than human physiology or psychology [9]-[12]. In electrical and electronic engineering, CVNNs are the networks dealing with complex amplitude. That is, they are coherent neural networks, that treat phase information in a consistent manner. They are very powerful in phase-sensitive physics and systems such as satellite/air-borne synthetic aperture radar (SAR) [13]-[16], ground penetrating radar (GPR) [17]-[20], quantum computing [21], [22], and lightwave computing [23], [24]. Analyses of the performance and dynamics in such systems are also very active in various fields such as communication systems [25] and related mathematics in complex manifold [26], [27]. Similarly, QNNs are widely effective in polarization processing when we represent its state on/in the Poincare sphere because of the high generalization ability in three-dimensional space [28]-[30]. Many ideas have been proposed on supervised land classification [31]-[34], unsupervised classification [35], [36], and also self-organizing codebook generation [37]. Moreover, device applications are also within the CVNN applications. For example, a reservoir computing device using spin waves without hard wiring has also been proposed [38]–[42].

We can observe these trends from an inverse-directional sight as follows. Adaptive signal processing based on learning and self-organization has a long history in such fields as adaptive antennas and inverse problems [43]. In linear systems, the complex least mean square (LMS) algorithm was proposed in 1975 [44] and used in various steepest-descent and conjugate-gradient systems [45]. CVNN is a more generalized framework by introducing nonlinearity, layers and various connection forms. Recent special issue on CVNNs [46] also focuses on such aspects of CVNNs in their theories [47], [48] and applications [49]–[52].

People just using neural networks as convenient tools may sometimes say, "a neural network solves problem, but it is a black box that presents nothing about the logical deduction or reasoning." It is true that a neural network is not a "logic' machine, but actually it presents the reasons. In reality, there exist many networks which analyze problems and/or extract features, namely, deep-learning networks based on classical hourglass-structure neural networks [53], [54], and convolutional neural networks including the Neocognitron [55], [56]. Neural networks reflect the external world in their connections. Then, it is also possible for us to prepare a network suitable for respective tasks to be dealt with. A good example is adaptive antennas. CVNN is a network incorporating the wave physics. It is also a combination of things (physics) of electromagnetic wave propagation and scattering and things (matters, affairs) of adaptive and statistical information processing.

CVNNs show the merits in, for example, earth observation from space, several hundred kilometers above land, by using microwave transmitted from a satellite and processing the scattered wave data with the synthetic aperture process, to obtain earth surface information included basically in the amplitude, phase and polarization [57]. The raw data often suffer from the distortion generated by propagation, diffraction, refraction, interference and spatial and temporal discreteness in observation. Restoration of the real information included in the distorted data requires good estimation and prediction methods. Here the CVNNs, supposing wave physics, play an

A part of this work was presented in IEICE Transactions on Electronics (Japanese Edition) [1]. This work was partly supported by JSPS KAKENHI Grant Number 15H02756 and 18H04105.



Fig. 1. (a) A set of teacher input-output signals, (b) a look-up table and (c) processing with one (solid) or another (broken) generalization characteristic for estimation and prediction [1].

important role [58], [59]. Their estimation and prediction are the result of the combination of mathematics of adaptation and physics of waves. The application fields are expanding continuously. The precise and global observation of earth surface forms and motions will realize dense and quasi-realtime monitoring of volcano activities, earthquake damages, changes in glaciers and pole ices as well as forest biomass in relation to global warming, agricultural product based on polarization changes, and so on and so forth. All of them are related to the key issues to be solved in the present society, namely, sustainable development goals (SDGs).

This paper discusses the essence of CVNNs by focusing on the generalization ability. Then we check the merit briefly in communication channel prediction. In the presentation, we discuss also QNNs as well.

II. GENERALIZATION ABILITY

Generalization ability is the most important concept in neural networks and other adaptive signal processing. We review and discuss the key points in the generalization ability of CVNNs [58].

A. Sparsity and degree of freedom

In general, a neural network realizes its desired function by adjusting the connection weights among neurons depending on fed signals and environments. The adjustment is preformed through supervised learning, unsupervised learning or reinforcement learning. The more the weights exist, the higher the degree of freedom is.

Fig. 1(a) illustrates a function-approximation task which a neural network is expected to learn by supervised learning with the data having three teacher points in one-dimensional

input x and one-dimensional output y. We express these teacher signal points as (·), i.e., $(\hat{x}, \hat{y}) = (1, 4), (3, 5), (6, 2).$ This is a common situation in an adaptive antenna to adjust its weights by using pilot signals and an inverse problem to learn examples. If we expect appropriate outputs only at these teacher inputs, we do not need any learning process. Instead, we have only to make a table of (memorize) the input-output values shown in Fig. 1(b). However, what we actually expect is the input-output relationship shown by the curve(s) in Fig. 1(c), generating outputs even for inputs that were not used in the learning process. This is done by the generalization ability, and the curve shows the generalization characteristics. A good generalization characteristics are not obtained by simply nullifying the square error $\sum (y - \hat{y})^2$ where y is the output value varying through the learning process. Moreover, we cannot say what curve is ideal before an actual task is given.

A small degree of freedom (DoF) often prevents the convergence of the error value to zero. Too large DoF, however, makes the curve wind, resulting in inappropriate estimation or prediction. Appropriate DoF is desired, which is realized by pruning or addition of network connections. We have also to check the meaningfulness of the alive connections by examining the independence among the signals going through the connections existing in parallel so that every connection works meaningfully without redundancy. In this sense, complex-valued connections and neurons hold a great advantage in dealing with waves and wave-related information. This is shown as follows.

B. Real 2×2 matrix representation

There are multiple manners to represent complex numbers. In 1835, Hamilton defined complex number z in an algebraic manner as ordered pair of real numbers, x and y, as $z \equiv (x, y)$ with the arithmetic of addition and multiplication as

$$(x_1, y_1) + (x_2, y_2) \equiv (x_1 + x_2, y_1 + y_2) \tag{1}$$

$$(x_1, y_1) \cdot (x_2, y_2) \equiv (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$$
(2)

where the multiplication has a special characteristic. We can also use real 2×2 matrices to represent complex numbers [60], [61]. With every complex number c = a + ib, we associate the *C*-linear transformation

$$T_c: C \to C, \quad z \mapsto cz = ax - by + j(bx + ay)$$
 (3)

If we identify C with R^2 by

$$z = x + jy = \begin{pmatrix} x \\ y \end{pmatrix}$$
(4)

it follows that

$$T_{c}\begin{pmatrix} x\\ y \end{pmatrix} = \begin{pmatrix} ax - by\\ bx + ay \end{pmatrix}$$
$$= \begin{pmatrix} a & -b\\ b & a \end{pmatrix} \begin{pmatrix} x\\ y \end{pmatrix}$$
(5)

That is, the linear transformation T_c determined by c = a + ib is represented by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Generally,



Fig. 2. A mapping that maps x^{IN} to x^{OUT} [58].



Fig. 3. A simple case to learn the mapping in Fig. 2: (a) A real-valued singlelayered two-input two-output feedforward network, (b) a possible degenerate learning result that is mostly unuseful, (c) a complex-valued neural network seemingly identical to (a), and (d) the learning result obtained in this smalldegree-of-freedom case [58].

a mapping represented by a 2×2 matrix is noncommutative. However, in the present case, it becomes commutative. What is most important is that this matrix representation clearly expresses the function specific to complex numbers—that is, the rotation and scaling as

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$
(6)

III. COMPLEX NUMBERS IN FEEDFORWARD NEURAL NETWORKS

Fig. 2 shows an example situation. We assume that a network should realize a mapping that transforms an input x^{IN} to an output x^{OUT} through supervised learning that adjusts the synaptic weights w_{ji} . Simply, we have only a single teacher pair of input and output signals.

Here we consider a case shown in Fig. 3(a), where we have a single-layer 2-input 2-output feedforward neural network in real number. For simplicity, we omit nonlinearity at the neurons, i.e., activation function. The neurons have no threshold either. Then we can express a general input-output relationship

as

$$\begin{pmatrix} x_1^{\mathbf{OUT}} \\ x_2^{\mathbf{OUT}} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1^{\mathbf{IN}} \\ x_2^{\mathbf{IN}} \end{pmatrix}$$
(7)

There exist a variety of possible mapping obtained through the learning because the number of parameters to be determined is larger than that of the conditions; i.e., the learning task is an ill-posed problem. The functional difference emerges as the difference in the generalization characteristics. For example, learning can result in a degenerate mapping shown in Fig. 3(b), which is often unuseful in practice.

In contract, let us consider the mapping learning task in the one-dimensional complex domain, which transforms a complex value $\mathbf{x^{IN}} = (x_1^{IN}, x_2^{IN})$ to another complex value $\mathbf{x^{OUT}} = (x_1^{OUT}, x_2^{OUT})$. Fig. 3(c) shows the complex-valued network, where the weight is a single complex value. The situation is expressed just like in (7) as

$$\begin{pmatrix} x_1^{\mathbf{OUT}} \\ x_2^{\mathbf{OUT}} \end{pmatrix} = \begin{pmatrix} |w|\cos\theta & -|w|\sin\theta \\ |w|\sin\theta & |w|\cos\theta \end{pmatrix} \begin{pmatrix} x_1^{\mathbf{IN}} \\ x_2^{\mathbf{IN}} \end{pmatrix}$$
(8)

where $\theta \equiv \arg(w)$. There are only two parameters |w| and θ . The learning result is given in Fig. 3(d). The DoF is lower, and the arbitrariness of solution is reduced, resulting in a simple combination of phase rotation and amplitude scaling. This result is interpreted also as the effectiveness of *sparsity* embedded in the number itself. This result is compatible with wave phenomenon.

The size of the above network is so small that the relationship between the parameter number and DoF is estimated very simply. In general, the dynamics of a neural network is determined by various parameters such as network structure, input–output data dimensions, and teacher signal numbers. However, the above characteristics of phase rotation and amplitude modulation are embedded in the complex-valued network as a universal elemental process of weighting.

It is true that the essential merit of neural networks in general lies in the high degree of freedom in learning and selforganization. However, if we know *a priori* that the objective quantities include "phase" and/or "amplitude," we can reduce possibly harmful portion of the freedom by employing a complex-valued neural network, resulting in a more meaningful generalization characteristic. The "rotation" in the complex multiplication works as an elemental process at the synapse, and it realizes the advantageous reduction of the degree of freedom. This is a type of sparsity built in the complex number itself.

IV. CHANNEL PREDICTION IN MOBILE COMMUNICATIONS: CVNN DEALS WITH THE COMPLEX CHANNEL AS AN ACTUAL ENTITY

We present an example of the improvement of neural dynamics by introducing physical picture into neural networks. The example is channel prediction in mobile communications. The details are given in Ref. [62].



Fig. 4. Conceptual illustrations representing (a) the chirp z-transform to convert the past channel information into the frequency domain for the extraction of the frequency-domain channel parameters, namely, (b) amplitude a_m and Doppler frequency f_m , and (c) phase ϕ_m respectively [62].



Fig. 5. Geometrical setup including two scatterers, a base station and a mobile user in the line of sight, which can cause severe fading [62].

Channel prediction is indispensable for future effective preequalization and/or transmission-power control in mobile communications to avoid harmful fading phenomenon in multipath environment. There have been many proposals in time domain and frequency domain including super-resolution methods. But they are not useful practically because of the poor prediction performance and/or the large calculation cost. Recently, how-



Fig. 6. Layered complex-valued neural network having a small size to reduce the calculation cost [62].



Fig. 7. BER curves obtained for TD-Linear prediction, TD-RTRL-CVNNbased prediction, CZT only, CZT-Linear prediction, CZT-Lagrange-based prediction, CZT-AR-model-based prediction, CZT-RTRL-CVNN-based prediction, CZT-ML-CVNN-based prediction, and actual channel characteristics when the geometry causes strongly channel changes with severe fading [62].

ever, we solved this problem by proposing a new method by combining the chirp z-transform (CZT) and a complex-valued neural network [62].

Fig. 4 is the schematic diagram showing our channel prediction. We transform the observation signal illustrated in Fig. 4(a) by using the CZT to obtain the spectrum with mitigation of discreteness as shown in (b). We realize the path separation, by marking the peaks in the magnitude, based on so-called Jakes model. We obtain the estimates of Doppler shift frequency \hat{f}_m and the amplitude \hat{a}_m of path m, and also

the phase $\hat{\theta}_m$ by referring to the phase spectrum shown in (c).

Previously before the proposal, we predicted the channel based on the respective past parameters \hat{f}_m , \hat{a}_m , $\hat{\phi}_m$ individually to obtain the present (t = 0) and future $(t \ge 0)$ parameters \tilde{f}_m , \tilde{a}_m , $\tilde{\theta}_m$ as

$$\tilde{f}_m, \tilde{a}_m, \tilde{\phi}_m \stackrel{\text{Predict}}{\longleftarrow} \hat{f}_m, \hat{a}_m, \hat{\phi}_m$$
(9)

$$\tilde{c}(t) = \sum_{m} \tilde{a}_{m} e^{(j2\pi \tilde{f}_{m}t + \tilde{\phi}_{m})}$$
(10)

It worked well mostly. However, it sometimes failed when the situation is very difficult. An example is shown in Fig. 5 where the receiver runs just beside two scatters located near to each other.

In our newly proposed method, we consider the channel as a unified complex-valued entity, and treated it by using the complex-valued layered neural network shown in Fig. 6 in combination with CZT generating a series of past channel states $\hat{c}_m(t-1), ..., \hat{c}_m(t-I_{\rm ML})$ as

$$\hat{c}_m(t) \equiv \hat{a}_m(t)e^{(j\hat{\theta}_m(t))} \\
(\hat{\theta}_m(t) \equiv 2\pi \hat{f}_m(t)t + \hat{\phi}_m(t)) \quad (11)$$

$$\tilde{c}_m(t) \stackrel{\text{Predict}}{\leftarrow} \hat{c}_m(t-1), ..., \hat{c}_m(t-I_{\text{ML}})$$
 (12)

We slide the prediction window by repeating this prediction to come to the future time slot. Finally we unite the respective path states $\tilde{c}_m(t)$ to predict the channel state in the future as

$$\tilde{c}(t) = \sum_{m} \tilde{c}_{m}(t) \tag{13}$$

Prediction based on the estimated path states $\hat{c}_m(t)$ appeared more difficult than the prediction on the respective parameters \hat{f}_m , \hat{a}_m , $\hat{\theta}_m$. However, our physical experiments showed a channel state with near-circular time evolution, which matches the representation with the sparsity built in the complex number itself. This fact suggests that the complexamplitude prediction is better than the prediction based on respective Doppler frequency, magnitude and phase. Fig. 7 shows the bit error rate (BER) obtained for the difficult case shown in Fig. 5 when we employ various prediction methods, namely, time-domain linear prediction, time-domain real-time recurrent-learning CVNN-based prediction, CZT only without prediction process, CZT frequency-domain linear prediction, CZT frequency-domain Lagrange-based prediction, CZT frequency-domain auto-regression-model-based prediction, CZT frequency-domain real-time recurrent-learning CVNN-based prediction (the proposal), CZT frequencydomain multilayer-CVNN-based prediction, and ideal case where we assume that we could obtain actual channel state (impossible in reality), respectively.

It is shown that our proposed method that combined the CZT and the multilayer CVNN realizes the lowest BER. The calculation cost is small because of its successive online learning. The result directly presents the high performance of the CVNN when it deals with wave information.

V. SUMMARY

This paper discussed the origin of the advantages of CVNNs. We focused on the generalization ability originating from the sparsity existing in the complex number itself. The merit is clearly presented in the performance comparison in channel prediction for mobile communications.

REFERENCES

- A. Hirose and T. Ding, "Neural networks to deal with complex amplitude and its strength in electronics," *IEICE Transactions on Electronics* (*Japanese Edition*), vol. J98-C, no. 10, pp. 184–192, October 2015.
- [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. Nitta, Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters, T. Nitta, Ed. Pennsylvania: Information Science Reference, 2009.
- [6] E. Bayro-Corrochano, Geometric Computing for Wavelet Transforms, Robot Vision, Learning, Control and Action. Springer, 2010.
- [7] I. Aizenberg, *Complex-Valued Neural Networks with Multi-Valued Neurons*, ser. Studies in Computational Intelligence. Springer, 2011.
- [8] A. Hirose, "Complex-valued neural networks for more fertile electronics," *Journal of the the IEICE*, vol. 87, no. 6, pp. 447–449, June 2004, (in Japanese).
- [9] A. Hirose (Ed.), Complex-Valued Neural Networks: Theories and Applications, ser. Series on Innovative Intelligence. Singapore, New Jersey, London: World Scientific Publishing, 2003, vol. 5.
- [10] A. Hirose, "Advances in applications of complex-valued neural networks," *Journal of the Society of Instrument and Control Engineering*, vol. 51, no. 4, pp. 351–357, April 2012, (in Japanese).
- [11] S. Suresh, R. Savitha, and N. Sundararajan, "A sequential learning algorithm for complex-valued self-regulating resource allocation network – CSRAN," *IEEE Transactions on Neural Networks*, vol. 22, pp. –, 2011.
- [12] A. Hirose, "Complex-valued neural networks (review)," Journal of the Institute of Electrical Engineers of Japan, vol. 131, no. 1, pp. 2–8, January 2011, (in Japanese).
- [13] 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.
- [14] —, "Adaptive noise reduction of insar image based on complex-valued mrf model and its application to phase unwrapping problem," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 3, pp. 699–709 (followed by publisher's errata on Fig.12), 2002.
- [15] R. Yamaki and A. Hirose, "Singular unit restoration in interferograms based on complex-valued Markov random field model for phase unwrapping," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 1, pp. 18–22, January 2009.
- [16] Y. Sunaga, R. Natsuaki, and A. Hirose, "Proposal of complex-valued convolutional neural networks for similar land-shape discovery in interferometric synthetic aperture radar," in *International Conference* on Neural Information Processing (ICONIP) 2018 Siep Reap, to be presented, December 2018.
- [17] T. Hara and A. Hirose, "Plastic mine detecting system using complexvalued self-organizing map that deals with multiple-frequency interferometric images," *Neural Networks*, vol. 17, no. 8-9, pp. 1201–1210, November 2004.
- [18] Y. Nakano and A. Hirose, "Improvement of plastic landmine visualization performance by use of ring-CSOM and frequency-domain local correlation," *IEICE Transactions on Electronics*, vol. E92-C, no. 1, pp. 102–108, January 2009.
- [19] 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, 2005.
- [20] Y. Nakano and A. Hirose, "Taper-walled linearly tapered slot antenna," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 4, pp. 779–784, April 2011.

- [21] S. Sato, M. Kinjo, O. Takahashi, Y. Nakamiya, and K. Nakajima, "A study on neuromorphic quantum computation," in *International Joint Conference on Neural Networks (IJCNN) 2004*, vol. 4, 2004, pp. 3253– 3256.
- [22] E. C. Behrman, R. E. F. Bonde, J. E. Steck, and J. F. Behrman, "On the correction of anomalous phase oscillation in entanglement witnesses using quantum neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1696–1703, September 2014.
- [23] S. Kawata and A. Hirose, "Coherent optical neural network that learns desirable phase values in frequency domain by using multiple opticalpath differences," *Optics Letters*, vol. 28, no. 24, pp. 2524–2526, 2003.
- [24] —, "Frequency-multiplexing ability of complex-valued Hebbian learning in logic gates," *International Journal of Neural Systems*, vol. 12, no. 1, pp. 43–51, 2008.
- [25] X. Hong and S. Chen, "Modeling of complex-valued Wiener systems using B-spline neural network," *IEEE Transactions on Neural Networks*, vol. 22, no. 5, pp. 818–825, May 2011.
- [26] S. Fiori, "Extended hamiltonian learning on riemannian manifolds: Theoretical aspects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 22, no. 5, pp. 687–700, May 2011.
- [27] —, "Extended hamiltonian learning on riemannian manifolds: Numerical aspects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 22, no. 5, pp. 7–21, May 2012.
- [28] 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.
 [29] S. Buchholz and G. Sommer, "Quaternionic spinor MLP," in *European*
- [29] S. Buchholz and G. Sommer, "Quaternionic spinor MLP," in European Symposium on Artificial Neural Networks (ESANN) 2000 Bruges, 2000.
- [30] S. Buchholz, "A theory of neural computation with Clifford algebras," Ph.D. dissertation, Christian-Albrechts-Universität Kiel, May 2005.
- [31] 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.
- [32] 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.
- [33] F. Shang and A. Hirose, "PolSAR land classification by using quaternion-valued neural networks," in Asia-Pacific Conference on Synthetic Aperture Radar (APSAR) 2013 Tsukuba. IEEE Geoscience and Remote Sensing Society Japan Chapter / IEICE Electronics Society, September 2013, pp. 593–596.
- [34] S. Buchholz and N. Le Bihan, "Polarized signal classification by complex and quaternionic multi-layer perceptrons," *International Journal of Neural Systems*, vol. 18, no. 2, pp. 75–85, 2008.
- [35] H. Kim and A. Hirose, "Unsupervised fine land classification using quaternion auto-encoder-based polarization feature extraction and selforganizing mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1839–1851, March 2018.
- [36] —, "Polarization feature extraction using quaternion neural networks for flexible unsupervised PolSAR land classification," in *International Geoscience and Remote Sensing Symposium (IGARSS) 2018 Valencia*, no. TUP2.PD.3, July 2018.
- [37] —, "Codebook-based hierarchical polarization feature for unsupervised fine land classification using high-resolution PolSAR data," in *International Geoscience and Remote Sensing Symposium (IGARSS)* 2018 Valencia, no. WEP1.PD.9, July 2018.
- [38] R. Nakane, G. Tanaka, and A. Hirose, "Reservoir computing with spin waves excited in a garnet film," *IEEE Access*, vol. 6, pp. 4462 – 4469, 2018.
- [39] 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*, 2017, pp. 449–456.
- [40] Y. Katayama, T. Yamane, D. Nakano, R. Nakane, and G. Tanaka, "Wavebased neuromorphic computing framework for brain-like energy efficiency and integration," *IEEE Transactions on Nanotechnology*, vol. 15, no. 5, pp. 762–769, September 2016.
- [41] A. Hirose, S. Takeda, D. Nakano, T. Yamane, S. Nakagawa, R. Nakane, and G. Tanaka, "Complex-valued neural networks to realize energyefficient neural networks including reservoir computing," in *International Symposium on Nonlinear Theory and Its Applications (NOLTA)* 2017 Cancun, A2L-E-3, 2017.

- [42] A. Hirose, G. Tanaka, S. Takeda, T. Yamane, H. Numata, N. Kanazawa, J. B. Heroux, D. Nakano, and R. Nakane, "Complex-valued neural networks for wave-based realization of reservoir computing," in *International Conference on Neural Information Processing (ICONIP) 2018 Siem Reap*, to be presented, 2018.
- [43] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143– 2159, December 1967.
- [44] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," Proceedings of the IEEE, vol. 63, pp. 719–720, 1975.
- [45] K. Takao and N. Kikuma, "An adaptive array utilizing an adaptive spatial averaging technique for multipath environments," *IEEE Transactions on Antennas and Propagation*, vol. AP-35, pp. 1389–1396, 1987.
- [46] A. Hirose, I. Aizenberg, and D. P. Mandic, "Guest Editorial: Special Issue on Complex- and Hypercomplex-Valued Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1597–1599, September 2014.
- [47] G. M. Georgiou, "The field of values of a matrix and neural networks," *IEEE Transactions on Neural Networks and*, vol. 25, no. 9, pp. 1613– 1620, September 2014.
- [48] Z. Zhang, C. Lin, and B. Chen, "Global stability criterion for delayed complex-valued recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1704–1708, September 2014.
- [49] K. Subramanian, R. Savitha, and S. Suresh, "A metacognitive complexvalued interval type-2 fussy inference system," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1659–1672, September 2014.
- [50] S. Chen, X. Hong, J. Gao, and C. J. Harris, "Complex-valued B-spline neural networks for modeling and inverting Hammerstein systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1673–1685, September 2014.
- [51] K. Terabayashi, R. Natsuaki, and A. Hirose, "Ultrawideband directionof-arrival estimation using complex-valued spatiotemporal neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1727–1732, September 2014.
- [52] R. Song, W. Xiao, H. Zhang, and C. Sun, "Adaptive dynamic programming for a class of complex-valued nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1733–1739, September 2014.
- [53] M. K. Fleming and G. W. Cottrell, "Categorization of faces using unsupervised feature extraction," in *IJCNN International Joint Conference* on Neural Networks (IJCNN) 1990, vol. 2, June 1990, pp. 65–70.
- [54] G. W. Cottrell, "Face recognition using unsupervised feature extractor," in *International Conference on Neural Networks (ICNN)*, 1990, pp. 322– 325.
- [55] K. Fukushima, "Neocognitron A neural network model for a mechanism of pattern recognition unaffected by position shift," *IEICE Transactions on Fundementals (Japanees Edition)*, vol. J62-A, no. 10, pp. 658–665, October 1979, (in Japanese).
- [56] —, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.
- [57] W. M. Boerner, "Recent advances in extra-wide-band polarimetry, interferometry and polarimetric interferometry in synthetic aperture remote sensing and its applications," *IEE Proceedings – Radar, Sonar and Navigation*, vol. 150, no. 3, pp. 113 – 124, June 2003.
- [58] A. Hirose and S. Yoshida, "Generalization characteristics of complexvalued feedforward neural networks in relation to signal coherence," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, pp. 541–551, 2012.
- [59] 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.
- [60] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert, *Numbers*. (Chapter 3, Section 2): Springer-Verlag, 1983.
- [61] E. T. Copson, An Introduction to the Theory of Functions of a Complex Variable. Oxford: Clarendon Press, 1935.
- [62] 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.