Journal of Industrial Information Integration xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Industrial Information Integration

journal homepage: www.elsevier.com/locate/jii

High-dimensional information processing through resilient propagation in quaternionic domain

Sushil Kumar*, Bipin Kumar Tripathi

Department of Computer Science and Engineering, Harcourt Butler Technical University, Kanpur, India

A R T I C L E I N F O

Keywords:
Quaternionic-valued neural network
Quaternionic back-propagation learning
Quaternionic resilient propagation learning
Linear transformation
3D motion and interpretation
3D and 4D chaotic time series predictions

A B S T R A C T

This paper proposes a fast but effective machine learning technique for the high-dimensional information processing problems. A quaternion is the hyper-complex number in four dimensions which possesses four components in a single body with embedded phase information among their components. Thus, a neural network with quaternion as unit of information flow has ability to learn and generalize magnitude and argument of high-dimensional information simultaneously. The slow convergence and getting stuck into bad minima are the main drawbacks of the back-propagation learning algorithm; therefore learning technique with faster convergence is the demand of quaternionic neurocomputing. The basic issues of quaternionic back-propagation (ℍ-BP) algorithm have been combated by proposed resilient propagation algorithm in quaternionnic domain (ℍ-RPROP). The efficient and intelligent behavior of the proposed learning algorithm has been vindicated by the wide spectrum of benchmark problems. The simulations on different transformations demonstrate its ability to learn 3D motion, which is very desirable in viewing of objects through different orientations in many engineering applications. The critical thing is the training through mapping over straight line having a small number of points and eventually generalization over complicated geometrical structures (Sphere, Cylinder and Torus) possessing huge amount of point cloud data. The 3D and 4D time series prediction as benchmark applications are also taken into the account for the significant contribution of the proposed work. The substantial edge of the proposed algorithm is its quicker convergence with better approximation capability in high-dimensional problems.

1. Introduction

In the last decades, the information and communications technology (ICT) community has explored Industrial Information Integration Engineering (IIIE) as a new scientific sub-discipline which interrelates with various engineering disciplines in terms of scientific and engineering techniques for different industrial sectors [23]. IIIE collaborates with different industrial applications in the field of communication engineering, aerospace engineering, material engineering, bio engineering, energy engineering, and civil engineering, and provides to build the initial methodologies that assist the process of industrial information integration. Furthermore, it encompasses techniques for solving the complicated problems when building the infrastructure of information technology [24]. In the field of artificial neural networks as one of the specific engineering discipline of IIIE, the integration of high-dimensional information is also a challenging but complicated task. In the recent literature, it is also mentioned that the high-dimensional information is often confusing to human brains that do not possess the natural neural network to recognize and realize the objects in high-dimension [25]. In this direction, a learning machine with artificial neural networks is required to process the information in high-dimension. Although, the neural networks has been constructed for the integration of high-dimensional information through conventional real-valued neural network but it requires a large number of conventional neurons because it processes individual components of high-dimensional information [1]. Such networks are not only the complex and unnatural in structure but they also lead to slow processing and incapable to converge along their magnitude and phase components of high-dimensional information. Simultaneously, the conventional back-propagation learning method in neural network have issues like slow convergence and getting stuck local minima, hence degrade the efficiency of intelligent systems [2–3]. The recent publications in neural network bearing complex number as a unit of information [4–6] overcome from these problems in two-dimensional frame works. The quaternion is the four-dimensional hyper-complex number system that possesses four components and phase information along different components are embedded within it. This phenomenon provides the beauty of a quaternion that reflects the quality of information

* Corresponding author.
E-mail address: sushil0402k5@gmail.com (S. Kumar).

https://doi.org/10.1016/j.jii.2018.01.004
Received 4 November 2017; Received in revised form 19 January 2018; Accepted 19 January 2018

integration. This beauty provides an idea to develop a fast but effective machine learning technique for the high-dimensional information processing problems. This technique may help to researchers working in the field of industrial information integration where the fast processing of high-dimensional information is required because it is highly demanding in current researches. Thus, a technique based on neural network with quaternion as a unit of information flow efficiently learns and generalizes in high-dimension with fewer neurons. Although, the multiplication operation on quaternions is non-commutative in nature, but intelligent mathematical formulation will lead to its best use in high-dimensional neurocomputing. The back-propagation learning algorithm in quaternionic domain (ℍ-BP) is developed in recent past [7–12] using the concept of error gradient-decent optimization, but it suffers with its basic issues like getting stuck into local minima and slow convergence. Thus, an investigation of fast and efficient learning algorithm for the processing of high-dimensional information integration is highly demanding to overcome from the issues of back-propagation algorithm.

The concept of resilient propagation was floated by Riedmiller in 1993 [13], which had been exploited in single [14] and two dimension [5,15] problems, where it proved its momentousness. This paper proposes a quaternionic domain resilient propagation algorithm (ℍ-RPROP) for multilayered feed-forward networks in quaternionic domain and presents its exhaustive analysis through a wide spectrum of benchmark problems containing three or four dimension information and motion interpretation in space. The propagation of this procedure is based on the sign of partial derivatives of error function instead of its value as in back-propagation algorithm. The basic idea of the proposed algorithm is to modify the components of quaternionic weights by an amount Δ (update value) with a view to decrease the overall error and the sign of gradient of error function indicates the direction of weight update. Without increasing the complexity of algorithm, the proposed ℍ-RPROP algorithm is boosted by error-dependent weight backtracking step, which accelerates the training speed appreciably and provides better approximation accuracy. The neural network [7–12] and back-propagation algorithm in quaternionic domain (ℍ-BP) [15–18] has been widely applied in problems dealing with three and four dimensional information; recently its comparison with quaternionic scaled conjugate gradient (ℍ-SCG) learning scheme is presented in [11]. This paper proposes an ℍ-RPROP algorithm and compare with ℍ-BP and ℍ-SCG algorithms through application in 3D imaging and chaotic time series predictions. Though, ℍ-BP and ℍ-SCG learning algorithms can solve the typical class of 3D and 4D dimensional problems, but the proposed ℍ-RPROP algorithm has demonstrated its superiority over ℍ-BP and ℍ-SCG in all respects, which is reported by different statistical parameters.

The organization of the paper is as follows: Section 2 mainly focuses on the proper formulation of proposed resilient propagation algorithm in qiuaternionic domain (ℍ-RPROP). Section 3 discusses the comparative performance of the ℍ-RPROP algorithm with ℍ-BP and ℍ-SCG. The learning of various transformations in quaternionic-valued neural network through a straight line in 3D dimension and its subsequent generalization over complex geometric structures convey a way for efficient examination of the 3D images. This also shows an inherent ability to learn and generalize 3D motion with ℍ-RPROP. The chaotic time series prediction problems are also taken up to convince its applicability in high-dimensional applications. The final inferences of the proposed work are described in the Section 4.

2. Fast learning in quaternionic domain

In this paper, we propose a quaternionic resilient propagation (ℍ-RPROP) as fast learning scheme for multi-layered neural networks in the quaternion domain. A three-layered $(L - M - N)$ network is considered in quaternionic domain that employs L inputs, M hidden neurons and N output neurons respectively. All the input-output, connection weights and biases are quaternion. Quaternion algebra is incorporated to derive learning algorithm through gradient-decent optimization into a neural network. Let consider q_l to be lth, $(l = 1 \ldots L)$ quaternionic input at the input layer of the network, and presented as

$$\boldsymbol{q}_l = \Re(\boldsymbol{q}_l) + \Im_1(\boldsymbol{q}_l)\boldsymbol{i} + \Im_2(\boldsymbol{q}_l)\boldsymbol{j} + \Im_3(\boldsymbol{q}_l)\boldsymbol{k} \in \mathbb{H} \,. \tag{1}$$

where, $\Re(\boldsymbol{q}_l)$ denotes the real component and $\Im_1(\boldsymbol{q}_l)$, $\Im_2(\boldsymbol{q}_l)$, and $\Im_3(\boldsymbol{q}_l)$ denote the imaginary components of quaternionic variable $(\boldsymbol{q}_l)$. Its bases $(\boldsymbol{i},\ \boldsymbol{j},\ \boldsymbol{k})$ follow the vector cross product properties [19] as $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$ and $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k},\ \boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i},\ \boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$ respectively. The output of m^{th} neuron $(m = 1 \ldots M)$ in the hidden layer of the network can be obtained by employing the activation function on internal potential $(\boldsymbol{V}_m)$:

$$\boldsymbol{Y}_m = f_{\mathbb{H}}(\boldsymbol{V}_m) = f_{\mathbb{H}}\left(\sum_{l=1}^{L} \boldsymbol{w}_{lm} \otimes \boldsymbol{q}_l + \ \boldsymbol{\alpha}_m\right). \tag{2}$$

where, the symbol ⊗ denotes the non-commutative multiplication of quaternions based on Hamilton properties [20,21]. Let $\boldsymbol{w}_{lm}$ be the weight that connects lth input to mth hidden neurons, $\boldsymbol{\alpha}_m$ is the bias of mth neuron. Let $f_{\mathbb{H}}(\ )$ be the split type [10–13] activation function in quaternionic domain which is obtained as an extension of split type function in complex domain [3,5,7], over any general function (f) applied on components independently and expressed as

$$f_{\mathbb{H}}(\boldsymbol{q}) = f(\Re(\boldsymbol{q})) + f(\Im_1(\boldsymbol{q}))\boldsymbol{i} + f(\Im_2(\boldsymbol{q}))\boldsymbol{j} + f(\Im_3(\boldsymbol{q}))\boldsymbol{k} \,. \tag{3}$$

Let $\boldsymbol{w}_{mn}$ be the synaptic weight from mth hidden to nth output neuron and $\boldsymbol{\beta}_n$ is the bias of nth output neuron. Let V_n be the internal potential of nth, $(n = 1 \ldots N)$ neuron at output layer, then output of neuron can be defined as:

$$\boldsymbol{Y}_n = f_{\mathbb{H}}(\boldsymbol{V}_n) = f_{\mathbb{H}}\left(\sum_{m=1}^{M} \boldsymbol{w}_{mn} \otimes \boldsymbol{Y}_m + \ \boldsymbol{\beta}_n\right). \tag{4}$$

2.1. The back-propagation algorithm in quaternionic domain

The back-propagation learning in quaternionic domain for feed-forward neural network is originated from the gradient descent approach. Let error $\boldsymbol{e}_n = \boldsymbol{Y}_n^D - \boldsymbol{Y}_n = \Re(\boldsymbol{e}_n) + \Im_1(\boldsymbol{e}_n)\boldsymbol{i} + \Im_2(\boldsymbol{e}_n)\boldsymbol{j} + \Im_3(\boldsymbol{e}_n)\boldsymbol{k}$ be the difference between desired $(\boldsymbol{Y}_n^D)$ and actual $(\boldsymbol{Y}_n)$ output of n^{th} neuron. The weight update equation can be obtained by the minimization of real valued error function:

$$E = \frac{1}{2N}\sum_{n=1}^{N} \|\boldsymbol{e}_n\|^2 = \frac{1}{2N}\sum_{n=1}^{N} (\boldsymbol{e}_n \otimes \boldsymbol{e}_n^{\star}) \,. \tag{5}$$

where, $\|\boldsymbol{e}_n\| = \sqrt{(\boldsymbol{e}_n \otimes\ \boldsymbol{e}_n^{\star})} = \sqrt{(\Re(\boldsymbol{e}_n))^2 + (\Im_1(\boldsymbol{e}_n))^2 + (\Im_2(\boldsymbol{e}_n))^2 + (\Im_3(\boldsymbol{e}_n))^2}$ is the norm of the error function and $\boldsymbol{e}_n^{\star}$ is the conjugate of quaternionic variable. The weight update rules at the output layer of the network can be computed by the gradient of error function with respect to the corresponding weight $(\nabla_{(\boldsymbol{w}_{mn})}E)$ and bias $(\nabla_{(\boldsymbol{\beta}_n)}E)$:

$$\begin{aligned}\Delta\boldsymbol{w}_{mn} &= -\eta\nabla_{(\boldsymbol{w}_{mn})}E = -\eta\,(\nabla_{\Re(\boldsymbol{w}_{mn})}E + \nabla_{\Im_1(\boldsymbol{w}_{mn})}E + \nabla_{\Im_2(\boldsymbol{w}_{mn})}E \\ &\quad + \nabla_{\Im_3(\boldsymbol{w}_{mn})}E) \\ &= \tfrac{\eta}{N}(\boldsymbol{e}_n \odot f'_{\mathbb{H}}(\boldsymbol{V}_n)) \otimes \boldsymbol{Y}_m^{\star} \,.\end{aligned} \tag{6}$$

$$\Delta\boldsymbol{\beta}_n = -\eta\nabla_{(\boldsymbol{\beta}_n)}E = \frac{\eta}{N}\boldsymbol{e}_n \odot f'_{\mathbb{H}}(\boldsymbol{V}_n) \,. \tag{7}$$

where $\eta \in \mathbb{R}^{+}$ is the learning rate and ⊙ denotes the Hadamard product. The derivative of quaternionic activation function $(f'_{\mathbb{H}}(\boldsymbol{q}))$ can be obtained by the derivative function (f') that applies to each component