Optics Communications 399 (2017) 1–12

Contents lists available at ScienceDirect

Optics Communications

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

System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm

Danshi Wang*, Min Zhang*, Ze Li, Chuang Song, Meixia Fu, Jin Li, Xue Chen

State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Keywords:
Fiber optics communications
Digital signal processing
Machine learning
Artificial neural network

ABSTRACT

A bio-inspired detector based on the artificial neural network (ANN) and genetic algorithm is proposed in the context of a coherent optical transmission system. The ANN is designed to mitigate 16-quadrature amplitude modulation system impairments, including linear impairment: Gaussian white noise, laser phase noise, in-phase/quadrature component imbalance, and nonlinear impairment: nonlinear phase. Without prior information or heuristic assumptions, the ANN, functioning as a machine learning algorithm, can learn and capture the characteristics of impairments from observed data. Numerical simulations were performed, and dispersion-shifted, dispersion-managed, and dispersion-unmanaged fiber links were investigated. The launch power dynamic range and maximum transmission distance for the bio-inspired method were 2.7 dBm and 240 km greater, respectively, than those of the maximum likelihood estimation algorithm. Moreover, the linewidth tolerance of the bio-inspired technique was 170 kHz greater than that of the k-means method, demonstrating its usability for digital signal processing in coherent systems.

1. Introduction

The combination of advanced modulation formats and digital signal processing (DSP)-assisted coherent detection has significantly impacted the field of optical communication [1]. A great deal of researches have focused on multilevel quadrature amplitude modulation (QAM) since it can simultaneously achieve relatively high spectral efficiency and relatively long transmission distance [2]. However, the performances of QAM systems are significantly deteriorated by various impairments, such as the phase noise from the transmitter laser and local oscillator (LO) [3], the imbalance between the in-phase (I) and quadrature (Q) components incurred by I/Q modulator imperfection [4], and, in particular, the fiber nonlinearity caused by a large launch power [5]. Among the various fiber nonlinearities, the nonlinear phase noise (NLPN) is regarded as a dominant distortion factor; it is induced by the interaction between the signal and the amplified spontaneous emission (ASE) noise from inline optical amplifiers via the fiber Kerr effect, which is known as self-phase modulation (SPM) [6]. To address these problems, various DSP-based methods have been proposed: laser linewidth estimation [7], I/Q imbalance compensation [8], maximum likelihood estimation (MLE) [9], digital back-propagation [10], inverse Volterra-based equalization [11], and so on. However, some of these methods are quite complex; in addition, the nonlinear tolerance is dependent upon the particular transmission scenario. Therefore, the use of advanced DSP algorithms for nonlinearity compensation merits further research.

Machine learning techniques offer powerful tools to solve various problems in many areas [12,13]. Recently, optical communication researchers have also shown interest in machine learning algorithms, such as the Bayesian filter [14], hidden Markov model [15], and k-means algorithms [16], as well as our recently proposed k-nearest neighbors [17] and support vector machine method [18,19]. However, these machine learning methods each solve only one of the numerous system impairments, which means that different transmission situations require different algorithms. Moreover, as is well known, modern optical networks are often hybrid networks consisting of various fiber links, including dispersion-shifted, dispersion-managed, and dispersion-unmanaged links, but the abovementioned fiber nonlinearity compensation algorithms each focus on only one type of fiber link. Since dynamic routes across multiple fiber spans are becoming more important and are replacing traditional semi-static single-type fiber connections, these algorithms may become weak or invalid. Hence, it is beneficial to investigate algorithms that can not only combat various system impairments simultaneously, but also be usable in different

* Corresponding authors.
E-mail addresses: wang_danshi@163.com (D. Wang), mzhang@bupt.edu.cn (M. Zhang).

http://dx.doi.org/10.1016/j.optcom.2017.04.050
Received 21 November 2016; Received in revised form 20 April 2017; Accepted 21 April 2017

fiber links. To realize these objectives, another class of machine learning algorithms that are inspired by biological systems and are referred to as bio-inspired computing has been developed [20].

Nature is considered to be the most remarkable machine for solving complex, large-scale problems with robust, adaptable, and efficient solutions, and it has inspired the development of evolutionary algorithms including genetic algorithm (GA) [21], particle swarm optimization [22], ant colony optimization [23], and simulated annealing [24]. Among the various bio-inspired methods, artificial neural networks (ANNs), which are among of the most popular algorithms, have been widely applied in image processing, data mining, pattern recognition, and artificial intelligence [25]. Recently, a computer program AlphaGo from Google DeepMind, who for the first time beat a professional player at the board game Go, has attracted worldwide attention [26]. The core idea of AlphaGo comes from the technique of deep learning, which is also based on ANN [27].

In optical communication systems, ANNs have realized modulation format recognition [28], optical performance monitoring [29], and photonic label routing [30]. Meanwhile, ANNs based on multiplayer perception are also considered to be useful DSP tools in wireless channels, such as those employed in mobile communication [31], visible light communication [32], and satellite communication [33]. The usability of ANNs in generalized nonlinear communication channels has also been analyzed and discussed [34]. In optical fiber communication systems, the ANN-based nonlinear equalizers have been fully investigated in orthogonal frequency division multiplexing (OFDM) system [35–40]. However,in addition to fiber nonlinearity mitigation, linear noise, laser phase noise, and I/Q imbalance correction are also important applications of ANNs that require further investigation.

In this paper, a bio-inspired detector based on an ANN and a GA is proposed to overcome various system impairments in a 16QAM coherent optical system. The theory of ANN is described in detail, and the 16QAM signal detection method is proposed. Without any prior information or heuristic assumptions, the ANN is found capable of learning and capturing impairment information from observed data and of regenerating the I and Q components as output vectors. The numerical results obtained in this study demonstrate that the ANN can effectively mitigate laser phase noise and I/Q imbalance. Moreover, three types of fiber links, namely, dispersion-shifted, dispersion-managed, and dispersion-unmanaged links, are investigated. Compared with conventional methods, the ANN achieves higher nonlinear tolerance in all three cases, which demonstrates the feasibility of using the ANN for fiber nonlinearity mitigation. Finally, the results are discussed and analyzed.

2. Basic principle of artificial neural network

The term "neural network" has its origins in attempts to find mathematical representations of information processing in biological systems, which are built of a lot of interconnected neurons. In rough analogy, ANNs are built out of a densely interconnected set of simple units, where each unit takes a number of real-valued inputs (possibly the outputs of other units) and produces a single real-valued output (which may become the input to many other units) [41].

2.1. Neuron

One type of the ANN system is based on a unit called as *neuron*, as illustrated in Fig. 1. The neuron works as a nonlinear threshold function with multi-input and single-output. The input vector $\boldsymbol{X} = [x_1, x_2, \ldots, x_n]$ and weight vector $\boldsymbol{W} = [w_1, w_2, \ldots, w_n]$ are used to calculate an input weighted sum $\sum_{i=1}^{n} w_i x_i$. The symbols θ and o denote the threshold value and output value, respectively. If the input weighted sum is greater than the threshold value, the neuron is activated. Then the output computed by the neuron is

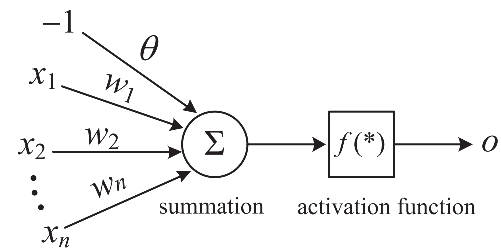

Fig. 1. The schematic diagram of a neuron. It works as a nonlinear threshold function with multi-input and single-output.

$$o(x_1, \ldots, x_n) = f\left(\sum_{i=1}^{n} w_i x_i - \theta\right) \tag{1}$$

where $f(*)$ represents the mapping relation between the input and output, referred to as *activation function*. Four types of activation functions are widely used for the ANN, as given in Fig. 2.

Different function types have different characteristics and thus fit for different applications, as shown in Fig. 2. Through selection of the suitable activation function and adjustment of the weight and threshold values, in theory, the neuron can generate desired output value from the input vector. Actually, the purpose of the training process is to acquire the appropriate weight and threshold values for the given activation functions through training data.

2.2. Neuron training rule

Let us begin by understanding how to learn the weight vector for a single neuron. To simplify the notation, we imagine an additional constant input $x_0 = -1$, and replace the threshold θ with w_0. Then we can rewrite the neuron output function (1) as:

$$o(\boldsymbol{X}) = f\left(\sum_{i=0}^{n} w_i x_i\right) \tag{2}$$

The space H of the candidate hypotheses is the set of all possible real-valued weight vectors. Then the precise learning problem is to determine a weight vector in H that drives the neuron to produce the desired output value.

The *gradient descent algorithm* is one of the most useful tools to solve this learning problem. The key idea of this algorithm is to use a gradient descent method to search for the hypothesis space of possible weight vectors until the weights that best fit the training data are found. To derive a weight learning rule for a neuron unit, we specify a measure for the training error of a hypothesis (weight vector), relative to the training data. Although there are many ways to define this error, the conventionally used convenient measure is

$$E(\boldsymbol{W}) \equiv \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2 \tag{3}$$

where D is the set of training data, t_d is the target output for training data d, and o_d is the actual output of the neuron for training data d. By this definition, $E(\boldsymbol{W})$ is simply half the squared difference between the target output t_d and the neuron output o_d, summing over the entire training set. Here we characterize E as a function of $\boldsymbol{W}$ because the neuron output o depends on the weight vector. Under certain conditions, the hypothesis that minimizes E is also the most fitting hypothesis in H.

To understand the gradient descent algorithm, it is helpful to visualize the entire hypothesis space of the possible weight vectors and their associated E values. Fig. 3 shows an example of hypothesis space in two-dimension (2D). Starting with an arbitrary initial weight vector and subsequently modifying it iteratively in small steps, the gradient descent search determines a weight vector that minimizes E. At each step, the weight vector is altered in the direction that produces the