Photonic Nonlinear Transient Computing with Multiple-Delay Wavelength Dynamics

Romain Martinenghi, Sergei Rybalko, Maxime Jacquot, Yanne K. Chembo, and Laurent Larger

UMR CNRS FEMTO-ST 6174/Optics Department, University of Franche-Comté, 16 Route de Gray, 25030 Besançon Cedex, France

(Received 1 November 2011; published 15 June 2012)

We report on the experimental demonstration of a hybrid optoelectronic neuromorphic computer based on a complex nonlinear wavelength dynamics including multiple delayed feedbacks with randomly defined weights. This neuromorphic approach is based on a new paradigm of a brain-inspired computational unit, intrinsically differing from Turing machines. This recent paradigm consists in expanding the input information to be processed into a higher dimensional phase space, through the nonlinear transient response of a complex dynamics excited by the input information. The computed output is then extracted via a linear separation of the transient trajectory in the complex phase space. The hyperplane separation is derived from a learning phase consisting of the resolution of a regression problem. The processing capability originates from the nonlinear transient, resulting in nonlinear transient computing. The computational performance is successfully evaluated on a standard benchmark test, namely, a spoken digit recognition task.

DOI: 10.1103/PhysRevLett.108.244101 PACS numbers: 05.45.−a, 05.90.+m, 07.05.Mh, 89.75.−k

In early 2000, the brain research and neural network computing communities independently proposed novel computational principles [1] suspected to mimic actual calculation and processing tasks that have been observed and studied in the brain. These computational principles, referred to as echo state network [2], liquid state machine [3], and the generic term reservoir computing (RC), are definitely different with respect to the standard Turing machine principles widely implemented in electronic digital processors. Instead of processing the calculation tasks step-by-step with static states stored in memories, this new principle is based on computational power performed by complex nonlinear transient motion developed in the high dimensional phase space of a nonlinear system excited by an input signal representing the information to be processed. The complex dynamics is usually materialized by a network of neurons (as in the brain) or by any spatially extended network of coupled nonlinear dynamical nodes.

The corresponding generic architecture is depicted in Fig. 1(a), where strong similarities can be seen compared to standard recurrent neural networks (RNN): an input layer is dedicated to the injection of the input information (input connectivity matrix W^I) into a complex interconnected network of dynamical nodes (internal network connectivity matrix W^D); an output layer (readout matrix W^R) is dedicated to the extraction of the result, computed from the nonlinear transient developed by the network dynamics consequent to the injected input signal. Since one of our aims is to transpose these concepts into physics and into a real-world experimental demonstrator, our system will be referred in the remaining part of the Letter as nonlinear transient computing (NTC) [4]). NTC is suggested with the intention to reflect more clearly the actual physical origin of the approach, in a way which is expected to be more meaningful for physicists and for the nonlinear dynamics community, although historically the computer or brain science community are referring to echo state network, liquid state machine, or RC.

There are at least two strong differences in NTC compared with RNN. (i) The internal network structure W^D is assumed to be fixed and not to be optimized according to a learning procedure as it is for RNN. This feature is particularly suited to a physical implementation, since reconfiguration flexibility typically brought by a numerical implementation of the RNN is not required anymore. (ii) A second difference compared with RNN consists in the fact that the learning procedure in NTC is strongly simplified. It is reduced to the determination of readout coefficients (matrix W^R), which can be easily calculated from a simple regression technique.

A significant innovation in NTC was recently proposed and implemented electronically [5] via the introduction of a delay dynamics instead of the classical network of dynamical nodes. This implemented original theoretical solution appeared also as a convenient technical one, since it provides a simpler way for real-world physical implementation, compared to the usually adopted topology of a network [6]. Such an approach can be justified by a known analogy between delay dynamics and spatiotemporal dynamics (as a RNN). Both are infinite-dimensional, and the space-time representation of a delay dynamics was already proposed 20 years ago [7], introducing on the one hand a discrete time variable corresponding to one delay step forward and, on the other hand, a virtual continuous space variable corresponding to the short time-scale fluctuations within a time-delay interval. As illustrated in Fig. 1(b), virtual nodes in delay dynamics can be defined as temporal positions within a time-delay interval τ_D, which are separated by a "node distance" $\delta\tau$ [5]. This fixes the number of virtual nodes within a time-delay interval, $N = \tau_D/\delta\tau$

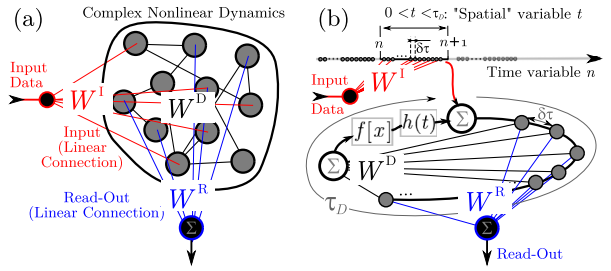

FIG. 1 (color online). Principles of an NTC. (a) With a spatially extended dynamical network. (b) With multiple delay feedback dynamics.

(150 in our case). The connectivity of the resulting virtual spatiotemporal network is achieved via two dynamical mechanisms. The neighboring nodes are linearly coupled via the characteristic time τ of the oscillator impulse response $h(t)$ ($\delta\tau \simeq \tau/5$ for an optimal adjacent node spacing, as found in Ref. [5]). Additionally, to this "short-distance" coupling, long-distance ones are also present. The proposed multiple-delay topology results in a denser internal dynamics connectivity of the equivalent network, with however still a sparse connectivity involving only N_I nodes among the N possible ones. Each node is thus subject to a nonlinear transformation (function $f[x]$) applied to a linear combination, with random weights w_i^D, of a few N_I "previous nodes" randomly defined among the N possible ones. Each corresponding delay is defined as $k_i\delta\tau$, $k_i \in [1, \ldots, N]$ for $i = 1$ to $N_I(= N/10)$.

Compared to the very recent NTC demonstrations with delay dynamics [5,8,9], we propose an original photonic implementation of RC where the dynamical variable is the wavelength of a tunable laser, and also a multiple-delayed feedback topology intended to enhance the internal connectivity of the equivalent dynamical network. While evaluating the processing capability of our approach through its performance on a standard benchmark test of spoken digit classification, we found that our multiple-delay photonic system with a reduced number of nodes exhibits a comparable computational efficiency, with word error rate (WER) of the same order as the best results achieved so far for the same test.

The operating principles of our NTC can be summarized as follow [see Fig. 1(b)]. Input information to be processed is exciting each virtual spatial node of the NTC dynamics; this node addressing procedure is practically performed by temporal division multiplexing, consisting of a random spreading of each input sample over all the nodes within a time-delay interval. The spreading is ruled by a sparse and random input connectivity matrix W^I applied to each of the original information samples. Each sequence of a spread input information undergoes (i) a nonlinear transformation, operating on a randomly weighted linear combination of N_I delayed feedback, and (ii) a linear filtering via an impulse response $h(t)$, which is limiting the rate of change in the feedback loop. The computed output of the NTC is obtained by a readout procedure, corresponding to a linear separation of the transient motion in the N_D-nodes dynamics phase space. The readout coefficients w_k^R are determined according to a learning phase, which consists of solving a regression problem from known pairs of transient-output pairs.

The experimental setup is depicted in Fig. 2. It is based on an optoelectronic architecture originally designed to display an Ikeda-like nonlinear delay dynamics [10]. The nonlinear function $[f[x] = \sin^2(x + \Phi_0)]$ is provided by a tunable interference phenomena: an imbalanced birefringent interferometer is seeded by the light emitted at $\sim$1.5 μm by a tunable two-electrode distributed Bragg reflector (DBR) laser diode. The wavelength is varied through an injection current I_{DBR} thus changing the interference condition, while another electrode (current I) serves as the usual input for setting the output power. The interferometer output is detected by a photodiode, and an electronic feedback circuit performs the bandpass filtering ruling the dynamics of the oscillator [impulse response $h(t)$]. This filtered signal is the one used for the NTC readout. The input information is added to the filtered signal, the resulting sum being nonlinearly transformed through $f[x]$, and multiple-delayed by an FPGA (field programmable gate array) board. The FPGA is programmed to perform a flexible and reconfigurable multiple-delay line, in which several elementary first in first out memories are implemented together with weights $w_{k_i}^D$ for each delay. The weighted sum of the multiple delayed signal is amplified, combined with an offset, and serves finally as the input drive I_{DBR} of the laser wavelength tuning electrode. The normalized dynamics can be written as follows:

$$\frac{1}{\theta}\int_{t_0}^{t} x(\xi)d\xi + x(t) + \tau\frac{dx}{dt}(t) = x_0 + \beta\sin^2\left\{\sum_{i=1}^{N_I} w_{k_i}^D[x(t - k_i\delta\tau)] + u(t) + \Phi_0\right\}, \quad (1)$$

where x is the output of the bandpass filter (also the signal used for the readout), u is the input data, Φ_0 is an offset

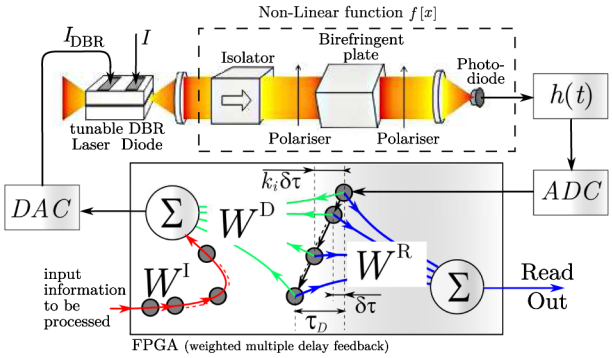

FIG. 2 (color online). Optoelectronic wavelength setup performing an NTC as multiple-delay feedback dynamics.

phase ruling the operating point along the nonlinear $\sin^2$ function, β is the loop gain set by the feedback amplification, and x_0 is a formal integration constant which guaranties a zero mean value for the bandpass filter output. The delayed feedback coefficients $w_{k_i}^D$ are programmed in the FPGA and they are randomly defined from a uniform distribution; as already stated, they are introduced for an enhanced internal connectivity compared to the single-delay case [5,8,9]. The response times $\theta = 1.59$ ms and $\tau = 7.95$ μs are associated to the low and high cut-off frequencies (100 Hz and 20 kHz) of the bandpass filter, respectively, and $\delta\tau \simeq 1.59$ μs is the elementary time delay determining the spacing between two adjacent virtual nodes. There are $N = 150$ different nodes and $N_I = 15$ randomly distributed delayed feedbacks, the largest delay being $N\delta\tau \simeq 238$ μs.

In order to ensure a stable fixed point solution in an input-free operation, the feedback gain β has to be set below the oscillation threshold defined as 1, calibrated experimentally as the minimum gain leading to a rising oscillation, while the nonlinear function is operating around its maximum linear slope. In the experiment β is set to ca. 0.5. The amplitude of the input information is set so that it induces a large scan of the nonlinear function (around 1–2 extrema, typically $\Delta u \simeq \pi$).

Each single input information consists of a sequence of samples, each of which has to be spanned over all the nodes within a time delay. The spanning is performed by an input connectivity matrix W^I, randomly but uniquely defined for each input information. The temporal waveform to be injected in the setup is derived from a computed matrix product between W^I and M_c, the latter being a 2D representation of the input information (the digit cochleagram in the case of the spoken digit recognition task, see below). The resulting matrix, after being horizontally unfolded, results in a 1D waveform $u(t)$, which is injected into the multiple-delay dynamics via the programming of an arbitrary waveform generator (AWG, Lecroy ArbStudio 1102). For each input information signal, the full transient response is recorded by a digital scope. An off-line postprocessing is then performed for both the training and testing stages. The training consists in the resolution of a regression problem and results in the definition of N readout coefficients w_k^R that are leading to an optimally correct output, for each input information belonging to a training subset. The efficiency of such an optimal readout is finally evaluated on a complementary subset of input-output pairs. In principle, the FPGA could also be programmed to implement a direct on-line readout, as soon as the coefficients are known. This configuration would result in a real-time processing, for which the processing time would be only limited by the analogue bandwidth of the delay dynamical system. For the sake of simplicity, this testing phase was fully processed off-line, right after the training phase.

In order to evaluate the processing efficiency, we performed a standard task typically used in other RNN or NTC reports [5,8,9] and consisting of spoken digit recognition. The main goal of this standard classification test is to recognize a pronounced digit among the ten possible ones from 0 to 9. The spoken digit database corresponds to 500 digits extracted from the TI46 speech corpus [11]. The digits are pronounced by 10 different female speakers uttering the 10 digits 5 times. Following a standard preprocessing task typically performed in many similar acoustic speech recognition tasks, the 1D acoustic waveform sampled at 12.5 kHz is transformed into a 2D frequency-time representation (matrix M_c), a so-called cochleagram (the Lyon ear model), which provides a monitoring of the average acoustic spectrum ($N_f = 86$ frequency channels) evolving during the spoken digit pronunciation time (N_s is typically 80 samples, varying between 32 and 130). The cochleagram represents the input information signal to be spread over the nodes of the NTC dynamics. This spreading is achieved via a sparsely and randomly determined connectivity matrix W^I with dimensions $N \times N_f$ (sparsity 0.1, and non-zero elements being ± 1 randomly distributed). The 2D input data to be injected in the NTC thus consists of an $N \times N_s$ matrix M_u (see Fig. 3), with a number of columns representing the digit duration (variable number of samples N_s). Each value of the N elements (the index of a virtual node within τ_D) of a column is thus built according to W^I. The action of W^{I} on the cochleagram can be interpreted as a random contribution of some selected frequency components at some selected time. The resulting input matrix $M_u = W^I \times M_c$ is converted in a 1D signal, $u(t)$ (programmed in the AWG), simply by unfolding horizontally each successive column of M_u. The input $u(t)$ in Eq. (1) thus consists of a sequence of N_s time intervals of length τ_D, each of which consisting of N samples defining the N virtual node amplitudes to be addressed. The 1D transient $x(t)$ of the NTC, ruled by Eq. (1), can be represented in the same 2D way, with a matrix M_x corresponding to the spatiotemporal response of the node amplitudes. The readout of this transient response consists then of a matrix product $(W^R)^t \times M_x = B$, expected to result in a target output

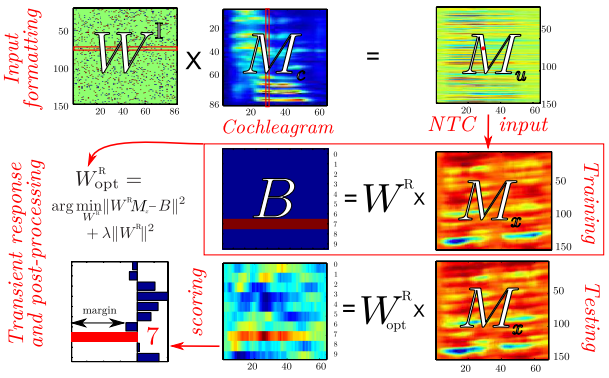

FIG. 3 (color online). Illustration of the NTC processing on the spoken digits; example of the spatiotemporal representation for the digit "seven."

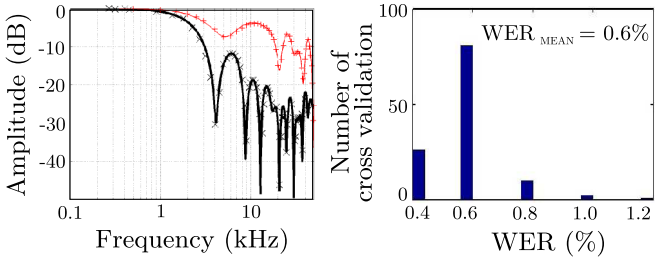

FIG. 4 (color online). Experimental results. Left: Transfer functions for 15 (red, thinner line) and 150 (black, thicker line) delayed feedbacks with random weights (theoretical curves in full lines, experimental data in symbols). Right: WER histograms with 15 delayed feedbacks and a total of 120 cross validations.

easily revealing the right digit (matrix B). The optimal matrix W^R_{opt} is calculated after a standard ridge-regression procedure (regression parameter $\lambda = 10^{-3}$),

$$W^R_{\text{opt}} = \text{argmin}_{W^R} ||(W^R)^t M_x - B||^2 + \lambda ||W^R||^2, \quad (2)$$

for which a training subset of 475 digits is used among the 500 available in the database. The complementary subset of 25 digits is used to test the performance of the learnt W^R_{opt}. This training and testing procedure is repeated for the 20 different possible partitions (cross validation). This results in the calculation of a WER, which is statistically limited by the size of the 500 tests actually performed during the cross validation.

Excellent performances (within the limit of statistical significance of 0.2%) have been obtained in terms of WER for such a complex classification test. Figure 4 displays the theoretical and experimental transfer functions of our multiple-delay feedback loops with random weights. Their complex frequency dependence confirms that the number of feedback loops, as well as the correct determination of the readout weighting coefficients is a critical process for the classification task. As shown in Fig. 4, the average WER performance of our setup is $0.6 \pm 0.2\%$ with only 15 feedback delay lines, thereby implying that multiple feedbacks enable us to achieve performance comparable to state-of-the-art systems with a relatively limited number of readout parameters to compute. This result validates the computational efficiency of our photonic NTC setup performed by a multiple-delay wavelength dynamics, and it opens the way to the future optimization of delay dynamics NTC processors, via a more accurate conceptual analogy between multiple-delay dynamics and classical network dynamics.

We have demonstrated experimentally the efficiency of a photonic neuromorphic processor. The proposed setup is based on a multiple delayed feedback wavelength dynamics providing an enhanced dynamical connectivity. In addition to the recent success in applications of complex dynamics [12–14], photonic nonlinear delay dynamics are confirmed as an efficient and flexible solution for the practical implementation of NTC. Future work will be devoted to ultrafast photonic versions making use of standard optical telecommunication devices and principles [13]. Processing speed for the recognition of a single spoken digit is around 20 ms, but this can be improved down to 100 ns with telecom-grade devices. Photonic NTC based on multiple-delay dynamics should offer an efficient and ultrafast hardware solution for future neuromorphic computers. Many fundamental issues still need to be addressed (quantitative connections between delay dynamics features and computational power, or optimal topology and system architecture for advanced functionalities such as plasticity and integrated learning capability). We also anticipate that information theory and nonlinear dynamics should be explored together to provide a theoretical framework capable of describing and optimizing the computational efficiency provided by delay dynamics-based NTC [15].

This work was supported by the PHOCUS project (FP7-240763) and the Région Franche-Comté. L. L. thanks I. Fischer and C. Mirasso for helpful discussions.

[1] H. Jaeger, W. Maass, and J. Principe, Neural Netw. **20**, 287 (2007).
[2] H. Jaeger and H. Haas, Science **304**, 78 (2004).
[3] W. Maass, T. Natschlager, and H. Markram, Neural Comput. **14**, 2531 (2002).
[4] N. Crook, Neurocomputing **70**, 1167 (2007).
[5] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Nat. Commun. **2**, 468 (2011).
[6] K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. Campenhout, Opt. Express **16**, 11182 (2008).
[7] F. T. Arecchi, G. Giacomelli, A. Lapucci, and R. Meucci, Phys. Rev. A **45**, R4225 (1992).
[8] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Opt. Express **20**, 3241 (2012).
[9] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Sci. Rep. **2**, 287 (2012).
[10] J.-P. Goedgebuer, L. Larger, and H. Porte, Phys. Rev. Lett. **80**, 2249 (1998).
[11] Texas Instruments, N.I.S.T., http://www.ldc.upenn.edu/Catalog/docs/LDC93S9/ti46.readme.html (1991).
[12] L. Larger and J. M. Dudley, Nature (London) **465**, 41 (2010).
[13] R. Lavrov, M. Jacquot, and L. Larger, IEEE J. Quantum Electron. **46**, 1430 (2010).
[14] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, Nature Photon. **2**, 728 (2008).
[15] D. Woods and T. J. Naughton, Nature Phys. **8**, 257 (2012).