COMPUTING WITH TRANSIENT STATES GENERATED BY DELAY-COUPLED SYSTEMS

L. Appeltant[1], M.C. Soriano[2], G. Van der Sande[1], J. Danckaert[1], I. Fischer[2] and C.R. Mirasso[2]
[1]Applied Physics Research Group (APHY),Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium
[2]Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (UIB-CSIC), Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain
claudio@ifisc.uib-csic.es

Abstract. Inspired by the way the brain process information, the recently introduced paradigm of reservoir computing has been proven to be very successful in solving certain tasks that are inherently difficult for traditional computers. Reservoir computing based on delay-coupled systems has been shown to perform as well as traditional reservoir-computing methods allowing to replace the large networks by a small number of delayed-coupled dynamical elements. Its simplest manifestation, a single nonlinear node subject to delayed feedback, provides excellent performance in certain tasks as spoken digit recognition and time series predictions.

Keywords: information processing, reservoir computing, delay-coupled systems, transient dynamics.

I. Introduction

Nonlinear systems subject to delayed feedback and/or delayed coupling, have attracted a lot of attention, both, from the fundamental and application point of view [1]. It has been shown that delay has an ambivalent impact on the dynamical behaviour, either stabilizing or destabilizing them. Simply by tuning two parameters, the feedback coupling time and the feedback strength, a variety of behaviours, ranging from stable via periodic and quasi-periodic oscillations to deterministic chaos can be realized. Applications of delay systems are resulting more and more interesting; chaos-based communications, random number generation or rainbow refractometry are only few examples.

One of the simplest possible delay systems, and the easiest to implement, consists of a single nonlinear node whose dynamics is influenced by its own output a time τ in the past. Well-studied examples include a semiconductor laser whose output light is fed back to the laser by an external mirror at a certain distance or an electronic circuit whose output voltage is also fed back after a certain time τ. In this article we show that the transient responses generated by such systems can be useful to process information. In particular we prove their effectiveness in spoken digit recognition and time series prediction tasks.

Reservoir computing (RC) is a bio-inspired, machine-learning paradigm that exhibits state of the art performance for processing empirical data. Computationally hard tasks, such as speech recognition, object recognition or chaotic time series prediction, amongst others, can be efficiently performed. RC bases in the brain ability to process information, by generating patterns of transient neuronal activity excited by input sensory signals. In this sense, RC mimics the way our brain process information. Traditional RC implementations [2-4] are generally composed of three distinct parts: an input layer, the reservoir, and an output layer, as shown in Fig. 1 left panel. Through the input layer the input signals are fed into the reservoir via fixed random weight connections. The reservoir usually consists of a large number of randomly interconnected nonlinear nodes, constituting a recurrent network (a network with internal feedback loops). The input signals generate transient responses of the networks that are then read out at the output layer via a linear weighted sum of the individual node states. Contrary to recurrent networks that are notoriously difficult to train, RC solves this problem by keeping the reservoir connections fixed and training only the output weights. As a result of this training procedure the system is capable to generalise, i.e. process unseen inputs or attribute them to previously learned classes.

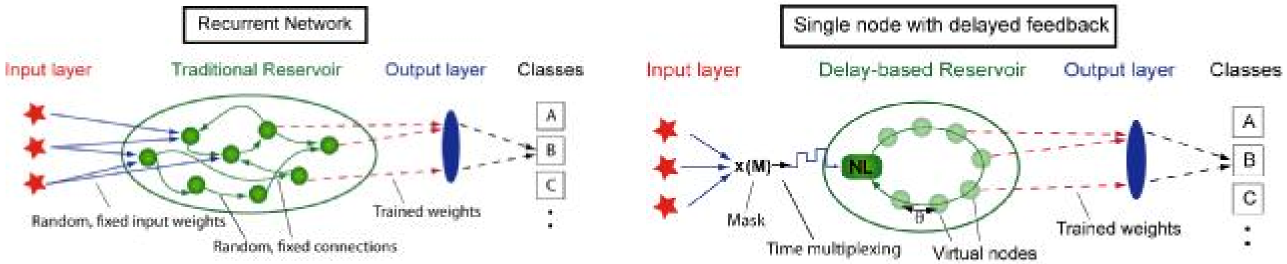

Figure 1. Left panel: sketch of a traditional reservoir computing setup; Right panel: reservoir computing based on a single nonlinear element subject to delayed feedback.

Some key properties are essential for RC to properly process information. Firstly it must nonlinearly transform the input signal into a high dimensional state space. The commonly high number of used nodes in the reservoir (hundreds/thousands) allows this transformation to be effective. Secondly, the dynamics of the reservoir has to exhibit a short-term memory such that the reservoir state is influenced by inputs from the recent past. Additionally, the results of RC computations must be reproducible and robust against noise. Different classes must generate different transient states but, at the same time, similar inputs should not be associated to different classes. Reservoirs typically depend on few parameters that must be adjusted to satisfy the above-mentioned conditions. Experience shows that these requirements are satisfied when the reservoir operates (in the absence of input) in a stable regime, but not too far from a bifurcation point.

A RC setup based on delay-coupled systems [5,6,7] is depicted in Fig. 1 right panel. Within one delay interval of length τ we define N equidistant points separated in time by $\theta=\tau/N$. These N equidistant points are called "virtual nodes". They play a role analogous to the nodes in traditional reservoirs. The values of the delayed variable at each of the N points define the states of the virtual nodes. Under certain inputs, these states characterize, at a given time, the transient response of the reservoir. The separation time θ among virtual nodes plays an important role and can be used to optimise the reservoir performance. It has been shown that a value of $\theta=0.2T$, with T being the characteristic time scale of the nonlinear node, yields best performance [5]. Via this choice, the states of the virtual nodes become dependent on the states of neighbouring nodes and emulate a network serving as reservoir. In the following we show that the single nonlinear node with delayed feedback performs comparably to traditional reservoirs.

.

II. Results

To prove the abilities of the reservoir based on delay-coupled systems, we concentrate in two tasks that are considered benchmark tasks in traditional RC: spoken digit recognition and time series prediction.

The spoken-digit recognition task was introduced by Doddington and Schalk [8]. The input dataset consists of a subset of the NIST TI-46 corpus[1] with ten spoken digits (0...9), each one re-

[1] Texas Instruments-Developed 46-Word Speaker-Dependent Isolated Word Corpus (TI46), Sep-

corded ten times by five different female speakers. We solve this task by using a Mackey-Glass nonlinearity whose dynamics is described by:

$$\dot{X}(t) = -X(t) + \frac{\eta\cdot\left[X(t-\tau)+\gamma\cdot J(t)\right]}{1+\left[X(t-\tau)+\gamma\cdot J(t)\right]^{p}}$$

with *X* denoting the dynamical variable, $\dot{X}$ its derivative with respect to a dimensionless time t, and *t* the delay in the feedback loop. The characteristic time scale of the oscillator, determining the decay of in the absence of the delayed feedback term, has been normalized to T=1. The parameters *h* and *g* represent feedback strength and input scaling, respectively.

The Mackey-Glass system is solved numerically and using electronic circuits [5]. Electronic circuits can be easily implemented and are excellent testbeds to prove the robustness of the results. The virtual node separation is set at θ=0.2 while the total number of virtual nodes is N=400 (the total delay time is then 80 when T is normalized to 1). As an example, we plot in figures 2 the classification performance as a function of the feedback strength ηfor the input scaling factor γ=0.5, which has been chosen such that input and feedback signals are of the same order of magnitude (see ref. [5]). The classification performance is quantified by the word error rate (WER), the percentage of words that were wrongly classified, and the margin (distance) between the reservoir's best guess of the target and the closest competitor. The WER is depicted in Fig. 2 left panel while the margin is depicted in Fig. 2 right panel. By comparing the two figures, it can be seen that an increase in margin is accompanied by a decrease in the WER. It is worth highlighting that there is a broad range of values of ηwhere excellent performance is observed. However, the optimum value for both quantifiers is around η = 0.8. At this optimum value, we obtain a WER of 0.2% in experiments and 0.14% in numerical simulations, which corresponds to only one misclassification in 500 words. These performance levels are comparable to those obtained with traditional RC.

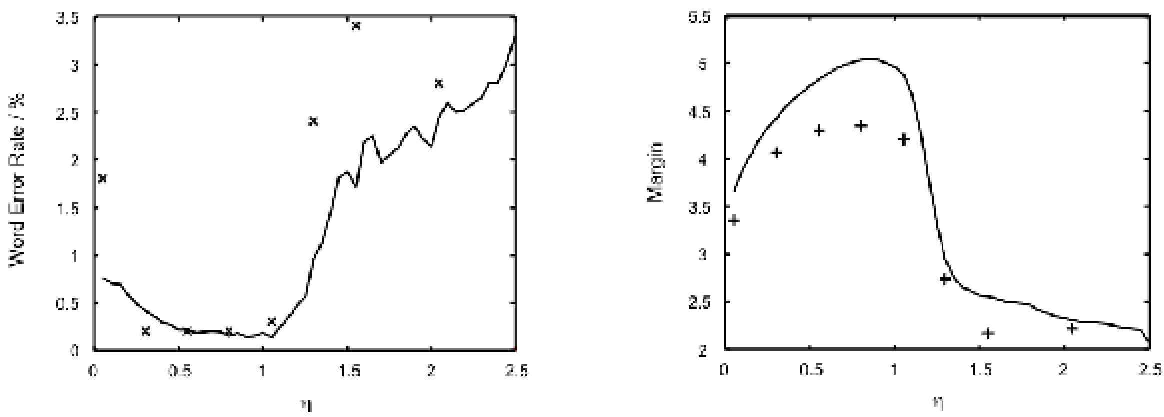

Figure 2. Word error rate (left panel) and Margin (right panel) vs. the feedback strength η.

As a second example we performed a one-time-step prediction task on a time series record-from a far-infrared laser operating in a chaotic state [9]. To evaluate the performance of our RC approach we computed the normalized mean square error (NMSE) between the predicted point and the target. The corresponding results, obtained with the electronic circuit implementation of the Mackey-Glass system, are depicted in Fig. 3. It can be seen that the performance is optimum for values of the feedback strength η between 1 and 1.5, where a NMSE value smaller than 0.1 is obtained.

III. Conclusion

We have shown that the transient states generated in delayed-coupled systems could be used

tember 1991, NIST Speech Disc 7-1.1 (1 disc).

to process information. In particular, we have demonstrated, both numerically and using electronic circuits, that a simple nonlinear element, in this case a Mackey-Glass system, subject to delay feedback is capable of performing spoken digit recognition tasks as well as time series prediction tasks as efficiently as the traditional reservoir computing.

Besides the fundamental aspects, our results also offer practical advantages. The reduction of a large network to a single node facilitates implementations enormously, allowing at the same time, the utilization of inexpensive and easy to handle components.

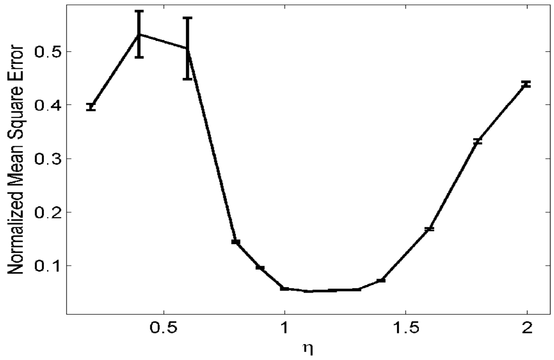

Figure 3. Normalized mean square error vs. the feedback strength η

IV. Acknowledgments

This research was partially supported by the Belgian Science Policy Office, under grant IAP P6-10 ' photonics@be ', by FWO and FRS – FNRS (Belgium), MICINN (Spain) and FEDER under projects FISICOS (FIS2007-60327) and DeCoDicA (TEC2009-14101) and by the European project PHOCUS (EU FET-Open grant: 240763). L.A. and G.VdS. are a PhD Fellow and a Postdoctoral Fellow of the Research Foundation-Flanders (FWO).

V. References

1. Erneux, T., Applied delayed differential equations. *Springer Science + Business Media* (2009).
2. Maass, W., Natschläger, T., and Markram, H., Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comp.* **14**, 2531-2560 (2002).
3. Jaeger, H., Haas, H., Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* **304**, 78–80 (2004).
4. Buonomano, D.V. and Maass, W., State-dependent computations: spatiotemporal processing in cortical network, *Nat. Rev. Neurosc.* **10**, 113 (2009).
5. Appeltant, L., Soriano, M. C., Van der Sande, G., Danckaert, J., Massar, S., Dambre, J. Schrauwen, B., Mirasso, C. R. and Fischer, I, Information processing using a single dynamical node as complex system, Nature Commun. (London) **2**, 1–6 (2011).
6. Larger, L, Soriano M. C., Brunner D., Appeltant L., Gutierrez J. M., Pesquera L., Mirasso, C. R. and Fischer, I., Photonics computing beyond Turing: an optoelectronic implementation of reservoir computing, Optics Express **20**, 3241-3249 (2012).
7. Paquot Y., Duport F., Smerieri A., Dambre J., Schrauwen B., Haelterman M. and Massar, S., Optoelectronic Reservoir Computing, Scientific Reports 2, Article number: 287 (2012).
8. Doddington, G., Schalk, T., Speech Recognition: Turning Theory to Practice. *IEEE Spectrum* **18**, 26-32 (1981).
9. Weigend, A. S. and Gershenfeld, N. A., "Time series prediction: Forecasting the future and understanding the past" ftp://ftp.santafe.edu/pub/Time-Series/Competition (1993).