

Optics Letters

Reservoir computing based on transverse modes in a single optical waveguide

CHARIS MESARITAKIS^{1,*}  AND DIMITRIS SYVRIDIS²

¹University of the Aegean, Department of Information & Communication Systems Engineering, 2 Palama Str., Karlovassi, Samos 83200, Greece

²National & Kapodistrian University of Athens, Department of Informatics & Telecommunications, Panepistimiopolis Ilisia 15784, Greece

*Corresponding author: cmesar@aegean.gr

Received 20 December 2018; accepted 28 January 2019; posted 30 January 2019 (Doc. ID 355898); published 26 February 2019

In this work, a passive reservoir computing scheme is presented, which relies on the excitation and guiding of transverse modes in a large cross-section photonic waveguide. The optical modes act similar to conventional waveguide-based reservoir nodes, whereas random defects act as the operational equivalent of internode connections. Therefore, the proposed scheme's number of nodes and connections can scale with the waveguide's volume, contrary to planar implementations. The principle of operation, alongside the role of mode defects to performance, is validated through a numerical model, whereas the classification of time-dependent, 2D analog/digital image streams is used as an application-relevant benchmark test. © 2019 Optical Society of America

<https://doi.org/10.1364/OL.44.001218>

The concept of reservoir computing (RC) has garnered considerable attention, due to the success of RC schemes to address complex time-dependent machine-learning tasks with increased efficiency [1]. Fields like communications, robotics, economics, and neuroscience have been proliferated by the rise of RC paradigms [2]. A recent, pivotal advance includes the realization of RC architectures with photonic components. Based on virtues like high bandwidth and efficiency, photonic devices have outperformed conventional RCs and have infiltrated demanding multi-GHz applications [3,4]. Photonic RCs can be branched into two main categories: single-node multiplexed systems [3,5–7] and integrated RCs. Typical implementations of the latter include RCs based on optical amplifiers [8], microring resonators [9], waveguides [10], or on photonic crystal cavities [11] and spatial light modulators [12]. Especially [10] exhibits particular interest, due to the simplicity of the RC's nodes that consists of passive waveguides, accompanied by optical couplers, without the inclusion of any nonlinear mechanism. The necessary nonlinearity is triggered through the intrinsic square law of photodiodes (PDs) at the output layer. The richness of RC's dynamics stems from the simultaneous existence of phase and amplitude components in the RC's signals. Despite the simplicity of [10], its size scalability is capped by planar technology limitations and by the losses introduced by optical couplers [10,13]; couplers are

mandatory for signal mixing and for realizing the RC's input/outputs [11].

Here, we present numerical results concerning a photonic implementation that, although it is similar to [10], does not share the aforementioned vulnerabilities. The proposed RC at its core consists of an optical waveguide that allows the random guiding and coupling of a significant number of transverse optical modes. The transverse modes act as an operational equivalent of RC nodes, whereas the random mode-coupling events can provide similar functionality to optical couplers [10]. Compared to [10], the increased number of optical modes that can be invoked by appropriate waveguide design and the wavelength size of the scattering centers can enable the scaling of the number of nodes (modes) and internode connections (defects) with the device's volume. Such an arrangement offers increased number of nodes/connections. On the other hand, the proposed scheme's operation deviates from typical RCs; the utilization of optical modes as nodes enables at the output layer the sampling of their superposition, thus unlocking a phase-to-space mapping computational process that is similar to the scattering-based approach used in [14]. The fundamental difference here is the inclusion of memory effects and thus the ability to process time-dependent signals, whereas the analogy between random mode coupling and RC's inherent randomness is exploited.

The RC consists of an optical waveguide [e.g., polymer optical fiber (POF)] [Fig. 1(a)]. The optical source is a laser whose beam is spatially modulated through the use of a liquid crystal display (LCD), thus providing 2D spatial input for the RC. These motifs feed-excite the waveguide's supported optical modes by controlling their power distribution. The reflectivity, induced by the refractive index difference of the air-POF interfaces create a short fiber cavity [Fig. 1]. A longer external cavity is also assumed, with a delay time larger than the duration of each spatial input, through the use of external beam splitters. After propagation, the guided modes are projected to an imaging device (PD array), where all modes superimpose, forming an optical speckle-like pattern, thus transferring phase information to space. In particular, the guided modes share the same wavelength but exhibit unique propagation constants, and thus accumulate varying phase with propagation, equivalently to random-length planar waveguides. In [10], the signals transmitted through the RC interact through couplers, whereas in our case, the hypothetical

existence of intentional/unintentional scattering centers in the waveguide enables variations of intermodal power distribution (mode coupling). Nonlinear effects are also introduced through the intrinsic square law of the PDs. In terms of RC's fading memory effect [1], the presence of the two cavities provide optical delay, thus contrary to [14] our scheme, is not an extreme learning machine (ELM) but is able to process time-dependent input.

The adopted numerical modeling strategy is depicted in Fig. 1. It comprises a travelling wave tool that treats the POF cavity as a ring cavity and a 2D mode estimator that allows the computation of the spatial profile of each supported transverse mode [15]. In detail, the spatial distribution of random modes (LP_j) is computed, for wavelength $\lambda = 652$ nm and for a 1 mm diameter, 980 μm core POF. We assume that a random number of modes is excited, exhibiting a random intramode power distribution. This distribution is utilized as the equivalent of coupling coefficient in the POF cavity (κ_j). The POF cavity has a total length $d = 12$ cm, while the external cavity is 25 cm long. Propagation losses in the POF are assumed to be $\alpha = 100$ dB/Km, whereas the refractive index of the POF is set to 1.485. Therefore, following a typical ring cavity formalism, the time-dependent spatial distribution of the field entering the cavity [Fig. 1(b)] can be written as using a computational step of $dt = 29$ ps for our calculations,

$$A_1^j(x, y, t) = -j\kappa_j \text{In}_l(x, y, t) + (1 - \kappa_j) E_j(x, y, m, t) e^{\pi \cdot n_{\text{eff}}^j \cdot d / \lambda} \cdot \Phi(x, y) + (1 - R_{\text{ext}}) A_2^j, \quad (1)$$

where In_l is the integral overlap of the initial random modes (LP_j) and the spatial input offered by the LCD [Fig. 1], $E_j(x, y, m, t)$ is the optical field guided in the m th part of the cavity. $\Phi(x, y)$ accounts for the random amplitude and phase impact of the facet's defects on the injected optical field and acts similar to a random input mask for processed signals. The term $\exp(\pi \cdot n_{\text{eff}}^j \cdot d / \lambda)$ accounts for the phase accumulated due to propagation, $(1 - R_{\text{ext}}) A_2^j$ is the optical contribution of the external cavity, and n_{eff}^j is the refractive index for each LP_j mode. Fiber propagation is

$$E_j(x, y, m, t) = E_j(x, y, m - 1, t) \cdot e^{-\frac{\alpha}{2} dz}, \quad (2)$$

neglecting free-carrier-related nonlinear effects due to low injected optical power. The field exiting the cavity is

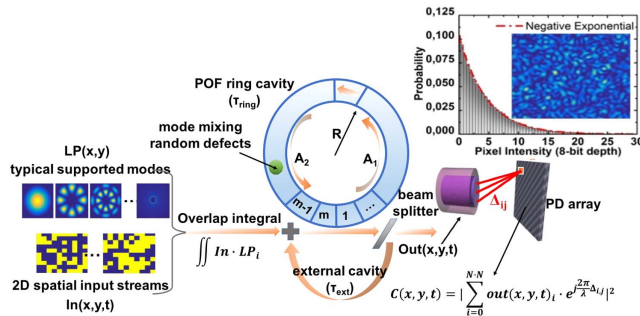


Fig. 1. Schematic of the numerical model including typical radial modes supported by the waveguide and examples of 2D inputs. Inset: probability density function, and typical output assuming 200 \times 200 pixels for visualization purposes at the RC's output.

$$A_2^j(x, y, t) = -j\kappa_j E_j(x, y, m, t) \cdot e^{\pi \cdot n_{\text{eff}}^j \cdot d / \lambda} + (1 - \kappa_j) \text{In}_j(x, y, t), \quad (3)$$

and the output field is given by

$$\text{out}_j(x, y, t) = R_{\text{ext}} \cdot A_2^j. \quad (4)$$

It is worth mentioning that the cavity is not considered resonant, meaning that due to the random coupling coefficient of each transverse mode (κ_j) and the different refractive index (n_{eff}^j), there is no set of global parameters that maximizes cavity enhancement effects. Therefore, the ability of past inputs to influence future results is governed by an exponentially decaying expression, linked to the coupling coefficient. Furthermore, we expanded our travelling-wave model with the inclusion of scattering defects that allow mode coupling [15]. We assumed the existence of (Q) hypothetical defects at random positions ($0 < L_i < d, i = 1 \dots Q$) inside the ring cavity that change the intramode power distribution. Thus, Eq. (2) is modified to

$$E_j(x, y, m, t) = \sum_{i=1}^Q \left(g_i E_i(x, y, m - 1, t) \cdot e^{\frac{j\pi L_i}{\lambda}} \right) \cdot e^{-\frac{\alpha}{2} dz}, \quad (5)$$

where for each defect, the mode-coupling coefficients that govern power exchange follow a normal distribution with the additional condition that

$$\sum_{i=1}^k g_i \leq (1 - \alpha_{\text{def}}), \quad (6)$$

implying scattering losses at each defect. The term α_{def} corresponds to the percentage of losses per defect and depend on the numerical aperture of the structure. The final part of the numerical model includes the projection of the spatial distribution of each mode at the output facet $F(x, y, t) = \sum_j^K \text{out}_j(x, y, t)$ to a square-shaped PD array with $V \times V$ pixels (10×10) and length of 4 mm, positioned at a distance $D_x (= 7$ cm) away from the POF's facet, assuming Fraunhofer diffraction. In particular, the power at each point of the PD array (pixel) is given by

$$P(x, y, t) = \left| \sum_{xi} \sum_{yi} F(xi, yi, t) \cdot e^{-\frac{2j\pi \Delta}{\lambda}} + n(t) \right|^2, \quad (7)$$

where Δ corresponds to the Euclidean distance between each x, y at the POF facet and x_i, y_i pixel at the imaging apparatus,

$$\Delta = \sqrt{(xi - x)^2 + (yi - y)^2 + D_x^2}. \quad (8)$$

Equations (7,8) imply that at each PD element, all modes superimpose, having a phase difference [Eq. (8)] dictated by their Euclidean distance between the facet and the PD array [Eq. (8)], forming an optical speckle with feature size proportional to the distance D_x [15].

Through the term $n(t)$, we introduce noise effects, assuming that it is thermal noise at the PDs. Taking into consideration that each voxel at the POF's output facet acts as a secondary radiation source emitting at a wavelength λ , the projection to the imaging apparatus is a highly complex image [inset of Fig. 1] that, due to the coherent nature of the process, exhibits typical negative exponential distribution [15,16]. From the 10×10 pixels at the imaging device, we exploit the dimensionality increase offered by the RC and use only 10 uniformly

chosen pixels $P(x_c, y_c, t)$, and feed them to a linear regression-based training scheme [17]. The application scenario considered in this work is equivalent to the one used in [9], where a planar RC was used and training involved classifying 1D images and is equivalent to [14]. Here, 2D image streams are assumed, thus elevating the complexity compared to previous works. In detail, each stream comprises five discrete monochrome 2D images of 10×10 pixels [see Fig. 2(a)], each used for 1.38 ns. For example, stream $S_1 \in (I_{x=1 \dots 4, X})$, where I_x are individual images in the stream, and whereas in the next stream, images are perturbed but always the fifth image (I_x) remains constant $S_2 \in (I_{x=2 \dots 5, X})$, etc. The reason for this arrangement is to allow minimum variation among streams and thus assume a stringer task, whereas emphasizing the memory capabilities of our RC. Aiming to generate a training set, we exploit the noise term $n(t)$ in Eq. (7), and we generated input/output instances with different signal-to-noise ratio (SNR), varying from 3 dB to 40 dB. In Fig. 2(a), four typical time traces that are generated by the RC and are recorded by a single PD are presented. It can be seen that although the fifth image is the same for all streams, due to the RC's memory, the time traces $[I_x$ in Fig. 2(a)] differ in power. In particular, the duration of the whole stream was set to 8.46 ns, while the delay induced by the external cavity that acts as the main memory mechanism was set to 3.38 ns (2/5 of the stream's duration) ($\tau_{\text{ext}} \gg dt$). The duration of the stream is linked to the memory that the RC should exhibit and the available bandwidth of the PDs. Therefore, ultrafast streams would demand shorter cavities but will push for higher bandwidth at the PDs. Therefore, the time scale used here is only indicative (1.38 ns per image plus an initial 1.56 ns zero intensity frame in each stream) and can be easily scaled to the Gf/s regime. Alternatively, the optical cavity can be replaced by an electro-optic feedback, linking the LCD and the PDs. In this

configuration, the RC exhibits increased memory at the expense of the processing rate.

Aiming to evaluate the classification capabilities of our scheme, we divided the data set to three subsets, the first consisted of 50% of the time traces and was utilized for training, 20% for neural network validation, while the remaining 30% was used for testing. Furthermore, aiming to eradicate the dependence of system's performance on the initial training set, we repeated the process 100 times so as to extract mean performance. In Fig. 2(b), the classification error is computed versus the minimum SNR of the samples used and assuming that 10 random transverse modes (nodes) are excited in the fiber (black squares). The minimum SNR at the x axis corresponds to the lower SNR added in the training procedure, and thus it can be seen that by incorporating samples with increased noise (SNR decreasing), the mean classification error exponentially increases. The worst case scenario that includes samples with SNR = 3 dB provides a classification error in the order of 9%, whereas for SNR > 10 dB, the error drops to 1.7%. Obviously, if more classes were assumed, it is expected that the SNR requirements for low error rate increase, thus leading to the need for a larger RC similar to [10]. Aiming to investigate the analogy of the number of transverse modes and RC's nodes, we reevaluated the classification error by increasing the number of excited modes. In Fig. 2(b), it can be seen that by increasing the modes from 10 (black square) to 20 (red circles) and finally to 30 (blue triangle), the classification error diminishes to 0.2%.

Towards the same direction, we validated the impact of the defects, which in our case act similar to internode connections. In Fig. 2(c), the classification error versus the number of defects is presented. It can be seen that the absence of any defects diminishes the classification capabilities, offering an error close to 75%. In the case that the number of mode-mixing defects increases beyond 20, implying at least the presence of a defect per 6 mm propagation length, the RC's performance is radically boosted, providing error below 2.5%. Especially in the case of 40 defects (1 defect per 3 mm), the RC provides practically no errors, even by including very noisy samples. The aforementioned performance is extracted by using randomly excited modes from the available mode pool. Aiming to check the impact of the mode's spatial complexity on performance, we assumed the excitation of only 10 modes (10 defects, full range of SNR) and calculated the RC's error and normalized mean square error (nmse) versus the order of the LP modes, meaning that in Fig. 3(a), a range of 10 equals to $LP_{0 \dots 2, 1 \dots 5}$, whereas 20 corresponds to $LP_{10 \dots 12, 10 \dots 15}$, etc. It can be seen that as we employ higher order modes, nmse drops, but at the same time, the standard deviation increases. The same trend is observed when computing the classification error; the high-order modes provide an error below 0.6%.

A critical parameter of our configuration is the distance between the POF and the PDs (D_x), because it regulates the mean speckle size and consequently controls the spatial complexity of the generated response [15]. This leads to the observation that limited D_x offers enhanced spatial complexity [see inset (i) of Fig. 3(b)], while increasing D_x results to [see insets (ii)–(iii) of Fig. 3(b)] speckle features expanding over multiple pixels. Aiming to investigate this parameter, we kept all the parameters constant except the POF-PD distance that ranged from 1 mm to the extreme of 1 m. Figure 3(b) demonstrates a region

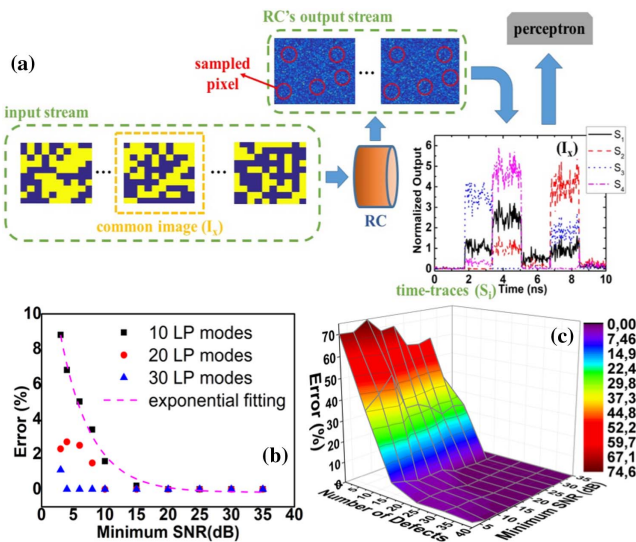


Fig. 2. (a) Typical input streams consisting of 2-bit images, the corresponding speckles (output stream), and the intensity time traces for different streams (S_i) from a single PD at the output. (b) Classification error versus the minimum SNR included in the training set, for different numbers of optical modes. (c) Classification error versus the number of defects and SNR.

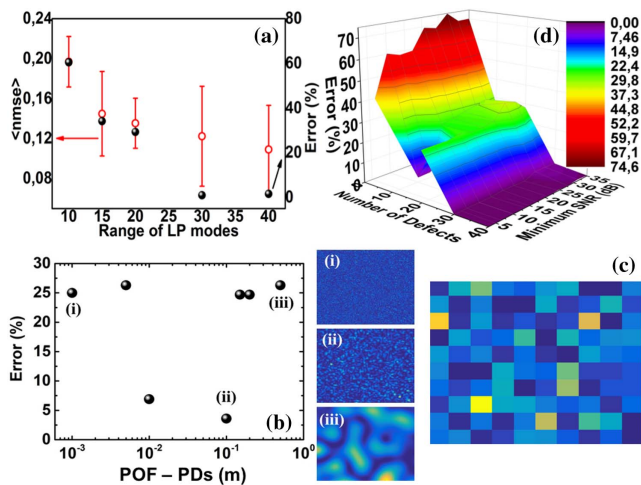


Fig. 3. (a) Normalized mean square error alongside error versus the order of the excited 10 LP modes. (b) Error rate versus the POF-PD distance, whereas the insets demonstrate speckle size for each regime. (c) Typical 8-bit images used at the RC's input. (d) Classification error versus the number of defects and SNR.

extending from 1 cm to 10 cm where classification errors are minimum ($<5\%$), while smaller or larger values increase error ($>20\%$). These results confirm that small distances offer enhanced dimensionality increase (high spatial complexity), but the RC is sensitive to PD's thermal noise, while larger distances, apart from being impractical, wash out spatial features and shatter the dimensionality increase associated with RCs. Finally, aiming to investigate the performance of the proposed RC under more stringent classification tasks, we used images with 8-bit color depth [Fig. 3(c)]. The classification performance was recomputed for the same system parameters used in Fig. 2(c) and are presented in Fig. 3(d). It is clear that the performance trends are preserved, while the system exhibits increased errors due to the fact that streams include more complex features. For example, for 20 defects and $\text{SNR} = 3$ dB, the 2-bit streams exhibit error of 2.3%, while the 8-bit streams allow a significant higher error equal to 26.3% [Fig. 3(d)]. Importantly, this error rate can be reduced to 0.3% by increasing the defects to 40.

The proposed RC relies on the number of optical modes in a single waveguide and in the mode-mixing effect of defects, whereas the benchmark task used is a more elaborate version of [9] and similar to [14] so as to enable comparison. The proposed scheme has the same nodes with previous planar RCs, but against a stringer task, provides similar performance. Furthermore, although only a marginal number of nodes are used, their number can easily scale up to the order of 10^6 , thus enabling to address more complex problems. This is achieved through the fact that the nodes (modes) scale following the number of plane waves [16], whereas internode connections

(defects) can have size similar to the used wavelength [15,16]. In addition, power consumption is marginal, whereas the RC's bandwidth capabilities depend on PD performance, and thus our scheme emerges as a high-rate low-power video recognition modality suitable for digital/analogue data and for medical or security applications. Finally, the proposed RC relies on an intensity/phase-to-space mapping that, through random mode mixing, the input's dimensionality is boosted. Similar to previous ELMs, the output samples include a mixture (superposition) of all modes (nodes) and not pure RC states. Furthermore, the mode-mixing-induced dimensionality increase is high enough so as to allow the spatial undersample of the output, implying that our RC could be also considered equivalent to photonic compressing sensing modalities [18], allowing for reduced PDs at the output layer.

Funding. Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT) (2247 NEBULA).

REFERENCES

1. H. Jaeger and H. Haas, *Science* **304**, 78 (2004).
2. M. Lukoševičius, H. Jaeger, and B. Schrauwen, *KI-Künstliche Intelligenz* **26**, 365 (2012).
3. L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, *Phys. Rev. X* **7**, 011015 (2017).
4. M. C. Soriano, *Physics* **10**, 12 (2017).
5. 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).
6. Y. Kuriki, J. Nakayama, K. Takano, and A. Uchida, *Opt. Express* **26**, 5777 (2018).
7. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, *Nat. Commun.* **4**, 1364 (2013).
8. K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, and J. V. Campenhout, *Opt. Express* **16**, 11182 (2008).
9. C. Mesaritakis, A. Bogris, A. Kapsalis, and D. Syvridis, *Opt. Lett.* **40**, 3416 (2015).
10. K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, *Nat. Commun.* **5**, 3541 (2014).
11. F. Laporte, A. Katumba, J. Dambre, and P. Bienstman, *Opt. Express* **26**, 7955 (2018).
12. J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, *Optica* **5**, 756 (2018).
13. A. Katumba, J. Heyvaert, B. Schneider, S. Uvin, J. Dambre, and P. Bienstman, *Sci. Rep.* **8**, 2653 (2018).
14. A. Lugnan, J. Dambre, and P. Bienstman, *Opt. Express* **25**, 30526 (2017).
15. C. Mesaritakis, M. Akriotou, A. Kapsalis, E. Grivas, C. Chaintoutis, T. Nikas, and D. Syvridis, *Sci. Rep.* **8**, 1 (2018).
16. A. P. Mosk, A. Lagendijk, G. Leroose, and M. Fink, *Nat. Photonics* **6**, 283 (2012).
17. M. Lukoševičius and H. Jaeger, *Comput. Sci. Rev.* **3**, 127 (2009).
18. Y. Chen, X. Yu, H. Chi, X. Jin, X. Zhang, S. Zheng, and M. Galili, *Opt. Lett.* **39**, 2222 (2014).