EFFICIENT CLASSIFICATION OF CHAOTIC SIGNALS WITH APPLICATION TO SECURE COMMUNICATIONS

Francesco Gianfelici, Claudio Turchetti, and Paolo Crippa

DEIT - Università Politecnica delle Marche, Via Brecce Bianche 12, I-60131 Ancona, Italy

ABSTRACT

This paper presents an exhaustive study on the classification capabilities of an efficient algorithm, which is able to accurately classify non-deterministic signals generated by chaotic dynamical systems, without estimating their probability density function (pdf). Experimental results were compared to other existing techniques such as Hidden Markov Model (HMM), Vector Quantization (VQ), and Dynamic Time Warping (DTW). Classification performance is higher than current best practices for chaotic signals. A better noise rejection was also achieved, and a reduction of two orders of magnitude in training-times compared with HMM was obtained, thus making the proposed methodology one of the current best practices in this field. As an application example, the recognition of encrypted chaotic-signals in a secure-communication context, is reported and discussed.

Index Terms— Signal classification, Karhunen-Loève Transform, cryptography, chaotic signals, non-probabilistic algorithm.

1. INTRODUCTION

During recent decades, many classifiers based on several probabilistic techniques such as, i.e. Hidden-Markov Model (HMM) [1], Vector-Quantization (VQ), Dynamic Time Warping (DTW) have been developed; a review of statistical recognition can be found in [2]. However, it is possible to classify the state-of-the-art according to the intrinsic limitations of the abovementioned practices: *(i)* the high computational complexity of probability density function (pdf) estimations *(ii)* the low-recognition performance in unsupervised cases *(iii)* the large number of constraints on signal features, and/or the assumptions on system properties and *(iv)* the high elaboration times for training phases.

In this work an exhaustive study of the classification capabilities of an our efficient algorithm[1] [4] is proposed. This algorithm analyzes the proximity-measures between the trajectories and the projections of a signal which has to be recognized, over all the eigenfunctions calculated in the training phase. In fact, the decision technology on which the recognition procedure is strongly based, is achieved using a non-probabilistic methodology that takes into account both the principal and the minimal components. Moreover, the recognition of chaotic signals is achieved without the probability density function (pdf) estimation.

The analyzed stochastic processes (SPs) are dynamical systems based on non-linear maps that generate non-deterministic signals from several initial conditions and random parameters. Several systems such as oscillating integrated circuits affected by random device variations or secure communication systems using secret keys, can be represented by these SPs. Exhaustive experimentation showed high recognition performance with a limited number of signals used in the training phase. Several comparisons with HMM, VQ, and DTW showed recognition performance that is higher for chaotic signals. Moreover, a better noise rejection for several Signal Noise Ratios (SNR), and a reduction of two orders of magnitude on HMM training-times have been achieved. As a case study, the recognition of encrypted chaotic-signals applied to secure communications [5] was considered.

[1] A patent application of the proposed algorithm has been deposited [3] by F. Gianfelici and C. Turchetti according to the copyright laws of the Italian Government.

2. THE KLT-BASED ALGORITHM

Let us briefly recall the KLT-based algorithm that was recently presented in [4]. Given two distinct SPs, $\mathbf{x}$ and $\mathbf{y}$ with autocorrelation matrices R_{xx} and R_{yy} respectively, the two bases $\{u_1,\ldots,u_N\}$ and $\{w_1,\ldots,w_N\}$ are defined in terms of the eigenvectors of the corresponding eigenproblems[2], namely $R_{xx}U=\Lambda U$, $R_{yy}W=\Sigma W$, where $U=[u_1,\ldots,u_N]$ and $W=[w_1,\ldots,w_N]$, with $U,W\in\mathbb{R}^{L\times N}$, and Λ, Σ are diagonal matrices containing the corresponding eigenvalues. By projecting all the realizations $x^{(i)}$ and $y^{(i)}\in\mathbb{R}^L$, $i=1,\ldots,N$ onto the bases U and W, the vectors of KLT coefficients $a^{(i)}=U^Tx^{(i)}$, $b^{(i)}=W^Ty^{(i)}$ $i=1,\ldots,N$, and the matrices $A=[a^{(1)}\ a^{(2)}\ \ldots\ a^{(N)}]$ and $B=[b^{(1)}\ b^{(2)}\ \ldots\ b^{(N)}]$, with A and $B\in\mathbb{R}^{N\times N}$, are defined. Similarly, by deriving the cross-projections, i.e. the projections of each SP onto the base of the other, we have $e^{(i)}=U^Ty^{(i)}$ and $f^{(i)}=W^Tx^{(i)}$. Thus we can define the two matrices: $E=[e^{(1)}\ e^{(2)}\ \ldots\ e^{(N)}]$, and $F=[f^{(1)}\ f^{(2)}\ \ldots\ f^{(N)}]$ with E, $F\in\mathbb{R}^{N\times N}$. By defining $X=[x^{(1)},\ldots,x^{(N)}]$ and $Y=[y^{(1)},\ldots,y^{(N)}]$ as the matrices of realizations so that $X,Y\in\mathbb{R}^{L\times N}$, the projections and the cross-projections can be rewritten in compact form as a collection of $N\times N$ parameters, which can be represented in matrix form as:

$$P_\Phi=\begin{bmatrix}P_U\\P_W\end{bmatrix}=\begin{bmatrix}A & E\\F & B\end{bmatrix} \tag{1}$$

where $P_\Phi\in\mathbb{R}^{2N\times 2N}$ is the non-symmetric matrix of extracted features. Letting ζ be the realization that has to be recognized, the first step is the definition of two vectors $l=U^T\zeta$ and $m=W^T\zeta$, where l and $m\in\mathbb{R}^N$. As a second step let us define a transformation $\mathcal{T}:\mathbb{R}^{N\times K}\rightarrow\mathbb{R}^{N\times 2K}$ acting on the columns of an $N\times K$ matrix

[2] The KLT is calculated by means of the Fast-KLT Algorithm that can be found in [6].

Table 1. Recognition with different training set dimensions ($M = 2, TS = 100$)

N	$\mathbf{x}$ Rec.	$\mathbf{y}$ Rec.	Rec. Perf.
5	90%	86%	88%
10	54%	100%	77%
20	100%	98%	99%
30	100%	96%	98%
40	100%	100%	100%
50	100%	100%	100%

as:

$$\mathcal{T}v = \begin{bmatrix} v_1 & v_1 \\ v_1 & v_2 \\ \vdots & \vdots \\ v_1 & v_N \end{bmatrix} = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(N)} \end{bmatrix}, \quad (2)$$

where $v = [v_1 \;\; v_2 \;\; \cdots \;\; v_N]^T$ is a generic column vector and $v^{(1)}, \ldots, v^{(N)}$ are elements of $\mathbb{R}^2$. Applying $\mathcal{T}$ to l, m, P_U, and P_W we obtain: $\mathcal{T}l, \mathcal{T}m \in \mathbb{R}^{N\times 2}$, and $\mathcal{T}P_U, \mathcal{T}P_W \in \mathbb{R}^{N\times 4N}$.

Thus we compute the matrices $D \in \mathbb{R}^{N\times 2N}$ and $H \in \mathbb{R}^{N\times 2N}$ whose generic ik-th elements are $[D]_{ik} = dist\left(l^{(i)}, P_{Ui}^{(k)}\right)$ and $[H]_{ik} = dist\left(m^{(i)}, P_{Wi}^{(k)}\right)$ with $i = 1, \ldots, N$ and $k = 1, \ldots, 2N$, where $dist$ is the Euclidean-distance between vector pairs. By defining another transformation $\mathcal{S} : \mathbb{R}^{N\times 2N} \to \mathbb{R}^{N\times 2N}$ such that, when applied to a matrix Q, results in a novel matrix $\widetilde{Q} = \mathcal{S}Q$, with same dimensions, whose elements are:

$$[\widetilde{Q}]_{ik} = \begin{cases} 1, & [Q]_{ik} = \min_l [Q]_{il} \\ 0, & \text{elsewhere}. \end{cases} \quad (3)$$

In such a way the minimum distance in the rows of matrices D and H is determined. Therefore we can define two vectors: $c = [c^{(1)}, c^{(2)}, \ldots, c^{(2N)}]$, and $p = [p^{(1)}, p^{(2)}, \ldots, p^{(2N)}]$ whose elements are: $c^{(k)} = \sum_{i=1}^{N} [\widetilde{D}]_{ik}$ and $p^{(k)} = \sum_{i=1}^{N} [\widetilde{H}]_{ik}$, where $\widetilde{D} = \mathcal{S}D$ and $\widetilde{H} = \mathcal{S}H$ and $c, p \in \mathbb{R}^{2N}$. The terms $c^{(k)}$, $p^{(k)}$ can be rewritten as elements of a novel matrix $\Pi \in \mathbb{R}^{2\times 2}$ as:

$$\Pi = \begin{bmatrix} \sum_{k=1}^{N} c^{(k)} & \sum_{k=N+1}^{2N} c^{(k)} \\ \sum_{k=1}^{N} p^{(k)} & \sum_{k=N+1}^{2N} p^{(k)} \end{bmatrix} \quad (4)$$

that can be summed by columns, thus obtaining following numbers: $\mu_x = [\Pi]_{11} + [\Pi]_{21}$ and $\mu_y = [\Pi]_{12} + [\Pi]_{22}$, which represent the *likelihood-scores* of ζ respect to $\mathbf{x}$, and $\mathbf{y}$. Finally the recognition of ζ is performed as follows: $\zeta \in \mathbf{x}$ if $\mu_x = \max[\mu_x, \mu_y]$ or $\zeta \in \mathbf{y}$ if $\mu_y = \max[\mu_x, \mu_y]$.

3. EXPERIMENTAL RESULTS ON CHAOTIC SIGNALS

In order to characterize the recognition performance and the noise rejection of the recognizer, chaotic SPs that are characterized by border collision bifurcations (C-bifurcations) and/or eigenvalue bifurcations were taken into account. Methodologically, the experiments were organized, according to the definition of SPs, as follows: *(i)* given a chaotic map, the realizations of each SP had random initial conditions defined in specific closed intervals which had no intersection, and *(ii)* each SP was generated by one different chaotic-map. Throughout this paper case *(i)* will be indicated as *single-map based SPs* and case *(ii)* as *multi-map based SPs*. In order to characterize the *single-map based SPs*, let us to define the piecewise linear map:

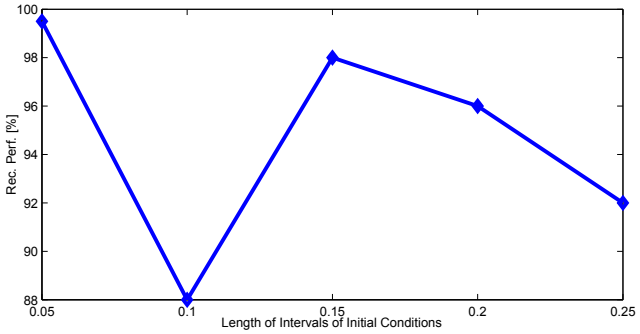

Fig. 1. Performance as a function of distance between intervals of random initial conditions.

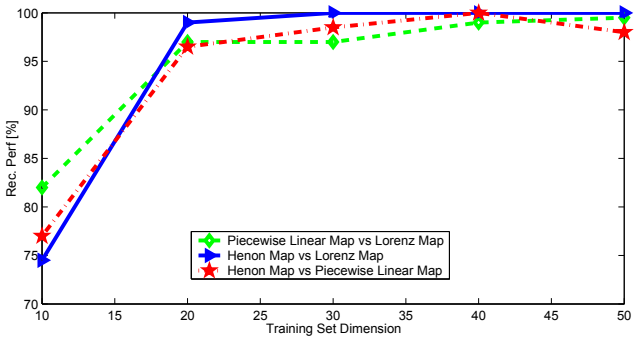

Fig. 2. Performance as a function of training dimensions.

$$f : x \to f(x) = \begin{cases} \alpha_1 x + \beta_1, & \text{if } 0 \le x < \gamma \\ \alpha_2 x + \beta_2, & \text{if } \gamma \le x < \delta \\ \alpha_3 x + \beta_1, & \text{if } \delta \le x \end{cases} \quad (5)$$

where $\alpha_1 = 1.4$, $\alpha_2 = -2$, $\alpha_3 = -0.8667$, $\beta_1 = 0.1$, $\beta_2 = 3.5$, $\gamma = 1$, and $\delta = 3$. Two SPs, $\mathbf{x}$ and $\mathbf{y}$, whose realizations have random initial conditions in [0, 0.01], and [0.5, 0.51], respectively, were considered. For this couple of SPs recognition performance with different training set dimensions varying from 5 to 50 is proposed in Tab. 1. Figure 1 gives the recognition performance as a function of the distance between intervals of random initial conditions. In order to characterize the *multi-map based SPs*, several chaotic maps were considered: *a)* the piecewise linear map described in *single-map based SPs*, *b)* the well-known Lorenz map (logistic map), and *c)* the Henon map. Their recognition performance as a function of different dimensions of the training set is indicated in Fig. 2. It is worth noting that for $N \ge 20$ the Rec. Perf.$> 95\%$, thus achieving high reliability of recognition of *chaotic signals*. Several comparisons with current best practices such as HMM, DTW, and VQ were considered. The results of the comparisons with HMM are shown in Fig. 5, where it is possible to note the recognition performance as a function of HMM states, its mean recognition value (90.9%), and the performance of our approach (99.3%). This comparison has $M = 2$ and $N = 50$, where $\mathbf{x}$ and $\mathbf{y}$ are generated by the normalized piecewise linear map with random initial conditions in [0, 0.01] and [0.5,

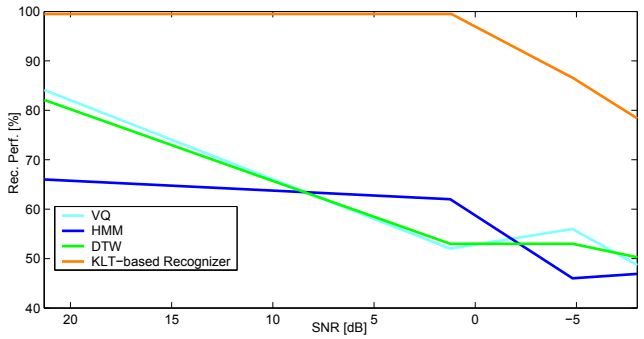

Fig. 3. Noise rejection: comparison of recognition performance.

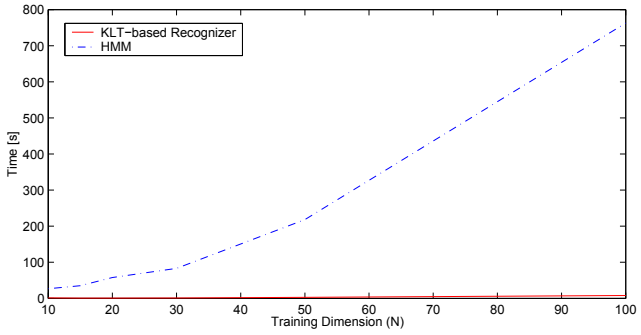

Fig. 4. Comparison between training times.

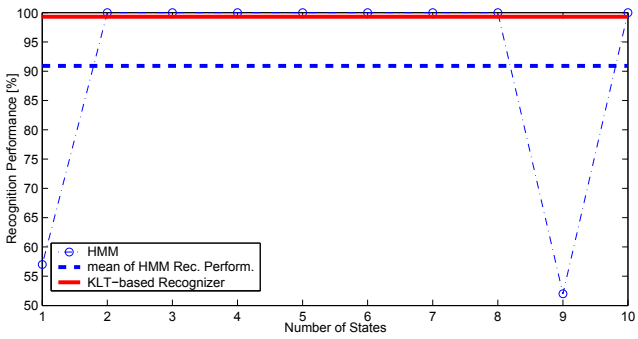

Fig. 5. Comparison between HMM and KLT-based recognizer.

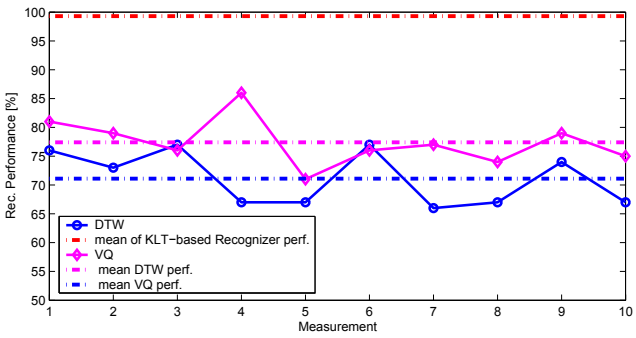

Fig. 6. Comparison between DTW, VQ, and KLT-based recognizer.

0.51], respectively. The comparison with DTW and VQ is shown in Fig. 6. In order to characterize noise rejection, the role of superposed noise was investigated. The results for $M = 2$ and $N = 50$, are shown in Fig. 3 where $\mathbf{x}$ and $\mathbf{y}$ are a normalized piecewise linear map with random initial conditions [0, 0.01] and [0.5, 0.51], and HMM is modelled with 10 states and 2 mixtures. It is worth noting that for chaotic signals with superposed noise, the KLT-based recognizer has a better noise-rejection than can be obtained using other techniques. In Fig. 4 the training time of the KLT-based recognizer and HMM respectively is shown for several training dimensions. It is worth noting that the KLT-based recognizer is able to perform this task with a reduction of two orders of magnitude compared with HMM-times.

4. APPLICATION: SECURE COMMUNICATION SYSTEMS

Secure communication systems, able to prevent abusive or illegal interceptions, are becoming more and more important in defence and in civil communications. In this context, the pioneering idea of Tang [7] on chaos synchronization applied to coherent demodulation, has stimulated the development of many scientific contributions during the last two decades. A recent paper [5], has reported a large-scale distributed application based on chaos synchronization. According to the classification proposed in [8], cryptosystems can be divided into: *(i)* chaotic switching, and *(ii)* chaotic modulation. A complete review of chaos-based technology applied to secure communications is in [9]. In this Section, an application to chaotic communications of the recognizer based on chaotic switching is proposed. The complete schema of the communication system being analyzed is shown in Fig. 7. In agreement with the well-known secure-communication theory, the secret key is the chaotic-map definition plus the random initial conditions of zeros and ones, the transmitted chaotic signal is the ciphered sequence of bits, and the receiver is represented by the recognition algorithm, whose structure can also be public, e.g. the RSA algorithm. Indeed the knowledge of the secret key allows us to train the algorithm and to effectively recognize the ciphered signals with the aim of reconstructing the original bit-sequences. Moreover, the knowledge of the recognition algorithm alone, without additional information on the secret key does not allow the calculus of training data, and therefore the message recognition. Figure 8 shows *(a)* the ciphered signal of the bit sequence 0101, where zeros and ones are encoded by a piecewise linear map described in eq. (5), with random initial conditions in [0, 0.01] and [0.5, 0.51], and *(b)* their spectra. In

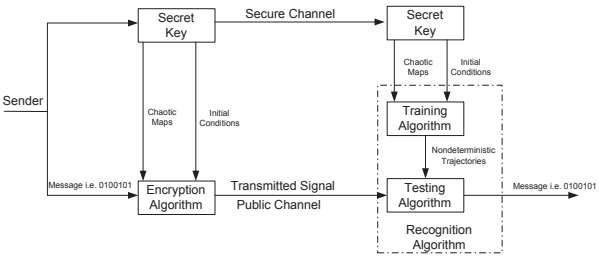

Fig. 7. Complete schema of the secure communication system.

Table 2. Different measurement of bit recognition ($M = 2$, transmitted bits= 100)

Par.	1st Meas.			2nd Meas.			3rd Meas.			4th Meas.			5th Meas.			Rec. Stat.	
N	0's	1's	Rec.	0's	1's	Rec.	0's	1's	Rec.	0's	1's	Rec.	0's	1's	Rec.	Mean	σ
10	49	45	94%	30	43	73%	47	49	96%	35	50	85%	49	49	98%	89.2%	10.33
20	49	47	96%	48	50	98%	50	50	100%	49	50	99%	49	49	98%	98.2%	1.48
30	49	49	98%	50	50	100%	50	50	100%	49	50	99%	50	50	100%	99.4%	0.89
40	50	50	100%	50	50	100%	50	50	100%	50	50	100%	50	50	100%	100.0%	0.00
50	50	50	100%	50	50	100%	50	50	100%	50	49	99%	50	49	99%	99.6%	0.55

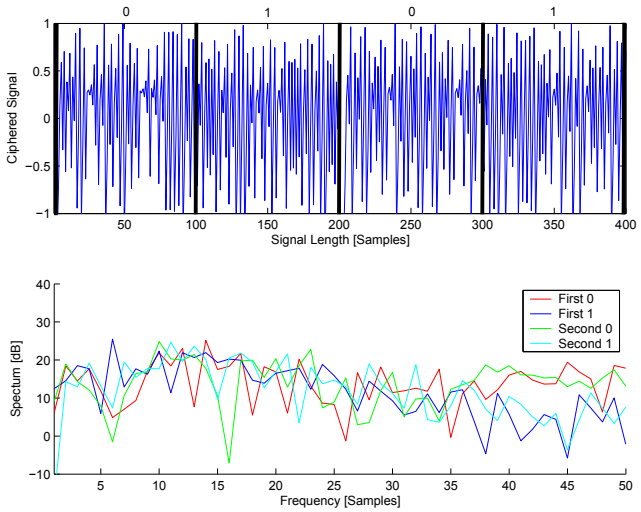

Fig. 8. Ciphered signal of the bit-sequence 0101 and their spectra.

order to validate the reliability of the receiver, several measurements are proposed in Tab. 2, which clearly shows a reduction in standard deviation σ with an increase in performance. Moreover, several malicious attacks on recognition procedure (assuming that the recognition algorithm is public, e.g. RSA) were considered. Figure 9 gives the recognition performance as a function of the exactness of the estimated training-set $N = 50$, conjecturing that the opponent is able to extract from the transmitted signal, the abovementioned percentage of realizations (with exact length) to train the recognizer. Where the exactness is represented as a percentage of realizations that are correctly generated by a specific SP, the inexact signals are generated by another SP, and the Testing Set has an exact time-support L.

5. CONCLUSIONS

In this paper, a novel KLT-based methodology for the efficient recognition of chaotic signals has been proposed. The results show a recognition rate which is close to 100%, thus demonstrating the validity of the proposed algorithm. A comparative evaluation with state-of-the-art techniques demonstrated the superiority of the proposed approach and a reduction of two orders of magnitude in training times.

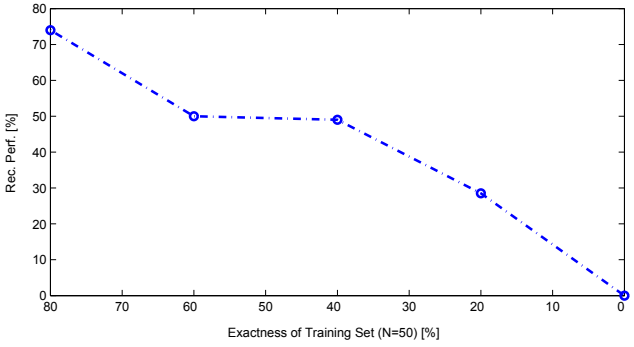

Fig. 9. Malicious attack.

6. REFERENCES

[1] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[2] A. K. Jain, P. R. W. Duin, and M. Jianchang, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[3] F. Gianfelici and C. Turchetti, *A Stochastic Process Recognizer*. Ufficio Brevetti - Ministero Attività Produttive, Nov. 2004, Italian Patent, Dep. Num. AN2004A000050.

[4] F. Gianfelici, C. Turchetti, and P. Crippa, "A non probabilistic algorithm based on Karhunen-Loève transform for the recognition of stochastic signals," in *IEEE Proc. Int. Symp. Signal Process. Inform. Technol. (ISSPIT 2006)*, Aug. 2006, pp. 385–390.

[5] A. Argyris and *alii*, "Chaos-based communications at high bit rates using commercial fibre-optics links," *Nature*, vol. 438, pp. 343–346, Nov. 2005.

[6] F. Gianfelici, G. Biagetti, P. Crippa, and C. Turchetti, "A novel KLT algorithm optimized for small signal sets," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2005)*, vol. 1, Mar. 2005, pp. 405–408.

[7] Y. Tang, A. Mees, and L. O. Chua, "Synchronization and chaos," *IEEE Trans. Circuits Systems*, vol. 30, pp. 620–626, Sept. 1983.

[8] N. Masuda and K. Aihara, "Cryptosystems with discretized chaotic maps," *IEEE Trans. Circuits Systems—I*, vol. 49, pp. 28–40, Jan. 2002.

[9] F. C. M. Lau and C. K. Tse, *Chaos-Based Digital Communication Systems*. Berlin: Springer-Verlag, 2003.