

NUEVA METODOLOGÍA PARA LA DIFERENCIACIÓN DE SEÑALES DISCRETAS UNIDIMENSIONALES MEDIANTE EL EMPLEO DE CARACTERÍSTICAS GEOMÉTRICAS

NEW METHODOLOGY FOR THE DIFFERENTIATION OF ONE-DIMENSIONAL DISCRETE SIGNALS THROUGH THE USE OF GEOMETRIC FEATURES

M. García Blesa^a, J. Vorobioff^b y W. Legnani^a

^a*Centro de Procesamiento de Señales e Imágenes, Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Medrano 951 (C1179AAQ) CABA, Argentina, hgarciablesa@frba.utn.edu.ar, <https://www.frba.utn.edu.ar/>*

^b*Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, <https://www.frba.utn.edu.ar/>*

Palabras clave: amplitud, ángulo de zenit, factor de forma, diferenciación de señales

Resumen. El presente estudio se centra en la aplicación de un algoritmo para la clasificación de sistemas dinámicos caóticos, incluyendo los sistemas de Henón, Chirikov, Schuster, Logístico y el modelo poblacional de Ricker. El algoritmo utiliza como mecanismo principal la técnica de barrido, para la cual se define una dimensión de *embedding* y un parámetro de discretización. La clasificación de los sistemas dinámicos caóticos se lleva a cabo mediante la extracción de características geométricas. La metodología utiliza tres elementos: la amplitud de la señal, y dos nuevos conceptos: ángulo de Zenit y un factor de forma, generando ternas y reduciendo aquellas que se repiten durante el proceso. A partir de estos datos se crea un espacio de características donde cada vector es único. Además de la clasificación, se lleva a cabo un análisis detallado del espacio generado por el algoritmo, el cual proporciona información adicional que complementa la clasificación realizada.

Keywords: amplitud, ángulo de zenit, factor de forma, diferenciación de señales

Abstract. This study focuses on the application of an algorithm for the classification of chaotic dynamical systems, including the Henón, Chirikov, Schuster, Logistic, and Ricker population model systems. The algorithm employs the sweep technique as its main mechanism, for which an embedding dimension and a discretization parameter are defined. Classification of chaotic dynamical systems is carried out by extracting geometric features. The methodology utilizes three elements: signal amplitude, Zenith angle, and a shape factor, generating triplets and reducing those that repeat during the process. From these data, a feature space is created where each vector is unique. In addition to classification, a detailed analysis of the space generated by the algorithm is performed, which provides additional information complementing the classification performed.

1. INTRODUCCIÓN

En los últimos años, los patrones ordinales han surgido como una herramienta ampliamente empleada en las técnicas de estudio de señales. Se han destacado por su eficacia en el análisis de señales discretas (Keller et al. (2017)). Su utilidad radica en la capacidad de detectar características, comportamientos y otros aspectos esenciales de las fuentes emisoras, como la identificación de fenómenos, la detección de anomalías o el análisis a largo plazo, lo cual es fundamental para diferenciar señales (Tangirala (2018)). En la construcción del modelo propuesto en este estudio, se toman algunas herramientas de la técnica de patrones ordinales, utilizando principalmente la técnica de barrido mediante la definición de dos simples parámetros: una dimensión de *embedding* y un factor de discretización. Junto a ello se tomaron en cuenta algunos conceptos básicos como la amplitud, y dos nuevas definiciones: una medida angular y un factor de forma que expresa un conjunto de datos en términos de variaciones de picos y valles. La hipótesis planteada sugiere que estas tres características son suficientes para lograr una descripción exhaustiva de la señal, algo similar a una huella digital. En este trabajo se analizarán sistemas dinámicos caóticos de determinadas señales ampliamente difundidos en la literatura como son los sistemas dinámicos de Henón, Chirikov, Schuster, Logístico y el modelo poblacional de Ricker con el objetivo de poner a prueba el modelo en sus capacidades para la clasificación. Para más antecedentes sobre la clasificación de este tipo de señales, se pueden consultar trabajos como el Keller et al. (2017), Olivares et al. (2019), Cencini et al. (2009) y Argyris et al. (1994). La importancia del análisis de estos sistemas radica en la posibilidad de obtener una diferenciación en el dominio del tiempo, ya que los sistemas caóticos no pueden analizarse óptimamente en el dominio de la frecuencia utilizando técnicas altamente difundidas como transformadas de Fourier y Onditas (Pinsky (2023), Strang (1993)).

2. MATERIAL Y MÉTODOS

Consideramos el espacio de características de una señal como el conjunto formado por vectores únicos, cuyas componentes definimos a continuación para $S = \{a_1, a_2, \dots, a_n\}$ una señal discreta unidimensional. Dada una dimensión de *embedding* d , se definen los segmentos de longitud d : $S_i = \{a_i, a_{i+1}, \dots, a_{i+d-1}\}$ para $i = 1, 2, \dots, n - d + 1$. Por otra parte se propone la definición de tres características: 1- La amplitud. 2- El ángulo de Zenit, que mide la distribución de los valores de la señal. 3- El parámetro de forma, que permite caracterizar la secuencia de picos y valles de la señal. A continuación detallamos las tres características.

2.1. Amplitud

Denotando como $\text{mín}(S)$ y $\text{máx}(S)$ a los valores mínimo y máximo de la señal S , respectivamente, calculamos la amplitud de S : $A(S) = \text{máx}(S) - \text{mín}(S)$.

2.2. Ángulo de Zenit

Definimos el ángulo de Zenit utilizando el valor medio p_S de los valores de S con el objetivo de determinar su posición geométrica dentro del rango acotado entre $\text{mín}(S)$ y $\text{máx}(S)$. Consideremos el siguiente intervalo en el espacio de dos dimensiones: $I_S = [0, 1] \times [\text{mín}(S), \text{máx}(S)]$. Definimos entonces el ángulo φ determinado por los segmentos \overline{PM} y \overline{PN} , donde $P = (0,5, p_S)$, $M = (1, \text{máx}(S))$ y $N = (1, p_S)$. Con el fin de comparar valores de φ de diferentes señales, normalizamos el intervalo I_S por $I = [0, 1] \times [0, 1]$ y el promedio p_S por $p = [p_S - \text{mín}(S)] / [\text{máx}(S) - \text{mín}(S)]$. Finalmente, calculamos $\varphi(S) = \arctan(2 - 2p)$. Debido a que el promedio no puede ser menor que el mínimo ni mayor que el máximo, sus

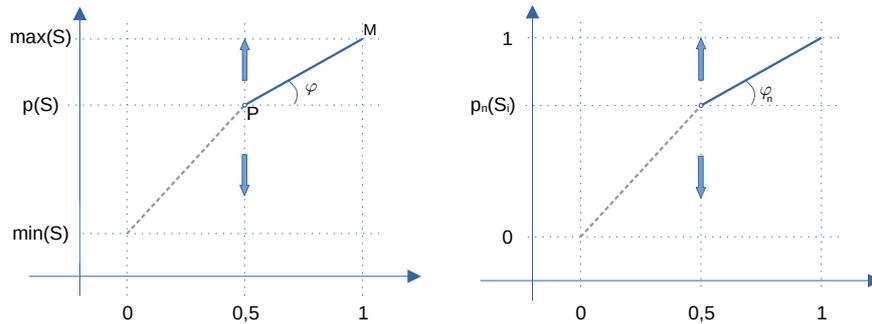


Figura 1: Representación gráfica del ángulo de Zenit original (izquierda) y normalizado (derecha).

valores oscilan entre 0 y 1, lo cual determina que el ángulo φ pertenece al intervalo $[0, 1,1071)$. La Figura 1 ilustra esta definición de forma genérica.

2.3. Factor de forma

Una señal de un fenómeno real puede tener una cantidad variable de valles y picos, en una determinada secuencia, que contribuya a su caracterización. Para formular la siguiente característica, proponemos una manera de obtener una medición de la forma de la señal teniendo como objetivo resumirla en un solo valor, a partir del siguiente algoritmo:

1. Consideramos la señal $S = \{a_1, a_2, \dots, a_n\}$.
2. Ordenamos los datos en orden creciente, obteniendo la secuencia $B = \{b_1, b_2, \dots, b_n\} \subseteq S$, donde $b_i \leq b_{i+1}$ para $i = 1, 2, \dots, n - 1$.
3. Eliminamos los elementos repetidos en B obteniendo la secuencia $C = \{c_1, c_2, \dots, c_m\} \subseteq S$, donde $m \leq n$ y $c_i < c_{i+1}$ para $i = 1, 2, \dots, m - 1$.
4. Definimos el vector $O = (o_1, o_2, \dots, o_n)$ como $o_i = j$ si $a_i = c_j$, para $i = 1, 2, \dots, n$.

Al generar el vector de números de orden O , es importante tener en cuenta que si hay valores repetidos en la secuencia que se está analizando, sus respectivos números de orden también se repetirán. Esta particularidad implica que las componentes del vector de orden están acotadas entre 1 y la longitud de la secuencia original. El vector O representa la forma de la secuencia en términos de picos y valles mediante números enteros positivos. Con el fin de cuantificar y resumir este vector, construiremos un número utilizando las componentes de O como exponentes de una base de números primos. Explícitamente, si $P_n = \{q_1, q_2, \dots, q_n\}$ es el conjunto de los primeros n números primos, definimos el valor que representa unívocamente el parámetro de forma como $\mathcal{F}(S) = \prod_{i=1}^n q_i^{o_i}$. Por ejemplo, para la señal $S = \{27, 44, 39, 71, 39\}$ y una dimensión $d = 4$, tenemos dos segmentos: $S_1 = \{27, 44, 39, 71\}$ y $S_2 = \{44, 39, 71, 39\}$, cuyos factores de forma son $\mathcal{F}_1 = 3241350$ y $\mathcal{F}_2 = 10500$. En la tabla 1 se muestra un ejemplo numérico básico de forma secuencial.

Debido al Teorema Fundamental de la Aritmética, la función que define al parámetro de forma resulta ser una biyección, lo que implica que se puede recuperar la estructura de valles y picos de la secuencia.

A partir de las características definidas anteriormente, para una señal S de longitud n y una dimensión de *embedding* d , podemos asociar a cada segmento S_i una terna formada por

Tabla 1: Ejemplo para el cálculo del factor de forma.

Acción	Resultado para S_1	Resultado para S_2
Señal original	$\{27, 44, 39, 71\}$	$\{44, 39, 71, 39\}$
Ordenamiento B	$\{27, 39, 44, 71\}$	$\{39, 39, 44, 71\}$
Eliminación repeticiones C	$\{27, 39, 44, 71\}$	$\{39, 44, 71\}$
Vector O	$(1, 3, 2, 4)$	$(2, 1, 3, 1)$
Primeros 4 primos	$\{2, 3, 5, 7\}$	$\{2, 3, 5, 7\}$
Factor de forma \mathcal{F}	$2^1 3^3 5^2 7^4$	$2^2 3^1 5^3 7^1$

$(A(S_i), \varphi(S_i), \mathcal{F}(S_i))$. Para simplificar la notación, indicamos $A_i = A(S_i)$, $\varphi_i = \varphi(S_i)$ y $\mathcal{F}_i = \mathcal{F}(S_i)$ y definimos el vector de características $v_i = (A_i, \varphi_i, \mathcal{F}_i)$. El conjunto de todos estos vectores es $V = \{v_i : i = 1, 2, \dots, n - d + 1\}$. Las primeras dos componentes del vector v_i –amplitud y ángulo de Zenit– son valores reales, por lo tanto la probabilidad de su aparición es muy baja. Para aumentar dicha probabilidad discretizamos ambos parámetros, asignando los valores calculados en un rango definido por un valor $k \in \mathbb{N}$ arbitrario.

Con el objetivo de reducir la información redundante, eliminamos los elementos repetidos del conjunto V . Así, consideramos el conjunto $V' = \{v_1, v_2, \dots, v_k\} \subseteq V$ está formado por elementos no repetidos.

Este conjunto, formado por vectores únicos, representa el espacio de características de S . Observamos que se satisface $\#V' \leq \#V$.

Consideramos el cardinal del conjunto $\#V'$ como un valor que contabiliza los vectores únicos de la señal que la caracterizan. Sintetizamos pues el proceso diciendo que por cada señal S se tiene un cardinal $\#V'$.

Luego, para $\mathcal{S} = \{S_1, S_2, \dots, S_t\}$ un conjunto de señales, se tiene un cardinal $\#V'$ para cada señal de \mathcal{S} . Definimos el conjunto de cardinales $\Omega = \{\#V'_1, \#V'_2, \dots, \#V'_t\}$. Finalmente, por cada grupo de señales \mathcal{S} se tiene un conjunto de cardinales $\Omega_{\mathcal{S}}$.

2.4. Sistemas dinámicos caóticos

Con el objetivo de estudiar el rendimiento de nuestro modelo, lo aplicamos a señales caóticas [Devaney \(2020\)](#) generadas utilizando los sistemas dinámicos de Henón, Chirikov, Schuster, Logístico y el modelo poblacional de Ricker (la figura 2 muestra a modo de ejemplo una gráfica de cada uno de ellos). Se generaron 6 señales de 10^6 puntos en cada conjunto. El cálculo se realizó para un total de 30 señales.

Los sistemas dinámicos analizados son los siguientes:

2.4.1. Mapa Logístico

Utilizado comúnmente para modelar el crecimiento de una población en función del tiempo

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n),$$

donde x_n representa la población en el tiempo n y r es el parámetro de crecimiento.

2.4.2. Mapa de Chirikov

Es un mapa conservador de área para dos variables dinámicas canónicas, es decir, momento y coordenada (p, x) . Está descrito por las siguientes ecuaciones:

$$\begin{cases} p_{n+1} &= p_n + K * \sin(x_n), \\ x_{n+1} &= p_n + x_n, \text{ mod}(2\pi), \end{cases}$$

donde (p_n, x_n) son las coordenadas en el plano fase y K es el parámetro de fuerza. Los mapas fueron generados con las siguientes condiciones iniciales: $x = 0,00, x = 0,11, x = 0,12, x = 0,13, x = 0,14$ y $x = 0,15$; $p = 0,1$ y $K = 1$. Para el análisis se utilizaron los valores correspondientes al eje x .

2.4.3. Mapa de Henón

Es un sistema dinámico altamente utilizado por su comportamiento caótico. Las ecuaciones que lo generan son las siguientes:

$$\begin{cases} x_{n+1} &= y_n + 1 - a \cdot x_n^2, \\ y_{n+1} &= b \cdot x_n, \end{cases}$$

donde (x_n, y_n) son las coordenadas en el plano fase y a y b son parámetros de control. Para generar las señales se establecieron las siguientes condiciones iniciales: $x = y = 0,1, b = 0,3, a = 1,4, a = 1,409, a = 1,41, a = 1,412, a = 1,415$ y $a = 1,42$. El análisis se ha realizado sobre los valores de la variable x .

2.4.4. Mapa de Schuster

Exhibe señales intermitentes con un espectro de potencia $P(f) \sim \frac{1}{f^z}$ (Boaretto et al. (2021)). Se define como:

$$x_{n+1} = x_n + x_n^z,$$

donde x_n representa la variable en el tiempo n y z es el parámetro de crecimiento.

2.4.5. Modelo Poblacional de Ricker

Es un sistema altamente utilizado en modelos de desarrollo demográfico y ecológico. May y McLean (2007)

$$N_{n+1} = N_n \cdot e^{r \cdot (1 - N_n/K)},$$

donde N_n es la población en el tiempo n , r es la tasa de crecimiento y K es la capacidad de carga del medio ambiente. Las siguientes condiciones iniciales se definieron de la siguiente manera: $K = 100, N = 10, r = 3,95, r = 3,951, r = 3,952, r = 3,9513, r = 3,954$ y $r = 3,955$. El análisis se ha realizado sobre los valores de la variable x .

Para los casos de Henón y Chirikov, se analizaron sólo los valores de uno de los ejes, debido a que estos sistemas dinámicos están definidos en \mathbb{R}^2 .

Para este experimento se definieron como valores de entrada para nuestro algoritmo $d = 4$ y $k = 50$, siendo d la dimensión de *embedding* y k el factor de discretización.

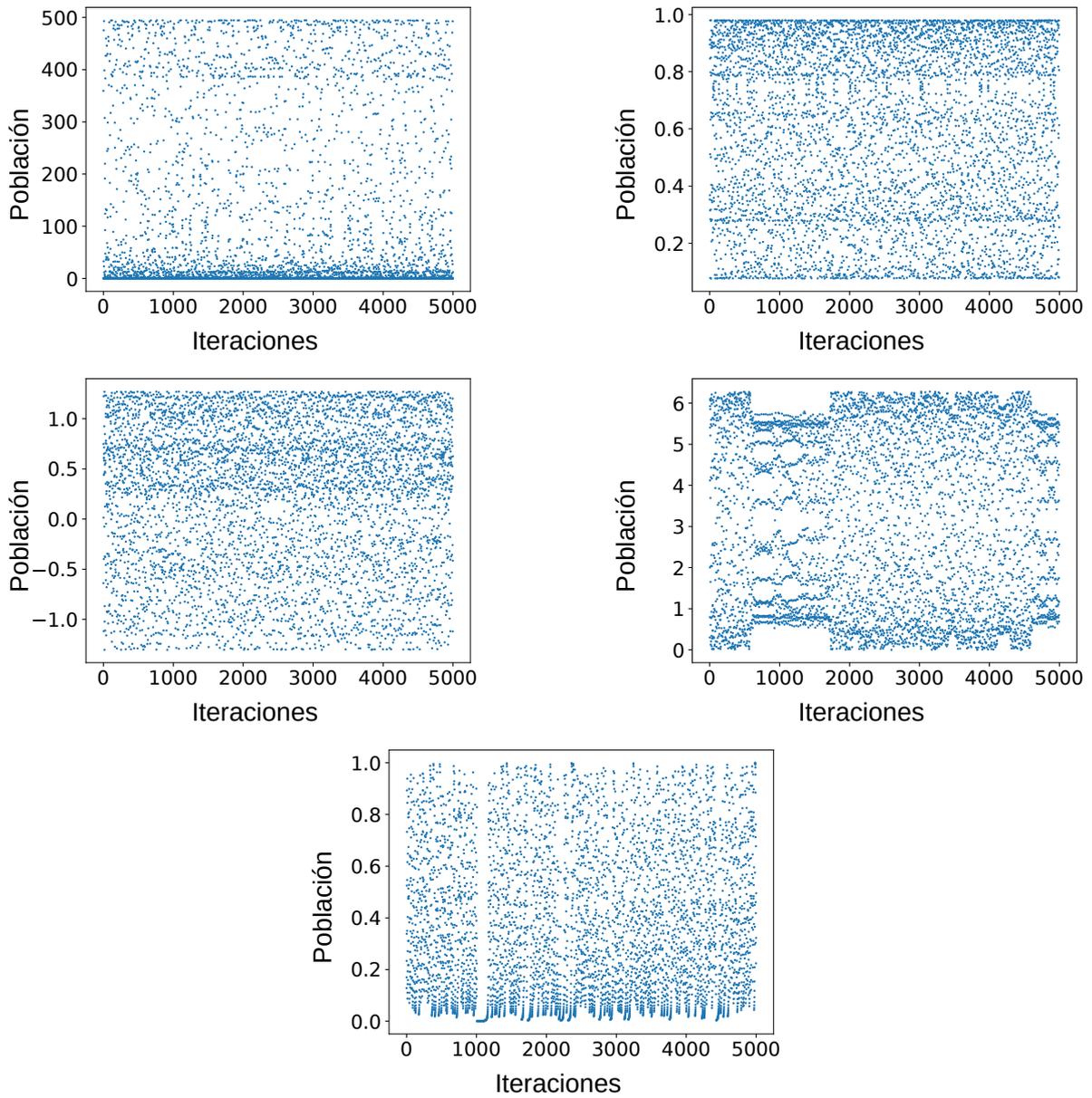


Figura 2: De arriba hacia abajo y de izquierda a derecha: mapa de Ricker (no inversible), mapa Logístico (disipativo), mapa de Henón (disipativo), mapa de Chirikov (conservativo) y mapa de Schuster (intermitente).

3. RESULTADOS

La tabla 2 muestra los promedios de los cardinales obtenidos de cada uno de los grupos de señales.

Modelo	Valor	Longitud
Logístico	266.8333	1,000,000
Schuster	288.3333	1,000,000
Ricker	496.0000	1,000,000
Henon	631.1667	1,000,000
Chirikov	1152.6667	1,000,000

Tabla 2: Tabla de cardinales promedio obtenidos de las de los cinco grupos de señales analizadas.

La Figura 3 muestra la distribución de los cardinales de los 5 grupos mediante *boxplots*. Se observa una separación clara entre los cinco grupos a excepción del mapa Logístico y el mapa de Schuster que muestran un solapamiento en el último cuartil del primer grupo. Las discrepancias y solapamientos observados en los *boxplots* subrayan la importancia de considerar tanto las medidas resumen como la variabilidad interna de los datos al analizar y comparar los conjuntos de señales.

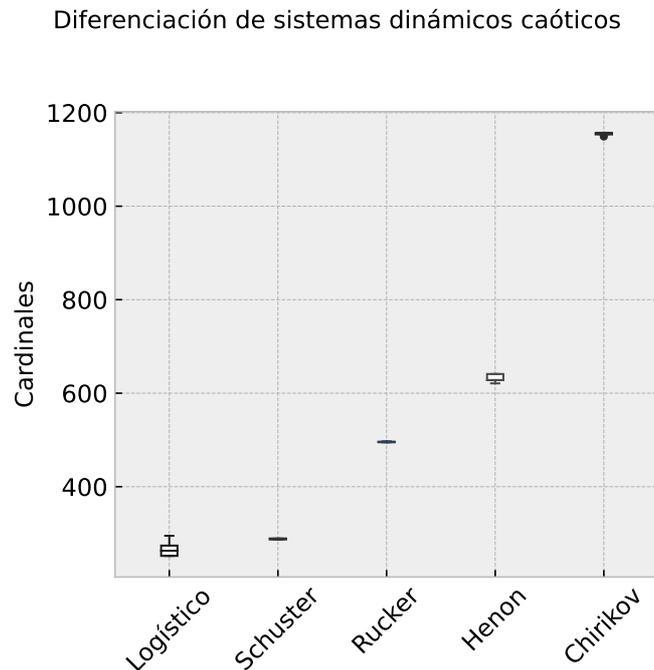


Figura 3: Distribuciones de los cardinales de señales de sistemas dinámicos caóticos.

4. DISCUSIÓN Y CONCLUSIONES

La premisa fundamental es que el conjunto de atributos propuesto en este estudio constituye una conjunto de características (como si se tratara de una huella digital) que posibilita la

discriminación entre señales de distintas procedencias mediante la presencia de ciertas configuraciones geométricas de la señal. Consideramos que los resultados expresan esta idea debido a la complejidad y a la longitud de las señales analizadas. Los cardinales promedio obtenidos son significativamente bajos respecto de la longitud de las señales estudiadas. Si realizamos una analogía con el concepto de entropía, en este sentido podríamos decir que el cardinal como herramienta de clasificación está indicando que las señales caóticas portan poca información ya el máximo valor obtenido de vectores (el caso del mapa de Chirikov) fue de 1152 sobre $10^6 - d + 1$ posibles vectores.

Podemos ver además que el cardinal pudo clasificar los cinco grupos de señales ya que se observa una separación clara expresada en los promedios (tabla 2). Un aspecto importante a recalcar es el hecho de que no fue necesario realizar ningún tipo de preprocesamiento.

Este aspecto resalta la noción de que el vector característico resulta una herramienta simple y eficiente para extraer atributos, al menos en los casos analizados. El cardinal resulta un herramienta de clasificación de lectura directa, lo que facilita la comprensión de los resultados y el posterior análisis. En consecuencia, el espacio de atributos generado de esta manera resulta adecuado para la discriminación o identificación de señales discretas o conjuntos de las mismas dentro del dominio del tiempo.

Una posible extensión de este trabajo es explorar el uso del espacio de características derivado del proceso de obtención del cardinal. Estos espacios contienen información única que permite identificar los grupos de señales analizados con precisión. Un análisis futuro podría enfocarse en utilizar estos espacios para caracterizar sistemas dinámicos caóticos, ampliando así las aplicaciones de la metodología propuesta.

REFERENCIAS

- Argyris J.H., Faust G., y Haase M. *An exploration of chaos: an introduction for natural scientists and engineers cover*, volumen 7. North Holland, 1994.
- Boaretto B.R.R., Budzinski R.C., Rossi K.L., Prado T.L., Lopes S.R., y Masoller C. Discriminating chaotic and stochastic time series using permutation entropy and artificial neural networks. 2021.
- Cencini M., Cecconi F., y Vulpiani A. *Chaos: from simple models to complex systems (Advances in Statistical Mechanics)*. World Scientific Publishing Company, 2009.
- Devaney R. *A first course in chaotic dynamical systems: Theory and experiment*. Chapman and Hall/CRC, 2020.
- Keller K., Mangold T., Stolz I., y Werner J. Permutation entropy: New ideas and challenges. *Entropy*, 19(3):134, 2017.
- May R. y McLean A.R. *Theoretical ecology: principles and applications*. Oxford University Press, 2007.
- Olivares F., Souza L., Legnani W., y Rosso O. Informational time causal planes: A tool for chaotic map dynamic visualization. En *Nonlinear Systems-Theoretical Aspects and Recent Applications*. 2019.
- Pinsky M.A. *Introduction to Fourier analysis and wavelets*, volumen 102. American Mathematical Society, 2023.
- Strang G. Wavelet transforms versus fourier transforms. *Bulletin of the American Mathematical Society*, 28(2):288–305, 1993.
- Tangirala A.K. *Principles of system identification: Theory and practice*. Crc Press, 2018.