

DIFERENCIACIÓN DE SEÑALES CAÓTICAS Y ALEATORIAS MEDIANTE CARACTERÍSTICAS GEOMÉTRICAS

M. G. Blesa, W. Legnani y A. Rey

Centro de Procesamiento de Señales e Imágenes, Facultad Regional Buenos Aires, Universidad Tecnológica Nacional hgarciablesa@frba.utn.edu.ar, walter@frba.utn.edu.ar, arey@frba.utn.edu.ar

Resumen: En este trabajo, proponemos un modelo basado en la construcción de un espacio de características utilizando descriptores geométricos innovadores, para identificar diferentes tipos de señales discretas 1D. Calculamos dichas características mediante el método de barrido con una dimensión de emcastre (*embedding*, en inglés), técnica ampliamente utilizada en aplicaciones de Teoría de la Información. Proponemos tres características: amplitud, ángulo de Zenit y factor de forma. A partir de las mismas, asociamos una terna a cada segmento de cada una de las señales de un grupo dado. Luego, reducimos el número de ternas concatenando la información repetida. El cardinal de este nuevo conjunto de ternas define la variable utilizada para la distinción de señales. Aplicamos la propuesta en un conjunto de señales sintéticas aleatorias y caóticas obteniendo muy buenos resultados para distinguir ambos tipos de señales.

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

2000 AMS Subject Classification: 94A12 - 97R50

1. INTRODUCCIÓN

La tarea de diferenciar señales es esencial por varios motivos, debido a que permite detectar características, comportamientos, conductas y otros aspectos importantes de las fuentes emisoras como la identificación de un fenómeno, la detección de una anomalía o el análisis a largo plazo [6]. Dentro de las técnicas de estudio de señales, se encuentran los patrones ordinales que son una herramienta ampliamente utilizada en los últimos veinte años, con muchas aplicaciones exitosas en casos de diversa naturaleza [7]. Para obtener más información sobre esta técnica, se puede consultar el libro de Amigó [1]. En la construcción del modelo propuesto en este estudio, se consideraron conceptos básicos como la amplitud, una medida angular y la forma de la señal en términos de variaciones de picos y valles. La hipótesis sobre los resultados al aplicar este modelo es que estas tres propiedades son suficientes para obtener una caracterización completa de la señal, como su huella digital. Probamos el modelo en la clasificación de sistemas dinámicos caóticos y ruido. Algunos antecedentes sobre la clasificación de este tipo de señales pueden encontrarse en trabajos tales como [7], [8], [9] y [10].

2. ESPACIO DE CARACTERÍSTICAS

Consideramos el espacio de características de una señal como el conjunto formado por vectores representativos 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 y con el fin de caracterizar la señal, se proponen 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.

Amplitud. Denotando como $\min(S)$ y $\max(S)$ a los valores mínimo y máximo de la señal S , respectivamente, calculamos la amplitud de S : $A(S) = \max(S) - \min(S)$.

Ángulo de Zenit. Calculamos 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 espacio acotado entre $\min(S)$ y $\max(S)$. Consideremos el siguiente intervalo en el espacio de dos dimensiones: $I_S = [0, 1] \times [\min(S), \max(S)]$. Definimos entonces el ángulo φ determinado por los segmentos \overline{PM} y \overline{PN} , donde $P = (0.5, p_S)$, $M = (1, \max(S))$ y $N = (1, p_S)$. Con el fin de comparar valores de φ de diferentes señales, normalizamos el intervalo I_S

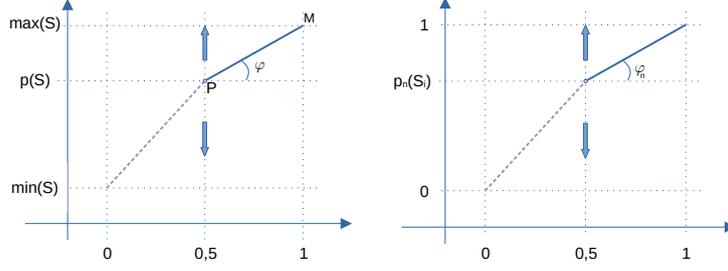


Figura 1: Representación gráfica del ángulo de Zenit original (izquierda) y normalizado (derecha).

por $I = [0, 1] \times [0, 1]$ y el promedio p_S por $p = [p_S - \min(S)] / [\max(S) - \min(S)]$. 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 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.

Factor de forma. Una señal de un fenómeno real puede tener una variedad de valles y picos en una determinada secuencia que contribuye 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 $S = \{11, 31, 28, 42, 28\}$ y una dimensión $d = 4$, tenemos dos segmentos: $S_1 = \{11, 31, 28, 42\}$ y $S_2 = \{31, 28, 42, 28\}$, cuyos factores de forma son $\mathcal{F}_1 = 3241350$ y $\mathcal{F}_2 = 10500$, respectivamente (ver Tabla 1 para los detalles de cálculo).

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 $(A(S_i), \varphi(S_i), \mathcal{F}(S_i))$. 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.

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

Acción	Resultado para S_1	Resultado para S_2
Señal original	{11, 31, 28, 42}	{31, 28, 42, 28}
Ordenamiento B	{11, 28, 31, 42}	{28, 28, 31, 42}
Eliminación repeticiones C	{11, 28, 31, 42}	{28, 31, 42}
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}	$2^1 3^3 5^2 7^4$	$2^2 3^1 5^3 7^1$

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 $\Omega_{\mathcal{S}}$.

3. CASOS DE ESTUDIO

Con el objetivo de estudiar el rendimiento de la propuesta, la aplicamos a dos grandes grupos de señales utilizando el conjunto de cardinales Ω como herramienta de clasificación. Para ello, consideramos un grupo de señales sintéticas de ruido blanco gaussiano, ruido blanco uniforme, ruido rosa, ruido azul, ruido rojo y ruido violeta [5], que agrupamos en una misma clase. El segundo grupo corresponde a señales caóticas [2] generadas utilizando la función logística $f(x) = \lambda x(1 - x)$, con $\lambda \in [3.65, 3.7]$ (para esta implementación nos referimos al trabajo de Lynch [4]), y los atractores caóticos Lorentz, Chua, Rossler y Rikitake (calculados mediante el código disponible en <https://github.com/hukenovs/chaospy>, última consulta realizada el 03-12-2022), que compilamos en un segundo grupo. En una segunda etapa, analizamos las señales de cada grupo por separado, obteniendo un total de 5 y 6 subgrupos, respectivamente. La metodología se llevó a cabo para un total de 100 señales de 10000 puntos por cada uno de los 11 subgrupo.

4. RESULTADOS

Las Figuras 2 y 3 muestran la distribución de los cardinales, mediante boxplots, de los dos grandes grupos y de sus correspondientes subgrupos, respectivamente. Podemos ver que el conjunto Ω , construido a partir del espacio de características propuesto, resulta una herramienta eficaz para la diferenciación de señales caóticas respecto de ruido, tal como se muestra en la Figura 2. La Figura 3 muestra que es capaz de distinguir incluso las diferentes señales caóticas y diferentes tipos de ruido entre sí. Para este experimento se eligieron $d = 5$ y $k = 10$, como valores de los parámetros de *embedding* y discretización, respectivamente.

5. DISCUSIÓN Y CONCLUSIONES

La hipótesis central se basa en considerar que el espacio de características propuesto en este trabajo conforma una huella digital que permite diferenciar señales de diferente origen sólo a partir de ciertas estructuras geométricas de la señal. Consideramos que los resultados expresan esta idea debido a la complejidad, la longitud y la cantidad de señales analizadas. En particular, los resultados de la clasificación de ruidos entre sí evidencian una alta efectividad en la caracterización. Este aspecto refuerza la idea de que el vector característico resulta una herramienta sencilla y efectiva para extraer cualidades, al menos en los casos estudiados. Como resultado de esto, el espacio de características así construido resulta apropiado para la diferenciación o distinción de señales discretas o grupos de ellas.

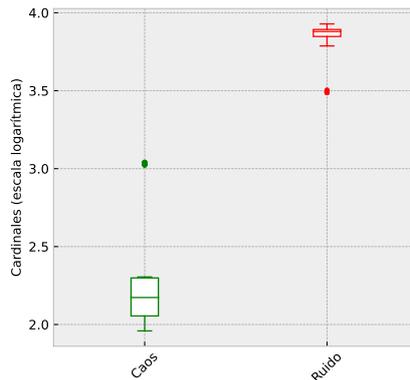


Figura 2: Distribuciones de los cardinales de señales de sistemas dinámicos caóticos y ruidos.

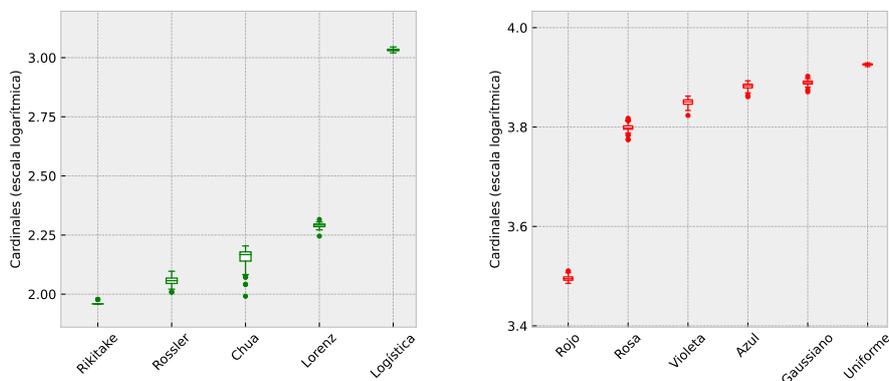


Figura 3: Distribuciones de los cardinales de los subgrupos de señales dentro de los grupos caóticos (izquierda) y de ruido (derecha).

REFERENCIAS

- [1] AMIGÓ, J., *Permutation complexity in dynamical systems: Ordinal patterns, permutation entropy and all that.*, Springer Science & Business Media, 2010.
- [2] DEVANEY, R., *A first course in chaotic dynamical systems: Theory and experiment.*, Chapman and Hall/CRC, 2020.
- [3] KELLER, K., MANGOLD, T., STOLZ, I., WERNER, J., *Permutation entropy: New ideas and challenges.* *Entropy*, 19(3), 134, 2017.
- [4] LYNCH, S., *Dynamical Systems with Applications using Python.*, Birkhäuser, 2018.
- [5] SHYNK, J. J., *Probability, random variables, and random processes: Theory and signal processing applications.*, John Wiley & Sons, 2012.
- [6] TANGIRALA, A. K. *Principles of system identification: Theory and practice.*, Crc Press, 2018.
- [7] GAO, J. B., HU, J., TUNG, W. W. AND CAO, Y. H., *Distinguishing chaos from noise by scale-dependent Lyapunov exponent*, American Physical Society, *Phys. Rev. E*, vol. 74, no 6, p. 066204, 2006.
- [8] OLIVARES, F., SOUZA, L., LEGNANI, W. AND ROSSO, O., *Informational time causal planes: A tool for chaotic map dynamic visualization*, *Nonlinear Systems-Theoretical Aspects and Recent Applications*, 2019.
- [9] CENCINI, M., CECCONI, F. AND VULPIANI, A., *Chaos: from simple models to complex systems (Advances in Statistical Mechanics)*, Series on Advances in Statistical Mechanics, World Scientific Publishing Company, 2009.
- [10] ARGYRIS, JOHN H., GUNTER FAUST, AND MARIA HAASE, *An exploration of chaos: an introduction for natural scientists and engineers cover*, Vol. 7. North Holland, 1994.