Journal of Theoretical Biology 364 (2015) 197–205

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Tortuosity entropy: A measure of spatial complexity of behavioral changes in animal movement

Xiaofeng Liu [a,b,*], Ning Xu [a,b], Aimin Jiang [a,b]

[a] College of IoT Engineering, Hohai University, Changzhou 213022, China
[b] Changzhou Key Laboratory of Robotics and Intelligent Technology, Changzhou 213022, China

HIGHLIGHTS

- Tortuosity entropy (TorEn), a novel measure for analyzing animal tracking data.
- TorEn can be used to analyze all the parameters of trajectory, such as heading, bearing, speed, and distance between successive track points.
- TorEn can be easily applied to arbitrary real world data, whether deterministic or stochastic, stationary or non-stationary.

ARTICLE INFO

Article history:
Received 24 December 2013
Received in revised form
16 September 2014
Accepted 17 September 2014
Available online 26 September 2014

Keywords:
Animal movement
Entropy
Spatial complexity
Trajectory

ABSTRACT

The goal of animal movement analysis is to understand how organisms explore and exploit complex and varying environments. Animals usually exhibit varied and complicated movements, from apparently deterministic behaviours to highly random behaviours. It has been a common method to assess movement efficiency and foraging strategies by means of quantifying and analyzing movement trajectories. Here we introduce a tortuosity entropy (TorEn), a simple measure for quantifying the behavioral change in animal movement data. In our approach, the differences between pairwise successive track points are transformed into symbolic sequences, then we map these symbols into a group of pattern vectors and calculate the information entropy of pattern vectors. We test the algorithm on both simulated trajectories and real trajectories to show that it can accurately identify not only the mixed segments in simulated data, but also the different phases in real movement data. Tortuosity entropy can be easily applied to arbitrary real-world data, whether deterministic or stochastic, stationary or non-stationary. It could be a promising tool to reveal behavioral mechanism in movement data.

1. Introduction

The movement of an organism plays a fundamental role in the structure and dynamics of populations (Nathan et al., 2008). In recent years, the study of movement of organisms has been a rapidly expanding field of research, and the emphasis of movement research has shifted from quantifying population redistribution to quantifying movement of individuals (Nathan et al., 2008). A growing body of work has been devoted to the movement of individual organisms due to technological advances of tracking devices, e.g., GPS and Argos (Tomkiewicz et al., 2010). Full range technological advances from the microelectronics, wireless telecommunication, satellite telecommunication, to high-capacity storage make possible the development of a high performance track logger with miniaturized size and ultra light weight. As a result, there has been an accumulation of movement data of various organisms (Nathan et al., 2008), which raises the challenges of analyzing movement data, including quantitatively characterizing the spatiotemporal trajectories and retrieving relevant information to reveal the causal links between animal movements and internal states and external factors.

A variety of methods have been developed for extracting behavioral information or for characterizing behavioral change in movement data (Batschelet, 1981; Bovet and Benhamou, 1988; Dicke and Burrough, 1988; Fauchald and Tveraa, 2003; Benhamou, 2004; Roberts et al., 2004; Horne et al., 2007; Barraquand and Benhamou, 2008; Gaucherel, 2011; Gurarie et al., 2009; Nams, 1996). Straightness index (SI) is the most straightforward measure of how behavior varies over spatiotemporal scales (Batschelet, 1981). It is defined as the ratio of the beeline distance between

* Corresponding author at: College of IoT Engineering, Hohai University, Changzhou 213022, China.
E-mail address: xfliubme@gmail.com (X. Liu).

http://dx.doi.org/10.1016/j.jtbi.2014.09.025
0022-5193/

the start and the end of a trajectory to the length of a traveling trajectory. It has been further extended to a multi-scale straightness index (Postlethwaite et al., 2013). Sinuosity is another index that measures the degree of how tortuous a random search trajectory is. It is determined by distributions of changes of both direction and step length (Bovet and Benhamou, 1988). Fractal dimension is also used to characterize the tortuosity of a trajectory (Dicke and Burrough, 1988; Nams, 1996; Atkinson et al., 2002; Fritz et al., 2003). All of them are a class of measure of the spatial features of trajectory. Another class of method is to quantify the temporal characteristics of movement data. First-passage time (FPT) (Fauchald and Tveraa, 2003), defined as the time required for an animal to cross a given area, is a scale-dependent metric which indicates search effort of animal along a path (Fauchald and Tveraa, 2003; Pinaud, 2008). Similar to FPT, residence time (Barraquand and Benhamou, 2008) in the vicinity of successive path locations is used to identify a profitable place. The periodicity of movement recursion has received attention too. Well-established Fourier and wavelet analyses are both used to extract periodic patterns or to detect the regimes shift (Polansky et al., 2010; Gaucherel, 2011; Riotte-Lambert et al., 2013). Unlike the traditional feature extraction methods, Roberts and his colleagues have proposed positional entropy (Roberts et al., 2004) to measure the uncertainty of a navigational trajectory by means of calculating the stochastic complexity of the trajectory (Rezek and Roberts, 1998).

Animal movements can be described as a time series of movement steps. Motivated by the method of analyzing time series using complexity and entropy, in this paper we introduce tortuosity entropy for analyzing animal trajectories. Here we focus on the pattern of local movement change by presenting an algorithm to quantify behavioral change in a small scale. We represent the difference between two successive track points by a symbol, then map the symbol series into a pattern vector. Finally, we calculate the information entropy of pattern vector, termed tortuosity entropy (TorEn). Our method can be applied to various parameters of trajectory, e.g., heading, speed, bearing, and position, and moreover, and it can be easily applied to arbitrary real-world data – deterministic or stochastic, stationary or non-stationary.

2. Methods

In this section, we introduce the definition of tortuosity entropy. The basic idea is that animal movements are mapped into a symbolic sequence through quantifying the relative change of consecutive points of a movement series. Then the information entropy of the symbolic sequence is defined as tortuosity entropy, which is an indicator of how tortuous the animal movement is. The procedure of TorEn consists of two main steps: (1) quantification and symbolization of the time-delay embedding representation of a movement path, and (2) calculation of information entropy of the reconstructed symbolic sequence. This procedure is summarized in Fig. 1. and described below in detail.

2.1. Quantification and symbolization

We consider a raw time series of N samples, $\{x(n) : 1 \le n \le N\}$ and form $N-(m-1)\tau$ vectors $X_j^{m,\tau}$ for $\{j|1 \le j \le N-(m-1)\tau\}$, where $X_j^{m,\tau} = x(j), x(j+\tau), \ldots, x(j+(m-1)\tau)$, where m and τ denote embedding dimension and time delay, respectively. The raw time series can be absolute positions, headings, or speeds of an animal movement. To simplify our calculation, we assume that each track point in movement data is recorded at a fixed time interval. Here we are interested in the difference between consecutive elements in reconstruction vector $X_j^{m,\tau}$. We firstly define the differences between two consecutive elements in $X_j^{m,\tau}$ as

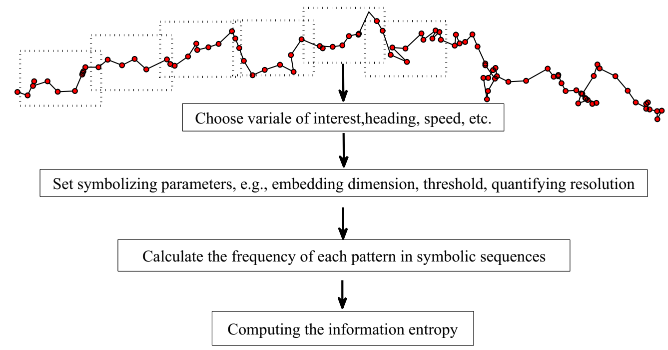

Fig. 1. Schematic description of the steps of the tortuosity entropy.

$\{d_j(i)\}_1^{m-1} = \{x(j+i\tau) - x(j+(i-1)\tau) : 1 \le i \le m-1\}$. There exist a variety of approaches to quantify $\{d_j(i)\}_1^{m-1}$. Here we set an integer, $\lambda > 0$, as the threshold and then compute $k(i) = \lfloor |d_j(i)|/\lambda \rfloor$, the largest integer not greater than $|d_j(i)|/\lambda$. We further quantify the element of $\{d_j(i)\}_1^{m-1}$ using the following equation:

$$s_j(i) = \mathrm{sgn}(d_j(i)) \cdot k(i), \tag{1}$$

where sgn() is a sign function defined by

$$\mathrm{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0. \end{cases}$$

For a special case $\{d_j(i)\}_1^{m-1}$ can be coarsely quantified only in terms of the sign of its elements for simplicity. To limit the range of a symbol, we let $k(i) = K, i = 1, \ldots, m-1$ when $|k(i)|$ is larger than a positive integer K. The quantifying resolution is $2K+1$. Then each vector $X_j^{m,\tau}$ is mapped into a word, $s_j(1) \cdots s_j(m-1)$. As a result, the raw time series is symbolized into $N-(m-1)\tau$ words of length $m-1$.

2.2. Entropy of reconstruction vector

Each vector $X_j^{m,\tau}$ has a corresponding word after symbolization. As a result, a time series of N samples, $\{x(n) : 1 \le n \le N\}$, can generate a word set of $N-(m-1)\tau$ words. The probability of each word, p_j, is defined as the relative frequency that it occurs in the entire word set

$$p_j = \frac{\text{The number of a word appears in the set}}{N-(m-1)\tau} \tag{2}$$

Shannon entropy is the average uncertainty in a random variable. The tortuosity entropy of order $m \ge 2$ is defined as the Shannon entropy of the word. Because there exist $(2K+1)^{m-1}$ possible patterns for each word of length $m-1$, the upper limit of the summation is $(2K+1)^{m-1}$

$$H(m) = -\sum_{j=1}^{(2K+1)^{m-1}} p_j \log p_j. \tag{3}$$

Obviously, $H(m)$ will reach the minimum, 0, for a fully localized probability distribution where $p_j=1$ for a specific word and $p_j=0$ for other words. For example, if a time series increases or decreases monotonously (e.g., an ordered sequence values), its tortuosity entropy will equal to zero. $H(m)$ will reach the maximum, $(m-1)\log(2K+1)$, for a uniform distribution. $H(m)$ is normalized by $m-1$ (Bandt and Pompe, 2002)

$$h(m) = \frac{H(m)}{m-1}. \tag{4}$$

Tortuosity entropy can be calculated for different embedding dimensions m. In practice, we recommend $m=3$, 4, 5. In general,