Extraction of movement features for cross-scale behavioral classification

Soleymani, Ali; Weibel, Robert

Abstract: Since movement takes place in space and time and behavioral movement patterns are likely to be valid within different unique ranges of spatial and temporal scales, it is imperative to apply a cross-scale analysis approach in both spatial and temporal domains to yield behavioral clues about movement. Particularly in a classification problem, where the main interest is to detect specific behaviors in movement data, relevant features are needed to be extracted in the spatial and temporal domains, respectively, in a coordinated fashion. In this paper, a methodology for joint spatial and temporal, cross-scale behavioral classification of movement data is introduced, and the potential features that can be used in this cross-scale classification problem as well as their integration will be discussed.

Extraction of movement features for cross-scale behavioral classification

Ali Soleymani, Robert Weibel

Department of Geography, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland
ali.soleymani@geo.uzh.ch
robert.weibel@geo.uzh.ch

1. Introduction

Since movement takes place in space and time and behavioral movement patterns are likely to be valid within different unique ranges of spatial and temporal scales, it is imperative to apply a cross-scale analysis approach in both spatial and temporal domains to yield behavioral clues about movement. Particularly in a classification problem, where the main interest is to detect specific behaviors in movement data, relevant features are needed to be extracted in the spatial and temporal domains, respectively, in a coordinated fashion. In this paper, a methodology for joint spatial and temporal, cross-scale behavioral classification of movement data is introduced, and the potential features that can be used in this cross-scale classification problem as well as their integration will be discussed.

2. Extraction of movement features

The selection of relevant input features and the way in which they are extracted crucially decides on the performance of any classification model. This general observation holds particularly in classification problems that rely on movement data, since feature extraction is further complicated by the presence of both spatial and temporal domains, and the need to extract the features at the scale range where the key characteristics of the particular movement behaviors are best captured. The scale at which the movement features are extracted is defined in several ways. First of all, movement features can be related to the level of individual fixes of the trajectory or aggregated to the trajectory as a whole. If the movement data have been sampled at sufficiently high temporal resolution, resampling and cross-scale extraction of movement features becomes possible (Laube and Purves 2011), or mid-range features can be extracted, for instance, by segmenting the tracking data into sub-trajectories of homogeneous movement characteristics (Dodge et al. 2009, 2012). Although in principle, many variations are possible in feature extraction for movement classification, very few studies have actually explored these to date. Hence, in Soleymani et al. (2014) a three-stage methodology was proposed for joint cross-scale spatial and temporal behavioral classification of movement (Figure 1). The main contribution of this work will be an extension of that framework, proposing concrete features that lend themselves for cross-scale movement classification.

2.1 The methodology

The methodology of Soleymani et al. (2014) consists of three main stages: extraction of features in the spatial domain, extraction of features in the temporal domain and integration of both feature sets for the final classification (Figure 1).

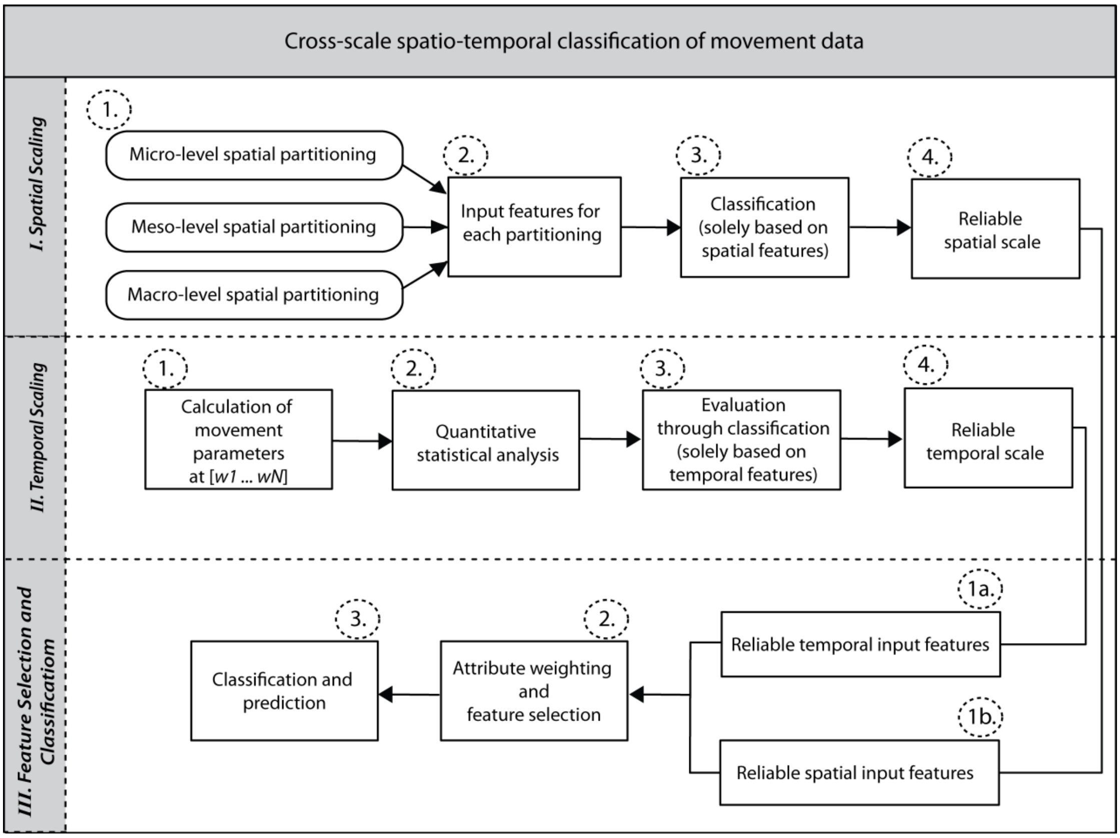

Figure 1. Overview of the methodology proposed by Soleymani et al. (2014) for cross-scale behavioral movement classification

The first two feature extraction stages lead to identifying the reliable spatial and temporal scales, respectively. The 'reliable scale' is the scale at which the best classification performance is achieved. The final integration stage is achieved by using a feature selection procedure, thus retaining the most relevant features contributing to known behavioral states and using these features to learn the classification model. In order to demonstrate the feasibility of this methodology, Soleymani et al. (2014) used the case study of classifying behaviors in zebrafish, expressed through different swimming patterns under different neuropharmacological drug treatments.

Cross-scale movement analysis in the *spatial domain* means that the underlying movement space is partitioned into different zones (e.g. the tank arena where the fish is swimming is divided into different zones) and the features related to each zone are extracted. As shown in Soleymani et al. (2014), partitioning of space strongly affects the classification results, as well as it can be used for linking the contextual data in other cases. However, besides the features which are developed based on partitioning of movement space, other features capable of measuring the response to the spatial scale may be considered, as will be shown below.

The features related to the *temporal domain* should be capable of investigating the variation of movement patterns engrained in the trajectory of a moving object across different scales. In the zebrafish case study, these features include statistical descriptors of movement parameters per trajectory, i.e. maximum, minimum, standard deviation and average values of velocity, acceleration, turning angle, meandering and sinuosity, computed at different temporal scales. Again, further features are possible that are more inherently multi-scale, as will be discussed in the following sections.

2.2 Features in the spatial domain

In the work by Soleymani et al. (2014), only one feature was used to explore the spatial domain, i.e. the time spent in different spatial zones using a hierarchical spatial partitioning scheme with different zoning levels. Other examples of spatial features that are based on hierarchical spatial partitioning include characteristics of movement parameters within zones (e.g. descriptive statistics of MP values for each zone); contextual information linked to the zones (e.g. certain zones might be more prone to specific behaviors or they might be related to particular food resources); and frequency of transitions between zones. Here, we introduce two other features — fractal dimension and first-passage time — capable of measuring the response of a movement trajectory to different spatial scales. These measures do not inherently consider the zoning levels, but depending on the research questions, can be measured for sub-trajectories placed in different zones.

Fractal dimension: Fractal dimension has been used for characterizing different mobility patterns in movement data. In movement ecology for example, since the patchy nature of habitat causes movement path structure to vary throughout space and the hierarchical nature of habitat causes movement path structure to vary with spatial scale, this measure has been used as an indicator of path tortuosity across different scales (Nams 2005). Different forms of fractal dimension (e.g. Fractal D and vFractal) can be used. The vFractal estimator is not only useful to have an idea of the 'path tortuosity', but of the scale level at which 'search patterns' occur (Nams 2005). So, computation of fractal dimension across different scales may provide relevant features for the classification. Because different scales are often associated with different driving processes, the fractal dimension may have the desirable feature of only being constant over a finite range of measurement scales, thus providing hints to important scale breaks.

First-passage time: The first-passage time has been studied well in movement ecology (Fauchald and Tveraa 2003). It is defined as the time required for an animal (or any other moving object) to cross a circle with a given radius. It is a measure of how much time an animal spends within a given area. First-passage time is scale dependent (through the chosen circle radius), and a plot of variance in first-passage time vs. spatial scale reveals the spatial scale at which the animal concentrates its search effort. By averaging the first-passage time on a geographical grid, it is possible to relate first-passage time to environmental variables and the search pattern of other individuals (Fauchald and Tveraa 2003).

2.3 Features in the temporal domain

The main focus in the temporal domain is on features that are based on movement parameters extracted from trajectories, incl. velocity, sinuosity, turning angle, and others. Laube & Purves (2011) have proposed a method for cross-scale analysis of these parameters, which has been also employed in the study by Soleymani et al. (2014). However, other than using only descriptive statistics of movement parameters assigned to entire trajectories, developing features that can investigate the variations in movement parameter values in a more meaningful way is crucial in trajectory classification problems. In the following, wavelet analysis and approximate entropy and the way that they can be used on profiles of movement parameters will be discussed. These profiles are achieved by plotting the sequential data of movement parameters against time and by using them, structural differences of different behaviors can be observed along the time axis (Dodge et al. 2009). Both wavelet analysis and approximate entropy are common methods in time series analysis, yet have been hardly used in movement classification.

Wavelet coefficients: Wavelet analysis is one of the methods for multi-scale time-frequency representation of a signal (Daubechies 1990). In movement analysis, the considered signal will be the profiles of movement parameters. Fourier transform, as another example of time-frequency representation, is helpful for studying stationary or periodic time series (Daubechies 1990). However, most movements are non-homogeneous, made up of a combination of discrete behaviors. Wavelet analysis can be used for analyzing such non-stationary time series where the frequency content changes over time or when transient types of activity occur (Wittemyer et al. 2008 and Gaucherel 2011). Wavelet analysis resolves the frequency content of a time varying variable (e.g. velocity, sinuosity, etc.) by successively analyzing the autocorrelation properties of the movement. Since the signals in movement analysis are in discrete form (sampled fixes over time), discrete wavelet coefficients of different scales of the movement parameter profiles may provide useful input features for a classification model.

Approximate entropy: Approximate entropy is a classic method in time-series analysis for quantifying regularities and fluctuations in sequential data (Pincus 1991). As a measure of system complexity, higher values of approximate entropy suggest a more random behavior, while a smaller value implies less complexity and more regularity. Thus, approximate entropy of movement parameter profiles shows how the structural complexity of movement parameter varies over time (Li 2014). For cross-scale analysis, the main interest is to determine how the values of approximate entropy vary when they are applied over movement parameter profiles computed at different temporal scales. This may then reveal how different behaviors manifest themselves at different scales.

2.4 Integration of features

The corresponding feature sets extracted from both the spatial and temporal domains must be integrated in order to perform the final classification. The main idea in this step is to investigate the importance of different features and evaluating their contribution to identifying particular behaviors. Following the methodology of Figure 1, automatic dimensionality reduction techniques may be used to define the contribution of the individual features and optimize the feature selection process in the classification. As explained in Soleymani et al. (2014), the cross-scale analysis in the spatial and in the temporal domains may be necessary, as it pays off even more when we combine the features from both these domains. As a further work, it would be interesting to see if this step of the proposed methodology can be used to identify the reliable spatio-temporal scale.

3. Conclusions

Extraction of relevant features in a trajectory classification problem is a crucial step for the detection of behaviors in movement data. This paper proposed several measures that could be used as input features in cross-scale classification of movement behaviors, as opposed to the commonly used single-scale classification models. These features were suggested since they appear to be intuitive candidates for addressing cross-scale problems, yet haven’t been used in movement classification so far. Owing to the interactive and informal nature of this workshop, we would be interested to discuss the usefulness of the proposed features with the participants of the workshop. In particular, have other participants already used any of these features in their research? How have they addressed cross-scale issues? And do they know any alternative features that might be useful in cross-scale movement behavior classification?

References

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5), 961–1005

Dodge, S., Laube, P., & Weibel, R. (2012). Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science*, 26(9), 1563–1588.

Dodge, S., Weibel, R., & Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems*, 33(6), 419–434.

FAUCHALD, P., & TVERAA, T. (2003). USING FIRST-PASSAGE TIME IN THE ANALYSIS OF AREA-RESTRICTED Reports. *Ecology*, 84(2), 282–288.

Gaucherel, C. (2011). Wavelet analysis to detect regime shifts in animal movement. *Computational Ecology and Software*, 1(2), 69–85.

Laube, P., & Purves, R. (2011). How fast is a cow? Cross-Scale Analysis of Movement Data. *Transactions in GIS*, 15(3), 401–418.

Li, X. (2014). Introducing Complexity Measures to Trajectory Classification Modeling to Automatically Detect Different Movement Types with Unknown GPS Trajectories. *American Journal of Geographic Information System*, 3(2).

Nams, V. O. (2005). Using animal movement paths to measure response to spatial scale. *Oecologia*, 143(2), 179–88.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6), 2297–301.

Soleymani, A., Cachat, J., Robinson, K., Dodge, S., Kalueff, A. V, & Weibel, R. (2014). Integrating cross-scale analysis in the spatial and temporal domains for classification of behavioral movement. *Journal of Spatial Information Science*, 8(8), 1–25.

Wittemyer, G., Polansky, L., Douglas-hamilton, I., & Getz, W. M. (2008). Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proceedings of the National Academy of Sciences*, 105(49), 19108–19113.