Computers and Geosciences 121 (2018) 53–66

Contents lists available at ScienceDirect

Computers and Geosciences

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

A framework for natural phenomena movement tracking – Using 4D dust simulation as an example

Manzhu Yu[a], Chaowei Yang[a,*], Baoxuan Jin[b,1]

[a] NSF Spatiotemporal Innovation Center, Department of Geography and Geoinformation Science, College of Science, George Mason University, Fairfax, VA, 22030, USA

[b] Yunnan Provincial Information Center for Natural Resource, Kunming, Yunnan, China

ARTICLE INFO

Keywords:
Feature tracking
Spatiotemporal feature detection
4D
Dust simulation

ABSTRACT

Natural phenomena evolve in space and time are often highly dynamic. Numerical simulations and earth observations have provided the capability to capture and study the complex evolvement of natural phenomena in a discrete fashion. It is demanding but challenging to extract events from these datasets automatically. Based on the previous research on feature identification, this research presents a movement tracking framework to analyze evolvements and dynamic movements of detected events. The framework consists of three components: feature identification, movement tracking, and track simplification. Based on the proposed framework, dust storm events are systematically detected and analyzed concerning their dynamic movements from a 4D (x, y, z, and t) simulation dataset over North Africa, the Mediterranean, and the Middle East from December 2013 to November 2014. The systematic research includes single event, multi-event, and seasonal analyses. Evaluation of the detected dust events shows that the tracked dust events align well with observations, with ~80% identification accuracy and consistency in the movement pattern. To briefly demonstrate its capability, we adopted the proposed framework to detect precipitation events from 3D (x, y, and t) precipitation observation data.

1. Introduction

Natural phenomena evolve in space and time and can be highly dynamic (Yang et al., 2011). With the improvement of numerical simulations and earth observations in spatiotemporal resolution and coverage, scientists and researchers can capture and study complex physical processes and evolution patterns in a discrete fashion. These simulations and observations can be three-dimensional (3D: x, y, and t) or four-dimensional (4D: x, y, z, and t), allowing the investigation of the movement patterns of natural phenomena in the temporal and vertical dimensions. The obtained knowledge or insights may include "where and when natural phenomena happen," "how long a natural phenomenon lasts," or "what the common transport pathway is for a natural phenomenon."

GIScience methodologies and techniques assist the understanding of dynamic geographic changes over space and time, but challenges remain in handling complex natural phenomena, especially for data with higher dimensions (Yuan, 2001; Worboys, 2005; Pultar et al., 2010). The increasing spatiotemporal resolution of simulations and earth observations has become more complicated for scientists to examine manually. Although numerical simulations and earth observations provide the spatiotemporal data source, researchers and scientists still need to develop algorithms to identify and track the movement of features (e.g., thunderstorm, hurricane, ocean eddy). Automatically identifying and tracking features are challenging; because features are moving with changing boundaries and capable of splitting and merging, and these movement patterns distribute over space and time. Therefore, providing an efficient way to detect these movement patterns is essential to the natural phenomena analysis. Besides, tracking features at different thresholds convey different information about the phenomena. It is essential to be able to track the movement of events at various thresholds efficiently.

The objectives of our research are threefold: 1) identify features based on out previous work (Yu and Yang, 2017) and introduce the tracking framework to connect the identified features in consecutive time steps; 2) apply the framework to a 4D simulation dataset; and 3) analyze the evolvements and dynamic movements of the events. Dust events are chosen as case studies to illustrate how this tracking approach can be used to represent and analyze the dynamic movements of natural phenomena. For an individual dust event, it is essential to

* Corresponding author.
E-mail address: cyang3@gmu.edu (C. Yang).

[1] Manzhu Yu co-designed the research, conducted experiments, and wrote the manuscript. Chaowei Yang formed the research idea, co-designed the research, analyzed the results and revised the manuscript. Baoxuan Jin contributed to methodology design.

https://doi.org/10.1016/j.cageo.2018.10.003
Received 7 April 2018; Received in revised form 24 August 2018; Accepted 1 October 2018
Available online 04 October 2018

understand the physical process of dust up-lift from arid and semi-arid regions, transport in the air, and deposition back to the ground. Besides, one of the major topics towards mineral dust is the spatiotemporal patterns of dust transport from desert source regions (Israelevich et al., 2003; Prospero and Lamb, 2003; Borbély-Kiss et al., 2004). Dust events that originate from a specific source show region-specific patterns of transport pathways (Israelevich et al., 2003; Moulin et al., 1998). Natural phenomena tend to interact with each other during transport in the atmosphere, such as split and merge (El-Askary et al., 2002; Ammar et al., 2014; Stein et al., 2015). These complex physical processes can only be adequately addressed in a 3D or 4D environment.

Related research of 3D feature tracking is reviewed in Section 2. Section 3 introduces the movement tracking approach. Section 4 analyzes dust storm events, their dynamic changes, and transport pathways, and Section 5 evaluates the resulting dust events with visibility observations, and archived dust events in NASA Earth Observatory. Finally, Section 6 offers conclusion from this research followed by potential future work.

2. Related works

Tracking natural phenomena includes two categories of methods: centroid- and overlapping-based. Centroid-based tracking methods normally treat the features in consecutive time steps whose centroids are within a certain radius in the same track. Johnson et al. (1998) made a guess on the cell centroid locations at t_{n-1} to where they would be at t_n based on its positions at several previous times, and assigned each cell at t_n to the closest unassigned centroid within a certain search radius. Lakshmanan et al. (2009) tied projected cells within a size-based radius (given by $\sqrt{(A/\pi)}$, where A is the area of the storm, when dealing with 2D) across different time steps.

Tracking methods using overlapping mechanism require that spatial and temporal frequencies be high enough regarding the expected size and speed of the features to track (Samtaney et al., 1994). Otherwise, unlinked features need to be associated using additional information in a second iteration of tracking. Choi et al. (2009) calculated the degree of association between overlapping storm features using an inverse cost function. The degree of association reflects the size similarity and moving speed of two associated features. Han et al. (2009) treated cells at t_n that have 50% or more significant overlap with cells from t_{n-1} as first matched, while unmatched cells are associated using a global cost function or assigned a new ID. A better approach is that of Dixon and Wiener (1993), where it utilized a combined approach of areal overlapping and centroid matching. First, storms that overlap significantly at two successive times are likely to be from the same storm. Then, an optimization scheme determines the most likely match between storms identified at successive scans. The optimization selects the track paths with shorter lengths, connects storms with similar characteristics (such as size and shape), and eliminates those tracks that exceed the maximum expected speed of storm movement.

Choosing the right approach is closely related to the spatiotemporal scales of datasets. Overlap-based methods are more suitable for tracking more substantial features with higher temporal resolution, and centroid-based methods may not be suitable for features with variable sizes (Lakshmanan et al., 2009). Therefore, a tracking approach needs to be designed and developed specifically for the natural phenomenon.

3. Feature identification and tracking methodology

The procedure of constructing the four entities of the framework includes the following: 1) identification of static dust storm features; 2) track features over consecutive time steps; and 3) composition of the event through tracking results (Fig. 1a).

3.1. Data and implementation

Dust simulation outputs were obtained from BSC-DREAM8bv2.0 (Pérez et al. 2006a, 2006b; Basart et al., 2012), a dust forecast operational system with the updated version of the former Dust Regional Atmospheric Model (DREAM; Nickovic et al., 2001) maintained by the Barcelona Supercomputing Center. The simulated dust concentration data include 12 months from December 2013 to November 2014 and cover a standard latitude/longitude grid of approximately 0.3° × 0.3° resolution for the broad north African and European domain (25.7W–59.3E, 0.76S–64.3N). The temporal resolution is hourly, and each time step contains voxel number of 256 × 196 × 24 (latitude, longitude, pressure level) (Fig. 1b).

The implementation was conducted in Python (Van Rossum and Drake, 1995) as a prototype, including the feature identification (Section 3.2) and tracking algorithms (Section 3.3). Visualization was implemented with the assistance of third-party libraries, including Numpy (Van Der Walt et al., 2011) and Scikit-learn (Pedregosa et al., 2011).

3.2. Identifying static meteorological features

The identification of meteorological features at each time step is conducted using a region-grow based algorithm that integrates the idea of the region-grow algorithm (Zucker, 1976) into 3D context; the simplified pseudo-code version of the algorithm is illustrated in Fig. 2. The computational complexity of this algorithm is quadratic concerning the number of voxels at two consecutive time steps. This algorithm is based on Yu and Yang (2017). The original algorithm has a multi-thresholding approach, which facilitates the identification of multiple high-concentration substorms within a larger low-concentration system. In this research, it is simplified to a single-thresholding approach, so that the tracked dust event is of the same level of concentration. A meteorological feature is specified as a contiguous volume with a concentration/intensity value greater than a threshold (*Dth*), while its volume is greater than a threshold (*Vth*). For each identified meteorological feature, the geometry is calculated using the boundary extraction method - Marching Cubes (Lorensen and Cline, 1987). Besides, associated attributes are calculated (e.g., concentration-weighted centroid in degree, speed in degree/hour, number and position of pixels in the dust storm object, area in degree*degree, maximum and average concentration or intensity). For the 4D dust simulation data, a dust concentration threshold of 360 μg/m3 and a volume threshold of 10 voxels were used.

3.3. Tracking the linkages of features over consecutive time steps

After the feature identification strategy (Section 3.2) is applied to the meteorological data, a list of feature objects exists for each time step. The tracking algorithm associates these objects across time to track the progress of the features as they form, move, and dissipate.

3.3.1. Overlapping strategy

Since the volumetric size of dust feature varies from 10 to 500 voxels, an overlap-based method was developed with an additional check, detailed in the pseudo-code (Fig. 3). In the first overlap check, the dust features over consecutive time steps are checked to track the potential linkages, based on the assumption that meteorological features from a later time step have partial overlap with those from an earlier time step. This overlap approach performs a matching test on features extracted from one timestep with all of the features extracted from the subsequent time step, and all combinations of features from dataset t_{i+1} (for amalgamation/bifurcation). The best match is selected by minimizing the cost function defined as follows: $\beta = 100 - O(C^t, O^{t-1})$, where O is a function measuring the percentage of overlap between the candidate object C^t and the object O^{t-1}.

The second check considers the rare cases when a feature is small in size and moving fast compared to the spatial and temporal resolution of