applied sciences

MDPI

Article

Visualization of the Strain-Rate State of a Data Cloud: Analysis of the Temporal Change of an Urban Multivariate Description

Lorena Salazar-Llano [1,2] and Camilo Bayona-Roa [2,3,*]

1 Sustainability Measurement and Modeling Lab., Universitat Politecnica de Catalunya—Barcelona Tech, ESEIAAT, Campus Terrassa, Colom 1, 08222 Barcelona, Spain
2 Centro de Ingeniería Avanzada, Investigación y Desarrollo—CIAID, Bogotá 111111, Colombia
3 Departamento de Ingeniería Civil y Agrícola, Facultad de Ingeniería, Universidad Nacional de Colombia, Bogotá 111321, Colombia
* Correspondence: cabayonar@unal.edu.co

Received: 11 June 2019; Accepted: 19 July 2019; Published: 22 July 2019

check for updates

Abstract: One challenging problem is the representation of three-dimensional datasets that vary with time. These datasets can be thought of as a cloud of points that gradually deforms. However, point-wise variations lack information about the overall deformation pattern, and, more importantly, about the extreme deformation locations inside the cloud. This present article applies a technique in computational mechanics to derive the strain-rate state of a time-dependent and three-dimensional data distribution, by which one can characterize its main trends of shift. Indeed, the tensorial analysis methodology is able to determine the global deformation rates in the entire dataset. With the use of this technique, one can characterize the significant fluctuations in a reduced multivariate description of an urban system and identify the possible causes of those changes: calculating the strain-rate state of a PCA-based multivariate description of an urban system, we are able to describe the clustering and divergence patterns between the districts of a city and to characterize the temporal rate in which those variations happen.

Keywords: dimensionality reduction; pattern identification; three-dimensional data-cloud; strain-rate; Finite Element Method (FEM); trajectory visualization; deformation analysis

1. Introduction

One challenging problem in data analysis is the representation of three-dimensional discrete data [1]. This analysis becomes harder when a time-dependent change of the data takes place, introducing a new temporal dimension. In the present article, we explore formal approaches to quantify the temporal change of discrete three-dimensional data. Specifically, we build a methodology to assess the transformation of a data cloud that is derived from a *Principal Component Analysis* (PCA): a 13-year multivariate description in [2] that provides a reduced description of an urban system given only by the first three principal components. Since the points represent an abstraction of an urban system, one main goal is to understand the temporal variation of the multivariate description of the districts in order to analyze the behavior of the overall city in the time-span. Our main hypothesis is that these three-dimensional datasets can be thought of as a cloud of points that gradually deforms.

However, the challenging issue is that deformation between consecutive times cannot be visualized accurately. There are some methods to overcome this difficulty. One is the vector plot of three-dimensional displacements or velocities, which is typically used to visualize results in computational mechanics applications [3–5]. For example, these methods are used in the kinematic visualization of motion [6–10], but they are restricted to display the relative motion among the data, and are not able to identify the most dynamic regions of the dataset. Another is the parallel coordinate technique that successfully exhibits the temporal change of highly-dimensional statistical and information datasets [11]. However, multi-dimensional data are mostly segmented into two-dimensional subsets, which are easier to deal with, Similar to the computer tomographic scans of medical imaging [1,12,13]. However, none of these methods are suitable for understanding the patterns of diversification or conformation, which are closely related to the temporal change of the differences between join data values: the maximum and minimum magnitudes of variation and the evaluation of their direction can be significantly helpful to identify differentiation patterns in the data [14]. Conversely, they can be beneficial for locating uniformity in a dataset that was previously differentiated.

One common approach to understanding the temporal change of the dataset is to use *confidence ellipses* in dimensionally-reduced representations (i.e., PCA) [15–17]. This type of approach has been applied to several time-dependent problems (e.g., [18–21]). Confidence ellipses quantify the dispersion of the results: the main directions and intervals of variance accounting for the spatiotemporal distribution are calculated first, and then, the ellipses (or ellipsoids in 3D) are drawn using those intervals as the main axes. A temporal analysis can be achieved by evaluating the geometrical characteristics of the ellipses, especially their area, to which the change of the distribution can be associated. Their main drawback lies in the failure to locate the most dynamic areas in the dataset [22].

The field of continuum mechanics provides a measure of the temporal variation of the distance in between points: *the strain-rate tensor* (see, for instance, [23]). The continuum mechanics theory—which arises from the classical Newtonian mechanics—analyzes the causes and effects of motion for a deformable media composed by an infinite group of particles. When a continuous media is being deformed in various directions at different rates, the strain-rate of a certain position in the medium cannot be expressed by a scalar value solely. It cannot even be expressed by using a single vector. Instead, the rate of deformation must be expressed by the rank-two strain-rate tensor with its components determined by the positional derivatives along each spatial dimension. Hence, the mathematical framework of tensors can determine exactly the deformation that is accumulated in a location inside the medium—which is typically subject to the imposition of displacements or loads. This tensor is commonly used to detail the amount of elastic energy in the physical descriptions of multiple materials, such as solids or fluids (see [24] for a complete mathematical exposition). Most of those models are formulated as the product of a constitutive tensor and the strain-rate tensor, giving the stress condition of the material that is balanced in the kinetic equations. In the present study, the calculation of the strain-rate tensor is not related to the kinetics of any material, and, thus, it can only be a mathematical tool that supports the examination of the deformation rates in the discrete statistical data.

However, the strain-rate tensor arises from the continuum assumption, and discrete displacements of points rather than continuous distributions take place in the variation of the data cloud. Typically, the issue of applying derivatives to discrete displacements of points is solved by several different approaches. Few statistical techniques use co-variance functions to represent directly the strain-rate field (see, e.g., [25]). The common approach is to compute a continuous version of the displacement—or velocity—field, so that the derivatives can be applied to the continuous displacements. Some methods in this line have interpolated the discrete displacements by minimizing the residual—or distance—between the continuous interpolation and the discrete version [26]. Other interpolation techniques weight the distance between an interpolated piece-wise continuous field and the discrete displacement field (see, e.g., [27]). This method results in a minimization technique where a continuous strain-rate field can be derived. For example, the piece-wise continuous field can be defined as being splines or the widely used linear polynomials in variational formulations [28]. These techniques have been applied in Earth science and medical imaging works [29–31], as well as in the strain-rate calculation of geodetic observations [32,33].

Another fundamental issue is the representation of the strain-rate state. One of the possible techniques that can help to visualize the deformation rate of the dataset is to plot the main components of the tensor using *strain-rate diagrams*, where concentrations of strain-rate patterns can be displayed as vector fields (see, for example, the ones in geodetical observations of the Earth's mantle [34–36]). The main drawback of strain-rate diagrams is that the strain-rate components are visualized as the projection of three-dimensional vector fields into the two-dimensional framework, and, therefore, the third-dimension component has to be necessarily neglected. Another method, more suitable to two-dimensional plots, is the contour graph of principal stresses, where the stress patterns in structural elements [37,38] and tectonics [39] are visualized using continuous lines that depend on the stress magnitude. That method overcomes the three-dimensional issue, but it does not give insights about the orientation of the principal stresses. A dual form of the contour graph is to calculate the family of curves that are instantaneously tangent to the *extension* and *contraction* components of the strain-rate tensor: the so-called *trajectory* curves in the continuum mechanics field [23]. The trajectories of each principal component of the stress can be depicted in a separated plot with the stress magnitude colored along the trajectory line such that stress patterns are completely shown in a two-dimensional framework.

Since a robust methodology that describes the temporal change of the urban system—represented by a multivariate dataset—has not been carried out before, we perform a quantitative analysis by including the strain-rate tensor as the fundamental metric. In this work, we calculate the strain-rate state of the discrete dataset without a priori assuming the mechanisms by which the system experiences transformation. To apply the continuum mechanics principles into the discrete dataset, we use interpolation methods, such as the ones applied in discrete variational formulations (i.e., *Finite Element Methods* (FEM), Particle Methods, Collocation Methods, Mesh-less methods, etc.). Specifically, we derive the three-dimensional strain-rate tensor from a FEM interpolation of the discrete velocity field, as introduced in [40–42]. We include the trajectory lines of the principal strain-rate components as the methodology that can be used in a computational (two-dimensional) framework for visualizing the main patterns of change in any time-dependent data cloud. This approach overcomes the three-dimensional representation problem by separating the strain-rate state of the cloud in several two-dimensional plots—one for each principal component of the strain-rate—which exhibit the magnitudes and orientations of the deformation patterns.

The remaining parts of this document are organized as follows. In Section 2, the methodology to compute the discrete version of the strain-rate tensor is presented. Since the main problem involves the calculation of the derivatives of discontinuous—discrete—velocities and the visualization of the strain-rate patterns, we extensively detail the numerical techniques that are adopted in the present approach. Next, in Section 3, we present the application of the methodology to the case study—the urban system of Barcelona—by visualizing its main strain-rate patterns, meaning the city's environmental, social and economic change. In Section 4, some conclusions of the proposed methodology close this article.

2. Methods

We begin this section with a review of the strain-rate tensor calculation provided a discrete three-dimensional data cloud. For doing so, the formal problem of the time-dependent dataset is introduced first. Then, we explain the numerical techniques that transform the discrete dataset into a mathematical framework by which the strain-rate tensor can be computed. Most of the ideas rely on the geometrical analysis of the discrete dataset by computing the spatial discretization of the dataset into geometric elements through a *Delaunay Triangulation*. After doing that, the computation of the strain-rate is performed with a FEM interpolation of the velocity field. Finally, we address the *eigen-problem* for the strain-rate tensor, such that the solution of the eigenvalues/eigenvectors gives the extrema strain-rates at each geometric element. The flow chart diagram of this methodology is abstracted in Figure 1, including the main outcomes of each step. The extended explanation of the methodology is developed along the present section.

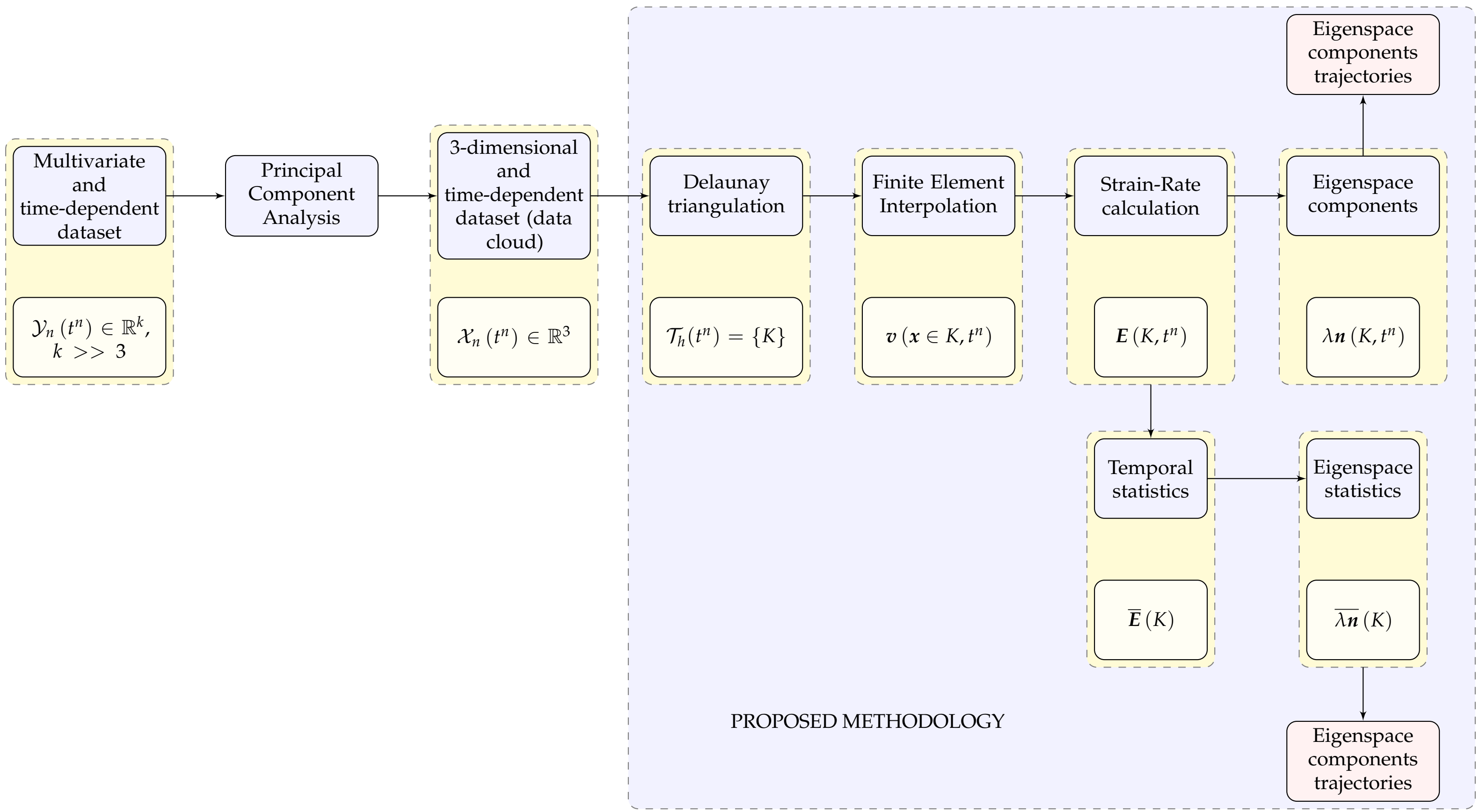

Figure 1. Flow chart of the visualization methodology.

2.1. Time-Dependent Three-Dimensional Dataset

Since the main objective of this work is to reveal the temporal transformation of a three-dimensional and time-dependent dataset, let us first introduce some notation in order to clarify the mathematical ideas to be used. Let us define the discrete time-dependent data to be the set of points $\mathcal{P} = \{p_i\}$, with $i = 1, 2, \ldots, m$, being m the total data. The values in each one of the three dimensions can be seen as scalar coefficients for a set of basis vectors. This tuple of components compose the vector that we call the *position* or *coordinate* $\boldsymbol{x}_i = [x_{i,1} \quad x_{i,2} \quad x_{i,3}]^\top$, with the superscript $\top$ denoting the transpose operation, the first subscript referring to the point i and the second to the dimension. Hence, we call $\mathcal{X}_n(t) = \{\boldsymbol{x}_1(t), \boldsymbol{x}_2(t), \ldots, \boldsymbol{x}_i(t), \ldots, \boldsymbol{x}_n(t)\} \in \mathbb{R}^3$ the *positions* of the points in a certain time t. Let us consider a uniform partition of the time span $t \in [t^d, t^g]$ in a sequence of discrete time-steps $t^d = t^0 < t^1 < \ldots < t^n < \ldots < t^N = t^g$, with $\delta t > 0$ the time-step-size defining $t^{n+1} = t^n + \delta t$ for $n = 0, 1, 2, ..., N$. Thereby, we use the superscripts to denote the discrete time-steps, with the only exception of denoting the transpose operation with the superscript $\top$.

Since the time-dependent dataset of the case study comes from a PCA reduction of a higher-dimensional multivariate dataset $\mathcal{Y}_n(t) = \{\boldsymbol{y}_1(t), \boldsymbol{y}_2(t), \ldots, \boldsymbol{y}_i(t), \ldots, \boldsymbol{y}_n(t)\} \in \mathbb{R}^k$, with $k >> 3$ and $t \in [t^d, t^g]$, into a lower-dimensional one $\mathcal{X}_n(t)$, $t \in [t^d, t^g]$ that possesses only three independent dimensions, namely Principal Component 1 (PC1), Principal Component 2 (PC2), and Principal Component 3 (PC3), we use the Cartesian coordinate system straightforwardly with each principal component being a dimension. That is, $\boldsymbol{x}_i(t^n) = [x_{i,\mathrm{PC1}}(t^n) \quad x_{i,\mathrm{PC2}}(t^n) \quad x_{i,\mathrm{PC3}}(t^n)]^\top$. Hence, the discrete time-dependent dataset can be thought of as a cloud of points in the three-dimensional space that deforms gradually over time.

2.2. Finite Element Method Interpolation

The main idea of the present approach is to transform the discrete cloud of points into a mathematical framework—similar to a deformable medium—by which the strain-rate tensor can be computed. To do so, we generate a mesh $\mathcal{T}_h(t) = \{K\}$ from the set of points $\mathcal{P}$ that is composed of non-overlapping and conforming geometrical elements K of diameter h. There are several methods to generate a mesh from a set of points, all which are studied in the computational geometry field [43]. Here, we apply Delaunay triangulation $DT(\mathcal{P})$ for several reasons: (1) the aspect ratio of the triangulated elements produce a high-quality mesh; and (2) fast Delaunay triangulation algorithms have been developed recently (see, for example, the one in [44]).

The result of applying the Delaunay triangulation over the set of points is a discrete mesh $\mathcal{T}_h := DT(\mathcal{P})$ that possess the following characteristics: it covers exactly the convex hull Ω of the point set, no point p_i is isolated from the triangulation, and all the elements $\{K\}$ are four-point tetrahedrons, which are completely defined by the position of their four corner points $K := \{\boldsymbol{x}_j\}$, with $j = 1, 2, 3, 4$. The generated mesh $\mathcal{T}_h = DT(\mathcal{P})$ can be seen as a—material—domain Ω that suffers deformations from the displacements of the points between consecutive time-steps. Since only discrete displacements between consecutive time-steps are known for the set of points, we now explain how the continuous velocity field inside the mesh is calculated.

Even though the FEM has been used to perform interpolation using the point-wise data (see [41,42] for applications in one and two-dimensional data), in this work, we apply this well-known method in a three-dimensional setting. In FEM, the finite interpolating space $\boldsymbol{\mathcal{V}}_h$ is defined as made of continuous piece-wise polynomials $N(\boldsymbol{x})$ in the mesh $\mathcal{T}_h$, where the discrete approximation $\boldsymbol{F}_h(\boldsymbol{x}, t) \in \boldsymbol{\mathcal{V}}_h$ of any multi-dimensional function $\boldsymbol{F}(\boldsymbol{x}, t)$, $\boldsymbol{x} \in \Omega$, can be written as

$$\boldsymbol{F}(\boldsymbol{x}, t) \approx \boldsymbol{F}_h(\boldsymbol{x}, t) := \sum_{i=1}^{n} N(\boldsymbol{x}_i) \boldsymbol{F}_i(t), \qquad \boldsymbol{x} \in \Omega. \tag{1}$$