Computers, Environment and Urban Systems 61 (2017) 1–12

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

Contextual mapping: Visualization of high-dimensional spatial patterns in a single geo-map

Vahid Moosavi

Chair for Computer Aided Architectural Design, Institute of Technology in Architecture, Faculty of Architecture, ETH Zurich, HPZ F/John-von-Neumann-Weg 9, CH-8093 Zurich, Switzerland

A R T I C L E I N F O

Article history:
Received 6 April 2016
Received in revised form 17 August 2016
Accepted 17 August 2016
Available online xxxx

Keywords:
Multivariate geo-mapping
Spatial clustering
Pattern recognition
Self-organizing map
Urban planning

A B S T R A C T

In this study, we proposed a generic methodology for combining high-dimensional spatial data to identify and visualize the hidden spatial patterns in a single-layer geo-map. By using the less explored one-dimensional self-organizing maps, we showed how the high-dimensional data can be transformed into a spectrum of one-dimensional ordered numbers. These numbers (codes) can index a high-dimensional space with the important property that similar indices refer to similar high-dimensional contexts. Thus, the high-dimensional vectors will be attributed to single numbers, and this one-dimensional output can be easily rendered as a new single data layer in the original geographic map. As a result, it simultaneously identifies the main spatial clusters and visualizes the high-dimensional correlations (if any) in a single geographic map. Further, because the output of the proposed method is a set of ordered indices, there is no need to define a fixed number of clusters in advance. Because these composite spatial layers are identified on the basis of the selected context (i.e., the selected features or aspects of the spatial phenomena), they are called *contextual maps*.
Finally, we showed the results of applying the proposed methodology to several synthetic and real-world data sets.

1. Introduction

With the current rapid growth in the amount of digital data, we must address the challenge of finding appropriate techniques to harness the power of these data streams. For example, in many cities across the world, no longer does anyone lack access to digital spatial maps; instead, the current challenge is, considering the amount and diversity of these digital data regarding different aspects of the cities, how one can picture his/her own map of the space as a combination of several factors of interest.

Toward this direction, there have been several interesting cases such as peoplemaps[1] or Livehood projects (Cranshaw, Schwartz, Hong, & Sadeh, 2012), which are explorations and mapping of activities within cities based on data available from online social networks. One of the cases most similar to our work is a project called Whereabout,[2] where by applying the K-means data-clustering algorithm to a collection of spatial data consisting of >200 different aspects of each ward in the city of London, a fixed number of groups were created by grouping based on informational similarities (not physical locations). Then, on top of the classical map of London, people get an impression of different regions on the basis of their similarities in all of these categories. In a similar manner, but only based on demographic information, a new coding system of London called LOAC was developed (Longley & Singleton, 2014).

The classical clustering algorithms divide the high-dimensional data space into a predetermined number of groups, where each will be given a label (usually an arbitrary number). Then, these cluster labels attributed to each spatial data point can be visualized on the geographic map with a specified color code. However, despite the fact that standard clustering methods such as K-means are easy to use, they have some limitations in the domain of spatial pattern recognition. One of the main problems is that they divide the space into a small number of categories. Instead, it would be preferred to have a continuous and smooth changing pattern on top of the high-dimensional data. Further, one needs to select the number of clusters in advance, which is a critical decision (Tibshirani, Walther, & Hastie, 2001). In addition, in the context of spatial clustering, because the cluster labels are not ordered according to their high-dimensional similarities, the colored visualization of clusters in the geographic map is not directly helpful. Therefore, similar colors in a clustered geo-map do not necessarily refer to similar high-dimensional patterns. As a result, increasing the number of clusters with different colors may result in final spatial visualizations that are not helpful, but having too few clusters produces results that are too aggregated. One current solution to this problem is to create an RGB (red, green, blue) pattern after data clustering by reducing the high-dimensional vectors of the cluster centers to their first three principal components (Mahinthakumar, Hoffman, Hargrove, & Karonis, 1999). However, in

E-mail address: svm@arch.ethz.ch.

[1] http://peoplemaps.org/.
[2] http://whereaboutslondon.org/#/about.

http://dx.doi.org/10.1016/j.compenvurbsys.2016.08.005
0198-9715/

addition to losing some information (by selecting only three principal components), the color interpretations will need an additional step.

The main hypothesis of this study is that if we find a method to sort the clusters in a way such that similar cluster indices refer to similar contexts (i.e., similar high-dimensional patterns), we can make a direct projection from high-dimensional spatial data to a one-dimensional vector and visualize the high-dimensional patterns in the geographical maps using a simple color spectrum. In this manner, by having many indices instead of dividing the high-dimensional data into a few distinct groups, one can create a spectrum of high-dimensional patterns that are visualized with a colored spectrum on spatial maps. Because the high-dimensional patterns would change gradually, this would also solve the problem of distinct cluster borders and the fixed number of clusters. As we show in Section 2, our proposed approach can be discussed from the viewpoint of dimensionality reduction and manifold learning (Bengio, Courville, & Vincent, 2013), where one of the best methods that satisfies these requirements is self-organizing maps (SOMs) (Kohonen, 2013).

2. SOMs in the domain of spatial analysis

SOM is a general-purpose machine-learning method that offers interesting solutions to different data-driven modeling tasks (Kohonen, 2013).

SOM is a nonlinear space transformation method that tries to preserve the topology of high-dimensional data, while transforming them into a low-dimensional space. This means that SOM projects the high-dimensional data points to a lower-dimensional space (normally a two-dimensional grid) in a manner such that neighboring objects in high-dimensional space remain neighbors in low-dimensional space. This topology-preserving transformation unfolds the nonlinear and high-dimensional patterns into a low-dimensional space that can be easily visualized.

Nevertheless, a major difference between SOM and other data-unfolding and dimensionality reduction methods such as locally linear embedding (Roweis & Saul, 2000), complete isometric feature mapping or ISOMAP (Tenenbaum, De Silva, & Langford, 2000) and *t*-distributed stochastic neighbor embedding, known as *t*-SNE (Van der Maaten & Hinton, 2008) is that it creates an abstraction of the data into new prototypes, while in typical dimensionality reduction methods, there is always a one-to-one relationship between all the observed points in the high- and low-dimensional space. In SOM algorithm, these identified abstract prototypes (usually called nodes or codebooks) are essential elements for the pattern recognition and data reduction tasks such as clustering. These nodes have a dual representation, including a low-dimensional vector, showing the location of the node in the lower-dimensional space, and a high-dimensional weight vector, which has the same dimensionality as the original high-dimensional data. Therefore, these nodes, as distributed models of the training data, can be used separately in different modeling problems. This property of SOM makes it very attractive for many tasks such as data visualization, function approximation, and data clustering in general. In the domain of spatial analysis, SOM has been used in several applications (Delmelle, Thill, Furuseth, & Ludden, 2013; Frenkel, Bendit, & Kaplan, 2013; Agarwal & Skupin, 2008; Skupin & Esperbé, 2011; Wang, Biggs, & Skupin, 2013 and Arribas-Bel, Nijkamp, & Scholten, 2011; Spielman & Thill, 2008) and is well-known as a tool for visual data mining and exploration of high dimensional spatial interactions (Yan & Thill, 2009).

Because data clustering is an exploratory activity, the high-dimensional maps resulting from the clustering of high-dimensional vectors using SOM are commonly visualized on two-dimensional colored maps known as component planes (see Fig. 3 for an example). However, in the context of spatial clustering, there is normally an extra constraint on projecting the final outputs of the pattern recognition algorithms onto the original spatial map. Considering this requirement, the main problem of classical SOM is that it loses the spatial index of data that are not part of the training data (Bação, Lobo, & Painho, 2005).

Therefore, one of the main concerns of spatial clustering is how to consider the effect of spatial coordinates of data points alongside the other attributes (Bação et al., 2005 and Hagenauer & Helbich, 2013).

Spatial autocorrelation is one of the underlying concepts in spatial data modeling, which states that physically nearby objects are more likely to exhibit similar properties (Tobler, 1970).

To address this issue, there are two modifications to the original SOM. The first approach is to consider the similarity of spatial objects as a weighted sum of similarity between high-dimensional attributes and their physical proximity. However, since spatial coordinates are not semantically comparable with other attributes, this approach is not widely accepted (Bação et al., 2005). The second approach leads to a class of spatial variants of SOM such as GeoSOM (Bação et al., 2005; Henriques, Bacao, & Lobo, 2012), where, the algorithm forces the training data that spatially similar observations are placed in similar regions of the low-dimensional map of SOM. Therefore, spatial coordinates and spatial attributes are contributing next to each other, but not at the same time in one single distance measure.

A more recent method in this approach is contextual neural gas (CNG) (Hagenauer & Helbich, 2013), which is based on similar idea to GeoSOM, but implemented in the context of the neural gas (NG) algorithm (Martinetz & Schulten, 1991). The NG algorithm is a modified version of SOM that unlike SOM there is no defined low dimensional grid and the nodes are dynamically distributed in the high-dimensional input space. In the case of CNG, the geographical map is used as the lower-dimensional representation.

The main contribution of these two spatial variants of SOM (i.e., GeoSOM and CNG) is that, to an extent, they replace the original synthetic topology of SOM with a spatial map with the cost of n based method discussed above, the final two-dimensional SOM grid can be used as a bivariate color code (Guo, Gahegan, MacEachren, & Zhou, 2005). Although a two-dimensional SOM performs better than the PCA-based map coloring method, which is a linear dimensionality reduction approach, because we are dealing with high-dimensional data, we need another diagram to connect these bivariate color codes of clusters on top of the SOM grid to show the characteristics of the clusters in terms of their high-dimensional vectors. As a result, one needs to select a small number of clusters for better visualization in most of these applications.

As an alternative approach to those mentioned above, the principle idea of this study is to view the problem of spatial clustering from the perspective of manifold learning and dimensionality reduction (Bengio et al., 2013). In the context of spatial pattern recognition, this implies that if there exist some underlying spatial similarities in high-dimensional data that are not easy to track in the original spatial maps, there should be an appropriate manner of encoding the data from high-dimensional space to lower-dimensional codes (specifically to a one-dimensional vector) while preserving the patterns in the encoded data. These low-dimensional codes should index a high-dimensional state space with the important property that similar regions in the high-dimensional state space receive similar codes. Specifically, if we could encode the high-dimensional vectors as a single-dimensional code, we can treat these codes as a type of numerical value, and they can be treated as a single layer of spatial data in the same way that, for example, we can render the surface temperature in a geographic map. Therefore, if there are spatial patterns in the high-dimensional data, one can quickly see them visualized in the geographical maps. This will solve the abovementioned problems of the current spatial clustering approaches.

In addition, because we transform the high-dimensional space into single-dimensional numbers, these numbers can be seen as abstractions of those high-dimensional spaces that they refer to. Therefore, as we will show in the following sections, one can combine the results of several clustering steps in a systematic and hierarchical manner.