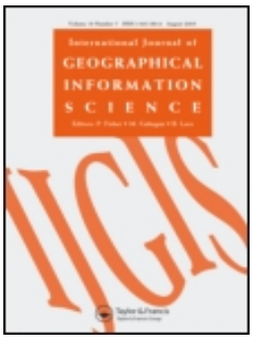

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China

Bailang Yu, Song Shu, Hongxing Liu, Wei Song, Jianping Wu, Lei Wang & Zuoqi Chen

To cite this article: Bailang Yu, Song Shu, Hongxing Liu, Wei Song, Jianping Wu, Lei Wang & Zuoqi Chen (2014) Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China, International Journal of Geographical Information Science, 28:11, 2328-2355, DOI: 10.1080/13658816.2014.922186

To link to this article: https://doi.org/10.1080/13658816.2014.922186

Published online: 27 May 2014.

International Journal of Geographical Information Science, 2014
Vol. 28, No. 11, 2328–2355, http://dx.doi.org/10.1080/13658816.2014.922186

Taylor & Francis
Taylor & Francis Group

Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China

Bailang Yu[a,b]*, Song Shu[a,c], Hongxing Liu[c], Wei Song[d], Jianping Wu[a], Lei Wang[e] and Zuoqi Chen[a]

[a]*Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai, China;* [b]*Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration, East China Normal University, Shanghai, China;* [c]*Department of Geography, University of Cincinnati, Cincinnati, OH, USA;* [d]*Department of Geography and Geosciences, University of Louisville, Louisville, KY, USA;* [e]*Department of Geography & Anthropology, Louisiana State University, Baton Rouge, LA, USA*

(*Received 8 January 2014; final version received 4 May 2014*)

Previous studies have demonstrated urban built-up areas can be derived from nighttime light satellite (DMSP-OLS) images at the national or continent scale. This paper presents a novel object-based method for detecting and characterizing urban spatial clusters from nighttime light satellite images automatically. First, urban built-up areas, derived from the regionally adaptive thresholding of DMSP-OLS nighttime light data, are represented as discrete urban objects. These urban objects are treated as basic spatial units and quantified in terms of geometric and shape attributes and their spatial relationships. Next, a spatial cluster analysis is applied to these basic urban objects to form a higher level of spatial units – urban spatial clusters. The Minimum Spanning Tree (MST) is used to represent spatial proximity relationships among urban objects. An algorithm based on competing propagation of objects is proposed to construct the MST of urban objects. Unlike previous studies, the distance between urban objects (i.e., the boundaries of urban built-up areas) is adopted to quantify the edge weight in MST. A Gestalt Theory-based method is employed to partition the MST of urban objects into urban spatial clusters. The derived urban spatial clusters are geographically delineated through mathematical morphology operation and construction of minimum convex hull. A series of landscape ecologic and statistical attributes are defined and calculated to characterize these clusters. Our method has been successfully applied to the analysis of urban landscape of China at the national level, and a series of urban clusters have been delimited and quantified.

Keywords: urban spatial clusters; object-based method; DMSP-OLS data; Gestalt theory; minimum spanning tree

1. Introduction

The rapid urbanization at global scale has transformed a large quantity of the Earth's natural land surfaces into urban landscapes at unprecedented rates, which results in the physical expansion and growth of built-up area in a single city as well as the increasing coalescence among cities. Owing to modern transportation systems, efficient communication technology, and improved infrastructure, interactions among neighboring cities become much more frequent than before, reflected by increased commercial cooperation and movements of people, goods, and information between adjacent cities. When several

*Corresponding author. Email: blyu@geo.ecnu.edu.cn

large cities are in sufficient geographical proximity, these cities and their surrounding areas would be considered as a single coherent urban complex – urban cluster. Gottmann (1957) used the term 'megalopolis' to describe the huge metropolitan area along the eastern seaboard of the US extending from Boston to Washington, DC. Other terms such as *Conurbation*, *Urban agglomeration*, *Urban Cluster*, and *Metropolitan interlocking region* have been also used in the literature to denote this large urban landscape phenomenon (Geddes 1915, Gottmann 1957, 1961, Zhou 1991). Despite the inconsistency in terminology, this type of extensive, multi-centered, and multi-city urban landscape has been well recognized. In this paper, we refer to the urban phenomena as *Urban Spatial Cluster* in order to emphasize the geographic agglomeration and linkages between individual cities. As Batten (1995) depicted, urban spatial cluster has a polycentric configuration, in which cities with complementary functions strive to cooperate and achieve significant scale economies, facilitated by fast and reliable corridors of transportation. Many researchers attempted to explain the formation of urban spatial clusters from the perspective of geographical conditions, economic bases, policy background, traffic conditions, and human capital (Alonso-Villar 2002). A better understanding of this kind of large urban phenomenon is critical for intelligent management and governance of these cities (Vogel *et al.* 2010).

The spatial pattern and structure of an urban spatial cluster evolves since its emergence (Gottmann 1957, Friedmann 1973, Zhu 2004, Chen *et al.* 2006, Ma *et al.* 2008, Fang *et al.* 2010). Gu *et al*. (1999) identified two types of urban spatial clusters in China based on spatial shape: blocky and linear shapes. The blocky urban spatial clusters tend to be located in the well-developed regions, while linear-shaped clusters are always associated with convenient transportation corridor. Zhu (2004) identified urban agglomerations with various shapes in China, including '>', 'Δ', '^', 'H', and '*Φ*' shapes. In most cases, the nodes in those shapes are the leading cities in the urban agglomerations. Some metrics, such as fractal dimension and compactness, were also used to characterize urban spatial clusters (Chen *et al.* 2003, Chen *et al.* 2006, Fang *et al.* 2008, Ma *et al.* 2008, Fang *et al.* 2010). Fang *et al*. (2008) found a positive correlation between the spatial compactness and the degree of development of an urban agglomeration in mainland China. Examining the spatial pattern and structure of urban spatial clusters is crucial for understanding the dynamics of urban spatial clusters as well as the driving forces for their formation and expansion.

Traditionally, urban spatial clusters are defined and identified based on statistical analysis of social-economic variables, such as urban area, population size and density, and economic interaction between cities (Gottmann 1961, Zhou 1991), in which individual cities are simply represented as point features. Geographically, an urban spatial cluster is the spatial coalescence regions of cities, which can be visually perceived and delineated on satellite imagery at a large scale. The geographical proximity between cities can be a key factor affecting the interaction and integration of cities in a region, and hence can serve as a proxy variable for identifying an urban spatial cluster. For example, the Japanese government defines the area within 50 km from the central city Tokyo as Metropolitan region where there is a high degree of spatial concentration of cities (Okamoto 1997). In the 2000 US Census, cities with higher commuter rates between them are grouped into the same metropolitan area (Reschovsky 2004). Commuter rate tends to be higher between the cities with a short distance apart. The geographical proximity may thus serve as a proxy for the commuter rate if no commuting data are available.

Previous studies have demonstrated that remote-sensing technologies have been useful in urban extent delineation, urban impervious surface mapping, and urban growth monitoring. For the purpose of studying urban clusters at the national or continental scale,

satellite images with relatively coarse resolution but a broad spatial coverage are more suitable and efficient. The nighttime light images from Defense Meteorological Satellite Program – Operational Linescan System (DMSP-OLS) – have been available since 1973 (Brandli 1978). The National Oceanic and Atmospheric Administration's National Geosciences Data Center (NOAA/NGDC) has produced annual digital stable nighttime light image data set at global scale since 1992, by processing, rectifying, and aggregating daily nighttime DMSP-OLS images (Elvidge *et al.* 1997a, Elvidge *et al.* 2007). Despite its relatively coarse spatial resolution (about 1 km), the stable nighttime light images have been widely used in mapping urban settlements (Imhoff *et al.* 1997, Elvidge *et al.* 1999, Lo 2002, Sutton 2003, Small *et al.* 2005, He *et al.* 2006, Zhang *et al.* 2011, Letu *et al.* 2012, Liu *et al.* 2012), estimating population (Elvidge *et al.* 1997b, Sutton *et al.* 2001, Roychowdhury *et al.* 2011), and other socioeconomic indicators (Elvidge *et al.* 1997b, Doll *et al.* 2006, Shi *et al.* 2014b) at the national or continental scale. Lo (2002) pointed out that there was a potential for detecting urban spatial clusters besides extracting urban extents by using nighttime light image data, and he visually identified 10 most notable urban spatial clusters in China without giving a formal numerical definition and detection procedure. A more objective and quantitative method is needed to identify and characterize urban spatial clusters.

This paper presents a novel framework for deriving and quantifying urban spatial clusters using nighttime light satellite images. The contributions of this paper are summarized as follows:

- Our new object-based method provides an effective way for detecting and characterizing urban spatial clusters quantitatively.
- The minimum spanning tree (MST) is used to represent spatial proximity relations among urban objects.
- Urban spatial clusters in China at the national scale are numerically identified and analyzed.

In the following sections, we will first describe the case study area and the data sets used in the study. Then, we will present the methodological framework for object-based urban spatial cluster analysis, followed by a report and discussion of analytical results. In the last section, we will summarize key research findings and draw conclusions.

2. Study area and data sets

Our urban spatial cluster analysis covers mainland China at the national level. Hong Kong, Macao, and Taiwan are not included in this analysis due to lack of relevant statistical data. Since the economic reform and the adoption of open-door policy in 1979, China has been undergoing rapid economic development. With the relaxation of the internal population migration control, China has been experiencing a rapid urbanization, and the size and number of towns and cities have increased significantly.

The stable nighttime light image data in 2005 processed and archived by the NOAA/NGDC (http://www.ngdc.noaa.gov) are used in this study. The data values represent the light intensity from cities, towns, and other sites with persistent lighting, sensed and recorded by DMSP-OLS sensors. The annual stable nighttime light images are cloud-free composites made by using all the available archived DMSP-OLS smooth resolution data for the calendar year. Ephemeral events, such as fires, have been discarded. Pixel value of the image data ranges from 0 to 63, and 0 represents the dark background (Baugh *et al.* 2010).

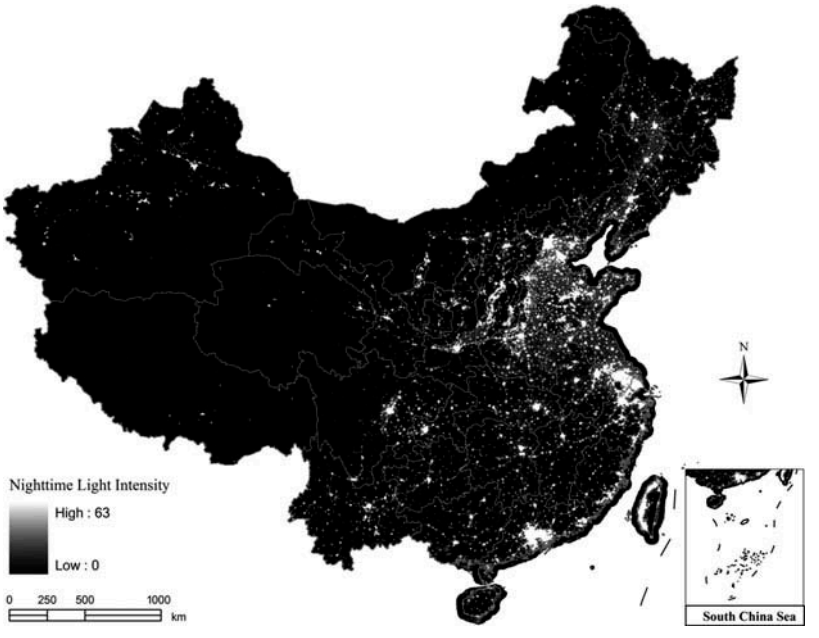

Figure 1. Stable nighttime light image of China (2005).

The nighttime light image of China in 2005 (Figure 1) was extracted from the global data set by using a mask polygon of the national boundary of China with a 50 km buffer, considering the light 'blooming' effect (Small *et al.* 2005) and the possible georeferencing error (Henderson *et al.* 2003). The image was projected into the Lambert Azimuthal Equal Area Projection referencing to WGS84 datum with the spatial resolution of 1 km.

Eleven scenes of Landsat 7 ETM+ images from US Geological Survey (http://www.usgs.gov) were processed. The Landsat images were acquired in 2005 and covered the Yangtze River Delta region. The urban areas manually interpreted from the Landsat data were used as the ground truth to estimate the accuracy of urban areas extracted from DMSP-OLS data.

Statistical data set of urban built-up areas was obtained from the China Land & Resources Almanac, which was compiled by the Ministry of Land and Resources of the People's Republic of China (2005). This ancillary data set was used for determining the threshold to extract the urban built-up area from the nighttime light image.

3. Methods

Our urban spatial cluster detection technique consists of several technical components: (1) to extract urban built-up areas as urban objects from nighttime light image and derive spatial attributes for each urban object; (2) to compute the minimum distance between urban object boundaries and represent the spatial proximity relationships between urban objects through a MST; (3) to partition the MST for identifying urban clusters; and (4) to derive spatial attributes to characterize urban clusters. The data-processing procedure and technical steps are shown in Figure 2.

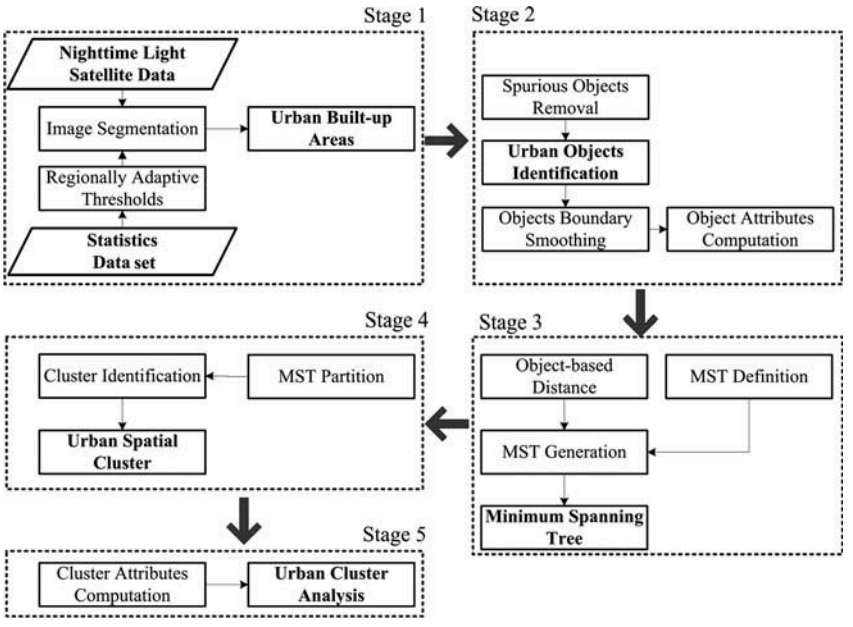

Figure 2. Flowchart of object-based urban spatial cluster analysis from DMSP-OLS stable light image.

3.1. *Delineation of urban built-up areas from nighttime light satellite images*

Since urban built-up areas in modern societies are illuminated artificially at night, their corresponding pixels in nighttime stable light images have larger digital number (DN) values than the surrounding rural areas. In previous studies, the threshold-based segmentation method had been proven to be effective and efficient to distinguish urban pixels from rural pixels on DMSP-OLS nighttime light images (Imhoff *et al.* 1997, Small *et al.* 2005, He *et al.* 2006, Liu *et al.* 2012). Pixels with a DN value equal to or larger than the threshold value DN_0 will be considered part of an urban built-up area. On the contrary, other pixels will be denoted as rural area. Namely, with a threshold value, the original stable nighttime light image is segmented into a binary image by applying Equation (1):

$$g(i,j) = \begin{cases} 1, l(i,j) \geq DN_0 \\ 0, l(i,j) < DN_0 \end{cases} \quad (1)$$

where $l(i, j)$ is the DN value of the pixel at row i and column j in stable nighttime light image, and $g(i, j)$ is the pixel value at location (i, j) in the transformed binary image. In the binary image, the DN value 1 represents urban areas and DN value 0 indicates rural or sub-urban areas.

He *et al.* (2006) compared the urban built-up areas from the government's statistical yearbook with the urban areas segmented from the nighttime light image by varying the threshold values. And the threshold value that minimizes the difference between the image-derived urban area and statistical urban built-up area is selected as the optimal value to extract urban areas from the nighttime light images. Since this study mainly focused on the urban spatial cluster analysis after detecting urban built-up areas, we

directly adopted a similar approach of He *et al.* (2006) to determine the optimal threshold value for urban object extraction from the nighttime light images.

Due to the regional difference in atmospheric condition, climate, landscape, as well as socioeconomic conditions (e.g., population density, business density, and GDP per capita), the annual DMSP-OLS nighttime light image DN values may not be comparable at regional and national scales. Therefore, no single threshold value is suitable for segmenting urban areas from DMSP-OLS nighttime light data for different regions or countries (Henderson *et al.* 2003). In this paper, we subdivided China to 10 regions (Yang 1990), and determined the optimal threshold value for each region using the approach in He *et al.* (2006). The determined optimal threshold values for these regions are listed in Table 1, and the urban built-up areas extracted from nighttime light image with these optimal threshold values are shown in Figure 3. The relative error between the image-derived

Table 1. Optimal threshold values for the extraction of urban built-up areas in different districts of China.

District	Threshold	Statistical area (km^2)	Extracted area (km^2)	Relative error (%)
Northeast China	8	31035	31741	2.27
North China	15	47259	47700	0.93
South China	19	22483	22341	−0.63
East China	22	28578	28476	−0.36
Central China	7	51260	50638	−1.21
Southwest China	7	28254	28025	−0.81
Northwest China	7	19823	20187	1.83
Inner Mongolia	6	11836	12508	5.68
Xinjiang	9	9702	9656	−0.47
Tibet	11	385	380	−1.47

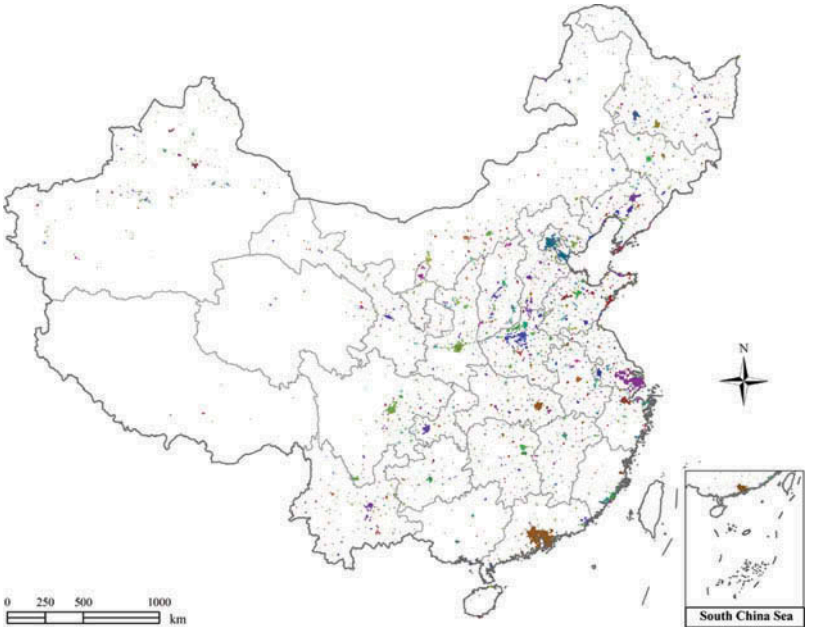

Figure 3. Urban objects extracted from the nighttime light image of China. Each object is coded in a unique color.

urban area and statistical urban built-up area is less than 5% for all regions, except for the Inner Mongolia region where the relative error is 5.68% (Table 1). The Yangtze River Delta region was selected as the case study area to estimate the accuracy of urban areas extracted from DMSP-OLS data. The urban and non-urban areas interpreted from Landsat data were regarded as the ground truth. The evaluation showed that the overall accuracy of DMSP-OLS-derived urban areas was 90.60% and the Kappa index was 0.79. The estimated accuracy values were comparable with previous studies (Liu *et al.* 2012, Yang *et al.* 2013, Shi *et al.* 2014a). We thus believe that the extracted urban areas in this study can well represent the urban landscape patterns of China.

3.2. *Identifying urban objects*

3.2.1. *Urban objects identification*

The urban built-up areas extracted from DMSP-OLS nighttime light image comprise a series of urban patches. Each urban patch consists of spatially connected urban pixels. Urban patches detected in the nighttime light image can be regarded as the generalized urban built-up areas at a spatial resolution of 1 km. Small urban patches may correspond to separate urban districts of the same city, and large urban patches may cover the entire city or several cities in a metropolitan area. In this study, an urban patch is treated as the basic spatial unit in the analysis and referred to as an urban object. We adopted an object-based approach to explicitly represent and characterize urban patches as urban objects. The object-based approach is able to extract semantic information (size, shape, texture, and contextual) of urban objects and the spatial relationships (e.g., proximity relationship) between urban objects.

A recursive connected-region labeling algorithm (Liu *et al.* 2004, Yu *et al.* 2010) was employed to explicitly mark out the urban objects based on the spatial 4-connectivity of the foreground urban pixels. The derived urban objects are each indexed incrementally with a unique ID number. Two types of morphological operations, a filling operation and a closing operation, are applied to the urban objects to smooth urban object boundaries. After that, a threshold value of the object size is selected to remove small spurious objects, which might be induced by data noise. In total, 2052 urban objects were detected in China finally (Figure 3).

3.2.2. *Attributes of urban objects*

The size, shape, and other geometric attributes of urban objects can be used to analyze urbanization progress and urban spatial clusters. We calculate planimetric attributes and shape attributes for characterizing each urban object. The planimetric attributes depict the geographical location, horizontal dimensions, and size of an urban object, including centroid location ($\bar{x}$, $\bar{y}$), perimeter (*P*), area (*S*), length (*LEN*), and width (*WID*). The shape attributes for characterizing the planar geometric morphology of urban objects include compactness index (*CI*), elongatedness (*ELG*), orientation (Φ), rectangularity (*REC*), ellipticity (*ELP*), and triangularity (*TRI*). The definitions of those spatial attributes for urban objects are shown in Table 2. More detailed definitions are also available in Liu *et al.* (2010). Minimum bounding rectangle and best-fit ellipse are also fitted to graphically represent the shape of each urban object (patch) (Liu *et al.* 2010).

The 10 largest urban objects by size are shown in Figure 4, with their minimum bounding rectangles and best-fit ellipses. The locations of main cities are overlaid with the urban objects. Table 3 shows the attributes of the 10 largest urban objects. It should be

Table 2. Definitions of spatial attributes for an urban object.

Attributes	Definition
Centroid location $(\bar{x}, \bar{y})$,	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ n is the number of cells consisting of an urban object, x_i and y_i are the row and column coordinates of the ith cell of the object.
Perimeter (P)	$P = m_1 CS + \sqrt{2} m_2 CS$ m_1 is the number of boundary cell in horizontal or vertical orientation, m_2 is the number of boundary cell in diagonal step, and CS is the grid cell size.
Area (S)	$S = nCS^2$
Length (LEN)	Length of the minimum bounding rectangle surrounding the urban object.
Width (WID)	Width of the minimum bounding rectangle surrounding the urban object.
Compactness index (CI)	$CI = \frac{4\pi S^2}{P}$
Elongatedness (ELG)	$ELG = \frac{LEN}{WID}$
Orientation (Φ)	$\Phi = \frac{180}{2\pi}\arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right)$ $\mu_{pq} = \sum_{i=1}^{n} (x_i - \bar{x})^p (y_i - \bar{y})^q$ μ_{pq} are the central moments, Φ is defined as an angle in degree between the x-axis and the major axis of the best-fit ellipse measured counterclockwise [0, 180°].
Rectangularity (REC)	$REC = \frac{S}{LEN} \cdot \mathrm{WID}$
Ellipticity (ELP)	$ELP = \begin{cases} 16\pi^2 \cdot AMI, \text{if } AMI \le 1/16\pi^2 \\ \frac{1}{16\pi^2 \cdot AMI}, \text{otherwise} \end{cases}$ $AMI = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4}$ AMI is the affine moment invariant.
Triangularity (TRI)	$TRI = \begin{cases} 108 \cdot AMI, \text{if } AMI \le 1/108 \\ \frac{1}{108 \cdot AMI}, \text{otherwise} \end{cases}$

Source: Liu *et al.* (2010).

noted that many large urban objects contain more than one city, thus being referred to as urban agglomerations. For example, the urban object (ID 1966) in the Pearl River Delta contains eight municipal cities, including Guangzhou, Shenzhen, Dongguan, Foshan, Zhongshan, etc. Shanghai, Suzhou, Jiaxing, Wuxi, and Changzhou are contained in the urban object (ID 1266) in the Yangtze River Delta. The third largest urban object (ID 511) encloses three municipal cities, including Beijing, Tianjing, and Langfang. Meanwhile, several major cities each form a separate urban object, such as Wuhan (urban object ID 1425) and Chongqing (urban object ID 1553).

3.3. *Representing spatial proximity relations between urban objects*

3.3.1. *Computing minimum distance between the boundaries of urban objects*

Different types of distance between cities have been used to measure their proximal relations. Conventionally, cities are often simplified as point features and their spatial locations are represented by the geographical positions of city centers, which are designated either by city halls (municipal government) or by the centroid points of municipal

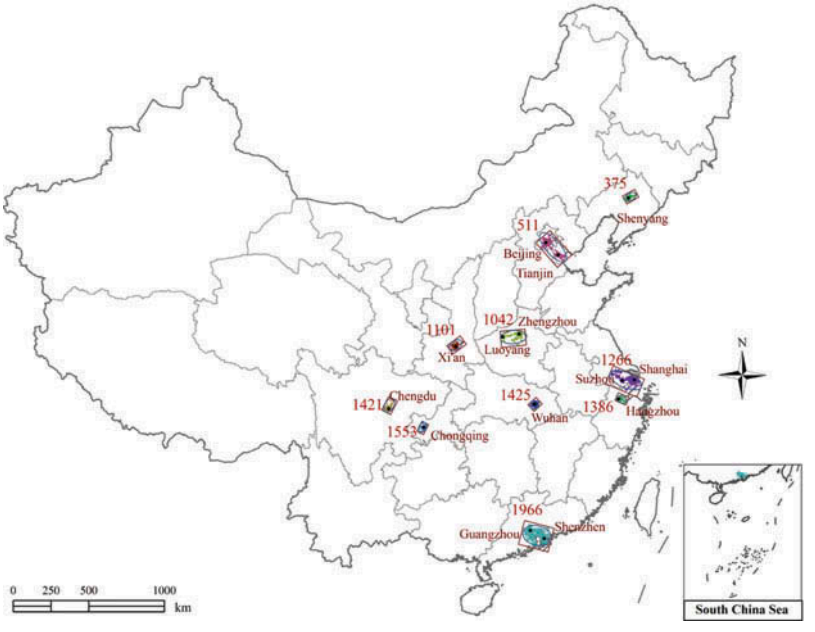

Figure 4. The 10 largest urban objects with their minimum bounding rectangles and best-fit ellipses.

administrative areas for the sake of convenience (Zhou 1991, Batten 1995, Dematteis 1997, Assuncao *et al.* 2006, Chen *et al.* 2006). The distance between the city centers was computed to quantify their spatial proximity in previous studies (Portnov *et al.* 2004, Luo *et al.* 2009, Terzi *et al.* 2011).

However, in reality, a city is a two-dimensional space with various shapes. Social, economic, and cultural activities, along with material, energy, information, and population flows occur not merely in the city center, but also across the whole built-up area. For instance, some large manufacturing plants may be located near the boundary of a metropolitan area due to low land price. Meanwhile, some residential districts of a city may house a large number of people who work in other nearby cities. The everyday journey-to-work between cities supports the operations of those manufacturing plants and many other urban economic activities. Although conceptually and computationally easy and simple, a point-based representation may greatly distort the proximal relation of polygonal objects (Okabe *et al.* 1996) and hence result in spurious and unrealistic urban clusters.

As shown in Figure 5(a), there are seven hypothetical urban objects (polygons) in the region. If the polygonal urban objects are simplified as point features and represented by their centroid points (the stars in Figure 5(b)), four small spatial clusters of points are identified using the distance between points (Figure 5(b)). This clustering result is very different from that visually grouped based on the proximity of polygonal urban objects (Figure 5(a)). To obtain reliable spatial clustering result, it is critical that urban objects are treated as polygonal features rather than as point features and the proximity between urban objects should be measured by the distance between their boundaries rather than by the distance between their centroid points.

The MST in graph theory is used in this study to represent spatial proximity relationships between urban objects. An edge-weighted linear graph (G) is an ordered pair $G = (V, E)$

Table 3. Spatial attributes of the 10 largest urban objects in the order of area.

Rank	Object ID	S (km^2)	P (km)	$\bar{x}$	$\bar{y}$	*LEN* (km)	*WID* (km)	*CI*	*ELG*	Φ	*REC*	*ELP*	*TRI*
1	1966	14480	1935.82	891843.36	2412733.23	200.08	132.11	0.05	1.51	14.65	0.55	0.56	0.82
2	1266	9795	1879.79	1489132.44	3441909.08	237.52	103.02	0.03	2.31	23.28	0.40	0.30	0.44
3	511	8527	1740.69	994771.00	4317446.15	254.30	99.32	0.04	2.56	48.76	0.34	0.21	0.31
4	1042	5600	1595.80	742744.83	3725737.47	181.06	90.67	0.03	2.00	174.74	0.34	0.22	0.32
5	1101	3207	678.12	355598.42	3664274.26	96.69	61.44	0.09	1.57	143.91	0.54	0.54	0.79
6	1421	3180	554.47	−84135.57	3259794.64	106.22	61.69	0.13	1.72	117.14	0.49	0.44	0.64
7	375	2387	440.49	1511008.19	4647094.61	86.91	49.27	0.15	1.76	150.77	0.56	0.58	0.85
8	1425	2301	378.66	880664.19	3275034.16	64.71	54.23	0.20	1.19	137.54	0.66	0.80	0.85
9	1553	2096	415.37	141916.71	3123371.24	72.11	50.64	0.15	1.42	117.41	0.57	0.61	0.90
10	1386	2063	417.30	1452577.86	3309217.59	77.94	45.51	0.15	1.71	23.68	0.58	0.63	0.92

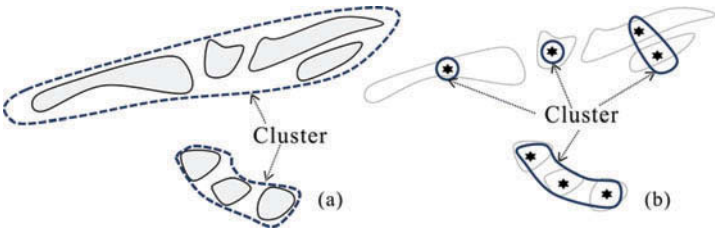

Figure 5. Object clustering based on different criteria: (a) clusters based on object boundaries and (b) clusters based on distance between object centroids.

comprising a set of vertices called *nodes* (V) and a set of node pairs called *edges* (E). In our case, a node represents an urban object. An edge E_{mn} between urban objects O_m and O_n is associated with a *weight* (W_{mn}), which is the minimum distance between the boundaries of those objects (Equation (2)). A *path* in a graph is a sequence of edges joining two nodes (objects) and a *circuit* is a path where the first and the final nodes are the same. A *tree* is a connected and undirected graph with no circuits. A *spanning tree* is a tree in which all the nodes of G are connected together by edges. A MST is a spanning tree whose total weight of all edges is the smallest among all possible spanning trees generated from the nodes.

$$W_{mn} = min\{d_{mn}\} \tag{2}$$

where d_{mn} is the distance between the pixels located at the boundaries of urban objects O_m and O_n.

Each edge in the MST of urban objects links one urban object with its nearest urban object, and the weight of the edge represents the shortest distance between the boundaries of the two nearest urban objects. The MST, with the minimum weight among all possible spanning trees, represents the nearest spatial proximity of all urban objects in the whole region.

3.3.2. *Generating MST to represent proximal relations among urban objects*

Different algorithms (Pettie *et al.* 2002, Assuncao *et al.* 2006) have been proposed to generate MST for point features. As illustrated above, it is critical to use the distance between objects' boundaries to generate and recognize clusters of urban objects. The computation of the distance between objects' boundaries and associated MST is much more complicated than the creation of the MST directly based on the distance between objects' centroid points. Since there is no computational tool readily available, we developed a new algorithm to calculate the minimum distance between urban objects and create the MST simultaneously.

Our new algorithm is based on a recursive competitive propagation scheme. The d_∞-propagation algorithm proposed by Eggers (1998), a type of distance transformation method, is used to expand the border pixels of each object step by step to calculate the minimum distance between a pair of urban objects. The algorithm starts with detecting and recording the inner boundary B_k for each urban object O_k in the region using an inner boundary tracing algorithm. Then, the urban object whose ID is 0 (O_0) is regarded as the first node and included into the MST. The pixels in the boundary of urban object O_0 are propagated outwards recursively using the d_∞-propagation algorithm (Eggers 1998). The propagation will pause when encountering the first other urban object s (O_s). The urban object s (O_s) is considered a new node of the MST, and O_s as well as the edge between

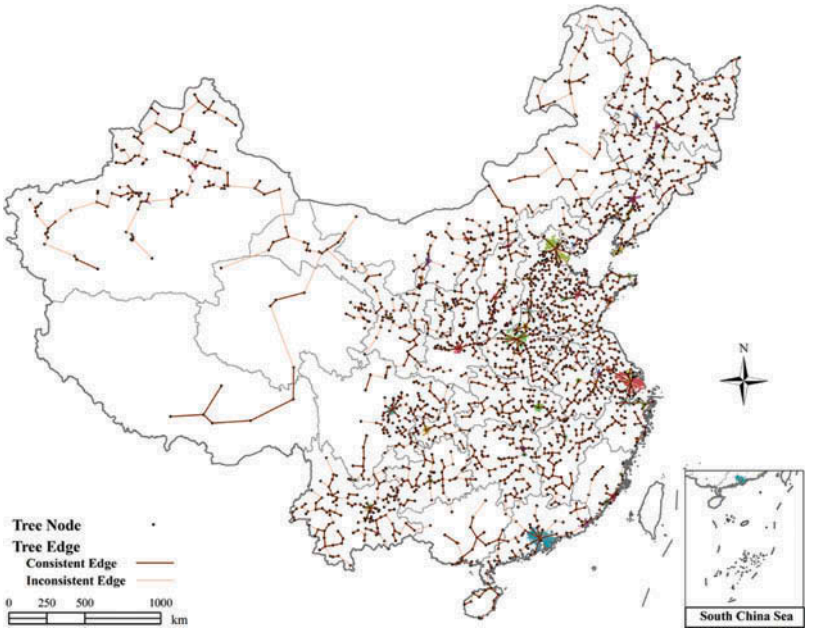

Figure 6. The generated and partitioned minimum spanning tree for urban objects based on the minimum distance between their boundaries.

urban objects O_0 and O_s are included into the MST. The distance between O_0 and O_s is recorded as the propagated distance of O_0. Then, the objects with minimum propagated distance are selected from those objects (nodes) already included in the MST to propagate simultaneously. The propagated distance values of objects are updated immediately after propagation. If the propagation of any propagating object meets first urban object not included in the MST, the object and the corresponding pair of objects (edge) will be added to the MST. The recursive competitive propagation will stop when all the objects are included in the MST.

The constructed MST for the urban objects of mainland China is shown in Figure 6.

3.4. *Identifying urban object clusters by partitioning MST*

3.4.1. *MST partition based on the Gestalt theory*

Spatial clustering can be realized by partitioning the MST appropriately into different subtrees. Each subtree stands for a targeted spatial cluster. The only principle for the partitioning process is to find the tree edge with a significantly heavy weight that can distinguish itself from other subtree edges.

As Wertheimer (1958) enunciated, proximity is the most rudimentary element of all Gestalt principles that govern the way of our perceptual process in organizing the raw sensory data of various spatial patterns presented to our eyes. Zahn (1971) developed an effective graph-theory method for detecting and describing Gestalt clusters in an arbitrary point set. Based on MST, his method detects inherent spatial separations (inherent spatial distance) between subsets (clusters) of a given point set and then extracts the grouping patterns that are compatible with our own visual perception of the two-dimensional point sets (Zahn 1971).

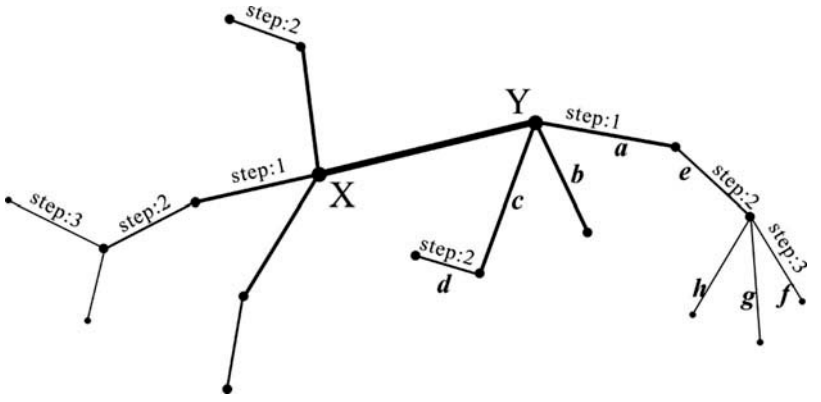

Figure 7. Definition of nearby edges of the node in MST.

We adapted Zahn's (1971) algorithm to handle polygon type of features, urban objects, in contrast to the point data set. We formulated the urban object clustering problem into an MST partitioning problem. In the MST, the weight for an edge E_{xy} is represented by the minimum distance between urban object O_x and O_y. Following the Gestalt principle, we adopt a similar strategy as in Zahn (1971) to partition the MST. The tree partition is realized by removing edges satisfying certain conditions in an MST. Removing an edge will divide the tree into two disjoint non-empty subtrees. After removing all the edges satisfying certain conditions, the resulting subtrees with more than 2 nodes (objects) can be considered as spatial clusters.

For an edge E_{xy} in a tree, if its weight W_{xy} is significantly larger than the average weight of nearby edges, it is labeled as *inconsistent* edge and will be removed. Similar to Zahn (1971), we use two indicators to define an inconsistent edge to be cut off. One indicator is the ratio of the weight W_{xy} for edge E_{xy} to the average weight of its nearby edges. The other is the ratio of the weight difference between W_{xy} for edge E_{xy} and the average weight of its nearby edges to the standard deviation of those edges' weights. When one of these two ratios is larger than a threshold value, the edge is considered to be inconsistent and then cut off from the tree. The standard deviation and average of weights for each node object (O_x or O_y) in edge E_{xy} are calculated based on a sample of nearby edges, which are the ones within a specific step length to that node. In our study, the step length is 2. As shown in Figure 7, edges *a, b*, and *c* are within 1 step to node O_y whereas edges *e* and *d* are within 2 steps to node O_y.

The inconsistency of an edge is evaluated based on the following rules in our algorithm:

For edge E_{xy},
If the $W_{xy} > W_0$ Then
  If $W_{xy} > A_0 \cdot \overline{W_x(ST)}$ Or $W_{xy} > A_0 \cdot \overline{W_y(ST)}$ Then
    the edge *XY* is *inconsistent*
  Else If $W_{xy} > \overline{W_x(ST)} + A_1 \cdot \mathrm{STD}(W_x)$ Or $W(XY) > \overline{W_y(ST)} + A_1 \cdot \mathrm{STD}(W_y)$ Then
    the edge E_{xy} is *inconsistent*
  Else
    the edge E_{xy} is *consistent*
Else
  the edge E_{xy} is *consistent*

W_0 is a user-defined threshold weight of the edge. As demonstrated by Portnov (2004), clear similarities can be detected between the core city and neighboring towns within 20–40 km distance range. However, Portnov's distance is calculated between the geometric centers of two cities. As we use the distance between two city's boundaries, the threshold value is set as 15 km, i.e., W_0 = 15 km. If the minimum distance between two cities is less than 15 km, the edge connecting them will be directly assigned as a consistent edge. $\overline{\boldsymbol{W_x}(\boldsymbol{ST})}$ and $\overline{\boldsymbol{W_y}(\boldsymbol{ST})}$ are the average weights for nodes O_x and O_y, respectively, in edge E_{xy} based on the sampled edges. STD(W_x) and STD(W_y) are the corresponding standard deviations of weights for nodes O_x and O_y in edge E_{xy}. The sampled edges consist of all the edges that have ST step length to the node. A_0 and A_1 are the ratios of the average and standard deviation. The values of ST, A_0, and A_1 are set to 2, 2, and 3, respectively.

All the edges in the MST will be checked through the processes described above. Figure 6 shows the subtrees after removing the inconsistent edges.

3.4.2. *Urban spatial cluster identification*

After the MST partitioning operation, the urban objects connected by the same subtree will be considered as an urban spatial cluster. Each urban spatial cluster (subtree) is assigned a unique ID, and all the urban objects connected by the same subtree will be identified as the members of an urban cluster. In total, we detected 521 urban spatial clusters in China as shown in Figure 8.

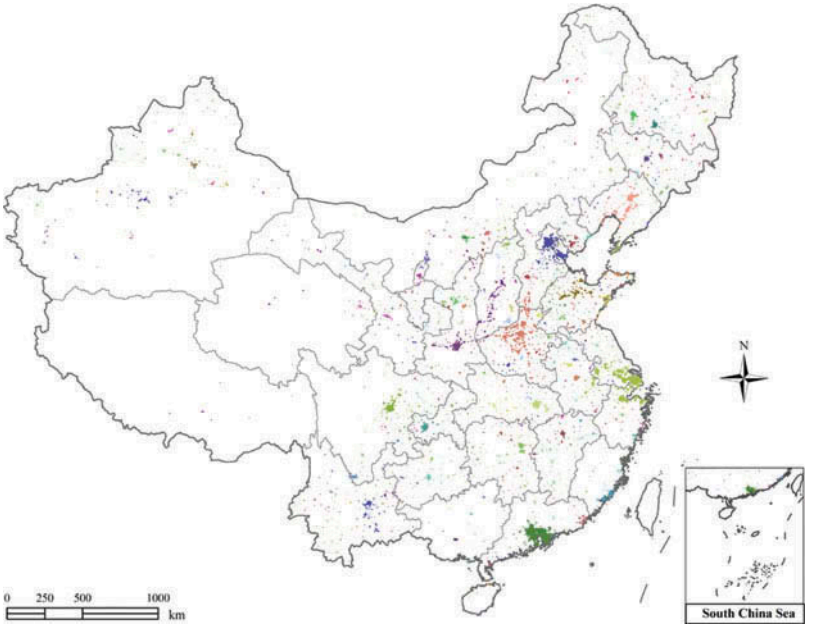

Figure 8. The urban clusters classified according to the partitioned minimum spanning tree; urban objects belonging to an urban cluster are coded in the same color.

3.5. Derivation of spatial attributes of urban spatial clusters

A set of attributes are defined and calculated to characterize the urban spatial clusters. The planimetric and shape attributes are also useful for characterizing the morphology and the development level of urban spatial clusters. The planimetric attributes of the urban spatial cluster include centroid location ($\overline{x_c}$, $\overline{y_c}$), area (S_c), length (LEN_c), and width (WID_c). The definitions are the same as those of urban objects (Table 2). The 10 largest urban spatial clusters by area are shown in Figure 9. The shape attributes of the urban spatial cluster include elongatedness (ELG_c), orientation (Φ_c), fractal dimension (D_c), rectangularity (REC_c), ellipticity (ELP_c), and triangularity (TRI_c). Except for fractal dimension, the definitions of other shape attributes are the same as those of urban objects (Table 2). For an urban spatial cluster, the fractal describes the homogeneity of the spatial distribution of urbanized areas in a region (Tannier *et al.* 2005). The fractal dimension of the urban spatial cluster (D_c) is defined in Equation (3).

$$D_c = \lim_{r \to 0} \frac{\ln C(r)}{\ln(r)} \tag{3}$$

where $\boldsymbol{r}$ is a distance scale and $\boldsymbol{C(r)}$ is a spatial correlation function (Equation (4)) of a specific urban spatial cluster.

$$C(r) = \frac{1}{(Num\_Obj - 1)^2} \sum_{i=0}^{Num\_Obj-1} \sum_{j=0}^{Num\_Obj-1} H\left(r - d_{ij}\right) i \neq j \tag{4}$$

$$H\left(r - d_{ij}\right) = \begin{cases} 1\, r - d_{ij} \geq 0, i \neq j \\ 0\, r - d_{ij} < 0, i \neq j \end{cases} \tag{5}$$

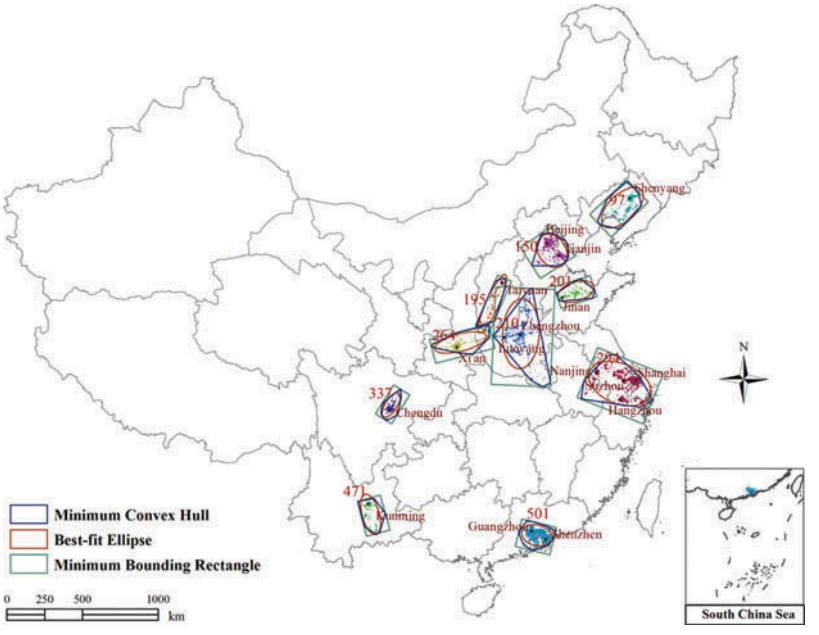

Figure 9. Ten largest urban spatial clusters and their fitted geometries.

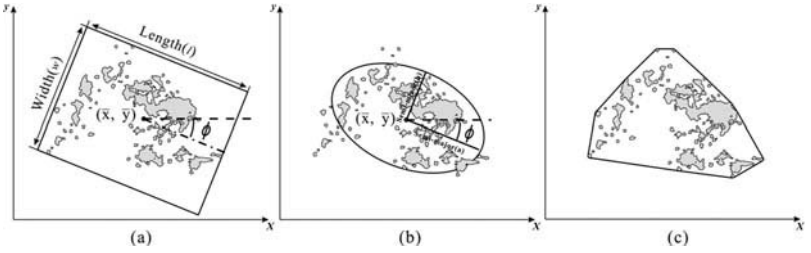

Figure 10. The fitted geometry of an urban spatial cluster: (a) minimum bounding rectangle, (b) best-fit ellipse, and (c) minimum convex hull.

where ***Num_Obj*** is the number of urban objects in an urban spatial cluster, d_{ij} is the minimum distance between the boundaries of urban objects O_i and O_j. $H(r - d_{ij})$ is the Heaviside step function (Equation (5)), which is used to count the number of urban object-pairs whose distances are less than the distance scale r. r is an adaptable distance scale decreasing from $d_{max} = max\{d_{ij}\}$ to $d_{min} = min\{d_{ij}\}$ with a decrement of Δd each time. In our analysis, Δd is set as 10 km.

For each distance scale $r_k = d_{max} - k*\Delta d$, the number ($n_k$) of urban object-pairs whose distance is less than r_k will be counted. Consequently, at this distance scale, $C(r_k) = n_k/(Num\_Obj - 1)^2$. After the numbers of object-pairs at all distance scales are counted, we get two series of corresponding data sets, $\bar{\mathbf{r}} = \{r_0, r_1, r_2, \ldots, r_k\}$ and $\overline{C(r)} = \{C(r_0), C(r_1), \underline{C(r_3)}, \ldots, C(r_k)\}$. Then D_c is given by the slope of a linear regression applied to ($\ln \bar{\mathbf{r}}$, $\ln \overline{C(r)}$) (Equation (3)).

The minimum bounding rectangle (Figure 10(a)) and best-fit ellipse (Figure 10(b)) are also utilized to fit and represent urban spatial clusters. A minimum convex hull (MCH) is defined as the smallest polygon containing all the sampling points, in which any line connecting any two points within the polygon must lie entirely within it. In addition, an MCH has been proved to be a higher-quality geometric approximation of geographical features (Frontiera *et al.* 2008). In our method, the MCH of an urban spatial cluster (Figure 10(c)) is used to delineate the relatively regular shape of the cluster. The minimum bounding rectangles, best-fit ellipses, and MCHs of the 10 largest urban spatial clusters by area are shown in Figure 9.

In this study, an urban spatial cluster is identified as a subtree in the MST partition process. The properties of each subtree also reflect the constituent structural characteristics of the urban spatial cluster. We define a set of properties, including the number of urban objects (*NUM_OBJ*) and the number of edges (*NUM_EDGE*) in a cluster, maximum (*MAX_OBJ_S*), minimum (*MIN_OBJ_S*), average (*AV_OBJ_S*), and standard deviation (*STD_OBJ_S*) of objects' area in a cluster, maximum (*MAX_EDGE_W*), minimum (*MIN_EDGE_W*), average (*AV_EDGE_W*), standard deviation (*STD_EDGE_W*), and sum (*SUM_EDGE_W*) of edges' weights in a cluster, compactness index (*MCH_CI*) of MCH, the ratio of the area of urban objects in a cluster to the area of MCH (*R_OBJ_MCH_S*), and the primacy ratio of urban objects in an urban spatial cluster (*PRI_R_OBJ*). The numerical definitions of these attributes are shown in Table 4.

Main municipal cities contained by these clusters are identified and listed in Table 5. We also computed the morphological attributes (Table 6) and constituent structural attributes (Table 7) of the 10 largest urban spatial clusters. These attributes are combined to provide a comprehensive description of the urban spatial clusters.

Table 4. The definitions of the attributes for the constituent structure of an urban spatial cluster.

Attributes	Definition
Number of urban objects in a cluster (*NUM_OBJ*)	$NUM\_OBJ = Count\{O\}$ O is the urban objects in an urban spatial cluster.
Number of edges in a cluster (*NUM_EDGE*)	$NUM\_EDGE = Count\{E\}$ E is the edges in the subtree of an urban spatial cluster.
Maximum area of objects in a cluster (*MAX_OBJ_S*)	$MAX\_OBJ\_S = max_{i=0}^{NUM\_OBJ-1}\{S_i\}$ S_i is the area of the ith urban object in an urban spatial cluster.
Minimum area of objects in a cluster (*MIN_OBJ_S*)	$MIN\_OBJ\_S = min_{i=0}^{NUM\_OBJ-1}\{S_i\}$
Average area of objects in a cluster (*AV_OBJ_S*)	$AV\_OBJ\_S = \frac{1}{NUM\_OBJ}\sum_{i=0}^{NUM\_OBJ-1} S_i$
Standard deviation of objects' area in a cluster (*STD_OBJ_S*)	$STD\_OBJ\_S = \sqrt{\frac{\sum_{i=0}^{NUM\_OBJ-1}(S_i - AV\_OBJ\_S)^2}{NUM\_OBJ}}$
Maximum weight of edges in a cluster (*MAX_EDGE_W*)	$MAX\_EDGE\_W = max_{i=0}^{NUM\_EDGE-1}\{W_i\}$ W_i is the weight of the ith edge in an urban spatial cluster.
Minimum weight of edges in a cluster (*MIN_EDGE_W*)	$MIN\_EDGE\_W = min_{i=0}^{NUM\_EDGE-1}\{W_i\}$
Average weight of edges in a cluster (*AV_EDGE_W*)	$AV\_EDGE\_W = \frac{1}{NUM\_EDGE}\sum_{i=0}^{NUM\_EDGE-1} W_i$
Standard deviation of edges' weight in a cluster (*STD_EDGE_W*)	$STD\_EDGE\_W = \sqrt{\frac{\sum_{i=0}^{NUM\_EDGE-1}(W_i - AV\_EDGE\_W)^2}{NUM\_EDGE}}$
Sum weight of edges in a cluster (*SUM_EDGE_W*)	$SUM\_EDGE\_W = \sum_{i=0}^{NUM\_EDGE-1} W_i$
Compactness index (*MCH_CI*) of MCH	$MCH\_CI = \frac{4\pi {S_{MCH}}^2}{P_{MCH}}$ S_{MCH} is the area of the MCH and P_{MCH} is the parameter of the MCH of an urban spatial cluster.
Ratio of the area of urban objects in a cluster to the area of MCH (*R_OBJ_MCH_S*)	$R\_OBJ\_MCH\_S = \frac{S_c}{S_{MCH}}$
Primacy ratio of urban objects in an urban spatial cluster (*PRI_R_OBJ*)	$PRI\_R\_OBJ = \frac{S_{1st}}{S_{2nd}}$ S_{1st} and S_{2nd} are the areas of the 1st and 2nd largest urban objects in an urban spatial cluster.

4. Discussions

4.1. Urban spatial cluster analysis

We analyzed the development and spatial morphology of urban spatial clusters in terms of their spatial extent, members, and inner connections from our object-based analysis. The urban spatial cluster located in the Yangtze River Delta is used to illustrate the spatial organization or structure of an urban cluster (Figure 11 and Table 8).

Table 5. The main municipal cities lying in the 10 largest urban clusters in the order of area.

Rank	Cluster ID	Area (km^2)	Number of cities	Main cities				
1	294	20887	20	Shanghai	Shaoxing	Zhoushan	Hangzhou	Guichi
				Ningbo	Huzhou	Tongling	Xuanzhou	Jiaxing
				Suzhou	Wuhu	Wuxi	Chaohu	Maanshan
				Changzhou	Nantong	Zhenjiang	Yangzhou	Nanjing
2	210	17855	14	Zhengzhou	Luoyang	Luohe	Pingdingshan	Xuchang
				Xingtai	Zhumadian	Kaifeng	Xinyang	Jiaozuo
				Xinxiang	Hebi	Anyang	Handan	
3	501	15460	10	Guangzhou	Zhuhai	Jiangmen	Zhongshan	Qingyuan
				Shenzhen	Foshan	Dongguan	Zhaoqing	Huizhou
4	150	11042	4	Beijing	Tianjing	Langfang	Baoding	
5	97	8653	11	Shenyang	Yingkou	Anshan	Panjing	Liaoyang
				Benxi	Fushun	Huxin	Tieling	Jingzhou
				Huludao				
6	264	6081	8	Xi'an	Shangzhou	Xianyang	Weinan	Baoji
				Sanmenxia	Yuncheng	Tongchuan		
7	201	5704	5	Jinan	Weifang	Zibo	Binzhou	Dongying
8	195	5034	3	Taiyuan	Yuci	Linfen		
9	337	4662	3	Chengdu	Mianyang	Deyang		
10	471	3785	3	Kunming	Yuxi	Gejiu		

Table 6. Morphological attributes of the 10 largest urban clusters in the order of area.

Cluster ID	S_c (km^2)	$\overline{x_c}$	$\overline{y_c}$	LEN_c (km)	WID_c (km)	ELG_c	Φ_c	D_c	REC_c	ELP_c	TRI_c
294	20887	1444.39	3420.54	445.69	274.44	1.62	22.00	0.95	0.17	0.05	0.08
210	17855	774.41	3752.91	480.81	250.10	1.92	92.75	0.84	0.15	0.04	0.06
501	15460	888.58	2414.31	211.42	141.64	1.49	13.75	0.64	0.52	0.50	0.73
150	11042	988.05	4306.35	243.65	187.77	1.30	56.95	0.83	0.24	0.11	0.16
97	8653	1458.05	4585.91	353.68	182.77	1.94	136.86	0.93	0.13	0.03	0.05
264	6081	368.55	3680.50	367.47	118.59	3.10	165.73	0.82	0.14	0.04	0.05
201	5704	1139.91	4023.44	257.65	132.55	1.94	162.19	0.72	0.17	0.05	0.08
195	5034	593.35	3956.69	437.54	89.15	4.91	113.66	0.70	0.13	0.03	0.05
337	4662	−80.25	3268.46	193.57	81.65	2.37	122.20	0.73	0.29	0.16	0.24
471	3785	−217.42	2560.61	286.25	117.52	2.44	74.68	0.66	0.11	0.02	0.03

The artificial light in night sensed in the space indicates intensive human activities. Urban objects derived from the nightlight image represent discrete urban built-up patches. Some large urban objects cover the entire built-up area of a single municipal city, while other large objects may contain urban built-up areas of several adjacent municipal cities. The urban built-up areas of the municipal cities contained in the same urban object appear as a geographically continuous lit urban region in DMSP-OLS nighttime light image. The urban objects are the basic components of an urban spatial cluster, and we define the large urban object that contains two or more municipal cities as an Urban Agglomeration (UA).

Table 7. The constituent structure attributes of the 10 largest urban clusters.

Cluster ID	*NUM_OBJ*	*NUM_EDGE*	*MAX_OBJ_S* (km^2)	*MIN_OBJ_S* (km^2)	*AV_OBJ_S* (km^2)	*STD_OBJ_S*	*MAX_EDGE_W* (km)	*MIN_EDGE_W* (km)	*AV_EDGE_W* (km)	*STD_EDGE_W* (km)	*SUM_EDGE_W* (km)	*MCH_CI*	*R_OBJ_MCH_S*	*PRI_R_OBJ*
294	61	60	9795	20	342.41	1284.68	25.50	0.00	8.23	5.63	493.96	0.79	0.20	4.75
210	72	71	5600	20	247.99	707.87	39.62	0.00	9.05	8.44	642.61	0.61	0.15	3.61
501	15	14	14480	27	1030.66	3720.96	16.42	1.00	6.10	5.02	85.46	0.89	0.48	73.90
150	36	35	8527	20	306.74	1410.72	14.36	1.00	6.98	4.08	244.39	0.85	0.27	25.67
97	47	46	2387	20	184.10	399.63	28.53	0.00	8.06	6.22	370.85	0.83	0.16	1.69
264	38	37	3207	20	160.03	513.75	24.76	1.00	9.59	6.76	354.80	0.60	0.15	9.49
201	26	25	1529	22	219.39	376.58	11.63	0.00	6.25	2.86	156.37	0.75	0.24	1.20
195	31	30	862	20	162.39	211.17	13.83	0.00	6.24	3.90	187.31	0.53	0.18	1.04
337	17	16	3180	21	274.23	753.65	21.19	0.00	7.97	6.35	127.60	0.73	0.28	8.78
471	28	27	1450	20	135.17	272.27	36.82	0.00	11.06	10.37	298.53	0.82	0.14	3.77

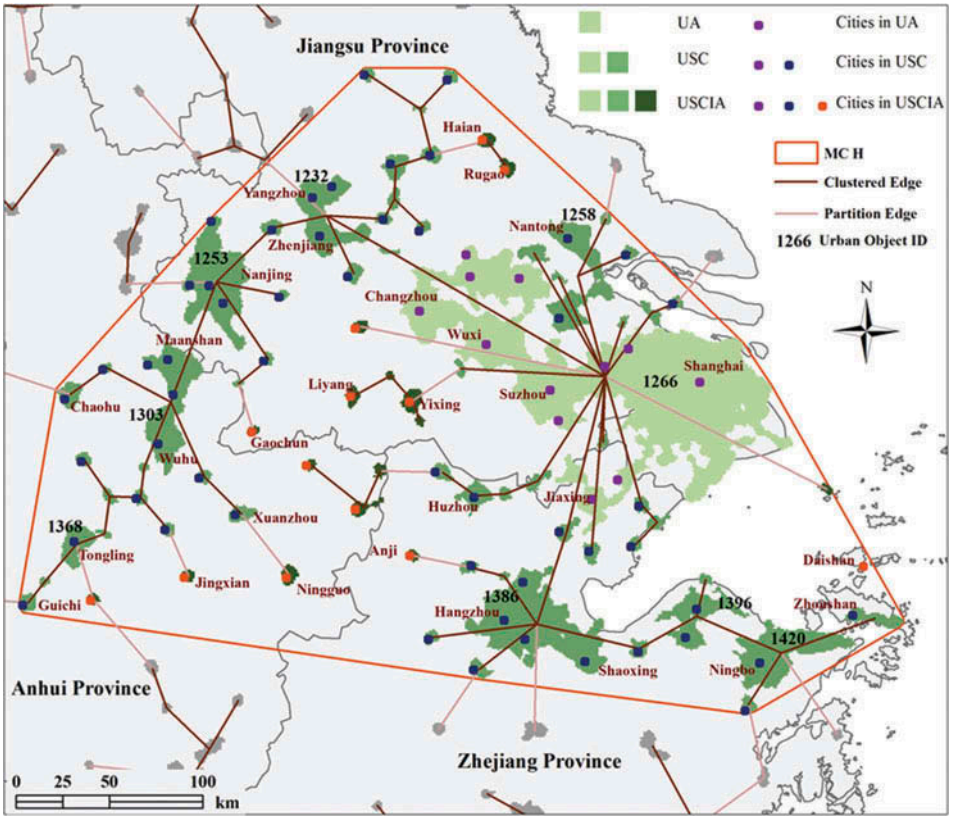

Figure 11. The three levels of urban spatial units located in the Yangtze River Delta.

The largest urban object (object 1266, see Figure 11) in the Yangtze River Delta, which contains five main municipal cities, including Shanghai, Wuxi, Suzhou, Changzhou, and Jiaxing, and eight other cities (see Table 8), is a typical urban agglomeration. With the top economic competitiveness in China, Shanghai serves as the main driving force for economic development and urbanization in the Yangtze River Delta area (Jiang *et al.* 2010). In order to lower the operation cost, many companies in Shanghai select the neighboring cities, such as Suzhou and Wuxi, as the locations for manufacturing plants and warehouses. Meanwhile, new residential districts of the neighboring cities of Shanghai are booming not only around their own city cores but also in the area adjacent to Shanghai in order to house residents from the nearby metropolises. Since the built-up areas of the cities in the region are geographically so close together, they cannot be separated in the nighttime light satellite image, and thus form a large urban agglomeration.

An Urban Spatial Cluster (USC) contains two or more urban objects. Although there is apparent spatial separation between urban objects, these urban objects have strong spatial interaction due to the relatively close geographical proximity and therefore can be considered as an integrated urban spatial unit.

Figure 12 is the tree view of the urban objects in an integrated urban spatial cluster in the Yangtze River Delta region. Large urban objects, namely urban agglomerations, are shown in the tree as nodes that have multiple edges connecting to other urban objects. For example, besides the largest urban object-Shanghai (object-1266), urban object-Nanjing

Table 8. Cities of the three-level urban spatial units in Yangtze River Delta.

Level	Number of cities	Cities						
1 Urban agglomeration	12	Shanghai	Suzhou	Wuxi	Changzhou	Jiaxing	Jinjiang	Zhangjiagang
		Kunshan	Jiashan	Wujiang	Jiangyin	Taicang		
2 Urban Spatial Cluster	43	Cities in Urban Agglomeration						
		Hangzhou	Nanjing	Ningbo	Huzhou	Wuhu	Xuanzhou	Tongling
		Shaoxing	Zhoushan	Guichi	Chaohu	Maanshan	Nantong	Zhenjiang
		Yangzhou	Linan	Haining	Yuhang	Shangyu	Yuyao	Cixi
		Fenghua	Dongtai	Changxing	Xinghua	Jiangning	Danyang	Xiaoshan
		Taizhou	Changshu	Lishui				
3 Urban Spatial Cluster Influencing Region	56	Cities in Urban Agglomeration and Urban Spatial Cluster						
		Ningguo	Qingyang	Anji	Rugao	Huaian	Jintan	Jingxian
		Daishan	Guangde	Langxi	Gaochun	Yixing	Liyang	

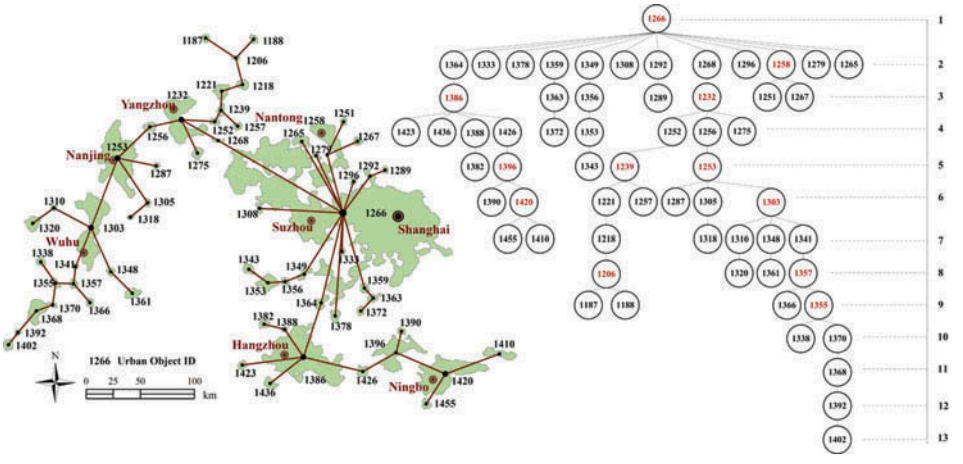

Figure 12. The tree view of urban spatial cluster in the Yangtze River Delta.

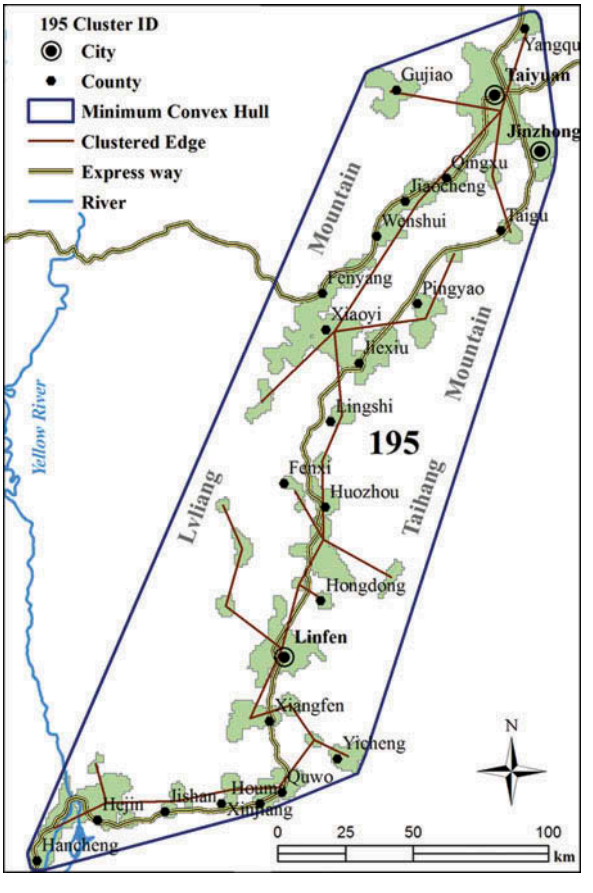

Figure 13. Urban spatial cluster developing along with expressway.

(object-1253) and urban object-Hangzhou (object-1386) have four and five linked edges, respectively. Those urban agglomerations (nodes) serve as the economic and social hinges and have a wide influencing domain (Liang 2009) in the urban cluster.

Urban spatial clusters detected by our method show close resemblance to the regional network of cities delineated by Gu *et al.* (2008) based on the strength of territorial relations, which are defined as the social and economic proximity relationship between neighboring cities (Dematteis 1997). For instance, the Shanghai urban system identified by Gu *et al.* (2008) includes most of the cities in the detected urban spatial cluster (Table 8) except for City Guichi. For the remaining urban clusters listed in Table 5, their member cities (urban objects) are largely in line with those cities in the corresponding urban systems described in Gu *et al.* (2008). Unlike the economic and social proximity, our urban cluster analysis based on spatial proximity can depict the spatial patterns and links of cities in a more objective and direct manner. The hierarchical level of the cities in the region can be identified from the geographical-proximity-based graph. The nodes in the graph have locational advantages in the development of the region.

In our regionalization, an MCH is used to delimit a continuous region that contains all urban objects in an urban spatial cluster and their immediately surrounding rural areas. Some small urban objects that do not belong to the urban spatial cluster may also be contained in the MCH. The MCH of urban spatial cluster is referred to as the Urban Spatial Cluster Influencing Region (USCIR) in this study. The area inside the MCH is strongly influenced and supported by the core urban agglomerations (objects). The urban spatial cluster influencing region (see Figure 12) defined by the MCH of the Yangtze River Delta urban cluster exhibits geometric similarities with the urban influence domain defined by Liang (2009) based on the gravity model and social-economic statistical data (such as GDP and population) of each city. All the urban spatial clusters in Figure 9(b) also appear in Liang's delineation of the 15 largest urban influence domains (Liang 2009).

4.2. Form and evolution of urban spatial clusters

The spatial distribution and evolution of urban spatial clusters are influenced jointly by various factors, such as topography, stream drainage system, economy, and transportation. Among them, natural environment, for example topography, usually determines the form and expansion direction of an urban spatial cluster (Thapa *et al.* 2011). For cities located in a mountainous region, their spatial expansions are surely more restricted. Improved transportation condition often stimulates urban expansion outward (Duranton *et al.* 2012).

The numerical attributes derived for urban spatial clusters provide a useful set of indicators for diagnosing their development stage, driving forces and possible problems. As shown in Figure 13, the largest urban object in the urban cluster (ID 195) (see Table 5) in Shanxi Province, China, contains City of Taiyuan, the capital of Shanxi Province. The expressways apparently played an important role in the formation and evolution of this elongated urban spatial cluster. The high elongatedness value (*ELGc* = 4.91) reflects the formation of the urban spatial clusters along a transportation corridor. In addition, it is apparent that the spatial pattern of the clustered edges of this cluster is consistent with that of the expressways in the region. Therefore, the MST can reflect the structure of transportation network to a certain degree.

In an urban spatial cluster, the leading urban agglomeration (the largest urban object) is the core and focus of the cluster. The primacy ratio of urban objects in an urban spatial cluster (*PRI_R_OBJ*) indicates the level of dominance of the leading urban agglomeration in that cluster. The urban spatial cluster (ID 501) in the Pearl River Delta region has the

highest *PRI_R_OBJ* value (73.90) among the 10 largest clusters (see Table 7). The largest urban object (ID 1966) (see Table 3), which contains Guangzhou, Shenzhen, Zhuhai, and other cities, plays the dominant role in the formation of the Pearl River Delta urban cluster.

The sum (*SUM_EDGE_W*) and average (*AV_EDGE_W*) of edge weights in a cluster are two indicators for quantifying the overall spatial proximity and the corresponding level of spatial integration and interaction between urban objects in a region. Normally, the urban spatial cluster with lower *SUM_EDGE_W* and *AV_EDGE_W* values has higher spatial integration level than others. As shown in Table 7, the urban spatial cluster (ID 501) in the Pearl River Delta region has the lowest *SUM_EDGE_W* as well as *AV_EDGE_W* values. This is because the Urban Agglomeration of Guangzhou (No. 1966 urban object) is the unrivaled leading unit in the development of the region. It exerts a powerful centripetal force to other cities in the region, which effectively results in an increasing spatial concentration and interconnectedness of urban built-up areas of those cities.

4.3. *Spatial distribution of cities in an urban spatial cluster*

Spatial form and organization of an urban spatial cluster has a direct impact on the efficiency of utilizing various natural or social resources in this area (Williams *et al.* 2000). Although there is no consensus on whether the compact city and compact urban spatial cluster are sustainable forms for urban development, quantifying compactness is unquestionably a useful way to assess the internal structure of an urban spatial cluster. The level of urbanization and integration of an urban cluster can be measured by the fractal dimension (D_c) and average edge weight (*AV_EDGE_W*). The fractal dimension for an urban spatial cluster, ranging from 0 to 2.0, reflects the homogeneity of urbanization and spatial integration in the region (Tannier *et al.* 2005). A small value of D_c indicates a higher level of spatial concentration of urbanized areas in the region, and accordingly a lower level of urbanization and spatial integration. A large fractal dimension value, on the contrary, means a more dispersed and even spatial distribution of urbanized areas, indicating a higher level of urbanization and spatial integration. The urban spatial cluster (ID 501) in Table 6 with a relatively low fractal dimension (D_c = 0.64) indicates its relatively lower level of urbanization and spatial integration, while the cluster (ID 294) with D_c = 0.95 shows a relatively more balanced urbanization and higher degree of spatial integration.

The fractal dimension and other attributes calculated in our study can be used to quantify many other characteristics of the urban clusters, such as self-similarity, fragmentation of spatial patterns at different scales, hierarchical organization, and nonlinear dynamics (Tannier *et al.* 2005). Those characteristics are useful for us to understand the development and the evolution of the urban spatial clusters.

5. Conclusions

Cities are interlinked and interdependent, based on the movement or exchange of goods, services, materials, people, money, and information. Improving technologies in transportation and communication have drastically enhanced interactions between cities and between cities and their hinterlands. Cities and towns are thus organized into an integral urban system or network in a region or a country. The spatial structure of urban system is successively transformed by the processes of innovation diffusion, control, migration, and investment (Yeates *et al.* 1997), leading to the formation and development of extensive

multi-nuclei and hierarchical urban spatial clusters. With the increasing pace of urbanization worldwide, more and more urban spatial clusters have been emerged and expanded. The scientific knowledge and critical information about urban spatial clusters are important for monitoring and managing this mega-scale urbanization process. In this study, we demonstrate that the object-based spatial clustering method along with synoptic remotely sensed data provides an effective quantitative method for detecting and analyzing urban spatial clusters.

This paper presents a novel object-based method for detecting and characterizing urban spatial clusters from nighttime light satellite images. The method uses the MST to represent spatial proximity relationships among urban objects. Unlike previous studies, the distance between urban objects (i.e., the boundaries of urban built-up areas) rather than urban centers are adopted to quantify the spatial proximity. We developed an object-based competing propagation algorithm to compute the minimum distance between urban objects and to construct the MST simultaneously. A method based on the Gestalt theory is employed to partition the MST and to identify subtrees as urban spatial clusters. Compared with the manual visual interpretation method, the urban spatial clusters detected by our automated method are more objective and repeatable.

We have applied our method to the nighttime light satellite imagery and numerically identified urban spatial clusters in China at the national scale. A set of morphological attributes has been derived to depict the shapes and development levels of urban spatial clusters. In addition, some structural attributes of urban spatial clusters are defined to quantify the interconnectedness between urban objects in a cluster. Three levels of urban spatial unit, including urban agglomeration, urban spatial cluster, and urban spatial cluster influencing region, are delineated and used to characterize the relative strength of linkages between cities. The empirical findings of this research portray the spatial organization of China's national urban system. Regional urban clusters or subsystems with different spatial extents are clearly recognizable.

Our analysis results are consistent with previous studies, and also provide much needed and multi-dimensional quantitative measures of the spatial distribution, constitution, hierarchical structure, and pattern of China's national urban system. These measures are of great value for researchers, city planners, administrators, and others alike to better understand and manage the dynamics of this type of spatially extensive yet internally connected and integrated urbanized areas. In the future, we will analyze the spatiotemporal variations of urban spatial clusters in China by applying our object-based spatial clustering method to time series DMSP-OLS nighttime light images. We believe that our method would find wide applications in urban geography studies with various remote-sensing images.

Acknowledgments

The authors thank three anonymous reviewers for their constructive comments and suggestions.

Funding

This study was supported by the Specialized Research Fund of Key Lab of Geographic Information Science, Ministry of Education (Grant No. KLGIS2011C04), the Grant from Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration (Grant No. SHUES2012A03), and the Fundamental Research Funds for the Central Universities of China.

References

Alonso-Villar, O., 2002. Urban agglomeration: knowledge spillovers and product diversity. *The Annals of Regional Science*, 36, 551–573. doi:10.1007/s001680200090.

Assunção R.M., *et al.*, 2006. Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20, 797–811. doi:10.1080/13658810600665111.

Batten, D.F., 1995. Network cities: creative urban agglomerations for the 21st century. *Urban Studies*, 32, 313–327. doi:10.1080/00420989550013103.

Baugh, K., *et al.*, 2010. Development of a 2009 stable lights product using DMSP-OLS data. *Proceedings of the Asia-Pacific Advanced Network*, 30, 114–130. doi:10.7125/APAN.30.17.

Brandli, H.W., 1978. The night eye in the sky. *Photogrammetric Engineering and Remote Sensing*, 44, 503–505.

Chen, Y.G. and Zhou, Y., 2006. Reinterpreting central place networks using ideas from fractals and self-organized criticality. *Environment and Planning B: Planning and Design*, 33, 345–364. doi:10.1068/b31131.

Chen, Y.G. and Zhou, Y.X., 2003. The rank-size rule and fractal hierarchies of cities: mathematical models and empirical analyses. *Environment and Planning B: Planning and Design*, 30, 799–818. doi:10.1068/b2948.

Dematteis, G., 1997. Globalisation and regional integration: the case of the Italian urban system. *GeoJournal*, 43, 331–338. doi:10.1023/A:1006833200307.

Doll, C.N.H., Muller, J.P., and Morley, J.G., 2006. Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics*, 57, 75–92. doi:10.1016/j.ecolecon.2005.03.007.

Duranton, G. and Turner, M.A., 2012. Urban growth and transportation. *Review of Economic Studies*, 79, 1407–1440. doi:10.1093/restud/rds010.

Eggers, H., 1998. Two fast Euclidean distance transformations in Z^2 based on sufficient propagation. *Computer Vision and Image Understanding*, 69, 106–116. doi:10.1006/cviu.1997.0596.

Elvidge, C.D., *et al.*, 1997a. Mapping city lights with nighttime data from the DMSP operational linescan system. *Photogrammetric Engineering and Remote Sensing*, 63, 727–734.

Elvidge, C.D., *et al.*, 1997b. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing*, 18, 1373–1379. doi:10.1080/014311697218485.

Elvidge, C.D., *et al.*, 1999. Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment*, 68, 77–88. doi:10.1016/S0034-4257(98)00098-4.

Elvidge, C.D., *et al.*, 2007. The Nightsat mission concept. *International Journal of Remote Sensing*, 28, 2645–2670. doi:10.1080/01431160600981525.

Fang, C., Qi, W., and Song, J., 2008. Researches on comprehensive measurement of compactness of urban agglomerations in China. *Acta Geographica Sinica*, 63, 1011–1021.

Fang, C., Song, J., and Lin, X., 2010. *Theory and practice of urban agglomeration in China*. Beijing: Science Press.

Friedmann, J., 1973. *Urbanization, planning, and national development*. Beverly Hills, CA: Sage Publications.

Frontiera, P., Larson, R., and Radke, J., 2008. A comparison of geometric approaches to assessing spatial similarity for GIR. *International Journal of Geographical Information Science*, 22, 337–360. doi:10.1080/13658810701626293.

Geddes, P., 1915. *Cities in evolution*. London: Williams.

Gottmann, J., 1957. Megalopolis or the urbanization of the Northeastern Seaboard. *Economic Geography*, 33, 189–200. doi:10.2307/142307.

Gottmann, J., 1961. *Megalopolis: the urbanized North Eastern Seaboard of the United States*. Cambridge, MA: The MIT Press.

Gu, C., *et al.*, 1999. *Urban geography of China*. Beijing: The Commercial Press Library.

Gu, C., Yu, F., and Li, W., 2008. *China's urbanization: pattern, process, and mechanism*. Beijing: Science Press.

He, C.Y., *et al.*, 2006. Restoring urbanization process in China in the 1990s by using non-radiance calibrated DMSP/OLS nighttime light imagery and statistical data. *Chinese Science Bulletin*, 51, 1614–1620. doi:10.1007/s11434-006-2006-3.

Henderson, M., *et al.*, 2003. Validation of urban boundaries derived from global night-time satellite imagery. *International Journal of Remote Sensing*, 24, 595–609. doi:10.1080/01431160304982.

Imhoff, M.L., *et al.*, 1997. A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area. *Remote Sensing of Environment*, 61, 361–370. doi:10.1016/S0034-4257(97)00046-1.

Jiang, Y. and Shen, J., 2010. Measuring the urban competitiveness of Chinese cities in 2000. *Cities*, 27, 307–314. doi:10.1016/j.cities.2010.02.004.

Letu, H., *et al.*, 2012. A saturated light correction method for DMSP/OLS nighttime satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 389–396. doi:10.1109/TGRS.2011.2178031.

Liang, S.M., 2009. Research on the urban influence domains in China. *International Journal of Geographical Information Science*, 23, 1527–1539. doi:10.1080/13658810802363614.

Liu, H. and Jezek, K.C., 2004. Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. *International Journal of Remote Sensing*, 25, 937–958. doi:10.1080/0143116031000139890.

Liu, H.X., *et al.*, 2010. An object-based conceptual framework and computational method for representing and analyzing coastal morphological changes. *International Journal of Geographical Information Science*, 24, 1015–1041. doi:10.1080/13658810903270569.

Liu, Z.F., *et al.*, 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning*, 106, 62–72. doi:10.1016/j.landurbplan.2012.02.013.

Lo, C.P., 2002. Urban indicators of China from radiance-calibrated digital DMSP-OLS nighttime images. *Annals of the Association of American Geographers*, 92, 225–240. doi:10.1111/1467-8306.00288.

Luo, J. and Wei, Y.H.D., 2009. Modeling spatial variations of urban growth patterns in Chinese cities: the case of Nanjing. *Landscape and Urban Planning*, 91, 51–64. doi:10.1016/j.landurbplan.2008.11.010.

Ma, R.H., *et al.*, 2008. Mining the urban sprawl pattern: a case study on Sunan, China. *Sensors*, 8, 6371–6395. doi:10.3390/s8106371.

The Ministry of Land and Resources of the People's Republic of China, 2005. *China land & resources Almanac*. Beijing: Geological Publishing House.

Okabe, A. and Miller, H.J., 1996. Exact computational methods for calculating distances between objects in a cartographic database. *Cartography and Geographic Information Science*, 23, 180–195. doi:10.1559/152304096782438846.

Okamoto, K., 1997. Suburbanization of Tokyo and the daily lives of suburban people. *In*: P.P. Karan and K.E. Stapleton, eds. *The Japanese city*. Lexington: The University Press of Kentucky, 79–105.

Pettie, S. and Ramachandran, V., 2002. An optimal minimum spanning tree algorithm. *Journal of the ACM*, 49, 16–34. doi:10.1145/505241.505243.

Portnov, B.A. and Wellar, B., 2004. Development similarity based on proximity: a case study of urban clusters in Canada. *Papers in Regional Science*, 83, 443–465. doi:10.1111/j.1435-5597.2004.tb01917.x.

Reschovsky, C., 2004. *Journey to work: 2000* [online]. Washington, DC: US Census Bureau. Available from: http://www.census.gov/prod/2004pubs/c2kbr-33.pdf [Accessed 11 November 2012].

Roychowdhury, K., *et al.*, 2011. A comparison of high and low gain DMSP/OLS satellite images for the study of socio-economic metrics. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4, 35–42. doi:10.1109/JSTARS.2010.2053022.

Shi, K., *et al.*, 2014a. Evaluation of NPP-VIIRS nighttime light composite data for extracting built-up urban areas. *Remote Sensing Letters*, 5, 358–366. doi:10.1080/2150704X.2014.905728.

Shi, K., *et al.*, 2014b. Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: a comparison with DMSP-OLS data. *Remote Sensing*, 6, 1705–1724. doi:10.3390/rs6021705.

Small, C., Pozzi, F., and Elvidge, C.D., 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment*, 96, 277–291. doi:10.1016/j.rse.2005.02.002.

Sutton, P., *et al.*, 2001. Census from Heaven: an estimate of the global human population using night-time satellite imagery. *International Journal of Remote Sensing*, 22, 3061–3076. doi:10.1080/01431160010007015.

Sutton, P.C., 2003. A scale-adjusted measure of 'urban sprawl' using nighttime satellite imagery. *Remote Sensing of Environment*, 86, 353–369. doi:10.1016/S0034-4257(03)00078-6.

Tannier, C. and Pumain, D., 2005. Fractals in urban geography: a theoretical outline and an empirical example [online]. *Cybergeo: European Journal of Geography*, Document 307. Available from: http://cybergeo.revues.org/3275 [Accessed 19 March 2013]. doi:10.4000/cybergeo.3275.

Terzi, F. and Kaya, H.S., 2011. Dynamic spatial analysis of urban sprawl through fractal geometry: the case of Istanbul. *Environment and Planning B: Planning and Design*, 38, 175–190. doi:10.1068/b35096.

Thapa, R.B. and Murayama, Y., 2011. Urban growth modeling of Kathmandu metropolitan region, Nepal. *Computers Environment and Urban Systems*, 35, 25–34. doi:10.1016/j.compenvurbsys.2010.07.005.

Vogel, R.K., *et al.*, 2010. Governing global city regions in China and the West. *Progress in Planning*, 73, 1–75. doi:10.1016/j.progress.2009.12.001.

Wertheimer, M., 1958. Principles of perceptual organization. *In*: D. Beardslee and M. Wertheimer, eds. *Readings in perception*. Princeton, NJ: Van Nostrand.

Williams, K., Burton, E., and Jenks, M., 2000. *Achieving sustainable urban form*. London: E & FN Spon.

Yang, S., 1990. *Research on economic regionalization of China*. Beijing: China Expectation Press.

Yang Y., *et al.*, 2013. Timely and accurate national-scale mapping of urban land in China using defense meteorological satellite program's operational linescan system nighttime stable light data. *Journal of Applied Remote Sensing*, 7, 073535–073535. doi:10.1117/1.JRS.7.073535.

Yeates, M. and Garner, B.J., 1997. *The North American city*. New York, NY: Addison Wesley.

Yu, B., *et al.*, 2010. Automated derivation of urban building density information using airborne LiDAR data and object-based method. *Landscape and Urban Planning*, 98, 210–219. doi:10.1016/j.landurbplan.2010.08.004.

Zahn, C.T., 1971. Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions on Computers*, C-20, 68–86. doi:10.1109/T-C.1971.223083.

Zhang, Q.L. and Seto, K.C., 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment*, 115, 2320–2329. doi:10.1016/j.rse.2011.04.032.

Zhou, Y., 1991. The metropolitan interlocking region in China: a preliminary hypothesis in the extended metropolis. *In*: N. Ginsburg, ed. *The extended metropolis, settlement transition in Asia*. Honolulu: University of Hawaii Press, 89–111.

Zhu, Y., 2004. *Analyzing the economic space of urban agglomeration*. Beijing: Science Press.