International Journal of *Geo-Information*

MDPI

*Article*

# Detecting Urban Commercial Districts by Fusing Points of Interest and Population Heat Data with Region-Growing Algorithms

**Bingbing Zhao , Xiao He , Baoju Liu, Jianbo Tang, Min Deng and Huimin Liu ***

Department of Geo-Informatics, Central South University, Changsha, 410083, China
* Correspondence: lhmgis@csu.edu.cn

**Abstract:** Reasonable urban commercial planning must clarify the location and scope of urban commercial districts (UCDs). However, existing studies typically detect spurious UCDs owing to the bias in a single data source while ignoring the continuity and ambiguity of commercial district boundaries. Therefore, in this study, we designed a two-stage approach for detecting UCDs. First, points of interest and population heat data were fused through hotspot and overlay analyses to detect core commercial areas. The boundaries of the UCDs were then identified by considering adjacent blocks using adjusted cosine similarity and region-growing algorithms. Finally, an experiment was conducted in Xiamen, revealing concentrated businesses on Xiamen Island and sparse businesses outside Xiamen Island. An experimental comparison with other strategies confirmed the improved modeling ability of this approach for the edge ambiguity of UCDs. This framework provides tools for urban commercial planning and helps recognize urban commercial patterns in a timely manner.

**Keywords:** urban commercial districts; region-growing algorithms; POIs; population heat data

check for updates

**Citation:** Zhao, B.; He, X.; Liu, B.; Tang, J.; Deng, M.; Liu, H. Detecting Urban Commercial Districts by Fusing Points of Interest and Population Heat Data with Region-Growing Algorithms. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 96. https://doi.org/10.3390/ijgi12030096

Academic Editors: Wolfgang Kainz and Mingshu Wang

Received: 11 December 2022
Revised: 6 February 2023
Accepted: 24 February 2023
Published: 26 February 2023



## 1. Introduction

Urban commercial districts (UCDs) refer to the concentrated areas of commercial buildings that provide products and services to the surrounding area [1–3]. With the rapid growth of the urban economy, the urban form is moving from monocentric to polycentric [4], such that multiple UCDs have gradually appeared in cities. The spatial layout of multiple UCDs determines the connectivity and development of urban economic networks. For example, the cooperation of neighboring UCDs can promote economic benefits, while the layout of UCDs also affects the traffic flow within the city by forming traffic flows from residential areas to UCDs and from UCDs to UCDs. For reasonable urban planning, pinpointing the distribution of UCDs within a city is essential [2].

However, UCDs may shift and renew with economic development needs [5], i.e., their location and scope may change, which poses a challenge to urban planning. Many studies have attempted to identify UCDs [2,6,7]. Initially, these studies collected information on residents' perceptions of commercial districts through questionnaires to determine the geographical boundaries of UCDs [8,9]. However, rapid urbanization has accelerated the creation, variation, and extinction of UCDs. Frequently conducting city-wide large-scale surveys is difficult and costly. Subsequently, the rise of geographical big data (such as points of interests (POIs), Open Street Map (OSM), and travel data) has yielded new insights into UCD recognition [7,10,11]. Kernel density estimation and clustering methods can be used to identify high-density areas based on the spatial distribution of geographical data as UCDs. A threshold is usually set to delineate high-density areas of population or commercial POIs. Areas with high population density, facility density, and road network connectivity are usually considered UCDs [12]. However, single POI data or human activity data are limited in their ability to portray the function of an area. Specifically, the depression of commercial districts may be hidden by high-density POIs because POIs are not real-time. Areas with a high population density, such as residential estates, may be identified as commercial districts when only using population data. Additionally, this method of identifying commercial areas using high-density thresholds ignores the ambiguity of UCD boundaries, which stems from differences in the development of commercial districts. A high-density threshold would focus excessively on the central commercial district at the expense of the secondary commercial districts, but a low-density threshold would limit the spatial extent of the central commercial district. Additionally, most existing studies are based on regular grid statistics to synthesize multi-source data [10,13], which are easy to collect, but separate from daily activity spaces. This separation increases the information error in urban planning. For example, two areas with different functions may belong to the same grid.

To address these questions, this study proposes a method for detecting UCDs by combining POIs and population heat data at the block scale. Its main contributions are as follows: (1) The combination of POIs and population heat data fused the functional characteristics of blocks in human activity and physical spaces, which avoided spurious UCDs as represented by a single data source. As there was a gap between commercial facilities and commercial activities, some areas with a high density of unused commercial facilities were identified as UCDs when only using POI data. (2) Considering that blocks formed by road division are basic spatial units where people live [14], we conducted our study based on the block scale; the features of adjacent blocks were weighted and aggregated to smooth the occurrence of functional characteristics caused by irregular block shapes. The extreme block size may affect the judgment of the blocks' commercial function. For example, a block with an area that is excessively small may present a localized high-density commercial function characteristic, whereas a block with an area that is excessively large may present a low density. This smoothing ensured the functional continuity of the UCDs. (3) Region-growing algorithms were introduced to identify the boundaries of UCDs by measuring the similarity of attributes between adjacent blocks. This adjacency comparison allowed for different UCDs to grow individually; UCDs with different levels of development and spatial extents could be detected. The results under different growth threshold settings provided data support for the fuzzy boundary perception of UCDs.

The remainder of this article is organized as follows. Section 2 introduces related studies and Section 3 describes the method proposed in this study. Section 4 presents the specific study area and data. Section 5 analyzes and discusses the results. Section 6 presents our conclusions and future work.

## 2. Related Work

### *2.1. Concept of Commercial Districts*

Many concepts are related to urban commercial structures, such as central business districts (CBDs), central activity zones (CAZs) [12], and UCDs. The CBD is derived from the "concentric circle theory" proposed by Burgess [15]: land use in a city is a concentric circle structure around the CBD. Proudfoot's study [16] first analyzed commercial districts in cities, where commercial districts no longer referred to the city center, but the gathering districts of businesses. CBDs are distinguished by their size and function, typically having a higher commercial function and larger size [2]. However, rapid economic development has narrowed the gap between these two concepts, with more than 30 cities in China attempting to establish CBDs in 2003 [17]. "CBD" has gradually become widely used to indicate the gathering of business buildings, which is similar to the concept of the UCD, rather than being confined to the city center [12]. Compared to these two concepts, a CAZ is more functional, with the addition of leisure and sports [12]. It is an extension and renewal of the CBD in response to the growing needs of residents. UCDs provide the more basic components of the urban commercial structure and a basis for the development of CAZs.

### *2.2. Data Used for UCD Detection*

Initially, scholars identified UCDs through field research or economic census, using the experiential perceptions of residents and the physical environment to determine the commercial structure. For example, Lamb [18] conducted field research in 40 villages and identified the components of their commercial structures. Lüscher and Weibel [8] examined urban centers by collecting questionnaires from citizens. Ye et al. [19] compared the existing situation with commercial planning through a survey of buildings in Changsha. Zhang [20] used economic census data to analyze the distribution characteristics of commercial outlets in Beijing. However, the economic census and field research requires significant manpower and time, which hinders our understanding of commercial structures promptly. The rise of geographical big data(such as commercial building areas and human activity data) provides an opportunity to address this challenge. The area of commercial buildings [21], the density of commercial facilities [22,23], and the density of origins and destinations for floating vehicles [24] were used to characterize commercial districts. Geographical big data provides more timely feedback on changes in urban space than questionnaire data and is also more readily available, but these studies ignore the bias from the single source data. For example, POI data can reflect the spatial distribution of urban facilities but it is difficult to reflect dynamic attributes in real-time [25].

### *2.3. Geographical Units of UCD Detection*

The spatial grid is usually established as the smallest unit for urban analysis, due to its high regularity and easy delineation. He [26] evaluated the ecological health of land use and its spatial differentiation pattern in Yibin City based on a kilometer grid. Hu [27] quantified the functional areas of Guangzhou city at the scale of 250 m grid by POI data. Similarly, the spatial grid is widely used for UCD detection. Chen [23] gridded the commercial facilities in the urban space and then used the kernel density estimation method to identify the agglomeration centers of commercial facilities. And the density of the starting and ending points of the floating vehicles in the 500 m grid was used to identify the extent of the commercial centers [24]. However, it is difficult to determine the right size of the spatial grid. An overly large grid ignores the complexity of the city, while a grid that is too small can lead to over-analysis [28]. And the rigid fragmentation of natural geographic space by the regular grid is also inconsistent with the continuity of living space [29], which may lead to the fragmentation of continuous commercial areas. Considering the close connection between road networks and urban development [30], blocks surrounded by roads are also often used as the basic unit of urban analysis, showing adaptive sizes with the distribution of urban functions compared to a grid of fixed area sizes. For example, Wang [31] identified block-scale commercial areas in Beijing by calculating the amount of commercial activity, which ensures consistency in commercial distribution and traffic structure. However, existing studies of block-scale urban commercial district analysis are limited, which hinders the development of urban commercial facility planning.

### *2.4. UCD Detection Methods*

UCDs are often identified by detecting the high concentration of commercial activities [7]. Murphy and Vance [21] proposed the Central Business Height Index (CBHI) and Central Business Intensity Index (CBII) to delineate CBDs based on the building area of commercial activities. The intensity of commercial facilities has been calculated by Carol to identify CBDs [22]. Chen et al. [23] applied kernel density estimation and Getis–Ord (Gi*) spatial statistics to identify retail gathering areas using POIs. However, detecting the boundaries of these commercial areas is difficult owing to the functional ambiguity of the boundaries [32]. Boundaries are essential to the planning of commercial areas because they define the geospatial scope of policy implementation [12]. Thus, some studies have proposed methods for identifying the boundaries of commercial districts. Wang et al. [11] constructed a fuzzy affiliation function based on the point density of stores to determine the affiliation between a location and commercial districts. A Standard Deviational Ellipse was built into the high-value aggregation areas of check-in data to analyze the direction and extent of commercial districts [33]. Convex hulls of POIs have also been built to identify the scope of trade areas in Zhao's research [34]. These methods provide references for commercial district planning, but they are mostly based on a single data source, which can be susceptible to bias. These methods identify boundaries by setting a global parameter threshold that ignores the diversity of commercial areas in the city. Specifically, the intensity of commercial functions in each commercial area decreases from the center to the edge according to the distance decay effect [35], but there are variations in the density of commercial facilities at the edge of large and small commercial areas.

## 3. Study Area and Data

### *3.1. Study Area Description*

Xiamen, an important port and scenic tourist city located on the southeast coast of China, was one of the earliest areas in China to open to the world. In October 1980, the State Council approved the establishment of the Xiamen Special Economic Zone (SEZ). Since then, Xiamen's economy has continued to grow; many commercial centers have recently emerged. However, the unbalanced development of commercial space has become increasingly prominent, such as the contradiction between the excess commercial space on Xiamen Island and the shortage of commercial space outside Xiamen island [36,37]. Thus, we selected Xiamen as the study area to analyze the urban commercial structure. Our results can offer more useful suggestions for urban commercial planning. Figure 1 shows a map of Xiamen.

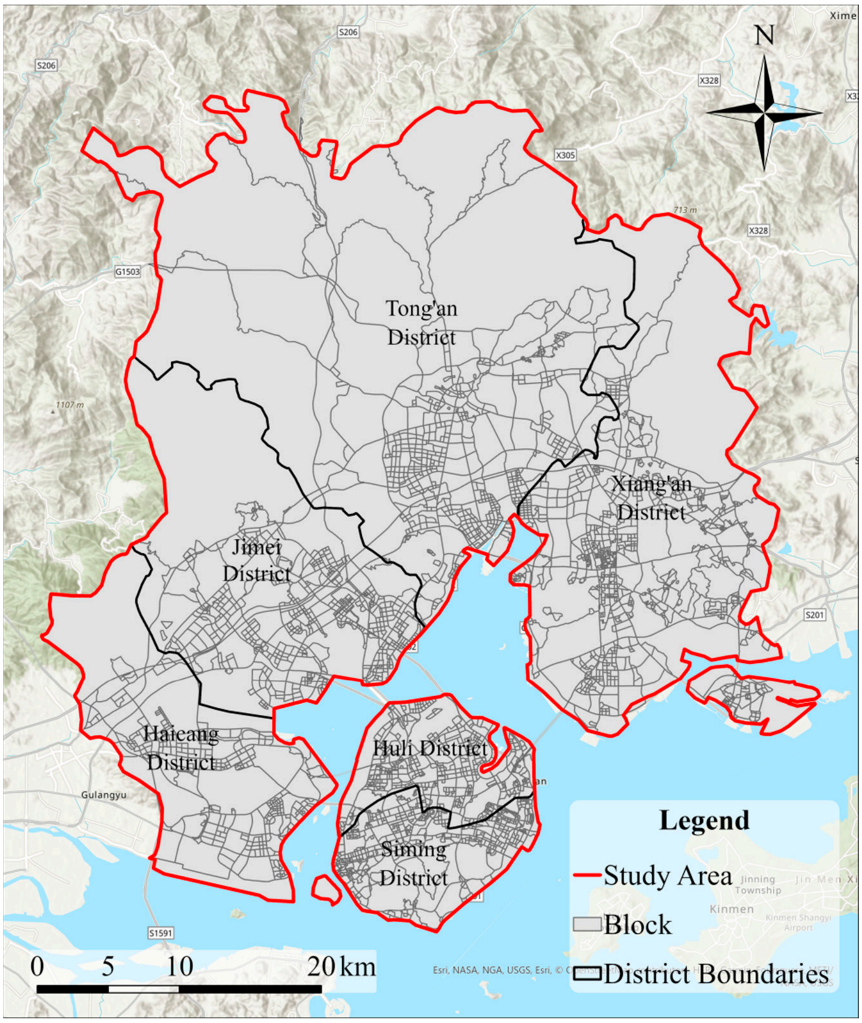
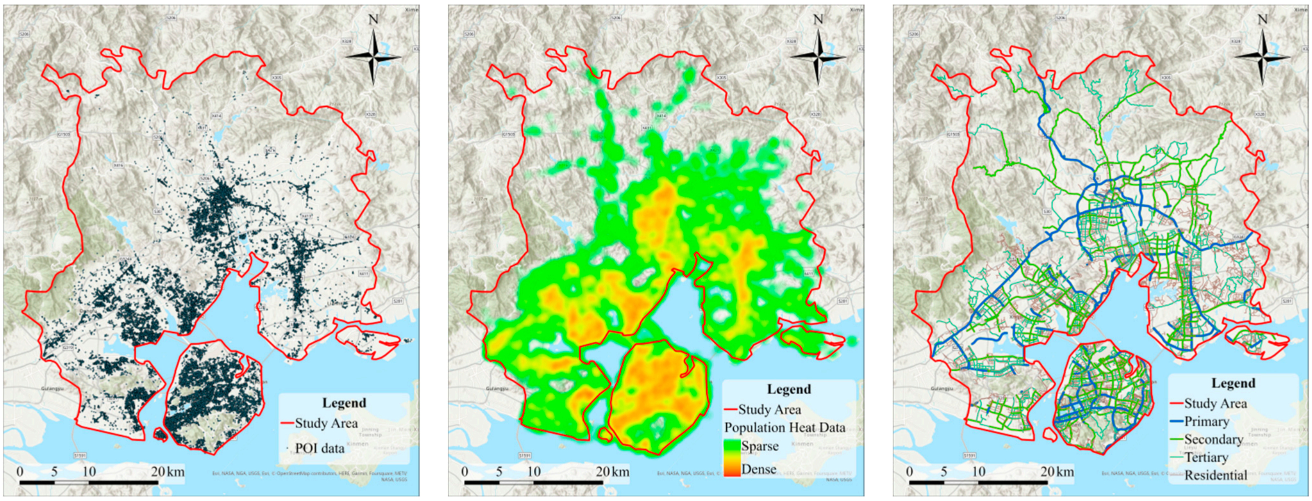



**Figure 1.** Study area.

### *3.2. Research Data*

In this study, we collected the road network of Xiamen to construct the blocks. Blocks are the basic form of urban organization in which people conduct their daily activities [38]. POIs and population heat data were collected to describe the commercial functions of the blocks. Figure 2 shows all datasets.




(**a**) (**b**) (**c**)

**Figure 2.** Visualization of the datasets: (**a**) POIs, (**b**) population heat data, and (**c**) road networks.

POIs record the spatial attribute information of a city, which is widely used to characterize urban functions [39,40]. In this study, 244 764 POIs in Xiamen city in September 2020 were obtained using the Gaode Open Platform (lbs.amap.com). Each POI contains attributes including name, address, coordinates, and category. However, there is a discrepancy between the POI categories and those perceived by residents. According to [41], we re-aggregated the commercial-related categories into POIs, resulting in 105,434 POIs to characterize the commercial functions of regions.

Baidu Map Huiyan (https://huiyan.baidu.com/, accessed on 2 November 2020) is a spatio-temporal big data service based on Baidu Map, which provides population heat data, i.e., the positioning data generated when terminals using Baidu services call Baidu location services, which can reflect the population activities and distribution in the region. Population heat data in Xiamen City from 1–2 November 2020 were obtained from Baidu Maps Huiyan (https://huiyan.baidu.com, accessed on 2 November 2020), which recorded the number of people per 200 $m^2$ per hour in Xiamen city.

Open Street Map (OSM) is an open-source platform that offers a variety of road networks. We collected road networks in Xiamen from OSM, including primary roads, secondary roads, tertiary roads, residential roads, and trunks. The primary, secondary, tertiary, and residential roads were aggregated to build blocks, yielding 3641 blocks.

## 4. Methods

This paper proposes a method for identifying UCDs at the block scale. First, the density of POIs and the population heat value of each block were calculated to describe the commercial function intensity and human activity characteristics, respectively. Getis-Ord $G_i^*$ statistic was used to identify high-value clusters of POIs and population heat in urban spaces as core commercial areas. Second, the commercial function and human activity characteristics on adjacent blocks were aggregated by weighting to quantify the features of the blocks. Finally, region-growing algorithms were introduced to identify the boundaries of commercial districts. Figure 3 shows the workflow of this study.

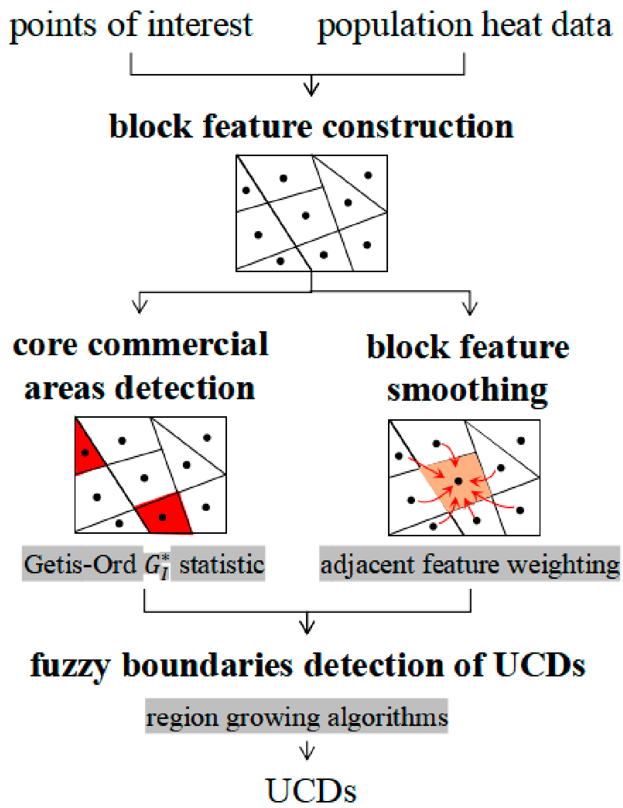



**Figure 3.** Workflow of UCD detection. (Words in the gray background indicate the method applied for that step. The detected core commercial areas are shown in red. For each block, such as the orange block in the figure, its adjacent blocks are used for weighting and are indicated by red arrows in the figure).

### 4.1. Detecting Core Commercial Areas

Core commercial areas usually have high human traffic and a high density of commercial facilities [41], such that they are easily identified. We attempted to identify core commercial areas by detecting the clustering of commercial facilities and human traffic. The density of POIs related to the commercial facilities in the block was calculated to quantify the commercial function intensity of block $i$, denoted as $x_i^{poi}$. Relevant studies have shown that there are two peak periods of human traffic in commercial areas: the peak work period, such as 9:00 am and 6:00 pm, and entertainment activities in commercial districts from 8 to 10 pm [42,43]. Thus, we calculated the human traffic density in these two periods to indicate human activity characteristics, denoted as $x_i^{work}$ and $x_i^{commercial}$, respectively.

The Getis-Ord $G_i^*$ statistic is a type of hotspot analysis method that has been widely used to identify spatial clusters with high or low values [44]. It identifies clusters by comparing the value of the block with that of adjacent blocks, as follows:

$$G_i^* = \frac{\sum_{i=1}^{n} w_{ij}x_i - \overline{X}\sum_{i=1}^{n} w_{ij}}{s\sqrt{\frac{n\sum_{i=1}^{n} w_{i,j}^2 - \left(\sum_{i=1}^{n} w_{i,j}\right)^2}{n-1}}}, \tag{1}$$

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}, \text{ and} \tag{2}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\overline{X}\right)^2}, \tag{3}$$

where $x_i$ represents the attributes of block $i$, including $x_i^{poi}$, $x_i^{work}$, and $x_i^{commercial}$; $w_{ij}$ denotes the weight between blocks $i$ and $j$; and $n$ indicates the number of blocks. When block $i$ is adjacent to block $j$, $w_{ij}$ between them is equal to 1; otherwise, it is 0. Considering the characteristics of the commercial district mentioned above, high-value aggregation regions were detected from the distribution of these three attribute values. Overlay analysis was then used to explore common aggregation regions for these three types of attributes as core commercial areas.

### 4.2. Block Description Considering Adjacency and Distance Decay

The strength of an area's commercial function may diminish as it moves away from the core commercial areas [45], which determines ambiguity at the edges of UCDs, such that it is difficult to identify boundaries. We must compare the functional similarity of adjacent blocks when attempting to identify the boundaries of functional decay, i.e., the boundaries of commercial districts. However, the attributes of individual blocks are often affected by chance, such as a greenbelt in a commercial district that is functionally different from its blocks but still belongs to that commercial district. To eliminate this possibility and amplify differences between adjacent blocks, we aggregated adjacent blocks in a weighted manner for feature construction. The spatial inverse distance weighting method was used to determine the weights between the adjacent blocks. The proximity of neighborhoods results in a stronger spatial influence. The aggregated commercial function of block $i$ was calculated as follows:

$$x_i^{ag_{poi}} = \sum_{j=1}^{k} x_j^{poi} w_{i,j} \text{ and} \tag{4}$$

$$w_{i,j} = \frac{\frac{1}{\sqrt{\left(loc_{x_i} - loc_{x_j}\right)^2 + \left(loc_{y_i} - loc_{y_j}\right)^2}}}{\sum_{j=1}^{k} \frac{1}{\sqrt{\left(loc_{x_i} - loc_{x_j}\right)^2 + \left(loc_{y_i} - loc_{y_j}\right)^2}}}, \tag{5}$$

where $(loc\_x_i, loc\_y_i)$ represents the spatial coordinates of the center of mass location of block $i$, $w_{i,j}$ indicates the weight between blocks $i$ and $j$, and $k$ indicates the number of adjacent blocks of block $i$. Similarly, the human traffic attributes were aggregated as $x_i^{ag_{work}}$ and $x_i^{ag_{commercial}}$. These three attributes were concatenated as vector $x_i$ to describe block $i$:

$$x_i = \left(x_i^{ag_{poi}},\ x_i^{ag_{work}}, x_i^{ag_{commercial}}\right)^T \tag{6}$$

### 4.3. Fuzzy Boundaries Detection of UCDs Supported by Region-Growing Algorithms

The region-growing algorithm is widely applied in the image processing field, which attempts to identify an object by assembling pixels that satisfy the predefined growth criterion [46,47]. This ensures spatial and attribute continuity of the identified objects. Comparing our city as an image, the core commercial areas were set as the seed pixels; the growth criterion was defined according to the adjusted cosine similarity [48] between the two blocks. This is because the original cosine similarity focuses too much on the directional consistency between the vectors and ignores the differences in the norm of the vectors, which may conceal the differences between blocks:

$$sim_{i,j} = \frac{(x_i - \overline{x}) \cdot (x_j - \overline{x})}{|x_i - \overline{x}|\left|x_j - \overline{x}\right|} \text{ and} \tag{7}$$

$$f^I(j) = \begin{cases} 1,\ if\ sim_{i,j} > \varepsilon\ and\ i \in I \\ 0,\ else \end{cases}, \tag{8}$$

where $\overline{x}$ indicates the average of the vectors of all blocks, $\varepsilon$ indicates the growth threshold, $I$ indicates the set of known core commercial areas, and $i$ refers to the block that belongs to set I. Considering each core block detected in Section 3.1 as a seed neighborhood, the adjusted cosine similarity between the seed blocks and adjacent blocks was traversed. Blocks satisfying the growth threshold were marked as UCDs and added to the sequence of seed blocks for traversal until the region no longer grew. The growth results of all core blocks were then overlaid to generate the final UCD detection results.

## 5. Results

### *5.1. Core Commercial Areas Identification and Verification*

After overlaying the Getis-Ord $G_i^*$ statistic analysis results from the POIs and population heat data, 35 blocks were detected as core commercial areas, as shown in Figure 4. Most occurred on Xiamen Island, reflecting its pillar position as an economic center and unevenness in Xiamen's economic development. The core commercial areas on Xiamen Island mainly developed from the former old town, located around Zhongshan Road, Hexiang West Road, and Xiamen Railway Station. Additionally, some new economic development areas have been detected, including shopping areas centered around large brick-and-mortar stores (e.g., METRO Mall and SM City Plaza) and government-planned commercial districts (e.g., Huli Pedestrian Street). A few core commercial areas were also detected outside Xiamen Island, but no large commercial entities were found around them, indicating that there are initial commercial districts in these three districts: other types of commercial districts are driven by daily basic business needs. However, no core commercial areas were detected in the Tongan District, which may be related to its economic development. According to the National Economic and Social Development Statistical Bulletin of Xiamen [49], the gross domestic product (GDP) of Tongan reached 59.121 billion yuan in 2020, which is lower than that of the other districts in Xiamen.

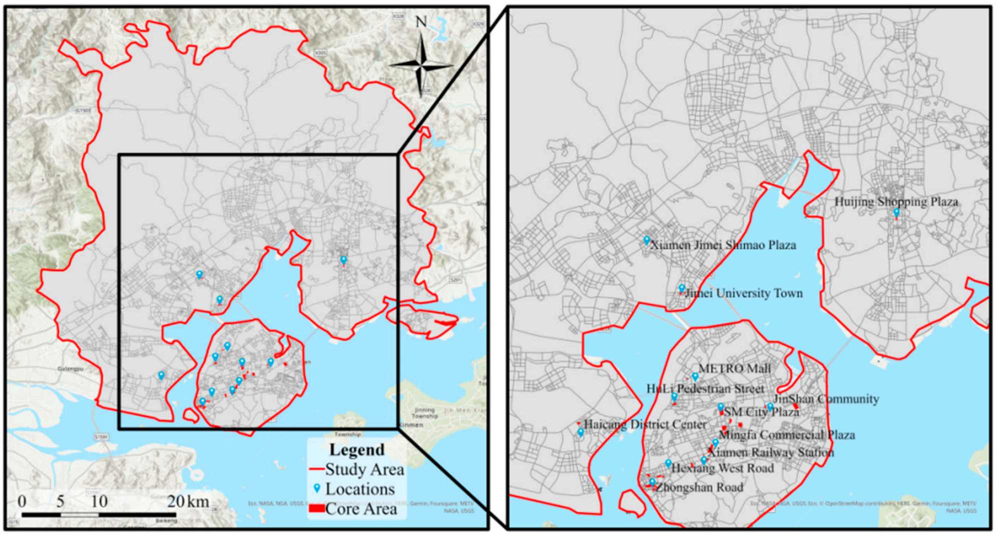



**Figure 4.** Results of the core commercial areas detection.

Furthermore, to verify the correctness of the core commercial areas, we collected the distribution of UCDs in Xiamen from previous studies. According to government reports, such as 'City's level to be upgraded again and cross-island development to be accelerated again' [50] and commercial reports, most of the well-known commercial districts in Xiamen are located on Xiamen Island, concentrated in the southwest and central areas (Figure 5), which is consistent with the results shown in Figure 4. Additionally, emerging commercial districts outside Xiamen Island were identified in this study. These UCDs are usually neglected by business reports due to their low economic development, but they carry out major economic activities outside of the island and are an indispensable component in urban planning.


**Figure 5.** UCDs reported in previous studies [51–54].

## 5.2. UCD Boundary Detection Results

Figure 6 plots the frequency distribution of the adjacency similarity to determine the appropriate growth threshold. A significant turnaround in the frequency of similarity occurred after 0.9 in Figure 6, indicating that several blocks showed a high degree of similarity in commercial functions with their neighbors. With the turning point distinguishing between commercial functionally continuous and fragmented areas, the growth threshold was initially determined (>0.9).

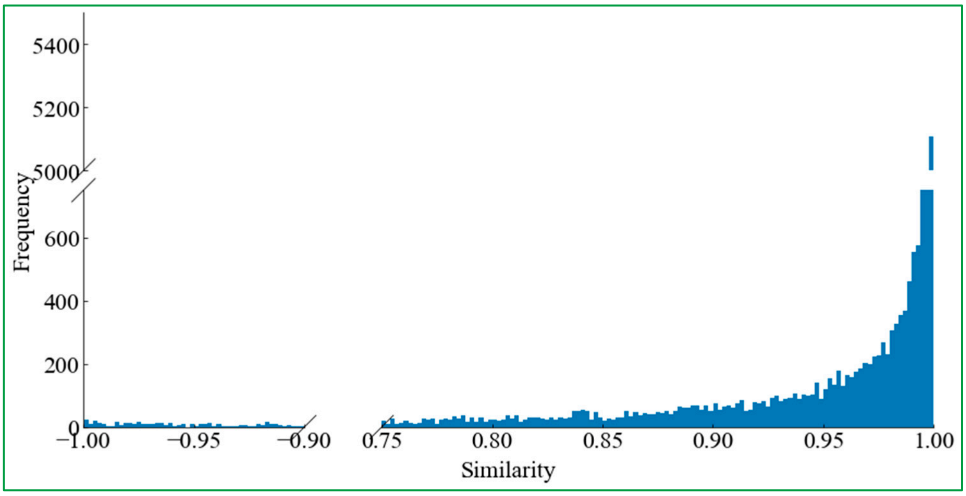



**Figure 6.** Frequency distribution of the adjacency similarity.

Four thresholds (from 0.92 to 0.98, with an interval of 0.02) were then set to detect the boundaries of the commercial districts, as shown in Figure 7. The number of UCDs detected increased as the threshold increased, with 5 UCDs detected at a growth threshold of 0.92 and 13 UCDs detected at a growth threshold of 0.98. This revealed that the internal composition of the UCDs may also be polycentric: a large commercial district may form via the development of multiple smaller commercial districts connected from earlier periods. For example, *UCD_1* in Figure 7a can be regarded as a merger of *UCD_1-1* and *UCD_1-2* in Figure 7b. This emphasizes the necessity for the pre-detection of core commercial areas in UCDs because the distance decay around a single center cannot identify such polycentric structures. The area of the UCDs decreased as the growth threshold increased. For example, the area of *UCD_2* shrunk by one-third when the threshold increased from 0.92 to 0.94; it continued to shrink by one-half when the threshold increased from 0.94 to 0.96. Finally, only a few coastal blocks were identified as *UCD_2* when the threshold reached 0.98. This shrinkage reveals the circled growth structure of urban commercial areas and provides a reference for urban planning, where planners can effectively distinguish between areas of commercial development and areas of commercial concentration. Differences in UCD development were also found in the comparison of the effects of growth thresholds on different commercial districts. For example, each growth threshold change affected the spatial morphology of *UCD_2*, whereas *UCD_3* changed its morphology considerably only when the growth threshold was changed from 0.94 to 0.96. This suggests that the government must consider the development stages of all UCDs in a balanced manner when planning.

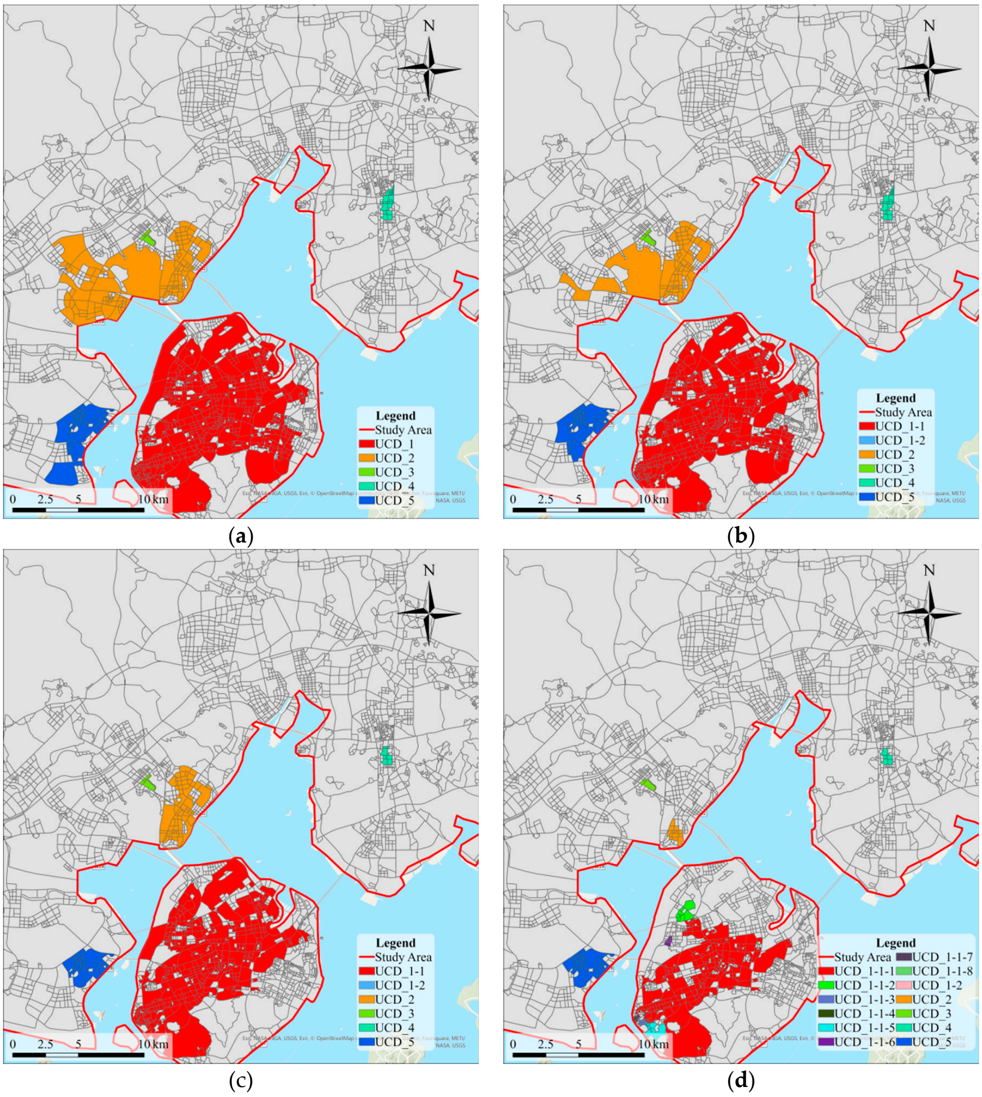


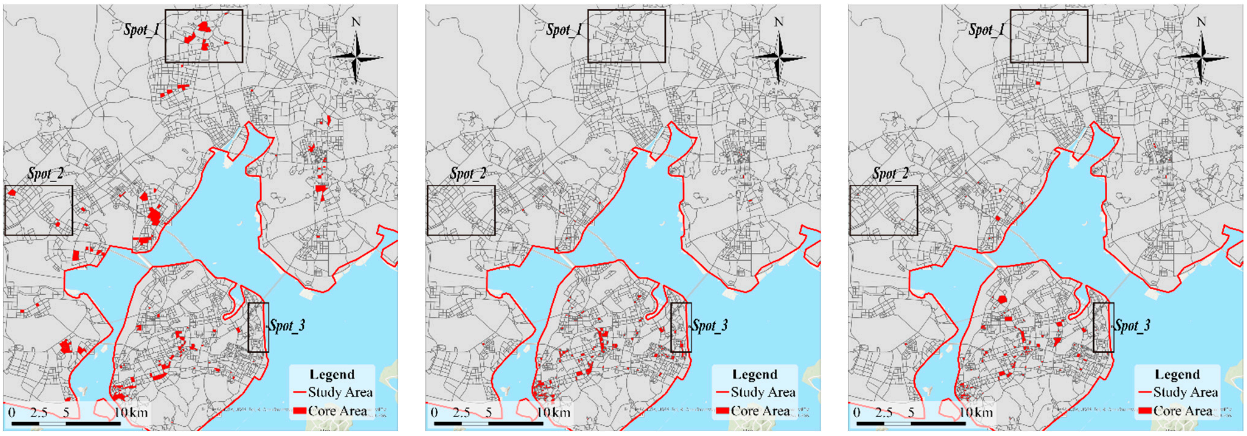

**Figure 7.** UCDs detected by different growth thresholds: (**a**) growth threshold of 0.92, (**b**) growth threshold of 0.94, (**c**) growth threshold of 0.96, and (**d**) growth threshold of 0.98.

The ambiguity of the boundaries of commercial districts stems from the heterogeneity in their development and geographical conditions. The ambiguity of commercial district boundaries was explored by combining the commercial district structures presented under different growth thresholds and those recorded in previous studies. Generally, the economy on Xiamen Island is more developed than outside of the island and has been developed for a longer period; therefore, commercial districts within the island are extremely well connected and are usually identified as large commercial districts at the scale of Xiamen City. Zhongshan Road commercial district (*UCD*_1-1-5), located near Gulangyu Island (a famous tourist attraction), is one of the most prosperous areas in Xiamen, whose commercial value is determined by the tourist resources it benefits from. Therefore, Zhongshan Road commercial district was separated from the island commercial districts when the growth threshold was increased to 0.98. Similarly, the Metro shopping district (*UCD*_1-1-2) and Huli pedestrian street commercial district (*UCD*_1-1-6) were also divided, both of which are located at the edge of Xiamen Island and near communities and factories. The Hexiang West commercial district, Xiamen Railway Station-Fushan commercial district, and Lianhua-SM commercial district were identified as commercial districts (*UCD*_1-1-1) owing to their compact transportation connections and proximity to the geometric center of Xiamen Island. All of these UCDs gradually expand their spatial extent with a decrease in the growth threshold, which is exactly what we perceived to be a fuzzy boundary. UCDs do not have exact boundaries; depending on the growth thresholds, we can identify the circles of the decreasing functionality of UCDs.

### *5.3. Comparison of UCDs Detection Results between Different Strategies*

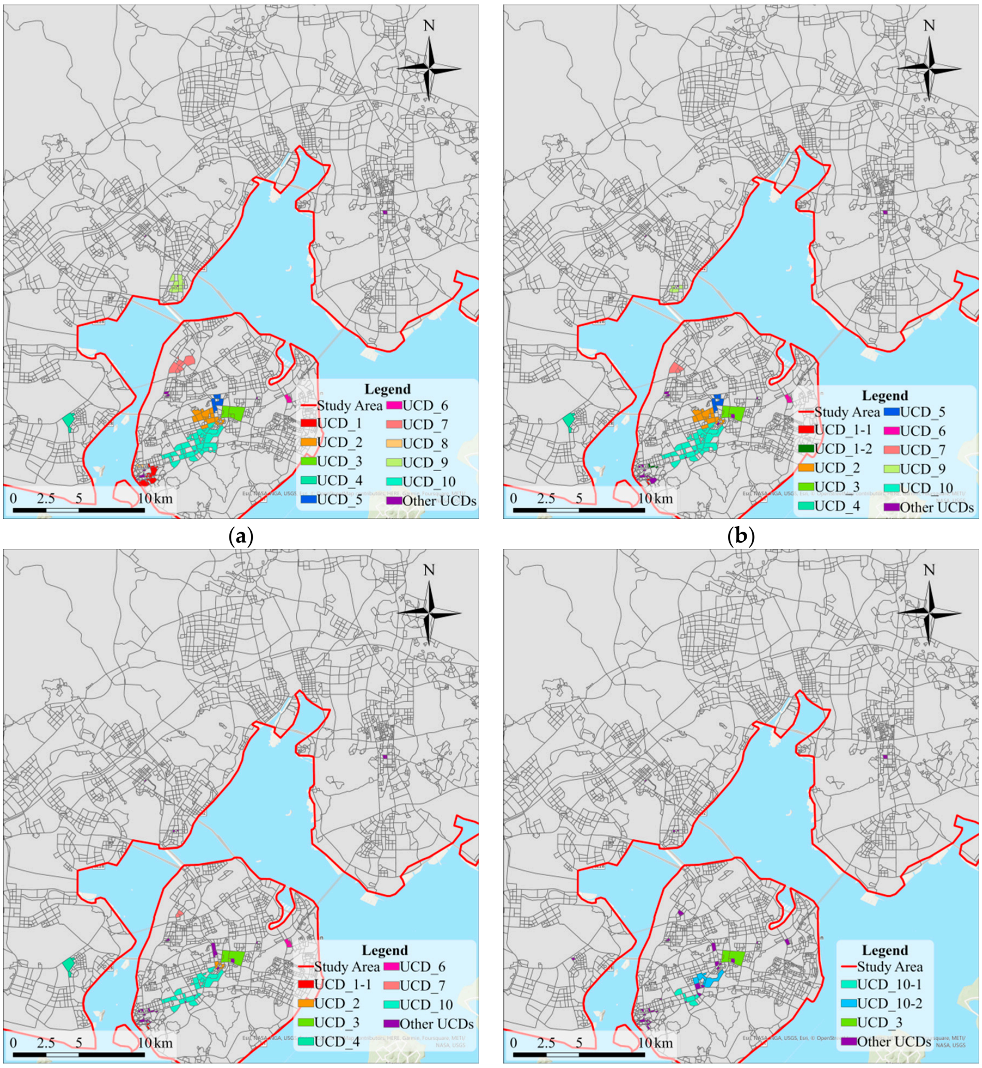
To validate the need for multisource data fusion, Figure 8 compares the results of the hotspot analysis using single and multiple data. Many hotspot districts were detected based on POIs rather than population heat data, thereby revealing their differences. Specifically, some significant differences were labeled with numbers, including *Spot*_1, *Spot*_2, and *Spot*_3. *Spot*_1 had a high concentration of POIs, but no clustered population traffic. By reviewing local information and maps, there were no large commercial spaces where the spatial distribution of commercial spaces was scattered, which may lead to a low volume of people and prevent the formation of sizeable urban commercial districts. Similarly, *Spot*_2 was clustered in Figure 8a, but not clustered in Figure 8b. This is because dense POIs were detected around the housing communities. *Spot*_3 was located on Guanyin Mountain, an area recently developed and constructed by Xiamen, where important commercial facilities, such as the International Convention and Exhibition Center, Tennis Center, and Yacht Club, have been established. This already occurs in the commercial district reports, but its development is not yet mature, especially when supporting small equipped commercial facilities without a gathering of commercial POIs, which fail to attract residents. These comparisons demonstrate that the integrated consideration of multiple data features can effectively avoid the impact of biased data from a single data source on the identification of urban core commercial areas.

Additionally, Figure 9 shows the results considering only the attributes of blocks instead of adjacent blocks to validate the need for considering adjacent blocks. The frequency distribution of similarity is plotted in Figure 10 to identify the range of growth thresholds. A similar distribution, as shown in Figure 10, allowed us to set the same growth threshold. Many UCDs with small areas were detected on Xiamen Island (Figure 9), indicating that considering only the features of the block itself without adjacent features will result in the fragmentation of the detection results; the continuity of the functional space cannot be well portrayed. A high commercial density was not always uniformly present in UCDs owing to the presence of vegetation and mixed commercial and residential areas, among others. Growth thresholds can prematurely interrupt the growth of core commercial areas owing to differences between blocks when only the characteristics of the block itself are considered.




(a) (b) (c)

**Figure 8.** Results of the core commercial area detection using single data: (**a**) hotspot analysis results of POIs, (**b**) hotspot analysis of human traffic in the peak work period, and (**c**) hotspot analysis of human traffic from 8 to 10 pm.




(a) (b) (c) (d)

**Figure 9.** UCDs detected without considering the features of adjacent blocks. (**a**) Growth threshold of 0.92, (**b**) growth threshold of 0.94; (**c**) growth threshold of 0.96, and (**d**) growth threshold of 0.98.

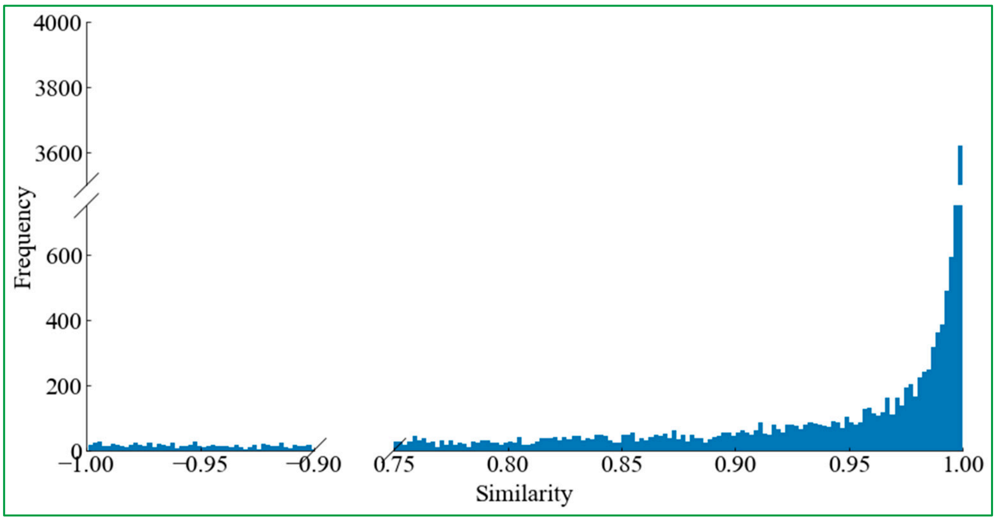



**Figure 10.** Frequency distribution of adjacency similarity (without weighting adjacent blocks).

## 6. Discussion

The perception of UCD spatial structure provides fine details for urban commercial development planning. Firstly, the results of core commercial areas detection reveal the planar pattern of commercial distribution in Xiamen. The southwestern part of Xiamen Island is still an important commercial center area, which is consistent with the study of Chen [55] and Huang [56]. The spatial pattern of the commercial areas along the periphery is different from that of other cities with a circular commercial distribution that decreases from the center outward, such as Kraków [57] and Nanjing [58]. It may relate to the urban characteristics of Xiamen, where tourism is an important industry. Famous attractions such as Gulangyu Island and Nanputuo Temple are located in the southwestern part of Xiamen Island, which attracts a large number of commercial resources. In addition, the emergence of core commercial areas outside the island indicates that the imbalance in commercial distribution, with excess resources on the island and insufficient resources outside the island [59], is gradually improving. Secondly, the spatial extent of each UCD is quantified through the regional growth strategy with multiple growth thresholds, which provides a solution to recognize the ambiguity of geographical entities. The fuzziness of a geographic entity is reflected in the fact that its boundary may not be a clear line demarcation but a two-dimensional transition region [60]. The geospatial extent of the same UCD presented at different growth thresholds simulates the hierarchical nature of human perception in the spatial extent of geographic entities, and the two-dimensional transition region was revealed in this way. Thirdly, the joint relationships between UCDs were revealed, which facilitates the government to promote cooperation and win-win among UCDs. Several small UCDs can be jointly considered as one large UCD due to their spatial proximity and close communication, and similar results occur with different bandwidths of kernel density estimates [7].

In addition, multi-source data fusion provides methodological support for the accurate identification of UCDs. Point-of-interest data has demonstrated great value in providing information on the distribution of UCDs in cities, but the dynamic information it can reveal is limited [25]. Similarly, while population heat data provides information on the dynamic dynamics of urban space, it lacks information on functional attributes [61]. In commercially prosperous areas, such as Xiamen Island in Figure 8, the distribution of commercial areas and the distribution of population heat data largely match, which is consistent with the findings of Huang [56]. However, the distribution of population heat data and point-of-interest data outside Xiamen Island(in Figure 8) shows significant differences, indicating that the degree of construction and population vitality are not synchronized in the initial stage of commercial areas, and a single data source cannot reflect their real commercial intensity.

Although our approach is effective for detecting UCDs, experimental analysis has identified some limitations that must be addressed in future studies. First, commercial areas may exhibit functional differences depending on how they form and their geographical location. Their functional semantics must be further understood for urban planning purposes. Second, our results may be limited by the accuracy of the population heat data; more data must be collected to compensate for errors caused by data serendipity. Thirdly, more ground truth data such as survey data need to be collected to quantify and compare our results.

## 7. Conclusions

Given that the detection of UCDs is affected by bias from a single data source and boundary ambiguity, this study proposed a new framework for discovering UCDs by fusing POI data and population heat data with adjusted cosine similarity and region growth algorithms. The results at different growth thresholds revealed the inconsistent development and uneven spatial distribution of UCDs. The connection between commercial districts was also revealed by the split-merger relationship under different thresholds. Most UCDs in Xiamen are concentrated on Xiamen Island. The development of UCDs on the island is relatively stable, i.e., different growth thresholds have less of an impact on their scope. This finding suggests that the government should focus more on the development of off-island commercial districts. Additionally, the comparison between UCD detection without considering adjacent blocks and using only a single data source validated the effectiveness of the strategy proposed in this study.

**Author Contributions:** Conceptualization, Bingbing Zhao and Huimin Liu.; methodology, Bingbing Zhao and Xiao He; data curation, Xiao He; writing—original draft preparation, Bingbing Zhao; writing—review and editing, Baoju Liu, Jianbo Tang, and Huimin Liu.; supervision, Huimin Liu and Min Deng. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by grants from the National Natural Science Foundation of China (42101471, 42171441) and the Hunan Provincial Natural Science Foundation of China (No. 2021JJ40727, No. 2022JJ30701).

**Data Availability Statement:** Publicly available datasets were analyzed in this study and their links were shared in the manuscripts.

**Acknowledgments:** This work was carried out in part using hardware and/or software provided by the High Performance Computing Center of Central South University.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Duan, X.; Hu, Q.; Zhao, P.; Wang, S.; Ai, M. An Approach of Identifying and Extracting Urban Commercial Areas Using the Nighttime Lights Satellite Imagery. *Remote Sens.* **2020**, *12*, 1029. [CrossRef]
2. Wang, F.; Niu, F. Urban Commercial Spatial Structure Optimization in the Metropolitan Area of Beijing: A Microscopic Perspective. *Sustainability* **2019**, *11*, 1103. [CrossRef]
3. Yaguang, S. Development and Characteristics of Central Business District Under the Philosophy of Health. *Procedia Eng.* **2011**, *21*, 258–266. [CrossRef]
4. Huang, D.; Liu, Z.; Zhao, X.; Zhao, P. Emerging Polycentric Megacity in China: An Examination of Employment Subcenters and Their Influence on Population Distribution in Beijing. *Cities* **2017**, *69*, 36–45. [CrossRef]
5. Alperovich, G. Density Gradients and the Identification of the Central Business District. *Urban Stud.* **1982**, *19*, 313–320. [CrossRef]
6. Şikoğlu, E.; Kaya, H.S.; Arslan, H. Identification of Central Business District (CBD) Boundaries by Space Syntax Analysis and the Case of Elazığ (Turkey). *J. Fac. Archit.* **2020**, *17*, 115–125. [CrossRef]
7. Yu, W.; Ai, T.; Shao, S. The Analysis and Delimitation of Central Business District Using Network Kernel Density Estimation. *J. Transp. Geogr.* **2015**, *45*, 32–47. [CrossRef]
8. Lüscher, P.; Weibel, R. Exploiting Empirical Knowledge for Automatic Delineation of City Centres from Large-Scale Topographic Databases. *Comput. Environ. Urban Syst.* **2013**, *37*, 18–34. [CrossRef]
9. Lu, C.; Wang, H.; Liu, L. Retail District Range Compare Model Construction and Shanghai Districts Demonstration. *Econ. Geogr.* **2015**, *11*, 133–137. (In Chinese)

10. Ponce-Lopez, R.; Ferreira, J., Jr. Identifying and Characterizing Popular Non-work Destinations by Clustering Cellphone and Point-of-Interest Data. *Cities* **2021**, *113*, 103158. [CrossRef]
11. Wang, S.; Liu, Y.; Chen, Z.; Shi, L.; Zhang, J. Representing Multiple Urban Places' Footprints. Available online: Dianping.com Data. *Acta Geod. Cartogr. Sin.* **2018**, *47*, 1105–1113. (In Chinese)
12. Liu, L.; Liu, Z. The Recognition of CAZ in Shanghai Based on Evaluated POI. In *Big Data Support of Urban Planning and Management*; Springer: Cham, Switzerland, 2018; pp. 99–112. [CrossRef]
13. Tan, R.; Zhang, Y. Road Network-Based Region of Interest Mining and Social Relationship Recommendation. *Soft Comput.* **2019**, *23*, 9299–9313. [CrossRef]
14. Liu, X.; Long, Y. Automated Identification and Characterization of Parcels with OpenStreetMap and Points of Interest. *Environ. Plann. B Plann. Des.* **2016**, *43*, 341–360. [CrossRef]
15. Burgess, E. The Growth of the City: An Introduction to a Research Project. In *The Trend of Population*; Routledge: England, UK, 1925; pp. 85–97.
16. Proudfoot, M.J. *Intra-City Business Census Statistics for Philadelphia, Pennsylvania*; United States Department of Commerce, Bureau of the Census: Suitland, MD, USA, 1937.
17. Chen, W. Investigation on CBD Development of Chinese Cities. *City Plan. Rev.* **2003**, *12*, 18. (In Chinese)
18. Lamb, R.F. The Morphology and Vitality of Business Districts in Upstate New York Villages. *Prof. Geogr.* **1985**, *37*, 162–172. [CrossRef]
19. Ye, Q.; Tan, Y.; Zhao, X.; Luo, L.; Chen, N.; Xiang, H. Assessment of Implementation Effect of Commercial Network Planning Based on GIS:A Case Study of Changsha. *Geogr. Res.* **2013**, *32*, 317–325. (In Chinese)
20. Zhang, X.; Zhang, X.; Zhong, E.; Wang, S.; Zhang, J. Study on the Spatial Structure and Driving Force of Urban Areas in Beijing Based on Building Space Structure. *Geogr. Res.* **2013**, *32*, 2055–2065. (In Chinese)
21. Murphy, R.E.; Vance, J.E. Delimiting the CBD. *Econ. Geogr.* **1954**, *30*, 189. [CrossRef]
22. Carol, H. The Hierarchy of Central Functions Within the City. *Ann. Assoc. Am. Geogr.* **1960**, *50*, 419–438. [CrossRef]
23. Chen, W.; Liu, L.; Liang, Y. Retail Center Recognition and Spatial Aggregating Feature Analysis of Retail Formats in Guangzhou Based on POI Data. *Geogr. Res.* **2016**, *35*, 703–716. (In Chinese)
24. Zhou, S.; Hao, X.; Liu, L. Validation of Spatial Decay Law Caused by Urban Commercial Center's Mutual Attraction in Polycentric City: Spatiotemporal Data Mining of Floating Cars' GPS Data in Shenzhen. *Acta Geogr. Sin.* **2014**, *69*, 1810–1820. (In Chinese)
25. Yu, B.; Wang, Z.; Mu, H.; Sun, L.; Hu, F. Identification of Urban Functional Regions Based on Floating Car Track Data and POI Data. *Sustainability* **2019**, *11*, 6541. [CrossRef]
26. He, W.; Li, Z.; Jing, D. Evaluation of Health Level of Land-use Ecosystem Based on GIS Grid Model. *Nat. Environ. Pollut. Technol.* **2020**, *19*, 1475–1482. [CrossRef]
27. Hu, Y.; Han, Y. Identification of urban functional areas based on POI data: A case study of the Guangzhou economic and technological development zone. *Sustainability* **2019**, *11*, 1385. [CrossRef]
28. Luo, S.; Liu, Y.; Du, M.; Gao, S.; Wang, P.; Liu, X. The Influence of Spatial Grid Division on the Layout Analysis of Urban Functional Areas. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 189. [CrossRef]
29. Musiaka, Ł.; Nalej, M. Application of GIS Tools in the Measurement Analysis of Urban Spatial Layouts Using the Square Grid Method. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 558. [CrossRef]
30. Liu, J.; Zhang, Z.; Yang, C.; Xu, S.; Chen, C.; Qiu, A.; Zhang, F. Personalized city region of interests recommendation method based on city block and check-in data. *Acta Geod. Cartogr. Sin.* **2022**, *51*, 1797–1806. (In Chinese) [CrossRef]
31. Fang, W.; Xiaolu, G.; Zening, X. Identification and classification of urban commercial districts at block scale. *Geogr. Res.* **2015**, *34*, 1125–1134. (In Chinese) [CrossRef]
32. Hollenstein, L.; Purves, R. Exploring Place Through User-Generated Content: Using Flickr Tags to Describe City Cores. *J. Spat. Inf. Sci.* **2010**, *1*, 21–48.
33. Hu, Q.; Wang, M.; Li, Q. Urban Hotspot and Commercial Area Exploration with Check-In Data. *Acta Geod. Cartogr. Sin.* **2014**, *43*, 314–321. (In Chinese) [CrossRef]
34. Zhao, Y.; Zhou, Z.; Wang, X.; Liu, T.; Liu, Y.; Yang, Z. CellTradeMap: Delineating Trade Areas for Urban Commercial Districts with Cellular Networks. In Proceedings of the IEEE INFOCOM 2019—IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; Volume 2019, pp. 937–945. [CrossRef]
35. Young, W.J. Distance Decay Values and Shopping Center Size. *Prof. Geogr.* **1975**, *27*, 304–309. [CrossRef]
36. Zhang, Y. Analyses of the Evolution and Causes of Urban Commercial Space Development in Xiamen. *Knowl. Econ.* **2014**, *6*, 132–133. (In Chinese)
37. Li, T.; Han, P. Land Use/Cover Change Detection and Analysis with Remote Sensing in Xiamen City. *Sci. Geogr. Sin.* **2001**, *21*, 537–543. (In Chinese)
38. Oikonomou, M. The Urban Block as a Potential for Sustainable Urban Design. *WIT Trans. Ecol. Environ.* **2015**, *194*, 69–77. [CrossRef]
39. Wang, Y.; Gu, Y.; Dou, M.; Qiao, M. Using Spatial Semantics and Interactions to Identify Urban Functional Regions. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 130. [CrossRef]
40. Zhang, X.; Li, W.; Zhang, F.; Liu, R.; Du, Z. Identifying Urban Functional Zones Using Public Bicycle Rental Records and Point-of-Interest Data. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 459. [CrossRef]

41. Yang, J.; Zhu, J.; Sun, Y.; Zhao, J. Delimitating Urban Commercial Central Districts by Combining Kernel Density Estimation and Road Intersections: A Case Study in Nanjing City, China. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 93. [CrossRef]
42. Zhang, Z.; Xiao, Y.; Luo, X.; Zhou, M. Urban Human Activity Density Spatiotemporal Variations and the Relationship with Geographical Factors: An Exploratory Baidu Heatmaps-Based Analysis of Wuhan, China. *Growth Chang.* **2020**, *51*, 505–529. [CrossRef]
43. Wu, Z.; Ye, Z. Research on Urban Spatial Structure Based on Baidu Heat Map: A Case Study on the Central City of Shanghai. *City Plan. Rev.* **2016**, *4*, 33–40. (In Chinese)
44. Xu, H.; Croot, P.; Zhang, C. Discovering Hidden Spatial Patterns and Their Associations with Controlling Factors for Potentially Toxic Elements in Topsoil Using Hot Spot Analysis and K-Means Clustering Analysis. *Environ. Int.* **2021**, *151*, 106456. [CrossRef]
45. Tobler, W.R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* **1970**, *46* (Suppl. 1), 234–240. [CrossRef]
46. Revol-Muller, C.; Peyrin, F.; Carrillon, Y.; Odet, C. Automated 3D Region Growing Algorithm Based on an Assessment Function. *Pattern Recognit. Lett.* **2002**, *23*, 137–150. [CrossRef]
47. Sato, M.; Lakare, S.; Wan, M.; Kaufman, A.; Nakajima, M. A Gradient Magnitude Based Region Growing Algorithm for Accurate Segmentation. In Proceedings of the International Conference on Image Processing, Vancouver, BC, Canada, 10–13 September 2000; Volume 2000, pp. 448–451.
48. Sarwar, B.; Karypis, G.; Konstan, J.; Reidl, J. Item-Based Collaborative Filtering Recommendation Algorithms. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; pp. 285–295. [CrossRef]
49. Xiamen Bureau of Statistics. National Economic and Social Development Statistical Bulletin of Xiamen in 2020. Available online: http://tjj.xm.gov.cn/tjzl/ndgb/202103/t20210319_2625973.htm (accessed on 5 January 2022).
50. Xiamen Daily News. City's Level to Be Upgraded Again and Cross-Island Development to Be Accelerated Again. Available online: http://dpc.xm.gov.cn/zl/ggzl/202102/t20210226_2521014.htm (accessed on 5 January 2022).
51. Ke, Z. A Brief analysis of the construction of urban commercial space—A case study of Zhongshan Road and Mingfa Commercial Plaza in Xiamen. *Market Mod. Mag.* **2010**, *15*, 68. (In Chinese) [CrossRef]
52. Chen, Y.; Chen, J. The Study of the Internal Structure and the Formation Mechanism in the City-level Commercial Center ——A Case of Zhongshan District. *World Reg. Stud.* **2008**, *27*, 104–114. (In Chinese)
53. Li, W.; Huang, Z. Commercial space structure and format evolution of Xiamen based on POI. *Mod. Urban Res.* **2018**, *4*, 56–65. (In Chinese) [CrossRef]
54. Xin Jing Xiang. Xin Jingxiang Xiamen Business District Analysis and Xianglu Flower City Commercial Pricing Report. Available online: http://jz.720ku.net/shangyeyunying/112981.html (accessed on 10 February 2022).
55. Chen, Y. SWOT Analysis of Xiamen Traditional Commercial Center District. *China Bus. (Jingji Lilun Yanjiu)* **2008**, *12*, 3–5.
56. Huang, Y.; Xu, J.; Gao, Q. Hotspot analysis and vitality assessment of commercial facilities under multi-source data—Xiamen Island as an example. In Proceedings of the 2020 China Urban Planning Conference, Chengdu, China, 25–30 September 2021.
57. Blazy, R.; Łabuz, R. Spatial distribution and land development parameters of shopping centers based on GIS analysis: A case study on Kraków, Poland. *Sustainability* **2022**, *14*, 7539. [CrossRef]
58. Hu, X.; Yang, H.; Zhang, Y. Research on the Spatial Arrangement of Retail Business in the Central City of China's Megacity—A Case Study of Nanjing. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2018; Volume 153, p. 032026.
59. Zhang, Y. Study on the Urban Commercial Space Development and Planning under the Background of Rapid Urbanization-Taking Xiamen City for Example. Master's Thesis, Jimei University, Xiamen, China, 2014. (In Chinese).
60. Hollenstein, L. Capturing Vernacular Geography from Georeferenced Tags. Master's Thesis, Geographisches Institut der Universität Zürich, Zürich, Switzerland, 2008.
61. Shi, P.; Xiao, Y.; Zhan, Q. A study on spatial and temporal aggregation patterns of urban population in Wuhan City based on Baidu heat map and POI data. *Int. Rev. Spat. Plan. Sustain. Devel.* **2020**, *8*, 101–121. [CrossRef] [PubMed]