Spatial Pattern Analysis of Zhuhai Urban Development Based on POI Big Data

Zili Zhao, Xiupeng Zhang(✉), Zhi'ao Zhang, and Xintao Peng

Zhuhai Institute of Urban Planning & Design, Zhuhai, Guangdong, China
zhangxp@zhghy.com

Abstract. This study is to obtain accurately the urban development & evolution characteristics and spatial development pattern information of Zhuhai, to analyze its urban development laws by using spatio-temporal big data, to improve its scientific entity and effectiveness of planning, as well as to help its urban construction and development. It makes research to the population density, catering service facilities and spatial distribution characteristics of work and residence in Zhuhai with Baidu Huiyan data and from the time-space macro analysis of the urban development of the whole city to the specific evaluations of people, land and things in the dimensions of districts, counties and streets. It aims to explore the internal texture, spatial pattern and evolution law of Zhuhai urban development. It proves that Baidu Huiyan big data under scientific mining and processing can not only provide a more dynamic perspective and method for urban space research, but also provide a good analysis and expression way for the facility planning and evaluation in the future urban planning.

Keywords: POI big data · Baidu Huiyan · spatial pattern · Zhuhai city

1 Introduction

Urban spatial structure has always been a key research object of urban planning, which has both physical and social attributes. Previous studies on urban spatial structure have focused on its physical attributes, while the current ones are gradually shifting to planning informatization [1]. Hence, the explicit physical characteristics have been difficult to reflect the actual operation state of urban space, instead the implicit social attributes, including the actual distribution of population and economy within the city, can better reflect the essence of urban spatial structure although it cannot be felt visually. Because of the increasing scale and complexity of urban space, it is time-consuming to analyze urban spatial structure through traditional urban research and investigation methods, and cannot provide real-time data [2–4]. In contrast, the use of artificial intelligence, "BICM" (big data, Internet of things, cloud computing, mobile Internet technology) and other technical means to study the internal texture, spatial pattern and evolution law of urban space-time development has the advantages of instant and high efficiency, low cost as well as wide coverage. Therefore, it has great application prospects in urban planning research. Among which, catering service is an important part of urban life, which is of

G. Guan et al. (Eds.): ICBDSS 2022, AHCS 8, pp. 794–807, 2023.
https://doi.org/10.2991/978-94-6463-064-0_82

great significance to understand the urban spatial structure by extracting the hot spots of urban catering service and identifying its spatial distribution pattern [5]. In view of this, based on spatio-temporal big data, this study uses reasonable research methods to examine and analyze the spatial distribution pattern of urban development in Zhuhai in recent years, and evaluate its planning and implementation. It is to provide a reference method for the extraction of urban spatial characteristic elements and for the vertical and horizontal structure analysis of urban spatial pattern, and has enlightenment significance for the research of urban spatial pattern in different themes based on point elements.

2 Overview, Data Sources and Research Methods of the Studied Areas

2.1 Overview of the Studied Areas

This study selects Zhuhai as the research area, including Xiangzhou district, Doumen disrict and Jinwan district. Its geographical coordinates are between 21°48′N and 22° 27′N, 113°03′E to 114°19′E. It faces Hong Kong across the sea in the East, Macao in the south, Xinhui and Taishan city in the West and Zhongshan in the North. The total area of the city covers 1711km^2, its GDP is about 348.194 billion yuan, and the general public budget revenue reaches 37.9 billion yuan with an average annual growth of 8.4%. The total investment in fixed assets is totaled 911.3 billion yuan, the total retail sales of social consumer goods totaled 453.7 billion yuan, the total import and export volume totaled 1462.9 billion yuan, the actually utilized foreign capital totaled 12.1 billion dollars, the per capita disposable income of residents reached 55936 yuan with an average annual growth of 9.1%. By the end of 2020, the permanent resident population has exceeded 2 million.

2.2 Data Sources

The research data of this study comes from Baidu map POI data captured through the network in March 2021 and vector data such as terrain, road and river downloaded from BiGMAP map. Because POI data contains the spatial locations and attributes of most entity objects in the city, which are their abstract representations on the map, so it considers that POI data contains all research objects in urban space. After cleaning, correcting and checking the acquired data, we obtained 117545 pieces of POI data. Combined with the POI classification system of Baidu Huiyan map and the different functional classification systems of cities, the POI data is divided to the following six categories: life service, business, finance and insurance, public service, leisure and entertainment and residence (Table 1).

Table 1. Types of POI data

Category	Sub-category
Public facilities	Sports facilities, medical facilities, public green space, education and training, government institutions
Transportation facilities	Transportation facilities
Tourism facilities	Scenic spots
Commmerce facilities	Factories and mines, companies and enterprises, financial facilities and office buildings
Business facilities	Catering, shopping, hotel, life service, leisure vacation, leisure and entertainment
Neighbourhood	Neighbourhood

2.3 Research Methods

1) **Nuclear Density Method**

The basic idea of nuclear density analysis is that there is a density at any point in the study area, not just at the event point. This method calculates the density of spatial point and of line elements in their surrounding neighborhood, simulates continuously the density distribution, and uses the nuclear density value of each grid in the image to reflect the distribution characteristics of spatial elements [6]. This study uses the nuclear density analysis method to explore the spatial distribution characteristics of POI data in Zhuhai, estimates the surrounding density according to the POI nuclear density value in each grid, and compares the nuclear density analysis results under different search radii in order to select the best search radius suitable for this study. In the study area R, the nuclear density estimation model takes any point S as the center, and estimates the density value of the target point through the set bandwidth R, its density is jointly determined by the number of event points within the bandwidth and the distance from the event point to the target point. The simplified process of calculating a nuclear density value is as follows: firstly, the study area is meshed and the number of event points in each grid is counted; secondly select the range based on the bandwidth R; finally, the density value of nuclear K is calculated according to the nuclear density formula [7]. The calculation formula is:

$$\mathrm{f(x,\,y)} = \frac{3}{nr^2\pi}\sum_{i=1}^{n}\left[1 - \frac{(x-x_i)^2 + (y-y_i)^2}{r^2}\right]^2$$

Where: f (x, y) is the density of the center point P (x, y) of the estimated target grid unit; r is the bandwidth. n is the number of samples within the bandwidth. x_i, y_i is the coordinate of sample point I. x,y is the center point coordinate of the estimated target

grid unit. $(x\text{-}x_i)^2+(y\text{-}y_i)^2$ is used to estimate the square of the Euclidean distance between the center point of the target grid and the grid sample i within the bandwidth.

2) **Average Nearest Neighbor Analysis**

Process of average nearest neighbor analysis: calculate the observation distance between each POI and its nearest POI, and calculate the average value of all nearest distances. If the average observation distance of one kind of POI is smaller than the expected average distance of the assumed random distribution, this kind of POI belongs to aggregation distribution, on the contrary, it belongs to dispersion distribution [8, 9]. This study uses the average nearest neighbor tool of ArcGIS 10.7 software for analysis, which results include five values: average observation distance, expected average distance, nearest neighbor index, Z score and P value. The smaller the R value, the higher the degree of aggregation. Since the null hypothesis in the average nearest neighbor statistical tool is that the input elements are randomly distributed, it is necessary to judge whether to reject the null hypothesis at a certain significance level according to the Z score and P value [13]. The calculation formula is as follows:

$$R = d_i/d_e$$

$$d_e = 0.5/\sqrt{N/A}$$

$$Z = (d_i - d_e)\sqrt{\frac{N^2}{A}}/0.26136$$

Where: A is the area of the study area; N is the total number of POI; When $|Z| > 2.58$ and $P < 0.01$, reject the null hypothesis (the calculation results of Z and P in this paper prove the aggregation distribution of POI data).

3) **Moran Index (MOran's I)**

Spatial autocorrelation indexes are the ones to measure the relationship between two adjacent elements or two elements with some spatial relationship [10–12]. There are two commonly used indexes: Moran index and Gary index. In this paper, Moran's index (Moran's I) is mainly used, and the value range is between (-1, 1). Moran's $I > 0$ indicates spatial positive correlation. The larger its value is, the more obvious the spatial correlation is. Moran's $I < 0$ indicates spatial negative correlation. The smaller the value is, the greater the spatial difference is. Otherwise, Moran's $I = 0$, and the space is random. Its calculation formula is as follows [13]:

$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}(x_i - x_m)\left(x_j - x_m\right)/\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}}{\sum_{i=1}^{n}(x_i - x_m)^2/n}$$

Where: x_i is the value of cell i, x_j is the value of cell j, x_m is the average value of the grid cells; W_{ij} is the coefficient; n is the total number of grid cells. If j is one of the four cells directly adjacent to i, the coefficient W_{ij} is one. If other cells or cells have no data, the coefficient W_{ij} is zero.

4) **Hot Spot Analysis**

The hotspot analysis tool can calculate Getis-Ord Gi* and make statistics for each element in the data set, which principle is to know the spatial clustering position of high-value or low-value elements through the obtained Z score and P value [14]. The goal of hot spot analysis is to identify the regions with statistical significance clustering and to analyze the spatial correlation of these factors. Getis-Ord General G and Getis-Ord G_i* indexes are used in this study to measure the global and local spatial characteristics, structural patterns and spatial clusters of Zhuhai City [15]. The final research results are shown by the spatial distribution of hot spots and cold spots.

3 Empirical Analysis

3.1 Analysis on Spatial Distribution Characteristics of Pedestrian Flow Density in Zhuhai

The construction of "smart city" will in urgent need of the rational urban planning method as "flow" determining "shape". Mining the flow elements under the surface form of the spatial distribution pattern of people flow density is very important for formulating and summarizing the urban development model and method [16]. Baidu Huiyan thermal map is a visualization product of Baidu big data, which starts from the geographical location data of mobile phone users on LBS Platform, and finally presents users with different degrees of crowd agglomeration through certain spatial expression processing.

This study investigated respectively the thermal degrees reflected in the thermal map of Baidu map on weekdays and weekends. The study shows that citizens' activities show periodic changes in units of days to the large extent. The number, area and location of heat intensity areas at all levels change significantly with the passage of time in a day. At the same time, there are certain differences in the distribution of people on weekends (Saturday and Sunday) and weekdays (Monday to Friday). There are two main factors causing this change: one is the change of spatial distribution caused by the movement of mobile client users in cities, and the other is the change of the proportion of mobile client users in the population over time (Figs. 1 and 2).

The pedestrian flow density in Zhuhai is divided into three levels based on the population thermal value and real-time passenger flow data of Baidu map Huiyan. The pedestrian flow density in Xiangzhou district in the East is highest at the first level, which clearly shows that the central district still has the strongest attraction for pedestrian flow. Tangjiawan Gaoxin district in the North and the ecological new town in the West, including Baijiao town (Binjiang neighborhood), Jing'an town (Binjiang neighborhood), Hongqi Town (aviation neighborhood) and Sanzao town, are at the second level where some strong pedestrian concentration areas have also been formed. Pingsha town (Fushan neighborhood), Nanshui town (Haigang neighborhood), Doumen town and Hengqin new district are at the third-level pedestrian density gathering areas (see Fig. 3 for details).

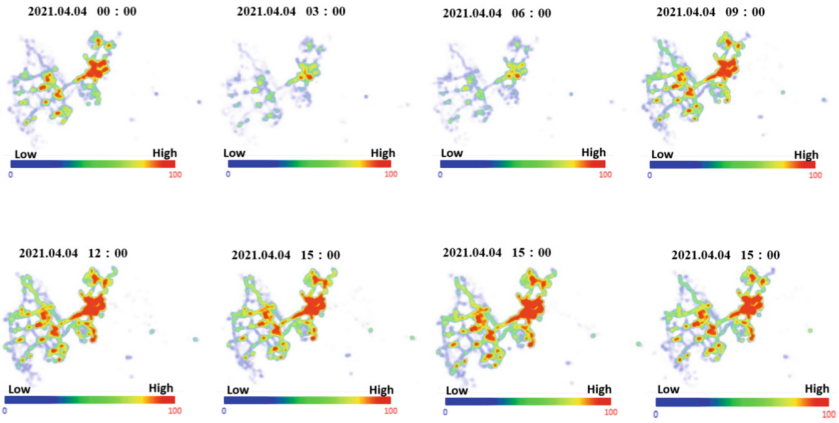

Fig. 1. Real time passenger flow heat map data (part) on the weekend (April 4)

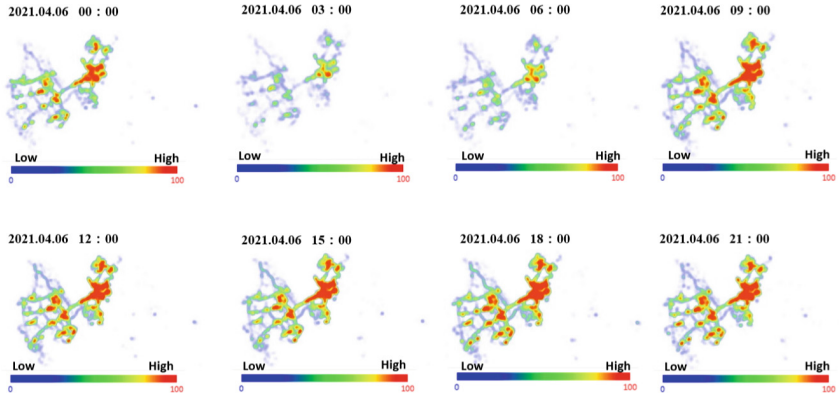

Fig. 2. Real time passenger flow thermodynamic diagram data (part) on the working day (April 6)

3.2 Analysis on Spatial Distribution Characteristics of Catering Service Facilities in Zhuhai

Catering service is an important part of urban life. Extracting the hot spots of urban catering service and identifying its spatial distribution pattern is of great significance to understand the urban morphological structure. The formation of the spatial density characteristics and spatial distribution pattern of catering service POI points are the result of the gradual development and evolution of a city under the joint action of many factors. The formation of the final distribution pattern corresponds to its main influencing factors and formation mechanism [17]. From this perspective, the spatial pattern of urban catering service hotspots can be used to scientifically explore the main driving factors of its pattern formation from a more detailed level. These factors may include historical and cultural driving, policy and system driving, social public opinion driving, planning

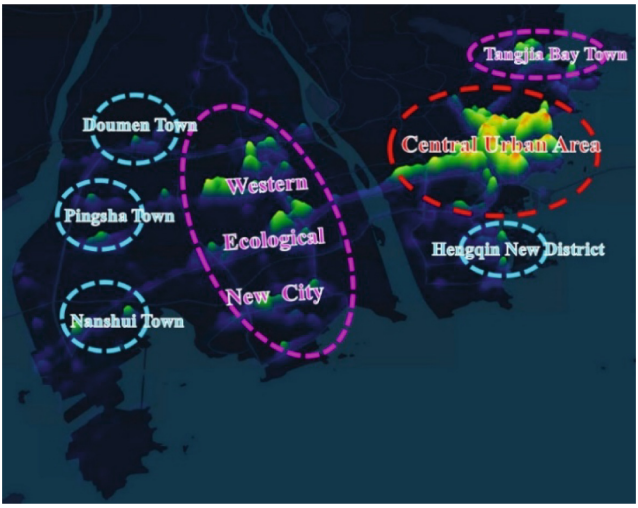

Fig. 3. Analysis of spatial structure of pedestrian flow density in Zhuhai

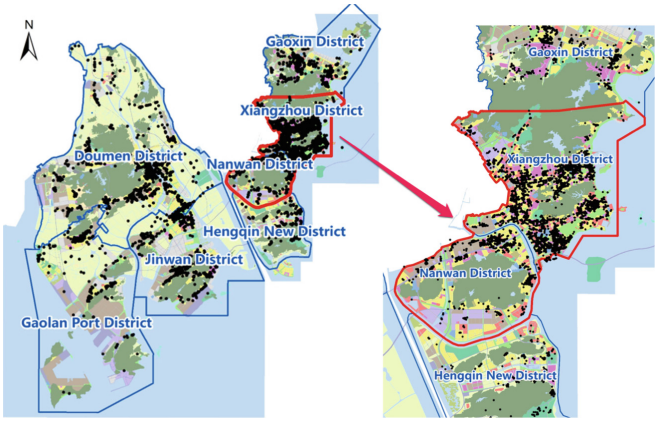

Fig. 4. Spatial distribution of catering service facilities in Zhuhai

strategy driving and so on to further explore the spatial distribution characteristics of urban structure.

The overall overview of the study area and the distribution of catering service POI data are shown in Fig. 4.

After the nuclear density analysis of 14426 pcs of POI catering data in Zhuhai, it concludes that the central urban area has the most of catering service facilities. It's proved by the agglomeration degree of the below areas are of the highest level: Gongbei block, Cuixiang block, Shishan block, Jida block and Nanping block of Nanping town, Shier village and Beishan area. The areas are of a relatively high level as below: Tangjia village, Guantang village and Shangce village of Tangjiawan town in Gaoxin district; Chengyu village, Laijia village, Dongan village of Baijiao Town in Doumen distrcit,

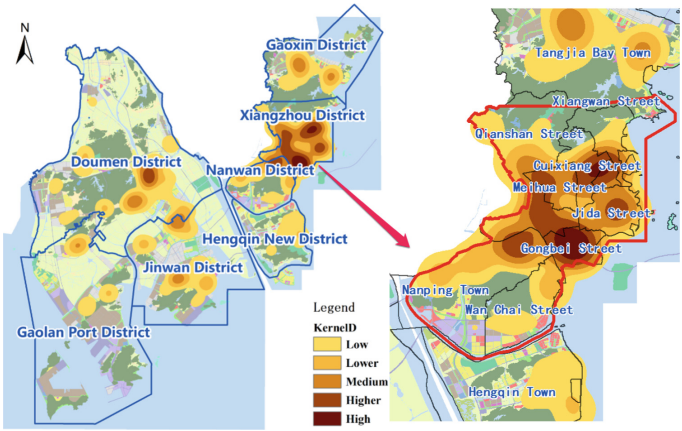

Fig. 5. Nuclear density analysis of catering service facilities in Zhuhai

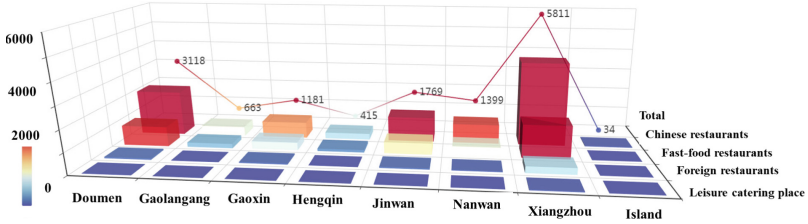

Fig. 6. Overall distribution of different types of catering service facilities in each planning area of Zhuhai

Xintang village and Xipu village of Jinan town in Doumen District; Sanzao community, Caotang community and Yuyue village of Sanzao town in Jinwan district. The catering service facilities in Gaolan Port area and Hengqin new area are relatively few and lack (Fig. 5).

The catering service facilities in Zhuhai are mainly Chinese restaurants, accounting for about 67% of the total, followed by fast food restaurants, accounting for about 29% of the total. From the perspective of spatial distribution, the total number of catering service facilities in Xiangzhou urban area is the most; the number of islands, Gaolan Port area and Hengqin new area is a few, and the maximum values of different types of catering service facilities all appear in the Xiangzhou district. The main reason is that Xiangzhou urban area has almost formed a relatively complete urban comprehensive functional area after the reform and opening up for more than 40 years (Figs. 6 and 7).

3.3 Analysis on the Spatial Distribution Characteristics of Job and Housing in Zhuhai

Based on the big data of Baidu Huiyan map points of interest (POI), this paper makes an empirical study on the living and employment space in Zhuhai, and analyzes the matching situation of living and employment space at the overall level and zoning scale in Zhuhai

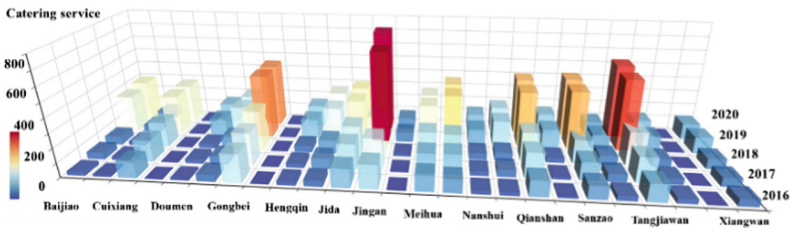

Fig. 7. Overall distribution of catering service facilities in Zhuhai from 2016 to 2020

Table 2. Calculation of job and housing balance of each block (town) in Zhuhai

Name	Job and housing deviation index	Name	Job and housing deviation index
Wanzi block	0.55	Tangjiawan town	0.81
Xiangwan block	0.58	Meihua block	0.87
Jin'an block	0.64	Jida block	1.1
Hengqin town	0.64	Doumen town	1.14
Sanzao town	0.68	Nanshui town	1.19
Gongbei block	0.7	Qianshan block	1.26
Hongqi town	0.71	Cuixiang block	1.27
Shishan block	0.73	Qianwu town	1.31
Baijiao town	0.75	Nanping town	1.68
Pingsha town	0.77	Lianzhou town	2.25

by constructing the employment and residence deviation index model. This study refers to the calculation method of employment and residence deviation index, adopts the ratio of company data and residential area data in Zhuhai, as well as standardizes the data to calculate the employment-residence deviation index of each block (town) in Zhuhai (Table 2).

The research shows that the spatial distribution of job and housing in Zhuhai is almost balanced. From the perspective of Zhuhai, the separation distance of job and housing space is in a reasonable range [18–20]. Compared with some large cities in China, the average commuting time in Zhuhai is relatively short, more than 70% of residents need about 30 min to go to work, and there is no obvious separation between job and housing space. Specifically, the deviation index of job and housing of Wanzi block, Xiangwan block, Jing'an town, Hengqin town, Sanzao town, Gongbei block, Hongqi town, Shishan block, Baijiao town, Pingsha town, Tangjiawan town and meihua street is less than 1, indicating that the matching degree of job and housing space is relatively high; the deviation index of job and housing in Jida block, Doumen town, Nanshui town, Qianshan block, Cuixiang block, Qianwu town, Nanping town and Lianzhou town is greater than

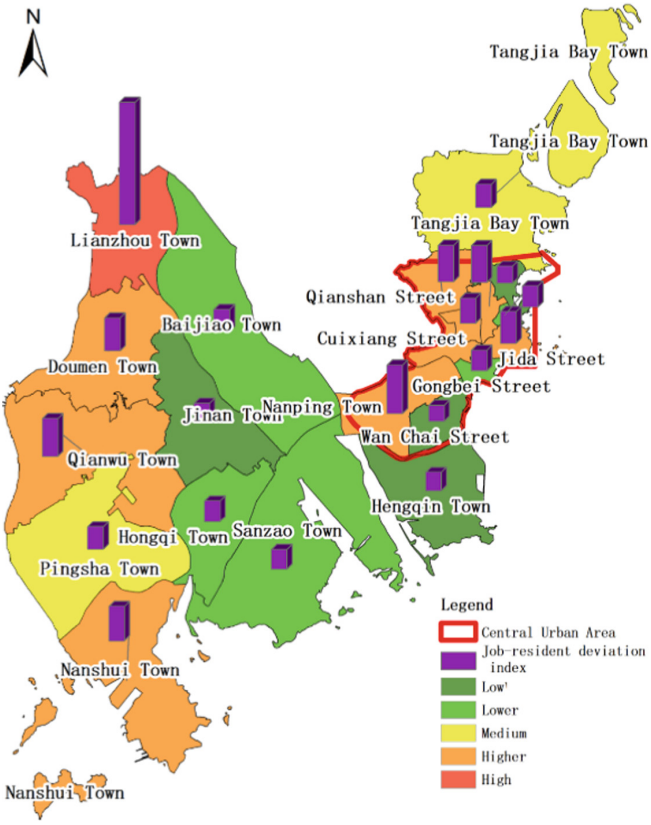

Fig. 8. Analysis on the balance of job and housing in various towns and blocks in Zhuhai

1, indicating that the spatial matching degree of job and housing is relatively low. (see Fig. 8 for details).

4 Discussion

4.1 Comparative Analysis of Spatial Distribution of POI Big Data and General Planning Spatial Structure

The urban master plan of Zhuhai (2001–2020) proposes a progressive and intensive group spatial structure of "1 main urban area - 5 new urban areas - 7 central towns". It finds that the spatial distribution characteristics of POI catering service facilities data are almost consistent with the spatial structure of the previous master plan. Spatially, except that Sanzao town and Qianwu town have new agglomeration areas that can be added as new central towns, the other agglomeration areas are completely consistent with the progressive and intensive cluster spatial structure of "main urban area - new urban area - central town", and also show the obvious characteristics of "one core, two centers and six areas" and the multipolar cluster structure of three special functional areas (see Fig. 9 for details).

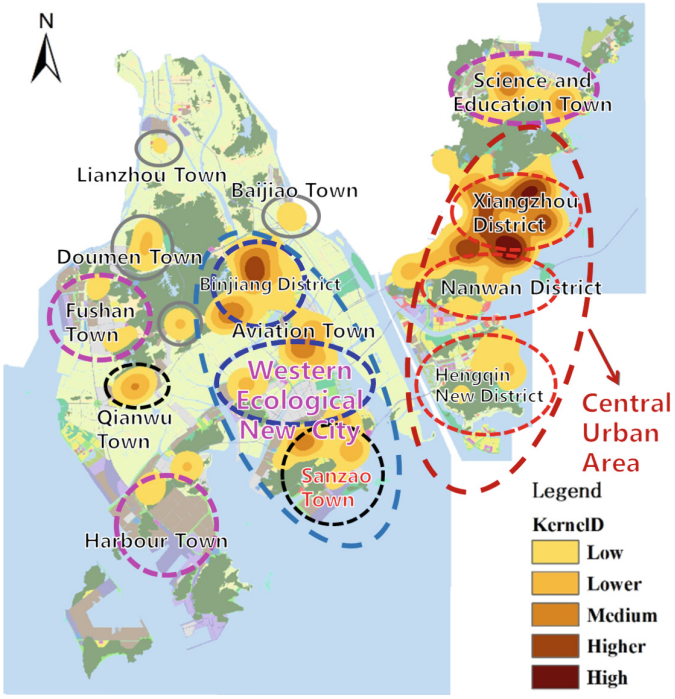

Fig. 9. Spatial structure analysis of catering service facilities in Zhuhai

4.2 Hotspot Analysis Based on Average Nearest Neighbor Index and Spatial Correlation Index

Through the analysis of the average nearest neighbor index of 117545 pcs of POI data in Zhuhai, it concludes that there is a large gap between the average observation distance and the expected average distance and the average observation distance is far less than the expected observation distance. It shows that the catering service POI in Zhuhai is of spatially concentrated distribution. The nearest neighbor index is 0.1230, far less than 1 and the degree of aggregation is large. When measuring the spatial autocorrelation of 117545 pcs of POI data in Zhuhai, the Moran index I calculation method is used. The results show that the Moran index I is 0.187920, greater than 0, indicating that there is a positive correlation in the spatial distribution of POI data in Zhuhai. With the aggregation of spatial distribution positions (distances) within the range, the correlation becomes more and more significant.

This study carries out hot spot analysis based on the POI data density of different administrative regions in Zhuhai. According to the scores of G statistics, the hot spot analysis results are divided into five levels as belows: core cold spot area, edge cold spot area, edge hot spot area, sub core hot spot area and core hot spot area. The results show that the POI hot spots in Zhuhai are mainly distributed in layers of high in the East, slightly low in the middle and the lowest in the West. The core hot spots and sub core hot spots are highly concentrated in Xiangzhou urban area and Nanwan urban area in

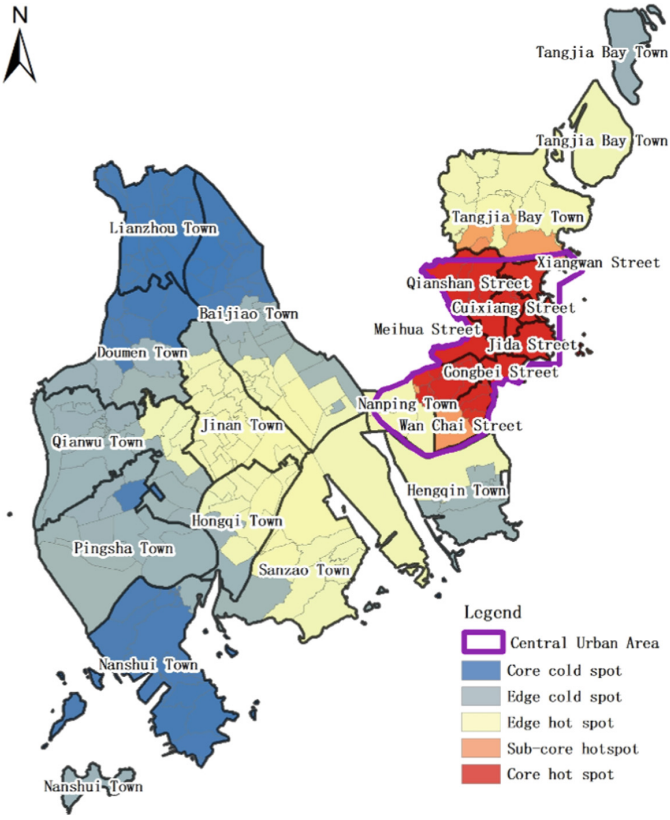

Fig. 10. Hot spot analysis of POI big data in Zhuhai

the central city. Marginal hot spots are mainly distributed in Tangjiawan town, Nanping town, Hengqin town, Sanzao town, Hongqi town and Jing'an town. The core cold spot area and edge cold spot area are mainly distributed in Lianzhou town, Doumen town, Qianwu town, Pingsha town, Nanshui town and Hengqin town on the western edge. It wholly presents the characteristic of cluster distribution (Fig. 10).

5 Conclusion

Baidu Huiyan big data has the advantages of timeliness, integrity and representativeness, which is of great significance to apply it to the research of planning informatization. It can provide urban researchers with an unprecedented new perspective, so that designers can use a dynamic perspective subdivided into hours or even minutes to see the crowd activities and the use of urban space in the city. From this perspective, people can see how the crowd concentration degree, dispersion and the location of crowd concentration change in a day or even a shorter period, and when the use intensity of urban space with different locations and functions is the highest. The organization of job and housing

space plays an important role in urban development. Therefore, exploring the development characteristics and matching relationship of job and housing space is not only conducive to people's understanding of the law of urban development, but also conducive to guiding the future spatial development of the city. Although this paper constructs and makes empirical analysis on the methods that can express the analysis of urban spatial structure, such as nuclear density analysis, average nearest neighbor analysis, spatial autocorrelation analysis and hotspot analysis, there are still many problems. In future research, we will focus on the analysis of urban morphological structure characteristics of more types of elements in order to provide a practical analysis model for facility planning and evaluation in urban planning.

Acknowledgment. This work is funded by Supported by Major State Basic Research Development Program (Item Number: 2016YFB0502300).

References

1. Wenhao Yu, Tinghua AI, Pengcheng Liu, et al. Network kernel density estimation method for facility POI distribution hotspot analysis [J]. Journal of Surveying and mapping, 2015,44 (12): 1378–13831400.
2. Kunyang Wang, Bowen Qiao, Jing Zhang. Current situation of urban system in Beijing, Tianjin and Hebei from the perspective of big data -- Taking POI data and mobile phone data as an example [J]. Journal of Capital Normal University (NATURAL SCIENCE EDITION), 2020,41 (6): 67-75.
3. Zeshuang Zheng. Study on spatial distribution characteristics and influencing factors of job and housing in Zhuhai -- Investigation and analysis based on Gaode map interest points and traffic sensing data [J]. Planner, 2019,35 (4): 51-56.
4. Jiliang Wang. Simulation of spatial distribution of population density in small area based on GIS -- Taking Tongzhou City, Nantong as an example [D]. Hunan: Central South University, 2009.
5. Liyan Wang, Yuchen Dai, zhenpo Wang. Research on housing rental demand preference in new urban areas based on POI big data -- Taking Tianjin Binhai New Area as an example [J]. Regional research and development, 2020,39 (6): 71–76.
6. Yaming Duan, Yong Liu, Xiuhua Liu, et al. Multi center identification of main urban areas of Chongqing Based on POI big data [J]. Journal of natural resources, 2018,33 (5): 788-800.
7. Zijiang Yang, Xiong He, Xin Sui, et al. Analysis of spatial evolution of urban center based on POI -- Taking the main urban area of Kunming as an example [J]. Urban development research, 2019,26 (2): 31 - 35.
8. Li Wei, Huang Zhengdong. Analysis on the evolution of commercial spatial structure and business format in Xiamen based on POI [J]. Modern urban research, 2018 (4): 56-65.
9. Tingting Huang, Xiaolong Li, Dajun Li. Analysis of commercial space agglomeration characteristics and business form evolution in Shenzhen based on POI [J]. Jiangxi science, 2020,38 (6): 909-914.
10. Jing Yang. Identification and extraction of urban business centers and clusters based on POI data -- Taking Nanjing as an example [D]. Jiangsu: Nanjing Normal University, 2018.
11. Lanlan Ying, Weiwei Niu. Evaluation of living convenience of residents in main urban areas of Wuhan Based on POI data [J]. Research on land and natural resources, 2020 (3): 14-17.

12. Gang Li, Lan Yang, Jianxiong He, et al. Study on spatial pattern and spatial relationship of express pick-up points in Xi'an Based on POI data -- Taking rookie post station as an example [J]. Geographical science, 2018,38 (12): 2024-2030.
13. Ruxue Xia, Xiaolan Tang, Gang Hu. Spatial distribution and impact analysis of wetland parks in the Yangtze River Delta Based on POI data [J]. Land and natural resources research, 2020 (3): 17-23.
14. Zhiqiang Wu, Zhongnan Ye. The study of urban spatial structure based on the Baidu map thermodynamic map -- Taking Shanghai central city as an example [J]. city planning, 2016,40 (4): 33–40.
15. Meijuan Wang. Analysis on the change of urban internal spatial structure in Dalian based on the transfer of urban center of gravity [J]. Research on land and natural resources, 2015 (4): 1-3.
16. Yuhuan Wang, Cheng Jin, Jiazhen Du. Analysis of temporal and spatial characteristics of residents' travel hotspots in Chengdu based on multi-source data [J]. Journal of Nanjing Normal University (Engineering Technology Edition), 2020,20 (2): 80-87.
17. Yanyang Wang, Yuzhe Liang, Weiling Luo, et al. Detailed investigation on the current situation of land spatial planning based on high score remote sensing images and poi [J]. Tropical geography, 2020,40 (4): 649-658.
18. Yu Xia, Shubo Yang, Zhonghai Zhang. Spatial hot spot analysis of urban catering industry based on meituan takeout POI -- Taking Nanchang as an example [J]. Jiangxi science, 2020,38 (1): 31-35.
19. Feng Zhen, Bo Wang, Yingxue Chen. Chinese urban network characteristics based on Cyberspace -- Taking Sina Weibo as an example [J]. Journal of geography, 2012,67 (8): 1031-1043.
20. Xiaoyu Xu. Study on temporal and spatial characteristics of urban hot spots based on location check-in and POI data [D]. Shandong: Shandong University of technology, 2018.