Spatial Distribution Dataset of 1598 More Chinese Traditional Villages

Yu, L.[1*]   Ding, Y. Q.[1]   Tang, M. J.[1]   Cao, Q. Y.[1]   Fu, M.[1]   Yang, X. H.[1]
Shen, J. X.[2]

1. School of Architecture, Soochow University, Suzhou 215123, China;
2. Zhejiang College, Tongji University, Jiaxing 314000, China

Abstract: The Ministry of Housing and Urban-Rural Development, the Ministry of Culture, and the Ministry of Finance published the fourth batch of 1,598 traditional Chinese villages in December 2016 after the first released 2,555 traditional Chinese villages (2012–2014). A total of 4,153 villages are currently included in the list of traditional Chinese villages. The spatial distribution dataset of the newly added 1,598 traditional villages were based on the research method of "The spatial distribution dataset of 2,555 Chinese traditional villages" (http://www.geodoi.ac.cn/WebEn/doi.aspx?Id=910). Referring to the Baidu map and the Google Earth image, the coordinate of the spatial position were extracted according to the text prompt of the village name, the geometric center of the village image, or the village location that is easily to recognize. In case that the village's name is unconventional and difficult to identify, measures such as querying relevant information or contacting directly with the village were taken to determine the name and location of the village. When there is no information and image about the village on the map, the coordinates of the upper administrative unit was used as the coordinates of the village. Finally, "The spatial distribution dataset of 1598 traditional villages" was obtained. The dataset is archived in .shp and .kmz formats, consisting of 7 data files with a total data size of 4.50 MB (compressed into 2 files, 258 KB).

Keywords: China; traditional village; directories; spatial distribution

1 Introduction

Traditional village, a polymeric form and carrier of human living space in early days, is a valuable historical and cultural heritage. It spreads all over the world and the geographical feature is obvious. As a necessary space form for life, it is the embodiment of the concept that human beings respect nature and harmoniously coexist with nature. Due to the needs of human activities, there are a large number of traditional dwellings in traditional villages. Traditional dwellings and villages cannot be separated, just like the relationship between

water droplet and pond. Figure 1 shows the residential buildings of Xu'ao village, Xiaocun town, Taishun county, Wenzhou city, Zhejiang province, and Nanjiao village, Nanjiao town, Fangshan district, Beijing, which were investigated by author. They are far apart and both are located in a mountainous area, and the construction of the buildings has common characteristics, such as they both use a large amount of local stone resources and have unique aesthetic standards, which shows the harmonious and orderly spatial pattern of the village. From this, we can see that when we are analyzing and comparing the villages' space, the form of unique residential buildings should not be neglected. By comparing the residential buildings in different villages, we can further sort out the spatial distribution characteristics of the villages in macroscopic view. Liu *et al.*[1] and Kang *et al.*[2] studied the spatial distribution of villages in the 3 batches of 2,555 villages that were published on the Ministry of Housing and Urban-Rural Development, respectively. The newly added 1598 traditional village spatial distribution datasets are the basic data for analyzing and interpreting the geographical location and spatial distribution characteristics of villages, and are reorganizations of the 2555 village spatial distribution dataset[3–4] in the previous period. It can provide more accurate and complete administrative subordination and naming from province's to village's levels for other researches. The village is not only a space field, but also a record of rural work and life scenes that human beings are adapted to nature. Through the systematic expression and comparative application of spatial data, it can show people outstanding rural culture better and also play an important role in promoting the rural revitalization.

Figure 1   Traditional village with distinctive regional characteristic (Left: Xu'ao village, Xiaocun Town, Taishun county, Wenzhou city; Right: Nanjiao village, Nanjiao town, Fangshan district, Beijing)[4]
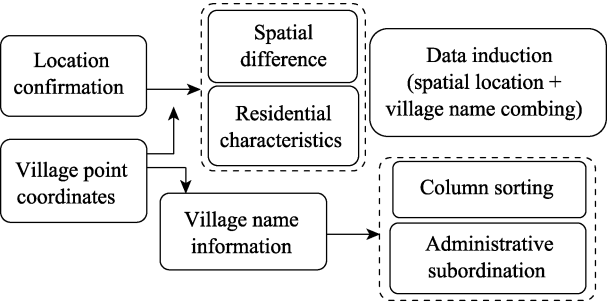
2 Metadata of Dataset

The metadata of "The spatial distribution dataset of 1598 more Chinese traditional villages"[5] is summarized in Table 1. It includes the full name, authors, data format, data size, data files, data publisher, and data sharing policy, etc.

3 Village Space Point Data Processing

It is extremely difficult to obtain data of the spatial points of the villages. The published village list does not show geographical coordinates, and sometimes distinguishing the administrative affiliation associated with the village name is difficult. From the analysis of the village shape in the space, we could know that the village is composed of people, which are both a spatial form and a social form. Within the cognitive range of the village space, except for natural object such as plant and pond, the largest and most prominent object is residential

building. As a necessary place for human living and working, the dwellings have an identifiable protrusion (mainly the roof) to shield the outside world, which is helpful to get a clue to the village space point data. When we acquainting village coordinate, the villages' name information is used to identify the coordinate position by discriminating the information features and spatial location identification. The principle of obtaining coordinate is basically the same as the previous one.

Although some processing is different in some places[3]. Figure 2 shows the thinking and procedural steps of determining the data of the village points.

Figure 2   Reflections and procedural steps for determining village location data

Table 1   Metadata Summary of "The spatial distribution dataset of 1598 more Chinese traditional villages"

Items	Description
Dataset full name	The spatial distribution dataset of 1598 more Chinese traditional villages
Dataset short name	VillagesChina1598
Authors	Yu, L. F-8099-2018, School of Architecture, Soochow University, yuliang_163cn@163.com Ding, Y. Q. O-6445-2018, School of Architecture, Soochow University, 442801422@qq.com Tang, M. J. O-6467-2018, School of Architecture, Soochow University, 361988267@qq.com Cao, Q. Y. O-6478-2018, School of Architecture, Soochow University, 604459330@qq.com Fu, M. O-6455-2018, School of Architecture, Soochow University, 821064405@qq.com Yang, X. H. School of Architecture, Soochow University, 654712015 @qq.com Shen, J. X. Tongji Zhejiang College, 409209740 @qq.com
Geographical region	China, 31 provincial-level administrative regions (excluding Hong Kong, Macao, and Taiwan)
Year	2012–2014 Data format .shp, .kmz Data size 4.50 MB
Data files	Two files (VillagesChina1598.kmz, VillagesChina1958.rar)
Foundation	National Natural Science Foundation of China (41371173)
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn
Address	No. 11A, Datun Road, Chaoyang district, Beijing 100101, China
Data sharing policy	*Data* from the Global Change Research Data Publishing & Repository includes metadata (Chinese & English), datasets (data products) (Chinese & English), and publications (in this case, in the *Journal of Global Change Data & Discovery* (Chinese & English)). *Data* sharing policy includes: (1) *Data* are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use *Data* subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute *Data* subject to written permission from the GCdataPR Editorial Office and the issuance of a *Data* redistribution license; and (4) If *Data* are used to compile new datasets, the 'ten percent principal' should be followed such that *Data* records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset[6]

3.1 Dataset Range and the Attribute of Villages

The dataset covers 31 provincial administrative regions, village raw data: 1598 villages in the "fourth batch of villages listed in the list of traditional Chinese villages" (Announced in 2016, based on the three batches of national traditional villages published by the Ministry of Housing and Urban-Rural Development, the Ministry of Culture, and the Ministry of Finance since 2012). The total number of villages in the list reached 4,153. It can be

seen from the number of villages in the series of different batches that there is an obvious trend of data increase from the first batch to the fourth (Figure 3).

Figure 3   Trends of the number of first to fourth batch of villages

3.2 Acquisition of Village Point Data and Village Names

The data acquisition of the village space points includes: first, spatial cognition, paying attention to the individuals in the space that are easy to identify, confirming the existence of the building or the dwelling, and distinguishing the difference from the natural objects to obtain the location of the village in the wide-area space; secondly, identifying the name of the village and determining the administrative subordination relationship between the village and the superior department to be classified into different data segments. As for the method, Baidu map is used to obtain the point data coordinates (latitude and longitude) of the village space. Taking into account the difference in location between Baidu map and Google earth, especially the dataset needs to be shared in Google Earth globally, on the basis of Baidu map data acquisition, we need to make a fine-tune of village locations for online images of Google Earth manually and visually, Figure 4 shows the fine-tuning technique of Dianqian village, Hetou town, Linhai city, Taizhou city, Zhejiang province. The online position of Google Earth was marked in the field before the adjustment, and after that, it was moved to the building. Regarding the online image of Google Earth, Shuai *et al.*[7] pointed out that it is widely used in the field of geosciences.
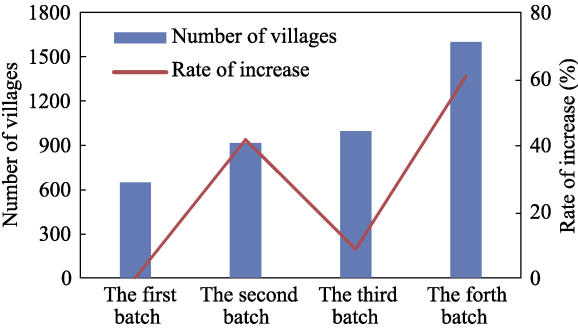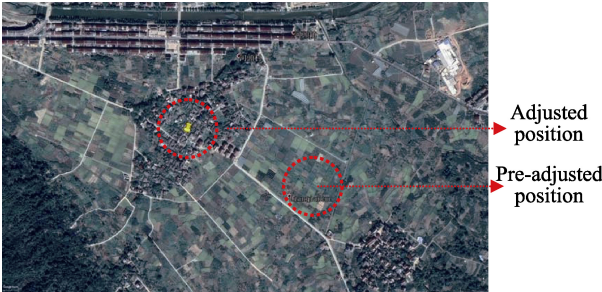

Figure 4   Fine-tuning the location of the village based on Google Earth imagery (Dianqian village, Hetou town, Linhai city, Zhejiang province)

3.3 Sorting out the Name of the Villages

The village's name is the text information of the village clues, the original name of the village, such as Nanjiao village, Nanjiao township, Fangshan district, Beijing, is not only indicating the name of the village, but also showing the administrative affiliation of the village. Generally, the subordinate structure from the village to the provincial level is: province -> city -> county -> town -> village. These affiliation needs to be expressed in different columns according to their respective administrative levels. For example, Nanjiao village, Nanjiao town, Fangshan district, Beijing, can be divided into four levels of city, district, township and village, because it is a direct-controlled municipality. In the general provinces, such

as Yuanqu village of Guxin town, She County, Handan city, Hebei province, there are 5 levels of provinces, cities, counties, towns and villages. In order to unify the number of columns of administrative subordination, this dataset is based on five columns, so some columns will be repeated. For example, Beijing, the direct-controlled municipality, the name of province must be the same as direct-controlled municipality, and there is no municipal name below the direct-controlled municipality. Similarly, there is Huangyaguan village, Xiaying town, Jizhou district, Tianjin, etc. This phenomenon is more common in the direct controlled municipality. Secondly, involving specific level textual representation, the subordinate relationship of the column is not easy to distinguish. Translating from Chinese to English is possible to be misunderstood, so the "county" level is retained and we usually use "Yuxian" directly. The village is not only the village-level organization at the grassroots level, but also the tail of the village lists, but it is not necessarily marked by the "village". Due to the characteristics of Chinese multi-ethnicity and multi-climate type, the village-level title is abundant, some southern villages and ethnic minority use "zhai", "group" and "tun". For example, the "group" of the Ailing Hongyao group of the Jianxin village Committee of Jiangdi township, Longsheng county, Guilin city. And the "gazha", such as the "gazha" of Morenga gazha, Xinhure sumu, Wulatehou qi, Bayanzhuoer city. It can be concluded that there are many representations of villages, communities, neighborhood committees, and village committees, "tai", "tun", "zhai", "zu" and "gazha" in 1,598 villages (Table 2).

Table 2   Inclusions of organizations at different administrative levels

Level	Abbreviation	Connotation and expression
1	Province	Autonomous regions, municipalities directly under the central government and special administrative regions
2	City	Regions, states, district and county of municipality directly under the central government, county
3	County	Municipal districts, county-level cities and banners
4	Town	Township, town, street, etc
5	Village	Village, community, neighborhood committee, village committee, tai, tun, zhai, gazha, etc

The village marked with "community" and "neighborhood committee" is often closer to the urban area, and the inherent village form is prone to be changed. For example, the author investigated the old street community of Luodai town, Longquanyi district, Chengdu city, Sichuan province. Because it is close to Chengdu and commercial activity is developed, it feels like the village is not a traditional village (Figure 5).

Figure 5   "Village" community around the city (left: Laojie community, Luodai town, Longquanyi district, Chengdu city, Sichuan province[7]; right: Wulin community, Xintang street, Jinjiang city, Quanzhou city, Fujian province[8])

4 Dataset Composition and Spatial Distribution of Villages

4.1 Datasets Composition

The dataset consists of two files: 1) VillagesChina1598.rar, which is composed of six data files in ArcGIS with data size of 4.36 MB; 2) VillagesChina1598.kmz, which is one file with data size of 136.17 KB in Google Earth.

4.2 Spatial Distribution of Villages

The data and spatial distribution of the 1,598 villages in the fourth batch are shown in Figure 6.

Figure 6   Spatial distribution of the fourth batch of 1,598 villages (Matching with Google earth Images)

5 Discussion and Conclusion

The village dataset can be used to know the spatial distribution relationship of the batch of traditional Chinese villages. Secondly, many villages in Zhejiang province are selected accounting for the largest proportion in 1,598 villages. And it proves the excellence of protection and utilization work. It is also the successful implementation of the ecological development concept of "Green water and green mountains are mountains of gold and silver" in Zhejiang, China. The protection work in Hunan province and Shanxi province is also very effective. On the contrary, there is no village in Shanghai included in the three batches of villages, indicating that the stronger the impact of urbanization is, the less of traditional villages are, and similar phenomenon is also appearing in Tianjin.

Author Contributions

Yu, L. made an overall design for the acquisition and development of datasets, and wrote data papers; Ding, Y. Q., and Tang, M. J. compiled the key data; Cao, Q. Y., Fu, M., Yang, X. H., and Shen, J. X. collected and processed the dataset.

References

[1]   Liu, D. J., Hu, J., Chen, J. Z., *et al*. Study of spatial distribution pattern of traditional villages in China [J]. *China Population, Resources and Environment*, 2014, 24(4): 157–162.

[2]   Kang, J. Y., Zhang, J. H., Hu, H., *et al*. Analysis on spatial distribution characteristics of traditional villages in China [J]. *Progress in Geography*, 2016, 35(7): 839–850.

[3]   Yu, L., Liu, J., Ding, Y. Q., *et al*. Spatial distribution dataset of 2,555 traditional villages in China [DB/OL]. Global Change Research Data Publishing & Repository, 2018. DOI: 10.3974/geodb.2018.04.06.V1.

[4]   Yu, L., Liu, J., Ding, Y. J., *et al*. The spatial dataset of 2555 Chinese traditional villages [J]. *Journal of Global Change Data* & *Discovery*, 2018, 2(2): 144–150. DOI: 10.3974/geodp.2018.02.03

[5]   Yu, L., Ding, Y. Q., Tang, M. J., *et al*. Spatial distribution dataset of 1598 more traditional villages in China [DB/OL]. Global Change Research Data Publishing & Repository, 2019. DOI: 10.3974/geodb.2019. 01.19.V1.

[6]   GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. DOI: 10.3974/dp.policy.2014.05 (Updated 2017).

[7]   Laojie community, Luodai town, Longquanyi district, Chengdu city, Sichuan province [OL]. https://image.baidu.com/search/.

[8]   Wulin community, Xintang street, Jinjiang city, Quanzhou city, Fujian province [OL]. http://qz.fjsen.co¬m/2017-¬08/07/¬conten¬t_ 19900355_all.htm.

[9]   Shuai, F., Xiao, G. R., Jie, Z. Q. Application of Google Earth in geoscience research [J]. *Journal of East China Institute of Technology* (*Natural Science*), 2011, 34(1): 89–93.