AUTOMATIC EXTRACTION OF URBAN BUILT-UP AREA BASED ON OBJECT-ORIENTED METHOD AND REMOTE SENSING DATA

Longlong Li[1], Huizhen Zhou[2,1], Qiang Wen[1,*], Ting Chen[1], Feng Guan[1], Baiyang Ren[1], Hong Yu[1], Zhiyong Wang[3]

[1] Twenty-first Century Space Technology Application Co., Ltd, Beijing, China - (lill, zhouhz, wenqiang, chenting, guanfeng, renby, yuhong)@21at.com.cn
[2] State Key Laboratory of Earth Surface Processes and Resource Ecology, School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing, China
[3] Beijing Engineering Research Center of Small Satellite Remote Sensing Information, Beijing, China - (wangzy@21at.com.cn)

KEY WORDS: Urban built-up area, Construction land, Object-oriented method, Remote sensing, Automatic extraction

ABSTRACT:

Built-up area marks the use of city construction land in the different periods of the development, the accurate extraction is the key to the studies of the changes of urban expansion. This paper studies the technology of automatic extraction of urban built-up area based on object-oriented method and remote sensing data, and realizes the automatic extraction of the main built-up area of the city, which saves the manpower cost greatly. First, the extraction of construction land based on object-oriented method, the main technical steps include: (1) Multi-resolution segmentation; (2) Feature Construction and Selection; (3) Information Extraction of Construction Land Based on Rule Set, The characteristic parameters used in the rule set mainly include the mean of the red band(Mean R), Normalized Difference Vegetation Index(NDVI), Ratio of residential index(RRI), Blue band mean(Mean B), Through the combination of the above characteristic parameters, the construction site information can be extracted. Based on the degree of adaptability, distance and area of the object domain, the urban built-up area can be quickly and accurately defined from the construction land information without depending on other data and expert knowledge to achieve the automatic extraction of the urban built-up area. In this paper, Beijing city as an experimental area for the technical methods of the experiment, the results show that: the city built-up area to achieve automatic extraction, boundary accuracy of 2359.65m to meet the requirements. The automatic extraction of urban built-up area has strong practicality and can be applied to the monitoring of the change of the main built-up area of city.

1. INTRODUCTION

Since the reform and opening up, China has started a rapid urbanization process. The rapid development of urbanization, the rapid expansion of urban space, and the large increase in the use of various construction lands have resulted in the occupation of a large amount of cultivated land and woodland resources, which pose a serious threat to the ecological environment and directly affect the sustainable development of the city (Xu et al., 2011). The use of land resources has become the most active resource allocation method in today's society. Monitoring and understanding the urban spatial characteristics and changes of cities and adjusting the land use pattern in this process have been considered by scholars as an effective way to reduce risks in the process of urbanization in China (Su et al., 2011). The built-up area marks the status of construction land in different development periods of the city, and accurate extraction is the key to studying the expansion of the city (Zhang et al., 2011). Timely grasping of the distribution of urban construction land and understanding of the expansion scope of urban built-up areas and their occupation of other land use are of great significance to land and resources planning departments and government decision-makers.

At present, there are two ways to calculate the area of built-up areas in China: The "two certificates" summary statistics and remote sensing technology act (Xu et al., 2005). However, there is a certain lag in the traditional summary statistics of "two certificates ", which can not meet the rapid expansion of the city. Satellite remote sensing has the advantages of rapid, macro and integrated access to the data of urban land use and its change,

and remote sensing technology has become the most effective means of monitoring land use change. At present, a variety of classification algorithms have been applied to the extraction of construction land for satellite data. Common methods include: (1) construction land extraction based on classification technology; (2) construction land extraction based on spectrum analysis and logical tree discrimination ; (3) construction land extraction based on construction land index. The classification accuracy of the above methods continues to improve, but the operation process is more complicated, and the model parameters are more. And all of the above methods are pixel-based construction site information extraction, the more use of spectral information and do not make full use of the many features of the target object. Compared with the above method, object-oriented remote sensing information extraction method has strong ability of extracting construction land information. The smallest unit of the information extraction of object-oriented method is the homogeneous image object from image segmentation, rather than a single pixel. The classification results of object-oriented contain richer semantic information and it reduce semantic information loss rate of the traditional pixel-based hierarchical classification method, so the object-oriented method shows great potential in the construction of land use information extraction capabilities.

The built-up area of a city generally refers to the area that can be covered by the outer contour of the built-up area, that is, the realm of construction for the city. Therefore, it is a closed and complete area. The extracted special information of construction land contains urban construction land and rural construction land, so we need to determine the land for urban construction

---

* Corresponding author

and define the scope of urban built-up area. The usual practice is to sketch the vector boundary of the built-up area by using artificial visual interpretation or to define the built-up area by means of other auxiliary data such as DEM, so that the construction area within the boundary is determined as the urban construction site. However, the visual interpretation method requires strong expert knowledge and a low degree of automation; and the method of aiding data is cumbersome and tends to fail because of lack of supporting data (Xu et al., 2011). In this paper, based on the object-oriented method extraction of construction land information, based on the degree of adaptability, distance and area of the object domain, the urban built-up area can be quickly and accurately defined from the construction land information without depending on other data and expert knowledge to achieve the automatic extraction of the urban built-up area.

2. TECHNICAL METHOD

2.1 Construction land information extraction

2.1.1 Multi-scale image segmentation: Using multi-scale segmentation algorithms (Ursula et al., 2004), image segmentation is performed to obtain image objects. This algorithm is based on the spectral features, geometric features, textures, and relationships with other objects of an image object. It is generated from the beginning of a pixel to the entire area. It is gradually merged from the first smaller, near-pixel object to a larger image object. A polygon image object of similar size is formed and the object heterogeneity in the resulting image is kept to a minimum.

2.1.2 Feature selection and construction: The object-oriented classification method divides the image object into an information carrier by segmentation (Ursula et al., 2004; Wu, 2010; Tang, 2013), thereby extracting a large amount of feature information. Making full use of the spectral characteristics (vegetation index, gray mean value, brightness value, standard deviation, etc.), texture features (qualitative, heterogeneity, contrast, and entropy, etc.), and spatial features (object shape, length, aspect ratio, etc.) of the image object can produce more accurate and detailed classification results. In addition, class-related features (topological relationships, contextual relationships), scene features, and process-related features are also used. Table 1 describes object feature information mainly used for classification.

In this study, the results of visual interpretation—Beijing 2014 land cover data as a sample, were used to calculate and analyze commonly used characteristic parameters. The results show that the construction land has a significant difference from other land types in the red light band mean value. Therefore, the effective extraction of construction land can be achieved through Mean R.

Feature	Feature Description and Formula
brightness value	$Brightness = \dfrac{1}{m}\sum\limits_{i=1}^{m}\left(\dfrac{1}{n}\sum\limits_{j=1}^{n}\overline{C_{i,j}}\right)$
Area	The sum of the real area of all pixels in the object
Spectral value	m bands of spectral values
length/width	LW=Long axis/Short axis
Maximum difference	$Maxdif_{i,j} = Max(C_{i,j}) - Min(C_{i,j})$
Index	NDVI, RRI, NDWI, etc.
Object location	The center of gravity of the object
Shape Index	$Shape = P / (4\sqrt{A})$
Gray Level Co-occurrence Matrix	GLCM
Standard deviation	$stdv_{i,j} = \sqrt{\dfrac{1}{s-1}\sum\limits_{i=1}^{s}(C_{i,j} - \overline{C_{i,j}})^2}$

Table1 Feature set description

As shown in Table1, m is the number of image bands, n is the number of objects in an image, s is the number of pixels in an object, and C_(i,j)is the gray value of band i and the jth object.

Statistics	Illustration	Formula
Mean	Refers to the average value of pixel grayscale in the object, reflecting the brightness characteristics of the image	$mean = \sum\sum (i \cdot p(i,j))$
Contrast	Reflects the thickness of the image texture. At the edges of objects and non-homogeneous areas, they all have higher features.	$con = \sum\sum (i-j)^2 \cdot p(i,j)$
Second moment of angle	Measure the uniformity of gray distribution and the thickness of texture. When the image is even and detailed, it has a higher characteristic value, and vice versa.	$asm = \sum\sum p(i,j)^2$
Homogeneity	For the measurement of the smoothness of the image	$hom = \sum\sum p(i,j) / \left[1 + (i-j)^2\right]$

	distribution, the homogeneous region has a higher eigenvalue, and vice versa.	
Variance	Reflects the inhomogeneity of the image, the characteristic value of the homogeneous region is low, and vice versa	$var = \sum\sum (i-u)^2 \cdot p(i,j)$
Entropy	One of the indicators reflecting the amount of image information, the coarse texture region image has good uniformity, and the feature value is low, and vice versa.	$ent = \sum\sum p(i,j) \cdot \lg(p(i,j))$
Correlation	Describe the degree of similarity of the texture under a certain positional relationship. If the similarity is high, the feature value is high, and vice versa.	$cor = \sum\sum \frac{[(i-u)\cdot(j-u)\cdot p(i,j)]}{\sigma^2}$
Dissimilarity	Similar to the variance, it also reflects the inhomogeneity of the image. When the area is more homogeneous, the characteristic value of the dissimilarity is low, and vice versa.	$dis = \sum\sum |i-j| \cdot p(i,j)$

Table2 GLCM Texture feature statistics（Li et al., 2015; Chen et al., 2013）

2.1.3 Automatic extraction of construction land: After multi-scale segmentation of remote sensing images, Object-based extraction of construction land is carried out. The specific process is shown in Figure 1. Detailed construction land extraction methods are as follows (Li et al., 2007; Che et al., 2010)：

1. Mean Red Band（*Mean R*）
Based on the statistical analysis of the sample-based feature parameters, *Mean R* can effectively extract construction land. Therefore, we first apply the characteristic parameter *Mean R* to extract rough construction land types, and set a threshold value A for distinguishing construction land from other land types. When *Mean R>A*, the object that satisfies this condition is construction land (including construction land, bare land and mixed vegetation), otherwise it is other land types.

2. Normalized Vegetation Index（*NDVI*）
Rouse et al. introduced the concept of normalized vegetation index (*NDVI*) in 1973. Through the ratio calculation and normalization process, the strongest reflection band and the weakest reflection band in the multi-spectral band were band-calculated. The strong is placed in the molecule, and the weak is placed in the denominator, thus widening the gap between the two, so that the features of interest are enhanced and other background features are suppressed.

$$NDVI=(NIR-R)/(NIR+R) \qquad (1)$$

In the formula, *NIR* is the brightness value of the pixel in the near infrared band, and R is the brightness value of the pixel in the red band. In this paper, the calculation is performed in units of objects.

Roughly extracted urban construction land can be divided into two categories using *NDVI*. One is construction land and bare land (*NDVI $\leq$ B*), and the other is mixed vegetation (*NDVI>B*) that needs to be removed.

3. Ratio of Resident-area Index（*RRI*）
The Ratio Resident-area Index (*RRI*) is similar to the Ratio Vegetation Index (*RVI*), which reflects the characteristics of residential areas.

$$RRI=Mean\ B/Mean\ NIR \qquad (2)$$

Through this characteristic parameter, the construction land and the bare land can be divided into two categories: construction land (*RRI $\geq$ C*) and bare land (*RRI<C*).

4. Mean Blue Band（*Mean B*）
In the process of removing mixed vegetation by *NDVI*, it is easy to remove the construction land of the blue roof. Therefore, the *Mean B* was used to reclassify the falsely rejected blue roof plant site (*Mean B>D*) to the construction land.
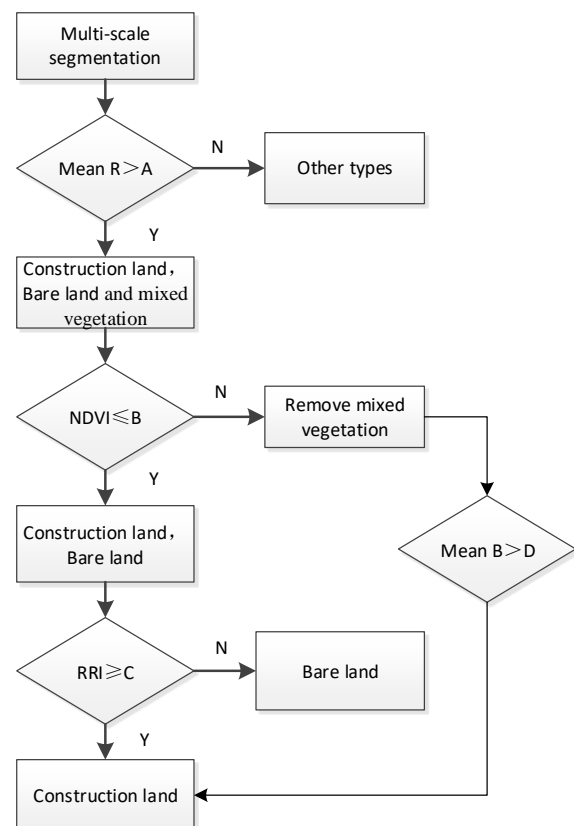

Figure 1. Technological process of construction land extraction

2.2 Automatic extraction of Urban built-up area

The construction land information extracted through the object-oriented method includes urban construction land and rural construction land. How to distinguish urban construction land from rural construction land has become the key to the extraction of urban built-up areas. In addition, the main urban built-up area is a closed area, which contains not only the urban construction land, but also the corresponding urban green space and other ancillary land. Therefore, another key point for automatically extracting the built-up area is to screen non-construction land within the built-up area, determine the scope

of land use in the main built-up area, and then extract the boundary of the main built-up area of the city.

In this paper, we use the proximity, distance and the size of the object to automatically achieve the boundary extraction of the main built-up area based on the object-oriented classification results. The results include two types: construction land and non-construction land. The specific steps and methods are as follows:

1. Screening non-construction land surrounded by construction land and classifying it as non-construction land in built-up areas, defining the conditions as follows：

$$\text{Rel.border to construction land} > 0.8$$

2. Screening the isolated construction land object area outside the built-up area and classifying it as a construction land in a non-built-up area, the rules are as follows：

$$\begin{cases} \text{Rel.border to Non-construction land} = 1 \\ \text{Distance to construction land} > 1000m \end{cases}$$

3. Merge all objects of built-up land which include construction land and non-construction land in built-up areas.

4. The largest patch in the object area of the built-up land was selected, and small patches outside the built-up area that did not meet the requirements were excluded.

5. Merge all objects of non-built-up land which include non-construction land and construction land in non-built-up areas.

6. The object area of land for non-built-up areas that is completely surrounded by the land for built-up areas is screened, and this part of the object area is classified as land for built-up areas. The filter conditions are as follows:

$$\begin{cases} \text{Rel.border to built-up land} = 1 \\ \text{Area} < M \end{cases}$$

The specific threshold M of the Area is set according to the manual judgment, which based on the smallest area of object in the non-built-up land area needs to be reserved.

7. All objects of built-up land are merged to generate the main built-up area.
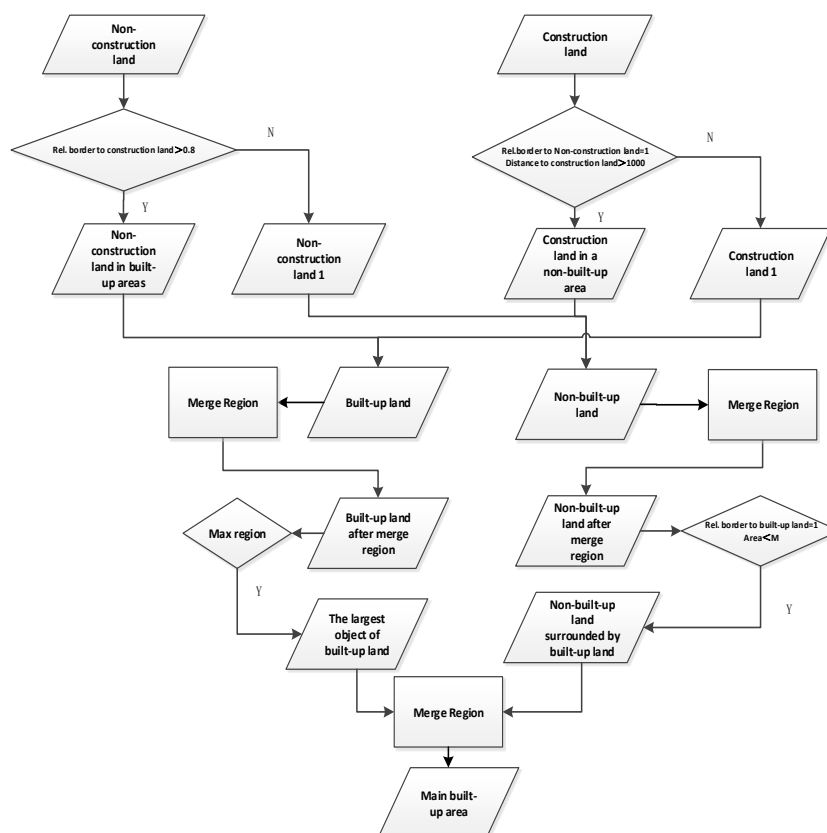

Figure 2. Automatic extraction process of urban main built-up area

3. EXPERIMENTS AND RESULTS

3.1 Study area and data

Beijing is the capital of China and is the political, cultural, educational and international exchange center of China. From the perspective of urban development, pattern and trend, Beijing, as the core city of the Capital Circle, has a large built-up area, complete infrastructure, and rapid development. It has

certain significance as a test and verification area. Therefore, this research will use Beijing as a experimental area to carry out a urbanization development pattern monitoring. The experimental area is shown in Figure 3. The blue area is the main 12 urban areas where the built-up areas of Beijing are located.

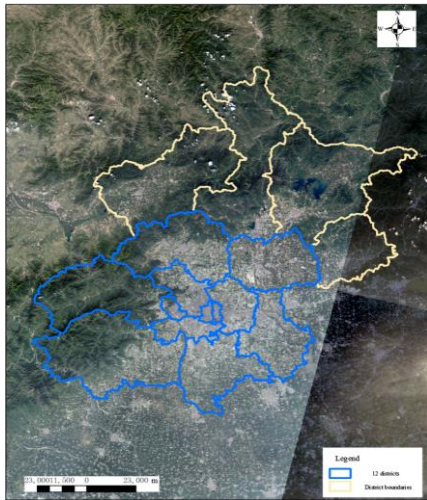The test data is GF-1 multispectral imagery with the resolution of 16m.

Figure 3. Experimental area

3.2 Result

1) The specific rules for the extraction of construction land are as follows:

a) Mean R
Use Mean R to distinguish between construction land and other land, the red band mean threshold is set as follows：R＞50.

b) NDVI
Apply NDVI to remove mixed vegetation from extracted construction land. The NDVI threshold is set as follows: NDVI ≤0.1.

c) RRI
Apply the RRI to remove the bare land from the construction land. The NDVI threshold is set as follows: RRI＞0.595.

d) Mean B
The Mean B was used to reclassify the falsely rejected blue roof plant site to the construction land. The Mean B threshold is set as follows: B＞60.
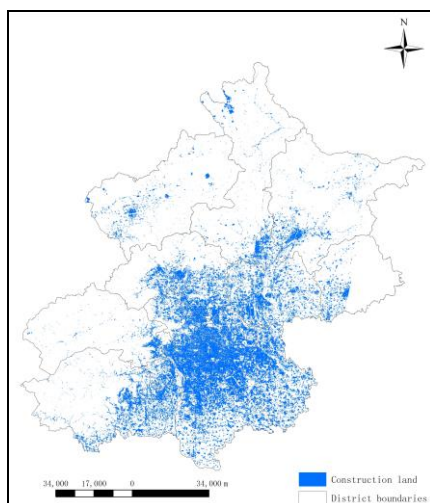

Figure 4. Extraction of construction land

2) According to the specific steps and methods in 2.2, the city's main built-up areas will be extracted. The specific M value is set to 76.8 km^2. The extracted results of the main built-up area are shown in Figure 5.
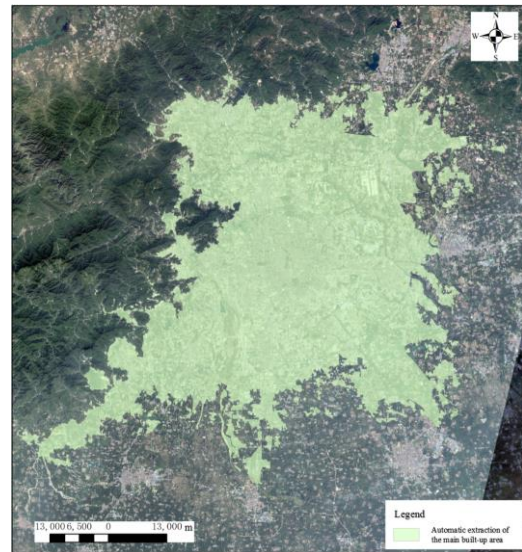

Figure 5. Extraction results of the main built-up area

3.3 Accuracy evaluation and analysis

The superposition analysis of the automatic extraction results of the main built-up area of Beijing and the manual extraction results was performed (Figure 6). Automatically generate 100 verification points (create random points, Figure 7) on the boundary of the manually extracted main built-up arelappoa, and calculate the distance between the point and the boundary of the automatically extracted built-up area (nearest neighbor analysis). Mean values, standard deviations, and other evaluation parameter values were calculated and accuracy evaluation analysis was performed (Figure 8).
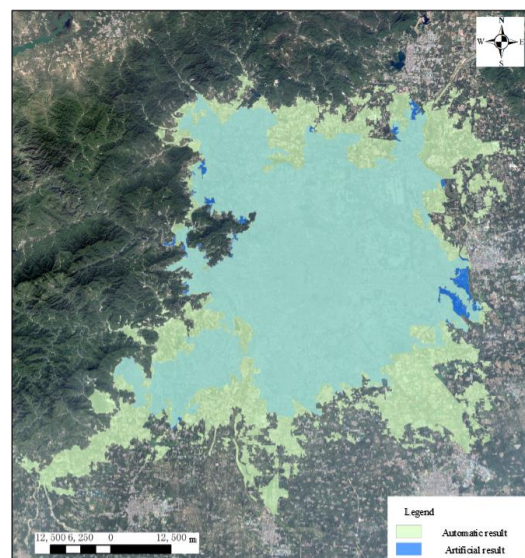

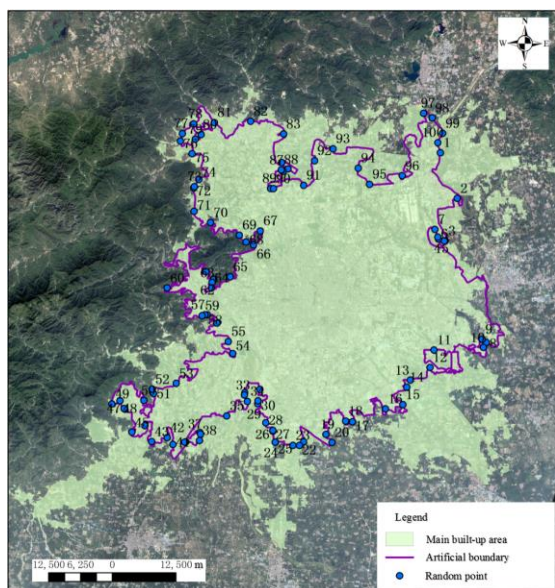Figure 6. Automatic extraction result and artificial extraction result
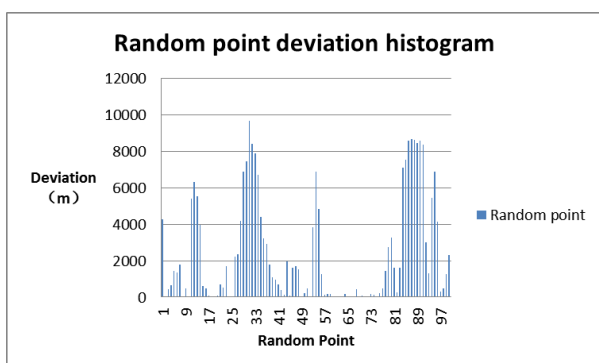
Figure 7. Accuracy verificaiton analysis

Figure 8 Random point deviation histogram

The statistical results show that the average boundary accuracy is 2359.65m and the standard deviation is 2856.96m. From Figure 7 and Figure 8, it can be seen that the accuracy of automatic extraction of the southern boundary of the main building area and part of the northern boundary is relatively large, mainly because the automatic extraction of built-up areas is to extract the contiguous construction land, and the construction land contains rural settlements areas. Therefore, in the extraction process of the main built-up area, it is easy to automatically extract the rural residential areas at the periphery of the main built-up area. The manually extracted main built-up areas, due to the influence of human experience, can effectively identify the rural settlements areas. Even rural settlements that are adjacent to the main built-up area will be removed. The urban-rural junctions in the northern and southern parts of Beijing have more rural settlements areas which closely adjacent to the main built-up areas. Therefore, the rural settlements are easily extracted as main built-up areas, resulting in certain deviations.

Combined with the relevant error analysis, the main problems are the following:

1. The main built-up area is the area formed by the contiguous construction land in the urban area. The scope of the main built-up area is the area that can be included in the outline of the built-up area. In the process of automatic extraction, it is easy to automatically extract the rural settlements area at the periphery of the main built-up area. However, the artificially extracted main built-up areas can effectively distinguish between the main built-up areas and rural residential areas due to the influence of human experience. Even rural residential areas that are adjacent to the main built-up area will be removed. Therefore, the boundaries of built-up areas extracted automatically are likely to be large.

2. In the automatic extraction of the main built-up area, the automatically extracted construction land includes the transportation land (road), and the road is connected to the main built-up area and the surrounding groups. Because the connection of the road is easy to cause the surrounding groups to be divided into the main built-up area, lead to accuracy deviations.

3. The obstruction of larger rivers in the city will result in the leakage of concentrated construction land in the main built-up area, and the leakage of small-scale areas will lead to automatic extraction of the deviation of the accuracy of the main built-up area.

4. CONCLUSIONS AND RECOMMENDATIONS

This paper proposes an automatic extraction method for urban built-up areas based on object-oriented method and remote sensing data, aiming at the actual needs of urban development monitoring. Firstly, based on the object-oriented method, automatic extraction of construction land information, and then based on the factors such as the proximity relationship, distance and area of the construction land object area, the built-up area can be defined quickly and accurately without relying on other data and expert knowledge. The results show that the boundary area of the main built-up area can be extracted automatically, the accuracy basically meets the requirements, and it has strong practicability. It can be applied to the monitoring of changes in the main urban built-up areas of the city. However, the method proposed in this paper has room for further improvement in the accuracy of the main built-up area. The existing problems are the focus of the next step in the study to further improve the precision of the main built-up area.

REFERENCES

Xu, H. Q., Du, L. P., and Sun, X. D., 2011. Index-based definition and auto-extraction of the urban built-up region from remote sensing imagery. *Journal of Fuzhou University (Natural Science Edition)*, 39(5), pp. 707-712.

Su, S., Yu, B. L., Wu, J. P., and Liu, H. X., 2011. Methods for Deriving Urban Built-up Area Using Night light Data: Assessment and Application. Remote Sensing Technology and Application. *Remote Sensing Technology and Application*, 26(2), pp. 169-176.

Zhang, X. F., Yang, Y. L., He, Z. W., Wang, R., and Zhang, D. H., 2011. Built-up area extraction by RS and urban expansion analysis of Suzhou，Wuxi and Changzhou. *Science of Surveying and Mapping*, 36(2), pp. 113-115.

Xu, Q., and Hua, C., 2005. Establishing a Uniform Standard for Calculation of Urban Built-up Area —With Hangzhou Central City Zone as an Example. *Planners*, 21(4), pp. 88-91.

Ursula, C. B., Hofmann, P., and Willhauck, G., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58(3-4), pp. 239-258.

Wu, J., 2010. The Research of Earthquack Information Extraction and Assessment based on Object-oriented Technology with Remotely-sensed Data. Wuhan University, Wuhan, China.

Tang, Y, Q., 2013. Object-oriented Change Detection with Multi-feature in Urban High-Resolution Remote Sensing Imagery. Wuhan University, Wuhan, China.

Li, H. J., Liu Y. L., Tan, R. H., Qiu, L. J., 2015. A New Algorithm of Urban Built-up Area Boundary Extraction Based on High-resolution Panchromatic Image. *Bulletin of Surveying and Mapping*, 2015(7), pp.36-41.

Chen, M. L., and Dai, S. K., 2012. Analysis on Image Texture based on Gray-level Co-occurrence Matrix. *Communications Technology*, 45(2), pp. 108-111.

Li, T. H., Ma, L., Yang, W. N., Zhang, P., Shao, H. Y., and Xia T., 2007. The analysis on urban expansion based on remote sensing and GIS- a case of Nanjing city. *Science of Surveying and Mapping*, 32(4), pp. 124-197.

Che, F., and Lin, H., 2010. Remote sensing information extraction of urban built-up land. *Science of Surveying and Mapping*, 35(4), pp. 97-99.