A Merging Approach for Urban Boundary Correction Acquired By Remote Sensing Images

P. L. Zhang[a, *], W. Z. Shi [b], X.Y. Wu[a]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China – (zpl, wxyandkay)@whu.edu.cn
[b]Joint Spatial Information Research Laboratory, The Hong Kong Polytechnic University and Wuhan University, Wuhan,China -john.wz.shi@polyu.edu.hk

KEY WORDS: urban expansion, urban boundary extraction, remote sensing detection, population distribution, boundary integration

ABSTRACT:

Since reform and opening up to outside world, ever-growing economy and development of urbanization of China have caused expansion of the urban land scale. It's necessary to grasp the information about urban spatial form change, expansion situation and expanding regularity, in order to provide the scientific basis for urban management and planning. The traditional methods, like land supply cumulative method and remote sensing, to get the urban area, existed some defects. Their results always doesn't accord with the reality, and can't reflects the actual size of the urban area. Therefore, we propose a new method, making the best use of remote sensing, the population data, road data and other social economic statistic data. Because urban boundary not only expresses a geographical concept, also a social economic systems.It's inaccurate to describe urban area with only geographic areas. We firstly use remote sensing images, demographic data, road data and other data to produce urban boundary respectively. Then we choose the weight value for each boundary, and in terms of a certain model the ultimate boundary can be obtained by a series of calculations of previous boundaries. To verify the validity of this method, we design a set of experiments and obtained the preliminary results. The results have shown that this method can extract the urban area well and conforms with both the broad and narrow sense. Compared with the traditional methods, it's more real-time, objective and ornamental.

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

Introduction

The urbanization in China is processes of urban spatial expansion, urban scale expansion and rural population assembles to the urban due to the development of social productive forces. At present, the urbanization process of China is at a stage of rapid development, and it is demonstrated by the expansion of built-up area. The built-up area reflects the level of development and comprehensive economic strength, and provides the scientific basis for urban management, planning and other research.(Hu Yi-dong,2008) At the same time, the rapid expansion of urban brings a lot of problems inevitably, such as resources waste, soil erosion, extinction ecological deterioration, especially the expansion of land use scale.(Shi Xiao-yun,2004) Grasping the information about the boundary of urban is meaningful for guiding the urban expansion, protecting resources and environment.

The expansion of urban is a complex process. And traditional methods of urban spatial morphological detection can't obtain the development information quickly. Remote sensing provides synoptic views in space and consistent data set with high temporal and spatial resolution. It can provide real-time efficient and accurate methods to get the information about the land use.(Sun Shan-lei,2008) It offers the capacity for dynamic monitoring of large urban agglomerations and proved to be an effect tool to map the land use and changes analysis. And today the increasing availability of satellite images is becoming instrumental in geographical applications, offering a reliable tool for decision making in territorial management.

Nowadays, aggregation indexes such as density of population, building density and urbanization indexes like population composition are major bases for dividing the urban area all over the world. For example, the United States and Japan use density of population as the main indicator to define urbanized region.(SCDPCG,2010) In the United Kingdom, the urban area is the prevailing conception. And it delimit the urban built-up area according to the distribution of buildings in the satellite image.(Healey,1998) In China, delimiting urban area always considers two aspects of administrative region and landscape regions.

Although urban area can be extracted by only using remote sensing technology, the region does not contain social economic information. We propose a method of using remote sensing image combined with demographic data, road data and other data. In this paper we select Landsat TM image, population and road as data resources which covered Hanyang district of Wuhan. On the basis of boundary of urban area extracting from satellite image, we use the boundaries from other data to rectify the former.

Study region

Hanyang District is located in the southwest of Wuhan, Hubei Province, China, bordering the Yangtze on the east, lying upon the Han River, 108 square kilometers. Jurisdiction over the region 11 administrative streets. Hanyang district regards urban industry, realty business, cultural tourism and modern service industry as focal points of development.

Data resources

The satellite earth observation systems has the characteristics of periodicity and multi-spectral which makes remote sensing images the best available data to research the time series variation of urban area.(Shen Run-ping,2002) Landsat can identify different objects effectively with its high spatial resolution and spectral resolution. In this research, we use the Landsat5 TM image. And the other datum resources contain census data, traffic map, administrative manual of Wuhan and so on.

Extraction based on TM Data

To analyze the objects in this research region, taking spectrum samples for typical objects (water, vegetation, building, farmland, low density vegetation and others). According to the difference between building and other objects in spectral response curve, the information about built-up area can be extracted from the image.

NDVI,() also called normalized difference vegetation index, is one of the most widely used indicators. There is difference between near infrared band and red band. And the corrected result is the difference corrected by the sum of the two, that is NDVI. It can be calculated by the following formula.

$$NDVI = (NIR - RED)/(NIR + RED) \quad (1)$$

The value ranges from -1 to 1. Zero represents non vegetation in this area. Negative value represents non vegetated area. But, when the value ranges from 0 to 1, the greater the value of a pixel in NDVI is, the higher is the possibility of the pixel being a vegetation area and the greater the amount of vegetation is. Cloud, water and ice have significant reflection in the red and near infrared wave band, so the NDVI value is negative. Soil and rock in the two band reflectance are nearly the same,so the NDVI value is close to zero.

An in-depth analysis of NDVI produced NDBI. Because the value of vegetation in mid infrared band is greater than it in near infrared band, and the DN value of other objects become smaller, the NDVI can extract the vegetation from the image. And the DN value of built-up area between mid infrared band and near infrared band is on the rise,the others are falling. So, for the same reason, NDBI can extract the built-up area, its formula is

$$NDBI = (MIR - NIR)/(MIR + NIR) \qquad (2)$$

However, the difference of built-up area between mid infrared band and near infrared band is far less than the vegetation. Except the built-up area, the DN value of bare land and low density vegetation which contains information of soil background are also have the same property: the value in mid infrared band is greater than it in near infrared band. Therefore, the accuracy of extracting built-up area would be affected.

BU index resulted from subtracting NDBI and NBVI.(Dai Jian-guang,2008) Due to NDBI index contains some low density vegetation area which brings some error to the built-up area extraction results, using NDVI index can remove the area of low density vegetation. Therefore, the greater the value of a pixel in BU is, the higher is the possibility of the pixel being a built-up area.

On the basis of NDWI index, MNDWI (Modified NDWI) has been proposed. MNDWI and NDWI are different in combination of bands. By comparing and researching experimental results, MNDWI is more effective than NDWI to extract water area from built-up area. The images of NDWI index always contains some built-up area, which cause the actual area of water expansion. (Xu Han-qiu,2005)The formula of MNDWI is

$$MNDWI = (Green - MIR)/(Green + MIR) \qquad (3)$$

In order to extract the built-up area efficiently, the images need to be classified. Before that, we use BU, NDVI and MNDWI results to produce a pseudocolor synthetic image. In the image, BU, NDVI and MNDWI represent the built-up area, vegetation and water respectively. By correlational analyses, they have weak correlation, even negative correlation. We carry out classification with this false color image to get the built-up area and extracted its boundary. This method reduces the possibility of spectral feature confusion, and is conducive to extract built-up area.

Extraction based on demographic data

Population distribution reflects the aggregation degree of urban, and to some extent, built-up area is densely inhabited area. In order to grasp the information about region and boundary of urban, firstly, demographic data need to be distributed to a specified size grid district according to a certain mathematical model. However, the publicly available demographic data is collected by units of administrative streets, villages or towns. The area weighted model is one of the main grid transformation models. These is a assumption that population evenly distributed in certain administration cells, which doesn't correspond to spatial characteristics of population distribution in urban. In order to distribute as much as makes sense, we use area weighted model with Thiessen Polygon.

A • H • Thiessen, a Dutch climatologists, proposed a polygonal model to calculate the average rainfall based on datum of discrete weather stations. The characteristics of this model are: each Thiessen polygon contains only one discrete point; all the points in the polygon are at a a minimum distance from the corresponding discrete point and the points on the edge of polygons is equidistant from both side discrete points.(Yan Qing-wu,2011)

We use Thiessen polygon to calculate the discrete demographic point data, get population in grid cell by population density and area ratio between Thiessen polygon cell and grid cell. It's efficient to express the spatial distribution and regularities. Through manual method and trial, based on a set threshold value to obtain a boundary.

Boundary Integration

In order to obtain a reliable and accurate urban boundary, it's necessary to synthesize all the boundaries we extracted from different datum and reduce errors made by manual operation of interpreters.(Jiang Jin-gang,2009) Therefore, we adopt a weighted average algorithm for integrating of the border information.

In this part, the boundary extracted from remote sensing image is chosen as the main data and assigned higher weight. The boundaries are all represented by a set of coordinated points. Every point of main data are set to a center and to find two points on other boundaries which have the min sum of Euclidean distance to the center point by spatial search. A triangle is formed by lines joining the two points and center point. Let's suppose *O* is the point on the main boundary, and points *A*,*B* on the other boundary have the min sum of Euclidean distance to the point O. We make a vertical line through the point *O* and intersects line *AB* at point *N* which involve in the weighted arithmetic.

$$\begin{aligned} x &= \omega_o x_o + \omega_n x_n \\ y &= \omega_o y_o + \omega_n y_n \\ \omega_o + \omega_n &= 1 \end{aligned} \tag{4}$$

where, x and y are the coordinates of this result. The ω_o and ω_n are weight of each boundary ,and the sum of them is 1.

Experiment and Results

Fig.1 shows Landsat5 TM image of the Hanyang District in 2010 and bands combination is 543.

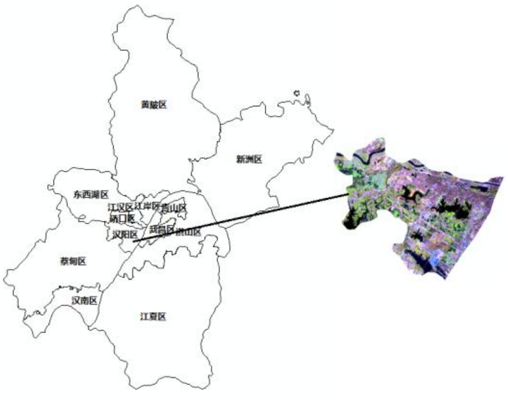

Fig. 1 TM image of study region

After calculating indexes, synthesizing false color map and classification, we get the preliminary result of built-up area. With the present land-use map, merger guidelines of discrete surface features and the definition of vegetation and water body, extracting the boundary from the result of classification. The Fig.2 show the white area is the boundary extracted from TM image.

Fig. 2 Boundary extracted from image

Based on the area of Hanyang district, we choose build 500m*500m grid. Through the overlay analysis of Thiessen polygon and grid, population distribution shows the spatial distribution patterns are aggregated in clumps. According to the statistic of city population density ,the threshold is set at 1000 people for every sq km. Fig.3 shows the boundary extracted from demographic data.

Fig. 3 Boundary extracted from demographic data

The Fig above shows, the two boundaries reflect a misalignment. The boundary integration is necessary. Using weighted average algorithm, we obtain the final result boundary.(Fig.4)

Fig. 4 The final boundary

The results have shown that index synthesis image can be classified well to extract the built-up area and decrease errors. And population data grid transform can provides strong support for built-up area and boundary revision. The final urban boundary contains both geographical and social economic information.

Summary

The article studied the technology of extracting the built-up area boundary from different data and integration. According to the features of the built-up area in remote sensing image, a specific project is presented which use Landsat image to extract the built-up area. It has been proved through a series of experiments. With the characteristics of the Landsat images bands, the use of NDVI,MNDWI and BU index to classify after false color composite processing, has laid the foundation for the accurate extraction of built-up area. At the same time, to ensure that the extraction of urban area is accurate and reasonable, we make use of social and economic statistical data such as population data to modify built-up area extracted. It is more real-time, objective and ornamental when compared with others.

However, there are still a lot for improvement. In the part of extraction from TM data, we need to merge scattered surface features by manual method according to the result of classification and present landuse map, like lakes, which adds up inaccuracy and mistakes to the boundary. Furthermore, the method of population distribution need to be perfect because of uneven population distribution.

References

Hu Yi-dong, 2008. Study on defining urban built-up area—a case of Wuhan. City planning review (4), 88-91.

Shi Xiao-yun, 2004. Study on the urban using land expansion in period of high-speed urbanization—a case of Nanjing. Nanjing Agricultural University.

Sun Shan-lei, 2008. Dynamic Monitoring of Urban Expansion in the Region around Hanzhou Bay. Journal of Natural Resources 23(2), 327-335.

The State Council Development Pesearch Center Group, 2010. International experience of urban and rural space division. China Development Observation 7, 54-57.

Healey, P., Collaborative planning in a stakeholder society. Town planning review 69(1),1.

Shen Run-ping, 2002. Research on Key Technologies of Land Use Monitoring by Remote Sensing and its Application—A Case Study of Poyang Lake Region in Jiangxi. Zhejiang University.

Dai Jian-guang, 2008. Methods for extraction of urban built-up land information based on TM Data. Modern Surveying and Mapping 31(6),34-36.

Xu Han-qiu, 2005.A Study on Information Extraction of Water Body with the Modified Normalized Difference Water Index(MNDWI). Journal of Remote Sensing 5,589-595.

Yan Qing-wu, 2011. Census Spatialization Based on Thiessen Polygons and Grids Smoothing: A case Study in Xuzhou. Geomatics and Information Science of Wuhan University 36(8),987-990.

Jiang Jin-gang, 2009. Fusion algorithm for the information of lake boundary integration from multi-temporal remote sensing images. Journal of Lake Sciences 21(2), 264-271.