

URBAN SHANTY TOWN RECOGNITION BASED ON HIGH-RESOLUTION REMOTE SENSING IMAGES AND NATIONAL GEOGRAPHICAL MONITORING FEATURES -- A CASE STUDY OF NANNING CITY

Yuqing He^{1,2}, Yongning He¹

¹ Geomatics Center of Guangxi, Nanning 530023, China - hyq1122@hotmail.com - hyn886@qq.com

² Guangxi Collaborative Innovation Center of Multi-source Information Integration and Intelligent Processing

Commission III: ICWG II/III

KEY WORDS: Gray-level co-occurrence matrix; shanty town; residential suitability; nearest neighbour

ABSTRACT:

Urban shanty towns are communities that has contiguous old and dilapidated houses with more than 2000 square meters built-up area or more than 50 households. This study makes attempts to extract shanty towns in Nanning City using the product of Census and TripleSat satellite images. With 0.8-meter high-resolution remote sensing images, five texture characteristics (energy, contrast, maximum probability, and inverse difference moment) of shanty towns are trained and analyzed through GLCM. In this study, samples of shanty town are well classified with 98.2% producer accuracy of unsupervised classification and 73.2% supervised classification correctness. Low-rise and mid-rise residential blocks in Nanning City are classified into 4 different types by using k-means clustering and nearest neighbour classification respectively. This study initially establish texture feature descriptions of different types of residential areas, especially low-rise and mid-rise buildings, which would help city administrator evaluate residential blocks and reconstruction shanty towns.

1. INTRODUCTION

Urban shanty towns are recognized as decrepit, dirty and disordered communities hidden in metropolises, which has high public safety risks such as fire accidents, public security crimes and public health problems. They are officially defined as urban regions with contiguous old and dilapidated houses (more than 2000 square meters built-up area or more than 50 households), high residential density, poor basic infrastructure, etc. Nanning is the capital city of Guangxi Zhuang Autonomous Region, China. Owing to the consistent and continuous efforts of city builders, shelters and squatter settlements are hardly found in Nanning's urban area; however, illegal construction and improper block plan are ubiquitous in urban villages, which makes them a new form of shanty towns.

The existence of shanty towns has adverse effect on a city's urbanization and identity. Even hidden behind mansions and busy streets and hardly seen by passers-by, their potential risks should not be overlooked. For better urban management, such areas need to be detected and they need to be carefully distinguished from surrounding blocks. In the practice of urban shanty town identification, city administrators rely heavily on field works, which wastes labour and time. Even later with the assists of remote sensing images, manually visual interpretation is still time-consuming and experience-requiring. Hence, an integrated shanty town detection method is necessary in macro city planning and local administration.

Shanty towns in Nanning have very unique spectral characteristic in remote sensing images. Many household of them utilize blue-painted iron sheet to set up their roofs. Disordered, contiguous and congested squares are shown in imagery. With such textural features, Gray-level co-occurrence matrix (GLCM) can be used to summarize and line out their signatures. GLCM is first put forward by R. Haralick et al. (1973) in early 70s. It is commonly applied to describe texture characteristics of images. GLCM has 14 different characteristics indexes, which quantitatively evaluate and distinguish different features texture in various aspects and help researchers to line out target objects automatically. Among these indexes, data range, mean, variance, entropy and skewness frequently are used in researches. The contiguousness and intensiveness of shanty town appearance in remote sensing image make GLCM a good choice in recognition and extraction experiment.

2. DATA

2.1 TripleSat satellite images

TripleSat satellite imagery is available at 0.8m high-resolution imagery products with a 23.4km swath. Both space and ground segments have been designed to efficiently deliver guaranteed timely information (Satellite Imaging Corporation, 2017). In this study, 0.8m TripleSat images is captured in October, 2017, which is latest remote sensing sources during time of study. Correction, color uniformity and cloud removal are proceeded prior to this study.

2.2 The First China's National Geography Census

The First China's National Geography Census passed acceptance check in 2017. This census investigated all land cover features via remote sensing visual interpretation and field

check. All land cover types are precisely and seamlessly classified into 8 major classes, 56 mid-classes and 126 subclasses (mainly with 5 pixels' plane precision). Built-up areas are classified into low-rise (mainly 3 floors and below), mid-rise (mainly 4 to 6 floors) and high-rise (7 floors or higher). According to the definition, shanty town belongs to the low-rise and mid-rise built-up area classes. This can be used to roughly extract areas that shanty towns potentially exist. Also, social elements are collected through this census, including road network, civil infrastructure, schools, hospitals, etc. Such data can be used to evaluate a community's residential suitability later on.



Figure 1. Zooming view of TripleSat and Residential Layer
© Satellite Imaging Corporation Copyright 2017

3. METHODOLOGY

3.1 Class determinations and sample selection

Among the low-rise and mid-rise building, there are majorly 4 types of building blocks: shanty towns, townhouses, ordinary residential blocks, and industrial buildings (Table 1). As defined before, shanty towns are residential blocks with contiguous old and dilapidated houses, mostly found inside urban village, and no visible interior alleys found from remote sensing images. Homestead area of a household is relatively small. Townhouses are also found in urban village sites, but unlike disordered shanty towns, townhouses are constructed in rows and interior alleys between rows are broad enough to distinguish them into small blocks. Hence, with more expedite road network and less residential density, townhouses have lower level of public safety risks than shanty towns. Ordinary residential blocks are commonly seen in urban area, dozens of households live in the same building unit but in different floors and different apartments. Their building areas are large, distance between buildings are long with greenbelt, and their roofs are irregular. Considering the floor restriction of low-rise and mid-rise class, such residential blocks could be ancient. Their construction standard might have been out of date so they could be fragile or dilapidated. Hence, involving such class into study is meaningful to urban renovation plan. The last one, industrial buildings are mostly seen in manufacturing districts. Similar to

shanty town and townhouses, industrial buildings have flat roof and the majority of them are covered by blue-painted iron sheets as roofs. What makes them unique is the large built-up area of a single building.

| Type Code | Type Name | Count |
|-----------|----------------------------|-------|
| 1 | Shanty town | 112 |
| 2 | Townhouses | 82 |
| 3 | Ordinary Residential Block | 103 |
| 4 | Industrial buildings | 103 |
| Total | | 400 |

Table 1. Classes of training residential areas

The study area is circled by Nanning's Express Loop, covering most of urbanized area and entire potential shanty town sites, and excluding rural village house-sites. In this study, low-rise and mid-rise build-up area parcels are extracted from the Land Cover Layer of the National Geography Census. Large parcels are split into small ones around 100m*100m ground area. TripleSat remote sensing images are clipped into 8710 quadrates with 125*125 pixels based on the small parcels. Training sample have to be chosen carefully, land cover type is exclusive within a quadrate. All 400 training samples saved as GeoTiff are labelled by their Type Code for references. The amount training samples of different types should be in approximately the same. Lacking of

3.2 Conduction of GLCM and derived features

Gray-level co-occurrence matrix (GLCM) is a statistic method of texture analysis, mainly reflects an object's roughness, contrast, fineness, and regularity by describing relationship between a pixel and its neighbouring pixels. Whereas, GLCM does not utilize an image's original greyscale values, it represents texture by calculating probability of joint criteria between image greyscale levels (Baraldi and Parmiggiani, 1995). GLCM needs to calculate the occurrence probability of (i, j) started from greyscale i in a given distance and a direction. Hence, all directions (0° , 45° , 90° , and 135°) need to be calculated in order to obtain full texture regularities. Considering the directions of residential blocks are different and irregular, direction is set as default (0°) in this study.

GLCM is a matrix of $2^k \times 2^k$ ($k > 0$) and the numbers seem meaningless at the first look. Afterwards, numerous of features conducted from GLCM help highlight texture characteristics. (Dasgupta et al, 2017). In this study, 5 features are selected, calculated and analyzed. (1) *Energy* is quadratic sum of GLCM elements, which measures stability of greyscale patterns, larger value means regulation is more stable. (2) *Contrast* is used to measure sharpness and of an image, higher value indicates texture steep, lower value means image is fuzzy. (3) *Maximum probability* shows the texture feature that occurs most often. (4) *Entropy* describes randomness of image information, and its value increases when a picture gets complex. (5) *Inverse difference moment (IDM)* reflects homogeneity of texture, higher value means less changes and uniform in local scale. The calculation formulas of these 5 features are shown below:

$$Energy = \sum_i \sum_j \{p(i, j)\}^2 \quad (1)$$

$$Contrast = \sum_i \sum_j (i - j)^2 p(i, j) \quad (2)$$

$$Max\ Probability = \max \sum_i \sum_j p(i, j)^2 \quad (3)$$

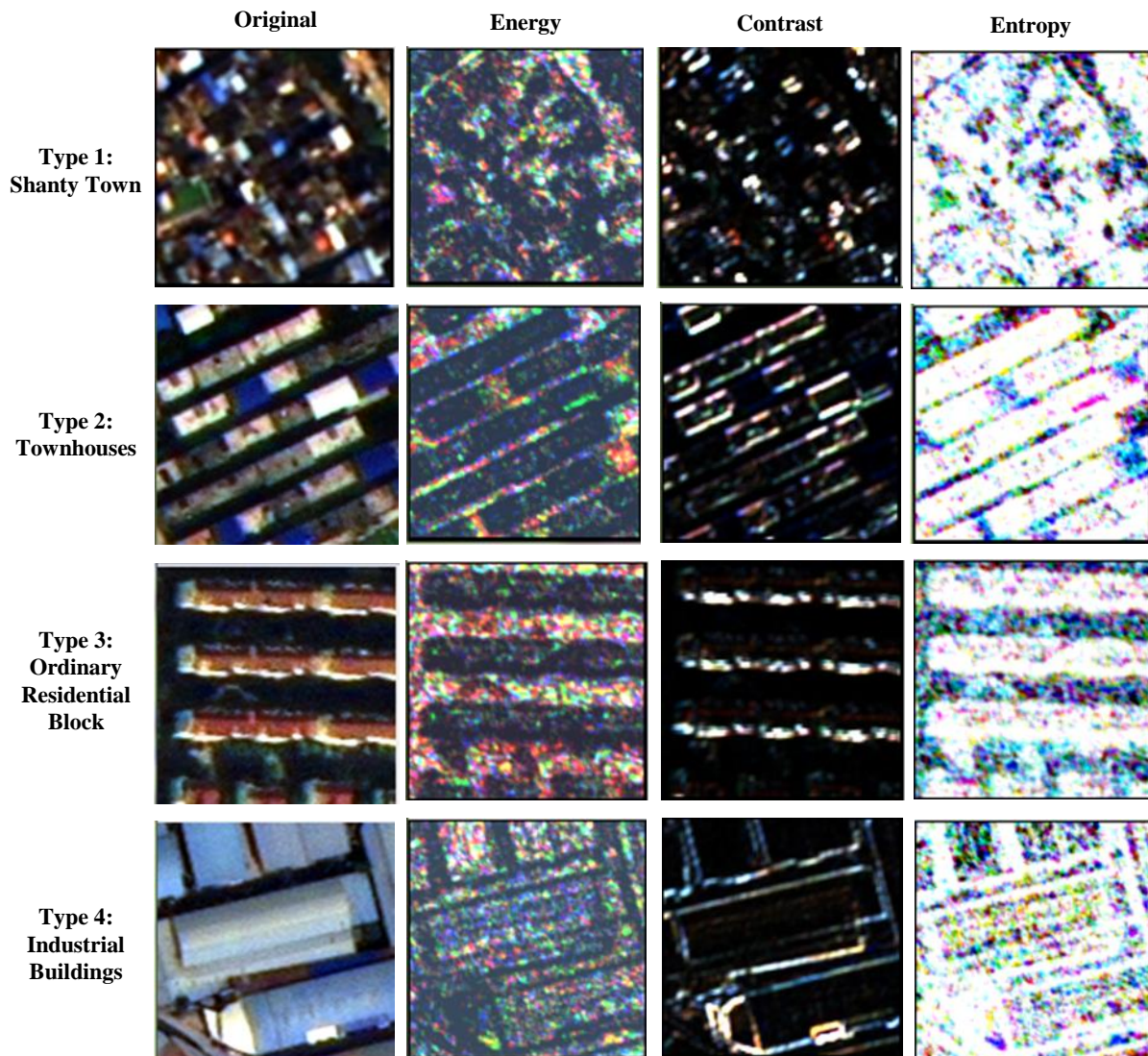


Figure 2. Features comparison of different residential types calculated based on pixels

$$Entropy = \sum_i \sum_j p(i,j) \log(p(i,j)) \quad (4)$$

$$IDM = \sum_i \sum_j \frac{p(i,j)}{1+(i-j)^2} \quad (5)$$

where $p(i,j)$ = GLCM matrix
 i, j = matrix coordinates

GLCM and its 5 features of 400 training samples are batch proceeded in Visual Studio 2015 (© Microsoft copyright 2017). The script is written in C++ and OpenCV is invoked (OpenCV team, 2018). Since the procedure only conduct numeral results, co-occurrence matrix module of ENVI (© ESRI copyright 2017) is utilized to proceed GLCM features in pixel level and give visual impressions. In Figure 2, energy, contrast and entropy of all 4 training block types are computed. Shanty town is the most scattered type among them. Patterns of townhouses and ordinary residential block are clearly seen. Edges of industrial buildings are highlighted though the Contrast feature.

3.3 Classifications

The original GLCM feature values are stored into a text file. To proceed classification modules, the result file is imported into SPSS (© Digital Millennium copyright 2017). Both unsupervised and supervised classification methods are used to test the 400 training samples.

K-means clustering: an unsupervised clustering algorithm. Assume there are a numbers of points in an x, y coordination and r as anticipated outputting group numbers, k-means initialized r points randomly (x_i, y_i) ($i \leq r$) as central points of every cluster (can also set up initial points manually). The module then use iteration method to assign samples to the nearest central point class. Weighted average is calculated and updated as the newest central point of each cluster group. Another iteration start to re-do the above process until iteration time reach pre-defined limit or the central no longer move. Without data training, priori knowledge and a series of parameters, k-mean is a simple way to classify unknown data. Nevertheless, it is not suitable for imbalanced dataset. If there is an obviously big class in the dataset, all the weights are taken

by the big class, and the small classes cannot stand out as independent groups.

Nearest Neighbour: supervised classification algorithm. Labelled training dataset is necessary and it belongs to memory-based learning. The module randomly divide dataset into training dataset and testing dataset. Training dataset is calculated for central points of every group, then for every testing point, find k nearest training points and assign label to the major group. Both k-means and KNN utilized nearest neighbour algorithm, normally using k-dimensional tree. Also, nearest neighbour algorithm does not perform well in imbalanced dataset.

4. RESULTS

4.1 k-means clustering classification

The final centrals of each cluster are shown in Table 2. Producer accuracy is marked in bold in Table 3, which is derived by calculating rations of correctly classified samples and total sample amount of a class. K-means recognize shanty town (98.2%) and industrial buildings (94.2%); however, Townhouses and ordinary residential block are indistinguishable via this cluster. User accuracy highlighted in Table 4 is conducted by calculating rations of correctly classified samples and total sample amount within new cluster. Comparing Table 3 and Table 4, high producer accuracy does not necessarily guarantee high user accuracy. For example, Samples of shanty town are mostly classified into k1, but shanty town is not exclusive in k1, some townhouse samples are also in k1. Hence, to evaluate a classifier, both producer accuracy and user accuracy need to take into consideration accordingly. Different amounts of clusters are proceed respectively; however, clusters of 4 perform the best in all aspects.

| | Cluster | | | |
|----------|-----------|-----------|-----------|-----------|
| | 1 | 2 | 3 | 4 |
| Energy | 1.20E+08 | 1.99E+08 | 1.58E+08 | 6.56E+07 |
| Contrast | 1718.88 | 759.85 | 1293.60 | 2024.56 |
| MaxPro | 10588.11 | 14063.65 | 12459.31 | 6263.18 |
| Entropy | 131798.55 | 142406.56 | 136776.82 | 124125.64 |
| IDM | 1576.85 | 654.74 | 1157.08 | 1838.15 |

Table 2. Final cluster centers of k-means

| | k1* | k2 | k3 | k4 |
|----------------------------|-------------|------|------|-------------|
| Shanty town | 98.2 | 0.9 | 0.9 | 0 |
| Townhouses | 54.9 | 0 | 41.4 | 3.7 |
| Ordinary Residential Block | 2.9 | 49.5 | 47.6 | 0 |
| Industrial buildings | 2.9 | 1.9 | 1.0 | 94.2 |

*k1, k2...stand for clusters group classified by SPSS.

Table 3. Producer accuracy of k-means classification (%)

| | Shanty town | Townhouse | Ordinary Residential Block | Industrial buildings |
|-----|-------------|-----------|----------------------------|----------------------|
| K1* | 68.3 | 27.9 | 1.9 | 1.9 |
| K2 | 1.9 | 0 | 94.4 | 3.7 |
| K3 | 1.2 | 40.0 | 57.6 | 1.2 |
| K4 | 0 | 3.0 | 0 | 97.0 |

*k1, k2...stand for clusters group classified by SPSS.

Table 4. User accuracy of k-means classification (%)

| | Nearest Neighbor Classification | | | | | | | | | Summary | | |
|-------------------------------|---------------------------------|-------------|-------------|-------------|-------|-------|-------|-------|-------|---------|-------|-------|
| | 1 | 2 | 3 | 4 | 1 - 2 | 1 - 3 | 1 - 4 | 2 - 3 | 2 - 4 | Correct | Vague | Wrong |
| 1. Shanty town | 73.2 | 9.8 | 1.8 | 0 | 15.2 | 0 | 0 | 0 | 0 | 73.2 | 15.2 | 11.6 |
| 2. Townhouses | 15.9 | 57.3 | 15.9 | 2.4 | 4.9 | 0 | 1.2 | 2.4 | 0 | 57.3 | 7.3 | 35.4 |
| 3. Ordinary Residential Block | 0 | 5.9 | 85.4 | 0 | 0 | 2.9 | 0 | 5.8 | 0 | 85.4 | 8.7 | 5.9 |
| 4. Industrial buildings | 3.9 | 0 | 2.9 | 92.2 | 0 | 0 | 0 | 0 | 1.0 | 92.2 | 1.0 | 6.8 |

Table 5. Classification accuracy of nearest neighbour (%)

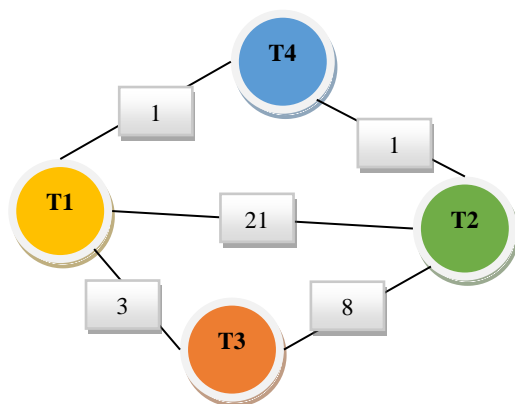


Figure 3. Undetermined samples between classes

4.2 Nearest Neighbor

In nearest neighbour classification, the system selected 74.5% training samples and 25.5% holdout samples. Unlike cluster algorithm of k-means, nearest neighbour method provides probabilities of a sample that might belong to a particular class. Samples are supposed to put into a class of the highest probability; however, in some cases, the probabilities are equally high, which make samples unable to be classified. In Table 5, some columns are combination of two classes, neither correct nor wrong, hence they belong to vague column. The undetermined part is not useless, sometimes it helps researchers to find out intermediate zone and connection between two classes. Seen from Figure 3, the vague samples are extraordinary large between T1 and T2, which indicates shanty towns and townhouses are not clearly divided. As mentioned before, the differences between shanty town and town houses

are width of internal alleys. If the house site area of each townhouse is small, the texture would be similar to that of shanty town. Also, small house site indicates higher density of household in a fixed region, which will also lead to shanty town problems. If a linear regression is proceeded among shanty town samples and townhouse samples, we could find disorderliness degree, potential risk level and more parameters scoring and ordering shanty town and townhouse samples form the worst to the best, then utilize the function to evaluate low-rise and mid-rise residential block in Nanning City.

There are 5 features in this study, which cannot be visualized by 3-D scatterplot. In Figure 4, 3 variables (contrast, energy, and maximum probability) are selected to display distributions of training samples.

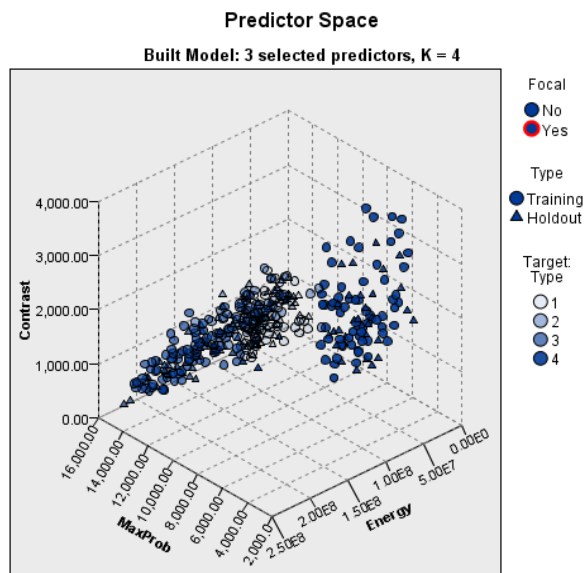


Figure 4. Predictor Space of three selected predictors*

*This chart is a lower-dimensional projection of the predictor space, which contains a total of 5 predictors.

4.3 Predictions

Both k-means clustering and nearest neighbour classification are used to predict 8710 residential parcels in Nanning City respectively. In k-means clustering, the final cluster centers of 400 training samples are set as initial cluster of entire dataset. Rates of change between initial and final centers (final center minus initial center divided by initial center) shown on Table 6 are not greater than 12%, which shows stability of the cluster centers. As results, the parcels are separated as Cluster 1 (2430), Cluster 2 (1680), Cluster 3 (2634), and Cluster 4 (1966).

| | Cluster | | | |
|----------|---------|------|------|------|
| | 1 | 2 | 3 | 4 |
| Energy | -4.7 | -1.9 | -2.5 | 11.1 |
| Contrast | -8.9 | -3.1 | -5.0 | -1.4 |
| MaxPro | -4.2 | -1.0 | -1.4 | 11.3 |
| Entropy | -0.1 | -0.3 | -0.3 | 1.0 |
| IDM | -7.1 | 3.4 | -2.0 | 1.0 |

Table 6. Changes between initial and final cluster centers in k-means clustering prediction (%)

In nearest neighbour classification, 400 labelled samples are taken as training data and the module predicts 3376 shanty towns, 494 townhouses, 2118 Ordinary residential blocks, and 2722 industrial buildings.

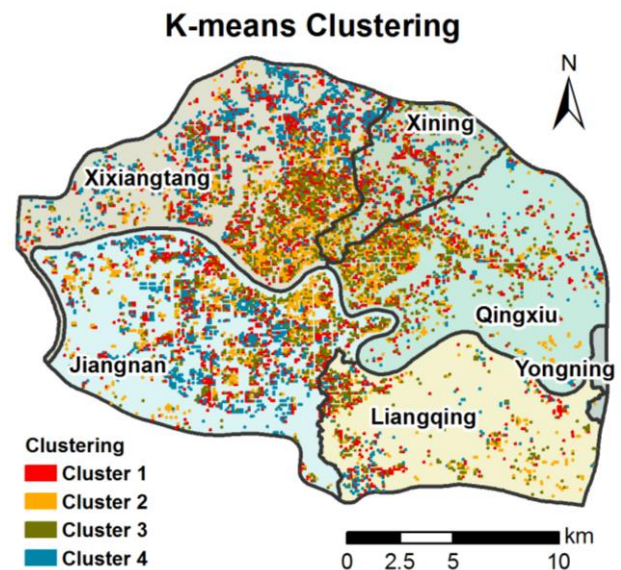


Figure 5. K-means clustering apply on Nanning

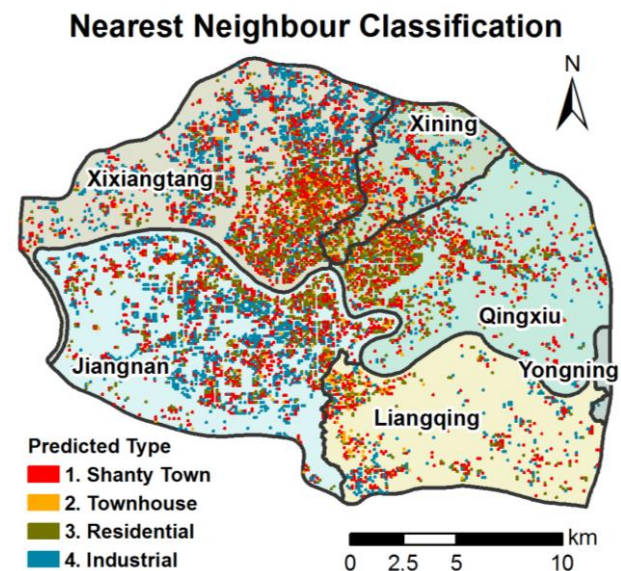


Figure 6. Nearest neighbour classification apply on Nanning

Spatially displayed on maps (Figure 5 and Figure 6), the classification results show similarities and dissimilarities. Cluster 4 in k-mean clustering, which is highly likely to represent type of industrial building found on 4.1, shares the same distribution area as the Industrial Type in nearest neighbour classification. They are found in north Xixiangtang and major part of Jiangnan, which consists with the fact that there are a lot of factories in these regions. As to differences, k-means clustering assigns more Cluster 2 and Cluster 3 in central Nanning City, while nearest neighbour classification predicts more on Shanty Towns.

The dissimilarities could be explain by several aspects. First, even with equivalent samples in each types during training process, nearest neighbour method only mark 5.7% parcels as Townhouse, which implies their scarceness. As mentioned in 3.3, k-means could misclassify parcels into Cluster 2 due to the type amounts imbalance. Moreover, in nearest neighbour classification, 927 (10.6%) parcels have equally high probabilities among 4 types, most of which are predicted as Shanty Town. It could explain the excess of Shanty Town type in nearest neighbour classification comparing to Cluster 1 in k-means.

5. DISCUSSIONS AND CONCLUSIONS

In this study, samples of shanty town are well classified with 98.2% producer accuracy of unsupervised classification and 73.2% supervised classification correctness. Linkage between shanty town and townhouse is found. The other sample class, such as industrial buildings, reaches 94.2% producer accuracy and 97.0% user accuracy.

Low-rise and mid-rise building areas in Nanning City are classified into 4 types respectively by k-means clustering and nearest neighbour classification, the results of which broadly consist with the urban landscape pattern of Nanning City. Whereas, Cluster 2 and Cluster 3 are interlaced and they are still unable to represent Townhouse or Ordinary Residential Block independently. New type definition should be scripted and more combinations of cluster amounts should be tested.

GLCM have been used to extract building boundary in many studies, yet texture characteristics of buildings and blocks in remote sensing images have not be widely used. Unlike pixel-based recognition and classification methods, GLCM is not tolerant to sample quality, land cover type must be exclusive throughout the sample area. Any disturbing types would lead to steeply decline of classification accuracy. Hence, sample selection is vital in this study.

The results of experiments demonstrate that texture features can be used to subdivide land cover and land use types. Former projects focused more on lining out individual buildings, omitted the solutions that can be easily found in a smaller scale. This study initially establish texture feature descriptions of different types of residential areas, especially low-rise and mid-rise buildings, which would help city administrator evaluate residential blocks and reconstruction shanty towns.

More studies should be done in the future on analysing the transforming relationship between shanty town and townhouses. The regression function between them can be used to determined residential type and evaluate degrees of decrepitness of a residential block. Also, more remote sensing data is needed to test whether GLCM is feasible in lower resolutions when recognizing shanty towns.

ACKNOWLEDGEMENTS

This paper was supported by Guangxi Collaborative Innovation Center of Multi-source Information Integration and Intelligent Processing.

REFERENCES

Baraldi, A., & Parmiggiani, F. (1995). An investigation of the textural characteristics associated with gray level cooccurrence

matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 293-304.

Dasgupta, A., Grimaldi, S., Ramsankaran, R., & Walker, J. P. (2017, July). Optimized glcm-based texture features for improved SAR-based flood mapping. In *Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International* (pp. 3258-3261). IEEE.

Haralick, R. M., Shanmugam, K. and Dinstein, I., 1973. Textural Features for Image Classification, in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973. doi: 10.1109/TSMC.1973.4309314

OpenCV team, 2018. Geographic Resources Open Source Computer Vision Library, OpenCV version 3.4 <https://opencv.org/> (22 March 2018).

Satellite Imaging Corporation, 2017. TripleSat Satellite Image of Nanning City, Guangxi, China. <https://www.satimagingcorp.com> (30 December, 2017).