

GAUSSIAN MIXTURE MODEL OF TEXTURE FOR EXTRACTING RESIDENTIAL AREA FROM HIGH-RESOLUTION REMOTELY SENSED IMAGERY

Juan Gu ^{a,b,*}, Jun Chen ^a, Qiming Zhou ^c, Hongwei Zhang ^a

^aNational Geomatics Center of China, Baishengcun, Zizhuyuan, China - julietgujuan@163.com

^bFaculty of Resource Science and Technology, Beijing Normal University, Beijing, China

^cDepartment of Geography, Hongkong Baptist University, Kowloon Tong, Kowloon, Hong Kong

KEY WORDS: Extraction of Residential Area, Gaussian Mixture Model, Texture, High-Resolution Remotely Sensed Imagery

ABSTRACT:

Using high-resolution remotely sensed imagery to timely detect distribution and expansion of residential area is one of most important jobs of national 1:5 spatial database updating. In view of complicated spatial characters of residential area and working disable of current automatic interpretation methods based on spectral features on high-resolution remotely sensed imagery, a classifier based on Gaussian Mixture Model (GMM) of texture is proposed. The combination of co-occurrence texture features (contrast, entropy, mean, standard deviation and correlation included) and edge density are used to substitute for spectral features as classification features. A mixture density function is employed to represent classes' distribution in texture spaces. And residential areas are extracted through the classification based on GMMs obtained through estimating using Expectation Maximum (EM) algorithm. The proposed method is examined by an IKONOS panchromatic imagery.

1. INTRODUCTION

Residential area occupies a relatively small portion of earth surface, but their exact extent, distribution and expansion have a great concern for governors and planners. In a rapid development region such as eastern china, timely updating spatial database that adequately reflect the change and expansion of residential area is widely recognized as one of the most challenging tasks for an operational GIS (Gu et al., 2005). Remotely Sensed data provide an efficient source of information for detecting and monitoring the change of residential area, especially when images with high-resolution (e.g. IKONOS, SPOT5 and QuickBird) become readily available (Jensen and Cowen, 1999; Zhang et al., 2002; Tatem et al., 2004).

However, with the increase of spatial resolution, between-class spectral confusion and within-class spectral variation were found to increasing for land cover/land use studies (Barsley and Barr, 1996; Shaban and Dikshit, 2001; Coburn and Roberts, 2004). It is impossible to define a spectral homogeneous class as 'residential area' (Zhang et al., 2002). Thus, current automatic interpretation methods assume that different surface materials have different spectral characteristics which work well on medium-low spatial resolution multi-spectral remotely sensed data lost its efficiency on high-resolution remotely sensed imagery (Barnsley and Barr 1996; Zha et al., 2003; Shackelford et al., 2003).

As more and more studies have addressed that the incorporation of spatial information in image classification procedures improves image classification, especially high-resolution PAN image are used (Guindon, 2000; Shackelford et al., 2003; Coburn et al., 2004). Karathanassi et al. (2000) classified built areas into high, medium and sparse density based on statistical measurement of texture using SPOT panchromatic remote sensing images. Shaban and Dikshit (2001) improved the classification in urban areas by use of texture features. Hofmann (2001) combined spectral and textural features to

extract residential area by using eCognition software. Statistical information of imagery was used in all above research, but structural information was overlooked. Zhang et al. (2003) performed a supervised classification based on a single Gaussian probability function based classifier using statistical and structural texture features when study of urban spatial patterns from SPOT panchromatic imagery.

A single Gaussian probability function based classifier assuming that the distribution of residential area in texture spaces is mono-modal. But in fact, it is poorly approximated using single Gaussians (Ünsalan, 2004) for different residential area in the same image may have different texture distribution. In this condition, a single Gaussian distribution fails to describe the probability function of residential area in texture spaces.

GMM is a type of density model which comprises a number of component Gaussian functions. These component functions are combined with different weights to result in a multi-model density. Indeed, if one is allowed an arbitrary number of components, any continuous density function of residential class can be approximated to any desired accuracy. In this paper, a mixture model of texture for extracting residential areas from high-resolution remotely sensed imagery is proposed. The following sections are organized as follows: section 2 introduces the extraction of texture features. Section 3 presents classification method of GMM of texture. Section 4 gives experiments and evaluates the extraction method. Finally, conclusions are given in Section 5.

2. TEXTURE MEASUREMENTS

The Haralick grey-level co-occurrence matrix is one of the most popular second-order statistics for texture processing (Coburn and Roberts 2004). Haralick and Shanmugam (1973) suggested fourteen textural features describe some characteristics of texture based on co-occurrence matrix are computed. Among the fourteen textural features, eight ones are mostly used in

remote sensing imagery analysis: contrast, homogeneity, correlation, entropy, dissimilarity, asymmetry, mean and standard deviation. For details about the GLCM method refers to Haralick *et al.* (1973).

Bayer (2000) divided these eight features into three groups: the 'contrast' group (Contrast, Dissimilarity and Homogeneity), the 'orderliness' group (Asymmetry and Entropy) and the 'descriptive statistics' group (Mean, Standard deviation and Correlation). The texture features in contrast group are correlated with each other, so are the features in the orderliness group.

This research investigates the significance of GLCM in measuring textures of three typical kinds of Chinese residential areas in high-resolution remotely sensed imagery. Nine usual land cover types were selected: new urban residential area, old urban residential area, rural residential area, water, grass land, wood land, road, bare farmland and bare land.

Figure 1 presents three texture features (Contrast, Dissimilarity and Homogeneity) of nine land cover types in contrast group. It is easy to find all three textures can reflect the contrast property of land cover. From the statistical results, it can be concluded that new urban residential area, road, rural residential area, old urban residential area and wood land have more contrast surface and water, grass land, bare farm land and bare land have relatively homogeneous surface. For three texture features are correlated with each other, contrast was selected to measure contrast of surface of residential area because that both three kinds of residential areas have the highest value, water and grass land have zero value in contrast space.

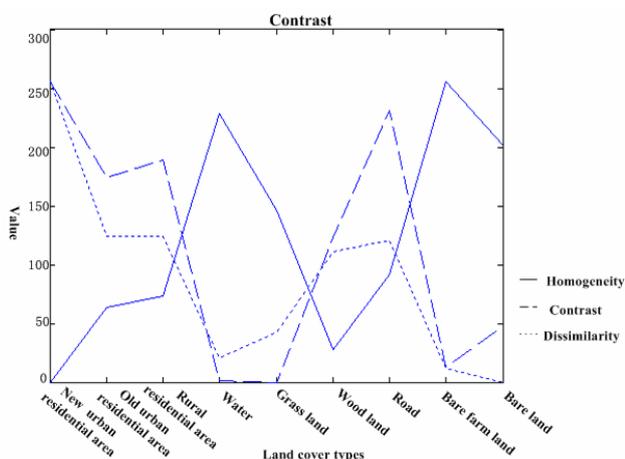


Figure 1 Contrast statistics of nine land cover types in high-resolution remotely sensed imagery

Figure 2 shows two texture features (Asymmetry and Entropy) of nine land cover types in orderliness group. All three kinds of residential areas have less order or uniform surface, wood land and road have medium uniform surface, and water, grass land, bare farm land and bare soil have uniform surface. For the correlation of two features, entropy was selected to measure orderliness of surface of residential area.

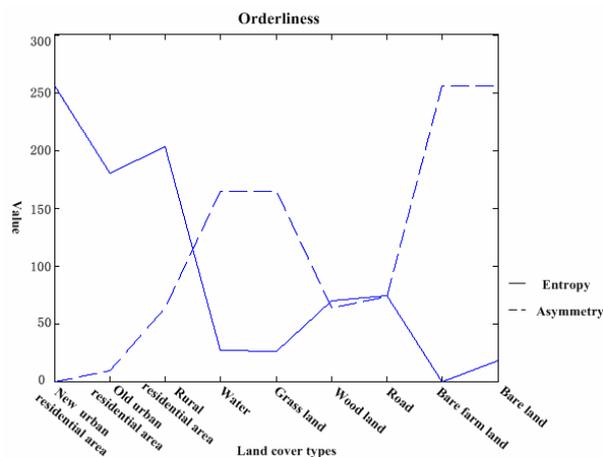


Figure 2 Orderliness statistics of nine land cover types in high-resolution remotely sensed imagery

Figure 3 indicates other texture features based on GLCM: mean, standard deviation and correlation. Mean feature reflect the bright degree of land cover in imagery. It can be found that all three kinds of residential areas have higher mean values than other land cover except bare land. This is somewhat similar to what we see by eyes. Residential areas have larger standard deviation values and smaller correlation values proved that pixels in this areas more discrete and less dependent.

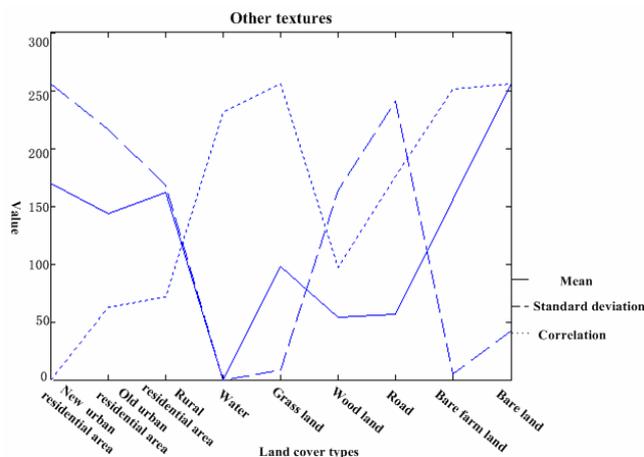


Figure 3 Other textures statistics of nine land cover types in high-resolution remotely sensed imagery

Based on above statistics and analysis, all texture features based on GLCM can describing some texture characters of residential area in high-resolution remotely sensed imagery from different aspect. In order to describe residential area from different aspects, this study selects contrast, entropy, mean, standard deviation and correlation instead of spectral signatures to act as classification features.

However, GLCM texture features cannot reflect edge information of residential area which is important for human eyes. Edge density is added to complement GLCM. Edge density extracted from SPOT XS imagery has been used to improve urban land use classification (Gong and Howarth 1990). In this study, for simplification, edge density was produced

through two steps. Firstly, edge detection by Canny edge detection operator (Canny, 1986) is performed which will produce a binary edge map coded as '1' and '0', with '1' representing edge pixels. Secondly, the count of edge pixels (pixel value is coded as '1') appear in a certain window is act as the edge density in this local region. Fig. 4 shows edge density statistics of nine land cover types. And it can be clearly seen that residential areas have larger edge density value.

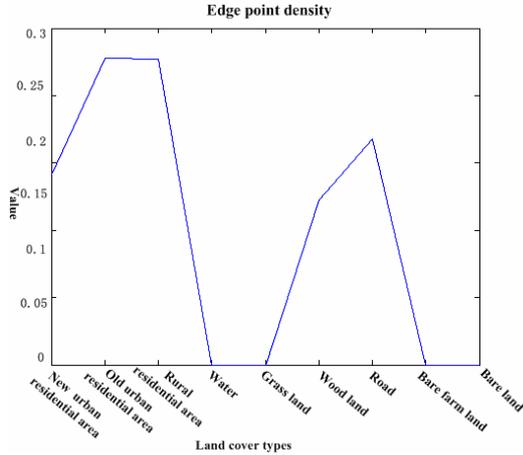


Figure 4 Edge density statistics of nine land cover types in high-resolution remotely sensed imagery

3. THE CLASSIFICATION METHOD BASED ON GMM OF TEXTURES

Since textures are available, we have to combine this information by using a GMM joint probability density function of texture measures for which enables one texture measure to support the other one if it becomes unreliable due to environmental changes.

3.1 GMM of Textures

Given texture vector is denoted as $x = \{x_1, x_2, \dots, x_n\}$

where n is the dimension of the feature vector, in this paper n equals to 6. We model the distribution of all samples by the following formula:

$$p(x | \lambda) = \sum_{i=1}^M \omega_i p_i(x) \quad (1)$$

Where $p_i(x)$ is a normal PDF which is a component of the GMM. It is parameterized by a mean vector μ_i , and a covariance matrix Σ_i :

$$p_i(x) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)\right] \quad (2)$$

ω_i is the weight of the component $p_i(x)$, $0 < \omega_i < 1$ for all components, and $\sum \omega_i = 1$. λ_i is the parameter vector of component i and will be estimated through given training data., $\lambda_i = (\omega_i, \mu_i, \Sigma_i)$. M is the number of mixture components. And $\lambda = (\lambda_1, \dots, \lambda_M) = \{\omega_i, \mu_i, \Sigma_i | i \in M\}$. Mixture model (1) is called the Gaussian Mixture Model (GMM).

In this study, two classes were defined: residential class and background class. Therefore, two GMMs need to be constructed:

$$p_r(x | \lambda_r) = \sum_{i=1}^{M_r} \omega_{ri} p_{ri}(x) \quad (3)$$

$$p_b(x | \lambda_b) = \sum_{i=1}^{M_b} \omega_{bi} p_{bi}(x) \quad (4)$$

Where formula (3) is GMM of residential class and formula (4) is GMM of background class. λ_r and λ_b are corresponding parameter vectors.

3.2 Estimation of Parameters of GMM

Parameters of GMM model are estimated from training samples for which we know the classes. Maximum-likelihood (ML) estimation and Bayesian estimation are commonly used approaches. While there are strong theoretical and methodological arguments supporting Bayesian estimation, in practice the ML estimation is simpler and, when used for designing classifiers, can lead to classifiers nearly as accurate (Paalanen *et al.*, 2005). Many implementation issues support the selection of maximum-likelihood estimation.

The expectation maximization (EM) algorithm is an iterative method for calculating maximum likelihood distribution parameter estimates from incomplete data (elements missing in feature vectors) (Bilmes, 1997). The algorithm can also be used to handle cases where an analytical approach for maximum likelihood estimation is infeasible, such as Gaussian mixtures with unknown and unrestricted covariance matrices and means. Although the EM algorithm has some limitations (e.g. it is not guaranteed to converge to a global rather than a local maximum of the likelihood), it is generally efficient and effective for the parameters' estimation of GMM.

EM algorithm is involving two steps: E-step and M-step. For a GMM and a feature vector $x = \{x_1, x_2, \dots, x_N\}$, suppose that

$\lambda^{(t)}$ denotes the estimation of λ obtained after the t th iteration of the algorithm. Then at the $(t+1)$ the iteration, the E-step calculates the expected sample data log-likelihood function

$$Q(\lambda, \lambda^{(t)}) = \sum_{k=1}^N \sum_{m=1}^M \{\log \omega_m p(x_k | \lambda_m) P(m | x_k; \lambda^{(t)})\} \quad (5)$$

Where $P(m | x_k; \lambda^{(t)})$ is a posterior probability and is computed as

$$P(m|x_k; \lambda^{(t)}) = \frac{\omega_m^{(t)} p(x_k | \lambda_m^{(t)})}{\sum_{l=1}^M \omega_l^{(t)} p(x_k | \lambda_l^{(t)})} \quad (6)$$

The M-step finds the $(t+1)$ th estimation $\lambda^{(t+1)}$ of λ by maximizing $Q(\lambda, \lambda^{(t)})$

$$\omega_m^{(t+1)} = \frac{\sum_{n=1}^N P(m|x_k; \lambda^{(t)})}{N} \quad (7)$$

$$\mu_m^{(t+1)} = \frac{\sum_{k=1}^K x_k P(m|x_k; \lambda^{(t)})}{\sum_{k=1}^K P(m|x_k; \lambda^{(t)})} \quad (8)$$

$$\sum_m^2 (t+1) = \frac{\sum_{k=1}^K P(m|x_k; \lambda^{(t)}) (x_k - \mu_m^{(t+1)}) (x_k - \mu_m^{(t+1)})^T}{\sum_{k=1}^K P(m|x_k; \lambda^{(t)})} \quad (9)$$

Check for convergence of the resulting log likelihood function (5) or a ‘sufficiently small’ change in the estimated parameters. If convergence is reached, then stop. Otherwise, repeat.

3.3 Classification Based on GMMs

As soon as the distributions and the associated parameter vectors of residential class and background class are known. Then the classification is performed based on Bayesian decision theory. An unlabeled pixel is assigned the class according to posterior probabilities or decision risks (10)

$$\begin{cases} P_r(x|\lambda_r) \geq P_b(x|\lambda_b) & x \in \text{residential class} \\ P_r(x|\lambda_r) < P_b(x|\lambda_b) & x \in \text{background class} \end{cases} \quad (10)$$

4. EXPERIMENTS AND RESULTS

4.1 Test Site and Image Data

The proposed method was evaluated by Wangjing District which is located in the north-east fringe of Beijing City. It is one of the biggest residential districts of Beijing city with a population over 120 thousands. It is experiencing rapid development with the increasing expansion of Beijing city. In the past, it was rural area. Most people lived in Chinese style courtyard houses- bungalows with small yard. With the development of urbanization, a lot of people working in downtown of Beijing City chose to install their houses in this places. A lot of new medium to high-rise residential and/or office buildings were recently built up. Some are still under construction. IKONOS Panchromatic imager acquired on 26 April 2001 was used for it is the highest resolution images available for us. Samples of different types of residential area

contained in three test sites were also magnified for authors to have a clear seeing, respectively.

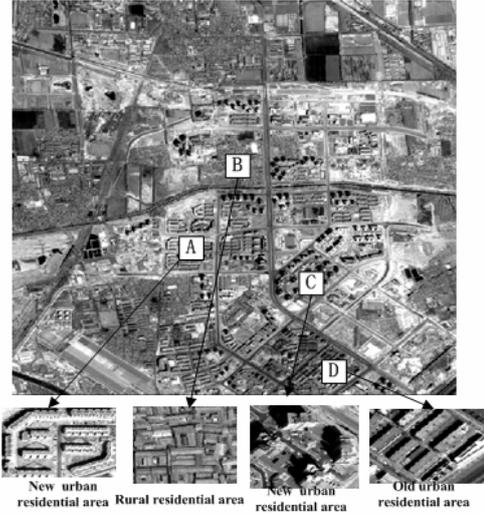


Figure 5 A sub-scene of Wangjing District (2719*2449 pixels)

4.2 Extraction of Texture Features

To compute second-order statistical features (entropy, dissimilarity, mean value, standard deviation and correlation) from a grey level co-occurrence matrix, firstly the quantization level (QL) of images was reduced to 5 bits (grey level 0-31). The reason for reducing QL was that most of the texture features are calculated by forming a matrix, the order of which is dependent on the QL of the image and this is expensive to store and process (Shaban and Dikshit, 2001). Marceau et al. (1989) have also shown that QL is not significant in affecting texture classification accuracies. The QL for test images, therefore, reduced to 5 bits for computing texture feature only by following a histogram equalization procedure explained by Haralick et al. (1973). Secondly for a trade-off between our ability to discriminate classes and the accuracy of boundary estimation, the window size 13*13 was used as window unit to compute texture values of a pixel. Thirdly the distance metric is setting to 1 pixel according to the observations (Rosenfeld, 1982). Finally the mean value of texture features in four principal directions, horizontal, vertical, right-diagonal and left-diagonal, was calculated and used as a texture feature (Haralick et al. 1973). The computation of edge density is also based on the same window unit.

4.3 The Estimation of Models' Parameters

To ensure better estimation of models parameters of residential area and background, sufficient training data were selected from test sites. Training sample set of residential area should include as many representative types as possible residential area in the test sites. Training sample set of background should include all kinds of land cover types of background such as bail soil, water, grass, trees and roads. Both samples were generated from contiguous groups of pixels belonging to distinct land cover classes. They were extracted manually with the help of IKONOS on the false color composite imagery. The training samples of residential class were 394615 pixels and those of background class were 398028 pixels.

Before the estimation of model's mixture parameters, the number of mixture components should be set. In order to find the best number of mixture components, a set of number 8, 16, 32, 64, 128, 256 and 512 were used to estimate the model's mixture parameters by training samples and extract residential area. And accuracy assessment was also made for each result. Relationship between number of mixture components and accuracy was presented in Figure 6. Accuracy of each test site improved greatly at with the increasing of number of mixture components until 256. When number is set to 512, the accuracy made little improvement compared to the accuracy of 256. Therefore, mixture parameters of all GMM models in this paper are set to 256. The mixture parameters of the models were obtained under the condition of 256 mixture components and training samples.

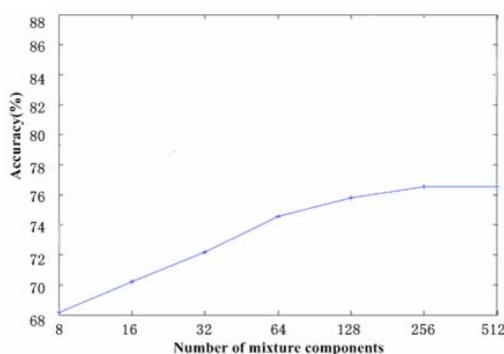


Figure 6 Relationship between number of mixture components and accuracy

4.4 Classification Results and Accuracy

We extracted residential areas from the image based on the method proposed above (Figure 7). In order to compare the performance of the proposed GMM method, the texture images were also classified using a single Gaussian based classifier. Table 1 presents the accuracy assessments and comparison of two classifiers. It can be seen that the accuracy based on GMM method is significantly higher than that based on a single Gaussian method.



Figure 7 Residential extraction result of Wangjing District by GMMs

	GMM based classifier	Single Gaussian based classifier
Residential area		
PA	77.82%	69.57%
UA	72.14%	61.54%
Background		
PA	75.50%	62.96%
UA	80.67%	70.83%
OA	76.53%	66.22%
Kappa	0.63	0.32

PA: Producer's accuracy UA: User's accuracy
OA: Overall accuracy

Table1 Classification accuracy assessments and comparison of GMM based classifier and single Gaussian based classifier

4.5 Post Processing of Classified Residential Data

In the classification results, those pixels have high texture values in background class such as edges of roads, barren land and noise were likely to be classified to residential class which made "salt and pepper" in the background. And at the same time, these land cover types in residential area were classified as the background class which made the residential area classified seemed fragmented. For this reason, post-classification based on mathematical morphological operation was implemented through three steps. The first step is to remove small parts and filling small holes by combination of erosion and dilation operations. The second step is to smooth residential area's boundaries. There are a lot of spurious peaks and pits of the boundaries of residential area. Opening operator removes small "necks" that connect larger regions, and the smaller spurious regions. Closing operator was used to remove the spurious pits. The third step is to implement area filtering and big holes filling. In this step, all connected components (objects) that have fewer pixels than threshold were removed. And holes inside residential area smaller than pre-defined threshold were also removed.

In order to make the classification results can be used to update GIS database, we converted the raster binary classification results into vector polygons through operators provided by ERDAS 8.6. For a clear eye seeing, we overlay vector results to corresponding image and it is shown in Figure 8. It can be seen that most residential area were extracted and their shape kept correctly.

5. CONCLUSIONS

This paper developed an extraction method of residential area from high resolution remotely sensed imagery using classification based on GMM of texture. Texture derived from GLCM: contrast, entropy, mean, standard deviation and correlation and edge density were selected to measure texture of residential area in high resolution remotely sensed imagery. GMM is used to instead of traditional single Gaussian model in classification procedure.

The method proposed in this research was tested over IKONOS panchromatic imagery. It can be found higher extraction result. And the overlay of vector polygons obtained from extraction results with original images also showed that the approach works well on extraction of complicated residential area from high spatial resolution remotely sensed image.

As present, experiments are being conducted to apply the method on IKONOS PAN imagery. This research needs to be tested in other environment. In the following studies, multispectral information will also be added to obtain a better result.



Figure 8 Overlay of vector polygons of residential areas extracted with original image

REFERENCES

- Barnsley, M.J., and Barr, S.L., 1996, Inferring urban land use from satellite sensor images using kernel-based spatial reclassification, *Photogrammetric Engineering and Remote Sensing*, 62, pp. 949-958
- Bayer M., 2000, GLCM texture: a tutorial: available on line at http://www.ucalgary.ca/~mhallbey/texture/texture_tutorial.html (22 July 2002)
- Canny, J.F., 1986, A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), pp. 679-698
- Coburn C.A., Roberts A.C.B., A multiscale texture analysis procedure for improved forest stand classification, *International Journal of Remote Sensing*, 20 October, 2004, 25(20), pp. 4287-4308
- Gong P. and Philip J. Howarth, 1990, The use of structural information for improving land-cover classification accuracies at the rural-urban fringe, *Photogrammetric Engineering and Remote Sensing*, 56(1), pp. 67-73
- Gu J., Chen J. and Zhou Q.M., 2005, A hierarchical object-oriented approach for extracting residential areas from high resolution imagery, ISPRS Hannover Workshop 2005: High-resolution Earth Imaging for Geospatial Information, Hannover, Germany, 17-20, May, 2005
- Guindon, B., Zhang, Y., Dillabaugh, C., 2004, Landsat urban mapping based on a combined spectral-spatial methodology, *Remote Sensing of Environment*, 92, pp. 218-232
- Haralick, R.M., Shanmugam, K. and Dinstein, I., 1973, Textural features for image classification. *IEEE Transactions on System, Man, and Cybernetics*. SMC-3, pp. 610-621
- Hofmann, P., 2001, Detecting informal settlements from IKONOS image data using methods of object oriented image analysis – an example from cape town (South Africa), In: Jurgens, Carsten (Editor): *Remote Sensing of Urban Areas*, Regensburg, pp. 107-118
- Jensen, J.R. and Cowen, D.C., 1999, Remote Sensing of urban/suburban infrastructure and socio-economic attributes, *Photogrammetric Engineering and Remote Sensing*, 65, pp. 611-622
- Karathanassi, V., Iossifidis, CH., and Rokos, 2000, A texture-based classification method for classifying built areas according to their density, *International Journal of Remote Sensing*, 21(9), pp. 1807-1823
- Shaban, M.,A., and Dikshit, O., 2001, Improvement of classification in urban areas by the use of textural features: the case study of Lucknow city, Uttar Pradesh, *International Journal of Remote Sensing*, 22(4), pp. 565-593
- Shackelford, A.K. and Davis, 2003, A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas, *IEEE Transactions of Geoscience and Remote Sensing*, 41(10), pp. 2354-2363
- Tatem A.J., Noor A.M. and Hay S.I., 2004, Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery, *Remote Sensing of Environment*, 93, pp. 42-52
- Ünsalan C. and Boyer K.L., 2004, Classifying land development in high-resolution panchromatic satellite images using straight-line statistics, *IEEE Transactions on Geoscience and Remote Sensing*, 42(4), pp. 907-919
- Zhang Q., Wang J., Peng X., Gong P. and Shi P., 2002, Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data, *International Journal of Remote Sensing*, 23(15), pp. 3057-3078
- Zhang, Q., Wang, J., Gong, P., and Shi., Pei, 2003, Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis, *International Journal of Remote Sensing*, 24(21), pp. 4137-4160

ACKNOWLEDGEMENTS

This project is financed by two items of the National Natural Science Foundation of China (Contract No. 40337055 and Contract No. 40501062)