

HIGH-RESOLUTION REMOTE SENSING IMAGE BUILDING EXTRACTION BASED ON MARKOV MODEL

Wenhao Zhao ^{1,2}, Li Yan ¹, Yutao Chang ³, Lincheng Gong ³

¹ Wuhan University School Of Geodesy and Geomatics, 430079, Wuhan, China - liyan@sgg.whu.edu.cn

² National Geomatics Center of China, 100830, Beijing, China - 280842460@qq.com

³ The Second Surveying and Mapping Institute of Xinjiang Uygur Autonomous Region, 830002, Xinjiang, China - (1139811285, 43901484)@qq.com

Commission III, WG III/1

KEY WORDS: Markov Model, Contourlet Transform, Map Clustering, Spectral Feature Index, Building Extraction, High Resolution Remote Sensing Image

ABSTRACT:

With the increase of resolution, remote sensing images have the characteristics of increased information load, increased noise, more complex feature geometry and texture information, which makes the extraction of building information more difficult. To solve this problem, this paper designs a high resolution remote sensing image building extraction method based on Markov model. This method introduces Contourlet domain map clustering and Markov model, captures and enhances the contour and texture information of high-resolution remote sensing image features in multiple directions, and further designs the spectral feature index that can characterize "pseudo-buildings" in the building area. Through the multi-scale segmentation and extraction of image features, the fine extraction from the building area to the building is realized. Experiments show that this method can restrain the noise of high-resolution remote sensing images, reduce the interference of non-target ground texture information, and remove the shadow, vegetation and other pseudo-building information, compared with the traditional pixel-level image information extraction, better performance in building extraction precision, accuracy and completeness.

1. INTRODUCTION

In recent years, with the rapid development of sensor technology, the resolution of remote sensing images has been continuously improved, and high-quality remote sensing images are playing an increasingly important role in monitoring global changes, regional changes, and environmental changes, as urban planning, population estimation, and Land use analysis, map updating, 3D modelling, digital cities and smart cities are important bases for the work. Building extraction has also become an important content of high resolution remote sensing image processing. However, high-resolution remote sensing images also bring more challenges while providing richer building information.

This is mainly due to the following reasons. One is that the efficiency and precision of building extraction cannot be effectively guaranteed. There are phenomena of "same spectrum foreign matter" and "homologous spectrum" in remote sensing images, which are more obvious in high-resolution remote sensing images. The results obtained by traditional extraction methods, such as image element classification, are relatively fragile and tend to appear, similar "pseudo-information" of noise, combined with a sharp increase in the amount of image data, greatly reduces the efficiency and accuracy of the extraction. The second is that the accuracy and integrity of building extraction are more vulnerable. Because of the significant increase in the amount of information on the types of surfaces and details of textures in high-resolution remote sensing images, not only the complex structural information of the buildings itself has caused many obstacles for extraction, but also other artificial features and natural features around buildings have become important disturbances affecting the accuracy and completeness of building extraction.

Based on the above problems, traditional pixel-level remote sensing image processing methods based on low-and medium-resolution remote sensing images cannot meet the needs of high-resolution remote sensing image data processing and application. It is necessary to fully consider the high-resolution remote sensing image-rich On the basis of texture features, the study of building extraction methods. There are various extraction methods based on texture features, which can be divided into four types of methods: statistics, model, signal processing and structure (LIU Li et al. 2009).

Among them, the model-based extraction method uses the method of modeling the texture image to reduce the image texture feature extraction to parameter estimation, which can take into account the local randomness of the texture and the overall regularity, and has a good extraction effect, which is one of research hotspots at home and abroad, and also is the research emphasis of this article. In the model-based extraction method, the Markov model is the most widely used. Li Xia et al. (LI Xia et al. 2014) proposed a multi-spectral image de-noising method based on wavelet-domain hidden Markov tree model, which fully embodies the non-Gaussianity and persistence characteristics of wavelet coefficients (Jiao Licheng et al. 2006), and accurately describes the statistical properties of the image data wavelet transform lay the foundation for image denoising, image segmentation and feature extraction. Tang Wenjing et al. proposed an image segmentation algorithm based on Markov model in the wavelet domain combined with Markov Random Field (MRF)-based segmentation algorithm, and established the corresponding boundary marker field, and with the help of Bayesian framework and SMAP criteria to achieve image segmentation and feature extraction (Tang Wenjing et al. 2014). Xu et al. proposed a method of image extraction based on the visual perception of the outside world's powerful ability to

perceive and recognize visual neural perception and combined Markov random field model (Xu Miao Zhong et al. 2015). Through experiments, this method has a good performance in high-resolution remote sensing image extraction tasks.

Based on the above problems and research, this paper designs a Markov model-based high-resolution remote sensing image building extraction method. Firstly, through the global optimization of the spectral space of high-resolution remote sensing images, the noise information is filtered, and the image contour and boundary information are enhanced; Then, the multi-scale Markov model-based remote sensing image segmentation is used to combine Contourlet transform, graph clustering and Markov model. With the fully consideration of the multi-directionality and anisotropy of the high-resolution remote sensing imagery feature information. Multi-scale Segmentation and Classification of Remote Sensing Images to Obtain Target Sets for Basic Features such as Buildings, Roads, Concrete Floors, Vegetation, Water, and Bare Soil; Then, using the K-adjacent classification method, the non-building areas such as roads, cement floors, and vegetation were removed and relatively independent building areas were extracted, as the basis for the subsequent extraction of the buildings, and finally, fully consider the spectral characteristics of buildings, shadows, and vegetation information that may exist in the building area and that are adjacent to each other, or even overlap, define and calculate the spectral characteristic index of the building that can significantly describe the degree of the building in the building area, and further remove the shadow, vegetation, and other pseudo-building information, and finally obtain accurate and complete buildings.

This method can suppress the noise of high-resolution remote sensing image and the interference of non-target texture information, effectively remove the shadows, vegetation, and other pseudo-building information adjacent to the building, and complete the fine extraction of buildings and the extracted building results has good precision, accuracy and completeness, which achieved good results.

2. THIS ARTICLE'S BUILDING EXTRACTION METHOD FLOW

As shown in Figure 1, the high-resolution remote sensing image building extraction method based on Markov model mainly includes

a. High-resolution remote sensing image spectral space optimization. The initial high-resolution remote sensing images were subjected to fine geometric correction, Gaussian low-pass filtering and image sharpening, filtering image noise, and enhancing the boundary information of the features to obtain a remote sensing image with optimized spectral space.

b. Remote sensing image segmentation based on multi-scale Markov model. There are three steps:

First, Contourlet transforms the image. The remote sensing image is transformed into a Contourlet domain with multi-resolution, local positioning, multi-directionality, near-neighbor sampling and anisotropy. The Contourlet coefficients in multi-scale and multi-directions are used to effectively capture the edge contours in the image.

Secondly, for the coarsest-scale low-frequency image after Contourlet transform, we use the spectral clustering algorithm

to segment the image to obtain a more reliable initial segmentation result.

Finally, image segmentation based on Markov model. Contourlet coefficients are used to construct the context model of image information, based on the obtained results of the coarsest-scale low-frequency image segmentation, the segmentation of high-resolution images is achieved by the transfer of segmentation information between the segments, and a set of fundamental object objects in the image is acquired.

c. Building area extraction based on K neighbor classification method. By using the K-adjacent classification method, a building area is extracted from the basic object collection obtained in the process 2 as a basis for the fine extraction of the building.

d. Building extraction based on spectral feature index. Through the use of spectral features of buildings, shadows, and vegetation, define and calculate the index of spectral characteristics of a building that can describe a significant degree of a building in a building area and eliminate shadows, vegetation, and other pseudo-building information, and finally obtain accurate and complete buildings.

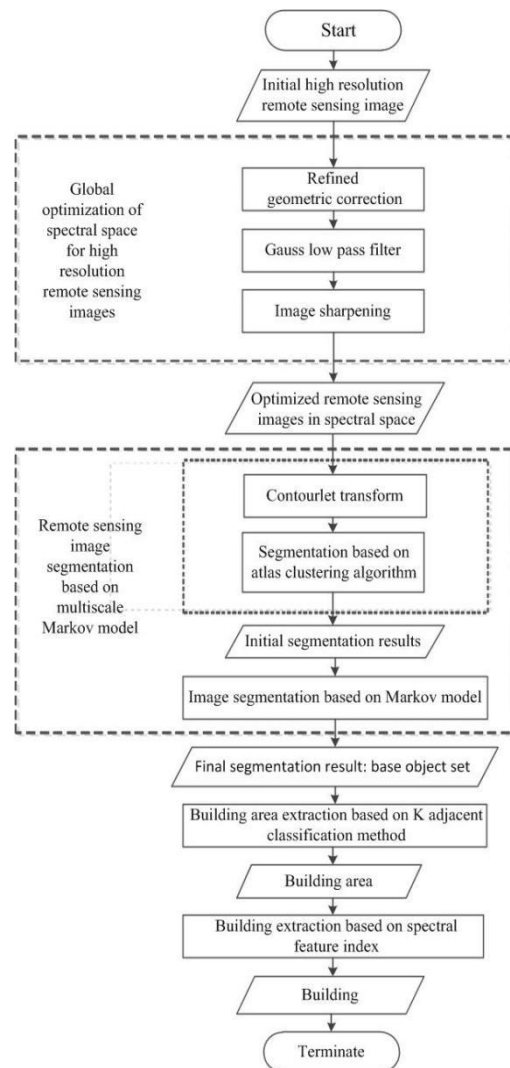


Figure 1. The process of building extraction in this paper

3. HIGH-RESOLUTION REMOTE SENSING IMAGE SPECTRAL SPACE OPTIMIZATION

The global optimization of spectral space is a necessary prerequisite for the analysis and processing of high-resolution remote sensing images. On the one hand, the original high-resolution remote sensing image generally contains a large amount of invalid noise information. If it is not filtered, it will easily cause unnecessary interference to subsequent image processing and reduce the accuracy and completeness of building extraction. On the other hand, due to the influence of atmospheric conditions and shooting conditions, the spectral information of remote sensing images is generally difficult to uniformly distribute, but often concentrated in certain spectral ranges, the contrast of the images is small and the contrast is poor, and it is not necessary to capture the texture information including the edges in the processing process, therefore, the global optimization of the spectral space must first be performed on the remote sensing image.

The specific method is:

a. Image Refinement Geometry Correction. The Near-point interpolation is used to resample high-resolution remote sensing images, by fine geometrical correction reduce the geometrical distortions such as uneven column, inaccurate correspondence between pixel size and ground size, and irregular shape change of the ground features due to sensors, terrain relief, etc.

b. Use Gaussian low-pass filtering to reduce noise. Gaussian low-pass filtering is used to process the image and the noise of filter high-frequency parts in the remote sensing image.

c. Image sharpening. Using Laplacian sharpening algorithm for image sharpening, enhance the edge and boundary information of the feature that is weakened after applying Gaussian low-pass filtering in the remote sensing image, increase the gray difference between the center and the edge of the feature, and highlight the effective feature information, to prepare for subsequent image analysis and processing. Among them, the transform formula of Laplacian sharpening algorithm is:

$$g(i, j) = 5f(i, j) - f(i-1, j) - f(i+1, j) - f(i, j+1) - f(i, j-1) \quad (1)$$

The sharpening operator is:

$$\begin{matrix} 0 & 1 & 0 \\ -1 & 5 & -1 \\ 0 & 1 & 0 \end{matrix} \quad (2)$$

By performing the above-mentioned spectral spatial global optimization operation on the high resolution remote sensing image to be processed, the following effects can be obtained:

First, a large number of invalid noises in the high frequency part of the image are filtered, which preserves and highlights the spectral features of the effective information in the image, ensuring the integrity and coherence of the texture information.

Second, enhanced the image edge information and gray transitions. So that the image contour becomes clear and the ground objects are prominent, which makes the ground objects easier to be identified and analyzed, and provides the basis for subsequent segmentation of remote sensing images based on multi-scale Markov models.

4. REMOTE SENSING IMAGE SEGMENTATION BASED ON MULTISCALE MARKOV MODEL

High-resolution remote sensing images provide abundant geographic and topographical information and spatial information, compared with medium and low-resolution remote sensing images, the information on their terrain is more abundant. In addition to buildings, there are roads, concrete floors, water conservancy facilities, electricity and other artificial features, the geometry structure, spectral information, texture information, and spatial location of these features are complex and changeable, which can easily cause interference in the extraction of buildings. Therefore, the effective identification and segmentation various feature categories in the high-resolution remote sensing imagery, set up of mutually independent ground areas, and remove of other artificial features and natural features that are not in the gathering area of the building are the key to ensure the accuracy and completeness of building extraction in high-resolution remote sensing images.

To solve this problem, this paper introduces the image segmentation technology into the building extraction process, and designs the image segmentation and classification as the key link. The remote sensing imagery is divided into several specific, unique, disjoint areas such as buildings, roads, concrete floors, vegetation, and water bodies (Pedro F. Felzenszwalb et al. 2004). To ensure that in the final building extraction process, the impact of the construction area features will be removed, and the building's extraction efficiency, accuracy and accuracy will be greatly improved.

In the process of remote sensing image segmentation based on multi-scale Markov model, the paper mainly includes Contourlet transform, map clustering and image segmentation based on Markov model. Specifically, as described below.

4.1 Contourlet Transform

Traditional pixel-level remote sensing image processing methods cannot meet the processing requirements of high-resolution remote sensing images. This is mainly due to two reasons. First, high-resolution remote sensing image has large data and high computational load; Second, there are abundant information on the geometry and texture of the objects, and the shape and directionality of the objects are complex and their correlation is large.

To solve these two difficulties, the building extraction method must be designed from two perspectives:

First, build a multi-level, multi-resolution remote sensing image space, calculate by layer-by-layer, resolution-by-resolution, reduce computational difficulty, and improve the efficiency and accuracy of high-resolution remote sensing imagery feature extraction.

Second, through the frequency domain processing method, describe the geometric information of the ground objects in the remote sensing image, especially the directional information that can characterize the geometric information, in order to judge the same type of ground objects and improve the accuracy of the high-resolution remote sensing image feature extraction. After research, the Contourlet transform is a two-dimensional image representation method with multi-resolution, local orientation, multi-directionality, near-neighbor sampling and anisotropic properties (Do MN et al. 2005). The basis function is distributed in multi-scale, multi-directions and a small

number of coefficients can effectively capture the edge contour in the image, and the edge contour is the main feature in the natural image (Li Hui et al. 2016). Therefore, the frequency domain processing method adopted in this paper is Contourlet transform, and based on this, the remote sensing image segmentation based on multi-scale Markov model is realized.

For specific method, first by the Laplacian filter structure (LP) to perform multi-scale decomposition on the image to capture singular points, and then apply a two-dimensional direction filter bank (DFB) to each of the LP decompositions. On the first-order high-frequency component, it can be decomposed at any scale to obtain the n-fold number of directional sub-bands. The high-pass sub-band input DFB generated by the LP sub-band decomposition each time gradually adds the points singularly into a linear structure, thereby capturing the outline in the image.

The transformation process is shown in Figure 2:

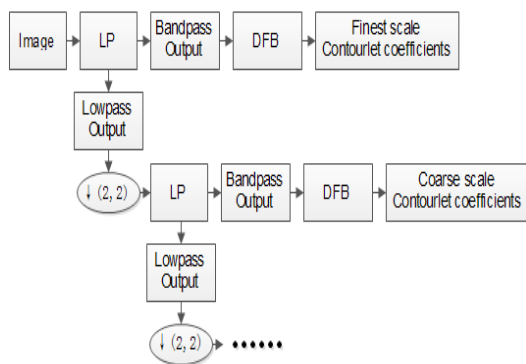


Figure 2. The Contourlet transform diagram

By applying the Contourlet transform to high-resolution remote sensing images, results can be provided for subsequent image processing:

One is to provide Contourlet domain image processing results for image segmentation based on Markov models and to construct multi-resolution, i.e., multi-scale image space, in order to build Markov models using Contourlet coefficients with adjacent correlations, thereby constructing an adaptive context, implement different resolution remote sensing image segmentation;

The second is to provide the most coarse-scale low-frequency remote sensing images, and to obtain more reliable initial segmentation results by applying a graph clustering algorithm, which provides a rough-to-fine calculation basis for the image segmentation process based on the Markov model.

4.2 Map Clustering

To perform image segmentation based on Markov model, in addition to constructing a multi-resolution image space by Contourlet transform, the segmentation results of the coarsest-scale low-frequency remote sensing image are also required as the original in the coarsest (low) resolution Contourlet domain image space. The basics of segmentation, from coarse to fine, are calculated through continuous loop iterations to finally obtain segmentation results of high-resolution remote sensing images.

The graph clustering algorithm is a high-performance data analysis method (Lijun Wang et al. 2012), this algorithm has the

ability to identify non-convex distributions, and has nothing to do with the scale (ie, the resolution in this paper) (CUI Weiwei et al. 2013). It is only related to the number of pixels, and the algorithm is simple and efficient, it is very suitable for the application of the coarsest low-frequency remote sensing images in this paper. The specific steps are as follows:

- Construct a weighted undirected graph with the sample point set as the vertex set, and construct a similarity matrix W of the data point set based on the pixel similarity measure;
- Calculate Laplacian matrix L ;
- Calculate the eigenvalues and eigenvectors of the matrix L ;
- Mapping sample points into a low-dimensional map space determined based on one or more eigenvectors;
- The classification is based on the representation of sample points in the map space.

In this process, the segmentation results of the coarsest-scale low-frequency remote sensing image obtained by applying the spectral clustering method will be used as a reliable initial segmentation result. It is applied in the segmentation process of the next section based on the Markov model, and serves as the starting classification step as the first step for layer-by-layer, resolution-by-resolution, scale-by-scale iterative calculations.

4.3 Image Segmentation Based on Markov Model

As described above, after obtaining the Contourlet domain image and the segmentation results of the coarsest-scale low-frequency remote sensing image, the key links of this paper are needed, image segmentation based on the Markov model is used to obtain buildings, roads, cement floors, vegetation, and water bodies. The collection of base objects that are not intersected and independent from one another, and further the acquisition of construction areas that are essential for building extraction unless the construction area interferes with the area.

In comparison, in the segmentation results obtained, the segmentation result on the coarse scale is more robust, while on the fine scale, the segmentation resolution is higher. In order to obtain a robust and high-resolution segmentation image at the same time, this article abstracts the ground feature information in remote sensing image as context information, references the adaptive context model for inter-scale fusion, and divides the images of different texture types (Fan Guo-liang et al. 2001).

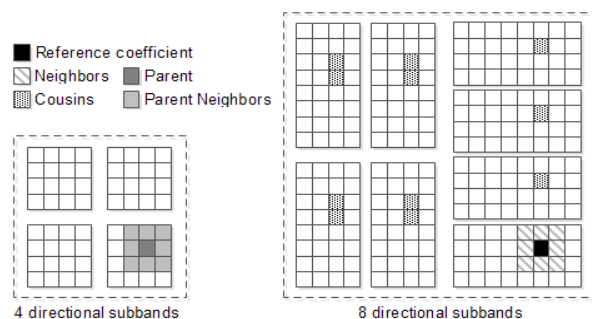


Figure 3. The relationship among Contourlet coefficients

First, we need to construct a context model. In the context model, the transfer of segmentation information between scales

is achieved based on Contourlet coefficient relationship. As shown in Figure 3, for any Contourlet coefficient X, the eight adjacent coefficients in the same sub-bands are defined as neighboring coefficients (NX), the coarse scale coefficients at the same position are defined as father coefficients (P), and the 8 adjacent coefficients of the father's coefficient are defined as the father's neighbor coefficient (NP). Because the Contourlet transform obtains multiple directional sub-bands, the Contourlet coefficients also have cousin coefficients (CX) compared to the wavelet coefficients, which are defined as coefficients in different directions at the same position on the same scale.

Second, we must construct the multi-scale Markov model of Contourlet domain.

Define a 2-dimensional context vector $\text{context} = (\text{context1}, \text{context2})$. Where context1 describes the statistical correlation between the scales and context2 describes the statistical correlation within the scales. If an image is first decomposed by a two-level pyramid and then decomposed into four and eight directional sub-bands, then for each Contourlet coefficient, context1 is determined by the category of father P and its eight neighbor coefficients NP, whose value c1 is taken as the majority of the categories in P and NP. The value c2 of context2 is determined by the neighbor coefficient NX and the cousin coefficient CX.

Assume $Y^{(j)} = \{y_s^{(j)}\}$, represents the set of all transform coefficients of j-scale, $X^{(j)} = \{x_s^{(j)}\}$, represents the set of all class labels of j-scale, $V^{(j)} = \{v_s^{(j)}\}$, represents the set of all context vectors for the j-scale. It is assumed that the joint distribution of the class index $X^{(j)}$ on the j scale is uniquely determined by the context vector $V^{(j)}$, and under the given $x_s^{(j)}$ conditions, $y_s^{(j)}$ and all of $y_t^{(j)}$ the $t \neq s$ are conditionally independent, then $v_s^{(j)}$, $x_s^{(j)}$ and $y_s^{(j)}$ form the Markov chain. The maximum a posteriori classification result of $Y^{(j)}$ can be obtained by maximizing the posterior probability.

$$p(X^{(j)}|Y^{(j)}, V^{(j)}) = \frac{f(Y^{(j)}|X^{(j)})p(X^{(j)}|V^{(j)})}{f(Y^{(j)}|V^{(j)})} = \frac{1}{f(Y^{(j)}|V^{(j)})} \cdot \prod_{s \in S^{(j)}} [f(y_s^{(j)}|x_s^{(j)})p(x_s^{(j)}|v_s^{(j)})] \quad (3)$$

then,

$$p(x_s^{(j)}|y_s^{(j)}, v_s^{(j)}) \propto f(y_s^{(j)}|x_s^{(j)})p(x_s^{(j)}|v_s^{(j)}) \quad (4)$$

among them, $f(y_s^{(j)}|x_s^{(j)})$ is the likelihood value of $y_s^{(j)}$ a coefficient given the class label. Prior probability $p(x_s^{(j)}|v_s^{(j)})$ provides category information $x_s^{(j)}$ for given $v_s^{(j)}$.

If the number of classifications is K, 2K different values can be taken for each context vector. To reduce model parameters, assume that the same scale of transform coefficients have the same model parameters, definitions $\gamma_{j,k}$ and $\alpha_{j,\text{context}_{\gamma,k}}$ are $p(x_s^{(j)} = k)$ and $p(x_s^{(j)} = \text{context}_{\gamma}|x_s^{(j)} = k)$, for all s belong to scale j, $k \in \{1, \dots, K\}$ and $k \in \{1, \dots, 2K\}$.

Finally, the image segmentation method based on the Markov model is calculated. The specific steps are:

a. Initialization $\gamma_{j,k}$ and $\alpha_{j,\text{context}_{\gamma,k}}$, calculate the context vector $v_s^{(j)}$ according to the segmentation result of the coarsest low-frequency image (or the previous coarse scale) and the Contourlet coefficients of the scale;

b. Calculated using equation (4);

$$p(x_s^{(j)} = k|y_s^{(j)}, v_s^{(j)}) = \frac{\gamma_{j,k} \alpha_{j,\text{context}_{\gamma,k}} f(y_s^{(j)}|x_s^{(j)} = k)}{\sum_{m=1}^K \gamma_{j,m} \alpha_{j,\text{context}_{\gamma,m}} f(y_s^{(j)}|x_s^{(j)} = m)} \quad (5)$$

c. Update parameters according to the following formula

$$\gamma_{j,k} = \frac{1}{2^{2j}} \sum_{s \in S^{(j)}} p(x_s^{(j)} = k|v_s^{(j)}, y_s^{(j)}) \quad (6)$$

$$\alpha_{j,\text{context}_{\gamma,k}} = \frac{1}{2^{2j} \gamma_{j,k}} \sum_{v_s^{(j)} = \text{context}_{\gamma}} p(x_s^{(j)} = k|v_s^{(j)}, y_s^{(j)}) \quad r \in \{1, \dots, 2K\} \quad (7)$$

d. If is received, go to step 5, otherwise go to step 1;

$$\text{Classify all the coefficients of scale J according to } \hat{x}_s^{(j)} = \arg \max_{k \in \{1, \dots, K\}} p(x_s^{(j)} = k|y_s^{(j)}, v_s^{(j)});$$

e. If , it ends, otherwise set and turn step a.

By performing the above operation with high-resolution remote sensing images, we can acquire disjoint and independent ground object sets such as buildings, roads, cement floors, vegetation, and water bodies, and on this basis, apply the next section of the K-neighbour classification method, the construction area is extracted and the impact of other non-building area features removed.

5. BUILDING AREA EXTRACTION BASED ON K NEIGHBORHOOD CLASSIFICATION METHOD

The remote sensing image segmentation method based on multi-scale Markov model provides a large number of rich and broken ground object sets. As mentioned above, due to the rich features of high-resolution remote sensing imagery and texture information, this generated base object set will contain a variety of complex base object units, in order to reduce the interference of the non-building area features as much as possible and further improve the extraction accuracy, a complete construction area must be extracted based on the above segmented results.

In this paper, the K-nearest neighbor classification method (Wang Lei et al. 2001) is adopted. At the same time, the texture, shape, and spatial features of the image are introduced into the feature space of the nearest neighbor supervised classification, by calculating the Euclidean distance between the ground object target and the adjacent ground object in the n-dimensional space, the same kind of merge is performed to obtain the complete building area image.

The main steps are:

a. Calculate the Euclidean distance. Calculate the distance between it and each object in the training set based on the object in the building area;

- b. Find neighboring pixels. Retrieve the nearest k training objects as neighboring pixels of the building area object;
- c. Classification. Based on the main categories to which the k neighboring cells belong, the building area objects are classified.

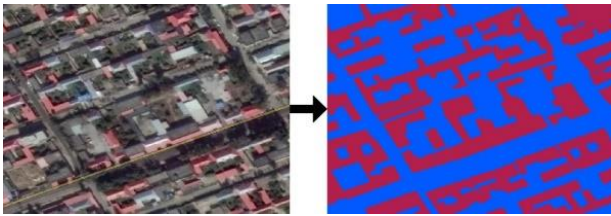


Figure 4. Building area extraction based on K adjacent classification method

As shown in Figure 4, the left picture shows the original high resolution remote sensing image, and the right picture shows the results of the extracted building area.

The high-resolution remote sensing image can be used to perform the above operations, and the building area can be completely extracted, for the fine extraction of buildings, it provides the environment of the construction area where the judgment conditions are relatively tolerant and the discriminating factors are relatively single.

6. BUILDING EXTRACTION BASED ON SPECTRAL FEATURE INDEX

The building should be inside the building area, therefore, the image of the building area extracted through the above steps can be taken as the building rough extraction result, and based on this, a more accurate and complete building can be extracted. However, in remote sensing images, shadows usually accompany buildings. Compared to low- and mid-resolution remote sensing images, buildings and architectural shadows are more clearly visible in high-resolution remote sensing images. At the same time, because urban man-made features are generally denser and overlap with vegetation, therefore, the fine extraction of buildings must remove information such as shadows, vegetation, and other pseudo-buildings to obtain pure buildings.

In order to make full use of the color characteristics of buildings in high-resolution remote sensing images and to form a greater degree of differentiation with other types of features, this paper proposes a method of feature space transformation of remote sensing image that takes into account the characteristics of buildings. It is used to describe the spectral features of buildings in high-resolution remote sensing images, calculate the spectral feature index, and accurately acquire buildings.

Buildings usually have high brightness in high-resolution remote sensing images, according to their characteristics, R , G , and B are the values of the image in the red, green, and blue bands. The feature transformation formula of the architectural feature dimension is:

$$f_1 = 2 \cdot \varphi_1 \cdot R + 2 \cdot (1 - \varphi_1) \cdot B - G \quad (8)$$

$$F_1 = \begin{cases} f_1 & \text{if } f_1 > 0 \\ 0 & \text{if } f_1 \leq 0 \end{cases} \quad (9)$$

In the formula, φ_1 is the feature weight, $\varphi_1 \in [0,1]$. When the image is dominated by reddish roofs, take $\varphi_1 > 0.5$; When the image is dominated by a bluish roof, take $\varphi_1 < 0.5$; When the color of the roof in the image is gray, take the default value $\varphi_1 = 0.5$.

Vegetation usually has a greenish hue and high saturation in high resolution remote sensing images. According to its characteristics, the characteristic transformation formula of the vegetation feature dimension is:

$$f_2 = \frac{G - \varphi_2 \cdot R - (1 - \varphi_2) \cdot B}{G + \varphi_2 \cdot R + (1 - \varphi_2) \cdot B} \quad (10)$$

$$F_2 = \begin{cases} f_2 & \text{if } f_2 > 0 \\ 0 & \text{if } f_2 \leq 0 \end{cases} \quad (11)$$

In the formula, φ_2 is the feature weight, $\varphi_2 \in [0,1]$. If the remote sensing image acquisition time is autumn, the hue of the vegetation in the image is shifted to the red band, the value is $\varphi_2 < 0.5$.

Shadows have lower brightness and higher hue in high resolution remote sensing images. According to its characteristics, the feature transformation formula of the shadow feature dimension is:

$$F_3 = \frac{H}{V} \quad (12)$$

In the formula, H and S are the hue and luminance components of the image in the HSV color space.

The transformed image feature space consists of three dimensions: normalized feature dimension, vegetation feature dimension, and shadow feature dimensions in this feature space, in order to better highlight the characteristics of buildings and simultaneously suppress vegetation and shadow characteristics, this paper proposes a building spectral feature index:

$$I_F = \frac{2F_1 - F_2 - F_3}{2F_1 + F_2 + F_3} \quad (13)$$

According to the above definition of the spectral characteristic index of buildings, the building extraction process of this paper mainly includes the following steps:

- a. Calculate the architectural feature dimension, vegetation feature dimension and shadow feature dimension index of each pixel area by traversing the HSV color space image of the building area, and further calculate the spectral characteristic index of the building;
- b. Set the empirical threshold m of the spectral characteristic index of a building, when the spectral characteristic index of a building in a certain area is less than m , the building is considered to be a building, and when it is greater than m , it is regarded as a pseudo-building and removed;
- c. The morphological closed calculations and open operations are performed on the extraction results, and the "glitches" in the extraction results are eliminated to obtain the final building;
- d. In the above steps, the empirical threshold m of the spectral characteristic index of the building is generally 0.8.

By building extraction of building based on spectral feature index, ultimately, accurate, precise and complete building information can be obtained.

7. EXPERIMENTAL RESULTS AND ANALYSIS

In order to further verify and analyze the building extraction effect of the proposed method on high-resolution remote sensing images, this paper uses the representative gray-level threshold extraction method, pixel-based classification extraction method and the proposed method for comparison. Two kinds of remote sensing images of multi-spectral bands and panchromatic bands in the same urban building area were selected as experimental data and tested. The test contrast effect is shown in the figure below.

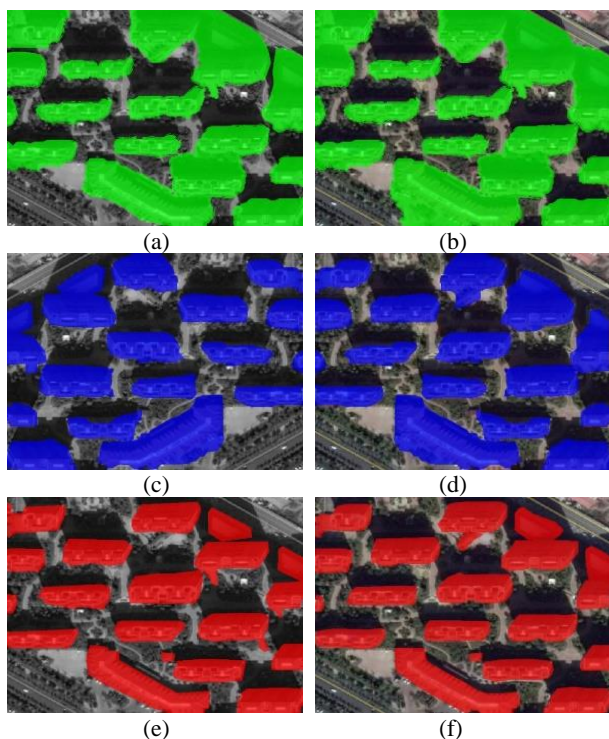


Figure 5. The Comparison of Building Extraction in The Same Urban Building Areas at Multispectral and Panchromatic Bands

Figure 5(a) and Figure 5(b) are building extraction results using a gray-level threshold-based extraction method from multi-spectral band image and full-color image. Figure 5(c) and Figure 5(d) are results of extracting a building using an extraction method based on pixel classification from multi-spectral band image and full-color image. Figure 5(e) and Figure 5(f) are the results of building extraction using the method proposed in this paper from multi-spectral band image and full-color image. After comparison, the proposed method can effectively eliminate roads, cement floors and other non-building areas. It can also remove information such as shadows, vegetation and other pseudo-buildings, in order to extract accuracy and completeness, the error score is lower, and the performance is better. In addition, due to the extraction method, the method also takes full account of the spectral features of the urban features, so the multi-spectral band images have better extraction effects.

After extraction experiment comparison and comprehensive statistics, the accuracy evaluation results of the proposed algorithm, the gray threshold segmentation method, and the pixel-based classification method are shown in Table 1.

Classification method	Overall classification accuracy (%)	Kappa coefficient	User accuracy (%)
The algorithm proposed in this paper	91.4436	0.8724	93.6235
Gray-level threshold extraction method	74.3211	0.7015	77.4937
Pixel-based classification method	77.5476	0.7381	81.2794

Table 1. The comparison of the overall classification accuracy of various classification methods

Among them, the overall classification accuracy is the sum of the correctly classified pixels divided by the total number of pixels. The Kappa coefficient is a commonly used measure of the classification accuracy based on the confusion matrix and the user accuracy is the ratio of the total number of pixels that are correctly assigned to the target class to the total number of pixels in the target class that the entire image pixel is divided into.

From Table 1, we can see that the gray-level threshold extraction method and pixel-based classification method are significantly lower than the proposed algorithm in the overall classification accuracy, Kappa coefficient and user accuracy. The proposed algorithm has obvious advantages in building extraction accuracy in high-resolution remote sensing images.

8. CONCLUSION

This paper proposes a building extraction method based on Markov model, and compares this method with the extraction method based on gray thresholds and the two commonly used extraction methods based on pixel classification. Separately targeted at two buildings of different high-resolution remote sensing images with multi-spectral bands and panchromatic bands in the same urban building area were tested for building extraction, and satisfactory results were obtained. Through comparative analysis and data statistics, it is proved that the proposed algorithm is superior to the above two extraction methods in the accuracy of building extraction.

In addition, experiments have shown that since the proposed algorithm fully considers the multi-directionality and anisotropy of high-resolution remote sensing imagery feature information, simultaneously with the construction of the Contourlet domain, the multi-scale segmentation and classification of remote sensing images is realized using the Markov model, based on this, the spectral characteristics of buildings, shadows, and a small amount of vegetation in the building area are fully considered and the fine extraction of buildings is finally achieved, so this method has the following features in addition to its high extraction accuracy:

- The integrity of building extraction is high, and effective and reliable building information can be obtained;
- The high accuracy of building extraction enables the effective removal of information about shadows, vegetation, and other pseudo-buildings that are adjacent to buildings, and allows for cleaner and standard building information.

REFERENCES

CUI Weiwei, TIAN Wei, ZHAO Wei. 2013, Adaptive cross-cutting nuclear model for image segmentation. *Computer Engineering and Applications*, 49(05): 190-194.

Do MN, Vetterli M. 2005, The Contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* . 14(2):2091-2106.

Fan Guo-liang, Xia Xiang-gen. 2001, A joint multi-context and multiscale approach to bayesian image segmentation. *IEEE Transactions on Geoscience and Remote Sensing* . 39(12):2680-2688.

Jiao Licheng, Sun Qiang. 2006, Perception and Recognition of Multi-scale Transform Domain Images: Progress and Prospects. *Chinese Journal of Computers*, 29(02): 177-193.

Li Hui, Jiang Chao. 2016, Image Adaptive Threshold Denoising Method Based on Contourlet Transform. *Computer and Digital Engineering*, 44(06): 1162-1166.

LI Xia, LUO Xin, XUE Hui, YANG Ting. 2014, Denoising of multispectral images using wavelet-domain Hidden Markov Tree Model. *Journal of Surveying and Mapping Science and Technology*, 31(03): 274-277+ 282.

Lijun Wang, Ming Dong. Multi-level Low-rank Approximation-based Spectral Clustering for image segmentation. *Pattern Recognition Letters*, 2012, 33(16).

LIU Li, ZHAI Gangyao. 2009, Overview of image texture feature extraction methods. *Journal of Image and Graphics*, 14(4): 622-635.

Pedro F. Felzenszwalb, 2004, Daniel P. Huttenlocher. Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision*, Vol.59 (2), pp.167-181.

Tang Wenjing, Xu Zhaoxin. 2014, Image Segmentation Algorithm Based on Boundary Markov Model in Wavelet Domain . *Sensors and Microsystems*, 33(09):145-147+154.

Wang Lei, Huang Chenxue. 2016, Improved hierarchical Markov random field color image segmentation algorithm. *Journal of Computer Applications*, 36(09): 2576-2579.

Xu Miaozhong, Cong Ming, Wan Lijuan, Jie Tianpeng, Zhu Xiaoling. 2015, High-resolution remote sensing image segmentation combined with Markov random field and visual perception. *Journal of Geomatics* 44(02): 198-205+ 213.