

Research Article

Spectral Residual Model for Rural Residential Region Extraction from GF-1 Satellite Images

Shaodan Li,^{1,2} Hong Tang,^{1,2} and Xin Yang²

¹Key Laboratory of Environment Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, China

²State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

Correspondence should be addressed to Hong Tang; hongtang@bnu.edu.cn

Received 22 December 2015; Revised 26 February 2016; Accepted 3 March 2016

Academic Editor: Fernando Torres

Copyright © 2016 Shaodan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Visual attention is an attractive technique to derive important and prominent information from a scene in natural pictures. As a visual attention approach, spectral residual (SR) model is adapted to extract the residential regions from GF-1 satellite images in this paper. Specifically, we analyzed the impact of both different combinations of GF-1 satellite image bands and threshold algorithms on rural residential region detection. In addition, the adapted approach is compared with related visual attention methods in terms of both quantitative and qualitative detection effectiveness. Experimental results showed that the SR model coupled with *red*, *green*, and *blue* bands in GF-1 images and Otsu threshold algorithm achieved the best results and is suitable to quickly extract rural residential regions from GF-1 images.

1. Introduction

The detection of residential regions has very important significance in urban development, map updating, and disaster management. In recent years, a wide range of automatic residential region detection techniques have been reported [1–7]. They could fall in the following three categories: index-based methods, image classification, and visual attention models.

In analogy to the widely known “vegetation index,” index-based methods aim to indicate residential regions by constructing an index of a combination of certain satellite image bands. Zha et al. (2003) presented a method based on Normalized Different Built-Up Index (NDBI), which was successfully applied to extract the urban residential regions using Landsat TM imagery [1]. The limitation of index-based method is that formula of the index must change with different satellite imagery. For example, a new index (PanTex) is proposed by Pesaresi et al. [2] when very high resolution (VHR) satellite images are used to extract residential regions. The index is based on fuzzy rule-based

composition of anisotropic textural measures derived from the remote sensing image by the gray-level cooccurrence matrix (GLCM).

Image classification approaches are often used to extract residential regions [3–7]. The methods for image classification can be roughly divided into two categories. The first category is supervised classification, which needs a set of specific training samples to learn features of the residential regions for detection. For examples, Benediktsson et al. use differential morphological profiles to extract structural features and neural network to classify the features [3]. The second category of such methods directly detects residential regions without using any training data, which include the approaches using local feature points [4–6] and the method using edge density feature [7].

Visual attention is a technique to derive important and prominent information from a scene in natural pictures. This type of method has the potential to discriminate the residential regions from the background in satellite images. From the view of the generation of saliency maps, these visual attention models could fall in two categories: decomposition model

and integration model. (1) The decomposition model consists of two independent steps, that is, feature extraction and saliency generation. For examples, Itti et al. [8] constructed the earliest visual attention model by using a biologically plausible architecture, which was proposed by Koch and Ullman [9]. In this model, feature maps are first derived by using center-surround differences from nine various spatial scales of Gaussian pyramids. Then intensity, color, and orientation conspicuity maps are obtained by the across-scale addition of the feature maps. The final saliency map is generated by summing and normalizing three conspicuity maps. However, the saliency map is computed by “across-scale operator” so that some information is lost. As an improvement of the Itti model, Harel et al. proposed a graph-based visual saliency (GBVS) [10]. The GBVS model obtains feature maps using the Itti model but improves their normalization by introducing the graph theory, that is, the Markov chain. The saliency map also has low resolution, and some spatial information is lost. (2) The integration model could directly generate the saliency map from the input image without the intermediate step. For instance, Achantay et al. presented the frequency-tuned (FT) saliency detection [11]. In FT model, the input image is first transformed to the *Lab* color space from the *RGB* color space, and then the saliency map is obtained by computing the Euclidean distance between the Gaussian-blurred image and the arithmetic mean image. The FT method could generate full-resolution saliency maps. Hou and Zhang attempt to obtain a saliency map in the transformed domain [12]. The model first analyzes the log spectrum of the images and extracts the spectral residuals (SR) in the spectral domain. Then the saliency map is created by transforming the spectral residuals to the spatial domain. Generally, visual attention models analyze the property of target objects, but the common properties of various types of objects can not exist. The SR model avoids this problem by exploring the properties of the backgrounds. The main advantage of the SR model is its generality. Based on this point, the SR model is focused on to extract the residential regions using GF-1 satellite images in this paper, and more details are described in next section. In this paper, the SR model is adapted to extract the rural residential regions using GF-1 satellite imagery by analyzing the impact of different band combinations and threshold algorithms.

The reminder of this paper is organized as follows. In Section 2, the SR model for extracting the residential regions using GF-1 satellite imagery is presented, including band combinations and threshold selection. In Section 3, the experiments of the SR model with different band combinations and different threshold algorithms are given, and related visual attention models are also compared in Section 3. Finally, the conclusion is drawn in Section 4.

2. The SR Model for Residential Region Extraction

In this section, the SR model is introduced to extract rural residential regions using GF-1 satellite imagery. As shown in Figure 1, the flow of the residential region extraction consists

of two main steps: (1) the SR model is used to obtain the saliency map of residential regions from the original GF-1 satellite imagery, and (2) the final residential regions are extracted from the corresponding saliency map by using the Otsu threshold algorithm.

2.1. The SR Model. By comparing the log spectra of a large number of natural images, Hou et al. found that the log spectra of different images share similar trends. Therefore, in different log spectra where considerable shape similarities can be observed, what deserves the attention is the information that jumps out of the smooth curves. It is believed that the statistical singularities in the spectrum may be responsible for anomalous regions in the image, where protoobjects are popped up. According to this observation, the SR model is proposed to derive a saliency map from transformed spectral residuals. Specifically, the framework of the SR model is shown in Figure 2. Given an input image $I(x)$, the saliency map is computed as follows.

Step 1. A gray image is computed by transforming the red (*R*), green (*G*), and blue (*B*) three-color channels of the input image $I(x)$ using

$$G(x) = 0.229 \times R + 0.587 \times G + 0.114 \times B. \quad (1)$$

Step 2. A frequency image f is obtained by $f = F(G(x))$, and the amplitude spectrum $A(f)$ and phase spectrum $P(f)$ are, respectively, computed by

$$\begin{aligned} A(f) &= \text{abs}(f), \\ P(f) &= \text{angle}(f), \end{aligned} \quad (2)$$

where F denotes the Fourier Transform.

Step 3. The spectral residual $R(f)$ of the input image is defined as

$$R(f) = L(f) - h(f) \times L(f), \quad (3)$$

where the log amplitude spectrum $L(f)$ is given by $L(f) = \log(A(f))$; $h(f) \times L(f)$ is the average log amplitude spectrum, which denotes the general shape of log spectra; and $h(f)$ is an 3×3 local average filter defined by

$$h(f) = \frac{1}{3 \times 3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (4)$$

Correspondingly, $R(f)$ denotes the residential regions in the input image.

Step 4. The final saliency map in spatial domain is constructed by Inverse Fourier Transform

$$S(x) = F^{-1} [\exp(R(f) + P(f))]^2, \quad (5)$$

where F^{-1} denote the Inverse Fourier Transform.

TABLE 1: The parameter of GF-1 satellite sensor.

Parameter	Index of band	Band name	Spectral range (μm)	Resolution (m)
Panchromatic band	—	PAN	0.45~0.90	2
Multispectral bands	B1	Blue	0.45~0.52	8
	B2	Green	0.52~0.59	8
	B3	Red	0.63~0.69	8
	B4	NIR	0.63~0.89	8

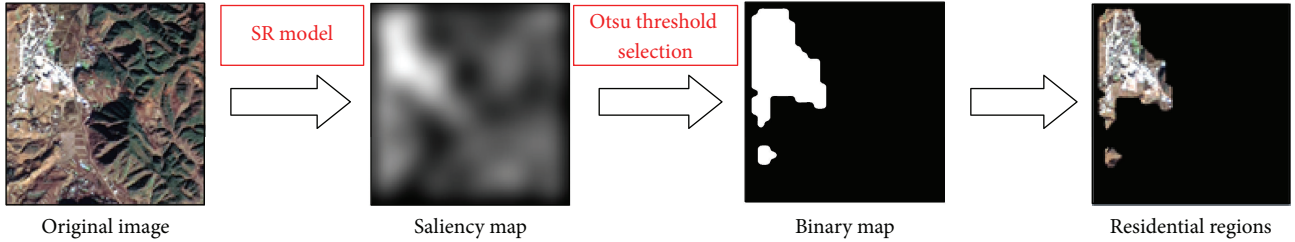


FIGURE 1: The flowchart of rural residential region extraction.

After the saliency map, the Otsu threshold algorithm is used to obtain a binary image and the final residential regions are obtained by multiplying the binary image with the input image.

2.2. Band Combinations of GF-1 Images. The input of the original SR model is red, green, and blue three-color channels in natural images, while the GF-1 satellite imagery has four multispectral bands and a panchromatic band. The parameter of GF-1 satellite sensor is shown in Table 1.

In order to find the optimal band combinations to produce the best result, several band combinations of GF-1 satellite images, including PAN_SR, MS_SR, RGB_SR, MS_SRs, and RGB_SRs, are considered in this paper. The flows of different band combinations for the SR model are shown in Figure 3.

- PAN_SR: in Step 1 of the program given in Section 2.1, the panchromatic band of GF-1 satellite image is used as the gray image, replacing the gray image in original SR model.
- MS_SR: in Step 1, the mean of four multispectral bands of GF-1 satellite image is used as the gray image to replace the original gray image.
- RGB_SR: in Step 1, the mean of red, green, and blue three bands of GF-1 satellite image is used as the gray image.
- MS_SRs: in Step 1, four multispectral bands of GF-1 satellite image are used as four gray images, respectively. Thus, four corresponding saliency maps, which are red saliency map, green saliency map, blue saliency map, and near infrared saliency map, are obtained. The final saliency map is computed by the mean of four saliency maps.

- RGB_SRs: this combination is similar to MS_SRs combination. The red, green, and blue bands of GF-1 satellite image are used as three gray images, respectively. Thus, three corresponding saliency maps are obtained. The final saliency map is computed by the mean of three saliency maps. In order to compare the differences of these band combinations, the corresponding experiments will be conducted in next section.

2.3. Threshold Selection for the SR Model. For the original SR model, the Otsu threshold algorithm is used to obtain a binary image from the saliency map. In this subsection, different threshold algorithms are compared to extract the residential regions, which are iterative threshold algorithm [13], Otsu threshold algorithm [14], KSW threshold algorithm [15], and moment-preserving threshold algorithm [16].

The iterative threshold algorithm is based on the approximation theory. The basic flow of the iterative threshold algorithm is described as follows. Firstly, an approximate threshold is randomly selected as an initial threshold, according to the distribution of the gray levels in the image. Thus the image is separated into two parts by the initial threshold. Secondly, we need calculate the means of gray levels of the two parts. u_1 is the mean of gray levels of the first part, and u_2 is the mean of the second part, respectively. Furthermore, the threshold is updated as $(u_1 + u_2)/2$. Finally, the optimal threshold is selected through multiple iterations, and the iterations would stop until u_1 and u_2 no longer change.

The Otsu threshold algorithm is proposed from the viewpoint of discriminant analysis, and the optimal threshold is automatically selected by the discriminant criterion. The basic idea of the algorithm is given as follows. The algorithm

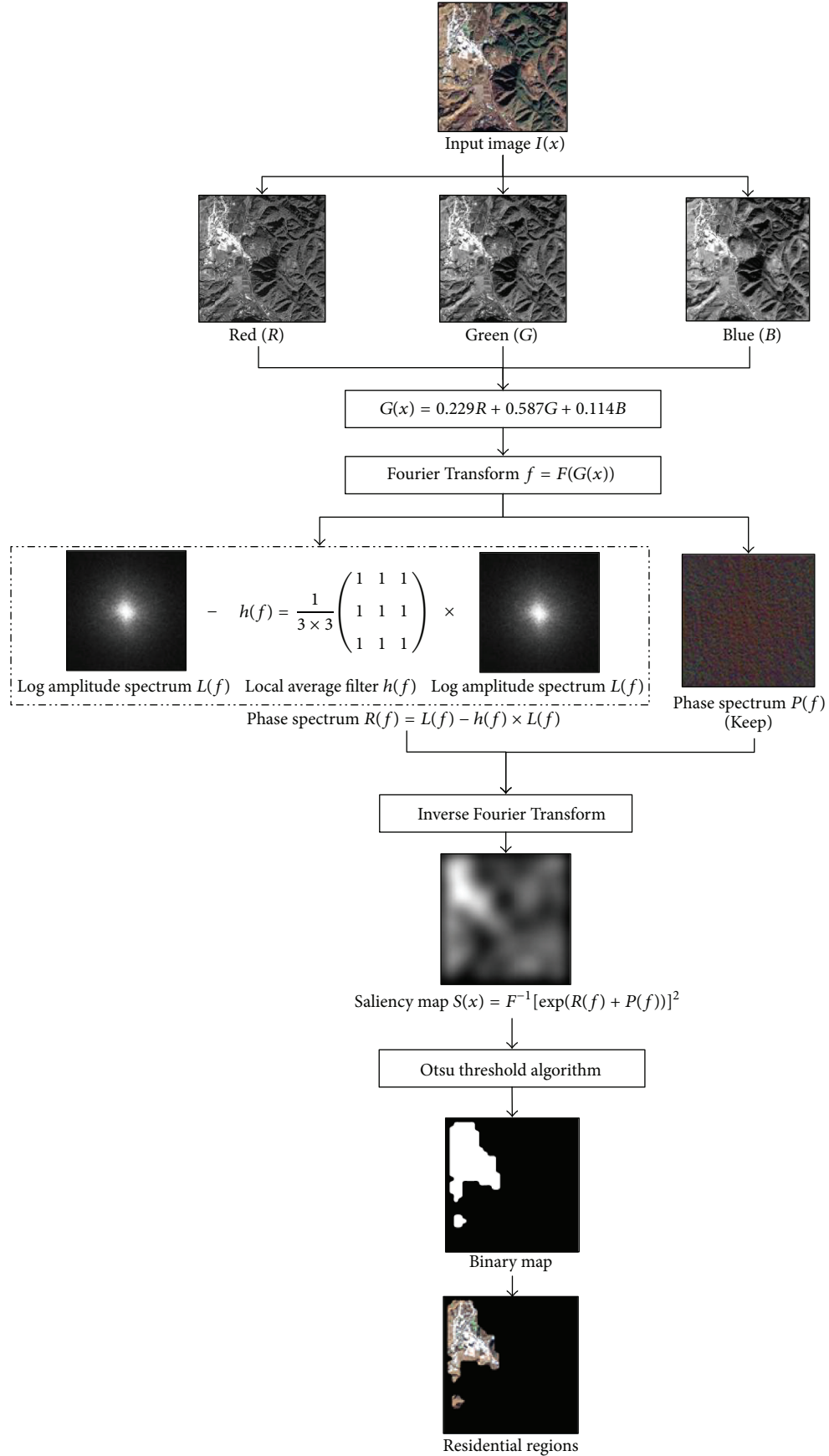


FIGURE 2: The framework of the SR model for residential region extraction.

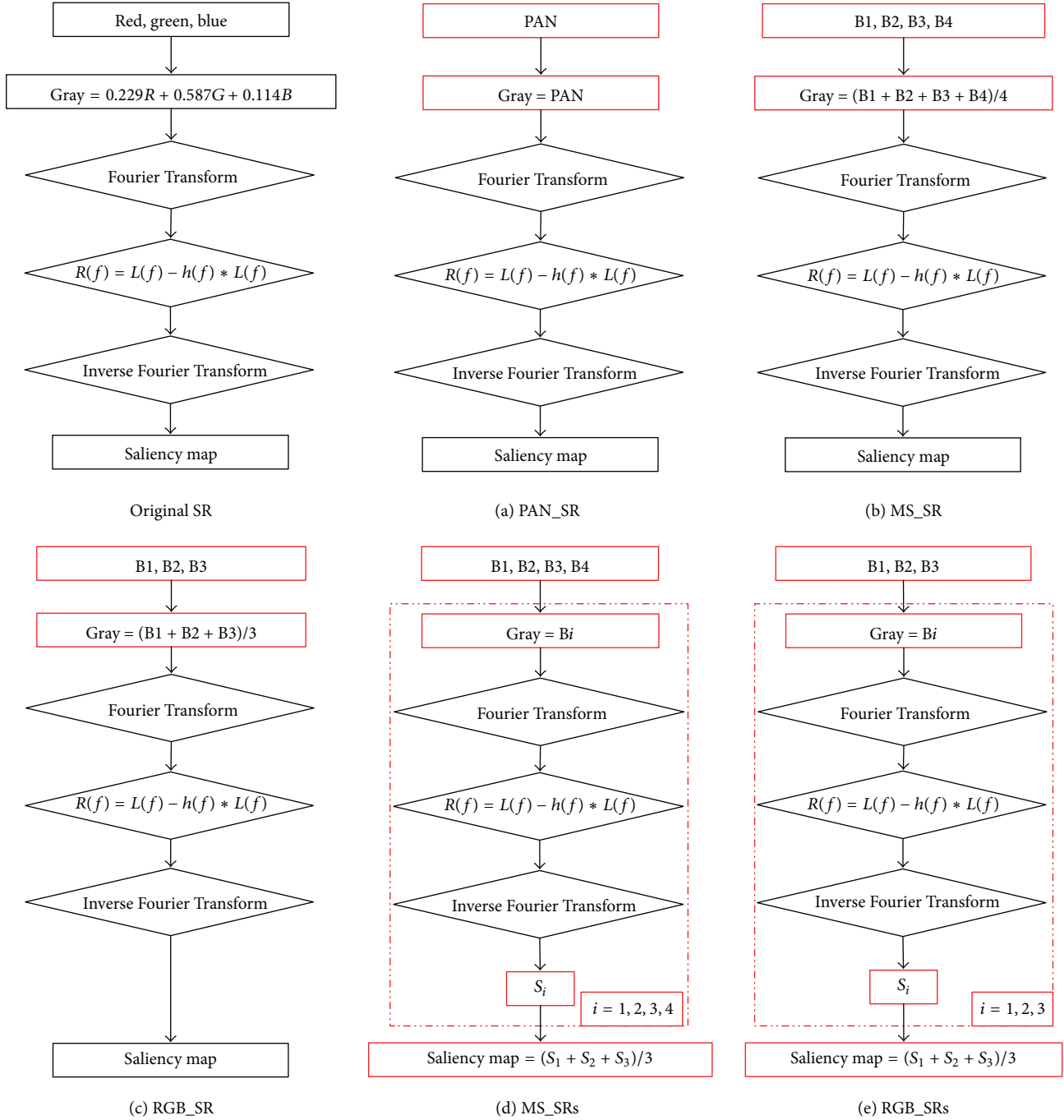


FIGURE 3: Flows of different band combinations for the SR model.

assumes that the image contains two classes of pixels following bimodal histogram, that is, object and background. Then it calculates the optimum threshold to separate the two classes so that their intraclass variance is minimal and their interclass variance is maximal. Since it is fast and effective, the Otsu threshold algorithm is widely used in image processing.

The KSW threshold algorithm for choosing a threshold is based on the entropy concept in information theory. The basic flow of the KSW threshold algorithm is described as follows. Firstly, the pixels of the image are divided into two parts (i.e., object and background) by an initial threshold. Then

the entropies of the object and background are computed. When the object is separated from the background in the optimal way, the sum of the entropies of the two parts should be maximal. Based on this theory, the optimal threshold is selected by maximizing the sum of the entropies of the two parts. The KSW threshold algorithm uses a global and objective property of the gray-level histogram of the input image.

The moment-preserving threshold algorithm is based on the moment-preserving principle. The basic idea of the algorithm is given as follows. The gray-level moments of the input

image are first computed, and then the threshold is obtained deterministically in such a way that the moments of the input image are preserved in the output image. This algorithm may be regarded as a moment-preserving image transformation which recovers an ideal image from a blurred version. The moment-preserving threshold algorithm can automatically and deterministically select the optimal threshold without iteration or search.

3. Experimental Results and Discussion

In this section, the experimental data, related visual attention models, and evaluation method are first described in Section 3.1. Then three experiments are conducted to evaluate the performance of the SR model using GF-1 satellite imagery to extract residential regions. In the first experiment, different band combinations for the SR model are compared to extract residential regions. In the second experiment, different threshold algorithms for the SR model are compared. In the last experiment, the SR model is compared with other visual attention models to extract residential regions.

3.1. Experimental Setting

3.1.1. Experimental Data. The experimental data sets used in this paper are taken from the GF-1 satellite imagery. The GF-1 satellite imagery has four multispectral bands (i.e., red, green, blue, and near infrared) and a panchromatic band. Specifically, the panchromatic image offers a resolution of 2 m, and multispectral image provides a resolution of 8 m. Since the resolution is inconsistent between panchromatic band and multispectral bands, the multispectral bands should be preprocessed. In order to improve the resolution of multispectral bands, the panchromatic and multispectral bands are fused by using Gram-Schmidt transformation in ENVI. Thus, the resolution of fused multispectral bands is updated to 2 m.

A total of five GF-1 satellite images are used to test the performance of the proposed method. For the number of the residential regions, there are one, two, and more than three residential regions in these test images. In terms of surrounding environment of the residential regions, the residential regions in Test-1 image are surrounded by bare soils and vegetation. The residential regions in Test-2, Test-3, and Test-4 images are surrounded by vegetation with different color and texture. The residential regions in Test-2 and Test-5 images locate near the water. For the pattern of the residential regions, the residential region in Test-4 image is concentrated, and the residential regions in Test-2 and Test-3 images are scattered. In addition, these test images are representative in our study areas, Western of China.

3.1.2. Related Visual Attention Models. Related visual attention models are compared with the SR model to extract residential regions using GF-1 satellite images in our experiment, which include Itti model, GBVS model, and FT model. These models used in our experiment are different from their

original models, because the input image is GF-1 satellite images rather than natural pictures.

In Itti model, the input is GF-1 satellite images with three multispectral bands (i.e., B1, B2, and B3) and a panchromatic band. In Itti model, three types of features are extracted from the input image, including the intensity, color, and orientation. For the intensity, the intensity map in our experiment is obtained by the panchromatic band of GF-1 satellite images rather than the mean of red, green, and blue used in original Itti model. For the color and orientation, the way of obtaining is the same as the original Itti model. Then the final saliency map is generated by summing and normalizing the intensity, color, and orientation maps.

In GBVS model, the input is also GF-1 satellite images with three multispectral bands (i.e., B1, B2, and B3) and a panchromatic band. Since the GBVS model obtains feature maps using the Itti model, the way of the feature maps obtained in our experiment is the same as the Itti model described above. Then activation maps are derived from the feature maps. An activation map is the stationary distribution of a Markov chain. Finally, the saliency map is generated by normalizing the activation maps.

In FT model, the input is GF-1 satellite images with three multispectral bands (i.e., B1, B2, and B3) in our experiment. The input image is first transformed to the *Lab* color space from three multispectral bands. Then the saliency map is obtained by computing the Euclidean distance between the Gaussian-blurred image and the arithmetic mean image.

3.1.3. Evaluation Methods. In our experiments, the overall precision (P), recall (R), and F -measure (F) [17] are used to evaluate the performance of the SR model using GF-1 satellite imagery to extract residential regions. P , R , and F -measure are defined as

$$\begin{aligned} P &= \frac{TP}{TP + FP}, \\ R &= \frac{TP}{TP + FN}, \\ F &= \frac{(1 + \beta^2) P \times R}{\beta^2 \times P + R}, \end{aligned} \quad (6)$$

where TP is correctly detected pixels by using models among the ground truth. FP is the pixels detected using models but not in the ground truth, and FN is the pixels which are not detected using models but in the ground truth. β^2 is a positive parameter for weighting the precision and recall (β^2 is chosen as 2 in this paper). F -measure is the harmonic mean of the precision and recall.

3.2. Different Band Combinations for the SR Model. In order to find the optimal band combination to extract the residential regions using GF-1 satellite imagery, several band combinations of GF-1 satellite imagery, including PAN_SR, MS_SR, RGB_SR, MS_SRs, and RGB_SRs, are compared in this experiment. Moreover, the performance of different band combinations for the SR model is evaluated in terms of both qualitative and quantitative aspects.

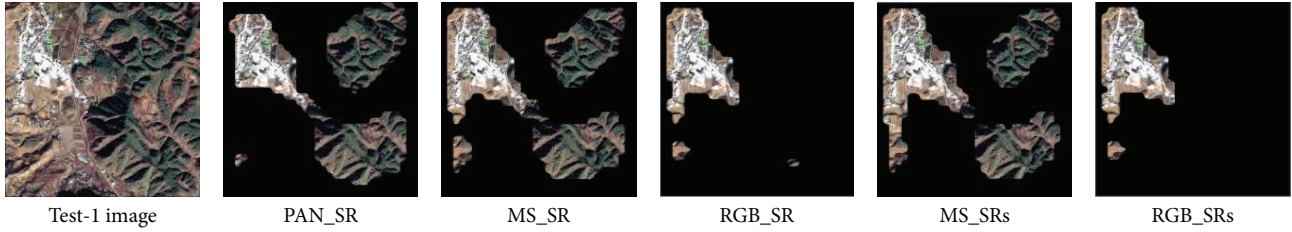


FIGURE 4: Comparison of different band combinations for SR model with Test-1 image.

3.2.1. Qualitative Evaluation. The comparisons of different band combinations with five test images are shown in Figures 4–8, respectively. As shown in Figures 4–8, the combinations RGB_SR and RGB_SRs provide a more accurate description of the residential regions than other combinations, compared with the ground truth (see Figures 16–20). Since both bare soils and residential regions have higher DN values in near infrared band (i.e., B4), some bare soils are wrongly extracted in the combinations MS_SR and MS_SRs which include near infrared band. From the visual perspective, the combinations MS_SR and MS_SRs extract more wrong regions than other combinations, and the combination PAN_SR has the worst performance due to the lack of color information.

3.2.2. Quantitative Evaluation. In order to evaluate the performance of different band combinations for SR model quantitatively, the precision, recall, and F -measure are used in this experiment. The quantitative comparison of different band combinations for SR model is shown in Figure 9.

As shown in Figure 9, for precision, the combinations RGB_SR and RGB_SRs have higher value than other combinations (see the first column in Figure 9). For recall, most of combinations reach about 90% (see the second column in Figure 9), because most of combinations extract more residential regions than the ground truth. Thus, the only recall can not indicate the performance of each combination. The F -measure, which balances the precision and recall, can comprehensively reflect the performance of each combination. The F -measure value of combinations RGB_SR and RGB_SRs is much higher than the combinations MS_SR and MS_SRs, and their F -measure values reach about 80% (see the last column in Figure 9). It means that the combinations RGB_SR and RGB_SRs are the best choices to extract the rural residential regions by using the SR model. In both Sections 3.3 and 3.4, the combination RGB_SR is used as a contrast option.

3.3. Different Threshold Algorithms for the SR Model. In this experiment, different threshold selection algorithms for the SR model are compared to extract residential regions. Furthermore, the performance of different threshold algorithms is evaluated in terms of both qualitative and quantitative aspects.

3.3.1. Qualitative Evaluation. The comparisons of different threshold algorithms for the SR model using five test images are shown in Figures 10–14. The residential regions extracted by using the Otsu and iterative threshold algorithms are

similar, but the Otsu threshold algorithm is much closer to the ground truth. In fact, the Otsu threshold algorithm is a global threshold method which has been widely used in threshold segmentation. Both KSW and moment-preserving threshold algorithms extract more wrong regions than the ground truth. Next, more details of the analysis are given from a quantitative comparison.

3.3.2. Quantitative Evaluation. The precision, recall, and F -measure are also used in this experiment to evaluate the performance of different threshold algorithms for the SR model quantitatively. The comparison of different threshold algorithms for the SR model is shown in Figure 15.

According to the definition, a higher precision means that the algorithm detects most of regions from the saliency map. Thus, the less detected the residential regions, the higher the precision value, such as the precision of the Otsu threshold (see the blue bar in the first column in Figure 15). Similarly, a higher recall means that the algorithm detects most of the residential regions from the ground truth. Thus, the more detected the regions, the higher the recall value, such as the recall of the moment-preserving threshold (see the purple bar in the second column in Figure 15). However, the ideal algorithm is neither an algorithm with highest precision value nor an algorithm with highest recall value, but the balanced one between precision and recall. The F -measure is the harmonic mean of the precision and recall, and thus the ideal algorithm we desired is the one with the highest F -measure value in all threshold algorithms (see the blue bar in the last column in Figure 15). So the optimal threshold algorithm is the Otsu threshold algorithm in our experiment.

3.4. Different Visual Attention Models. In this experiment, different visual attention models for extracting the rural residential regions using GF-1 satellite images are compared. Furthermore, the performance of different visual attention models is evaluated from both qualitative and quantitative aspects.

3.4.1. Qualitative Evaluation. The saliency map comparison experiments for different visual attention models with five test images are shown in Figures 16–20, respectively. The first rows in Figures 16–20 are the corresponding saliency maps of different visual attention models, and the second rows are the extracted rural residential regions by using Otsu threshold algorithm (Otsu threshold algorithm has the best performance, as described in Section 3.3). The saliency

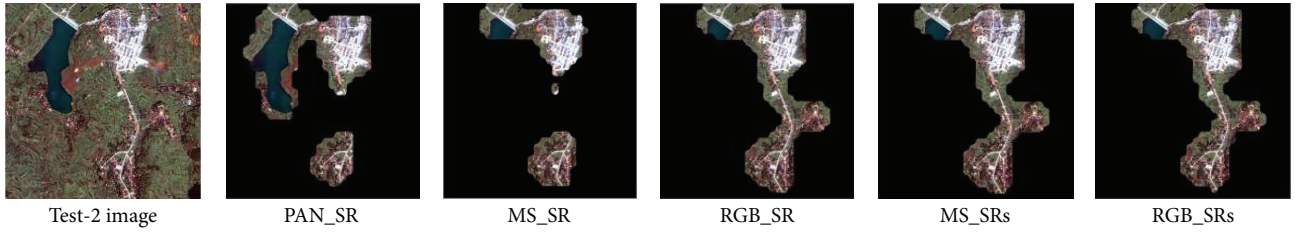


FIGURE 5: Comparison of different band combinations for SR model with Test-2 image.

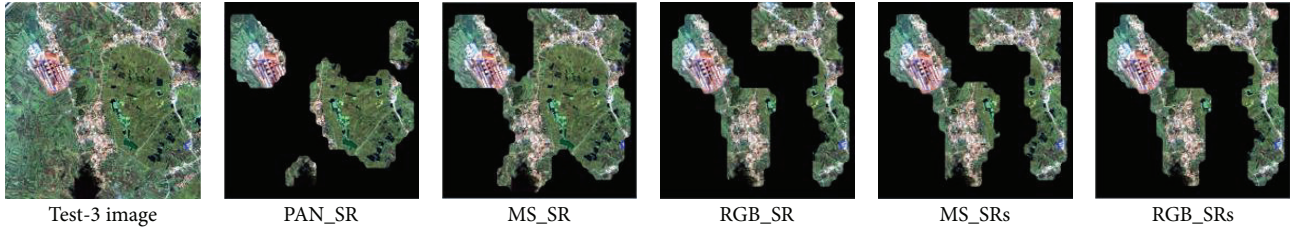


FIGURE 6: Comparison of different band combinations for SR model with Test-3 image.

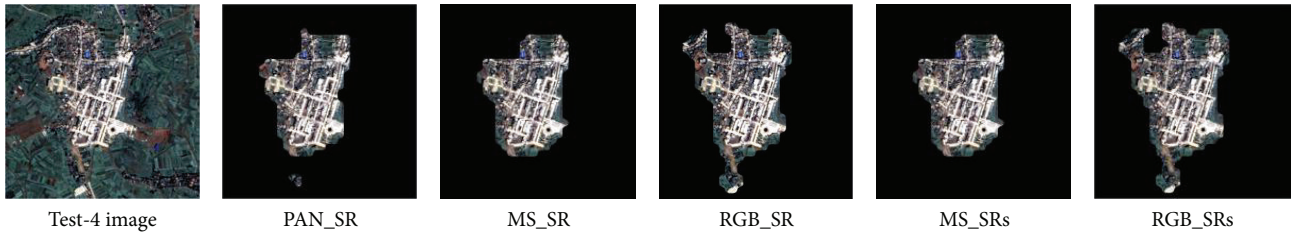


FIGURE 7: Comparison of different band combinations for SR model with Test-4 image.

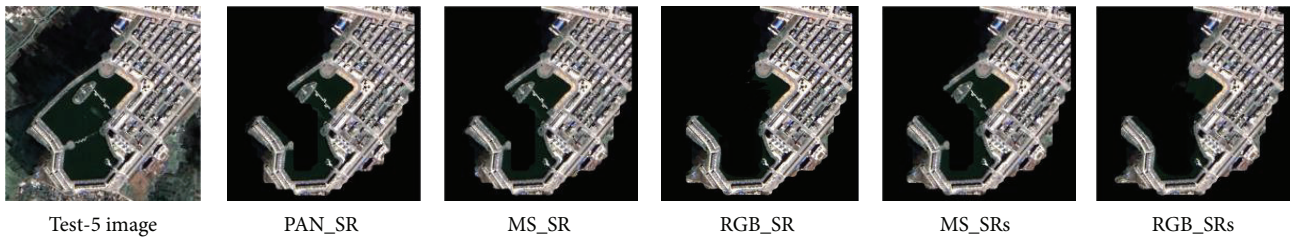


FIGURE 8: Comparison of different band combinations for SR model with Test-5 image.

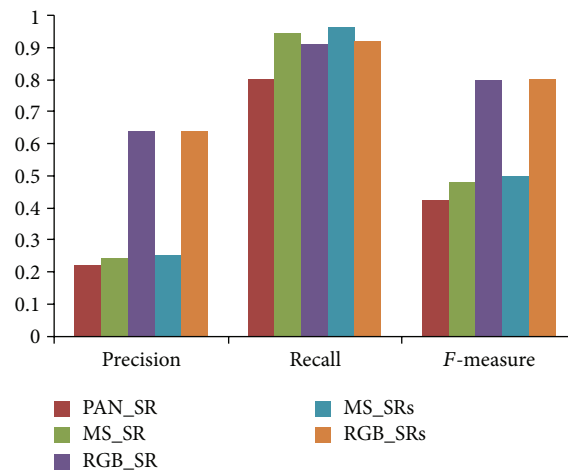


FIGURE 9: Quantitative evaluation of different band combinations for the SR model.

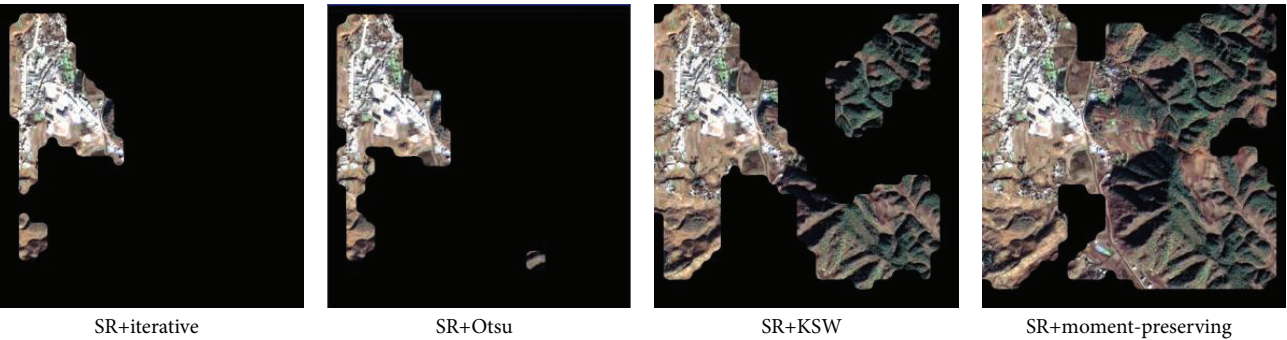


FIGURE 10: Comparison of different threshold algorithms for the SR model with Test-1 image.

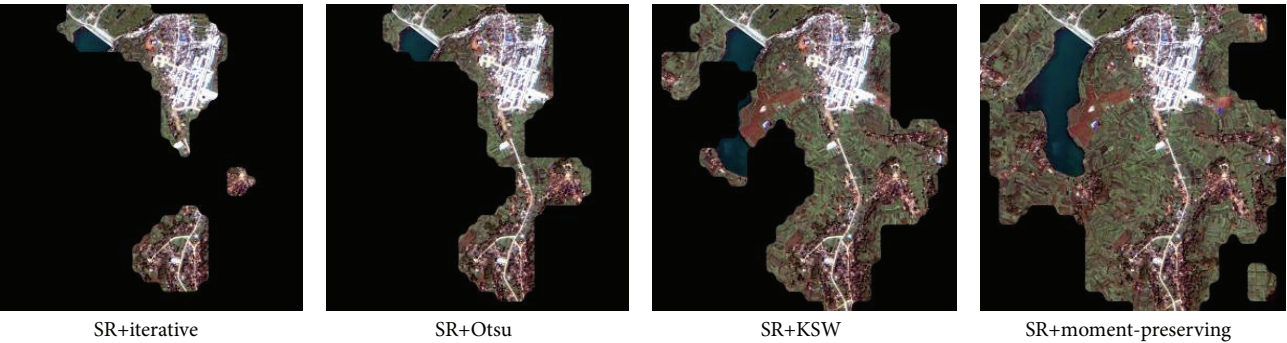


FIGURE 11: Comparison of different threshold algorithms for the SR model with Test-2 image.

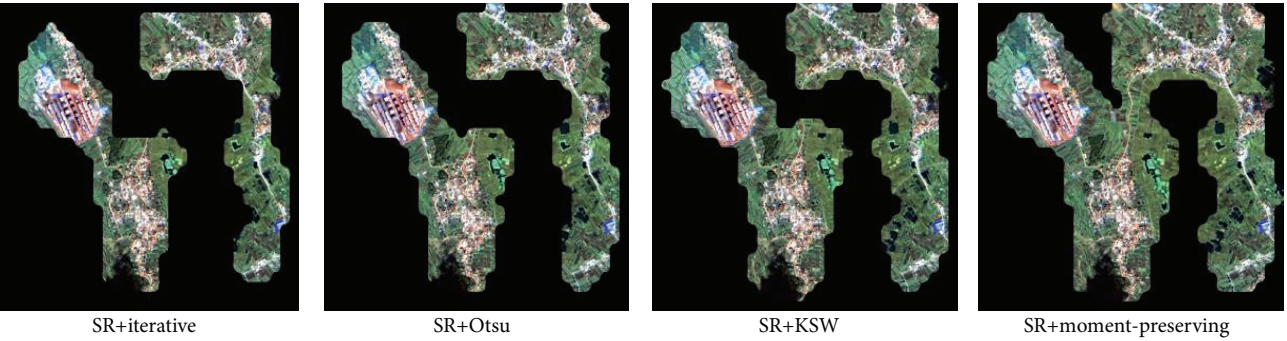


FIGURE 12: Comparison of different threshold algorithms for the SR model with Test-3 image.

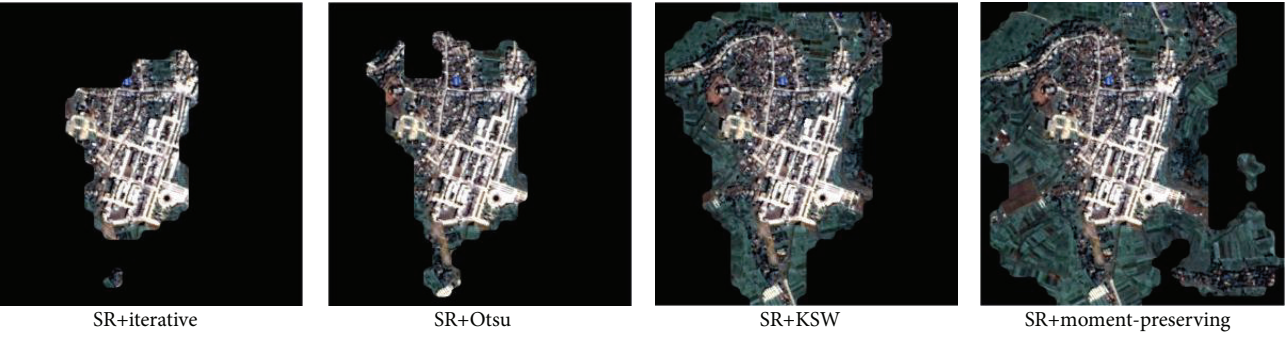


FIGURE 13: Comparison of different threshold algorithms for the SR model with Test-4 image.

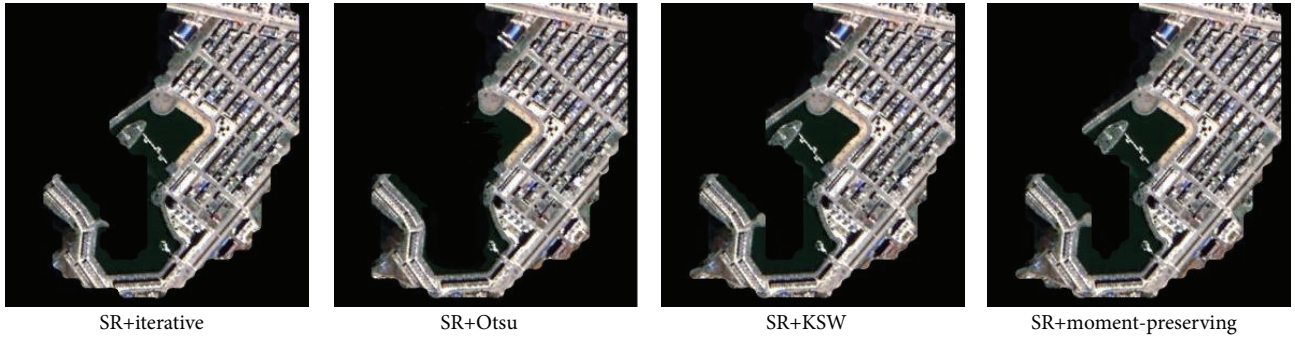


FIGURE 14: Comparison of different threshold algorithms for the SR model with Test-5 image.

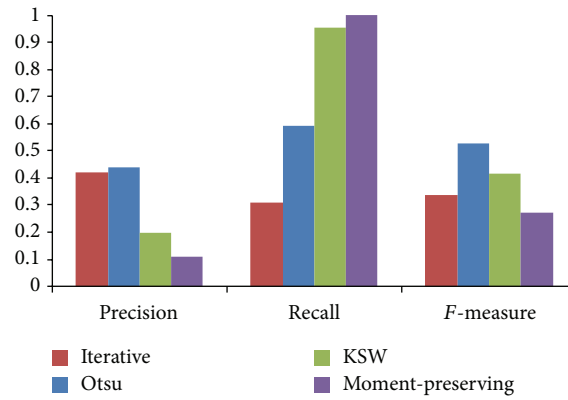


FIGURE 15: Quantitative evaluation of different threshold algorithms for the SR model.

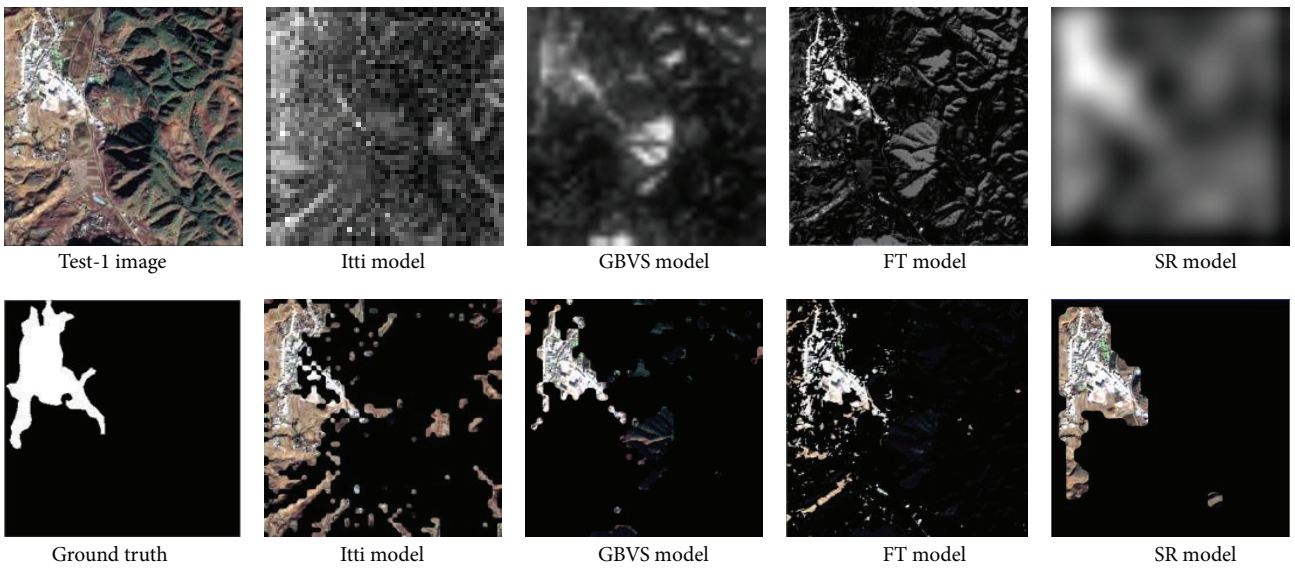


FIGURE 16: Comparison of different visual attention models using Test-1 image.

maps derived from Itti, GBVS, and SR models have lower resolution, and thus these saliency maps are blurred. The FT model provides the saliency map that has the same resolution as the input images, but it produces many fragmented regions from the background of the image. Experimental results

show that the residential regions extracted by using the SR model are much closer to the ground truth, and the SR model provides a more accurate description of the residential regions than other visual attention models. More details of the discussions are given in next subsection.

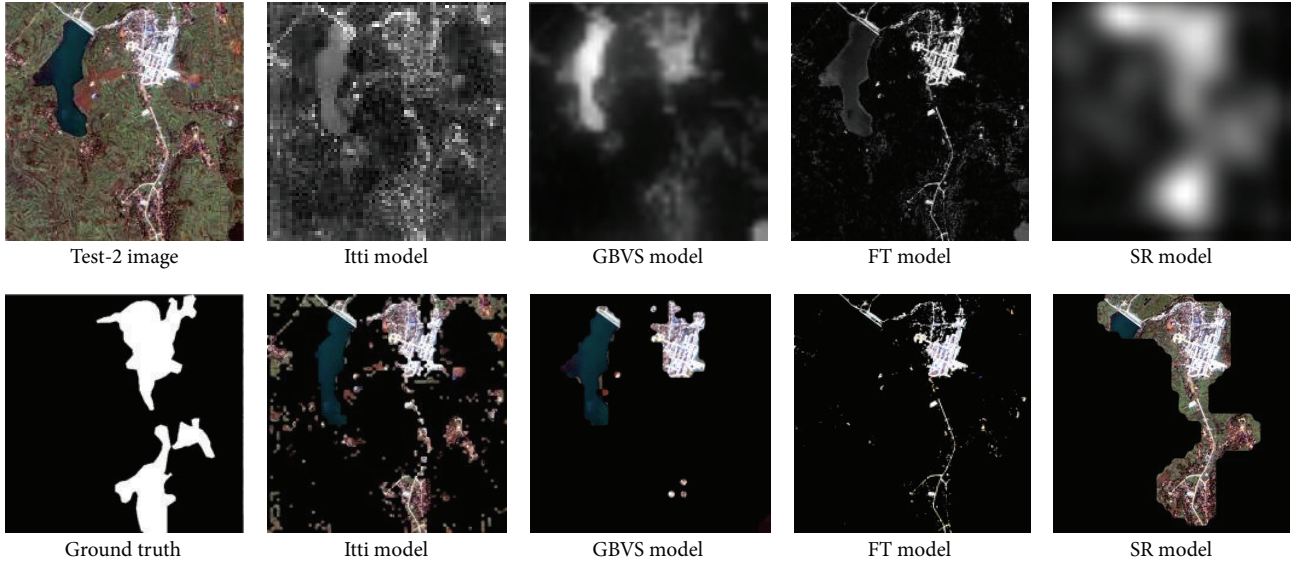


FIGURE 17: Comparison of different visual attention models using Test-2 image.

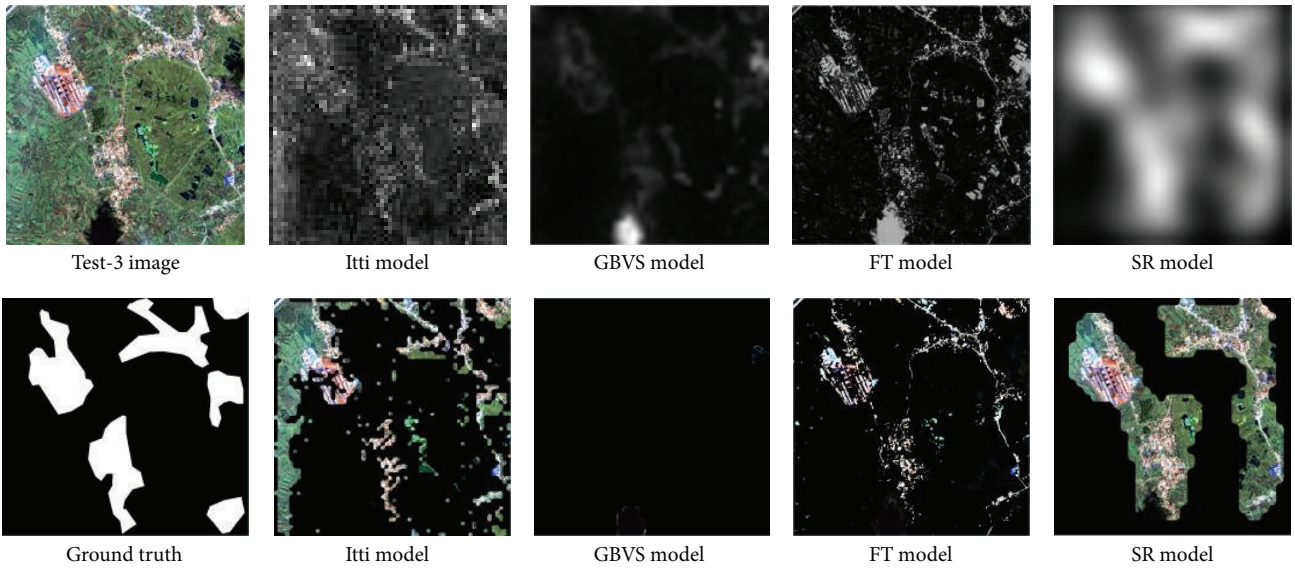


FIGURE 18: Comparison of different visual attention models using Test-3 image.

3.4.2. Quantitative Evaluation. In order to evaluate the performance of different visual attention models quantitatively, we computed the precision, recall, and F -measure for each model. The quantitative comparison of different visual attention models is shown in Figure 21.

As shown in Figure 21, the SR model has highest precision and recall value among other visual attention models (see the first and second column in Figure 21). The F -measure is the harmonic mean of the precision and recall, and thus the SR model has the highest F -measure value among others. The SR method is clearly outperformed in extracting the rural residential regions using GF-1 satellite images. The reason

why the SR model produced a more accurate result is that the SR model is based on the properties of the background in the satellite images, while other models focus on the property of target objects. As described in Section 3.1.1, the background of these test images in Southwest of China is anisotropic. Therefore, the SR model seems to be better at extracting rural residential regions from more complex background, such as the mountain area of Southwest of China. In fact, the rural residential regions extracted by using the SR model are satisfactory in most cases except when some cloud covered the residential regions. Furthermore, since both the residential regions and bare soils have higher

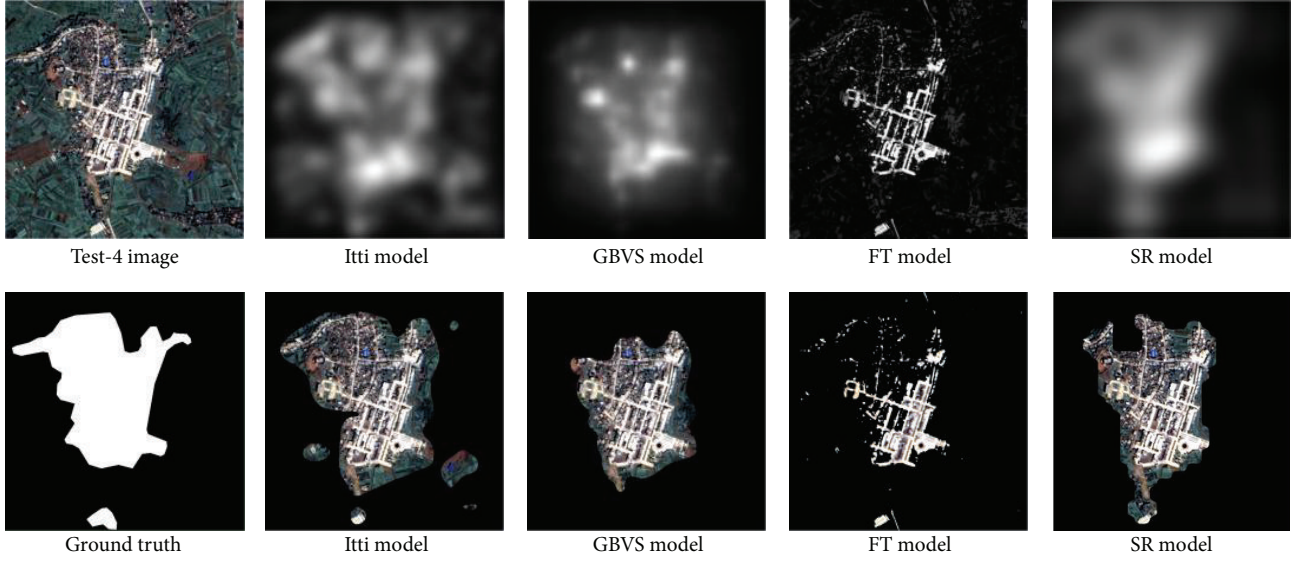


FIGURE 19: Comparison of different visual attention models using Test-4 image.

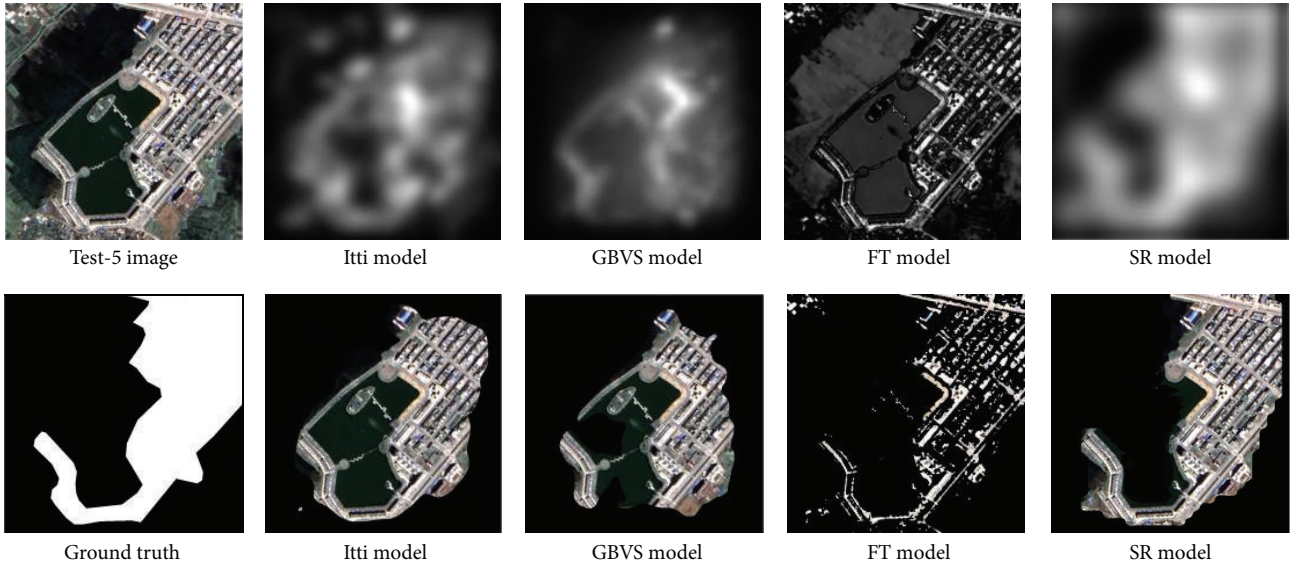


FIGURE 20: Comparison of different visual attention models using Test-5 image.

DN value in images, some background information, such as bare soils, is falsely extracted in other models based on the property of target objects.

4. Conclusions

In this paper, the SR model is adapted to the extract the rural residential regions using GF-1 satellite imagery. Specifically,

we analyzed the impact of both different combinations of GF-1 satellite image bands and threshold algorithms on rural residential region detection. In addition, the adapted approach is compared with other visual attention methods. Moreover, in order to quantitatively evaluate the performances, the precision, recall, and F -measure are presented in the experiments. Experimental results show that the SR model coupled with *red*, *green*, and *blue* bands in GF-1 images and Otsu threshold algorithm achieved the satisfactory results and is suitable to quickly extract rural residential regions from GF-1 images.

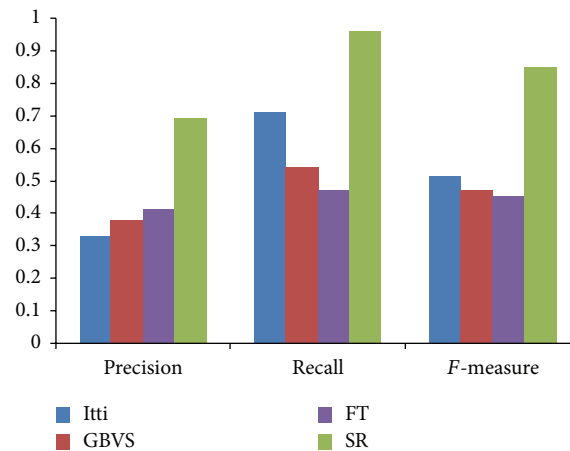


FIGURE 21: Quantitative evaluation of different visual attention models.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China (no. 41571334) and the Fundamental Research Funds for the Central Universities.

References

- [1] Y. Zha, S. Ni, and S. Yang, "An effective approach to automatically extract urban land-use from TM imagery," *Journal of Remote Sensing*, vol. 7, no. 1, pp. 37–40, 2003.
- [2] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 1, no. 3, pp. 180–192, 2008.
- [3] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 9, pp. 1940–1949, 2003.
- [4] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1156–1167, 2009.
- [5] B. Sirmacek and C. Unsalan, "Urban area detection using local feature points and spatial voting," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp. 146–150, 2010.
- [6] A. Kovács and T. Szirányi, "Improved harris feature point set for orientation-sensitive urban-area detection in aerial image," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 796–800, 2012.
- [7] H. Chen, C. Tao, Z.-R. Zou, and L. Shao, "Extraction of built-up areas extraction from high-resolution remote-sensing images using edge density features," *Journal of Applied Sciences*, vol. 32, no. 5, pp. 537–542, 2014.
- [8] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [9] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," in *Matters of Intelligence*, vol. 188 of *Synthese Library*, pp. 115–141, Springer, Dordrecht, The Netherlands, 1987.
- [10] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proceedings of the Advances in Neural Information Processing Systems (NIPS '07)*, vol. 19, pp. 545–552, September 2007.
- [11] R. Achantay, S. Hemamiz, F. Estraday, and S. Süssstrunk, "Frequency-tuned salient region detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1597–1604, Miami, Fla, USA, June 2009.
- [12] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, Minneapolis, Minn, USA, June 2007.
- [13] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 8, pp. 630–632, 1978.
- [14] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [15] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 3, pp. 273–285, 1985.
- [16] W.-H. Tsai, "Moment-preserving thresholding: a new approach," *Computer Vision, Graphics, & Image Processing*, vol. 29, no. 3, pp. 377–393, 1985.
- [17] C. J. van Rijsbergen, *Information Retrieval*, Butterworths, London, UK, 1979.

