

A Novel Segmentation Method for High Resolution Urban Remote Sensing Image Based on Improved JSEG Algorithm

Aihua Gu^{1,2*} and Yinyong Li³

¹*School of Information Engineering, Yancheng Teachers University, Yancheng 224002, China*

²*School of computer and information science, Hohai University, Nanjing 210098, China*

³*School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China*

¹*yctuaihuagu@163.com*

Abstract

In view of the over segmentation, insufficient segmentation and difficulty in accurately determining the boundary of the objects, this paper proposes a novel improved JSEG algorithm. Firstly, we adopted the multi-band fusion strategy based on information entropy to obtain a unified multi-scale J-images sequence to avoid the loss of detail information caused by excessively rough image quantization. On this basis, the locally homogeneous index J-value was utilized to mark the minimum at multiple scales. Finally, JSEG was adopted to carry out regional segmentation and merger. The experiments adopted high resolution IKONOS and WorldView-II images, which were compared the segmented results with JSEG and WJSEG. The results indicate that the proposed method can accurately locate the real boundary of the objects and effectively restrain the problem of less segmentation and over segmentation. Moreover, it has higher accuracy and reliability.

Keywords: *High resolution; Remote sensing image; Urban; Multi-scale segmentation; JSEG*

1. Introduction

With the continuous improvement of the spatial resolution of remote sensing images, Object-Based Image Analysis (OBIA) technology has been widely utilized in remote sensing image interpretation [1]. Image segmentation is the basis of OBIA, and it extracts the objects with geographical significance in the scene [2]. The quality of segmentation has a direct impact on the accuracy of subsequent image processing, such as classification and change detection [3].

Compared with medium and low resolution remote sensing images, high resolution remote sensing images have more abundant spectrum, texture and spatial context information, which are conducive to finely describing the contour features of the ground objects [4]. However, more significant spatial details also lead to prominent over segmentation, insufficient segmentation, and uncertainty of object boundary uncertainty of traditional method in high-resolution remote sensing image segmentation [5]. Particularly, the man-made objects with different shapes and diverse natural feature distribution in urban scene, image segmentation is likely to be affected by the shadow of surface features, cloud coverage and a large number of interference factors, which make it difficult to realize image segmentation with high accuracy and high reliability [6-7]. Meantime, in spite of the numerous remote sensing image segmentation methods with middle and low resolution, and the high reliability and segmentation accuracy, there are

various deficiencies in the systematic property and pertinency of the high resolution remote sensing image segmentation in urban scenes [8].

JSEG proposed by Deng *et al.* is one of the most popular segmentation algorithm for color image [9]. Locally homogeneous index J-value has a strong ability to detect the local spectral homogeneity [10]. However, JSEG algorithm also has inherent limitations such as inaccurate edge of the positioning object and over segmentation [11]. To this end, some scholars have carried out relevant research. For instance, Komati *et al.* proposed three kinds of improved JSEG algorithms, including Fractal-JSEG, Fractal-only and Local Fractal Dimension [12-14]. Fractal-JSEG and Fractal-only mainly extract the detailed features of the object boundary, while Local Fractal Dimension is applied in the scene with similar spectrum and texture features of the background area and the foreground target. These three improved strategies are only applicable to common image segmentation. The research focusing on high resolution urban remote sensing image segmentation is scarce. For example, Li *et al.* proposed to utilize the incremental growth mode to detect image details in the segmentation. At the same time, the integrated information of color and shape is utilized in the merging process to calculate the merging distance so as to better reflect the boundary of the surface features, but this algorithm has a more prominent problem of insufficient segmentation [16]. Wang *et al.* proposed an improved JSEG algorithm based on wavelet transform (WJSEG), which can locate the boundary of the object accurately and maintain the contour integrity of large size objects [17]. Nonetheless, these algorithms adopt excessively rough quantization before the region segmentation and ignore the complementary information of multi-band images, which lead to the loss of spectral and texture information conducive to locating object boundary. Also, the threshold values in extracting initial seed region by these methods are all determined at a single scale, and it is difficult to identify various surface features with complex shapes and sizes in urban scenes.

On account of the above analysis, this paper constructs a multi-band fusion strategy based on information entropy and multi-scale minimum mark strategy based on J-value. Moreover, we propose a novel segmentation method for high resolution urban remote sensing image based on the improved JSEG algorithm.

This paper is composed of four sections. The next section elaborates the basic principle and implementation process of the algorithm. In the third section, the experimental results are compared and analyzed, and the precision is further evaluated. The last section comes to the conclusion.

2. Method

In order to realize the high resolution urban remote sensing image segmentation with high accuracy and high reliability, the proposed method mainly consists of three steps: Multi-band image fusion based on information entropy; Multi-scale minimum mark extraction based on J-value; Multi-scale region segmentation and merging. The implementation process of the algorithm is shown in Figure 1:

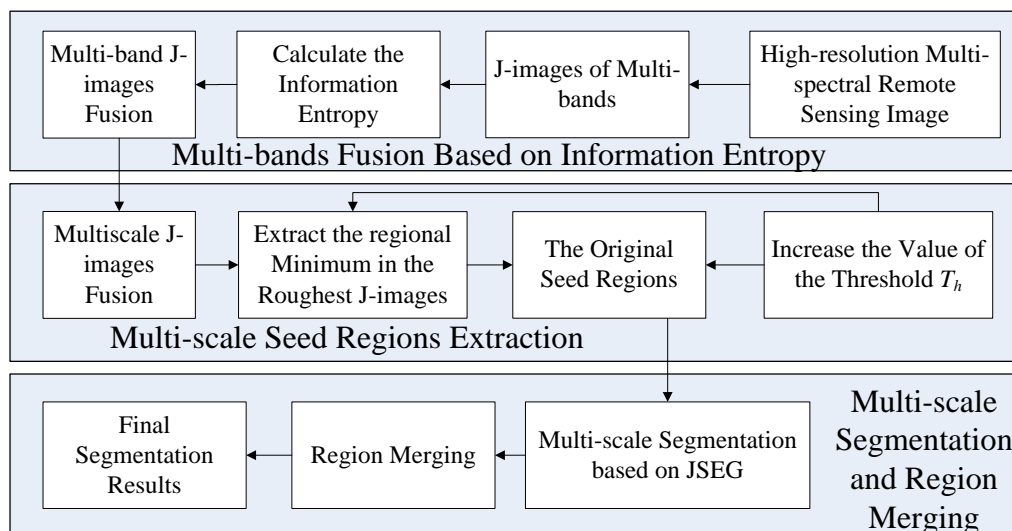


Figure 1. Algorithm Flow Chart

2.1. JSEG and Limitation Analysis

Traditional JSEG algorithm needs to quantize color of multi-band images before the segmentation so as to get a single band grayscale image. The image is transformed into LUV color space firstly. The gray level of the image conducted with Peer Group Filtering (PGF) smoothness is compressed through a clustering method similar to FCM so as to reduce the computational complexity.

On the basis of quantitative image, Deng *et al.* proposed a locally homogeneous index J-value, which can comprehensively reflect the spectral, texture and scale features of the local area. Its calculation process is: in the quantified image, the location $z(x, y)$ of each pixel z represents the corresponding pixel value. The specific size window Z is constructed for all pixels, and the size is $M \times M$ pixel. z is the central pixel of the window. Make m_p the mean value of the pixel with the gray level of p in Z . Z_p is the set of these pixels, and P is the total number of gray levels quantified. Define S_w as the variance of all pixels with the gray level of p in Z :

$$S_w = \sum_{p=1}^P \sum_{z \in Z_p} \|z - m_p\|^2 \quad (1)$$

Define S_T as the variance of all pixels in Z :

$$S_T = \sum_{z \in Z} \|z - m\|^2 \quad (2)$$

J-value is:

$$J = (S_T - S_w) / S_w \quad (3)$$

All pixels are traversed, and J-value is utilized to replace the original gray value to obtain a single scale J-image image. The multi-scale J-images sequence can be obtained by changing the scale parameter M . According to the definition of J-value, it can be seen that J-image is actually a gradient image. In the case of larger J-value, the pixel is more likely to locate on the edge of the object.

During the multi-scale segmentation, the connection regions with the threshold value smaller than $T_j = \mu_j + a\sigma_j$ are identified as the seed regions in the J-image

corresponding to the largest size window. μ_j and σ_j are the mean value and variance of J-image, respectively, $a \in [-0.2, 0.2]$. The mean value of all pixels in the non-seed regions is calculated. All pixels smaller than the mean value are incorporated into adjacent seed regions one by one. In the next fine scale, the same strategy is adopted to carry out the regional growing until the finest scale is calculated. Finally, the regions are merged based on the minimum Euclidean Distance criterion of adjacent regions to obtain the final segmentation results.

By analyzing the principle of JSEG, it can be found that there are two main limitations of JSEG in high resolution remote sensing image. Firstly, although image quantization is conducive to reducing the amount of computation, the quantized image usually contains only a few or more than a dozen of gray scales, which leads to the loss of a great number of image detail information. This is bound to have a negative impact on the accurate location of the object boundaries. In addition, the reasonable selection of threshold plays a key role in improving the over segmentation or insufficient segmentation. The single threshold extracted under the single scale cannot deal with the man-made objects with various sizes and shapes in urban scenes. As a result, the problems of over segmentation and insufficient segmentation are serious. Therefore, this paper will focus on improving these two limitations of the JSEG.

2.2. Multi-bands Fusion Based on Information Entropy

In order to avoid the loss of image details caused by image quantization and make full use of the complementary information of multi-band images, this paper carry out the feature level fusion strategy based on the information entropy of the J-value difference between the J-image of each band^[18]. The greater information entropy indicates higher separability of the band with the others. That is, a larger amount of information is contained and will occupy a greater weight in fusion.

Firstly, the values of a series of scale parameters $M_i (i=1,2...N)$ are determined so as to calculate the J-image sequence **J-image**_{*ij*} corresponding to each band in the original image, where *j* is the serial number of the band. The gray level difference of the J-value corresponding to a pixel between two bands is defined as *d*, and P_{j_1,j_2} is the probability that *d* occurs. The information entropy between band *j*₁ and band *j*₂ is

$$H_{j_1,j_2} = -\int_{-1}^1 P_{j_1,j_2}(d) \log P_{j_1,j_2}(d) \quad (4)$$

The sum of the information entropy of band *j* and other bands is $H_j = \sum H_{j,j_m} (l \neq m)$. Since greater H_j contains more information, the proportion of band *j* in feature fusion can be expressed as:

$$\omega_j = \frac{H_j}{\sum H_j} \quad (5)$$

Make the J-value value of each pixel in the scale *i*:

$$J_{z_i} = \sum w_j J_{z_j} \quad (6)$$

Where J_{z_j} is the J-value of pixel *z* in the first *j* band. So we can fuse J-image at all the scales in all bands in order to obtain a unified multi-scale J-image sequence, namely, J-image_{*i*} (*i*=1,2...*N*). The gray level of J-image_{*i*} is not compressed. As a result, the detailed information of the original image can be better preserved based on the complementary information of each band so as to provide a reliable multi-scale analysis platform for the subsequent image segmentation.

2.3. Multi-bands Fusion Based on Information Entropy

In the process of extracting the seed regions, in order to avoid over segmentation and insufficient segmentation as far as possible, the multi-scale minimum labeling method is adopted in J-image to detect the man-made objects and natural features with different size and internal texture complexity.

In the roughest scale J-image, the maximum threshold value $T_{J_{\max}}$ and minimum threshold value $T_{J_{\min}}$ of J-value is set firstly. $T_{J_{\min}}$ is corresponding to the object with the highest degree of homogeneity in the interior texture, and $T_{J_{\max}}$ is corresponding to the object with the highest degree of internal texture complexity. Make the labeling threshold value T_h gradually increase from $T_{J_{\min}}$ to $T_{J_{\max}}$ at the interval of 0.1 so as to get different labeling scales. In different scales, make the connection region with the J-value value less than T_h the initial seed region (*i.e.*, the labeled region). For the two adjacent labeled scales, the scales with large and small T_h are the upper parent label and lower son label. In the two labeled regions with the relationship of “father and son”, the centroid coordinates of the lower son labeled region shall be in the upper parent labeled region. On this basis, Structural Similarity (SSIM) is adopted to further screen the labeled areas with the relationship of “father and son”, namely, SSIM shall be greater than 0.6. The structural similarity of the two labeled regions R_x and R_y is calculated in J-image, and the formula is as follows [19]:

$$S(x, y) = \frac{(2\mu_x\mu_y + C_2)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

$\mu_x, \mu_y, \sigma_x, \sigma_y, \sigma_x^2, \sigma_y^2$ and σ_{xy} are the mean value, standard deviation, variance and covariance of x and y , respectively. C_1 and C_2 are the constants added in order to avoid the instability when the denominator is close to zero. In the case of the relationship of “father and son”, the upper labeled region will replace the lower labeled region. Otherwise, the lower labeled region will be retained to obtain the multi-scale seed region. Finally, based on the extracted seed region, the final segmentation results are obtained by the same region segmentation and merging strategy as the JSEG algorithm.

3. Experiment and Analysis

3.1. Experimental Data

In order to verify the accuracy and reliability of the proposed method, two high resolution remote sensing images with different sensor types are selected to carry out the segmentation experiments. Experiment 1 selects the Pan-Sharpned IKONOS image with the spatial resolution of 1m. The multi-spectral bands are composed of red, green, blue and near-infrared bands. The region is in Chongqing, China. The image size is 512×512 pixels, as shown in Figure 2. Experiment 2 selects the WorldView-II image with the spatial resolution of 1.8m. The multi-spectral band is composed of red, green, blue and near-infrared bands. The region is in Nanjing, China. The image size is 512 × 512 pixels, as shown in Figure 3.



Figure 2. The IKONOS Image



Figure 3. The WorldView-II Image

As shown in Figure.2 and Figure.3, both of the two images are typical urban scenes, including houses, roads and other man-made targets, as well as vegetation, wasteland and other natural features. Moreover, the image background is quite complicated due to the complex distribution of different types of objects, which is conducive to accurately evaluating the performance of the proposed algorithm.

3.2. Experimental Results and Visual Analysis

In addition to the proposed algorithm, we also selected the traditional JSEG algorithm and the WJSEG algorithm proposed by Wang *et al.* to carry out the comparative experiment. In the experiment, in order to evaluate the performance of different algorithms more objectively, the scale parameters of the three algorithms are set as $M \in [5, 7, 9, 12]$, and the other parameters are set according to the values suggested in literature 9 and literature 16. In the proposed algorithm, according to the trial-and-error method, preferentially set $T_{J_{\min}} = 0.1$, and $T_{J_{\max}} = 0.75$. The experimental results of the three algorithms are as shown in Figure 4 and Figure 5.



(a) The Improved JSEG Algorithm

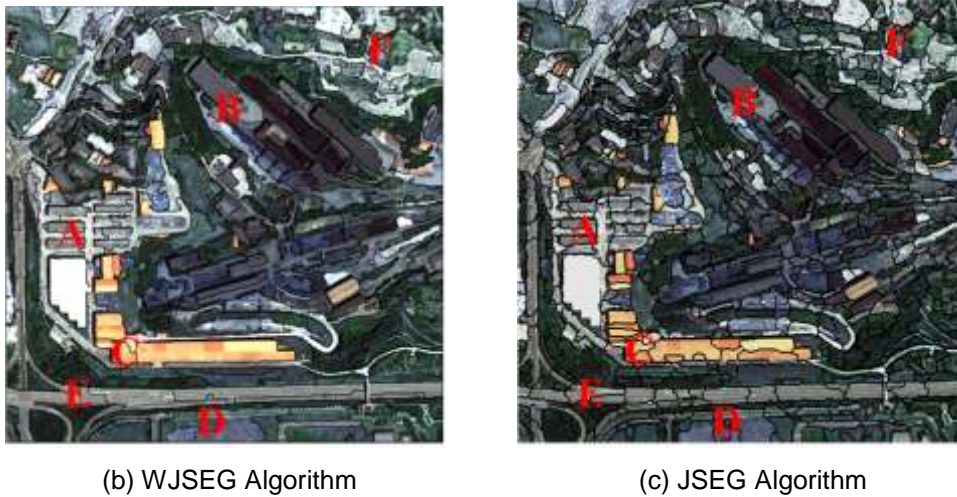


Figure 4. The IKONOS Image Segmentation Results



(a) The Improved JSEG Algorithm

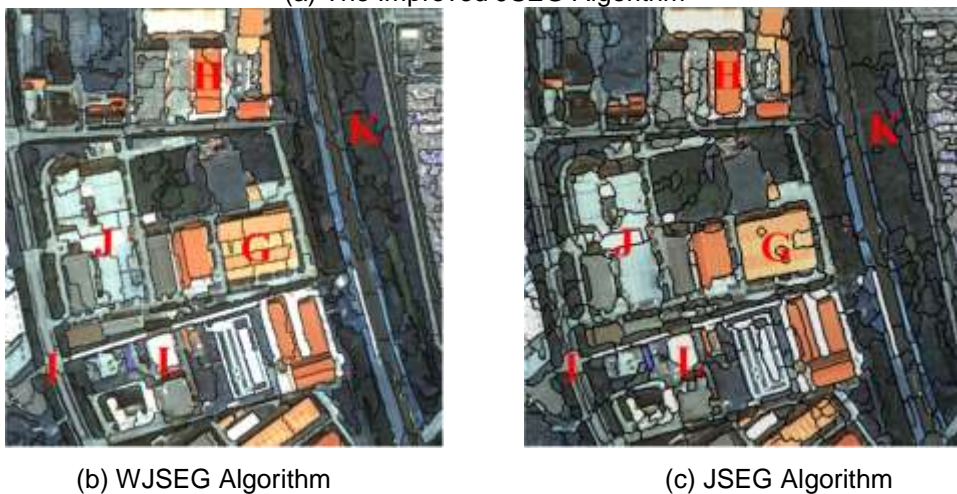


Figure 5. The Worldview-II Image Segmentation Results

As is shown in the figure, in order to facilitate the visual analysis, we label the typical features or locations in the image with letters. By comparing the experimental results of the three segmentation algorithms, it can be found that the segmentation results of the proposed algorithm and the WJSEG algorithm are significantly better than the JSEG algorithm. For instance, obvious error segmentation occurs in Position A, C and G. This

also proves that the JSEG algorithm with good performance in common image segmentation cannot be directly applied to the segmentation of high resolution remote sensing image. Based on the comparison between the proposed algorithm and WJSEG algorithm, it can be seen that WJSEG algorithm has a serious problem of over segmentation, which is particularly serious in location G and H, despite that both of the two algorithms can accurately locate the boundary of the large size objects with single internal texture feature, such as the roof of buildings of location B, C, G, H and the common ground region of location D. In the case that the boundary of the adjacent features is not quite significant, such as the location B, J, and L, the algorithm proposed in this paper is obviously more sensitive to the detailed information of the image and can accurately locate the weak edges. In contrast, WJSEG algorithm has the problem of insufficient segmentation. For urban roads, such as the location I and E, all of the three algorithms have certain problems of over segmentation. For vegetation regions, such as the location K and F, the three algorithms have achieved similar segmentation results.

To sum up, the segmentation effect of JSEG is the worst among the three algorithms. The algorithm proposed in this paper is significantly superior to WJSEG algorithm in suppressing over segmentation and insufficient segmentation. Also, it can better detect the weak edges, and can effectively maintain and accurately extract the detailed information of the image. However, the problem of over segmentation exists in extracting urban roads. In the two groups of high resolution remote sensing image experiments of different sensor types, the proposed algorithm shows similar segmentation capability and favorable reliability.

3.3. Accuracy Evaluation

In order to further evaluate the segmentation accuracy of the three algorithms, we select 1000 pixels as the sample set of the real boundary through the visual interpretation and evaluate the accuracy of the experimental results. The evaluation criteria is: The difference between the segmentation results and real boundary within 1 pixel is evaluated as “Very accurate”. The difference within 2 pixels is evaluated as “Relatively accurate”. The difference within 3 pixels is evaluated as “General”. The others are evaluated as “Poor”. The accuracy evaluation results are as shown in Table 1 and Table 2.

Table 1. The IKONOS Image Segmentation Accuracy Evaluation

	Very accurate%	Relatively accurate%	General%	Poor%
Improved JSEG	90.3	5.2	3.9	0.6
WJSEG	84.4	9.7	2.1	3.8
JSEG	75.6	14.3	5.5	4.6

Table 2. The WorldView-II Image Segmentation Accuracy Evaluation

	Very accurate%	Relatively accurate%	General%	Poor%
Improved JSEG	93.6	3.8	2.1	0.5
WJSEG	90.1	5.2	2.7	2
JSEG	85.9	7.4	2.5	4.2

As can be seen from the above tables, the boundary proportion that the algorithm proposed in this paper is judged to be “very accurate” and “relatively accurate” in the segmentation results of two groups of experiments reaches 95.5% and 97.4% respectively, which is significantly better than the other two algorithms. This is consistent with the results of visual analysis.

4. Conclusion

In terms of the problems of over segmentation and insufficient segmentation in high resolution urban remote sensing image segmentation, this paper proposed a novel improved JSEG algorithm. The improved algorithm gives up the excessively rough scale quantization based on the traditional JSEG algorithm, and carries out the feature level fusion strategy based on information entropy extracted from the difference of J-value between bands. While maintaining the detailed information of the image space, this strategy makes full use of the complementary information among multi-band images, which lays a good foundation for accurately locating object boundaries. In order to deal with the problem of over segmentation and insufficient segmentation, this paper also proposed a J-value based multi-scale minimum label extraction strategy, which can effectively describe the surface features with different sizes, shapes and texture complexity in urban scenes. Experimental results indicate that the proposed algorithm can achieve good results in high resolution urban remote sensing image segmentation of different sensors. The segmentation accuracy is significantly superior to JSEG and WJSEG algorithm, and has a favorable reliability.

Acknowledgments

This study is supported by the project funded by the the Natural Science Foundation of Jiangsu Province (Youth Fund, No.BK20160966), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No.16KJB510022), the National Natural Science Foundation of China (No.61601229).

References

- [1] H. Wu, Z. Cheng and W. Shi, "An object-based image analysis for building seismic vulnerability assessment using high-resolution remote sensing imagery", *Natural hazards*, vol. 71, no. 1, (2014), pp. 151-174.
- [2] C. Wang, W. Xu and X. Pei, "An unsupervised multi-scale segmentation method based on automated parameterization", *Arabian Journal of Geosciences*, vol. 9, no. 15, (2016), pp. 651-661.
- [3] J. Yuan, D. L. Wang and R. Li, "Remote Sensing Image Segmentation by Combining Spectral and Texture Features", *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 1, (2014), pp. 16-24.
- [4] G. Meinel and M. Neubert, "A comparison of segmentation programs for high resolution remote sensing data", *International Archives of Photogrammetry and Remote Sensing*, vol. 35, (2004), pp. 1097-1105.
- [5] Q. T. Vo, N. Oppelt and P. Leinenkugel, "Remote sensing in mapping mangrove ecosystems—An object-based approach", *Remote sensing*, vol. 5, no. 1, (2013), pp. 183-201.
- [6] T. Liu, L. P. Zhang and X. Lip, "Region segmentation of multi-spectral remote sensing images using improved JSEG algorithm", *J. Remote Sens*, vol. 13, no. 1, pp. 30-34.
- [7] D. Tuia, J. M. Marí and G. C. Valls, "Remote sensing image segmentation by active queries", *Pattern Recognition*, vol. 45, no. 6, (2012), pp. 2180-2192.
- [8] H. Sun, X. Sun and H. Wang, "Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model", *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 1, (2012), pp. 109-113.
- [9] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 8, (2001), pp. 800-810.
- [10] C. Wang, A. Shi and X. Wang, "A Novel Registration Method for High Resolution Remote Sensing Images Based on JSEG and NMI", *Computer Science & Information Systems*, vol. 12, no. 1, (2015), pp. 289-306.
- [11] A. F. M. G. Kibria and M. M. Islam, "Reduction of over segmentation in JSEG using canny edge detector", *International Conference on Informatics, Electronics & Vision. IEEE*, (2012), pp. 65-69.
- [12] K. S. Komati, E. O. T. Salles and M. S. Filho, "Fractal-JSEG: JSEG Using an Homogeneity Measurement Based on Local Fractal Descriptor", *SIBGRAPI*, (2009), pp. 253-260.
- [13] K. S. Komati and E. O. T. Salles, "Unsupervised color image segmentation based on local fractal descriptor and J-images", *Proceedings of the 2010 IEEE International Conference on Industrial Technology*, (2010), pp. 303-308.

- [14] K. S. Komati, E. O. T. Salles and M. S. Filho, "Unsupervised color image segmentation based on local fractal dimension", Proceedings of the 17th International Conference on Systems, Signals and Image Processing, (2010), pp. 243-246.
- [15] L. I. Nan, M. Liang and H. Huo, "Segmentation of high resolution remote sensing images based on improved JSEG algorithm", Journal of Xian University of Science & Technology, (2007).
- [16] C. Wang, A. Y. Shi and X. Wang, "A novel multi-scale segmentation algorithm for high resolution remote sensing images based on wavelet transform and improved JSEG algorithm", Optik-International Journal for Light and Electron Optics, vol. 125, no. 19, (2014), pp. 5588-5595.
- [17] W. Dai, C. Yu and Z. Jiang, "An Improved Hybrid Canopy-Fuzzy C-Means Clustering Algorithm Based on MapReduce Model", Journal of Computing Science and Engineering, vol. 10, no. 1, (2016), pp. 1-8.
- [18] K. Zhang, J. S. Kimball and S. W. Running, "A review of remote sensing based actual evapotranspiration estimation", Wiley Interdisciplinary Reviews: Water, vol. 3, no. 6, (2016), pp. 834-853.
- [19] W. Gao, S. Kwong and Y. Zhou, "SSIM-Based Game Theory Approach for Rate-Distortion Optimized Intra Frame CTU-Level Bit Allocation", IEEE Transactions on Multimedia, vol. 18, no. 6, (2016), pp. 988-999.