REGION OF INTEREST DETECTION BASED ON SALIENT FEATURE CLUSTERING FOR REMOTE SENSING IMAGES

Libao Zhang*, Xinran Lv, Jie Chen and Lan Zhang

College of Information Science and Technology, Beijing Normal University, Beijing, China

ABSTRACT

The region of interests (ROI) detection plays an important role in the remote sensing data processing and analysis. In this paper, a new region of interest detection method based on salient feature clustering for remote sensing images is proposed. Four steps are included in the proposed method. First, the information salient feature maps are constructed by computing the spectrum information and histograms of multispectral images. Second, a clustering strategy based on k-means is presented to generate the common salient feature maps in the CIE Lab color space. Third, the final saliency maps are generated by fusing the information salient feature maps with the common salient feature maps. Finally, we can get the ROIs by segmenting the final saliency map. Experimental results show that compared with five existing models, our model gets more accurate saliency maps without the basis of prior knowledge.

Index Terms—Remote sensing, image processing, salient feature clustering, spectrum information analysis

1. INTRODUCTION

With the improvement of ability to acquire high-resolution remote sensing images, the extraction of valuable targets from remote sensing images has become one of the most fundamental and challenging tasks for research [1, 2]. As an effective detection of the target area, the ROIs have become an alternative to improve the efficiency of processing. Thus, how to implement the ROI detection in remote sensing images accurately and quickly has become the problems to be solved currently.

Most traditional algorithms of ROIs detection in remote sensing images are global based or require prior knowledge. But they have high computational complexity and the establishment of prior knowledge is a complex issue.

In recent years, the saliency detection models based on visual attention have become the research hotspot, which covers many subjects, such as computer vision, image segmentation, video analysis, object recognition and remote sensing image processing.

Some models based on human visual system (HVS) have been developed [3-5]. Koch and Ullman [5] constructed a computational model which was later improved by Itti [4]. Itti’s model is based on biological architecture, which has been widely cited. This model adopts a Gaussian pyramid on three visual features: brightness, color and orientation to get the center-surround contrast. Harel et al. [6] proposed Graph Based Visual Saliency (GBVS) model, which complete the features extraction by Itti’s model, use the graph structure to represent the relation between the pixels of the image, and introduce Markov chains to compute saliency map. Hou et al. [7] extracted the spectral residual of an image in spectral domain by analyzing the log-spectrum of an image, and proposed a fast method to construct the corresponding saliency map in spatial domain. Achanta et al. [8] presented a frequency tuned model based on a Difference of Gaussian band-pass filter to compute full-resolution saliency maps by using low level color and luminance features.

Now, scholars have proposed several new algorithms for ROIs detection in remote sensing images. Zhang et al. [9] proposed a model based on frequency domain analysis and saliency region detection to extract ROIs of remote sensing images. They acquire the saliency maps using the quaternion Fourier transform, and get the ROIs using an adaptive segmentation algorithm based on the Gaussian pyramid. Besides, there is a new ROI detection model based on multiscale feature fusion [10]. The input image is processed along two feature channels: intensity and orientation, which are obtained by multi-scale spectrum residuals and integer wavelet transform, respectively. And they combine the conspicuity maps at different scales into one map by using a weighted across-scale fusion method.

The above models have their own pros and cons, but they all focus on single-image ROIs detection and ignore the inherent similarities among multi-images. When we get a group of remote sensing images with similar ROIs, we can take advantage of their similarities, thus excluding the salient regions that interfere with the ROIs detection.

Compared to the single-image saliency detection, multi-image saliency detection can cluster salient feature which means detecting the common salient objects in a group of images that contain similar goals.
2. METHODOLOGY

In this paper, we present a model based on salient feature clustering for detecting ROIs in remote sensing images. Firstly, we acquire the information salient feature maps by computing the Multi-spectral histograms, while clustering the group of the images in order to get the common salient feature maps. Then, we fuse the two saliency maps to be the final saliency map. Finally, we construct the mask of region of interest according to the final saliency map, which enable us to get the region of interest segmentation. Fig. 1 illustrates the framework of our model.

2.1 Generate the information salient feature maps

Remote sensing images have multiple spectrums, which can be used for the construction of one-dimensional histogram of intensities in different spectral channels based on spectrum analysis. Then the information content of the images is computed using the one-dimensional histogram.

For a group of remote sensing images with the size of $M \times N$, we first obtain the multi-spectral maps of the images and then construct the intensity histogram $H_c(i)$ of each image in different spectral channels.

$$H_c(i) = \sum_{x=1}^{M} \sum_{y=1}^{N} \delta_i(x,y)/(M \times N)$$

$$\delta_i(x,y) = \begin{cases} 1, & f_c(x,y) = i \\ 0, & \text{otherwise} \end{cases}$$

According to the histogram $H_c(i)$, we can get the information $In(i)$ of each pixel intensity in the spectrum $c$:

$$In(i) = -\ln(H_c(i))$$

Then the information $In(i)$ is assigned to the pixels in the spectrum $c$ to get the information map $LOG_c(x,y)$.

The saliency degrees $h_c$ of each spectrum are acquired by the information maps. Finally, the standardized saliency weight of each spectrum $w_c$ is calculated as is shown in equation (5):

$$h_c = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} LOG_c(x,y)}{\sum_{c=1}^{4} \sum_{x=1}^{M} \sum_{y=1}^{N} LOG_c(x,y)}$$

$$w_c = -\log\left(\frac{h_c}{h_1 + h_2 + h_3 + h_4}\right)$$

Now, the information salient feature maps $SS(x,y)$ can be constructed.

$$SS(x,y) = \sum_{c=1}^{4} w_c f_c(x,y)$$

2.2 Generate the common salient feature maps

In order to detect the common salient objects in the group of the remote sensing images, we carry out batch processing for saliency computation which starts from clustering.

Considering the various features of remote sensing images, color has been proven to be a useful and robust cue for distinguishing various objects. We choose to cluster the
color of the pixels, so it is necessary to select the appropriate color space. CIE Lab color space can remove luminance information to some extent thus we choose the CIE Lab color space for clustering in this paper.

For the input group of images, we implement clustering in the CIE Lab color space using the K-means method to get K clusters. The cluster number K is adaptively decided with regard to image content. In this paper, we partition the image into 3 clusters to obtain a more meaningful clustering map that matches the overall difference of the geomorphic features and at the same time maintain its efficiency.

After getting all the weights of the K clusters, we take advantage of the weights to calculate the saliency values of the clusters and then the common salient feature maps will be generated by assigning the values to all the pixels.

\[ o(l_i) = \text{weight of cluster } l_i \] and is calculated by the ratio between the number of pixels in cluster \( l_i \). \( D(l_i, l_j) \) is the distance of clusters \( l_i \) and \( l_j \), and \( CL(l_i) \) is the saliency value of cluster \( l_i \).

\[
D(l_i, l_j) = -\ln\left(1 - \frac{\sum_{s=t}^m (q_{ts} - q_{js})^2}{2 \sum_{s=t}^m (q_{ts} + q_{js})}\right) \tag{7}
\]

\[
CL(l_i) = \frac{\sum_{s} o(l_i) D(l_i, l_j)}{o(l_i)} \tag{8}
\]

Wherein, \( q_{ts} \) represents the probability of the color \( s \) appears in all kind of colors in cluster \( t \).

The saliency value of each pixel is equal to that of the cluster which it belongs to, and the common salient feature map \( CM(x, y) \) will be generated.

\[
CM(x, y) = CL(l_i) \tag{9}
\]

where \( SLab(x, y) \in l_i \) and \( SLab(x, y) \) represents the image in CIE Lab color space.

### 2.3 Salient feature map fusion

Considering two kind of salient feature maps, the information salient feature maps get the most salient objects in the images, but the disadvantage is that the fragmentation in the target regions and the paths can also be detected. However, in the common salient feature maps, the salient regions have good continuity, but cannot distinguish shadows. Therefore, it is necessary to fuse the two salient feature maps to get the final saliency map.

\[
S(x, y) = SS(x, y) \times CM(x, y) \tag{10}
\]

### 3. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluate the proposed model on the remote sensing dataset which contains 15 remote sensing images acquired by the SPOT 5 satellite with a resolution of 2.5 meters. Since there is no blue component in the remote sensing images acquired by the SPOT 5 satellite, we use near-infrared component instead.

We compare our model with five competing models through qualitative and quantitative experiments. The five saliency detectors are Itti et al. [4], Harel et al. [5], Hou and Zhang [7], Achanta et al. [8] and Zhang et al. [10], hereby referred as ITTI, Graph-Based Visual Saliency (GBVS), Spectral residual (SR), Frequency-tuned model (FT), and frequency domain analysis and salient region detection (FDA-SRD) respectively. Figure 2 demonstrates a qualitative comparison of the saliency maps and extracted ROIs created by the proposed model and the five competing models. ITTI model produces saliency maps that are just 1/256 of the original image size in pixels, and SR model outputs maps of size with 64×64 pixels for any input image size, thus, the boundaries are not well-defined. For GBVS model, the saliency maps include much background information and fail when objects touch image boundary to quite some extent. FT fails to highlight the entire salient area, which results in the incomplete description of ROIs. FDA-SRD fails to exclude the path from salient areas. Overall, the proposed model performs a complete ROI.

The receiver operator characteristic (ROC) curves were generated by classifying the locations in a saliency map into salient regions and non-salient regions with varying quantization thresholds. The area under the curve (AUC) of ROC reflects the accuracy of detection results. PRF is another commonly used evaluation index, including recall, precision and F-Measure which is introduced to give a more comprehensive evaluation of the testing model. Figure 3 and 4 give ROC curves and PRF of the proposed model and five competing models, respectively. The proposed model get the largest AUC and F-Measure than others, meaning its detection results is universally more accurate.

### 4. CONCLUSION

This paper proposes a model to detect a group of ROIs in remote sensing images, which contain the similar goal salient objects. In this model, we rely on the relevance information among a group of images to highlight the ROI of different images simultaneously. Experimental results demonstrate that the proposed model outperforms the other five models qualitatively and quantitatively and prove its pragmatic value in improving the efficiency and precision for ROI extraction from a set of remote sensing images.

### 5. ACKNOWLEDGMENTS

This work was sponsored by National Natural Science Foundation of China (Nos. 61571050 and 61071103), Beijing Natural Science Foundation (No. 4162033), and Open Fund of State Key Laboratory of Remote Sensing Science (OFSLRSS201621).
6. REFERENCES


