Biological Vision for Urban Vegetation Detection in Color Remote Sensing Imagery

Fang Xu

School of Geodesy and Geomatics
Wuhan University
Whuan, China
fxy@sgg.whu.edu.cn

Weiwei Jiang

School of Geodesy and Geomatics
Wuhan University
Whuan, China
626834986@qq.com

Abstract—**Urban vegetation, which scatters in gardens, settlements and streets etc., appears small, often fragmented, linear patches. Automatic vegetation detection in color remote sensing imagery is useful for obtaining more timely and accurate information. In this paper, a new biological vision methodology based on visual attention theory and adaptive resonance theory (ART) is presented to automatic detect urban vegetation in color remote sensing imagery. The central rationale of the method is that vegetation information is from the double-opponent saliency map and then gets object-based classification. Without a priori knowledge of image content, the image can be segmented into vegetation and other object through the unsupervised learning and self-organization fuzzy ART neural network. Experimental results indicate that our method performs much better than eCognition®.**

Index Terms—**urban vegetation; color remote sensing; visual attention; fuzzy ART**

I. INTRODUCTION

Urban vegetation is a critical component in urban ecosystems and urban landscape. It plays important roles in improving air, water and land quality, absorbing and mitigating carbon dioxide and many pollutants, lowering urban temperature [1], beautifying urban landscape and providing leisure facilities. Vegetation detection based on remote sensing data is by far the most widely used.

For vegetation detection, image spatial resolution is one factor to be considered. It is difficult or impossible to accurately map individual tree species and tree canopy by using moderate resolution satellite imagery especially in highly heterogeneous urban environment [1]. There are many small remnant and linear vegetation patches in city. Lechner et al. simulates the process of classifying small and linear vegetation on the appropriate spatial resolution remote sensing imagery [2]. This work proposes a biological vision method for urban vegetation detection in color remote sensing imagery.

II. VEGETATION CHARACTERISTICS IN COLOR IMAGE BASED ON VISION ATTENTION

Engel et al. suggested that spatial and chromatic opponency exists for the red/green, green/red, blue/yellow, and yellow/blue color pairs in human primary visual cortex in 1997. In the center of receptive fields, neurons are excited by one color (e.g., red) and inhibited by another (e.g., green), while the converse is true in the surrounding. The researchers proposed many simulation models of visual attention, in which the most representative is Itti model[3]. The model offered a basis for computational visual attention model using intensity, color, and orientation features. The algorithmic framework based on the color-double-opponent theory was used by Itti model [3].

A. Color saliency map

Four broadly-tuned color channels are created[3]: $R=r-(g+b)/2$ for red, $G=g-(r+b)/2$ for green, $B=b-(r+g)/2$ for blue, $Y=(r+g)/2-|r-g|/2-b$ for yellow (negative values are set to zero). Four Gaussian pyramids $R(\sigma)$, $G(\sigma)$, $B(\sigma)$, and $Y(\sigma)$ are created from these color channels. Center-surround differences ($\ominus$) between a "center" fine scale c and a "surround" coarser scale s yield the feature maps.

Maps RG and BY are created in the model to simultaneously account for red/green and green/red double-opponency Eq. 1 and for blue/yellow and yellow/blue double-opponency Eq. 2.

$$RG(c,s)=/(R(c)-G(c)) \ominus (G(s)-R(s))| \qquad (1)$$
$$BY(c,s)=/(B(c)-Y(c)) \ominus (Y(s)-B(s))| \qquad (2)$$

Where "center" fine scale $c \in \{0,1,2\}$, "surround" coarser scale $s=c+\delta$, $\delta \in \{2,3\}$. Six RG maps and six BY maps are created.

Feature maps are combined into the saliency map. They are obtained through across-scale addition, $\oplus$, which consists of reduction of each map to scale four and point-by-point addition: $\bar{C} = \bigoplus_{c=1}^{3} \bigoplus_{s=c+2}^{c+3} [N(RG(c,s)) + N(BY(c,s))]$. The result is illustrated in Fig. 1.

We also separately calculate RG, BY, color feature saliency map, see Fig. 1d, 1c.

$$RG = \bigoplus_{c=0}^{2} \bigoplus_{s=c+2}^{c+3} [N(RG(c,s))] \qquad (3)$$

$$BY = \bigoplus_{c=0}^{2} \bigoplus_{s=c+2}^{c+3} [N(BY(c,s))] \qquad (4)$$
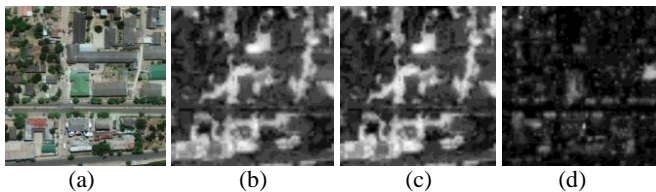
Fig. 1. Color feature saliency maps (a) the original image, (b) Itti saliency map, (c) BY saliency map, (d) RG saliency map.

The results show that the *BY* opponency saliency map comparison is larger in details. Most of the vegetation regions contrast clearly with the other objects in the *BY* saliency map, so that the majority of vegetation regions can be identified. While in *RG* saliency map, Fig. 1d, less comparison exists between vegetation regions and none-vegetation regions.

Many experiments support the validity of *BY* saliency map. Thus, we choose the *BY* saliency map as "Color saliency map".

B. Intensity saliency map

The brightness sensitivity of human visual is far stronger than the sensitivity of color shades. In HSI color space, hue, saturation and intensity are used to describe the color. The HSI color space is more in line with the human visual characteristics than the RGB color space. The intensity image is described as: $I=(r+g+b)/3$.

YUV color space is another color space, its luminance signal *Y* and chrominance signals *U, V* are separated. According to the standards of the National Television System Committee (NTSC), when the brightness of the white light is represented by *Y*, the relationship between *Y* and red, green and blue is described as: $Y=0.229R+0.587G+0.114B$.

The vegetation can't be highlighted by the intensity saliency map in HSI color space and intensity saliency map in YUV color space. This paper emulates and provides an opponent mode of the *I* and *Y*, and calculates opponent intensity saliency map. We use the *I* and *Y* instead of *R* and *G* in Eq. 1 and Eq. 3.

The experiment takes the Fig.1a as an example. Fig. 2c shows apparent difference between vegetation regions and none-vegetation regions. Although there is some confusion, less comparison area, on the whole it is better able to distinguish vegetation. So we select the opponent intensity saliency map as "intensity saliency map".

C. Saturation saliency map

Inspired by previous two successful vegetation features, this section also provides an algorithm of opponent saturation saliency map, which takes *S* in HSI color space and its histogram equalization result as an opponent feature.

We compute saturation $S0(\sigma)$ of each level of pyramid images, and calculate $S1(\sigma)$ through histogram equalization of $S0(\sigma)$. Then we use the $S0$ and $S1$ instead of *R* and *G* in Eq. 1 and Eq. 3.

$$S0(\sigma)=1-3/(r+g+b)[\min(r,g,b)] \qquad (5)$$

In the space of HSI, *S* expresses saturation. If we calculate saturation saliency map directly with only one expression, we haven't obtained an efficient feature. Unlike the color saliency

map and intensity saliency map, there are not obvious distinctions between the vegetation region and none-vegetation region in Fig. 2d. But it will be an irreplaceable role in the extraction process of vegetation, and it is auxiliary supplement to "color saliency map" and "intensity saliency map". In contrast with other two feature saliency maps, most vegetation regions are brighter than none-vegetation regions in saturation saliency map, see Fig. 2d. In Fig. 2a, none-vegetation regions in red circle can be identified only in saturation saliency map, but it isn't identified in color saliency map or intensity saliency map. And none-vegetation regions in blue circle are not identified from color saliency map, but it can be identified from any other two maps. These examples are just a small part in Fig. 2. It shows that the three features complement for each other, they will be very efficient vegetation features.
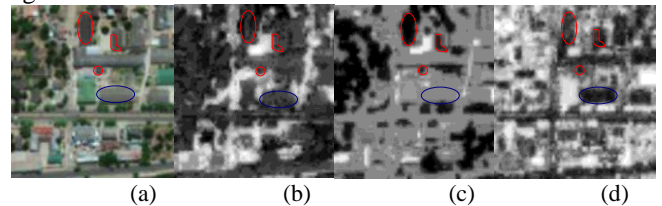

Fig. 2. Saliency maps (a) the original image, (b) color saliency map, (c) intensity saliency map, (d) saturation saliency map.

D. Other vegetation features

Besides these three primary saliency features, there are some other features which can play a little optimization role. From tests, it can be found that vegetation regions of some images have a larger comparison in blue channel than in red/green channel (as in Fig. 3b). And *H* channel of HSI color space, hue, also plays a role in some images' vegetation, Fig. 3c shows the distinction. Hence, blue channel and hue channel are taken as another two assistant features.
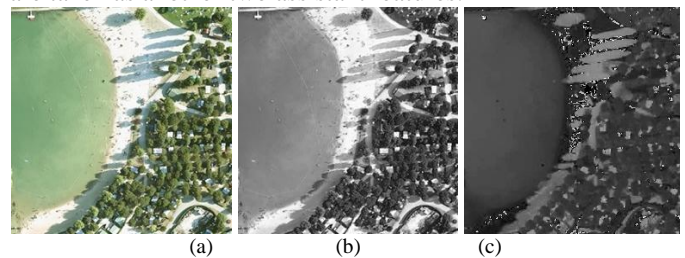

Fig. 3. Other features (a) the original image, (b) blue channel, (c) Hue channel in HSI color space.

III. VEGETATION DETECTION BASED ON FUZZY ART

Grossberg and Carpenter introduced fuzzy ART neural network in 1991[4]. Fuzzy ART networks are widely used in the clustering and pattern recognition etc. Carpenter et al. proposed fuzzy ARTMAP neural network for automatic mapping vegetation from Landsat thematic mapper (TM) and terrain data at the pixel level[5].

Fuzzy ART is an unsupervised learning and self-organization neural network. It completes a very complex nonlinear processing and information exchange by the interaction of the internal unit. Liu et al. point out that there are two major problems in the fuzzy ART algorithm: the saturation problem and similarity measure calculation[6].

We use the improved fuzzy ART algorithm[6] for vegetation object-based samples training and classification.

The identifying and detecting urban vegetation process occurs in two steps: a learning step and a detecting step. The learning phase uses the intensity saliency map to achieve object-based sample of vegetation and its memory. The detecting phase uses three saliency maps of double-opponency and two assistant features to identify and detect urban vegetation.

A. Object-based Sample collection

We extract the intensity saliency map of the series of images. It can be seen that vegetation's luminance is dark by contrast with other object. The dark areas almost cover all vegetation, and the vegetation proportion is much larger than the other objects. Fig. 2c shows that vegetation, roof, shadow and road are in the dark area, and vegetation pixel has a far greater number than that of others.

We use the threshold segmentation method to extract vegetation, and set them as the initial vegetation sample. Fig. 4a presents the threshold selection method of Fig. 1a. The initial sample contains few other objects and vegetation that account for the largest proportion.

Afterwards, we calculate five feature saliency maps. The five characteristics include color feature saliency (C), intensity feature saliency (I), saturation feature saliency (S), blue channel value (B), Hue channel value in HSI color space (H).

B. Networking training and classification

Firstly, we train the improved fuzzy ART network. It is an unsupervised classification process. While the five features of the initial sample are input to the improved fuzzy ART network for training, we set the vigilance parameter ρ, the learning rate β, the width of Gaussian membership function σ and the class number N=1. After training, we get several classes. The category of most number of pixels is vegetation. We only keep the vegetation as object-based class information, and delete other class information.

Secondly, we reset the network parameters such as ρ according to the characteristics of the vegetation category.

Finally, we use the network, which only contain the vegetation category, to classify the original imagery, and then extract the vegetation. This is a supervised classification process.
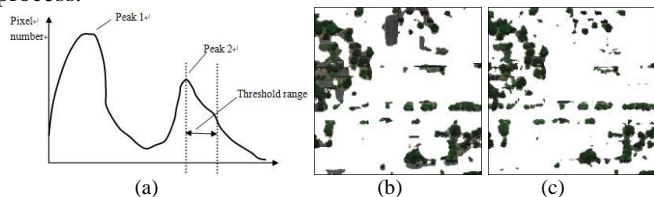

Fig. 4 Vegetation detection based on fuzzy ART (a) Threshold range, (b) Fuzzy ART, (c) improved fuzzy ART.

Fig. 4b is the vegetation detection result based on fuzzy ART, where β=1, ρ=0.86. Fig. 4c is the result of using the improved fuzzy ART, where β=1, ρ=05.

The experiments show that some houses were misclassified as vegetation in Fig.4b, the result, which based on improved fuzzy ART, is more accurate (Fig. 4c).

IV. VEGETATION DETECTION AND ANALYSIS

In order to evaluate the performance of our biological method, we have tested the algorithms on real high resolution color remote sensing imagery. Various color remote sensing images are tested with our method and eCognition®. The eCognition®, developed by Germany Definiens Imaging, is the mainstream software of remote sensing image processing. In these experiments, eCognition® needs to manually select the samples of vegetation and non-vegetation. It is a semi-automated process. The proposed method can automatically select sample, without human intervention.

In the following figures, the regions in red circle show misclassification, and the regions in blue circle are omitted.

A. Fragmented vegetation

Due to the complex structure and type of urban object, vegetation often presents small patches which scatter in gardens, settlements, streets, etc. It is difficult to detect the vegetation. So we select an image with typical vegetation characteristics. In the image, trees staggered with other objects and scattered orderlessly, which is interfered by the other objects with the same spectral characteristics.

Fig. 5b shows that the proposed method has the fewer misclassification and omissions, better anti-interference ability, and higher accuracy rate. The proposed method automatic extract vegetation class, and classification rely solely on the characteristics of the vegetation.
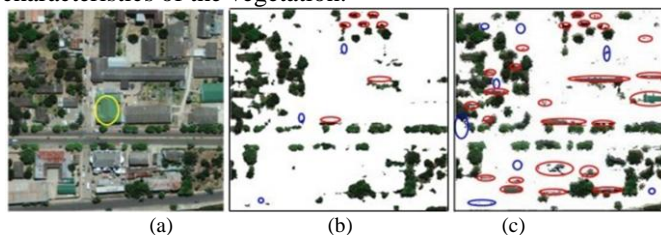

Fig. 5 Results comparison of vegetation (a) the original image of fragmentary vegetation, (b) detection using the proposed method, (c) detection using eCognition®.

If only set vegetation class and its samples, the result of eCognition® is relatively worse. So it needs to set non-vegetation class and its samples. In Fig. 5a, we manually set non-vegetation as region in yellow circle to avoid the phenomenon of same spectrum with different objects. But Fig. 5c still shows that the result is misclassification.

B. Same object with different spectrums and same spectrum with different objects

The same object with different spectrums and the same spectrum with different objects are two major difficulties in the interpretation of remote sensing images. There is the same object with different spectrums in the Fig. 6a. Some vegetation is reddish-brown, some vegetation is dark green, and other vegetation is emerald green. By using eCognition®, we need to set the reddish-brown vegetation, dark green vegetation and emerald green vegetation as vegetation samples. There is the same spectrum with vegetation and water in Fig. 6d. By using eCognition®, we need to set the water as non-vegetation sample. Otherwise it will be mistaken

for vegetation. In the proposed method, the automatic sampling has not added artificially information, without manual intervention.

Experiments show that the proposed method has the better results, but eCognition ® has more misclassification results. Even water has been set as non-vegetation sample in eCognition ® , Fig. 6f shows that water is classified as vegetation in the upper left corner. In addition, there is some shadow in Fig. 6d, the proposed method is better than eCognition ®. Section 4.4 demonstrates a similar phenomenon as tree shadow of Fig. 6f.
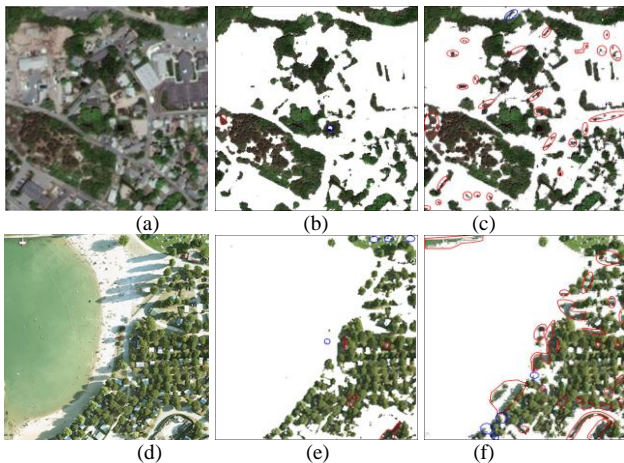

Fig. 6 Results comparison of same object with different spectrums and same spectrum with different objects (a) (d) the original image, (b)(e) detection using the proposed method, (c)(f) detection using eCognition®.

C. Shadow interference

In urban aerial images, shadows usually result in information loss or distortion of objects. In the research of vegetation extraction, the shadow is often misclassified as vegetation. The proposed method has better ability of anti-interference for shadow, and it reduces, in a way, the effect of shadow.
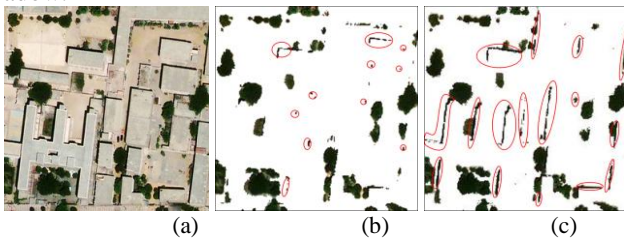

Fig. 7 Results comparison of shadow interference (a) the original image, (b) detection using the proposed method, (c) detection using eCognition®.

Fig. 7b shows that the proposed method perceives most of the shadows as non-vegetation. Fig. 7c shows that eCognition ® doesn't perceive shadow, and many houses shadows are misclassified as vegetation.

V. CONCLUSION

Automatic vegetation detection of high resolution color remote sensing image is widely used in urban environment and development. The challenge lies in the fact that many small remnant and linear vegetation, shadow patches, vegetation of different color are in city image, and vegetation indices are often obvious different in different region and time. In such method of visual saliency and fuzzy ART which are based on biological vision is a potentially powerful way for improving classification accuracy. This paper presents and evaluates the use of biological vision for vegetation detection. It is constructed that an urban vegetation automatic extraction model based on visual attention theory and adaptive resonance theory. The paper adopts an unsupervised learning and self-organization fuzzy ART neural network based on region algorithmic, in the network, we serve feature saliency map as the feature descriptor. Our contribution, as part of the research presented here, is the double-opponent saliency map of color, intensity and saturation. The threshold selection method simplifies the choice of the initial sample. The proposed method significantly improves vegetation detection accuracy as compared with using eCognition®.

We test our method by using various color remote sensing images of typical urban characteristics, the images contain fragmented vegetation, vegetation with different spectrums and same spectrum with different objects, and shadow interference.

REFERENCES

[1] R. Pu, S. Landry, "A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species". Remote Sensing of Environment.Vol. 124, pp. 516–533, 2012.

[2] A. M. Lechner, A. Stein, S. D. Jones, J. G. Ferwerda, "Remote sensing of small and linear features: Quantifying the effects of patch size and length, grid position and detectability on land cover mapping". Remote Sensing of Environment. Vol. 113, pp. 2194–2204, 2009.

[3] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis". IEEE transactions on pattern analysis and machine intelligence. Vol. 20(11), pp. 1254-1259, 1998.

[4] G. A. Carpenter, S. Crossberg, D. Rosen, "ART2-A:An adaptive resonance algorithm for rapid category learning and recognition". IJCNN-91-Seattle International Joint Conference on Neural Networks, IEEE Conference publications, 1991(2), pp. 151-156.

[5] G. A. Carpenter, M. N. Gjaja, S. Gopal, C. E. Woodcock, "ART Neural Networks for Remote Sensing: Vegetation Classification from Landsat TM and Terrain Data". IEEE Transactions on Geoscience and Remote Sensing. Vol. 35(2), pp. 308-325, 1997.

[6] Y. Liu, Q. Liu, Q. Ren, "Speech signal endpoint detection and separation based on improved fuzzy ART". Systems Engineering and Electronics. Vol. 26(8), pp. 1151-1154, 2004.