HIGH-RESOLUTION SATELLITE IMAGERY MODULATING BASED ON MULTI-SCALE OBSERVING FEATURE

Xu Fanjiang[a],Zhang Jinfang[a], Li Bangyu[a], Liu Jianhua[a], li Yaping[b]

[a] Science and Technology on Integrated Information System Laboratory, Institute of Software Chinese Academy of Sciences, 4# South Fourth Street, Zhong Guan Cun, Beijing 100190 P.R. China - (jinfang, fanjiang, bangyu, jianhua) @iscas.ac.cn

[b] Graduate University of the Chinese Academy of Sciences, No.80 Zhongguancun East Road, Haidian District, Beijing, China, 100190 - yaping10@iscas.ac.cn

KEY WORDS: High-Resolution Satellite Imagery, Multi-Scale Observing Feature, Local Feature Description, Scale Explicit Image (SEI), Difference of Gaussians (DoG)

ABSTRACT:

Objects extraction from high-resolution satellite imagery has raised extensive attention in the past twenty years. And it is still a hard problem, which is based on image feature representation. This paper describes a method based on the characteristics from observing by multi-scale. Firstly, multi-scale observing image features based on pixel and its significance to objects detecting are introduced; Secondly, the scale explicit images (SEIs) for scale feature description are derived, which generated by a method of modulating that can be explained through a lamp-standard model; Thirdly, the analysis of the SEIs characteristics is given; Finally, segmentation experiment based on the SEIs in commercialization software eCognition shows its application value.

1. INTRODUCTION

For object recognition, there are a lot of research work on global features extracting from images such as color and texture (Carneiro, 2005; Fan,2004; Niblack, 1993). An image can be represented by a global feature vector. And the problem of analysing image semantic is turned into the problem of supervised classifying. On the other way, local feature is studied as an improvement on global feature. DoG (Difference-of-Gaussian) is used to detect interest points from image and then SIFT (Scale Invariant Feature Transform) (Lowe, 2004) is used to extract a vector of feature from each of those points. There are also research on the integrating local feature and globe feature for object recognition (Lisin, 2005).

But the local features are all related to scale, which is an important characteristic. And the scale we referred is **Observing Scale**, not **Object Scale**. So the multi-scale segmentation (Baatz, 2000; Benz, 2004; Tzotsos, 2008)and their features are not similar to ours.

From atomic to galactic, the world is a structure of multi-levels; Accordingly, The Observation of the objects in it is also multi-level, and the observation tools from microscope for nano objects to telescope for objects measured by light years. Remote sensed imagery is an observation with a minimal scale of pixel size. Observing the image with a serial larger scale, we can find objects such as car, house, lake etc. For the same area, specified object can be detected within a specified scale bound. Out of the bound, the object cannot be detected (the atomic action cannot be observed with a telescope). The detectability of an object in a serial observation scale is a feature of object itself which is related to its size, shape, and contrast, named **Scale Feature**.

Objects in an image region, are recognized basing on distinguishability among objects and unity within an object with a specified observing scale. And the objects recognition is a balance between the largest internal unity and external difference.

A pixel's relative brightness varies with the observing area. In left image of figure 1, the center pixel appears darker within the 1 pixel neighborhood area, and appears brighter within the 2 pixel neighborhood area. The right curve in figure 1 illustrates the relative brightness of central pixel observed with different scale.

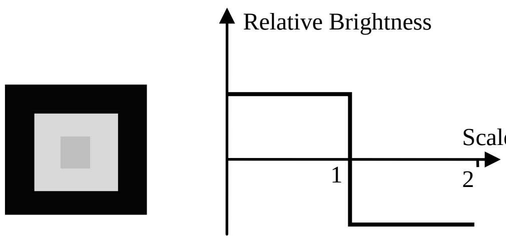

Figure 1. Relative brightness for center pixel observed with different observing scale

Other attributes based on the relationships of a pixel to its neighborhood can also be used. For instance, gradient and direction. In this article, the scale features are based on pixel brightness to its neighborhood.

Generally, pixels belong to the same ground object share common scale features. Scale features of an pixel is coded into pixel values, and so a new type image is formed which named as Scale Explicit Image (SEI). Pixel values of SEIs represent the scale semantics. SEIs possesses many excellent characteristics; the important one is that image objects segmented on SEIS are more pure and complete than on origin image.

2. THE GENERATION OF SCALE EXPLICIT IMAGE (SEI)

2.1 The principle of modulating the scale features into image

Pixel value of input image can be any single attribute such as brightness, gradient, direction. At first a serial single scale SEIs are produced. Then multiply single-scale SEI are synthesized into multi-scale SEI according to their weights (figure 2). We call this process scale feature modulating.

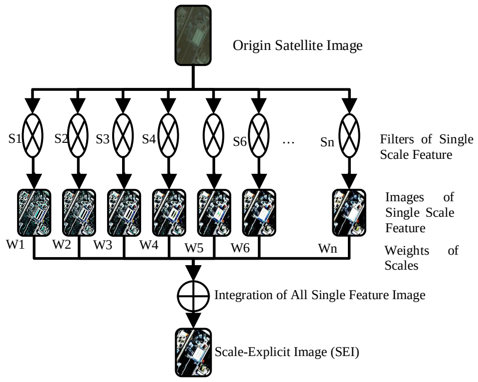

Figure 2. The process of modulating to build the scale-explicit image (SEI)

For image recognition, the relative features have better adaptability than the absolute features. For instance, the brightness of an object can vary with the light, the point of view and spectral bands used; however, the relative brightness of the object to its background is almost unchanged. And multi-scale relative features can depict comprehensive characteristics within a specific scale bound.

In order to conveniently represent the scale features, the relative feature values are binarized as 1 for those pixels holding the attribute, or 0 for those pixels not holding the attribute. And assuming that weight of relative feature of large-scale is twice as that in adjacent small one. So, the relative feature in one - scale can be expressed by a bit and that in multi-scale can be represented by continual bits which small scale locates in low bit; that is scale feature values which derived from the value of original pixel values have scale meaning and keeps relative brightness and hue in multi-channels.

This modulating process can be explained by a lamp-standard model which illustrated in Figure 3. On every pixel center stand the pillar with a serial lamp on it from bottom to top, represent one bit of the pixel value in turns (low lamp for low bit). And each lamp covers twice area as adjacent lower one. If the corresponding binary bit is 1, the lamp will be lit, contrarily extinguished; the total brightness of all lamps represents for the pixel value. The principle to judge the lamp to be lit or not is as follows: if the lamp belongs to relative light within its irradiating scope, turn on, otherwise turn off. So, the lamps being lit process the significance of relative lightness within its coverage range.

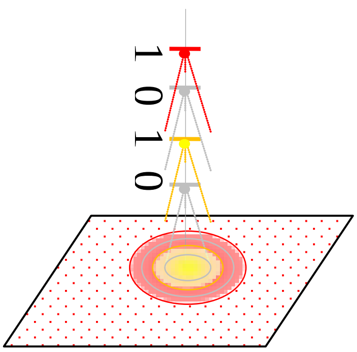

Figure 3. lamp-standard model

The pixel value on the SEI represents the relative features of specific scale bound; if one level lamp is opened by masking operation to keep the bit value, the image shows the relative features in the observed scale. When the lamp of the high lights in turn turns off, the scale of SEIs will be gradually reduced, and finally, the fine texture will be seen and the bigger scale features disappear. On the contrary, the small texture will be disappeared and the large object will be highlighted.

2.1 Implementation of SEIs

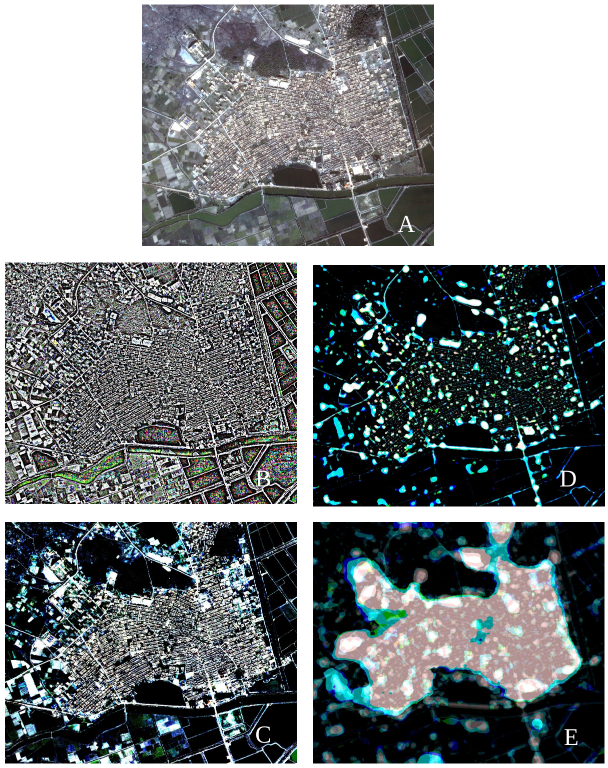

Figure 4. Two DoG sequences and their effect on SEIs. A - the origin image; B - SEI using DoGo method with scale from 0.1 to 12.8; C - SEI using DoGo method with scale from 0.4 to 51.2; D - SEI using DoGp method with scale from 0.1 to 12.8; E - SEI using DoGp method with scale from 0.4 to 51.2;

Scale features can be described by relative lightness of a pixel within its neighborhood, and the convenient method is utilizing convolution operation, but for its low efficiency to large convolution kernels, A DOG (Difference of Gaussians) method is introduced for this objective. Smoothing parameter σ of Gaussian filter is used for observing scale, and the scales form a geometric sequence with a common ratio of $\sqrt{2}$. The lightness of a pixel at specific observing scale is binaries to 0(light) or 1(dark). The SEI pixel value is calculated by accumulating all relative lightness of the origin pixel at every observing scale level with a weight of 2^{i} (i is scale level from 0). Apply the formula as follows:

$$S_v = \sum_{i=0}^{n} 2^{i} \cdot L(u, v, R)$$

Where: $R = 2^{0.5i} S_0$. n denotes the levels of scale. S_0 denotes the scale of level 0. $L(u, v, R)$ denotes the relative lightness of pixel at (u, v) within the neighborhood of R range(the observing scale), the function value is 0 or 1.Based on DOG method, R is equivalent to the smoothing parameter σ in the Gaussian filter, L (.) uses the following formula.

$$L(u,v,\sigma)=\begin{cases}1, DoG(u,v,\sigma)*f(u,v)-\tau\geq 0\\ 0, DoG(u,v,\sigma)*f(u,v)-\tau<0\end{cases}$$

Where τ is a small value to ensure the rightness of relative lightness, $f(u,v)$ is origin image. $DoG(.)*f(u,v)-\tau$ is the discriminant function, And the DoG(.) can takes the following forms:

$$DoG_o(u,v,\sigma)=1-\frac{1}{2\pi K^2\sigma^2}e^{-\frac{u^2+v^2}{2K^2\sigma^2}}$$

$$DoG_p(u,v,\sigma)=\frac{1}{2\pi K^2\sigma^2}e^{-\frac{u^2+v^2}{2K^2\sigma^2}}-\frac{1}{2\pi(k+1)^2\sigma^2}e^{-\frac{u^2+v^2}{2(K+1)^2\sigma^2}}$$

Where K is the ratio between two neighbor scale, in this case ,it is $\sqrt{2}$. DoG_o is origin scale based DoG, and DoG_p is previous scale based DoG. Though have similar effects , the DoG_o (figure 4 - B,C) is suitable for detecting detail texture, and the DoG_p (figure 4 – D,E) is suitable for image generalization.

3. CHARACTERISTICS OF SEI

Scale-explicit image described by relative lightness has characteristics as follows：

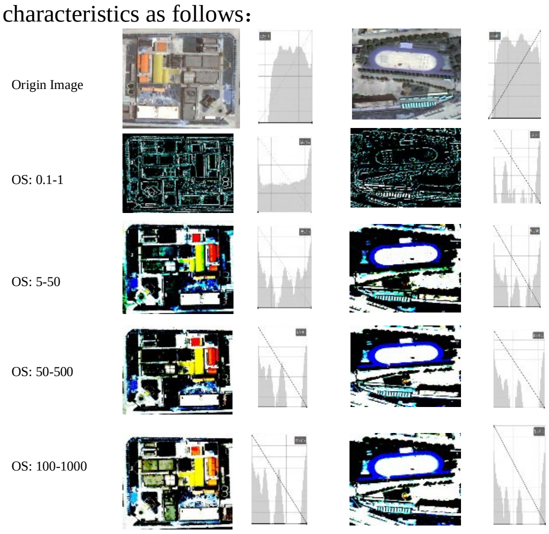

Figure 5. Landscape objects' appearance at different observing scale in SEI with logarithmic histogram. At the scale 0.1-1.0, smaller object s such as cars and edges are enhanced; at scale 5-50, salient buildings appear as blocks; At scale 50-500 large landscape objects such as course appear perfectly; And at scale 100-1000, some dim buildings appear

a) Pixel value has double meanings of lightness and scale;

b) A SEI has the same hue as origin image, with enhanced contrast and saturation (Figure 5);

c) The SEI's histogram assumes a form of discrete multi-peaks(Figure 5);

d) On a SEI, the feature of every scale reflects in different bits of the pixel value, ordered by scale level , so the objects belongs to specific observing scale can be enhanced by bit masking or assigning new weight values for every bits ;

e) Segmentation based on the SEI can decrease the mixing objects (Figure 6, Figure 7).

4. SEGMENTATION EXPERIMENT OF SEI

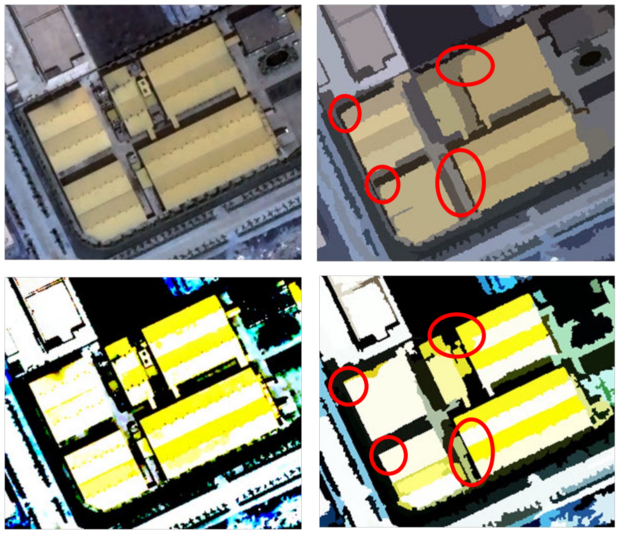

Figure 6. Integrality contrast between the objects segmented based on origin image and SEI. Left top: origin image. Right top: optimal segmentation based on origin image. Left bottom: SEI of the top left image. Right bottom: optimal segmentation based on SEI. (processed in eCognition 8.5)

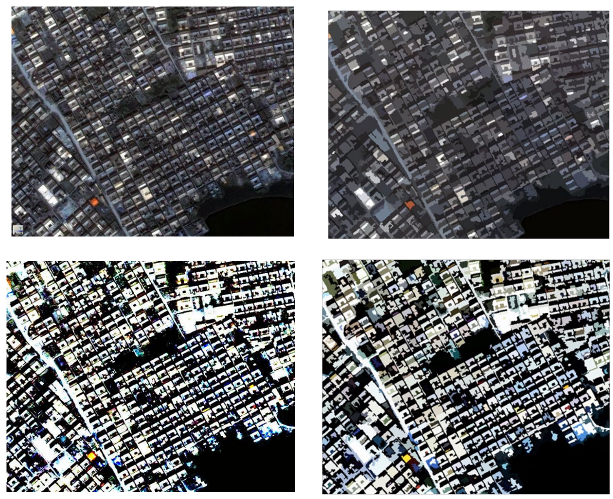

Figure 7. Contrast of Segmentation Results of origin image and SEI at whole area left top: origin image. Right top: optimal segmentation based on origin image. Left bottom: SEI of the top left image. Right bottom: optimal segmentation based on SFI. (Segmentation processed in eCognition 8.5)

In order to better observe the characteristics of the SEIs, segmentation experiment is done in the software eCognition (Martin, 2000), Two images are selected, one for low buildings in urban area (figure 6) and one for dense house area of village (figure 7). The image resolution is 0.51 meters. The experiments are illustrated in Figure 5 and Figure 6. From the SEI in figure 5, it can be seen that the difference between buildings and the ground is increase. The building edges are enhanced; From the SEI in figure 6,it can be seen that the houses are independent of the ground, however, A image object segmented based on original image are always a mixture of a house and its adjacent background.

5. CONCLUSION

The SEIs created through scale feature modulating embody scale semantic in pixel value while preserving origin relative lightness and hue. Objects in SEIs can easily be distinguished from background, and the completeness and purity of objects from SEIs get improved. For its explicit scale semantic, SEIs can easily be enhanced for objects of specific observing scale by bit masking or assigning new weight values for every bits of pixel value. On SEIs, edge detection has the same process as block object, i.e., edges are objects detected in small observing scale. Besides image objects, Pixels is another carrier for semantic information which should not be ignored.

6. REFERENCE

Baatz M., 2000. Multi-resolution Segmentation – an optimization approach for high quality multi-scale image segmentation. *In: Strobl, J. et al.(eds.): Angewandte Geographische Informations verarbeitung XII. Wichmann,Heidelberg*, pp. 12-23.

Benz U., 2004. Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS ready information. *ISPRS Journal of Photogrammetry and Remote Sensing 58*,pp. 239-258.

Carneiro, G.,2005. Formulating semantic image annotation as a supervised learning problem. *In Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 163-168.

Fan, J.,2004. Multi-level annotation of natural scenes using dominant image components and semantic concepts. *In Proceedings of the 12th annual ACM international conference on Multimedia*, pp.540-547

Lisin, D.A. 2005. Combining Local and Global Image Features for Object Class Recognition. *Computer Vision and Pattern Recognition - Workshops, 2005*. CVPR Workshops. IEEE Computer Society Conference，pp.1-8

Lowe, D.G. 2004. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, Vol. 60,pp. 91-110.

Martin Baatz, 2000. Multiresolution Segmentation : an optimization approach for high quality multi-scale image segmentation . *Journal of Photogrammetry and Remote Sensing* Volume: 58, Issue: 3-4, Publisher: Herbert Wichmann Verlag,pp: 12-23

Niblack, W., 1993. The QBIC project: Querying images by content using color, texture, and shape. *In Proceedings of SPIE on Storage and Retrieval for Image and Video Database*. pp. 173-187.

Tzotsos A., 2008. hybrid texture-based and region-based multi-scale image segmentation algorithm. *Object-Based Image Analysis Lecture Notes in Geoinformation and Cartography*, Section 2, DOI: 10.1007/978-3-540- 77058-9_12. pp.221-236