AUTOMATIC IDENTIFICATION OF CLOUD AND SNOW BASED ON FRACTAL DIMENSION

Haiyan DING, Lingling MA, Ziyang LI, Lingli TANG

Academy of Opto-Electronics, Chinese Academy of Sciences.
No.9 Dengzhuang South Road, Haidian District, Beijing 100094, P.R. China.
Tel: +8610 -82178614; E-mail: dingye43@163.com

KEY WORDS: Panchromatic image, Cloud and snow identification, Fractal dimension, Threshold segmentation

Abstract: The similarity of spectral feature in visible/near-infrared band between cloud and snow has been an important influence which degrades the recognition accuracy of cloud and snow, especially with the panchromatic images. In this paper, a novel and feasible method is presented to automatically identify cloud and snow from panchromatic images. The method makes full use of two different analytical techniques: the spectrum threshold segmentation and the texture analysis. These two approaches discriminate the image from two different aspects. At first, the cloud or snow is distinguished from the background utilizing the difference of spectral feature so the proportion of cloud or snow in the image was got. And then the samples' fractal dimension values which could reflect the texture features of cloud and snow from images are calculated to get the distribution of the fractal dimension values. At last, by comparing the proportion to the distribution, the automatic identification of cloud and snow is realized. The experimental results with the actual panchromatic images of Beijing-1 indicate the feasibility and accuracy of the method. The method could be also applied for other high resolution panchromatic images because of the universality of the texture feature.

1. INTRODUCTION

In the visible remote sensing images, the spectral feature of clouds and snow is strikingly similar, but they have different influences in earth observation remote sensing field. The presence of clouds not only directly affects the balance of radiation and energy between the earth and gas system, but also "contaminate" the radiation values in the satellite field. Therefore, cloud percent of each image is usually calculated to facilitate the user to determine whether the image is available or not in earth observation remote sensing data processing system. The spectral feature of clouds is far different from objects, but is strikingly similar to the snow in the visible band, as shown in Fig1 (Fenghua Wei, 2007), so they are easily confused in panchromatic images, especially in the processing of automatic cloud detection when the false detection is mainly caused by snow. In addition, as an important direction of remote sensing disaster mitigation, the accuracy and timeliness of snow monitoring are often seriously affected by clouds. Therefore, cloud and snow identification, especially objectively and stably automatic identification is a key problem in the preprocessing of optical remote sensing images or snow monitoring.

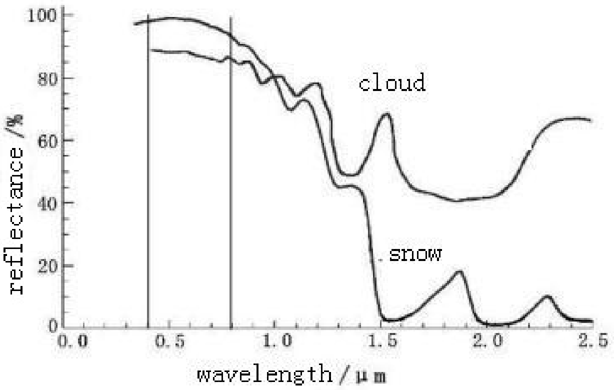


Fig.1: Reflection spectrum of cloud and snow

Currently, common methods of Cloud and snow identification are on basis of visible and shortwave infrared bands. The reflectance in the visible band is exactly similar, but is far different in the shortwave infrared band. They are usually simultaneously used to construct a divergence of cloud and snow. For example, Qingjun Yin took advantage of this feature to identify cloud and snow with NOAA/AVHRR images (Qingjun Yin, 2002). Xiaobo Zheng took advantage of the feature to establish normalized snow index to do snow monitoring with Modis images (Xiaobo Zheng, 2005). The results were both satisfied. However, for panchromatic images, specific method for automatic cloud and snow identification is relatively fewer. The spectral feature of clouds is so similar to the feature of snow in the visible remote sensing images that they could not be identified with the spectrum character. It's just one of the reasons that cloud coverage is rarely given for panchromatic images or visible multispectral images in domestic. With the appearance of high resolution remote sensing images, the application of texture feature has been a research direction in the field of image understanding. Texture feature is an important feature, which not only reflects the statistical gray scale, but also embodies the structure characteristics and the spatial relationship of objects. Therefore, it could be used to distinguish different objects. Several typical cloud images are showed in Fig.2, and several typical snow images are showed in Fig.3.

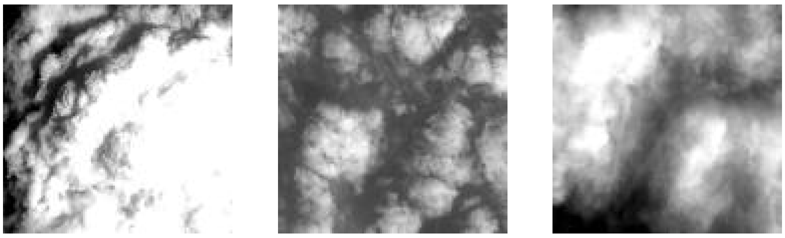


Fig.2: Typical cloud images

Fig.3: Typical Snow images

From Fig2, we know that texture feathers of clouds are random and variable, but they are different from those of objects on the earth and still self-similar at different scales which means the shape of a part is similar to that of the whole. The gray value of clouds is much greater than that of objects and also relatively uniform with a range of gray scale. In addition, the edge of the cloud area is usually fuzzier and the gradient changes slowly. From Fig3, we know that snow cover on the ground, due to the terrain (artificial objects, mountains, vegetation and so on), the edge of snow is usually sharp and gradient changes quickly (Wenxia Yu, 2006).

Therefore, the texture feature may be considered in the cloud and snow identification. Typical statistically analytical methods of texture consist of Gray level co-occurrence matrix, fractal dimension. Laws energy and so on (Shuhua Wang, 2002). Kittler first proposed the texture feature to detect clouds on the sea in 1985 (Kittler J, 1985). Rees studied the fractal dimension of snow and ice with Landsat-TM images, and the result proved the fractal dimension could well describe the surface of snow and ice (Rees, 1992). Wenxia Yu proposed that cloud and snow can be identified by calculating their fractal dimension values owing to the difference of irregularity (Wenxia Yu, 2006). Xianghang Liu interactively sheared the cloud/snow area which was already recognized and then calculated fractal dimension values of sub-images to identify cloud and snow (Xianghang Liu, 2006). However, limited by the texture is difficult and complex to describe, as well as other factors, the method based on texture features is still far from automation and high precision.

On the basis of training a large number of experimental samples，considering the statistical distribution of fractal dimension values of cloud and snow images, as well as the proportion of cloud or snow in the image, an automatic method is proposed in this paper to identify cloud and snow in panchromatic images or single image of visible band.

2. METHODS AND EQUATION

2.1 The calculation of fractal dimension

Since the fractal theory was founded in the seventh decade of the last century by a French mathematician named Mandelbrot, it was generally described as " the shape of a part is similar to that of the whole” (Fernández E, 2001), without a strict definition. Many things in nature are consistent with the characteristic of fractal, and the fractal dimension value is a statistical description of the texture (Shuhua Wang, 2002). Therefore, the fractal dimension can be used as an effective parameter to describe the texture feature of images. Common fractal dimension contains Similar dimension, Hausdorff dimension, Box Dimension, Information Dimension, Correlation dimension and so on. In this paper, Box dimension proposed by Sarkar and Chaudhari (Sarkar N, 1992; Chaudhuri B, 1995) which is relatively simple, practical and widely used is adopted.

The remote sensing image is a projection of the three-dimensional space in the two-dimensional surface, so the structure of a remote sensing image can be reflected by the spatial variation of a gray image. For the gray-scale image, we can consider the two-dimensional image as a surface (x, y, f(x, y)) in three-dimensional space, and f(x, y) is the gray value of the pixel at position (x, y). The variation of gray value reflects the roughness of the surface. The fractal dimension is calculated by using different scales to measure the surface. Firstly, we separate the M×M image into many r×r grids in accordance with scale r. On each grid, there is a line of r×r×h boxes. Symbol h presents the height of single box, and h:

$$h = \frac{r}{M} \times G \tag{1}$$

Given the grid (i, j), suppose that the minimum gray value is in box K and the maximum gray value is in box L. The minimum amount of boxes that can cover all the gray values in grid is:

$$n_r(i, j) = L - K + 1 \tag{2}$$

Then the amount of boxes that can cover all the images can be calculated by the following formula:

$$N_r = \sum_{i,j} n_r(i, j) \tag{3}$$

Fractal dimension value is:

$$D = \lim \frac{\log(N_r)}{\log(1/r)} \tag{4}$$

We need to define a series of r values, and then draw a scatter diagram with log(1/r) as abscissa, and log(Nr) as longitudinal coordinates. By simulating the sample points with a least square method, the slope of the line which is the value of the fractal dimension is got.

Fractal dimension values reflect the complexity of the texture feature. The greater the value is, the more complex the surface is, and vice versa. In the remote sensing images, the ground target has more texture details and snow covers over the ground, so its fractal dimension value is also greater.

2.2 The calculation of proportion of cloud or snow

To calculate the proportion of cloud and snow, the first thing is to detect the cloud and snow from the background utilizing the higher reflectivity of clouds and snow. We can use the method of threshold segmentation such as Histogram method, iterative method and Otsu method (Lu Sun, 2008). Considering the information of remote sensing images is rich and the texture is varied, an iterative method is used to detect cloud and snow. The iterative method means assume a threshold value in the initial condition and use a program to automatically update the assumed threshold. Thereby, the best segmentation threshold is obtained. The main algorithm is as follows:

1) Count the minimum gray value of the image $R_{\min}$ and the maximum gray value $R_{\max}$, and then the initial threshold value is got:

$$T_0 = \frac{R_{\min} + R_{\max}}{2} \tag{5}$$

2) Divide the image into two parts as the target and the background according to the threshold value, and then calculate the mean value of the two parts respectively:

$$R_0 = \frac{\sum_{R(i,j)<T_K} R(i,j) \times N(i,j)}{\sum_{R(i,j)<T_K} N(i,j)} \tag{6}$$

$$R_G = \frac{\sum_{R(i,j)>T_K} R(i,j) \times N(i,j)}{\sum_{R(i,j)>T_K} N(i,j)} \quad (7)$$

Here, R_0 is the average gray value of targets, R_G is the average gray value of the background, $R_{(i,j)}$ is the value of the point (i,j), $N_{(i,j)}$ is the weight coefficient of the point (i,j), generally is the number of (i,j), and T_K is the threshold.

3) Choose a new threshold value T_{K+1}, the new value is as follow:

$$T_{K+1} = \frac{R_0 + R_G}{2} \quad (8)$$

4) Cycle the second step to the third step until $T_K = T_{K+1}$, and then the best segmentation threshold to divide the image is obtained

We use the best threshold to segregate the image, and then compute the distribution of cloud or snow in the image. In traditional, spectrum threshold method processes the images pixel by pixel, but texture analysis method processes the images by N×N sub-image. So some modifications should be made to unify the processing unit. In this case, the concept of cloud threshold percent is put forward that presents the percentage of the cloud pixels in an N×N sub-image(Na Shan,2009). According to the cloud threshold percentage, the algorithm judges the sub-image into different categories: cloud, snow or the other and then static all the results of sub-images to get the distribution of cloud or snow in the image.

2.3 The discriminant mechanism of cloud and snow utilizing the characteristics of texture and distribution

According to the analysis above, fractal dimension values reflect the complexity of texture features. The ground target in the remote sensing images has more texture details and snow covers over the ground, so its fractal dimension value is also greater, for the cloud contains less texture details and its gray value changes smoothly. By training a large amount of experimental samples, we find the fractal dimension values of clouds are in the interval [m, n], but only part of fractal dimension values of snow are in the range. Considering the distribution characteristics of fractal dimension values of clouds and snow, as well as that clouds appear occasionally in time and space, we suppose a method as following to identify cloud and snow.

Firstly, use the iterative method to get the threshold(T) which distinguishes cloud/snow from the background, and then divide the input image into lots of sub-images. Secondly, calculate the threshold percentage of each sub-image and statistic the ones whose threshold percentage is greater than x%, assume the number is p. Hence, A is got by the formula: p divided by n. Thirdly, calculate fractal dimension values of each sub-image and statistic the ones whose fractal dimension is in the range of m and n, assume the number is p. Hence, B s got by the formula: q divided by n.

If it is a cloud image, when the cloud covers part of the image, the fractal dimension values of those cloud samples must be in range m and n, and some of the others may be in the range of m and n, or none in the range. Therefore, p is less or equal q， scilicet， A is less than or equal B. When the cloud covers the whole image, A is equal B and equal 1. Therefore, A is less or equal B when it is a cloud image. If it is a snow image, generally, snow covers the whole image, so A is nearly to 1. At this time, only some of the fractal dimension values of sub-images are in the range of m and n, with others out. Therefore, p is greater than q，scilicet，A is greater than B when it is a snow image.

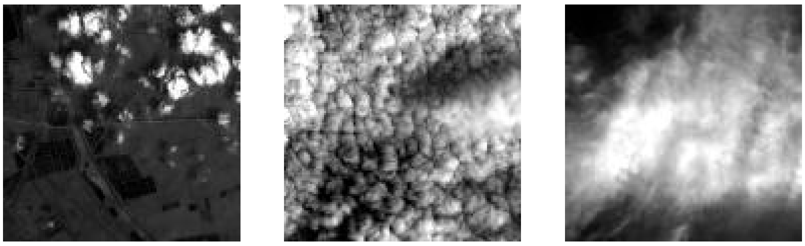
On the other hand, A is greater than B means that the number(p) of ones whose threshold percentage is greater than x% is greater than the number(q) of ones whose fractal dimension values is in the range of m and n, and also means only some of the ones whose threshold percentage is greater than x% are in the range of m and n, so the image is a snow image. A is equal B means that the number of whose fractal dimension values are in the range of m and n is also equal p, so the image is a cloud image. A is less than B means that the fractal dimension values of additional sub-images are in the range of m and n, so it is a cloud image.

Therefore, by comparing the proportion to the distribution, the automatic identification of cloud and snow was realized. If A is less than or equal B, the image is a cloud image. If A is greater than B, the image is a snow cloud.

3. EXPERIENCE &RESULTS

3.1 Data for experience

We choose data from high resolution panchromatic images acquired by “Beijing-1”, which theoretical spatial resolution is 4m. Texture analysis method processes the images usually by N×N sub-image. We choose 4596 cloud sub-images and 3065 snow sub-images in size of 64×64 as the samples to get the distribution of the fractal dimension values. In addition, choose 20 cloud images and 8 snow images in size of 2048×2048 as testing samples to test and verify the method proposed in the paper. Testing samples contain several kinds of cloud images (Cumulus, cumulonimbus, Cirrocumulus and strigiform) and snow images covering different objects (residential area, mountain, vegetation and so on). Fig4(a), 4(b), 4(c) are test images of cloud and Fig4(d), 4(e), 4(f) are test images of snow.



(a)Image of cloud 1 (b) Image of cloud 2 (c) Image of cloud 3

(d) Image of snow 1 (e) Image of snow 2 (f) Image of snow 3

Fig.4: Cloud and snow images acquired by "Beijing-1" Satellite

3.2 Training of experimental samples

Train those 4596 cloud sub-images and 3065 snow sub-images we choose and the training results are showed in Fig. 5 and Fig. 6.

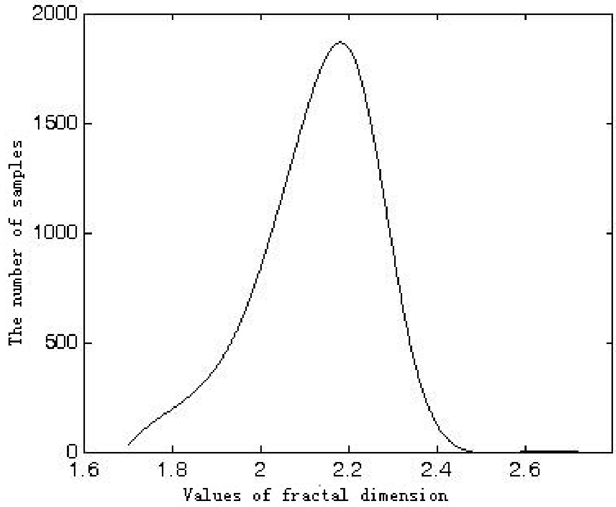


Fig.5: The distribution of the fractal dimension values of cloud images

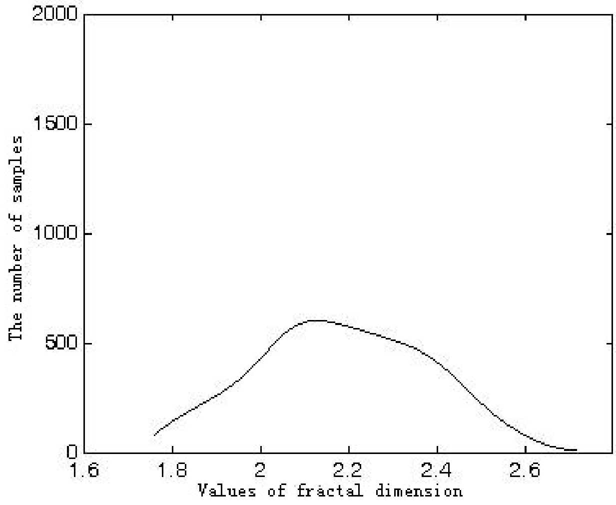


Fig.6: The distribution of the fractal dimension values of snow images

From Fig5 and Fig6, we can see that fractal dimension values of cloud sub-images are relatively concentrated, about 96.2% are in range of 1.8802 and 2.3381. In contrast, fractal dimension values of snow sub-images are in range of 1.7592 and 2.7152 and relatively dispersive. The training results indicate that values of clouds are mostly in the range of 1.8802 and 2.3381, only 58% of values of snow are in that range. In practice, the range of the fractal dimension value changes as the image data acquired by satellite increase.

3.3 Results and Analysis

Process the remote sensing images using the algorithm proposed in subsection 2.3 in this paper. In the experiment, the threshold percentage is assigned to be 60% which is an experience value. Table 1 records the testing results.

Table 1 Results of the automatic recognition of cloud and snow

	Threshold	Value of A	Value of B	Discrimination results	Actual results
Fig 4(a)	103	0.0600	0.9053	Cloud	Cloud
Fig 4(b)	138	0.6318	0.9219	Cloud	Cloud
Fig 4(c)	133	0.5713	0.9190	Cloud	Cloud
Fig 4(d)	128	0.7373	0.3193	Snow	Snow
Fig 4(e)	127	0.7548	0.2031	Snow	Snow
Fig 4(f)	118	0.7012	0.8701	Cloud	Snow

A is less than B in Fig 4(a), 4(b), 4(c), 4(f), so those images are judged to be cloud images. Among them, Fig 4(a), 4(b), 4(c) are consistent with the actual results, Fig 4(f) is an error of judgment. A is greater than B in Fig 4(d), 4(e), so those images are judged to be snow images and are consistent with the actual results. Test with all images and statist results, only Fig 4(f) is a miscarriage of justice. The critical evaluation criteria for the accuracy of the algorithm is the false alarming rate when the cloud is misidentified as the snow or the snow is misidentified as the cloud. The experimental results illustrate that the false alarming rate is 3.57%. False detection results from flat snow surface. The flat surface has few details, and its fractal dimension value is about 2.2. Above all, in most cases, the algorithm automatically classifies the cloud and the snow efficiently and accurately.

4. CONCLUSIONS & RECOMMENDATIONS

Spectral and texture features are general features of cloud and snow images. In this paper, on the basis of studying and analyzing the spectral and texture features of cloud and snow panchromatic images, an automatic method is proposed to identify cloud and snow. We test and verify the algorithm using remote sensing images acquired by “Beijing-1”, and the result shows a good accuracy. Though false detection exists because of the thickness of clouds or snow, the complexity of the surface of objects and the limitation of fractal theory, the result in this paper illustrates that the automatic identification which makes full use of the texture feature of images could achieve a good result. In addition, the method could be also applied for other high resolution panchromatic images because of the universality of the texture feature.

ACKNOWLEDGEMENTS

The author thanks the support of intelligent observation and verification of small Remote sensing satellite (2011BAH23B01) and Nestling Eagle Foundation of Academy of Opto-Electronics, Chinese Academy of Sciences.)

5. REFERENCES:

References from Journals:

Wei Fenghua, Li Caixing, Zhaxi Yangzong, 2007. Snow Monitoring based on CBERS Data. Remote Sensing for Land & Resources, 3, pp.31-35.

Zheng Xiaobo, Kang Weimin, Tian Pengju, 2005. Application of EOS/MODIS Remote Sensing to Snow Cover in Guizhou. Guizhou Agricultural Sciences, 33(3),pp.60-62.

Yin Qinjun,Yang Yinglian, Xu Weixin,2002. Research on Distinguishing between Cloud and Snow with NOAA Images. Plateau Meteorology, 21(5), pp.526-528.

Yu Wenxia, Cao Xiaoguang, Xu Lin,2006. Automatic Cloud Detection for Remote Sensing. Chinese Journal of Scientific Instrument, 6, pp.2184-2186.

Wang Shuhua, Zhao Yuming, Zhou Xiaosi, Zhou Hongmei,2002.Automatic classification of cloud and mist based on weighted fractal dimension. Infrared and Laser Engineering,31(1),pp.18-22.

Kittler J, Pairman D.Contextual,1985. Pattern Recognition Applied to Cloud Detection and Identification. IEEE Transactions on Geoscience and Remote Sensing, 23(11), pp.855-863.

Rees W G,1992.Measurement of the Fractal Dimension of Ice-sheet Surfaces using Landsat Data.International Journal of Remote Sensing,13(4), pp.663-671.

Fernández E, Jelinek HF, 2001. Use of Fractal Theory in Neuroscience: Methods, Advantages, and Potential Problems. Methods, 24(4), pp.309-321.

Sarkar N, Chaudhuri B B, 1992. An Efficient Approach to Estimate Fractal Dimension in Texture Image. Pattern Recognition, 25(9), pp.1035-1041.

Chaudhuri B B, Sarkar N, 1995. Texture segmentation using Fractal Dimension. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(1), pp.72-77.

Jiang Dong, Wang Jianhua, 2000. The Fractal Thought in Remote Sensing Information Science. Journal of Gansu Sciences, 12(1), pp.53-57.

Shan Na, Zheng Tianyao, Wang zhensong, 2009. High-speed and High-accuracy Algorithm for Cloud Detection and its Application. Journal of Remote Sensing, 13(6), pp.1147-1155.

Sun Lu, Bi Duyan, 2008. The Improved Iterative Algorithm based on Entropy for Image Threshold Segmentation. Computer Applications and Software, 25(10), pp.225-226.

References from Books:

Liu Xianghang, Cao Xiaoguang, Yu Wenxia. Cloud Detection based on Fractal Dimension//Annual Meeting of China Association of Remote Sensing Application in 2006, Taiyuan, 2006.