A New Automation Process of Clouds Removal for Hyper-Spectrum Images using Self-adaptive Frequency Homomorphism Filter

Xing LIU，Lixue LIU

Faculty of Surveying and mapping, Anhui University of Science and Technology, Huainan, Anhui, 232001, China

***Abstract* - The traditional homomorphism filter deals with the whole remote sensing image which damages the cloudless part when removes thin clouds, and the threshold of filter is needed to change to fit different bands. A proposed method aiming to solve these problems is introduced on the basis of geological statistic theory, by this method the cloudless part can be detected automatically and thereby reserved, also the threshold of filter is calculated to suit any band. The method is developed by IDL, the experiment result shows that the computer memory consumed becomes less; the speed is rapider than traditional way with less manual intervention.**

Keywords - Self-adaptive, Frequency homomorphism filter, Clouds Removal, Hyper-Spectrum Image

I. INTRODUCTION

In the most time, the clouds in remote sensing images are often regarded as a type of noise, which bring problems for images interpretation, analysis and application[1].The study of clouds removal is a key step in image processing because it should have improved the image quality for better application[2,3].There are some ways widely used in practice such as: replacing the clouds covering area by a new clear part[4], this way is out of hand if the new part is unachieved obviously. Some studies try to decrease the clouds by band transformation [5], but the useful information is also damaged at the same time. A number of researchers recover the clouds covering area by interpolation method based on geological statistic [6], small clouds area can be recovered effectively but it is reasonless when the clouds area become large, the interpolated result is usually unreal.

At present, homomorphism filter [7] is widely used to remove or weaken those large area clouds covering images, which can be executed at space domain or frequency domain. The frequency filter needs Fourier transform for the whole image primarily, which consumes too much computer's memory when the image size is big, another shortage of this way is that the part of cloudless area is also damaged contemporary because the image is filtered wholly[8]. When the homomorphism filter is carried out in space domain by a spatial filter, the computer memory spent becomes low, but the useful image parts are still damaged, and a new problem arises because the threshold of the filter for different band is varied, the thresholds should be changed to fit different bands one by one, the reason is that the transmission and transparency of air varies at any spectrum window, resulting in the clouds distribution different at every band [9].

A new way to resolve the above disadvantages is proposed in this paper. By this way, the frequency filtering is carried on in a sliding window, the size of window is determined on the basis of geological statistic; the filter threshold is calculated at different band automatically. The computer memory consumed falls and the cloudless part of image is reserved after processing.

II. THE LAW OF HOMOMORPHISM FILTER

An image can be represented by a 2-D function $f(x, y)$，which is often looked as the product of clouds incident ray $fi(x, y)$ and ground reflection ray $fr(x, y)$, the simplified formula is:

$$f(x, y) = fi(x, y) \times fr(x, y) \quad (1)$$

The component $fi(x, y)$ is bout the distribution of clouds, $fr(x, y)$ is related to the ground material distribution. The aim of clouds removal is to weaken $fi(x, y)$ and tone up $fr(x, y)$, but the popular filter is not suitable to remove the $fi(x, y)$ component [10].

The first step of homomorphism filter is to computer the logarithm of formula (1) :

$$\ln f(x, y) = \ln fi(x, y) + \ln fr(x, y) \quad (2)$$

After this step, the relation of incident ray and the reflection ray is translated from multiplicative to additive, then transforming the image into frequency domain by Fast Fourier Transformation (formula 3):

$$F(u,v) = \bar{f}\left|\ln f(x,y)\right| = \bar{f}\left|\ln fi(x,y)\right| + \bar{f}\left|\ln fr(x,y)\right| = I(u,v) + R(u,v) \quad (3)$$

Because clouds reflection varies slowly while ground material reflection varies quickly, this implies that the

DOI 10.5013/IJSSST.a.17.49.56 ISSN: 1473-804x online, 1473-8031 print

component of clouds accounts for low frequency yet the ground material accounts for high frequency in the frequency domain. Appling a high pass $H(u,v)$ filter on the $F(u,v)$ image can pass high frequency part and cut low frequency component, then bloom ground detail and weaken clouds incident ray:

$$G(u,v) = I(u,v)H(u,v) + R(u,v)H(u,v) \tag{4}$$

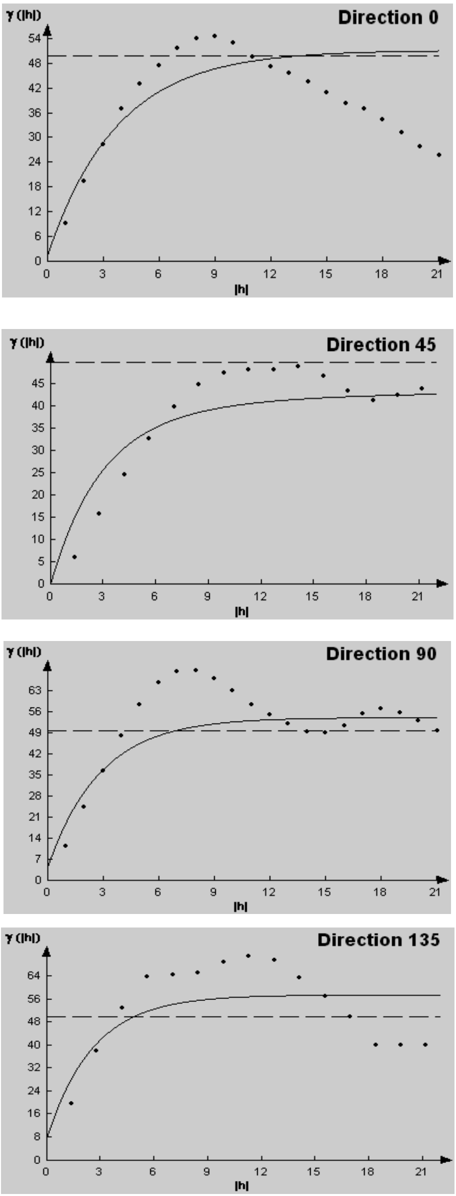

Figure 1 . Variation distance of an image at 4 directions (dot line is practical data, real line is theoretic line, x-axis is pixel distance, y-axis is variation distance)

Transforming the filtered image $G(u,v)$ back into space domain by inverse Fast Fourier Transformation (formula 5):

$$\ln f(x,y) = \bar{f}^{-1}\{G(u,v)\} = \bar{f}^{-1}\{I(u,v)H(u,v)\} + \bar{f}^{-1}\{R(u,v)H(u,v)\} \tag{5}$$

And then calculate the natural exponential value of filtered image $\ln f(x,y)$:

$$G(x,y) = \exp[\ln f(x,y)] = \exp[\bar{f}^{-1}\{I(u,v)H(u,v)\} + \bar{f}^{-1}\{R(u,v)H(u,v)\}] \tag{6}$$

By this route, the Flourier Transformations of the whole image consumes large time and memory [1], and cloudless part of image is damaged simultaneity because the way does not distinguish cloudless area from the cloud area [13], moreover, the above route brings edge effect which leads the filtered images hardly to interpret [8].

III. IDENTIFICATION OF CLOUDS

Clouds distribution determination is an important step. The original image pixel value should be transformed into 0-255 range by grey stretch. In succession, calculate the mean value and variance of each sliding window which stands for the centre pixel neighborhoods, and find out the minimum mean value and maximum variance from all the sliding windows for any band. After getting the minimum mean value and maximum variance of each band, calculate the absolute difference (D-value) between every pixel's mean value and and the maximum mean value, and calculate the absolute difference (D-value) between the pixel's variance and the minimum variance. Obviously, the two kinds of D-value are small if the pixel is in cloud area. Adding the two D-value together to construct a "cloud" criterion (scare) to determine clouds existing or not, the smaller the scare the thicker the clouds. The cloud thickness is classified into 10 classes according to the minimum scare and maximum scare by the following formula:

$$scale(x,y) = Class\{(scare(x,y) - \min\_scare) \times 10 / (\max\_scare - \min\_scare)\} \tag{7}$$

$scale$ represents the cloud thickness, $\min\_scare$ and $\max\_scare$ are the thinnest cloud scare and thickest cloud scare separately, $scare$ becomes small when the cloud becomes thick while $scale$ is big when cloud is thin.

There are some other high light ground material showing the similar characters as clouds such as glacier and snow, the previous studies point out that their pixels value is bigger than cloud's and usually over 200, and this paper declares that the removal method stated above is useless for too thick cloud, glacier or snow unless replacing those area by a new clear image.

DOI 10.5013/IJSSST.a.17.49.56 ISSN: 1473-804x online, 1473-8031 print

IV. THE SIZE OF FILTERING WINDOW DETERMINATION

Previous studies conclude that almost any spatial data is relying on in the space, which is called space correlation [15], also means that some geosciences phenomena or attributions depend on the neighboring value [14].The variogram map of an image shows that the pixel's values are often auto-correlated within certain distance at different orient (Fig.1), and the relationship disappears beyard the distance. By the principle, it is inferred that the homomorphism filter can be carried out in the window (the variable distance) rather than in the whole band [16].By this method, the memory space for FFT is saved, it is in favor of big data processing.

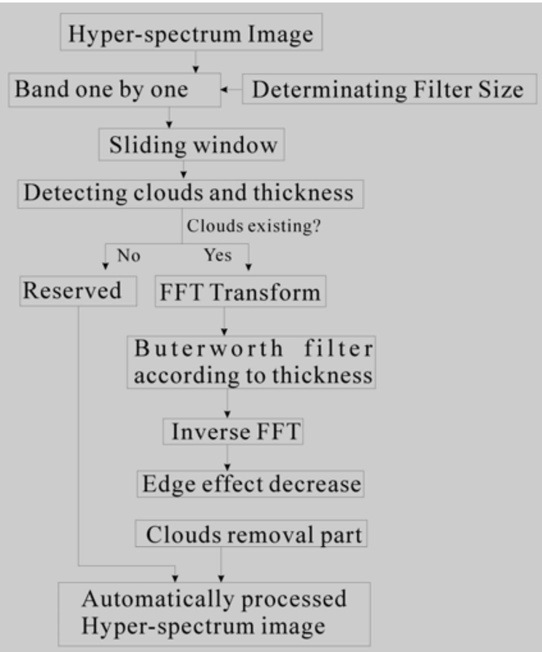

Figure 2. The automatic process chart of hyper-spectrum image clouds removal

The size of window (range) for filter is determined step by step. Calculate the variable distance of a bigger square block

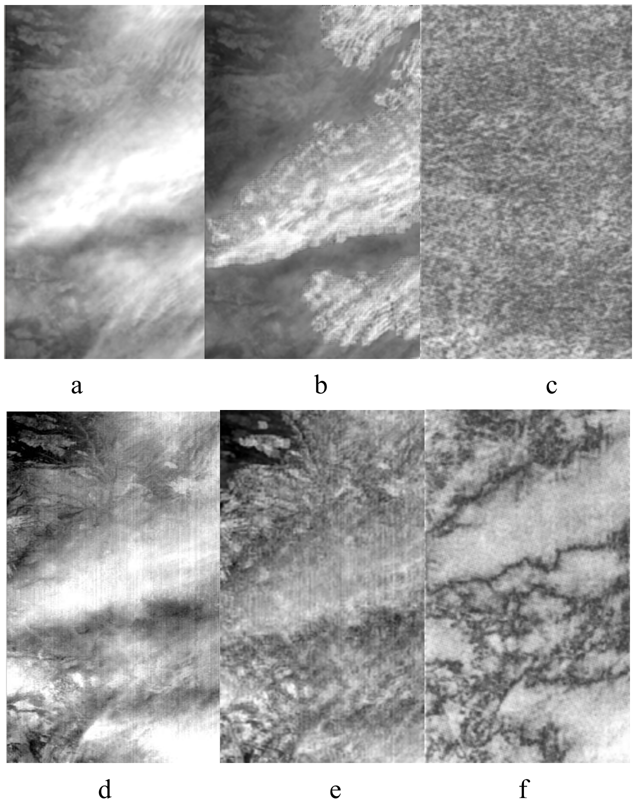

Figure 3. Contrast between an original and a clouds removal image by this study

（a,d are band 28 and band 183 of original image separately, b,e are band 28 and band 183 after cloud removal by the proposed method, c,f are band 28 and band 183 after cloud removal by the traditional method）

cut from a whole image, if the distance is less than the cut block, cut a smaller one until the variable distance equals to the block size. In this study area, the variable distances are about 10 pixels at 4 directions (Fig.1), implying that the high frequency and low frequency components are correlated in 10 pixels, so the size of window for homomorphism filter is 10 pixels.

V. ARITHMETIC IMPROVEMENT

The ideal method [13] of clouds removal is to delineate the clouds above all, and then weaken them partially by a suitable homomorphism filter so the cloudless image part can be reserved. We just do it.

A. *Adjustment of Cut-Off Frequency for Filter*

The clouds component occupies low frequency and the ground features occupies high frequency from the above analysis, the goal of clouds removal is to weaken low frequency and highlight high frequency component. Usually, a Butterworth high pass filter is adopted:

$$H(u,v) = \frac{1}{1+[\frac{D(u,v)}{D_0}]^{2n}} \quad (8)$$

In the formula, D_0 is the cut-off frequency (threshold), which is often fixed for the whole band in previous research. But D_0 should be changed to fit different clouds thickness, which is the reasonable result, and D_0 is direct ratio at clouds thickness. The formula (9) is suggested in this paper to calculate the threshold:

$$D_0 = scale \times 2 \quad (9)$$

the $scale$ of clouds is from 1 to 10, and the value of cut-off frequency D_0 is between 2～20.It is implied that D_0 is bigger when the clouds is thicker. With this variable D_0, the clouds of different thickness can be weakened respectively, it seems more reasonable than previous way. And, k is set to 0.141 or 1, n power of filter, usually set to 1.

B. The Workflow of New Method

- The pixel value of an original remote sensing image are adjusted to ground reflectivity after radiometric correction, and the pixel values are stretched from 0 to 255[17].
- Cutting a square block from the adjusted image, computing the range at four directions of 0°,45°,90°,135°,choosing a biggest range from them as the sliding size of filter window. If the range is greater than the cut square block, trying to cut a larger one until the variation distance equals to the cut image.
- Using the proposed way in part 3 to delineate clouds area and calculate its thickness. Usually, a sliding window with its mean value greater than 60 and its variance scare less than 0.68 is regarded as cloud covering zone.
- The left process is operated in the sliding window recognized as cloud zone. Above all, converting the pixels value of the window to a new matrix by logarithm function, and then carrying out Fourier transform to get its frequency distribution, using a high pass Butterworth filter to process the frequency image, then converting the filtered frequency image back into spatial domain through inverse FFT, computing its natural exponential value of the processed matrix at last.

C. Dealing with Edge Effect

The pixel values at the adjacency of cloud and cloudless area vary seriously, the directly homomorphism filter application will cause the processed result hardly uninterpretable.

A suitable strategy which can reduce the edge effect is suggested in the following. All the pixel values of a filtered sliding window decrease proportionally to its cloud thickness, the thicker the cloud, the more the pixel values decrease. The result brings out a grey balanced image for users to interpret. The above explanation is executed through formula (10):

$$fw(x,y) = \frac{scale \times (filtered\_win(x,y) - \min l)}{\max l - \min l} + \min o - scale \quad (10)$$

$fw(x,y)$ is the last image of a filtered sliding window, $scale$ is the value of cloud thickness of this window, $filtered\_win(x,y)$ is the pixel value after Butterworth high pass filter and inverse FFT, $\max o, \min o$ mean max value and min value of original image without any processing, $\max l, \min l$ are max value and min value of $filtered\_win(x,y)$. The edge effect weakens [8] and the result keeps the same contrast to no-processing area (Fig .3).

D. Dealing with the Multi-Band Image

The clouds distribution and related thickness is different at any two bands because the transmission of air window varies at different spectrum wavelength. So, the pixel values of each band should be stretched into the same range. The experiment shows that a sliding window can be recognized as cloud area if its mean value is greater than 60 and variance scare is less than 0.68 no matter which band. Fig.2 shows the whole flow chart.

VI. DISCUSSION OF EXPERIMENTAL RESULTS

By the previous way, the memory of computer needed for FFT is about $O(N^2)$ in case of an $N \times N$ image, there are $10 \times N$ assistant memory needed in addition [8],if an image is about 100M size, the computer memory needed exceeds 800M, the work is impossible to accomplish if an image reaching 1G by existing computer. But now the memory space falls at about $O(\gamma^2)$(γ is variation distance, about 10 pixels in this study), the time spent is about $O(\gamma^2) \times num\_win$ (num_win is the numbers of sliding window over image).This is a big image favorable way.

The experiment images are EO-1 satellite Hyperion Level1 data, covering Pulang County, Yunnan province, China with 196 effective bands after removing some overlaying spectrum. The radiant flux density is converted to ground reflectivity by radioactive correction model, the size of the image is 275×2500 pixels, there are some clouds distributing on the image centre (Fig.3a, d), the image is processed by the introduced method above and by the traditional way respectively. The result figure (Fig.3, b, e) shows that the new way not only can improve the computer efficiency but also can weaken the clouds when reserving cloudless area, with less edge effect and abundant texture. All those advantages are helpful to improve image quality and widen its application. Any more, the traditional way is focus on the one band image, but with our new method, the clouds of any band from the hyper-remote sensing image can be detected and then removed.

Contrast between the Fig3b and Fig3c (also between Fig 3e and Fig3f) shows good detail and texture by the suggested method whenever compared to the original image (Fig3a, d) or to the image processed by traditional method (Fig3, c, f).

The edge effect by Fourier transformation is weakened in the study but still existing, so the improvement of formula (10) deserves further study.

Fig.3 indicates that if the cloud is too thick the processed result may be imperfect; the suggestion is that for too thick clouds area any other methods are invalid unless replacing them by a new part.

The algorithm was developed by IDL language, which is an outstanding platform for matrix operation with powerful functions for image and signal processing by RSI company of USA, the platform gives user the focus only on data processing and let the data reading and saving of mass data to system itself.

Many functions were employed from ENVI dynamic library in this research, so the codes must be executed under ENVI environment.

VII. CONCLUSIONS

This paper deduces a route of adaptive Butterworth filter for thin clouds weakening. The method determines the correlative size(distance) by geological statistics at first, and filters the cloud frequency by an adaptive high pass filter in a sliding window one by one, the way consumes less memory space and littler time than previous way, brings big data processing easily. It is an automatic process for hyper-image cloud removal. The reason is that the new method can detect clouds area and its thickness, and adjusts the cut-off frequency of filter according to clouds thickness automatically. Without giving threshold manually, this method is suitable for any band of a hyper-spectrum image. The result by the suggested method shows good clods removal effect (Fig3b, e), the edge effect is lower than old way, the texture and details are remarkable in the clouds area without damaging cloudless area. But for those too thick clouds area there is no existing idea unless replacing them by a new piece of image over there because there is no reflection information from the surface.

The edge effect by Fourier transformation is weakened in the study but still exists, so the improvement of formula (10) deserves further study.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who helped me during the writing and publishing of this thesis. This thesis is sponsored by NSFC (NO.: 41372368), NSFC of Anhui(NO.:1408085MD761).

REFERENCES

[1] Shen Wenbin,Zhou XinZhi, "The algorithm of cloud removal of remote sensing image based on homomorphism filtering", Remote sensing of land and resources.,vol.22,no.1, 2010,pp.45-47.

[2] Song Xiaoning,Zhao Yingshi, "Cloud Detection and Analysis of MODIS Image",. Journal of Image and Graphics, vol.8A,no.9,2003,pp.1079- 1083.

[3] Liang S , Fang H , Chen M. "Atmospheric Correction of Landsat ETM+ Land Surface Imagery : I. Methods", IEEE Transactions on Geosciences and Remote Sensing .,vol,39,2001,pp.2490-2498.

[4] Wang Hui, Tan Bing, etal.,"The processing of Technology of removing Clouds Image Based on the Multi-resource RS Image", Journal of the Pla institute of surveying and mapping.,vol,18,no.3,2001,pp.195-198.

[5] Wang run, Liu Hongbing, Gong Rui, "A Method of Removal Cloud of Multi-spectral Satellite Image", Jisuanji Yu Xiandaihua., vol.6,2005,pp.13-15.

[6] Yu Yue, Gu Hua, Sun Weidong, "Spectral unmixing based image restoration for optical remote sensing images covered by thin cloud" , Journal of Image and Graphics. , Vol.15 , No.11,(2010),pp.1670-1680.

[7] Wu Ping, "Discussion on the method of think cloud removal by ENVI", Xin Jiang non-ferrous metal., 2005, pp.29-33.

[8] Xie Huamei,Li Rong yan, et al, "The Removing Clouds Method Based on Large Remote Sensing Image", Journal of Beijing Normal University (Natural Science).,vol.42,no.1,2006,pp.42-46.

[9] Lai Lifang,Yu Le, Zhang Hankui, "Improving the Quality of Optical Remote Sensing Images Based on Multi-fractal Technique" , Bulletin of Surveying and mapping.,vol.12,2010,pp.49-50.

[10] Zhao Zhongming,Zhu Chongguang, "Approach to Removing Cloud Cover from Satellite Image", Remote Sensing of Enviroment China.,vol.11,no.3,1996,pp.195-199.

[11] Wu Weilu, "Eliminating Cloud from Remote Sensing Images", Railway aerial survey.,vol.1,2003,pp.6-7.

[12] Yan Lijuan, "A method of thin cloud removal for remote sensing image based on wavelets", Mine Surveying.,2013,pp.61-65.

[13] Xie Hua mei,He Qi,pao,Zheng Nin,Bie Rong,fang, "The Improved Homomorphic Filter Algorithm for Removing Cloud of Remote Sensing Image Based on The Second Exploiture of Erdas Tool", Journal of Beijing Normal University (Natural Science).,vol.41,no.2,2005,pp.150-153.

[14] Shen Xiaole,Shao Zhenfen,Yan Beibei, "A Novel Dodging Algorithm for Remote Sensing Images Affected by thin Clouds", Geomatics and Information Science of Wuhan University., vol.38,no.5,2013,pp.543-547.

[15] Goodchild M F., "Spatial Autocorrelation", Norwich: GeoBooks,1986.

[16] Wu Gang, Chen Hui qiong, et al, "Multi-scale Edge Detection Based on Image Variation Function", Journal of HeFei University of Technology., vol.22,no.2,2000,pp.227-230.

[17] Li Haiwei, "Research on Removing Thin Cloud in Single Remote Sensing Image", A dissert of Centre and South university for master, 2012.