THE EFFICACY OF SEMI-AUTOMATIC CLASSIFICATION RESULT BY USING DIFFERENT CLOUD DETECTION AND DIMINUTION METHOD

C. Y. Hau [a, *], C. H. Liu [b], T. Y. Chou [a], L. S. Yang [a]

[a] Dept. of Land management, Feng Chia University, Wen-Hwa Rd., Taichung, Taiwan, China
[b] Graduate Institute of Civil and Hydraulic Engineering, Feng Chia University, Wen-Hwa Rd., Taichung, Taiwan, China-(liuhenry, st_how, jimmy)@gis.tw, lsyang@fcu.edu.tw

Commission VI, WG VIII/3

KEY WORDS: Image processing, Change detection, Image understanding, Matching, Air quality

ABSTRACT:

Passive Remote Sensing is multi-spectral, rapid and high resolution, but it is very easily influenced by intended to the atmosphere conditions, such as fog and haze. In the past, there have been many studies removed the noise of cloud and haze. This study proposes to use the differences between five filtering methods to judge the differences between their efficiencies. Fast Fourier transforms are used in all five methods to remove clouds. All five methods apply the high-pass filter concept. Standard reference data are used as a basis for companion. The results show that there is a trade-off between classification and appearance; high amounts of information smoothing worsen classification accuracy but improve appearance.

1. INTRODUCTION

1.1 General Instructions

Clouds and haze obstruct satellite imaging; they are the main source of noise in remote sensing. Remote sensing has become the essential information collection technique in environmental monitoring, resource investigation and other research. Their shortcoming is that minor atmospheric obstructions can easily cause major noise, even in multi-spectral, high-resolution images.

Cloud cover is a major source of noise in Asia and hard to resolve completely. Clouds reduce the inherent information of images; this has serious impact on monitoring. In this paper we review recent cloud removal research developed with fast Fourier transform filters, computer vision, and image processing. We can assume that sunlight was completely reflected when any given image was taken. Land surfaces lose information when clouds and haze reflect solar radiation and increase gray-level value.

The present research uses images filter and multi-data methods. We review earlier research which did not use these methods, include (1) the Dark-target approach method(e.g. Teillet, 1995; Chaze, 1988; Chaze, 1989); (2)the Histogram-match method(e.g. Richter, 1996; Artamonov, 1999); (3)the Liang algorithm method(e.g. Liang, 2001; Liang, 2002); (4)the HOT method(e.g. Zhang et al., 2002); (5)the Imagery filter method(e.g. Liu, 1984; Zhao et al., 1996) and (6)the Multi-data fusion method(e.g. Du, 2002; Wang, 1999). Each methods has its own advantage and shortcoming.

Current methods, including the abovementioned methods, use multispectral images and consider absorption and transmissivity characteristics. Two way to reduce cloud cover (in images) are the mosaic method and operation reducing. The mosaic method was time series to reduce atmosphere influence (McClain et al. 1985) operation reducing reduces atmospheric influence by file operations and image transformation (Lai et al 2004). While the mosaic method is unsuitable for this study, operation reducing can cause distortions. Because operation reducing process all the information of an image, operation reducing can distort alter, or eliminate any information in the image.

Dong et al., 2005; Li et al., 2006; Liu et al., 2008; Chen et al., 2007 have contributed to pre-processing and to processing, but they have given methods to evaluate the differences between images.

1.2 Study area and data

Satellite images with different wavelength and characteristic are affected by clouds in different ways. Different scanners and detectors also find slight variations. The Formosat II satellite obtained images on August 8, 2004 and September 29,2004 have been used in this study. The August 8, 2004 FS II images has three levels of cloudiness: cloud free, thin cloud (haze) and thick cloud three levels. The September 29,2004 FS II image, is cloud free and is used as reference data. The first image is subjected to filtration and image classification process to compare with its elfe before and after the filtration. the second image as the reference to determine the training sits and the check-points data. The two images belongs to the same paddy-work periods, so we can assume the land cover of this two images are very same alike..

The study area is located at Taichung county, area is about 110 hectare. The land cover types include bare land, river, grass land, paddy field, forest, road andurban land. The upper left

Corresponding author. Tel.: +886-424 516 669 # 110; faz: +886 424 519 278
e-mail adress: st_how@gis.tw (C. Y. Hau)

corner has coordinates (202809, 2688210); the lower right has coordinates (205521, 2684938). The study area images are shown as Figure 1..

The Formosat II (FS-II) was launched on May 21 2004; it entered a Sun-synchronous orbit at an altitude of 891 kilometres. It passes over Taiwan twice daily to collect earth-science information. The spatial resolutions of its Panchromatic (Pan) and Multi-Spectral (MS) layers are 2 m and 8 m, respectively. The wavelengths of the bands were 0.45～0.52μm，0.52～0.60μm，0.63～0.69μm，0.76～0.90μm.

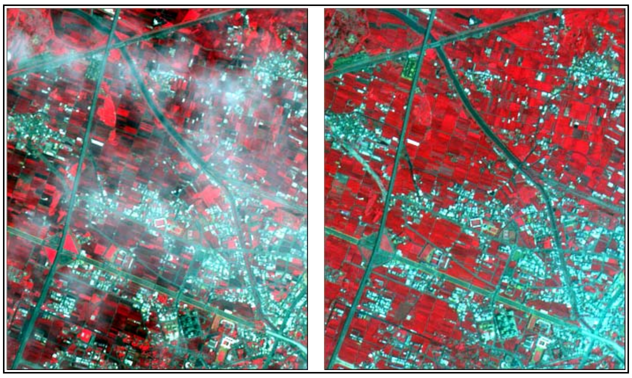

Figure 1. The image of study area. The Left side is the major image(08/20/04), and the right side is the reference image (09/29/04).

2. METHOD DESCRIPTION

Because these two images have similar spatial resolutions, times of observation and so on, one can assume they have identical land-cover. Under the assumption, the differences between these images are caused by atmospheric conditions such as the solar angle of incidence, solar intensity and cloud cover. The atmospheric conditions have the same effect in the whole image, so they will not influence the classification results unless cloud and haze interfere. Any two images of similar regions at comparably close times will produce highly similar classification results: the only differences are produced by cloud and haze.

The processing of a group of images begins with geocorrection to solve any geometric distortion form height, speed and sensor error of sensor. After that we classify the cloud and haze boundaries according to a histogram from each band. Filtrate the haze area and utilize the k-mean classification to evaluate and quantify the results.

2.1 Image pre-proceeding

Any satellite image produces both systematic and unsystematic error when it receives reflection from a surface. Geometric correction must use reference images to correct cloudy images. The process leaves a quantifiable residual error in the output image. In order to keep residual error as small as is practical, this research chooses Ground Control Points (GCPs) within the images that conform distinct characteristic. The Cubic Convolution is method used in our research to re-sample the image and to put identical feature into the same positions.

2.2 Cloud and Haze boundary detect

Because the thickness of cloud and haze in an image is not fixed, but gradually varies from the cloudiest region to the clearest region, different processing modes should be used in different regions (Chen et al., 2007). Various types of clouds present different reflection and transmission characteristics. One might extract a thick cloud boundary easily because solar radiation has reflected completely, but the surface information within that thick cloud might be difficult to extract; past methods have excised them (Liu et al., 2008). Systematic methods can filter haze and enhance surface information for haze region because there is a correlation between the penetrating solar radiation and surface reflection, the surface features themselves indistinctly display in an indistinct condition.

In isolation, either one of the addendum of filteration methods is not only very useful. However, separate processing for image region with different characteristics can produce better result. Therefore, we use the statistics from images to separate clean areas, haze areas, and cloudy areas. More details are shown in Equation [1].

$$f(x,y) = \begin{cases} I(x,y) < I_{mean} \in 0 \\ I_{mean} < I(x,y) < I_{mean+std.} \in 1 \\ I(x,y) > I_{mean+std} \in 2 \end{cases} \quad (1)$$

where $I(x,y)$ = cloud image value
I_{mean} = cloud image mean value
$I_{mean+std.}$ = cloud image mean+standard deviation value

Thick cloud areas totally reflect all spectrum information, and cover land surface with masses of clouds. Information is totally lost; in the past, the mosaic method was the major mean to remove clouds from images. By contrast, this study applies filtration and reclassification to thin cloud areas.

2.3 Haze filtration

We take the filtering method which presented by Lai et al., 2004 as an example in this research. This homomorphic filtering method uses as Fourier transform. The result is shown as Figure 2. During the homomorphic filtering, the results are strongly influenced by frequency, illumination and reflection,.Frequency has the greatest influence on the filtering results.

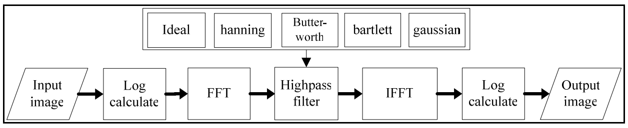

Figure 2. The Homomorphic Filtering process with different filters.

A fast Fourier transformation (FFT) is used to convert a raster image from the spatial domain into a frequency domain image. The FFT calculation converts the image into a series of two-dimensional sine waves of various frequencies. An analyst can edit the Fourier image to reduce noise or to remove periodic features (ERDAS Field Guide, 2005). A Fourier transform is a kind of orthogonal transformation. Moreover, it can enhance the image and extract characteristics simply and effectively in the frequency domain. Therefore, Fourier transforms have been applied to remote sensing. Fourier transforms have contributed to many kinds of product filters, such as High-Pass Filter, Low-Pass Filter and median filter. In this research, we take enhancement filter in the frequency domain from product theory and we use Fourier transforms as our foundation. Then, removing cloud noise effectively, we use low noise image production to obtain enhanced images.

Image transformation is the basis of image process technique. For effective analysis and processing, we need to transform the image into another domain, which can be used for easy processing, then transform it back to the spatial domain for our end user. When surfaces are covered by thin clouds, satellite sensors receive the solar radiation reflected by cloud and the planetary surface. The concept is shown by equation [2] (L.S. Kon 2004).

$$f(x,y) = [L, R(x,y)] = \alpha L R(x,y)\, N(x,y) + L(1-N(x,y)) \qquad (2)$$

Where f(x,y)= the sensor received image
R(x,y)= reflection from surface
N(x,y)=transmittance of cloud
L=solar intensity
α=atmosphere transformation attenuation coefficient
the α, N(x,y) and R(x,y) are between 0 to 1

The equation above can be simplified into equation [3], in which the image value can be regarded as the product of illumination and reflection.

$$f(x,y) = f_i(x,y)\, f_r(x,y) \qquad (3)$$

Where f(x,y)= the image received by the sensor
$f_i(x,y)$= illumination
$f_r(x,y)$= reflection

The image calculation process is a non-linear transform combined with a high-pass filter for illumination. The efficiency of filtration in a thin cloud area is decided by the filter function $H(u,v)$ and radius D_0. And then the procedure calculate and filter band by band with different radius, that is can get the various achievement of thin cloud filtration and follow the ringing artificial noise, Therefore, we have to consider not only the effect of reducing the cloud noise, but also increase of the fuzzy in satellite image. Fourier processing, five window functions are provided to achieve different types of attenuation: Ideal, Bartlett (triangular), Butterworth, Gaussian and Hanning (cosine).

2.4 Image aftertreatment

After image filtration, the gray value of image is distributed from 0 to 5 with which the format is float. For necessary of mosaic with cloud free and thick cloud area images, normalized the float to unsigned 8 bits and change value distribution from 0 to 255. The equation of normalized as shown as equation [4].

$$f(x,y) = \frac{X - X_{min}}{X_{min} - X_{max}} \times 255 \qquad (4)$$

High-pass filtering using the window with different function: All frequencies inside a circle of a radius D_0 are completely attenuated, and all frequencies outside the radius are retained completely (passed). The original data information can cause vanish during the process. For minimize this effect, we distinguished cloud into three different levels, one is cloud free area, another is thin cloud area and the other is thick cloud area. The cloud free area will extract from original image to keep information, the thick cloud area will extract from too, for the thin cloud area, and we exchange the filtration achievement into the original image to minimize the effect. And the process shown into figure3.

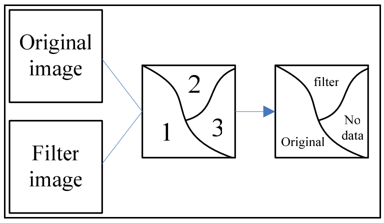

Figure 3. The image re-component process

2.5 Accuracy Comparison

In the past, most researches use the subjective vision asserting, but it doesn't have a effective quantization or assessment for achievement of cloud process. Although this way seems to obtain better image quality, sometimes the color of image is changed by the result of monitor display. Therefore, if we want to discover the image whether still had using value or recovery degree after cloud process, we have to measure the image by a reliable and objective method.

So, in this research, we use k-mean method in cloud process to prove the change vale of precision. However, the result of classification in Maximum likelihood method is easily influenced by the selecting of training site and the difference of random check point distribution. In order to analyze the changing value of cloud in two images, first we take reference image as an example to select the training site. Then, according to the resolution of image and the situation of land covered, we select the four kinds of AOI, Field & Veg., Roadway, Body of Water, Artificial area, as training pattern. Finally, after classification, we can produce 271 random check points and check them by using aerial photography and surface real data

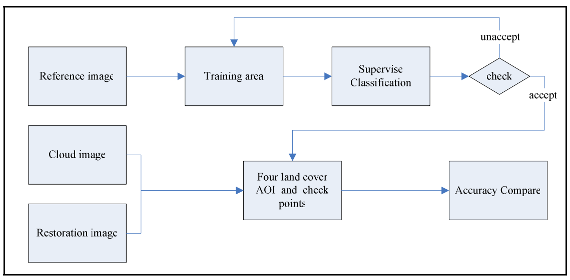

Figure 4. the procedure at assess the different filter methods by k-mean classification.

3. RESULT AND DISCUSSION

In the imaging stage, there are many indefinite factors like the systematic and non-systematic distortion (T.Y. Chou, L.S. Yan, 2006) needed to resolve. In our study, we set the projection of image in Transverse Mercator and the reference ellipse spheroid in GRS67, then selects 30 GCPs artificially according to the reference images. Re-projection by 3-order of polynomial, and the total RMS value is lower than 0.3362.

The description of experiment achievement as cloud body graduation achievement, thin cloud filtration achievement, mosaic image achievement and accuracy assessment comparison. The results are as follow.

3.1 Clouds class to individual category

The material of this research is FSII multi-spectral image, there are four bands in this image. The reflection situation of haze is various in each bands, may carry on the fog graduation from related research learning using the grey level statistics way, and based on the cloud distribution range general characteristics, is smaller than the area approximately 640 square meters (10 pixel) the sub-area regards as the miscellaneous news eliminates it, reduces in the image by this way the reflected value high surface object to the cloud stereoscope other influence.

Figure 5. shows that the influence by cloud in band 4 is much slight than other bands obviously because of the wave length is longer and penetrability of fog is stronger, causes to unable to obtain the good achievement on boundary extraction by using the threshold value. And this result is different from other bands, so, the band 4 does not include in any cloud processing in this study.

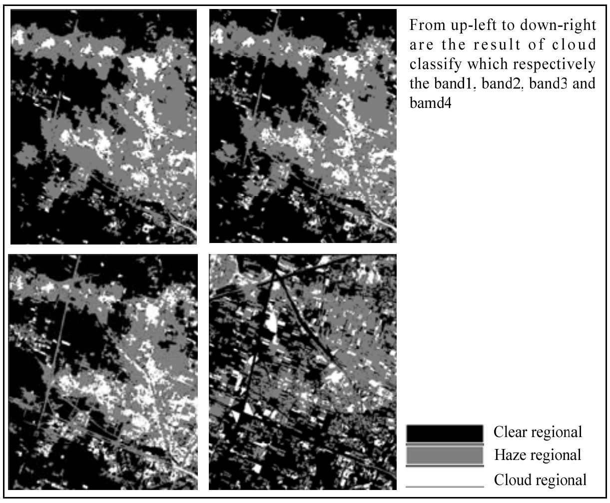

Figure 5. Different clouds type class result in each band.

3.2 Result of haze filtrate

Uses fast Fourier to transform the image to the frequency domain carries on the cloud image filter processing, and uses five different filter methods to carry on the thin cloud filtration separately, namely filter as ideal, Bartlett, Butterworth, Gaussian and Hanning, to red, green, the blue three kind of visible light wave band's image carry on the filter separately, uses the thin cloud scope the grey value change situation to discuss the results.

Because image filter processing will have the unnecessary value disturbance to the cloud free region, Causes image application in extraction information is not able to promote effectively while cloud filtration completely(e.g. effectively., Liu et al., 2008; Chen et al., 2007; Zhao et al., 1996). Therefore, discusses and compare the situation of histogram changes between original and completely the filtration image can get the filter achievement and efficiency. Because of fog may increase the mean value of image (Wang et al. 1999), The image grey value change may regard as the result of elimination fog disturbance. As the result, the Bartlett filter's achievement all has the obvious value decrease in each bands (the result as show as figure4); the achievement of other filter does not appear this situation.

The result of filtration for each filter as shown as figure 5, it shows that a fine achievement with the vision perception during the thin cloud area. The Bartlett and Ideal filter can get better and obvious filtration achievement than Butterworth and Gaussian filter during the thin cloud area. Although from this study can get much better achievement during thin cloud area, there is still increase the noise in the thick cloud area except Butterworth filter.

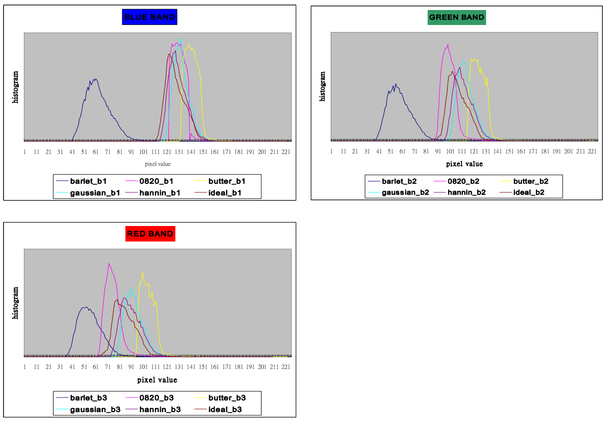

Figure 6. the histogram change in different filter methods which in haze regional. And the sequence is Blue band, Green band and Red band.

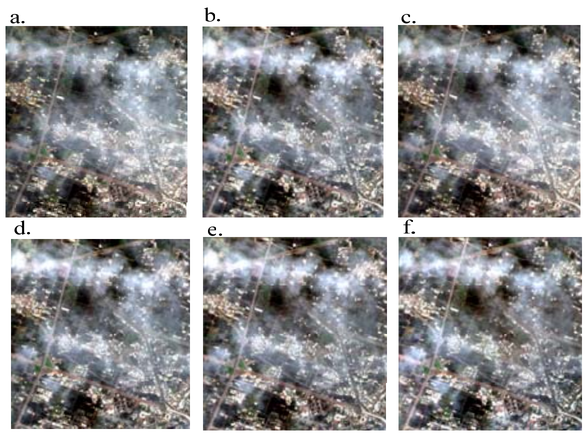

Figure 7. The filter results compare. And the image a is the original image、b is the Bartlett、c is the Butterworth、d is the Gaussian、e is the Hanning、f is the Ideal, all images band component shown by band3, band2, band1.

3.3 Accuracy compare of K-mean classification

We compare the supervise classification accuracy reports in this section to evaluate the cloud diminution results that process by above methods. The results of serious works describe as follow.

3.3.1 Training area and check points set: The supervised classification accuracy is affected by the training data selection and individual experience. In order to solve this problem, first step is that select the training sample data from the reference image which was shot at Sep. 29 2004, classifier the image and check classification accuracy, the checking point is selected by random.

The process is continuing and iteration till the overall accuracy is over 80%. Record the position of sample point and significant information by each land cover types as the standard reference sample data as shown as Table 1. the transformed divergence (TD) as the index to check the area of interesting (AOI) sample data whether cause the layer mixed up or not. Generally, the indicate lager than 1700 represent that the AOI sample date will be able to bring the good classified results. The TD index calculated results as shown as Table 2.

	The AOI amount of each class	total count(area)
Field & Veg.	25	1560
Road	15	752
Water	10	296
Artificiality	30	1316

Table 1 the amount and area of each category's training area

Signature Name	Artificiality	Field & Veg.	Road	Water
Field & Veg.	2000	0	2000	2000
Road	1688	2000	0	2000
Water	2000	2000	2000	0
Artificiality	0	2000	1688	2000

Table 2. Signature separability

Classification results compare: Because the filtration results involve to the whole image including the cloud-free area. For reduce the influence, we combine the cloud-free, haze and thick cloud information that according to the each boundary of cloud-coverage levels that we obtained in above section, and the new images are used to classify.

Table 3. shows the accuracy of supervised classification between referenced image, original image and five component images. From these results, we find that image filter methods can't promote the classify accuracy effectively. Through the visual observation, the obstructions of haze are reduced obviously. The lower accuracy of classify because the inner information and data construction that are compressed during the filter process is not enough to support the classification and cause omission and commission.

Compared the reports from image classification accuracy, we can find the achievement of Butterworth filter is better than other filters, and the results of Ideal, Bartllet, Hanning and Gaussian filter are alike. The values of user accuracy, producer accuracy, overall accuracy and kappa coefficient as shown as table 3.

	0929	0820	Bartlett	Butterworth	Gaussian	Hanning	Ideal
users accuracy	83.37%	78.20%	68.14%	67.60%	64.09%	64.59%	64.93%
producers accuracy	83.40%	72.49%	53.60%	70.07%	53.29%	53.83%	56.37%
overall accuracy	80.44%	74.91%	63.47%	68.63%	64.94%	65.31%	64.58%
kappa	0.7017	0.6145	0.4455	0.5103	0.4648	0.4686	0.4547

Table 3. The classification accuracy of image, respectively are the reference image, original image and five images from above methods.

4. CONCLUSIONS

The results of cloud filtration with five kinds of different filters have obtained a good result in the vision, but it actually reduces the accuracy of classification. And that is the reason that the cloud filtration is unable to be used generally. This research also uses the traditional supervised classification with reference images to conform the sample points and GCPs then evaluates the results of filtration through standardized classification method.

In this study, a comparative method is proposed to evaluate the cloud filtration in view of application. The methodology of the filtration is not changed. The clouds are processed in different ways based on the thickness of the clouds. In the concept, it is agreeable to common sense. The accuracy of classification does not be promoted after finishing cloud filtration because the methods of the cloud filtration destroy the original pixel values of the images extensively. Because the filters remove the noise in the images systematically and the cloudy noise distributes over the images randomly, the effect of filters is limited. It is the reason that the accuracy is lower than before instead of increase.

REFERENCES

Artamonov, E. S., K. T. Protasov, 1999. Restoration of fragments of satellite images of the Rarth's underlying surface shadowed by haze and clouds, Proc. of SPIE Sixth International Symposium on Atmospheric and Ocean Optics, vol.2983, pp.170-178.

Chaez, P. S., 1988. An improved Dark-object subtraction technique for atmosphere scattering correction of multispectral data, Remote Sensing of Environment , vol.24, No.3, pp. 459-478.

Chaez, P. S., 1989. Radiometric calibration of Landsat thematic and mapper multispectral images, Photorammetric Engineering and Remote Sensing, vol.55, No.9, pp. 1285-1294.

Chen, F., D. Yan, Z Zhao, 2007. Haze Detection and Removal in Remote Sensing Images Based on Undecimated Wavelet Transform. Geomatics and information science of Wuhan University, 32(1), pp. 71-74.

Chen, L. Y., J. Chen, C. X. Liu, X. Cao, 2006. An effective approach to remove cloud-fog cover and enhance remote sensing imagery. Journal of Cengdu university of technology Science & Technology Edition), 33(1), pp. 58-63.

Dong, Y. F., G.Q. Sun, Y. Pang, A. M. Fu, 2005. A comparative study of some atmospheric correction methods, Remote Sensing for Land & Resources, vol.2, pp. 16-20

Du, Y., B. Guidon, J. Cihlar, 2002. Haze detection and removal in high resolution satellite image with wavelet analysis, IEEE Trans. Geoscoence and Remote Sensing, vol.40, No.1, pp. 210-217.

Kon, L. S., 2004. Image Proceeding, Sciencept, Beijing, pp. 248-252.

Lai, K. Y., C. Y. Liu, S. C. Ku, 2004. The Remote Sensing cloud eliminates method based on ERDAS IMAGINE, Jiangxi normal university city and environmental science institute, vol.2

Leica Geosystem, 2005. ERDAS IMAGINE Field guide, Lica Geosystem, vol.6, pp. 189-202

Li, Y. C., J. Chen, C. X. Liu, X. Cao, 2006. An effective approach to remove cloud-fog cover and enhance remote sensing imagery, Journal of Chengdu University of Technology (Science & Technology Edition), vol.33, No.1, pp. 58-63.

Liang, S. L., H. Fand, M. Chen, 2001. Atmospheric correction of Landsat ETM+ Land surface imagery-Part I: Methods, IEEE Trans. Geoscience and Remote Sensing, vol.39, No.11, pp.2490-2498.

Liang, S. L., H. Fand, M. Chen, 2002. Atmospheric correction of Landsat ETM+ Land surface imagery-Part II: Validation and Applications, IEEE Trans. Geoscience and Remote Sensing, vol.40, No.12, pp.2736-2746.

Liu, Z. K., B. R. Hunt, 1984. A new approach to removing cloud cover from satellite imagery, CVIP, vol.25, pp. 252-256

Liu, C. H., M. L. Yeh, T. Y. Chou, L. S. Yang, 2008. The qualitative analyses of cloud cover on Optical Satellite image, Geoinformatrics 2008 , accepted.

Yang, L. S., T. Y. Chou, J. C. Lei, 2006. The application of theories and analysis in remote sensing, Kings information, pp.4-16~4-23

McMillin, L. M., D. S. Crosby, 1984.Theory and validation of the multiple window sea surface temperature technique. J. Geophys. Res., 89C, pp. 3655–3661.

Wang, B., A. Ono, K. Muramatsu, N. Fujiwara, 1999. Automated detection and removal of clouds and their shadows from Landsat TM images. IEICE Trans. INF. & SYST., vol. E82-D, No.2, pp. 453-460.

Richer, R., 1996. A spatially adaptive fast atmospheric correction algorithm, Int. J. Remote Sensing, vol.17, pp. 1201-1214.

Teillet, P. M. and G. Fedosejevs, 1995. On the dark target approach to atmospheric correction of remotely sensed data, Canadian Journal of Remote Sensing, vol.21, No.4, pp. 374-387.

Zhang, Y., B. Guidon, J. Cihlar, 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images, Remote Sensing of Environment, vol41, No.5, pp. 1082-1089.

Zhu, S., B. Zhu, C. Sha, Z. Zheng, 2005. Cloud Removal of IKONOS image using auxiliary image, Geomatic Science and Engineering, vol.24, No.4, pp. 24-28.