APPLICATION OF FOURIER PHASE-ONLY CORRELATION IN SATELLITE IMAGE REGISTRATION WITH DEM SHADING IMAGE

Hiroyuki Saito, Associate Professor
Akiyoshi Hosoi, Graduate Student
Graduate School of Science and Technology
Hirosaki University
Hirosaki, 036-8561, JAPAN
saitoh@cc.hirosaki-u.ac.jp, h09GS522@stu.hirosaki-u.ac.jp

ABSTRACT

There is an optimization method for the geometric correction of Landsat images, which uses a simulated direct solar irradiance image made from a digital elevation model (DEM) and a solar direction to determine parameters of the geometric transfer equations. The optimization method is quite simple, and also allows us high accuracy geometric correction in sub-pixel order. However, promising in promotion of satellite image use for extraction of geographical information as well as for topographic correction, the optimization method still requires some improvements. In this research, we are focusing on the algorithm of optimization method for the positional difference of scene centers between given in the system information and in actual. In the conventional method, the offset of satellite image that originated in the positional difference of scene centers is calculated by two-dimensional cross-correlation analysis with the DEM shading image. The maximum value of the correlation coefficients is searched by a method based on a steepest descent method, thus the solution tends to hang on to local minima when the beginning value is not set appropriately. To solve the problem, we propose a new optimization method using the Fourier phase-only correlation (POC). The POC function requires no beginning value, and its maximum of the correlation coefficients is a sharp peak, thus it may be easy to find the positional difference of scene centers. In this paper, we have applied the proposed method to the system corrected Landsat TM images to investigate the features of the new method comparing to the conventional method.

Key words: Geometric correction, Landsat TM, DEM shading image, Fourier phase-only correlation.

INTRODUCTION

Remote sensing images obtained by satellites provide us appropriate information used for update of land use map and multi-temporal monitoring of natural environment changes over wide area. In order to extract the information appropriately from satellite images, it is necessary to transform individual satellite images into orthorectified imagery because satellite images contain various types of geometrical distortion from tilt, topography, and curvature on the earth surface.

Several geometrical correction methods have been proposed, and all they can be categorized into two types: one is the direct method in which a rigorous solution of the spatial relationship between a satellite image and its conjugate ground point is calculated, and the other is the indirect method in which an image coordinate is modeled as functions of the map coordinate (Zhou *et al*., 2002). Both methods generally need to collect enough number of ground control points (GCPs), which is labor-intensive work.

Satellite image users in the application fields tend to prefer the indirect method. The following reasons may be considered: The affine transform in the indirect method is easy in handling for them, and they usually use the system-corrected satellite images (e.g., level 1B or 1G data) which contain geometric parameters for the affine transform such as the scene center positions both in the map and the image coordinates, the orientation angle and the scaling factor of the image. The geometric error amounts to several hundred meters when we apply the indirect method to satellite images such as system-corrected Landsat TM images with the given system parameters. The geometric error is mainly caused by inaccuracy of the orbit data. Owing to the geometric calibration efforts as well as the improvement of orbit data, recent satellite images such as Landsat ETM+ images show high geometric accuracy in flat areas. However, in case of rugged areas, there still remains significant topological relief displacement at both ends of an image, where the image point with a large off-nadir distance is shifted due to scan line perspective. Though Pala and Pons (1995) have developed an analytical geometric procedure for deriving the relief displacement due to scan line perspective, it is difficult to remove the relief effect through the system correction or the normal geocoding with polynomial transformation without elevations, which is supposed to have high geometrical accuracy.

One of the promising ways to remove the relief effect is a use of the digital elevation map (DEM). Iikura (2002) has modified the approach of Pala and Pons (1995), and proposed a practical relief displacement correction method for system-corrected Landsat TM images with DEMs which correspond to the satellite images. The relief displacement was successfully removed with elevations in DEM files. Horn and Bachman (1978) have proposed a method to register Landsat MSS images with DEM shading images (or simulated direct solar irradiance images). Their method required no GCPs because the DEM shading image was a reference. However, their method was not recognized well in those days because there was no authorized DEM. Iikura (2002) has proposed an optimization method for ortho-rectified Landsat TM images with DEM shading images. In order to remove the translations of the ortho-rectified Landsat TM images, which originated in the location error of the scene center, he adopted the two-dimensional cross-correlation technique, and searched the maximum of the cross-correlation function with a sub-pixel precision using the AMOEBA method (Press *et al*., 1989). However, the calculation was extremely time consuming, and the solution hanged on to local minima frequently when the threshold and the initial translation ware not set appropriately, as the AMOEBA method is a technique based on the steepest descent method.

The cross-correlation technique is the most common approach. However, the cross-correlation function with an ortho-rectified Landsat TM image and a DEM shading image has significant values even at rather large as well as rather a broad maximum, as shown in Fig. 1(a); therefore, it is difficult to locate the maximum. One of the alternatives to the cross-correlation function is the Fourier phase-only correlation (POC) function (Horner and Gianino, 1984; Chen et al., 1994; Takita *et al*., 2003). The POC function yields an even sharp maximum at the best match point, as shown in Fig. 1(b); therefore, it is easy to locate the maximum.

In this paper, we are focusing on the improvement of the optimization method by Iikura (2002). We use the POC technique instead of the cross-correlation technique to remove the technical issues of the conventional method. The new method has potentials to be not only a rectification method with high-accuracy and high-efficiency but also an automatic rectification method. In the following sections, the procedure of the new method is described. Then, the features of the new method are investigated with system-corrected Landsat TM images, comparing to the conventional method.

ORTHO-RECTIFICATION PROCEDURE

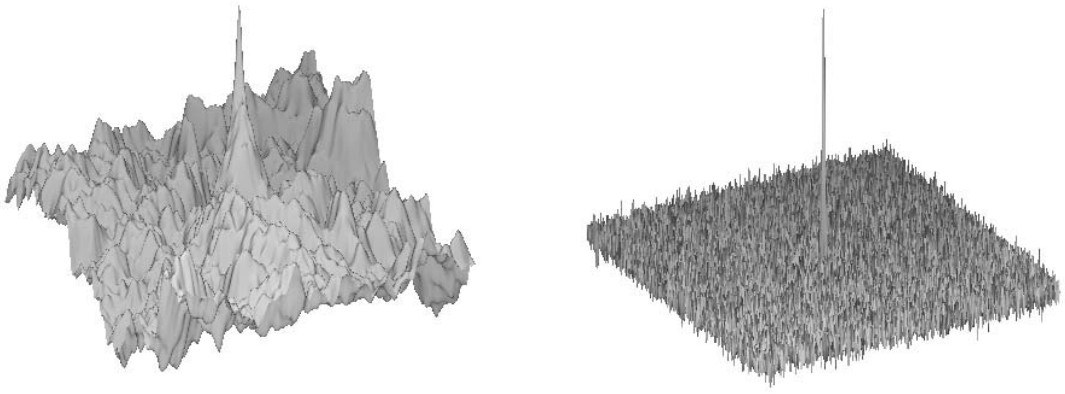
Relief displacement correction and geometric transfer equation with DEM

The displacement due to relief on a curved earth surface caused by scan line perspective is schematically illustrated in Fig. 2. The relief displacement of a target point is estimated using following four trigonometric equations (Pala and Pons, 1995):

$$SP^2 = R^2 + (R+H)^2 - 2R(R+H)\cos\alpha\,, \tag{1}$$

$$R/\sin\theta = SP/\sin\alpha\,, \tag{2}$$

Figure 1. Three-dimensional plots of the cross-correlation function and the POC function applied to an



ortho-rectified Landsat TM image (band 5) and a DEM shading image. The amplitudes are normalized by each maximum. (a) The cross-correlation function. (b) The POC function.

$$(R+z)/\sin(\theta+\delta\theta)=(R+H)/\sin(\alpha+\theta+\delta\theta), \quad (3)$$

$$R/\sin(\theta+\delta\theta)=(R+H)/\sin(\alpha+\delta\alpha+\theta+\delta\theta), \quad (4)$$

where SP is the off-nadir distance between the satellite and the target, R is the earth radius, H is the flight height of the satellite, z is the elevation of the target, θ is the off-nadir angle of the target, α is the angle between pointing vectors of the satellite and the target from the earth center, and $\delta\theta$ and $\delta\alpha$ are the angular displacements due to the elevation of the target. The angular displacement $\delta\alpha$ is given by

$$\delta\alpha=\sin^{-1}\left(\frac{R}{(R+H)\sin(\theta+\delta\theta)}\right)-\alpha-\theta-\delta\theta. \quad (5)$$

Thus, approximating the arc PP_2 as a rectilinear segment, the relief displacement δD is given by

$$\delta D = R\delta\alpha. \quad (6)$$

In the case of georeferenced images where the relief displacements occur only in the pixel direction (or a scan direction), the geometric transfer equation for ortho-rectification can be written as (Iikura, 2002)

$$\begin{aligned} p-p_0 &= \left((x-x_0)\cos\omega_0-(y-y_0)\sin\omega_0\right)/\Delta+\delta D(x,\,y)/\Delta, \\ l-l_0 &= \left(-(x-x_0)\sin\omega_0-(y-y_0)\cos\omega_0\right)/\Delta, \end{aligned} \quad (7)$$

where $(p,\,l)$ and $(x,\,y)$ are the satellite and the map image coordinates, $(x_0,\,y_0)$ is the scene center of the map corresponding to the satellite image's coordinate $(p_0,\,l_0)$, Δ is the spatial resolution of the satellite image, and ω_0 is the orientation angle of the system-corrected satellite image slanting to the axis of map coordinate system. These parameters are given as the system parameters in the system-corrected image.

Optimization of geometrical correction with DEM shading images

Comparing an ortho-rectified satellite image obtained by eq. (7) to the corresponding map carefully, a translation exists between them. The translation is mainly caused by positional error of the scene center given in the system information. As the DEM shading image resembles the ortho-rectified satellite image in appearance, it can be used as a reference to evaluate the translation (Horn and Bachman, 1978). The DEM shading image is a first degree as a function of how directly the sun's rays are incident upon the slope, that is, a cosine function of the solar incident angle (Ekstrand, 1996; Levin *et al.*, 2004):

$$\cos\beta=\cos\theta\cos e-\sin\theta\,\sin e\cos(\phi-A), \quad (8)$$

where β is the incident angle between the surface normal and the solar beam, θ is the solar zenith angle, e is the surface normal zenith angle, ϕ is the solar azimuth angle, and A is the aspect of the slope angle. The system information of the system corrected satellite image contains the two solar angles also. The topographic parameters, such as the aspect of the slope angle and the surface normal zenith angle, are calculated with the DEMs which correspond to the satellite images.

In order to evaluate the translation, or to determine the best scene center position, we have adopted the Fourier phase-only correlation (POC) technique. The POC function is given by the inverse Fourier transform of the function (Chen et al., 1994):

$$Q(u,\,v)=\frac{S(u,\,v)R^*(u,\,v)}{\left|S(u,\,v)\right|\left|R^*(u,\,v)\right|}, \quad (9)$$

where $S(u,\,v)$ and $R(u,\,v)$ are the Fourier spectra of the ortho-rectified satellite image and the DEM shading image, respectively, and $R^*(u,\,v)$ is the complex conjugate of $R(u,\,v)$.

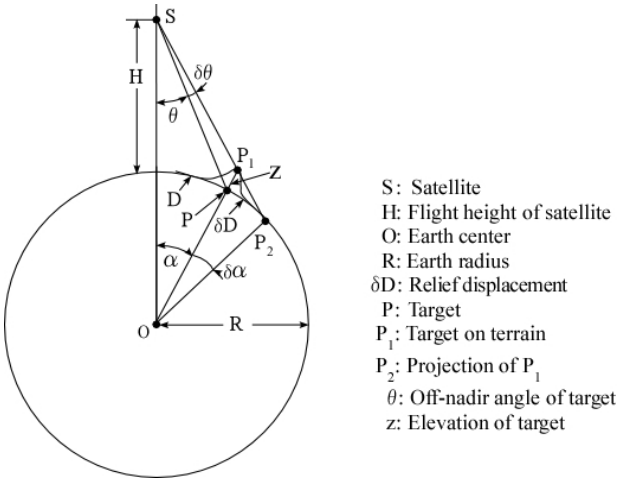


Figure 2. A schematic diagram of the displacement due to relief on a curved earth surface caused by scan line perspective.

The image registration flow for the ortho-rectification incorporating the POC technique consists of the following steps:

1) prepare the system collected satellite image and the corresponding DEM,
2) produce the DEM shading image using the DEM with the two solar angles (θ and ϕ) ,
3) transform the system corrected satellite image into the ortho-rectified image using eq. (7) with (x_0, y_0), ω_0 , Δ , and δD ,
4) calculate the POC function between the ortho-rectified image and the DEM shading image,
5) locate the position of maximum in the POC function, that is, determine the translation $(\delta x, \delta y)$,
6) if the translation $(\delta x, \delta y)$ exceeds some pre-set threshold, update (x_0, y_0) with $(x_0 + \delta x, y_0 + \delta y)$, and return to step 3.

As the calculated POC function is discrete in the step 4, we have fitted a quadratic function along the x and y coordinates of the POC function respectively to locale the maximum with a sub-pixel precision. The image registration is iterated to reach the pre-set threshold of translation, thus the final output of translation is defined as the sum of translations obtained in the iteration.

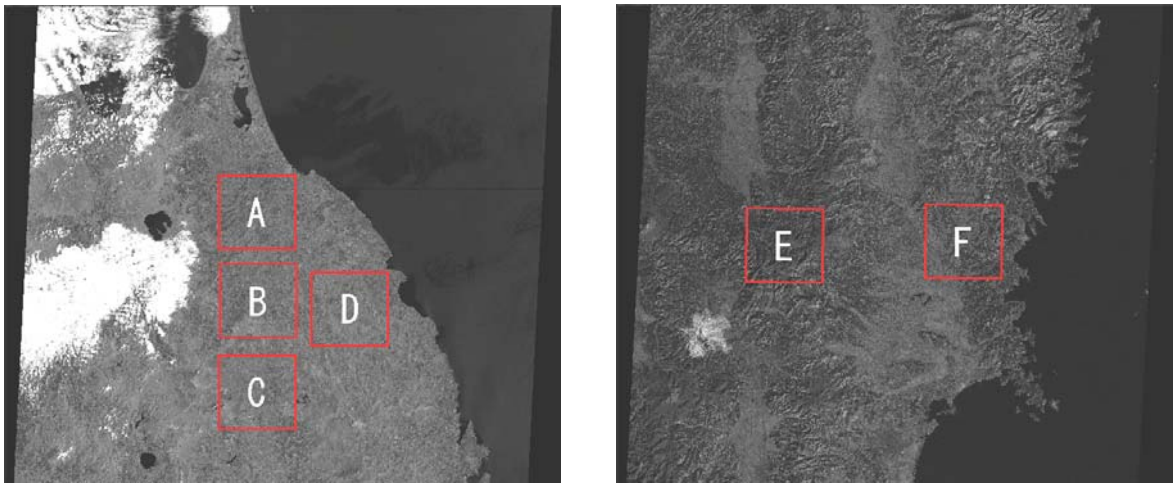
IMPLEMENTATION

Two system-corrected Landsat 5 TM images over the north-east area of Japan were used. Table 1 gives the specifications of the two images with the required parameters for the geometric correction expressed by eq. (7). The listed parameters, which were obtained from the image content, were used to avoid complicated calculations of the satellite orbit. In order to investigate the distribution of the geometrical distortion, six test sites were selected from the full scenes. The test sites are located in mountainous areas covered with vegetation, and are cloudless but contain rivers and river banks partially , as shown in Fig. 3(a) and (b). Each image of the test sites was divided into nine sub-images, and then the geometric correction was carried out for the band 3 and band 5. The spatial dimensions of the sub-images were set in 256×256 for the discrete implementation of the fast Fourier transform.

In the calculation of the relief displacement, the earth radius and the flight height of the satellite were assumed to be constant at 6377.397 km and 705 km respectively, and the elevations of the test sites were obtained from detailed DEMs issued by the Geographical Survey Institute of Japan (GSIJ). As the elevations in the GSIJ DEM files are given at gridded points on latitude and longitude coordinates with a spatial resolution of 50 m, we have transformed them into UTM coordinate system and resampled to a spatial resolution of 30 m to match the spatial

Table 1. Specifications of used images and parameters for geometric correction.

Date	Sensor Type	Path, Row	Scene Center[m]		Orientation Angle[deg]	Solar Angle[deg]	
			x0	y0		Zenith	Azimuth
Jul/11/2000	Landsat 5 TM	107, 32	534087.6	4462557.0	10.686236	61.0	121.0
Dec/26/1997	Landsat 5 TM	107, 33	504375.5	4303571.5	10.575941	23.0	155.0



(a) (b)

Figure 3. Two system-corrected Landsat TM images with the locations of test sites. The open squares in red show the test site locations. (a) The band 5 Landsat 5 TM image, Jul/11/2000. (b) The band 5 Landsat 5 TM image, Dec/26/1997.

resolution of Landsat TM images by using the bilinear interpolation method (Iikura, 2001).

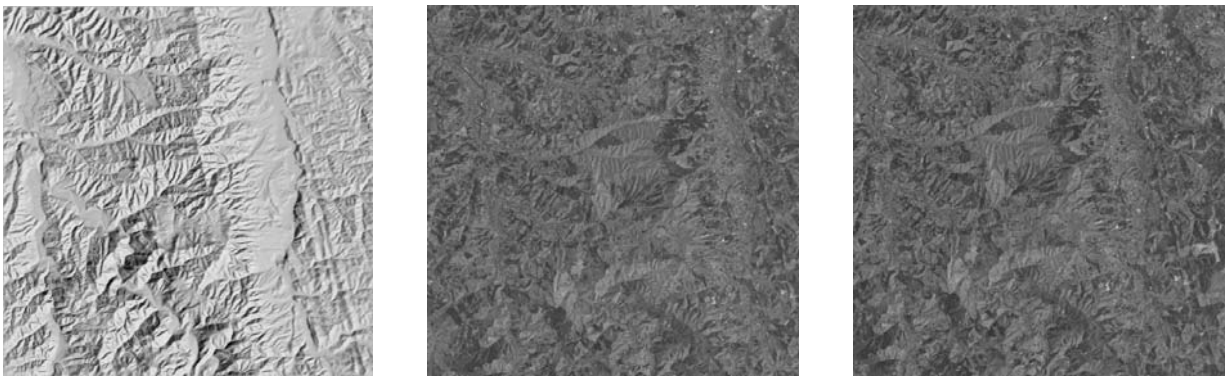
We have examined the conventional method as well as the present method to compare the two estimations each other. In the present method, a pre-set threshold of translation is required. This pre-set threshold is given by the translation offset:

$$Translation\ \ offset\left[\text{pixel}\right]=\sqrt{\delta x^2+\delta y^2}\ . \tag{10}$$

The pre-set threshold was set at 0.05 pixel throughout the calculation. In the conventional method, an initial location of translation and a pre-set threshold of gradient are required for the implementation of the AMOEBA method. The updated scene center location output obtained by the first iteration of the present method was given as the initial location to obtain stable solutions. The pre-set threshold of gradient was set at 1.0×10^{-5}.

The DEM shading image, the image after the geometrical correction, and the image after the optimization, which were produced in the present ortho-rectification procedure, are shown in Fig. 4(a), (b), and (c) respectively. Though a translation exists between Fig. 4(a) and (b), it can be seen from Fig. 4(c) that the translation is successfully removed by the present optimization process.

The estimated translations with band 3 and band 5, and the maxima of the cross-correlation function and the POC function are summarized in Table 2 and Table 3, respectively. Each value of the translation on the two tables is given as the average of nine sub-images in the individual test site. The standard deviations of translations and the differences in the standard deviations between the two methods are enough small to guarantee the adequacy of the pre-set thresholds as well as the stability of the estimated translations. The estimated translations derived from the two methods reached agreements within on the average 1.2 pixel in band 3 and 1.1 pixel in band 5. Note that the translations in Table 2 and 3 are supposed to be used in the eq. (7), that is, the real translations are given by $\left(-\delta x,\ \delta y\right)$.



(a) (b) (c)

Figure 4. Image samples obtained in the ortho-rectification procedure (One of the nine sub-images in the test site B). All dimensions of the images are 256×256. (a) The DEM shading image. (b) The image after geometric correction (band 5). (c) The image after optimization (band 5).

Table 2. Estimations of translation with band 3.

		Translation		SD of Translation		Max. of Correlation	SD of Max.
		δx [pixel]	δy [pixel]	δx [pixel]	δy [pixel]		
Site A	Conventional	49.142	10.809	0.281	0.194	0.202	0.027
	Present	48.076	9.981	0.119	0.109	0.038	0.005
	Difference	1.066	0.828	0.162	0.085	-	-
Site B	Conventional	49.509	11.014	0.099	0.113	0.301	0.015
	Present	48.490	10.486	0.127	0.175	0.056	0.006
	Difference	1.019	0.528	-0.028	-0.062	-	-
Site C	Conventional	49.712	11.112	0.089	0.142	0.292	0.057
	Present	48.535	10.507	0.142	0.319	0.060	0.011
	Difference	1.177	0.605	-0.053	-0.177	-	-
Site D	Conventional	49.720	12.418	0.218	0.215	0.342	0.030
	Present	48.443	11.952	0.195	0.140	0.067	0.007
	Difference	1.277	0.466	0.023	0.075	-	-
Site E	Conventional	28.053	16.662	0.172	0.360	0.367	0.066
	Present	27.019	15.915	0.158	0.269	0.094	0.012
	Difference	1.034	0.747	0.014	0.091	-	-
Site F	Conventional	27.206	18.742	0.057	0.111	0.626	0.084
	Present	26.572	18.454	0.283	0.134	0.126	0.025
	Difference	0.634	0.288	-0.226	-0.023	-	-

SD: standerd deviation, Max of Correlation: Maxmum of cross-correlation or phase-only correration.

Table 3. Estimations of translation with band 5.

		Translation		SD of Translation		Max. of Correlation	SD of Max.
		δx [pixel]	δy [pixel]	δx [pixel]	δy [pixel]		
Site A	Conventional	49.493	10.723	0.143	0.194	0.255	0.037
	Present	48.629	9.766	0.098	0.115	0.064	0.012
	Difference	0.864	0.957	0.045	0.079	-	-
Site B	Conventional	49.731	10.912	0.114	0.111	0.460	0.049
	Present	48.982	10.366	0.168	0.169	0.131	0.008
	Difference	0.749	0.546	-0.054	-0.058	-	-
Site C	Conventional	50.113	10.879	0.080	0.252	0.496	0.090
	Present	49.237	10.156	0.099	0.227	0.124	0.014
	Difference	0.876	0.723	-0.019	0.025	-	-
Site D	Conventional	50.188	12.234	0.111	0.149	0.577	0.022
	Present	49.323	11.503	0.089	0.185	0.138	0.006
	Difference	0.865	0.731	0.022	-0.036	-	-
Site E	Conventional	28.039	16.646	0.123	0.275	0.573	0.042
	Present	27.111	16.021	0.089	0.270	0.129	0.012
	Difference	0.928	0.625	0.034	0.005	-	-
Site F	Conventional	27.351	18.647	0.051	0.112	0.828	0.030
	Present	26.494	18.294	0.362	0.189	0.150	0.007
	Difference	0.857	0.353	-0.311	-0.077	-	-

SD: standerd deviation, Max of Correlation: Maxmum of cross-correlation or phase-only correration.

Fig. 5 shows the convergence behaviors of the translation offset for the two methods. The present method required about 10 iterations to reach the pre-set threshold, and the conventional method required more than 50 iterations. Comparing in the occupation time, the present method reduced the conventional to about one third, for example, the occupation times ware on the average 18 seconds and 60 seconds respectively when we carried out the two methods with a personal computer (CPU, Intel Core2Quad, 2.26 GHz).

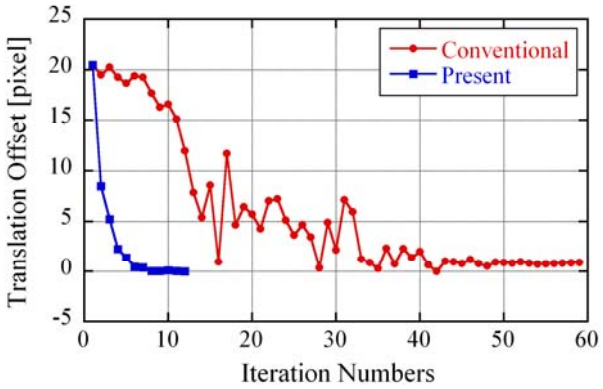


Figure 5. Convergence behaviors of the translation offset for the two methods. The red line with solid circles and the blue line with solid squares indicate the translation offset of the conventional method and the present method, respectively.

DISCUSSION

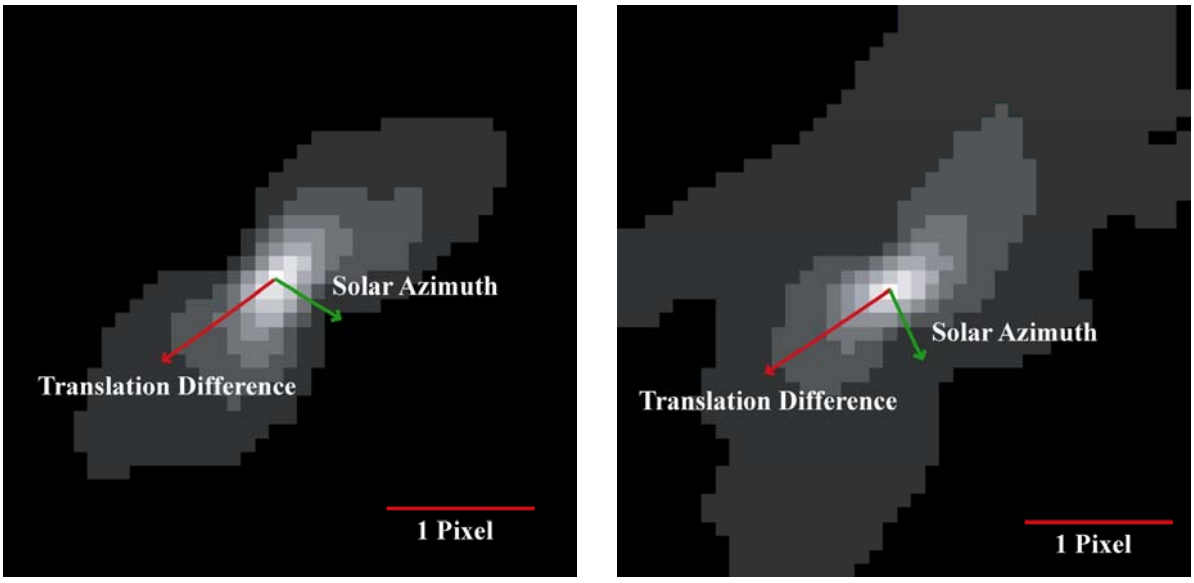
Accuracy of Translation Estimates

The differences in the estimated translations between the two methods are supposed to arise from the difference in the optimization processes because the geometrical correction is common. In case of the present method, the estimated translations have the possibility to include the error of 0.05 pixel due to the pre-set threshold. But, in case of the conventional method, the pre-set threshold does not always restrict the error of translations within 0.05 pixel since it is defined as the gradient of the cross-correlation function.

Table 4. Band 3 to Band5 registration

Scene	Conventional [pixel]	Present [pixel]
Jul/11/2000	0.393	0.718
Dec/26/1997	0.097	0.159

Fig. 6(a) and (b) show the relation between the translation difference and the configuration of the cross-correlation function. Fig.6 (a) is the relation of the test site B and Fig. (b) is that of the test site F. The translation difference with band 5 and the solar azimuth of the site are plotted on the contour map of the cross-correlation function. Each cross-correlation function in two figures is steepest in the direction of solar azimuth, which makes an almost right angle with the gentlest slope direction. The most striking feature in the two figures is that the direction of the gentlest slope corresponds well with the direction of the translation difference. This means that the translation difference is mainly caused by the estimation error of the conventional method.



(a) (b)

Figure 6. Relation between the translation difference and the configuration of the cross-correlation function. The translation difference with band 5 and the solar azimuth of the test site are plotted on the contour map of the cross-correlation function. (a) Test site B. (b) Test site F.

Differences in Band to Band Registration

The differences in band 3 to band 5 registrations are summarized in Table 4. The registration differences between the two bands are very small. However, it is unsuitable to examine the outputs of the conventional method because the average error of translations, which is mentioned above, exceeds the outputs. Thus, we only examine the outputs of the present method.

In the outputs of the present method, the registration difference of the scene on Jul/11/2000 (summer) is larger than that of the scene on Dec/26/1997 (winter). We investigated the maximum of correlation function, which is shown in Table 2 and 3, to see what feature reflects the registration difference between the two bands. In the differences in correlation maximum between the two bands, it can be seen that the difference of the summer scene is larger than that of the winter scene. This means that the band 3 to band5 registration with the present optimization method reflects not only the difference in spectral characteristics between two bands but also the seasonal changes in direct irradiance distribution of the DEM shading images.

SUMMARY

This paper has presented a practical ortho-rectification procedure with DEM shading images for Landsat TM images employing the Fourier phase-only correlation technique to improve the accuracy and efficiency of the conventional method (Iikura, 2001). The sensing accuracy of translation is 0.05 pixel. The occupation time for calculation is reduced to one third. In addition, the present method does not need the selection of GCPs and the consideration for setting of the pre-set thresholds. Therefore, it is easy to execute automatically.

We have optimized only scene center position, that is, the orientation angle and the scale factor in the system information were fixed. The optimization of the orientation angle and the scale factor is one of the important subjects.

REFERENCES

Bryant, N.A., A.L. Zobrist, R.E. Walker, and B. Gokman, 1985. An analysis of Landsat Thematic Mapper P-product and conformity to earth surface geometry, *Photogrammetric Engineering and Remote Sensing*, 51(9):1435-1447.

Chen. Q.S., M. Defrise, and F. Deconinck, 1994, Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition, *IEEE Trans. Pattern Anal. Machine Intell.*, 16(12):1156-1168.

Ekstrand, S., 1996, Landsat TM-based forest damage assessment: Correction for topographic effects, *Photogrammetric Engineering and Remote Sensing*, 62(2):151-161.

Horn, B.K.P., and B.L. Bachman, 1978. Using synthetic images to register real images with surface models, *Communications of the ACM*, 21(11):914-924.

Horner, J.L., and P.D. Gianino, 1984, Phase-only matched filtering, *Applied Optics*, 28(10):812-816.

Iikura, Y., 2001, Evaluation of interpolation methods for converting map projection of digital elevation model, *Journal of Japan Remote Sensing Society*, 21(2):150-157.

Iikura, Y., 2002, Optimization of geometric correction of Landsat TM images using digital elevation model, *Journal of Japan Remote Sensing Society*, 22(2):189-195.

Levin, N., E. Ben-Dor, and A. Karnieli, 2004, Topographic information of sand dunes as extracted from shading effects using Landsat images, *Remote Sensing of Environment*, 90:190-209.

Para, V., and X. Pons, 1995, Incorporation of relief in polynomial-based geometric corrections, *Photogrammetric Engineering and Remote Sensing*, 61(7):281-288.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1989, *Numerical Recipes* (FORTRAN Version), Cambridge University Press, Cambridge, pp.294-300.

Takita, K., T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, 2003, High-accuracy subpixel image registration based on phase-only correlation, *IEECE Trans. Fundamentals*, E86-A(8):1925-1934.

Zhou, G., K. Jezek, W. Wright, J. Rand, and J. Granger, 2002. Orthorectification of 1960s satellite photographs covering Greenland, *IEEE trans. Geosci. Remote Sensing*, 40(6):1247-1259.