

IMPROVEMENT OF THE STABILITY SOLVING RATIONAL POLYNOMIAL COEFFICIENTS

Xianyong Lin^{a,*}, Xiuxiao Yuan^a

^a School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
winterlinny@163.com

Commission I, WG I/5

KEY WORDS: High Resolution Satellite Imagery; Sensor Model; Orientation; Regularization; Accuracy

ABSTRACT:

The rational function model (RFM) utilized for high resolution satellite imagery (HRSI) provides a transformation from image to object space coordinates in a geographic reference system. Compared with the rigorous model based on the collinearity condition equation or the affine model, the RFM with 80 coefficients would be over parameterized. That would result in an ill-conditioned normal equation. Tikhonov regularization is often used to resolve this problem, and many applications have verified its serviceability. This paper will detail the method for regularization parameter selection. However, Tikhonov regularization makes the two sides of equation unequal, resulting in a biased solution. An unbiased method - The Iteration by Correcting Characteristic Value (ICCV) was introduced, and a strategy to resolve the ill-conditioned problem for solving rational polynomial coefficients (RPCs) was discussed in this paper. The tests with SPOT-5 and QuickBird imagery were accomplished. The empirical results have shown that our methodology can effectively improve the condition of the normal equations.

1. INTRODUCTION

Since the launch of the IKONOS II satellite, the rational function model (RFM) has gained considerable interests in photogrammetric community. SpaceImaging Company provides the RFM to users instead of the physical sensor model, subsequently, DigitalGlobe Corporation provides the RFM together with the strict geometric model in order to satisfy different users. The RFM has been universally accepted, and validated, as an alternative sensor orientation model for high resolution satellite imagery (HRSI). The RFM is an approximation of the rigorous sensor model, via a number of control points. Then it could be utilized in the photogrammetric process instead of the complex rigorous sensor model. It would be a part of the standard image transfer format, and it is becoming a standard way for economical and fast mapping from remotely sensed imagery.

The key of the RFM is to gain accurate rational polynomial coefficients (RPCs). Compared with the rigorous model based on the collinearity model or the affine model, the RFM with 80 coefficients would be over parameterized (Fraser et al., 2005). That may cause the design matrix to become almost rank deficient because of the complex correlation among RPCs. It may result in numerical instability in the least squares adjustment, or even producing wrong solutions. The regularization technique was often suggested to tackle the possible ill-conditioned problem during the adjustment (Tao and Hu, 2001a). It has been proved to effectively improve the condition of the normal equations. But the determination of the regulation parameter has still been considered to be a challenge.

Regularization parameter selection is crucial to the regularization technique. There are several methods for the optimal parameter determination, including ridge trace method, L-curve criterion, generalized cross validation (GCV) method, ordinary cross validation (OCV) method, and so on. The effects may be totally different when we use different methods. The

L-curve, GCV and OCV were compared by Choi et al. (2007). In practice, ridge trace method is widely used for its simpleness, where solutions are computed for a large number of different regulation parameters, selecting the best one by suitable heuristics (Tao and Hu, 2001a). However, ridge trace method can not obtain the optimal parameter, and it is inconvenient for automatic computation. The authors try different regulations for the RPCs computation. The L-curve criterion has been proved to be efficacious.

Despite regularization technique gives a good result, it makes the two sides of equation unequal by imposing constraints to the diagonal elements of the normal equation matrix, resulting in a biased solution. So, we will introduce an unbiased method, the Iteration by Correcting Characteristic Value (ICCV). This method is simple, and it was put forward more than ten years ago, but still not widely used. The initial values will be the main factor that affects the result. And this paper will suggest two ways to set initial values, just for RPCs computation.

Accurate RPCs are crucial to the RFM model, which directly determines whether it could replace the physical sensor model to accomplish the photogrammetric process. And the ill-conditioned normal equation would be the main problem. This paper aimed at finding a proper method to resolve the possible ill-conditioned problem, getting an accurate solution. Initially, we will review the basic model and the methods including the RFM solution, the terrain-independent and terrain-dependent computational scenarios. The regulation technique and the method of regularization parameter selection are then addressed, focusing on the L-curve method. The unbiased ICCV is followed. The results of experimental tests with SPOT-5 and QuickBird imagery are then discussed. Finally, we will suggest a strategy for RFM computation according to the experiments and comprehensive analysis of the characteristics of the various methods.

2. THE RATIONAL FUNCTION MODEL

2.1 The RFM Model

The RFM relates object point coordinates to image pixel coordinates in the form of rational functions that are ratios of polynomials. For the ground-to-image transformation, the defined ratios have the forward form (OGC, 1999):

$$\begin{aligned} l_n &= \frac{Num_L(U, V, W)}{Den_L(U, V, W)} \\ s_n &= \frac{Num_S(U, V, W)}{Den_S(U, V, W)} \end{aligned} \quad (1)$$

Where

$$\begin{aligned} Num_L(U, V, W) &= a_1 + a_2V + a_3U + a_4W + a_5VU \\ &+ a_6VW + a_7UW + a_8V^2 + a_9U^2 \\ &+ a_{10}W^2 + a_{11}UVW + a_{12}V^3 \\ &+ a_{13}VU^2 + a_{14}VW^2 + a_{15}V^2U \\ &+ a_{16}U^3 + a_{17}UW^2 + a_{18}V^2W \\ &+ a_{19}U^2W + a_{20}W^3 \\ Den_L(U, V, W) &= b_1 + b_2V + \dots + b_{19}U^2W + b_{20}W^3 \\ Num_S(U, V, W) &= c_1 + c_2V + \dots + c_{19}U^2W + c_{20}W^3 \\ Den_S(U, V, W) &= d_1 + d_2V + \dots + d_{19}U^2W + d_{20}W^3 \end{aligned} \quad (2)$$

Here, a_i, b_i, c_i and d_i are the 80 RPCs; b_1 and d_1 are commonly set to 1. (l_n, s_n) are the normalized line and sample index of the pixels in image space, while (U, V, W) are normalized object point coordinates. That is:

$$\begin{aligned} l_n &= \frac{l - LineOffset}{LineScale} \\ s_n &= \frac{s - SampleOffset}{SampleScale} \\ U &= \frac{\phi - LatitudeOffset}{LatitudeScale} \\ V &= \frac{\lambda - LongitudeOffset}{LongitudeScale} \\ W &= \frac{h - HeightOffset}{HeightScale} \end{aligned} \quad (3)$$

Here, (l, s) are the image line and sample coordinates; (ϕ, λ, h) represent latitude, longitude, height; the offsets and scales normalize the coordinates to $[-1, 1]$, minimizes the introduction of errors during computation.

The RFM has nine configurations with some variations, such as subset of polynomial coefficients, equal or unequal denominators. Also, it has forward and backward forms (Tao and Hu, 2001a). Generally speaking, the RFM refers to a specific case that is in forward form, has third-order polynomials with unequal denominators, and is usually solved by the terrain-independent scenario.

2.2 RFM Solution

Two methods have been developed to solve for the RFM, direct and iterative least-squares solutions (Tao and Hu, 2001a). Here, the direct least-squares solution of RFM is given as follows:

$$\begin{aligned} F_s &= Num_S(U, V, W) - s \cdot Den_S(U, V, W) = 0 \\ F_l &= Num_L(U, V, W) - l \cdot Den_L(U, V, W) = 0 \end{aligned} \quad (4)$$

$$\mathbf{V} = \mathbf{B}\mathbf{X} - \mathbf{L}, \quad \mathbf{P} \quad (5)$$

$$\mathbf{X} = (\mathbf{B}^T \mathbf{P} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{P} \mathbf{L} \quad (6)$$

where

$$\mathbf{B} = \begin{bmatrix} \frac{\partial F_s}{\partial a_i} & \frac{\partial F_s}{\partial b_j} & \frac{\partial F_s}{\partial c_i} & \frac{\partial F_s}{\partial d_j} \\ \frac{\partial F_l}{\partial a_i} & \frac{\partial F_l}{\partial b_j} & \frac{\partial F_l}{\partial c_i} & \frac{\partial F_l}{\partial d_j} \end{bmatrix},$$

$$(i = 1, \dots, 20; j = 2, \dots, 20)$$

$$\mathbf{X} = [a_i \quad b_j \quad c_i \quad d_j]^T,$$

$$\mathbf{L} = [d_1 s \quad b_1 l]^T = [s \quad l]^T.$$

\mathbf{P} is the weight matrix, and it is usually set as identity matrix.

2.3 Approaches of Determining RPCs

There are terrain-independent scenario using known physical sensor model and terrain-dependent scenarios using ground control points. The terrain-independent scenario is to use the onboard ephemeris and attitude data. With the physical model available, a virtual control grid covering the full extent of the image and the entire elevation is generated. The RPCs are estimated using a least-square solution with the image grid points and the corresponding object grid points (Tao and Hu, 2001a, 2001b).

For the terrain-dependent scenario, a number of ground control points are collected for the RPCs computation. At least 39 ground control points are needed per image to solve 78 RPC coefficients, excluding the constant parameters b_1 and d_1 . And the solution is highly dependent on the actual terrain relief, the distribution and the number of GCPs (Tao and Hu, 2001a, 2001b).

3. REGULARIZATION TECHNIQUE

The RPCs may display very high correlation between coefficients. That would be a potential problem for obtaining a stable solution. The design matrix is usually ill conditioned in the experiments (Tao and Hu, 2000). Even for the well conditioned observation equations, regularization can improve the accuracy of the RPCs, and help produce well-structured RPCs, especially for the third-order RFM (Hu and Tao, 2004).

3.1 Ridge Regression

Ridge regression (Ridge estimate), a part of regularization technique, is a biased estimation for nonorthogonal problems (Hoerl and Kennard, 1970). It carries out by adding a small positive quantity to the diagonal of $\mathbf{B}^T \mathbf{B}$. Ridge regression

obtains biased estimates with smaller mean square error. Ridge regression is defined as follow:

$$\hat{X}(k) = (B^T PB + kI)^{-1} B^T PL \quad (7)$$

Where

k is ridge parameter or regularization parameter, usually a small positive quantity;

I is identity matrix;

$\hat{X}(k)$ is ridge regression estimation.

3.2 Ridge trace method for parameter determination

Solutions are computed for a set of different k values. And the best k is selected by suitable heuristics, for the least error at check points (Tao and Hu, 1970). This method is very simple, and it is widely used.

3.3 L-curve criterion for parameter determination

The L-curve is a log-log plot of the norm of a regularized solution versus the norm of the corresponding residual (fitting error) as the regularization parameter is varied (Hansen, 1992; Rodriguez and Theis, 2005). L-curve is presented as:

$$(\eta(k), \xi(k)) = (\lg \|B\hat{X} - L\|_k, \lg \|\hat{X}\|_k) \quad (8)$$

The curve is L-shaped: approximately vertical for small k , and approximately horizontal for large k , with the corner providing the optimal regularization parameter. So the object is to find out the point with biggest curvature:

$$k = \arg \max \frac{\xi' \eta'' - \xi'' \eta'}{\sqrt{(\xi'^2 + \eta'^2)^2}} \quad (9)$$

Where

ξ' 、 ξ'' = the first and second derivative of ξ on k ;

η' 、 η'' = the first and second derivative of η on k .

In practical computation, curve fitting is often used to obtain the L-curve.

The L-curve criterion is able to recognize correlated errors, while the GCV method may fail to do so. That is essentially because the L-curve criterion combines information about the residual norm with information about the solution norm, whereas the GCV method only uses the information about the residual norm. The research done by Choi et al (2007) shows us that the L-curve method performed better than OCV or GCV, particularly for high noise levels. The L-curve method is found to be less susceptible to producing large reconstruction errors but it tends to over-regularize the solution in the presence of low noise, leading to under-estimates of the forces.

4. THE ITERATION BY CORRECTING CHARACTERISTIC VALUE

Regularization technique imposes constraints to the diagonal elements of the normal equation, resulting in a biased solution. So, we will introduce an unbiased method - the Iteration by Correcting Characteristic Value (Wang et al. 2001) for RPCs computation.

Here is the norm function,

$$B^T PB \hat{X} = B^T PL$$

Add \hat{X} to both sides,

$$(B^T PB + I)\hat{X} = B^T PL + \hat{X}$$

There are \hat{X} on both sides, so it should be resolved in iterative mode:

$$\hat{X}^{(k)} = (B^T PB + I)^{-1} (B^T PL + \hat{X}^{(k-1)}) \quad (13)$$

If we set $q = (B^T PB + I)^{-1}$,

Then the (13) could be written as:

$$\hat{X}^{(k)} = (q + q^2 + \dots + q^k) B^T PL + q^k \hat{X}^{(0)} \quad (14)$$

Where

$\hat{X}^{(0)}$ = initial values of the solutions.

Eqs. (13) and (14) are the expressions of the iteration by correcting characteristic value. The convergent and unbiased properties are discussed by Wang Xinzhou et al. (2001).

The ICCV carried by iteration, initial values should be offered for the iteration, and they have an important impact on the result or even determine the success of the method. The direct least-squares solutions are usually used as the initial values. Unfortunately, when the ill-condition happens, it is possible that the LS solution is so bad that the iteration is unconvergence. So, here we suggest another way special for the RPCs solution. Considering that the third-order RPCs are closed to zero, the initial values may set to zero. We will test it in the experiments.

5. TEST RESULTS AND EVALUATION

5.1 Design and Tests

The tests have been designed for these purposes:

To evaluate the numerical stability of the direct least squares solution. The condition number is cursorily employed to measure the condition of the design matrix. The number is much bigger when the function is ill-conditioned.

To compare the performances of the regularizations for the RPCs computation. We choose the widely-used ridge trace method and the L-curve criterion.

Mainly to evaluate the potential of the unbiased ICCV method for the RPCs computation, and to test the impact of the initial values. Initial values are set by zero and least square solution respectively.

To find out an effective strategy to tackle the possible ill-conditioned problem.

Here we confine the experiments to the third-order RFM with 80 coefficients, based on the terrain-independent scenario. With the rigorous sensor model established and the elevation range obtained from a cursory DEM, the 3-D grid of object points was generated, with 5 constant elevation planes each with 10 by 10 grid points. While the check grid consists of 10 constant elevation planes each with 20 by 20 grid points. So there are 500 control points and 4000 check points.

The fitting accuracy is measured in image both at control points and check points. Firstly, the image position of the grid points is

calculated by the obtained RFM. Then the differences between the pixel coordinates of the original grid points and those from the RFM are calculated for evaluation. The accuracy determination is quite the same as mentioned by Grodecki and Dial (2001).

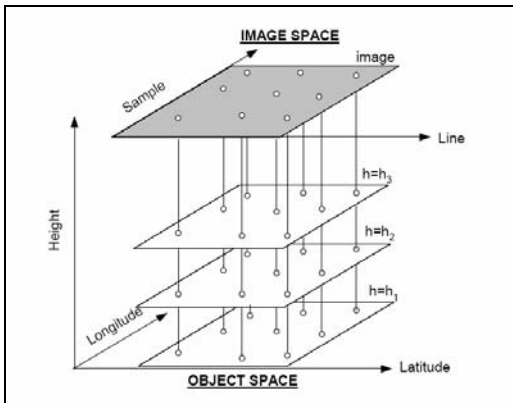


Fig. 1 3D object grid generated for solving RPCs

5.2 Test data sets

Tao and Hu (2001a) tested with the aerial photograph data and SPOT data with sizes of 6000 by 6000. In order to evaluate the fitting accuracy of the different methods for HRSI, we choose SPOT-5 and QuickBird imagery. Respective ground pixel sizes for testfield imagery were 5 m for SPOT-5, and 70 cm for the QuickBird. Further details regarding the test-range are given in Table 1.

Data set	Ground pixel Size (m)	Image size (pixel)	Elevation range (m)
SPOT-5	5	12 000×12 000	-2~327
QuickBird	0.7	27 552×22 700	340~1194

Table 1. Information of the data sets

5.3 Results and evaluation

All the methods are tested on both the SPOT-5 and QuickBird imagery. The RMSE and the maximal errors in the imagery at the control points and the check points are listed in Table 2 for SPOT-5 data, Table 3 for QuickBird data. The condition numbers of the norm function, before and after regularization, are also listed in the tables.

There is not an absolute criterion for exactly judging that the norm function is ill-conditioned or not, and how ill-conditioned it is. Generally speaking, the condition number is helpful, the bigger it is, the worse the condition is. Based on Table 2 and Table 3, the condition numbers are big for both images, 7.91×10^{14} for SPOT-5 data, and 1.13×10^{11} for QuickBird data. The direct least square solutions are not very good, out of sub-pixel, especially for the QuickBird data, the RMSE at check points arrives at 115.937 pixels, and the maximal error is as bigger as 7280.348 pixels. Therefore, the direct least square solutions here could not be the final RPCs which would substitute the physical sensor model.

Determining regularization parameter using ridge trace method is shown in Fig. 2, and L-curve method is shown in Fig. 3,

where SPOT-5 imagery is employed as an example. In the experiment, RPCs are computed for a number of k with different orders of magnitude varying from 10^{-10} to 10^{-1} , to determine the order of magnitude. Then employ more k around the order of magnitude, and choose the one that has the smallest error at check points. This method can not select the best parameter, and it is not convenient for automatic computation.

For the L-curve method, the curve is shaped like "L". And the corner point on the L-curve that has maximum curvature corresponds to the optimal parameter. This method can offer an exact parameter automatically, without the need to plot the L-curves. In the experiments, the parameter determined by L-curve criterion is 2.04×10^{-7} for SPOT-5 data, and 9.04×10^{-6} for QuickBird data.

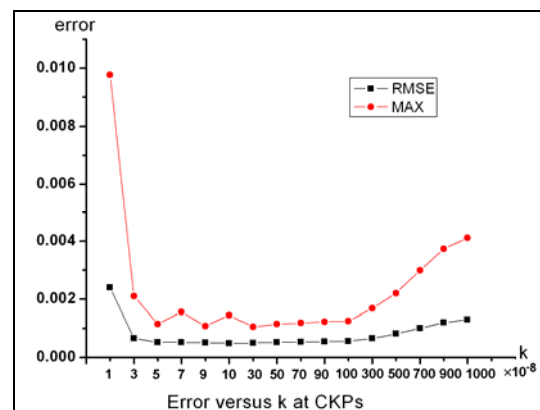


Fig. 2 Determining ridge parameter using ridge trace method

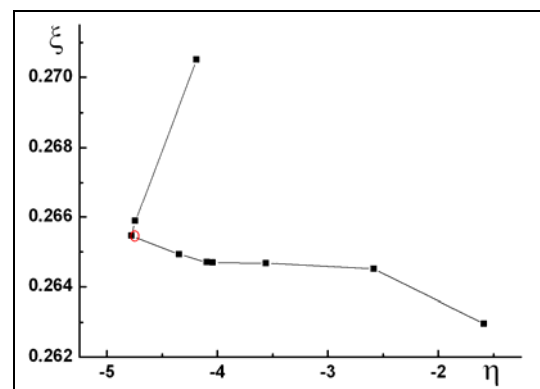


Fig. 3 Determining ridge parameter using L-curve method

By comparison, the regulation by L-curve criterion, made very significant improvements in terms of accuracy. After the use of regularization, the condition numbers are smaller than the original one, for the SPOT-5 data, reducing from 10^{14} to 10^9 , and for the QuickBird data, reducing from 10^{13} to 10^8 . Except the high accuracy, L-curve method shows very strong stability based on more data sets.

The ICCV carries out by iteration. In the experiments, the initial values are set by zero and least square solution respectively. The iterative threshold value set as 10^{-6} . For the zero initial values condition, the results are pretty good based on the tables that the accuracy is so close to and even better than the results

of the L-curve method. It converged after a few times iteration, even the iterative threshold value set to 10^{-9} , it converges quickly. For the least squares solution initial value condition, we can note that it is good for SPOT-5 data but invalid for the QuickBird data. It converged only after one time iteration even that the iterative threshold value is strict. Coercive iteration is also invalid for the improvement of the accuracy. More tests

should be done with the ICCV. It is worth of pointing out that the computation of the ICCV is very simple and fast, that is because there is no need to determine the regulation parameter, and no need to inverse the matrix every time during the iteration. And the structure of the solution by ICCV is as good as that by regulation. Anyway ICCV is a potential way to overcome the ill-conditioned problem for the RFM solution.

Approaches	Condition number	Iteration times	Errors at CNPs (pixels)		Errors at CKPs (pixels)	
			RMSE	Max	RMSE	Max
Least squares (LS)	7.91×10^{14}	-	1.774	4.190	1.609	4.188
Ridge Estimate						
L-curve criterion	4.26×10^9	-	0.000	0.001	0.000	0.001
ICCV						
LS solutions as initial values	-	7	0.000	0.001	0.000	0.001
Zero as initial values	-	12	0.000	0.001	0.001	0.001

Table 2. RMSE and Max errors in image with the SPOT-5 data

Approaches	Condition number	Iteration times	Errors at CNPs (pixels)		Errors at CKPs (pixels)	
			RMSE	Max	RMSE	Max
Least squares (LS)	1.11×10^{13}	-	4.893	77.168	115.937	7280.348
Ridge Estimate						
L-curve criterion	1.07×10^8	-	0.358	0.745	0.335	0.734
ICCV						
LS solutions as initial values	-	1	4.895	77.334	115.684	7264.072
Zero as initial values	-	8	0.357	0.728	0.335	0.710

Table 3. RMS and Max errors in image with the QuickBird data

The DigitalGlobal Corporation provide the RPC file to the users. As a comparison, we choose 9 ground control points to checkout the RPCs, calculating the differencec between the image coordinates of the GCPs and that from the RPCs. And the errors at the GCPs are listed in Table 4. From the table we can see the RPCs by the L-curve and the ICCV are so close according to the accuracy, and they are slightly better than the RPC provided by the corporation.

RPC		Errors at GCPs (pixels)		
		m_l	m_s	m_{ls}
By L-curve	RMSE	9.310	9.082	13.006
	Max	11.556	13.334	16.856
By ICCV	RMSE	9.309	9.081	13.005
	Max	11.550	13.335	16.855
Provided by DigitalGlobal	RMSE	9.391	9.592	13.423
	Max	11.563	13.867	17.646

Table 4. RMS and Max errors in image at GCPs

6. CONCLUSIONS

Since the SpaceImaging Company provided the RPCs to the end users and the service providers, the RFM has been with us for eight years, and a lot of researches show us that it is a useful tool for exploiting high resolution satellite images. Subsequently, the DigitalGlobal Corporation provides the RPCs together with the physical sensor model, and more imagery vendors may adopt the RFM, providing a way for economical and fast mapping from HRSI.

The aim of this paper is to suggest a proper way to resolve the ill-conditioned problem for the RFM solution. Regulations improve the stability of the inverse matrix evidently and produce a well structure RPCs. And the L-curve method performs well in the experiments, being accurate and stable. The ICCV is unbiased, simple, fast, and accurate, and the idea that set the initial values as zero acts well for the RFM solution. Both the methods show good effects, improving the accuracy of the solutions, and ameliorating the RPCs' structure. Considering that the L-curve method has the risk of over-regularizing the solution when the ill-condition is slight, though not happened in the experiments, the ICCV should be the first choice. Finally, we suggest a strategy, that L-curve method work for high level ill-condition and ICCV for low. Even for the well conditioned

design matrix, the ICCV is really helpful to improve the accuracy.

ACKNOWLEDGEMENTS

Thanks for the supporting from the 973 Program of the People's Republic of China under Grant 2006CB701302 and the National Natural Science of China under Grant 407721001.

REFERENCES

- Choi, H.G., Thite, A.N., Thompson, D.J. 2007. Comparison of methods for parameter selection in Tikhonov regularization with application to inverse force determination. *Journal of Sound and Vibration*, 304, (3-5), pp. 894-917.
- Hoerl, A.E., Kennard, R.W., 1970. Biased Estimation for Non-orthogonal Problems. *Technometrics*, 12(1), pp. 55-67.
- Hansen, P.C., 1992. Analysis of Discrete Ill-posed Problems by Means of the L-curve. *SIAM Review*, 34(4), pp. 561-580.
- Hu, Y., Tao, C.V., Croitoru, A., 2004. Understanding the rational function model: methods and applications, *IAPRS*, 12-23 July, Istanbul, vol. XX, 6 p.
- OpenGIS Consortium (OGC), 1999. *The openGIS abstract specification-Topic 7: Earth imagery*.
- Rodriguez, G., Theis, D., 2005. An Algorithm for Estimating the Optimal Regularization Parameter by the L-curve. *Rendiconti di Matematica*, 25(1), pp. 69-84.
- Fraser, C.S., Dial, G., Grodecki, 2005. Sensor orientation via RPCs. *ISPRS Journal of PRS*, 60, pp. 182-194.
- Tao, C.V., Hu, Y., 2000. Investigation of the rational function model. *Proceedings of ASPRS Annual Conference*, Washington, DC. May 22-26. 11 pages. <http://www.gis.usu.edu/docs/protected/procs/asprs/asprs2000/pdffiles/papers/039.pdf> (accessed 31 Mar. 2008).
- Tao, C.V., Hu, Y., 2001a. A comprehensive study on the rational function model for photogrammetric processing, *Photogrammetry Engineering and Remote Sensing*, 67(12), pp. 1347-1357
- Tao, C.V., Hu, Y., 2001b. The rational function model: a tool for processing high resolution imagery. *Earth Observation Magazine (EOM)*. 10(1). pp. 13-16
- Grodecki, J., Dial, G., 2001. Ikonos geometric accuracy, *Joint ISPRS Workshop on HRM from Space*, 19-21 Sept., pp. 77-86.
- Wang Xinzhou, Liu Dingyou, Zhang Qianrong, Huang Hailan. 2001. The Iteration by Correcting Characteristic Value and its application in surveying data processing. *Journal of Heilongjiang institute technology*. 15(2), pp. 3-6.