Defence Science Journal, Vol. 52, No. 3, July 2002, pp. 253–259
© 2002, DESIDOC

Image Registration using Polynomial Affine Transformation

Narayan Panigrahi

Centre for Artificial Intelligence & Robotics, Bangalore – 560 001

and

Smita Tripathy

Aeronautical Development Agency, Bangalore – 560 037

ABSTRACT

An error is generally introduced when an image is taken remotely by an imaging system. This is more prominent when an airborne system takes the image of the earth's surface. Generally the error, known as geometric error, is a composite of translation, rotation, scaling, and warping of the image. This paper discusses an algorithm to remove geometric error by applying geometric transformation and registering an image with its reference image, or the same image taken at some other instance of time. To remove these highly irregular distortions, translation, rotation, scaling, and warping to the image is applied simultaneously. This is done by establishing a polynomial affine transformation, which is applied to the image with geometric error to register it with the reference image, thus removing the error.

Keywords: Image registration, geometric error, polynomial affine transformation, ground correlation point, control point, barycentric coordinates

1. INTRODUCTION

An error is always introduced when an image is taken remotely by an imaging system. This is prominent when an airborne system takes the image of the earth's surface[1,2]. Geometric image transformation of an image includes translation, rotation, scaling, and warping of the image to remove highly irregular distortions, such as those caused by view angle change, scanner motion, camera motion during image acquisition or improper camera calibration[2–4]. These phenomena are, particularly, prominent in images taken by satellite or by a camera mounted on an airborne platform.

To remove geometric error and register an image with its corresponding reference map or image of the same area taken at some other instance of time, simultaneous application of translation, rotation, scaling and warping of the image is necessary[1,4-7]. These operations are collectively known as rubber sheet operation as the image is aligned with an existing digital vector map of the same area by stretching, shrinking or tearing like a rubber sheet. Also the term image registration is used to describe geometric transformation. For satellite image registration, a geo-coded, geo-referenced satellite image with highly irregular distortion has to be registered or aligned with an existing digital vector map of the same area, given three or more ground correlation points (GCPs)[2,6,8]. This emphasises the fact that the generated image should register with some standard or existing reference image. It is desirable to reduce the inherent error in the image so that proper alignment with the reference map is achieved.

Revised 07 July 2002

This paper discusses various aspects related to the geometric errors observed in a satellite image and geometric transformation used to remove these. A polynomial affine transformation (PAT) is established by at least three pairs of non-collinear GCPs in the satellite image and corresponding control points (CPs) in the reference image. The correlation of the GCP–CP pairs is put in the form of polynomials, which gives rise to the system matrix[5,6,9]. These system matrices are generally square matrices and highly inconsistent due to the choice of CPs, hence least square method is used to solve the system matrix and the coefficients of the polynomials, which establishes the deformation model or polynomial transformation, are obtained by solving the system matrix.

2. DESCRIPTION OF AFFINE TRANSFORMATION

A transformation f: $R^n \rightarrow R^n$ is affine if

$f[(1-t)P + tQ] = (1-t)\ f(P) + t\ f(Q)$, for $P, Q \in R^n$, and $t \in R$.

Geometrically, this means that the transformation preserves both collinearity and linear interpolation on each line, i.e., it maintains the ratio of segments on each line. Thus, two distinct points constitute an affine frame of the line in R. Three non-collinear points constitute an affine frame of the plane R^2, these points define a triangle of the plane. Four non-coplanar points of space R^3 constitute an affine frame. These points of the R^3 define a tetrahedral in the space. An affine frame enables one to introduce generalised coordinates on the affine space known as barycentric coordinates. Barycentric coordinates give a general formula for computing the coordinate of the space preserving the affine property. Each affine transformation is a linear combination of translation, scaling, rotation, and warping such that the affine property is satisfied. Thus, affine transformation enables one to obtain a technique to interpolate a transformation known on the vertices of a triangle[2,5,6]. Hence, affine transformation is a linear transformation followed by a translation. Given a matrix A and a vector B, the equation $X = AX + B$ defines the affine transformation of vector X.

$$X' = AX + B \Rightarrow \begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + B$$

For translation

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

For rotation

$$A = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

For scaling

$$A = \begin{pmatrix} s_{11} & 0 \\ 0 & s_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

3. POLYNOMIAL AFFINE TRANSFORMATION

The equation for a general affine transformation in R^2 is defined by M: $R^2 \rightarrow R^2$, and given by a simple equation

$$(K, L) = M(i, j)$$

where (i, j) is the coordinate of input image with error and (K, L) is the coordinate of reference image, the digital vector map in this case. M is the affine transformation, which transforms set of $(i, j) \in R^2$ to set of $(K, L) \in R^2$. In other words, for each pixel (i, j) in the output image, compute its corresponding location (K, L) in input image, obtain the pixel from input image and put it in output image. Since a reverse computation of pixel location is used, this process is also known as reverse transformation or inverse transformation or output-to-input transformation. This transformation can be expressed through a pair of polynomials as

$$K = Q(i, j) = q_0 + q_1 i + q_2 j + q_3 ij + q_4 i^2 \quad (1)$$

$$L = R(i, j) = r_0 + r_1 i + r_2 j + r_3 ij + r_4 i^2 \quad (2)$$

These polynomial equations can be represented in matrix form also. Since the affine transformation is represented through a set of polynomials, hence it is named as PAT. The unknown coefficients q_i and r_i are obtained after solving the below system matrix representing polynomial (1).

$$\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ \cdot \\ \cdot \\ k_m \end{bmatrix} = \begin{bmatrix} 1 & i_1 & j_1 & i_1 j_1 & \\ 1 & i_2 & j_2 & i_2 j_2 & \\ 1 & i_3 & j_3 & i_3 j_3 & \cdots \\ & & \cdot & & \\ & & \cdot & & \\ 1 & i_m & j_m & i_m j_m & \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ \cdot \\ \cdot \\ q_m \end{bmatrix}$$

i.e. $K = MQ$

The dimension and condition of the system matrix M depends upon the number of GCP–CP pairs selected and their distribution on the image. At least three pairs of non-collinear GCPs need to be selected to establish an affine frame in R^2. If three pairs of GCPs are selected, then a 3×3 square matrix representing the transformation is obtained, which can solve the translation, rotation, and scaling distortions in the input image. If the GCPs are collinear and densely populated, then the matrix is ill-conditioned and sometimes it leads to inconsistency and rank-deficiency[5,6]. Hence, choice of more number of GCP–CP pairs lead to remove highly irregular geometric distortions and with greater accuracy. Sometimes, least square method is used to avoid inconsistency and to solve the above matrix equation for robust result. By definition, least square method is the one that minimises

$\| K - MQ \|^2$, which when solved leads to

$$Q_{LSM} = [\ M^T M\ ]^{-1} M^T K$$

Similarly, a matrix equation for second polynomial (2) can be derived and solved for R resulting

$$R_{LSM} = [\ M^T M\ ]^{-1} M^T L$$

4. SOLVING SYSTEM EQUATION USING LEAST SQUARE METHOD

The problem above is analogous to solve the matrix equation $y = Hx$, where given the values of y and H and the relationship, one wishes to find x. The most obvious approach would be to use the inverse matrix method, i.e., a common problem encountered in many disciplines, the inversion of a linear system of equations.

$$X_{inv} = H^{-1} y$$

There are many well-known ways to invert a matrix. If the matrix is a symmetric non-negative definite one, a square root method (such as Cholesky decomposition) can be used to reduce the complexity of the computation. If the matrix is Toeplitz, then Levinson's recursion can be used. Also, there are several reasons why the inverse may not be useful in practice. The first reason applies to the case where the equations are inconsistent. The additional problem of rank deficiency may arise, which is described through the following set of equations.

Consider the following sets of inconsistent equations:

$$X_1 + 2X_2 + X_3 = 2 \tag{3}$$

$$X_1 - X_2 + X_3 = 1 \tag{4}$$

$$2X_1 + X_2 + 2X_3 = 2 \tag{5}$$

The first two equations can be added to obtain an equation with the same coefficients as found in the third equation. However, the value in the RHS is not the same, since addition of Eqns (3) and (4) yield $2X_1 + X_2 + 2X_3 = 3$ which is inconsistent with Eqn (5). The problem of inconsistencies within a set of equations can be circumvented by introducing extra variables to compensate for the error. Thus, the set of linear equations can be written as

$$X_1 + 2X_2 + X_3 + e_1 = 2 \tag{3a}$$

$$X_1 - X_2 + X_3 + e_2 = 1 \tag{4a}$$

$$2X_1 + X_2 + 2X_3 + e_3 = 2 \tag{5a}$$

The resulting set of equations is underdetermined, i.e., there are more unknowns than the equations. Thus, there are many solutions. Hence, by introducing error variables, a problem that had no solutions has been transformed into one that has an infinite number of solutions. The situation has improved by going to higher dimension (six unknowns instead of three). In order to select one of many solutions, one is free to choose some additional constraints. A good condition might be to constraint the error vector e to have the smallest energy. Ideally, one would like this quantity to be zero. In this case, one would like to find some x such that the following conditions are satisfied:

$$\text{Minimise} \quad \sum_{i=1}^{3} e_i^2$$

$$\text{Subject to } y = Hx + e$$

One may consider the minimisation of some other cost function, such as

$$\sum_{i=1}^{3} e_i$$

However, this cost function is not a good measure of the error, since there may be some large positive and negative errors that cancel each other, thus giving an incorrect indication of the size of the error. The solution to the least square problem with linear constraints can be solved exactly in closed form. The solution can be obtained by solving for e; i.e.

$$E = y - Hx$$

Substituting this expression into the objective function, one obtains:

$$E = \sum (y - Hx)^2$$

One can then take the partial derivatives wrt the unknown x_i and find the stationary point:

$$\frac{\partial E}{\partial x_i} = 0$$

For example, consider the problem given by

$$\text{Minimise} \quad \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$$

$$\begin{aligned} \text{Subject to} \quad y_1 &= h_{11}x_1 + h_{12}x_2 + \varepsilon_1 \\ y_2 &= h_{21}x_1 + h_{22}x_2 + \varepsilon_2 \\ y_3 &= h_{31}x_1 + h_{32}x_2 + \varepsilon_3 \end{aligned}$$

One can now substitute values of ε_i into the objective function to minimise

$$E = (y_1 - h_{11}x_1 + h_{12}x_2)^2 + (y_2 - h_{21}x_1 + h_{22}x_2)^2 + (y_3 - h_{31}x_1 + h_{32}x_2)^2$$

The resulting objective function is not constrained; it is a function of only the unknowns x_1 and x_2. Here, y and H are given. Thus, to find the optimum value for x by elementary calculus, i.e.

$$\frac{\partial E}{\partial x_i} = 0$$

$$2h_{11}e_1 + 2h_{21}e_2 + 2h_{31}e_3 = 0$$

$$2h_{12}e_1 + 2h_{22}e_2 + 2h_{33}e_3 = 0$$

putting in the matrix form, one obtains

$$\begin{pmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{33} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$H^T e = 0$$

This set of equations is called normal equations. The first equation can be interpreted as the dot product of first column vector of H with the error vector, i.e., $h_{c1}.e = 0$, similarly $h_{c2}.e = 0$. Thus, the error vector is normal or perpendicular to the column vector of matrix H. The normal equation can be rewritten in matrix form as

$$H^T [y - Hx] = 0$$

$$H^T y - H^T Hx = 0$$

$$H^T y = H^T Hx \quad \Rightarrow X = [H^T H]^{-1} H^T y$$

The authors started originally with $y = Hx$, then artificial variables were introduced, which were eliminated to obtain the least square solution. Hence, the above equation can be thought of as giving least square solution X_{LS} as

$$X_{LS} = H_{LS}\, y$$

where $H_{LS} = [H^T H]^{-1} H^T$ is the least square inverse matrix which is employed to solve the system matrix in the preceding section.

5. ERROR HANDLING & CHOICE OF GROUND CORRELATION POINT

The registration procedure is an interactive utility which offers the user an excellent facility for choice of GCP–CP pair and checking and discarding error. The user can have a trade-off between the best fit and the minimum error to choose or discard a particular registration. To access the overall quality of the fit and to identify erroneous points, the residual of the registration is calculated as

$$\Delta K_m = k_m - Q(i_m, j_m)$$
$$\Delta L_m = l_m - R(i_m, j_m)$$

Residue

$$r(m) = \sqrt{\Delta K_m^2 + \Delta L_m^2}$$

As the point with high residual may be relocated or simply omitted from the fitting, so the process of computing the polynomials Q and R are interactive. By excluding GCP with high residual and picking GCP with less residual, degree of the polynomial is changed interactively to achieve good registration.

6. ALGORITHM

Step 1. Display the image and its corresponding reference image overlapped.

Step 2. If not registered properly, then repeat.

Step 2.1 Choose the GCPs corresponding to the CPs on reference image.

Step 2.2 Compute residue r_m.

Step 2.3 If (r_m > threshold), then discard the GCP.

Step 2.4 If number of GCPs ≥ 3, then GOTO *step* 3.

Step 3. Establish the polynomial affine transformation.

Step 4. Solve for coefficients using least square method.

Step 5. Apply output-to-input transformation.

Step 6. If not registered properly, then GOTO *step* 2.

Step 7. Stop.

7. RESULTS & DISCUSSION

The entire process of image registration can be visualised from the results as shown in Figs 1–4. A raster image of geographical area and its corresponding reference digital vector map is displayed overlapped in Fig.1. It is evident that these are not properly aligned. Figure 2 depicts the choice of three pairs of GCPs and their corresponding CPs on the raster image and the reference digital vector map, respectively. Figure 3

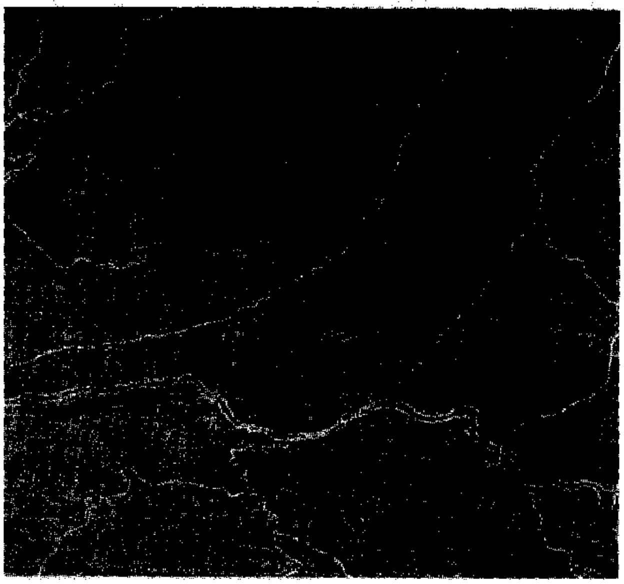


Figure 1. Digital vector map with its corresponding image of the same area displayed overlapped.

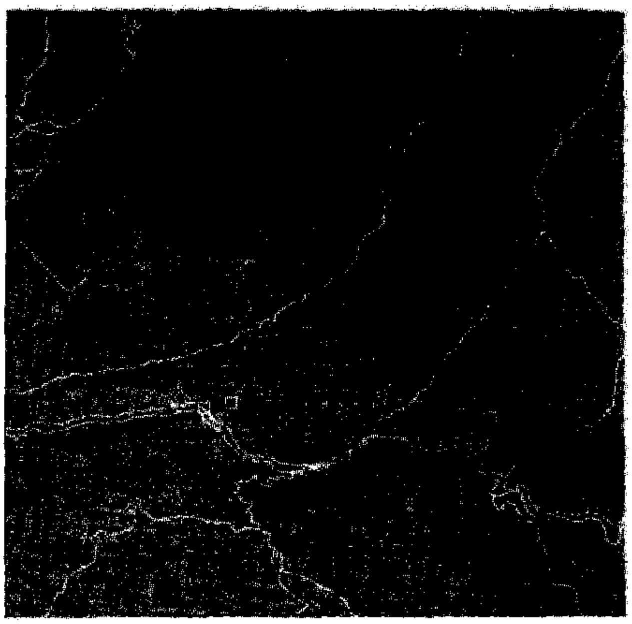


Figure 2. Control point on digital vector map and its corresponding ground correlation point on the image is depicted in the form of small squares. Middle of the square has the prominent landmarks on both the images.

depicts properly registered image with its reference vector map after applying PAT. Finally, Fig. 4 depicts a slightly improper alignment of the same set of image and the vector map.

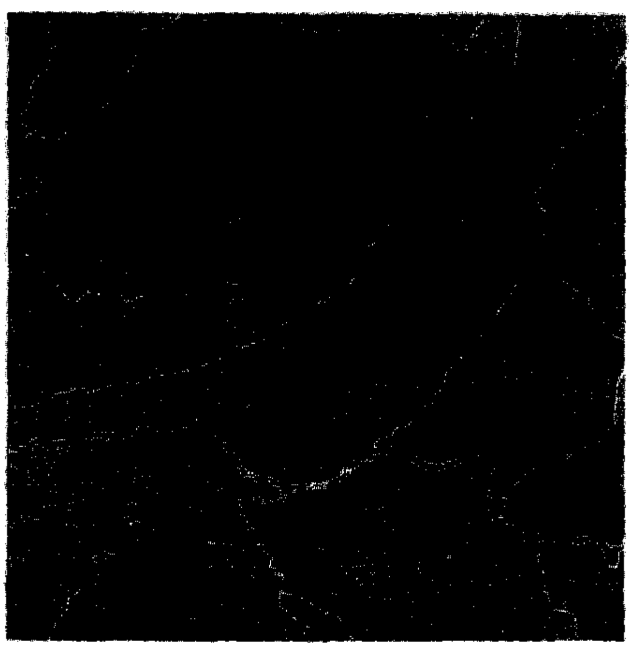


Figure 3. Properly registered image with its digital vector map after application of PAT.

The algorithm developed in section 6 was implemented for satellite image registration and for images captured through various remote sources. This algorithm takes a distorted image containing geometric error and its corresponding reference digital vector map as input. After displaying these images overlapped, the GCP–CP pairs are carefully chosen using mouse pointers and PAT is established. The PAT is applied to the distorted image to get a proper alignment with the reference map, thus removing the geometric errors. This process is known as image registration, and each registration is observed visually for a satisfactory alignment. The process is stopped once proper registration is achieved. The software has been developed using C-language and X-Windows (X11R6) and tested on HP-9000/720 workstation.

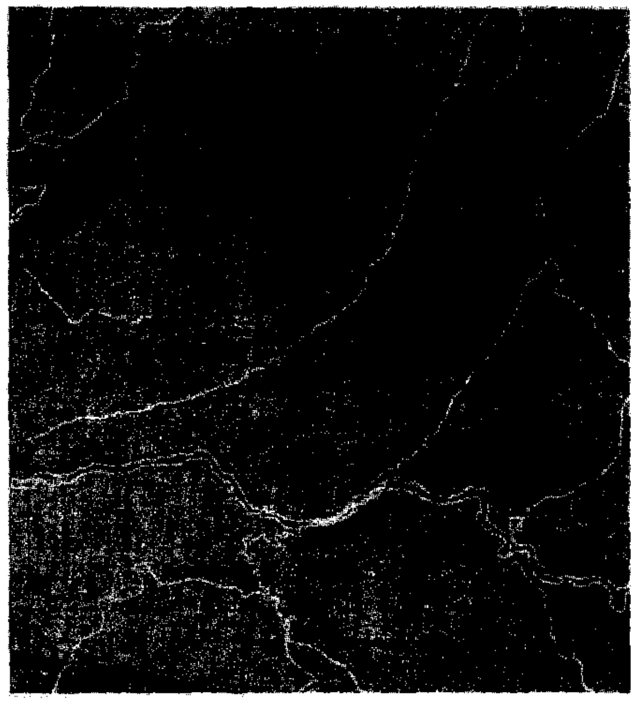


Figure 4. Set of less accurately registered pair of images after application of PAT.

REFERENCES

1. Enderle, G.; Kansy, K.; & Pfaff, G. Computer graphics programming. Springer Verlag, New York, 1984.

2. Shlien, S. Geometric correction, registration, and resampling of landsat imagery. *Canadian J. Remote Sensing*, 1979, **5**, 74–89.

3. Foley, J.D. & Dam, A.V. Fundamentals of interactive computer graphic. Addision-Weslley, Philippines, 1982.

4. Pratt, William K. Digital image processing, Ed. 2. John Wiley & Sons, 1991.

5. Gomes, Jonas, *et al.* Warping and morphing of graphical objects. Morgan Kaufmann Publishers, California, 1999.

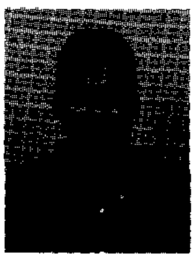
6. Niblack, Wayne. *In* An introduction to digital image processing. Prentice Hall International, New York, 1986. pp.129–63.

7. Richards, John A. & Jia, Xiuping. Remote sensing digital image analysis: An introduction. Springer, 1998.

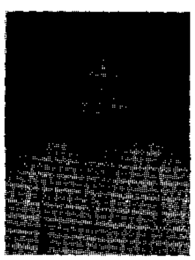
8. Orit, F. Optimal distribution of control points to minimise landsat image registration error. *Photogramm. Engg. Remote Sensing*, **47**, 101–10.

9. Turner, R.E. & Spencer, M.M. Atmospheric model for correction of spacecraft data. *In* Proceedings of the 8th International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan. pp. 895–34.

Contributors



Mr Narayan Panigrahi obtained his MSc (Computer Science) from the J.K. Institute of Applied Physics and Technology, the University of Allahabad in 1991 and MTech (Computer Science & Data Processing) from the Indian Institute of Technology, Kharagpur, in 2000. He joined at the Institute for Systems Studies & Analyses (ISSA), Delhi, in 1991 as Sci B. His areas of research include: Design and analysis of algorithms, computational geometry, image processing, wargaming and simulation, and geographical information system. Currently, he is working at the Centre for Artificial Intelligence & Robotics (CAIR), Bangalore.



Mrs Smita Tripathy obtained her MTech (Electronics and Telecommunication) from the J.K. Institute of Applied Physics and Technology, University of Allahabad in 1995. Her areas of research include: Signal processing and avionics system design. She is currently working as Scientist C at the Aeronautical Development Agency (ADA), Bangalore.