Jñānābha, Vol. 35, 2005

GEORECTIFICATION TECHNIQUE FOR REMOTELY SENSED HIERARCHICAL DATA IN GIS

By

N. Badal

Department of Computer Sciences & Engineering

Kamla Nehru Institute of Technology, Sultanpur-228118, Uttar Pradesh, India

e-mail : n_badal@hotmail.com

Krishan Kant

Department of Computer Sciences & Engineering

Motilal Nehru National Institute of Technology, Allahabad-211002,

Uttar Pradesh, India

e-mail : kk@mnic.ac.in

and

G.K. Sharma

Departmetn of Computer Science & Engineering

Indian Institute of Information Technology and Management,

IIITM, Gwalior-474005, Madhya Pradesh, India

e-mail : gksharma58@yahoo.co.in

(Received : Januwary 7, 2005)

ABSTRACT

HDF-EOS is a standard data format for Remotely Sensed Indian EOS program data products. Data stored in HDF-EOS swath structure is in raw sensor coordinate system and is not rectified to a map projection or geographical coordinate system. The geolocation information is stored as separate fields in a swath. Presently, no remote sensing image processing system has the capability to directly perform georectification of HDF-EOS swath data. Most even do not import such data. An Algorithimic and Mathematical approach has been developed to georectify HDF-EOS swath data. A description of algorithms is presented in this paper.

Key Words: Hierarchical Data Format(HDF), Earth Observing Systen(EOS), Geographic Information System (GIS)

1. Introduction. A computer provides the means to interact with data in ways that are impossible with printed information. Data can be changed, retrieved, recomputed, and displayed not just as a reflection of new information that has been collected from the real world, but also in response to numerical model of spatial and temporal processes.When geographical data are entered into a computer it would be very convenient if the GIS recognized the same conceptual models the user is accustomed to. This would allow interaction between the users and the computer to be both possible and intuitive. In addition GIS must display the data using graphical components and forms, such as color, symbols, shading, and line type and width that are interpretable by people. Therefore as a result of the nature, volume, and complexity of geographical data a number of logical computer schemata have been developed to pack data onto devices for efficient storage, updating, and retriever. These schemata are a continuance of the model/structure development process

The Hierarchical Data Format (HDF) is a widely used scientific data format because of its portability and multiple data model support. Earth Observing System (EOS) project extended HDF to HDF-EOS by adding three new EOS specific data models – point, swath, and grid. HDF-EOS is the standard format for EOS data and products. EOS project has since produced tremendous amount of data in HDF-EOS format and these data are highly demanded by a broad range of research and application communities. Among the three EOS specific data models, swath and grid are two- or multi-dimensional spatial data models while point is discrete point data model. The grid model is used for georectified data and grid data can be easily registered with other spatial data sets. Swath data are not georectified and cannot be combined directly with other georectified spatial data. Presently, no remote sensing image processing software and geographic information system (GIS) have georectification capability for HDF-EOS swath data. Most even do not import HDF-EOS data into their systems. This has significantly limited the public use of EOS products. Providing a generalized, efficient georectification software to the EOS data user community will greatly enhance user's ability to make use of the vast EOS data resources. A software package has been developed to perform georectification of HDF-EOS data. The software can be used to subset, resample, and georectify HDF-EOS swath data. Two georectification algorithms, bivariate polynomial fitting and piece-wide linear interpolation, are implemented in the GIS based software. This paper describes the two algorithms and implementation procedures, and demonstrates experiment results with different remotely sensed data.

2.HDF-EOS SWATH DATA MODEL. A swath structure contains two primary components: data fields and geolocation fields [1]. Data fields are the scientific data stored as two- or multi-dimensional arrays. Geolocation fields contain data representing the positions on the earth surface of the data points in data fields. Usually geolocation fields are composed of a latitude field and a longitude field. The geolocation fields and data fields are linked through common dimensions. For example, a swath may contain a data field, Reflectance, which is a three-dimensional array. The three dimensions of the data array are named as track, cross-track, and band. The geolocation fields in the swath, Latitude and Longitude, are two-dimensional arrays with their two dimensions named as track and cross-track. Thus, the latitude and longitude coordinates of a pixel in the Reflectance field, located at (track=i, cross-track=j), can be obtained by finding the values in the geolocation fields of the array elements located at (track=m, cross-track=n), where position (m,n) in geolocation fields corresponds to position (i,j) in the data field. The correspondence between (i,j) and (m,n) is defined by dimension mapping in the swath, which is described using an Offset and an Increment. Figure 1 shows a mapping relationship between data and geolocation dimensions. The offset and increment values in this example are 1 and 3, respectively.

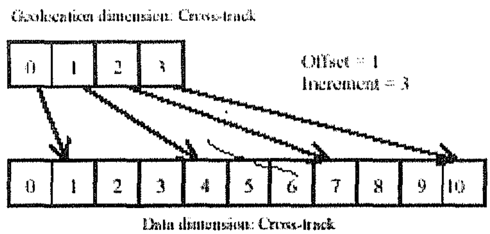

Figure 1 Dimension mapping in HDF-EOS swath

3.GEORECTIFICATION ALGORITHMS. The process of generating a remote sensing image in a map projection or earth coordinate system usually includes two steps. The first step is to determine the map projection or earth coordinate values for part or all pixels of an image and the second step is to transform the image from raw image coordinate system to map projection or earth coordinate system. We call the first step geolocating and the second step georectifying. The product of the first step is a geolocated image on which the map projection or earth coordinate values for part or all pixels are know but the image is still in image coordinate system (i.e., row/column) and the pixel shapes and sizes are different at different locations of the image. The product of the second step, a georectified image, is in a map projection or earth coordinate system and all pixels in the image have identical geometrical shape and size. The two steps are collectively referred as geometrical correction. There are two main approaches to geometric correction in literature: a parametric approach which utilizes sensor parameters and platform ephemeris data and a non-parametric approach which involves the establishment of mathematical relationships between pixel locations in remotely sensed images and the corresponding coordinate values of these pixels on the earth surface [2][3]. In fact, these two approaches are essentially for the two aforementioned steps of geometrical correction. The emphasis of the parametric approach is to derive geolocation information for all or a large part of pixels thus the rectification can be conducted through simple transformation between image and map/earth coordinates. The focus of the non-parametric approach is to perform coordinate transformation using relatively few tie points, often referred as ground control points (GCPs).

HDF-EOS swath data are geolocated and thus only the georectification step is needed to transform the swath data field(s) from raw image coordinate to map projection or earth coordinate. Traditionally, this transform is performed by using bivariate polynomials, e.g., [4][5], in the following form:

$$P' = a_0 + a_1p + a_2q + a_3p^2 + a_4pq + a_5q^2 + \ldots$$

$$Q' = b_0 + b_1p + b_2q + b_3p^2 + b_4pq + b_5q^2 + \ldots$$

where p and q are the pixel coordinate values in the georectified image and P' and Q' are corresponding positions in the original image coordinates (i.e., image column/row values). ai and bi (i=0,1,2,...) are coefficients derived from regression between image and map/earth coordinate values of GCPs.

This regression approach is effective and accurate when geometric distortions can be modeled by low order polynomials (usually two- to four-order) and adequate number of GCPs is available. For data with significant geometric distortions, such as airborne images, global polynomial fitting usually cannot generate satisfactory results. A remedy to this problem is using piecewise rectification approaches in which polynomial regressions are performed locally and transform piece by piece the original image to map/earth coordinate using locally derived fitting functions [5][6].

We implemented the global bivariate polynomial regression algorithm in our software. The algorithm produced satisfactory results for data with relatively small distortion but significant errors were observed for images having wide scan angles. Our software has a subsetting capability that allows a user to georectify a portion of a swath. This capability effectively gives a user the piecewise rectification alternative because a larger image can be constructed combining subsets of the image.

Many satellite data products contain very dense geolocated pixels, such as the level 1b products of the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Radiometer (MODIS). For those data, a piecewise linear interpolation algorithm is more appropriate. The algorithm utilizes every available geolocation pixel and produces more accurate rectified images in many cases. The algorithm was first developed to process AVHRR level 1b data stored in half inch, 9-track tapes [7]. It is modified and implemented in the current procedure based on HDF-EOS swath data structure. The algorithm involves a spatial interpolation procedure that transfers pixels of a data field from raw image coordinate to map/earth coordinate, and a data value interpolation that fills blank pixels in the map/earth coordinate. The value interpolation is needed because the pixel size in the georectified data (output image) is usually selected to be the nominal pixel size of a sensor and off-nadir pixel growth during scanning will result in blank pixels in the output image.

The primary procedures of the algorithm are as the following:

1) On the swath's geolocation fields, find the coordinates, in terms of row/column number, of the four corners of user supplied spatial bounding box. If no bounding box is supplied, it is assumed that the entire image region needs to be georectified. The corner coordinates in this case is the four corners of the geolocation fields.
2) Based on the corner coordinates in geolocation fields and the dimension mapping parameters defined in the swath (i.e., offsets and increments), determine the corresponding corner coordinates in the data field, also in terms of row/column number.
3) Read in one data row inside the bounding box from the data field and the corresponding rows from the geolocations fields. Because geolocation is usually not supplied for every data row (i.e., the row direction increment value is greater than 1), two geolocation rows need to be read in.
4) Determine the map/earth coordinate values of every pixel in the read-in data row. If there is a tie between a data pixel and a geolocation pixel (e.g., data pixel 1 and geolocation pixel 0 in fi gure 1), the data pixel is directly transformed to the output image, otherwise a bilinear interpolation is performed to derive the pixel's map/earth coordinate values.
5) Completing the spatial interpolation for all data inside the bounding box by repeat steps (3) and (4).
6) Perform data value interpolation for blank pixels in the output image. This interpolation can be performed using nearest neighbor, linear, bilinear or cubic convolution.

4.EXPERIMENTS. Tests of the bivariate polynomial algorithm data indicate that the algorithm works perfectly (error < 1 pixel) for entire ASTER scenes with approximately 60km swath width. However, significant errors exist when the same algorithm is applied to full MODIS scenes with approximately 2330km swath width of remotely sensed GIS data. Satisfactory results can be reached only when an entire scene is divided into 16 to 36 pieces (i.e., regressions are performed on 1/4 to 1/6 along both row and column directions) depending on the location of a subscene in the full scene of GIS image.

The piecewise linear algorithm produces satisfactory results for both MODIS and ASTER scenes. Because this algorithm directly transfers pixels from original image to output image using map/earth coordinate values either obtained or interpolated from geo-tie points, there is no positional error for pixels located at geo-tie points. The error of a non-tie point is introduced by the non-linearity of pixel distortion. The magnitude of this non-linearity error changes with the density of geo-tie points and the distance to the nadir point (or scan angle). The error increases toward the edge of a scan. In most cases, the error is well bellow a half pixel size. For example, in one of our 1km resolution MODIS data, the longitudes of the first three geo-tie points located at the upper left corner of the swath, where the maximum distortion occur, are -117.87901, -117.65004, and -117.32246. The dimension Increment value of this swath is 5 (i.e., the above three longitudes are for the first, the sixth and the eleventh pixels in the data fields). The pixel sizes after linear interpolation are 0.066932 degrees for second through the fifth pixels and 0.055166 degrees for the seventh through tenth pixels (i.e., the average pixel sizes between the geo-tie points). Apparently, positional error introduced using these average pixel sizes is smaller than one tenth of a pixel. Similar observations can be drawn from other data sets such as AVHRR [7].

5.CONCLUSIONS. The georectification technique provides a useful tool for rectifying HDF-EOS swath data. The bivariate polynomial regression algorithm is suitable for swath data with small swath width and relatively few geo-tie points such as ASTER data. When combined with the software's subset capability, the algorithm can also be used to rectify images with large geometrical distortions, provided that enough GCPs are presented in each subset of the image. The piecewise linear algorithm produces no positional error at geo-tie points and the maximum error for non-tie points is well bellow a half of pixel sizes for many remotely sensed GIS data. This algorithm is especially effective for swath data with dense geo-tie points such as MODIS level 1b data.

REFERENCES

[1] Kvamme, K.L. A Predictive Site Location Model on the High Plains: An Example with an Independent Test. *Plains Anthropologist* 37 (138): 19-40, 1992 .

[2] D. P. Roy, B. Devereux, B. Grainger , and S. J. White, "Parametric geometric correction of airborne thematic mapper imagery", *Int. J. Remote Sensing*, vol. 18, pp.1865-1887, 1997.

[3] A. J. De Leeuw, L. M. M. Veugen, and H. T. C. Van Stokkom, "Geometric correction of remotely-sensed imagery using control points and orthogonal polynomials", *Int. J. Remote Sensing*, vol. 9, pp.1751- 1759, 1988.

[4] N. K. Kardoulas, A. C. Bird, and A. I. Lawan, "Geometric correction of SPOT and Landsat imagery: a comparison of map- and GPS-derived control points", *Photogrammetric Engineering and Remote Sensing*, vol. 62, pp.1173-1177, 1996.

[5] M. Ji and J. R. Jensen, "Continuous piecewise geometric rectification for airborne multispectral scanner imagery", *Photogrammetric Engineering and Remote Sensing*, vol. 66, pp.163-171, 2000.

[6] E. J. Christensen, J. R. Jensen, E. W. Ramsey, H. E. Mackley, Jr., "Aircraft MSS data registration and vegetation classification for wetland change detection", *Photogrammetric Engineering and Remote Sensing*, vol.53, pp.521-529, 1988.

[7] L. Di and D. C. Rundquist, "A one step algorithm for correction and calibration of AVHRR level 1b data", *Photogrammetric Engineering andRemote Sensing*, vol.60, pp.165-171, 1994.