A Technique for Spatial Sampling and Error Reporting for Image Map Bases

Albert K. Chong

Abstract

This paper presents a technique for spatial sampling and error reporting for image base maps. The technique is based on the coverage of an image map base, the initial estimated accuracy, and the principle of error propagation to determine a number of check points. The location of these check points is randomly generated to obtain a non-biased evaluation of the overall image map base. A spatial error modeling formula is introduced to estimate the size of error at various locations over the image map base. Data from a recent orthoimage accuracy assessment project is used to show the procedure.

The implementation of this technique has resulted in an improvement in the checking procedure of image map base accuracy assessment.

Introduction

Digital image map bases (DIMBs) are geometrically rectified and geocoded raster images with minimal cartographic overlay (Offermann, 1993; Lerner and Denegre, 1994). It is common for a DIMB to contain one, or many, types of imagery. These imageries may have been acquired from satellite, aerial, or terrestrial platforms. The characteristics and size of errors associated with each type of imagery are different (Richards, 1986; Folving and Denegre, 1994). Having two or more types of imagery in an image map base can produce significant variation in the positional accuracy across the entire map base (Thapa and Bossler, 1992).

While many techniques are available to assess the spatial accuracy of cartographic maps and the data layers of GISs, an appropriate approach is needed for DIMBs. The purpose of this paper is to present a spatial accuracy assessment technique, which reflects the characteristics of errors of DIMBs.

Positional Accuracy Assessment

Positional accuracy is often expressed in two components: absolute and relative positional accuracy (Stanislawski *et al.*, 1996). Absolute positional accuracy addresses how closely all the positions on a map or data layer match corresponding positions of the features they represent on the ground in a desired map projection system (Stanislawski *et al.*, 1996). The relative positional accuracy of a map considers how closely all the positions on a map or data layer represent their corresponding geometrical relationships on the ground. Although absolute positional accuracy can be important and may also have a direct influence on the relative positional accuracy within a DIMB, only the latter is discussed in this paper.

A number of standards have been developed to provide for positional assessment of cartographic maps (Merchant, 1982; Vonderohe and Chrisman, 1985; Merchant, 1987; ACSM, 1988; ASPRS, 1990; Crosilla and Pillirone, 1995). A variation to the ASPRS standards are provided in Acharya and Bell (1992) and Ackermann and Rad (1996). Skidmore and Turner (1992) proposed a line intersect sampling technique which is directed at assessing land-cover class boundaries on cartographic land-cover maps. In 1998 the Subcommittee for Base Cartographic Data, Federal Geographic Data Committee published a National Standard for Spatial Data Accuracy (NSSDA).

There are a few factors to consider when one applies the above standards to DIMBs because they do not have similar spatial characteristics. For example, DIMBs are multi-scale (Loodts, 1993). A DIMB may cover a portion of a regular map sheet (e.g., 1:25,000-scale national map series) or they may cover many map sheets. In neither case do they make reference to any regular map series boundaries. Each image pixel in a DIMB is a unique real world feature on the ground (Baltsavias, 1996). Additionally, the positional error distribution and error size of a DIMB may vary throughout the map base due to a mixture of imageries used, large variations in terrain, and any additional residual errors in geometric rectification and geocoding (Richards, 1986; Folving and Denegre, 1994; Baltsavias, 1996).

This paper presents the detail of a positional accuracy assessment technique to provide testing and reporting of DIMBs. This approach can also complement other published techniques. A working example is provided for each step of the evaluation procedure.

Methodology

Check Point Sampling

Simple random sampling is a fundamental scientific selection method (Husch *et al.*, 1982). All other sampling procedures in the sciences are modifications of simple random sampling that are designed to achieve greater economy or precision (Husch *et al.*, 1982). The basic formula provided in Husch *et al.* (1982) is modified as follows:

$$n = \frac{t^2 s^2}{E^2}, \text{ or } n = \frac{t^2 c^2}{(E\%)^2} \tag{1}$$

where

n is the number of check points required for a DIMB;
t is the Student's "t" for the desired 95% probability level;
s is the standard deviation of the positional error factor within

Department of Geomatics, University of Melbourne, Parkville, Victoria 3052, Australia.

The author is presently with the Department of Surveying, The University of Otago, Dunedin, New Zealand (chonga@albers.otago.ac.nz).

Photogrammetric Engineering & Remote Sensing, Vol. 65, No. 10, October 1999, pp. 1195–1198.

0099-1112/99/6510–1195$3.00/0

TABLE 1. QUALITY PARAMETERS AND QUALITY STANDARDS FOR PHOTOGRAMMETRIC PROCESSES

Process	Output	Quality Parameters	Quality Standards
Field Survey	ground control points	RMS	0.10 m
Aerotriangulation	model control points	RMS	10–20 μm
Geometric rectification and Geocoding	interior orientation	mean residual	6 μm
		max. residual	10 μm
	relative orientation	mean parallax	6 μm
		max. parallax	10 μm
	absolute orientation	mean residuals	20 μm
		max. parallax	33 μm

a sample block of aerial photographs or within a block of satellite imagery and is expressed as σ_{dev} in Equation 3 below;

E is the allowable standard deviation of positional error of image features within a DIMB or the required confidence interval;

c is the coefficient of variation $= \frac{s}{\overline{x}}(100)$, where $\overline{x} = \frac{\sum_{i=1}^{n} X_i}{n}$ is the mean of the positional error of all image features within a stereomodel; and

$E\%$ is the allowable standard deviation of positional error of image features as a percentage of the mean positional error of image features within a DIMB.

The modification is made to satisfy the characteristics of DIMBs. Each pixel of a DIMB is essentially an individual, geometrically rectified and geocoded image pixel originated from aerial photographs, satellite imagery, or terrestrial photographs. Ackermann and Rad (1996) provide a list of quality parameters and quality standards for photogrammetric processes involving aerial photographs and satellite imagery (Table 1). The quality values quoted are consistent with present technology and are achievable in an analytical or digital photogrammetric environment. Consequently, it is appropriate to use this as the quality standard to determine the positional error of image features of geometrically rectified aerial photographs or geometrically rectified satellite imagery for large and medium scale map bases.

The value $\overline{x}$ can be computed by using the quality standards and the principle of error propagation for independent observations (Mikhail and Gracie, 1981; pp. 154-155). For example, the computation of a typical stereomodel may be as follows:

$$\overline{x} = \sigma_{total} = (\sigma^2_{io} + \sigma^2_{ro} + \sigma^2_{ao} + \sigma^2_{aero} + \sigma^2_{gd})^{1/2} \text{ micrometers} \quad (2)$$

where

σ_{total} is the total positional error budget for each image feature,

σ^2_{io} is the variance of interior orientation in micrometers,

σ^2_{ro} is the variance of relative orientation in micrometers,

σ^2_{ao} is the variance of absolute orientation in micrometers,

σ^2_{aero} is the variance of aerotriangulation in micrometers, and

σ^2_{gd} is the variance of error vector of ground survey expressed in micrometers.

σ^2_{gd} may be left out of Equation 2 because the variance of aerotriangulation is determined from the variance of ground survey errors.

Based on the assumption that errors vary from model to model due to error in aerotriangulation, terrain variation, and so forth, the value s can be computed as follows (Ackermann and Rad, 1996):

$$s = \sigma_{dev} = (\sigma^2_{aero} + \sigma^2_{terrain})^{1/2} \text{ (in micrometers)} \quad (3)$$

where $\sigma^2_{terrain}$ is the variance of terrain variation expressed in micrometers.

The coefficient of variation, c, may now be determined as a percentage of σ_{dev} over σ_{total}: i.e.,

$$c = \left(\frac{\sigma_{dev}}{\sigma_{total}}\right) \times 100 \text{ (in micrometers).} \quad (4)$$

Next, the $E\%$ can be determined according to the allowable standard error as a percentage of the mean error or it may be a pre-specified allowable error fraction. For example, if the mean positional error of an image feature is expressed as $\overline{X}_{err}$ (pixels or meters in ground units) and the permissible error variation is σ_{image}, then $E\%$ is calculated as follows:

$$E\% = \frac{t\sigma_{image}}{\overline{X}_{err}} \times 100 \quad (5)$$

where t is the Student's 't' for the desired 95 percent probability level (two-tailed test). The value of $E\%$ is a measure of the desired precision of sampling of the image map base, i.e., a high permissible error variation requires fewer sampling points. Table 2 provides a list of (n) for various common values of $E\%$ based on the Ackermann and Rad (1996) quality standards.

A Numerical Example for a Large Sample Size

Given (from Table 1)

$\sigma_{io} = 6$ μm; $\sigma_{ro} = 6$ μm, $\sigma_{ao} = 25$ μm, $\sigma_{aero} = 10$ μm,
$\sigma_{gd} = 7$ μm, $\sigma_{terrain} = 2$ μm,
$\overline{X}_{err} = 1.0$ m, and $\sigma_{image} = 0.06$ m.

Then, using the principle of error propagation (Equations 2 and 3), $\sigma_{total} = 29$ μm and σ_{dev} 10 μm.

Using Equation 4, $c = \frac{10}{29} \times 100 = 34$.

Next, using Equation 5, $E\% = \frac{1.96 \times 0.06\text{m}}{1.0\text{m}} \times 100\% = 12\%$.

The number of samples, assuming 't' = 1.96 (i.e., $n > 30$), is

$$n = \frac{(1.96)^2(34)^2}{(12)^2} = 31.$$

A Numerical Example for a Small Sample Size

Again, given $c = 25$ and $E\% = 14\%$, then the first estimate of 't' at $n > 30$ is 1.96,

$$n = \frac{(1.96)^2 \times (25)^2}{14^2} = 12, \text{ and}$$

the second iteration using 't' at $n = 12$, and from the Students 't' table, 't' = 2.18, and

$$n = \frac{(2.18)^2 \times (25)^2}{14^2} = 15.$$

TABLE 2. COMMON E% AND THE CORRESPONDING SAMPLE SIZE

E%	5	6	7	8	9	10	11	12	13	14	15
n_o									26	23	20
n	177	123	91	69	55	44	37	31	29	25	22

n_o is the initial estimate based on 't' at $n > 30 = 1.96$.

Check Point Selection Procedure

This new technique is ideal for the positional accuracy assessment of GIS image map bases because random point sampling is a simple GIS function achieved by using "Sample" in ARC/INFO GRID. Initially, the number of points required for a map base must be determined based on Equations 1. The suitability of the selected image pixel for identification on the ground is confirmed on the image map base. Any feature that is not identifiable on the ground is discarded and another image pixel is selected. Occasionally, an image may have a significant coverage of water or wet lands where there is no identifiable feature in the field. In such a case, a set of parameters may be set in the sampling GIS software to exclude the affected area.

Error Modeling and Error Analysis Technique

An error propagation algorithm used in photogrammetric analytical strip aerotriangulation (Ghosh, 1975; p. 70) is simplified to suit the error characteristics of image map bases. The modified algorithm is expressed as follows:

$$\delta_x = a_0 + a_1 (x - x_0) + a_2 (y - y_0) \tag{6a}$$

$$\delta_y = b_0 - a_2 (x - x_0) + a_1 (y - y_0) \tag{6b}$$

$$\delta_z = c_0 + c_1 (x - x_0) + c_2 (y - y_0) \tag{6c}$$

where δ_x, δ_y, and δ_z are coordinate differences in x, y, and z or height; a_0, a_1, a_2, b_0, c_0, c_1, c_2 are coefficients; x and y are image map base coordinates of each check point; and x_0, y_0 are coordinates of the centroidal point of the image map base.

The difference between map base coordinates and surveyed ground coordinates in x or in y can be modeled adequately by the above simple linear equations. Tests show that the use of high-order polynomials may not improve the quality of error modeling. For simplicity, the coefficients a_1 and a_2 are the same for the x and y equations because the scaling errors along the x and y axes are similar. However, the assumption may not hold true in the z axis because terrain variation (or slope) can result in different scaling errors along the x and y horizontal axes. Accordingly, a separate set of coefficients are used for the equation which models the z errors.

Next, the coordinate differences must be statistically tested for gross errors before they are used. Gross errors may be present in the image map base coordinates and in the ground coordinates. One source of error could be the inclusion of data from independent surveys. Any gross errors detected must be excluded from the computation of the coefficients. A useful statistical gross error detection method known as the TAU test was introduced by Pope (1976). This technique is available in many commercial surveying network adjustment packages and photogrammetric aerotriangulation software. Another simple gross error detection technique is the three-sigma (3σ) rule, which means that any observation having a residual larger than three sigmas (σ) is removed from the sample (Philip, 1994). For simplicity one can use the latter.

Computation of the Coefficients of Equation 6

A technique for solving the unknowns of a set of linear equations is provided in Moffitt and Mikhail (1980; pp. 606–616). This technique can be readily implemented in a GIS. For example, the linear regression command in ARC/INFO GRID can be modified for this purpose. In addition, Windows95 Excel mathematical functions are also available for the computation of the coefficients of a set of linear equations. To compute the coefficients, one must determine the variables for the equations. For example, the first check point (P) in Table 3 will have the following variables in a set of three equations:

in x, $-0.2 = a_0 + a_1 (384{,}170 - 384{,}910) + a_2 (726{,}386 - 726{,}800)$,
in y, $-0.1 = b_0 - a_2 (384{,}170 - 384{,}910) + a_1 (726{,}386 - 726{,}800)$,
in z, $2.5 = c_0 + c_1 (384{,}170 - 384{,}910 + c_2 (726{,}386 - 726{,}800)$,

where $x_p = 384{,}170$, $y_p = 726{,}386$, $x_0 = 384{,}910$ E, and $y_0 = 726{,}800$ N.

The Inverse Process

After the coefficients of Equation 6 are computed, it is possible to compute the value of the error at any given location on the image map base. The technique is good for any separated "partial" image map base as long as the coefficients are known. For example, the following computed coefficients and the coordinates of the centroidal point of an image map base are available: i.e.,

a_0 0.45 m, $b_0 = 0.44$m, $a_1 = -0.000042$, $a_2 = -0.000033$,
$c_0 = 2.45$ m, $c_1 = -0.00008$, $c_2 = -0.00009$, $x_0 = 384{,}910$ E, and
$y_0 = 726{,}800$ N.

To compute the error of a point at 726,500 N and 383,000 E, substitute the data into Equation 6: i.e.,

$$\delta_x = 0.45 + (-0.0000423)(383{,}000 - 384{,}910) + (-0.000033)(726{,}500 - 726{,}800)\text{ m} = 0.5\text{m}$$

$$\delta_y = 0.44 + (-0.000033)(383{,}000 - 384{,}910) + (-0.000042)(726{,}500 - 726{,}800)\text{ m} = 0.5\text{ m, and}$$

$$\delta_z = 2.45 + (-0.00008)(383{,}000 - 384{,}910) + (-0.00009)(726{,}500 - 726{,}800)\text{ m} = 2.6\text{ m}$$

Therefore, the positional error of point A is 0.5 m in easting (E), 0.5 m in northing (N), and 2.6 m in height, respectively.

Procedure for DIMB Evaluation

The following steps may be used to evaluate an image map base:

(1) Compute the sample size using Equations 1 through 5;
(2) Use the random sampling function to select features from the image map base for field identification and surveying;
(3) Obtain the ground coordinates of the selected features;
(4) Compute the differences between ground coordinates and image map base coordinates for all selected features;
(5) Compute the centroidal coordinates (x_0, y_0 of the image map base;
(6) Compute the coefficients a_0, a_1, a_2, and b_0 of the horizontal error propagation equations (Equations 6a and 6b); and
(7) Compute coefficients c_0, c_1, and c_2 of the vertical error propagation equations (Equation 6c).

Modified Positional Accuracy Reporting Format

DIMBs are multi-scale, which means that they may be used to create map products and to produce statistical data at various image/object ratios. It is essential that the positional accuracy reporting format be accurate and adequate. The modified format may help to eliminate a number of problems in the error modeling and error evaluation of GIS applications. For example, the error of the computed area of two

TABLE 3. DIFFERENCE BETWEEN IMAGE MAP BASE COORDINATES AND GROUND SURVEY COORDINATES

δ_n	−0.1	−0.6	1.9	0.9	−0.3	0.3	0.7	−1.7	−1.1
	−0.4	0.3	1.2	0.4	0.6	2.0			
δ_e	−0.2	0.3	0.5	1.1	2.3	0.9	1.2	0.8	−0.5
	0.4	0.3	−0.1	−0.9	−0.3	1.2			
δ_z	2.5	1.0	1.9	2.6	2.5	2.5	3.2	3.0	0.0
	2.2	1.9	2.3	7.0	3.1	2.8			

similarly sized plots from different locations on the image map base will have different values. Using an overall RMS value, on the other hand, will give similar error results. This format has the following components:

(1) RMS_x, RMS_y, RMS_z, and CMAS (ASPRS, 1990; Goodchild, 1991; FGDC, 1998),
(2) Number of and percentage of gross errors and size of largest gross error, and
(3) Value of the various error modeling algorithm coefficients, i.e.,

$$a_0,\ a_1,\ a_2,\ b_0,\ c_0,\ c_1,\ c_2,\ x_0,\ y_0$$

Conclusion

The proposed approach is statistically sound and is easy to follow. No complicated algorithms and computations are involved.

Generally, high quality image map bases do not require any positional evaluation because automation in most photogrammetric processes ensures high quality products. However, any digital map bases that come from an unknown or suspicious source may require an independent assessment.

Field checks are both laborious and expensive. The recent introduction of advanced real-time On-The-Fly GPS techniques may provide a cost-effective technique for feature identification and accurate positioning of indistinct objects, isolated shrubs, bushes, and trees — an abundant feature in digital images. A practical procedure has been implemented, and it appears to be cost effective.

Acknowledgment

The author wishes to acknowledge the helpful assistance of the *PE&RS* reviewers, and the staff and Head of the Surveying Department. Also, the author would like to thank the local regional council for supplying the project data.

References

Ackermann, R.J., and A. Rad, 1996. Quality control procedure for photogrammetry digital map, *International Archives of Photogrammetry and Remote Sensing*, 31(B4):12–17.

Acharya, B., and W. Bell, 1992. Universal accuracy standards for GIS, *Proceedings of ASPRS/ACSM/RT 92*, Washington, D.C., 3: 368–379.

American Congress on Surveying and Mapping (ACSM), 1988. The proposed standard for digital cartographic data, *The American Cartographer*, 15:9–140.

American Society for Photogrammetry and Remote Sensing (ASPRS), 1990. ASPRS accuracy standards for large-scale maps, *Photogrammetric Engineering & Remote Sensing*, 56(7):1068–1070.

Baltsavias, Emmanuel P., 1996. Digital ortho-images — A powerful tool for the extraction of spatial- and geo-information, *Photogrammetry and Remote Sensing*, 51:63–77.

Crosilla, F., and G. Pillirone, 1995. *Parametric and non-parametric* procedure for testing the metric quality of a digital map, *Manuscripta Geodaetica*, 20:231–240.

Folving, S., and J. Denegre, 1994. *Thematic Mapping from Satellite Imagery*, Pergamon, [place of publication], pp. 48–62.

Federal Geographic Data Committee, 1998. *National Standard for Spatial Data Accuracy (NSSDA)*, FGDC-STD-007.3-1998, Federal Geographic Data Committee, Washington, D.C., 25 p., (**http://fgdc.er.usgs.gov/fgdc.html**).

Ghosh, Sanjib K., 1975. *Phototriangulation*, Lexington Books, Lexington, Massachusetts, 241 p.

Goodchild, M., 1991. *Advances in Cartography* (J. Muller, editor), Elsevier Applied Science, London, 251 p.

Husch, B., C. Miller, and T. Beers, 1982. *Forest Mensuration*, John Wiley & Sons, N.Y., 401 p.

Lerner, J., and J. Denegre, 1994. *Thematic Mapping from Satellite Imagery*, Pergamon, [place of publication], pp. 63–73.

Loodts, J., 1993. Digital orthophotos and GIS: The perfect couple, *Photogrammetric Week'93*, Wichmann Verlag, Stuttgart, pp. 273–283.

Merchant, D.C., 1982. Spatial accuracy standards for large scale maps, *Proc. ACSM 42nd Annual Meeting*, pp. 222–231.

———, 1987. Spatial accuracy specification for large scale topographic maps, *Photogrammetric Engineering & Remote Sensing*, 53(7):958–961.

Mikhail, Edward M., and Gordon Gracie, 1981. *Analysis and Adjustment of Survey Measurements*, Van Nostrand Reinhold Company, New York, 340 p.

Moffitt, Francis H., and Edward M. Mikhail, 1980. *Photogrammetry, Third Edition*, Harper & Row, New York, 648 p.

Offermann, W., 1993. Cartographic finishing and production of image map, *Conference Proceedings, EGIS'93*, Italy, 2:1091–1095.

Philip, M., 1994. *Measuring Trees and Forests*, CAB International, UK, 310 p.

Pope, A., 1976. *The Statistics of Residuals and the Detection of Outliers*, NOAA Technical Report., NOS-65 NGS-1, National Geodetic Information Center, NOAA, 241 p.

Richards, J.A., 1986. *Remote Sensing Digital Image Analysis: An Introduction, Second Edition*, Springer-Verlag, New York, 281 p.

Skidmore, A., and B. Turner, 1992. Map accuracy assessment using line intersect sampling, *Photogrammetric Engineering & Remote Sensing*, 58(10):1453–1457.

Stanislawski, L., B. Dewitt, and R. Shrestha, 1996. Estimating Positional Accuracy of Data Layers within a GIS through Error Propagation, *Photogrammetric Engineering & Remote Sensing*, 62(4): 429–433.

Thapa, K., and J. Bossler, 1992. Accuracy of Spatial Data Used in Geographic Information Systems, *Photogrammetric Engineering & Remote Sensing*, 58(6):835–841.

Vonderohe, A.P., and N.R. Chrisman, 1985. Test to establish the quality of digital cartographic data: Some examples from the Dane County Land Records Project, *Proceedings AUTO-CARTO 7*, pp. 552–559.

(Received 30 December 1996; revised and accepted 03 November 1998)