Accuracy Assessment of Digitized Data with Randomized Block Model

By Wei-Hsin Ho[1] and Kuen-Tzung Chang[2]

Abstract: This study uses randomized block model variance analysis to assess factors affecting the quality of digitized data yielded by a digitizing tablet and desktop scanner. First, the geometric accuracy of a digitizing tablet and desktop scanner and the spectral response of the scanner were examined using a "calibration" reseau and a commercially available grayscale. The average planimetric positional error (μ_p) for the digitizer and the scanner were 0.249 and 0.119 mm, respectively. Experimental results showed that the planimetric position root mean square errors of both devices satisfy map accuracy requirements, and that scanner gray level response was homogeneous. Furthermore, this study used randomized block model variance analysis to assess factors affecting the quality of digitized data yielded by a digitizing tablet and desktop scanner. The effects of digitizing procedures for both devices were evaluated by this analysis. The ANOVA tables indicated that different map positions and widths of lines for a digitizing tablet, and the operator and scanner density for a scanner are shown to important elements in the data acquisition process.

INTRODUCTION

Analog maps are primary sources of data for geographic information systems (GIS). Those maps can be converted from analog to digital format. Low-cost and commercially available devices are used in this study. Among such devices, desktop scanners and digitizing tablets are widely used. However, data derived in this manner lead toward numerous errors. Further analysis in GIS using these data can yield questionable results due to error propagation (Lunetta et al. 1991). Therefore, the effects of propagation factors on accumulated error must be assessed. Also, controlling and maintaining data quality are highly desirable. As a result, more reliable descriptive data and the attainment of more accurate results for decision making are highly desired.

Assessment of geometric accuracy for digitizing equipment has received extensive attention (Lai 1988; Drummond and Bosma 1989; Bolstad et al. 1990; Carstensen and Campbell 1991). However, spatial accuracy represents only one kind of error outcome; many other error sources are produced by data-capture and manipulation procedures (Thapa and Bossler 1992). Assessing these error sources before digitizing can allow for data quality to be controlled and, ultimately, reliable products can be made.

A time-series analysis has been performed to measure errors in stream-mode digitizing; however, this approach is statistically unreliable (Brenton et al. 1991). Fernãndez et al. (1991) first derived an analysis method of variance to evaluate the statistical accuracy of coordinates retrieval from a U.S. Geological Survey (USGS) 7.5-min series map using a digitizing tablet (Fernãndez et al. 1991). Variance analysis is a useful technique in identifying statistically significant differences between sources of uncertainty. In this

[1]Assoc. Prof., Dept. of Land Economics, Nat. Cheng Chi Univ., Taipei, Republic of China, 11600.

[2]PhD Student, Dept. of Civ. Engrg., Nat. Chiao Tung Univ., Hsinchu, Taiwan, Republic of China, 30050.

Note. Discussion open until January 1, 1998. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on October 16, 1995. This paper is part of the ***Journal of Surveying Engineering***, Vol. 123, No. 3, August, 1997. ©ASCE, ISSN 0733-9453/97/0003-0087–0100/$4.00 + $.25 per page. Paper No. 11838.

study, we evaluate the effects of digitizing procedures for a digitizing tablet and desktop scanner using the randomized block model (RBM), which is commonly employed in variance analysis. An outline of RBM theory is given in the next section.

VARIANCE ANALYSIS USING RANDOMIZED BLOCK MODEL

Variance analysis is a highly effective technique that is capable of analyzing and interpreting several populations. This versatile tool partitions the total variation in a data set according to the sources of variation that are represented. The sources of variation, i.e., nonmetric independent variables that are referred to as "factors," are verified during the experimental procedures. The factorial effects are examined by forming groups based on all possible combinations of the levels of the various factors. Those combinations are referred to as subjects or treatments in experiments. Similar to the pairing of like subjects or experimental units called a "block" to improve upon the procedure of taking two independent samples, subjects can also be arranged into homogeneous groups of size b (the number of levels of a block) when comparing k factors. Then, if each factor is applied to exactly one unit in the block and if comparisons are only drawn between factorial responses from the same block, extraneous variability should be markedly reduced. This is the concept underlying the RBM (Richard and Gouri 1992).

The data, once obtained, can be arranged in rows according to the treatments and can be arranged in columns according to the blocks. By designating the measurement corresponding to treatment i and block j by y_{ij}, Table 1 displays the data structure of a randomized block design with b blocks and k treatments. The row and column means are shown in the rightmost column and bottom row of the table. An overbar on y indicates an average, and a dot in the subscript denotes that the average has been taken over the subscript appearing in that place. The sum of squares due to factorial deviations (S_f) is listed as

$$S_f = \sum_{i=1}^{k} \sum_{j=1}^{b} (\bar{y}_{i.} - \bar{y}_{..})^2 = b \cdot \sum_{i=1}^{k} (\bar{y}_{i.} - \bar{y}_{..})^2 \quad (1)$$

The number of distinct entries in the treatment-effects array = k; and the single constraint is that they must sum to zero. Thus, $(k - 1)$ degrees of freedom (DOFs) are associated with treatment sums of squares. In the same

TABLE 1. Data Structure of Randomized Block Design with *b* Blocks and *k* Treatments

Treatment (1)	Block 1 (2)	Block 2 (3)	Block — (4)	Block — (5)	Block — (6)	Block b (7)	Treatment means (8)
1	y_{11}	y_{12}	—	—	—	y_{1b}	$\bar{y}_{1.}$
2	y_{21}	y_{22}	—	—	—	y_{2b}	$\bar{y}_{2.}$
—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—
k	y_{k1}	y_{k2}	—	—	—	y_{kb}	$\bar{y}_{k.}$
[Block means]	$\bar{y}_{.1}$	$\bar{y}_{.2}$	—	—	—	$\bar{y}_{.b}$	$\bar{y}_{..}$

Note:—represents omitted items.

TABLE 2. ANOVA Table for Randomized Block Design

Source (1)	Sum of squares (2)	Degree of freedom (3)	Mean squares (4)	F-ratio (5)
Factors	S_f	$k - 1$	M_f	M_f/M_r
Blocks	S_b	$b - 1$	M_b	M_b/M_r
Residual	S_r	$(b - 1)\cdot(k - 1)$	M_r	—
[Total]	S_t	$b\cdot k - 1$	—	—

manner, block sums of squares (S_b) and residual sums of squares (S_r) are implied as in (1). Consequently, the decomposition of sums of squares and DOFs is typically represented in a tabular form referred to as the analysis of variance (ANOVA) table. Table 2 represents an ANOVA table for comparing k factors and b blocks. The mean-square items (M_f, M_b, M_r) in this table are defined as the sums of squares divided by their DOFs.

Implementing a formal statistical test for no difference among treatment or block effects requires having a population model for the experiment. A population model for a randomized block experiment is summarized as

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \text{ for } i = 1, \ldots, k \text{ and } j = 1, \ldots, b \tag{2}$$

where μ = overall mean; α_i = factorial effects; β_j = block effects; and e_{ij} = random error independently distributed as $N(0, \sigma)$.

Tests for the absence of factorial differences or the absence of differences in block effects can now be performed by comparing the corresponding mean squares using an F-test. The null hypothesis is proposed to be

reject H_0: $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$ (no factorial differences) if

$$\frac{M_t}{M_r} > F_\alpha[k - 1, (b - 1)\cdot(k - 1)] \tag{3}$$

The significant factors testing under this model is used to verify the deterministic elements effect on the digitizing result. If the F-test initially shows a significant difference in means, the individual confidence intervals are established to compare specific pairs of factorial differences.

EXPERIMENTAL DESIGN

A SummaSketch II Professional Plus digitizer and a Umax Vista-S8 scanner were used as digitizing devices in this work. The digitizer's mechanical resolution can reach 80 lines per mm (2,000 lines per inch) in which the active area is 458 × 305 mm. Its claimed standard accuracy is 0.245 mm (*SummaSketch* 1991). Meanwhile, maximum optical resolution vertical with scanning direction is 16 dots per mm [400 dots per inch (DPI)] for the scanner, and its active area is near 220 × 280 mm (*Umax* 1995). By assessing the measurement system's accuracy, geometric characteristics of the devices were initially realized by calibration; factorial analysis by a statistical model was then performed to assess the effect of error sources on digitizing. A reseau of cross-shaped points with a 1 cm interval functioned as a "calibration" sheet. It was drawn on mylar by a Wild TA10 plotter. The plotter's addressable resolution is 0.02 mm (*AVIOTAB* 1990) higher than measurement systems. The sheet size was nearly A4 size, and nine points on the corners of the reseau were chosen to be control points. On the other hand, 117 check points consisting of "H-V" and "Scatter" patterns were used to verify distortion of devices in the horizontal-vertical, and radial directions, respectively. Fig. 1 presents the distribution of all points and two kinds of patterns.

First, the aforementioned equipment was used to digitize and individually scan the reseau template for the calibration. Consequently, the coordinates of nine control points and 117 check points could be obtained by manual digitizing on the tablet or screen. Three control points labeled T_a, T_b, T_c in Fig. 1 were chosen to compute the angle between vector $\mathbf{T}_b\mathbf{T}_a$ and $\mathbf{T}_b\mathbf{T}_c$. The calculated angle from the original coordinates of the control points represents the angle (θ) between the X and Y axes because of the similarity of vectors parallel to the X and Y axes. After affine transformation, the ratio (λ_x/λ_y) of independent X and Y scale is computed from transformation parameters. The assumed transformation formula was

$$X = a_1x + b_1y + c_1; \quad Y = a_2x + b_2y + c_2 \tag{4a,b}$$

where a_1, b_1, c_1, a_2, b_2, c_2 = six affine parameters; (X, Y) = transformed coordinate; and (x, y) = original coordinate. Next, λ_x/λ_y was computed as follows:

$$\frac{\lambda_x}{\lambda_y} = \frac{a_1}{b_2} \text{ or } -\frac{a_2}{b_1} \tag{5}$$

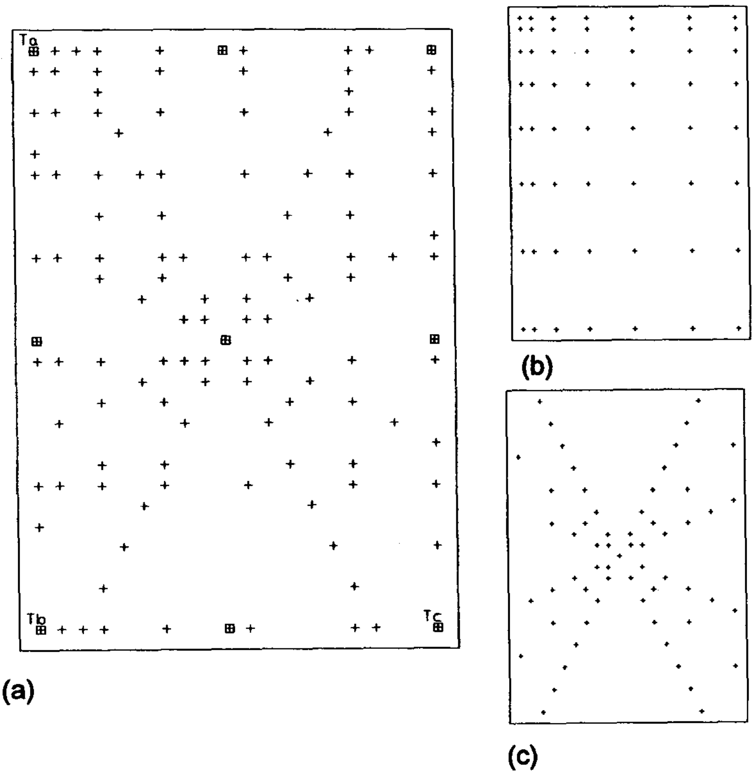

FIG. 1. Cross Dots Pattern Shapes: (a) Control Points; (b) "*H-V*"; (c) "Scatter"

Moreover, the error vectors of check points were also derived to provide essential information on the geometric characteristics of each device. Then, positional errors of planimetry (μ_p) were computed as

$$\mu_p = \sqrt{\frac{\sum_{i=1}^{n} (V_{x_i}^2 + V_{y_i}^2)}{n}} \tag{6}$$

where $V_{x_i} = X_i - X_i'$; $V_{y_i} = Y_i - Y_i'$; X_i, Y_i = measured coordinate of ith check point after affine transformation; X_i', Y_i' = referred coordinate of ith checkpoint; and n = number of checkpoints.

The distance between two checkpoints was calculated for comparison with the known distance between those two points. Consequently, the root-mean-square error (μ_d) between the differences in distances was also computed. The formula used was

$$\mu_d = \sqrt{\frac{\sum_{i=1}^{m} \Delta D_i^2}{m}} \tag{7}$$

where $\Delta D_i = D_i(p_j, p_k) - D_i'(p_j, p_k)$; $D_i(p_j, p_k) = \sqrt{(X_j - X_k)^2 + (Y_j - Y_k)^2}$; D_i, D_i' = measured and known distance between jth and kth checkpoint; and m = number of distance differences.

The previously mentioned transformation information and error analysis offer essential and detailed data for the geometric characteristics of each device. For quality control of geographic data, the positional error would fulfill the requirement of map accuracy standards.

The most common form of map accuracy standards is the U. S. National Map Accuracy Standard (USNMAS), as approved in 1947. It is currently used by the USGS to produce paper topographic quad sheets (Brenton et al. 1991). The USNMAS states that, for a given map, the deviation of well-defined points must be within a specific map distance for a specified percentage. This specification, called the circular map accuracy standard (ζ), can be conducted to evaluate the planimetric accuracy of desktop scanning and digitizing. In statistical terms, the planimetric accuracy can also be evaluated using the positional error of planimetry (μ_p). The relationship between ζ and μ_p is given as follows (ASPRS 1985):

$$\mu_p = \frac{\xi}{1.5174} \tag{8}$$

In this work, ξ was established so that more than 90% of well-defined points tested would not be in error by more than 0.5 mm in a horizontal plane. Thus, the implied equivalent allowable μ_p is

$$\mu_p = \frac{0.5 \text{ mm}}{1.5174} = 0.329 \text{ mm} \tag{9}$$

Except for the assessments mentioned before, a gray scale purchased from Kodak was used to verify the scanner's radiometric response. The gray scale had 20 steps, in 0.10 density increments, between a nominal "white" of 0.00 density and a practical printing "black" of 1.90 density. For comparative purposes, step numbers 1–20 were used to represent densities ranging from 0.00 to 1.90. After the gray scale was scanned, we computed the first moments of response, e.g., mean and standard deviation in each step area, and intervals between the logarithmic value of detected intensity and the next logarithmic value of intensity. The homogeneity and sensitivity of the gray level for the scanner could then be assessed by those statistics. Notably, radiometric characteristics of a scanner, such as linearity and stability of the sensor, are not included in this work.

Next, the digitizing behavior was analyzed to determine the factors affecting retrieval of coordinates. The analysis was primarily confined to manual digitizing, not a stream mode digitizing. The quality of digitized geographic data was related to the structural distribution of wire or sensors and the identification ability of operators. Therefore, factors affecting digitizing could be partitioned into two parts: "device" and "operator." The statistical analysis was performed whether the factors were significant or not. The template used for calibration was also seved as an experimental material for factorial analysis. However, it was drawn using 0.1, 0.3, and 0.5 mm widths of pens to realize the trend of identification of intersection points with different width of lines. The designed experiments are discussed in the following section to assess the factorial effects for both pieces of equipment.

Factors used in a randomized block model for the digitizer were considered to be three operators and three different line widths to evaluate the effects. Owing to the digitizer's active area being 458 × 305 mm, which is larger than templates, the map positions (i.e., mean left, middle, and right parts of the tablet) were treated as a block to verify the difference of digitizing results derived on different map positions. Fig. 2 displays the relationship between the active area of tablet and map positions. When digitizing, each operator placed templates at three map positions and recorded coordinates of desired points, respectively. Because equipment layout placed the tablet on the left side of the table, the left part of the digitizer was farther away from the monitor than was the right part. For the convenience of viewing the results on screen, the relationship between the operators' eyes and targets apparently did not remain orthogonally constant when the operators took data from the left map position. After transformation and μ_d computation, a two-way ANOVA table was organized. Next, the differences of operator and line width were verified to be significant or insignificant by the null hypothesis test.

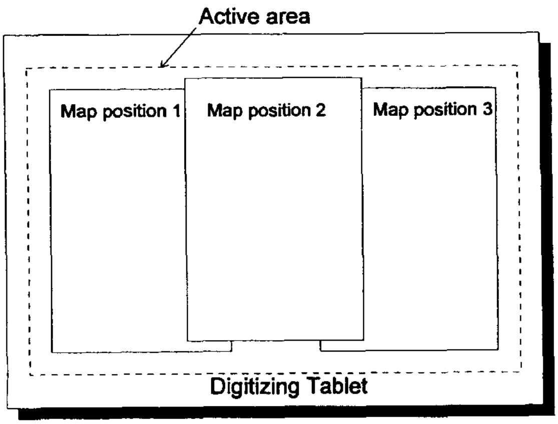

FIG. 2. Relationship between Active Area of Tablet and Map Positions

Thereafter, another RBM to analyze factorial effects on the scanner was set up with three scanning densities (100, 200, and 300 DPI) and three operators. The operator was treated as a factor, and different scanning densities were treated as a block. One of the templates (i.e., its line width is 0.3 mm) was chosen in the experiment to decrease measurements. First of all, the template was scanned into images by assigned densities. Next, each operator, on screen, retrieved the coordinates of nine control points and 117 checkpoints by locating the centers of cross-shaped targets. After affine transformation, μ_d for each image was calculated as observations by (7). Consequently, the ANOVA table was organized by those observations.

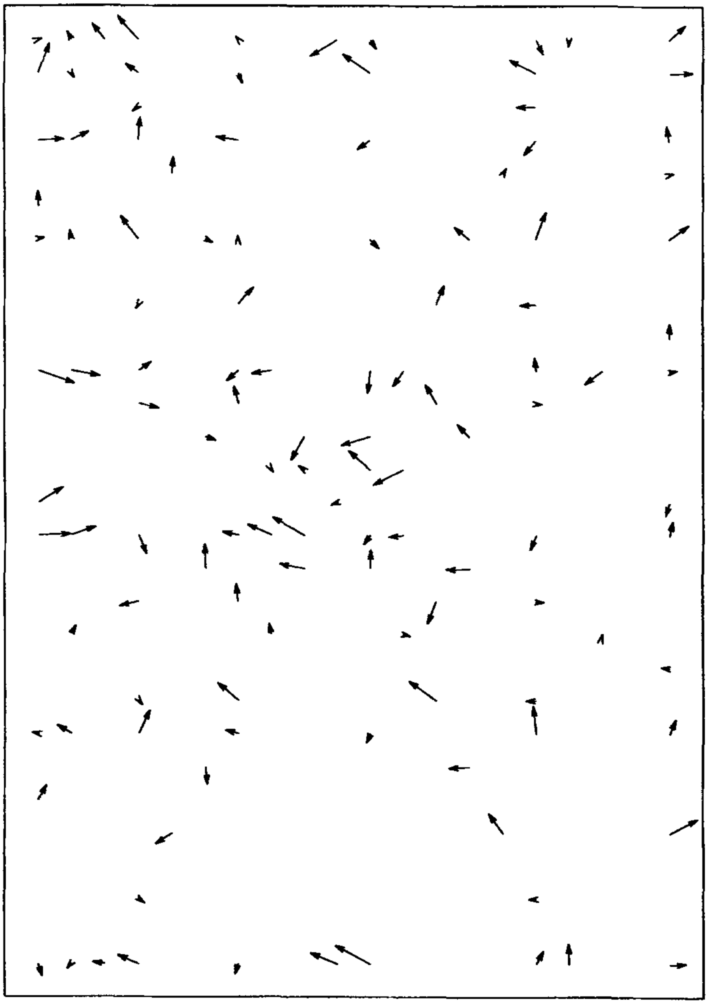

FIG. 3. Error Vectors Diagram for 117 Checkpoints Digitized by Scanner (For Clarify, Vectors Have Been Zoomed In 13 Times; Unit Is cm)

RESULTS AND ANALYSIS

Angle (θ) between the X and Y axes and the λ_x/λ_y ratio between the independent X and Y scale for the scanner are 89.95° and 1.0119, respectively. Moreover, angle θ and the λ_x/λ_y ratio for the digitizer are 90.05° and 0.9986, respectively. According to these results, the orthogonal attribute is well maintained and the scale ratio tends toward 1.0 for both pieces of equipment. Figs. 3 and 4 present error vectors of all 117 checkpoints for the scanner and digitizer, respectively. There are no systematic errors or specific trends on vector distribution. To analyze the geometric accuracy, Tables 3 and 4 list the positional errors μ_p and μ_d under different directions on H-V and Scatter patterns for both pieces of equipment. However, the average μ_p for the scan-

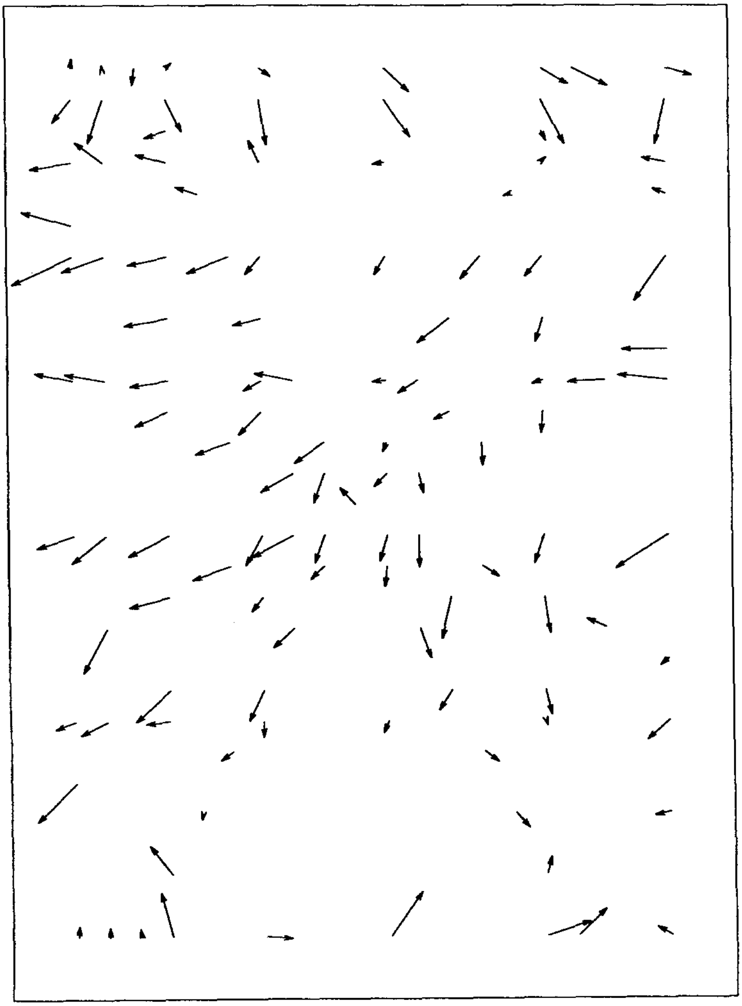

FIG. 4. Error Vectors Diagram for 117 Checkpoints Digitized by Digitizer (For Clarify, Vectors Have Been Zoomed In Eight Times; Unit Is cm)

TABLE 3. Error Statistic of Different Directions Assigned by *H-V* and Scatter Pattern for Digitizer

Orientation (degrees) (1)	Segments (2)	μ_p (mm) (3)	μ_d (mm) (4)
0°	168	0.283	0.246
45°	45	0.278	0.262
60°	105	0.193	0.244
90°	147	0.255	0.265
120°	105	0.216	0.261
135°	36	0.261	0.362
150°	45	0.236	0.326

TABLE 4. Error Statistic of Different Directions Assigned by *H-V* and Scatter Pattern for Scanner

Orientation (degrees) (1)	Segments (2)	μ_p (mm) (3)	μ_d (mm) (4)
0°	168	0.123	0.130
45°	45	0.122	0.128
60°	105	0.112	0.102
90°	147	0.107	0.098
120°	105	0.118	0.107
135°	36	0.102	0.077
150°	45	0.094	0.128

TABLE 5. Radiometric Response of Gray Level for Scanner

Step number (1)	Measured intensity (I) (2)	σ[a] (3)	Log (I)[b] (4)	Density (5)	Δ Log(I) (6)
1	255.00	0.000	2.4065	0.0	—
2	214.83	1.688	2.3321	0.1	0.07
3	174.08	1.492	2.2407	0.2	0.09
4	136.47	1.379	2.1350	0.3	0.10
5	110.31	1.157	2.0426	0.4	0.09
6	86.62	0.946	1.9376	0.5	0.10
7	70.08	0.929	1.8456	0.6	0.09
8	55.49	0.977	1.7443	0.7	0.10
9	44.72	0.808	1.6505	0.8	0.09
10	36.45	0.840	1.5617	0.9	0.09
11	29.68	1.040	1.4924	1.0	0.07
12	24.76	0.820	1.3937	1.1	0.10
13	20.28	0.802	1.3070	1.2	0.09
14	16.94	0.817	1.2289	1.3	0.08
15	14.00	0.851	1.1461	1.4	0.08
16	11.92	0.878	1.0764	1.5	0.07
17	9.95	0.918	0.9977	1.6	0.08
18	9.00	0.796	0.9541	1.7	0.04
19	8.12	0.852	0.9093	1.8	0.04

[a] The standard deviation of intensity is derived from the block area in the same density step.
[b] This component is the logarithmic value of intensity.

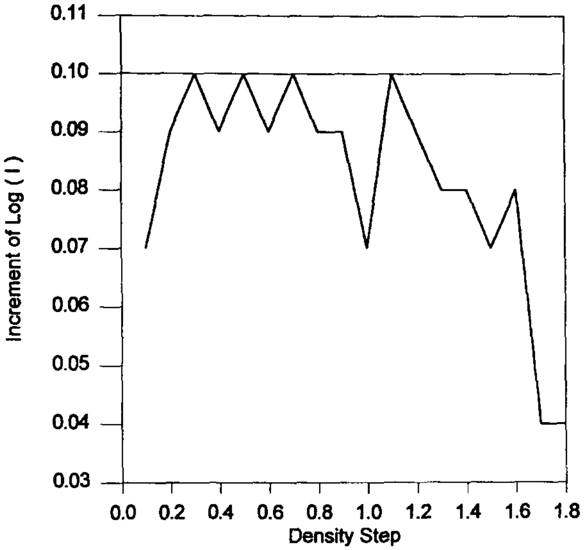

FIG. 5. Comparison between Density Step and Detected Increment of Gray Level for Scanner

TABLE 6. Two-Factor ANOVA Table of Operator and Linewidth for Digitizer

Source (1)	Sums of squares (2)	Degree of freedom (3)	Mean square (4)	*F*-ratio (5)
Op[b]	0.923	2	0.462	0.288
Lw[c]	49.027	2	24.514	15.304[a]
Op × Lw	0.539	4	0.135	0.084
Mp[d]	49.107	2	24.554	15.329
Residual	25.629	16	1.602	—
[Total]	125.225	26	—	—

[a]Null hypothesis (H_0) is rejected.
[b]Abbreviation of operator.
[c]Abbreviation of linewidth for reseau.
[d]Abbreviation of map position as a block.

ner and digitizer are 0.119 and 0.249 mm, respectively. According to those tables, positional errors for both pieces of equipment satisfy the requirement for the planimetric accuracy standard.

Table 5 summarizes the radiometric response results for the gray scale. According to this table, the standard deviation (0.00–1.688) of the radiometric response indicates that the response is quite homogeneous. The last column $\Delta\text{Log}(I)$ in this table is a logarithm value of the ith intensity minus the logarithm value of the $(i + 1)$th intensity. Fig. 5 defines the curve consisting of density and $\Delta\text{Log}(I)$ terms. Note that item $\Delta\text{Log}(I)$ in Fig. 5 is different from the original density increment 0.10 of the gray scale. This difference indicates that the larger difference appears as a trend on both ends of the gray scale, implying that the scanned result is unreliable when the target density is either too large (>1.6) or too small (<0.1).

TABLE 7. ANOVA Tables under Different Linewidth Constrained for Digitizer

Source (1)	Sums of squares (2)	Degree of freedom (3)	Mean square (4)	F-ratio (5)
(*a*) 0.1 mm Linewidth				
Mp[b]	0.008	2	0.004	11.009[a]
Op[c]	0.002	2	0.001	2.495
(*b*) 0.3 mm Linewidth				
Mp[b]	0.015	2	0.0075	8.233[a]
Op[c]	0.0004	2	0.0002	0.174
(*c*) 0.5 mm Linewidth				
Mp[b]	0.012	2	0.006	3.457
Op	0.006	2	0.003	1.716

[a]Null hypothesis (H_0) is rejected.
[b]Abbreviation of map position as a factor.
[c]Abbreviation of operator as a block.

TABLE 8. Two One-Way ANOVA Table for Scanner

Source (1)	Sums of squares (2)	Degree of freedom (3)	Mean square (4)	F-ratio (5)
(*a*) Three Operators				
Op[b]	0.003	2	0.002	5.678[a]
Sd[c]	0.010	2	0.005	15.720[a]
(*b*) One Operator				
Sd[d]	0.004	2	0.002	9.226[a]
Lw[e]	0.001	2	0.0005	1.425

[a]Null hypothesis (H_0) is rejected.
[b]Abbreviation of operator as a factor.
[c]Abbreviation of scanning density as a block.
[d]Abbreviation of scanning density as a factor.
[e]Abbreviation of linewidth as a block.

ANOVA results in Table 6 display the statistical experimental results for the digitizer. The F distribution table reveals that the tabulated 0.05 point of $F(2, 16)$ is 3.63. The value is less than the observed F-ratio in Table 6, i.e., 15.304, but is larger than 0.288. Thus, the hypothesis of no factorial differences for different line width is rejected. The block effects imply that the map positions are significant because the observed F-value of 15.329 is larger than the tabulated value $F_{0.05}(2, 16)$. Therefore, we can conclude that the line width and map position are statistically significant treatments. To verify the effects in detail, one-way ANOVA tables consisting of the map position as a factor and the operator as a block were reorganized. Table 7 summarizes the results having different line widths. According to this table, map position is a significant factor under a linewidth that is 0.1 or 0.3 mm, but it is insignificant when the linewidth is 0.5 mm. Meanwhile, the digitizing results in the left part of the tablet are worse than the ones in the right side due to error distribution. However, a digitizer's structural distribution should be uniform under the assessment of geometric characteristic for a digitizer.

According to the equipment layout and error distribution, the differences are believed to have resulted because three operators tended to locate targets on a tilt view when the map position stood at the left part of the tablet. Moreover, this effect would be more significant under different line widths. From the previous discourse, we can infer that factorial analysis is helpful in verifying the digitizing habits of operators. In addition to the geometric effects of device, if an operator's habits are corrected to abandon the wrong locating phenomenon, then quality of digitized data would be improved.

Table 8 shows the results of the ANOVA table for the scanner. The statistical result indicates that operator and scanning density are significant factors, but linewidth is insignificant from the *F*-test. No problem arose in locating targets projectively when the operators digitized points on screen; however, operators' perceptions of center point identifications did vary. Moreover, increasing the scanning density enhances the resolution. Consequently, the accuracy of target locating would be enhanced.

CONCLUSIONS

The average planimetric positional error (μ_p) for the digitizer and the scanner were 0.249 and 0.119 mm, respectively. Therefore, the geometric accuracy of both types of low-cost devices apparently satisfies the requirements of the mapping community. The radiometric response of the gray level derived from the standard deviations of intensity is quite homogeneous. Moreover, a circumstance in which intensity is close to printing ''black'' or nominal ''white'' implies that the density is higher than 1.6 or lower than 0.1. Therefore, the scanning results would be unreliable. Error analysis indicates no systematic errors or specific trends in geometric characteristics for the devices.

The RBM method can be effectively used in analyzing significant factors for digitizing equipment. Variabilities in map position and linewidth are significant for the digitizer. We believe that the primary factor influencing variance is that operators tend to locate targets on a tilted view. If the relationship between operator and targets remains orthogonally constant, errors would subsequently decrease. Also, scanning density and operator are significant elements for the scanner. Variance between factors is accounted for by the difference in the operators' perceptions regarding point identification. Except for the increasing resolution of devices, the source of the difference on identification is related to digitizing habits of operators. Moreover, elements affecting the digitizing behavior are verified by the statistical analysis. Such a behavior would then be reduced to enhance the quality of digitized data. An evaluation similar to that performed herein is necessary before digitized geographic data can be retrieved by those low-cost pieces of equipment.

This work has provided possible significant treatments for both the digitizer and scanner. However, further study is required to identify the other unknown factors that could also affect the outcome during digitizing.

APPENDIX I. REFERENCES

ASPRS (Am. Soc. for Photogrammetry and Remote Sensing). (1985). ''Committee for specification and standards: accuracy specification for large-scale line maps.'' *Photogrammetric Engrg. and Remote Sensing*, 51(2), 195–199.

AVIOTAB Wild TA10 digital plotting table instruction manual. (1990). Wild Leitz Ltd., Heerbrugg, Switzerland.

Bolstad, P. V., Gessler, P., and Lillesand, T. M. (1990). ''Positional uncertainty in

manual digitized map data." *Int. J. Geographical Information Syst.*, Hampshire, U.K., 4(4), 399–412.
Brenton, J. K., Smith, J. L., and Gregoire, T. G. (1991). "Modelling and evaluating the effects of stream mode digitizing errors on map variables." *Photogrammetric Engr. and Remote Sensing*, 57(7), 957–963.
Carstensen, L. W. Jr., and Campbell, J. B. (1991). "Desktop scanning for cartographic digitization and spatial analysis." *Photogrammetric Engr. and Remote Sensing*, 57(11), 1437–1446.
Drummond, J., and Bosma, M. (1989). "A review of low-cost scanner." *Int. J. Geographical Information Syst.*, Hampshire, U.K., 3(1), 83–95.
Fernãndez, R. N., Fabiãn, L. G., Deeds, G., and Johannsen, C. J. (1991). "Accuracy assessment of map coordinate retrieval." *Photogrammetric Engrg. and Remote Sensing*, 57(11), 1447–1452.
Lai, P. C. (1988). "Resource use in manual digitization—a case study of the Patuxent basin geographical information system database." *Int. J. Geographical Information Sys.* Hampshire, U.K., 2(4), 329–345.
Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwire, K. C., and Tinney, L. R. (1991). "Remote sensing and geographic information system data integration: error sources and research issues." *Photogrammetric Engrg. and Remote Sensing*, 677–687.
Richard, A. J., and Gouri, K. B. (1992). *Statistics: principles and methods*, 2nd Ed. John Wiley & Sons Inc., New York, N.Y., 541–571.
SummaSketch II series drivers/utilities user's guide. (1991). Summagraphics Co., Seymour, Conn.
Thapa, K., and Bossler, J. (1992). "Accuracy of spatial data used in geographic information systems." *Photogrammetric Engrg. and Remote Sensing*, 835–841.
Umax Vista-S8 scanner user's guide. (1995). Umax Data Systems Inc., HsinChu, Taiwan, Republic of China (in Chinese).

APPENDIX II. NOTATION

The following symbols are used in this paper:

$a_1, b_1, c_1, a_2, b_2, c_2$ = affine parameters;
D, D' = measured and known distance;
e = random error;
$F_\alpha(\ )$ = F value of tabulated α point;
H_0 = Null hypothesis;
I = intensity;
$k(b)$ = treatment (block) number;
$\mathrm{Log}(I)$ = logarithm value of intensity;
M = mean squares;
$n(m)$ = number of checkpoints (distance difference);
p = checkpoint;
S = sum of squares;
T_a, T_b, T_c = labeled control points;
$V_x(V_y)$ = difference between transformed and referred coordinate on $X(Y)$ axis;
y = measurement;
$\bar{y}$ = mean;
$\alpha(\beta)$ = factorial (block) effect;
ΔD = difference between measured and known distance;
θ = angle between X and Y axes;
$\lambda_x(\lambda_y)$ = $X(Y)$ scale;
μ = mean;

μ_d = root-mean-square error of differences of distances;
μ_p = positional error of planimetry; and
ξ = circular map accuracy standard.

Subscripts

f, b, r = factorial, block, and residual indices;
i, j, k = positive integer indices; and
= summation.