# Accuracy of a commercially available digitizer: A new method for assessment of errors in linearity

**Luc P.M. Tourne, DDS, MS**

Cephalometric analysis involves a number of technical manipulations and uses different kinds of equipment. From positioning the patient in the cephalostat through developing the radiograph, to the use of hardware to collect and measure the data, each process can introduce a certain amount of error. The size of this accumulated error has a direct bearing on the clinical and scientific reproducibility (reliability) and interpretation of the data.[1]

Some researchers have found that much of the error is due to errors in projection,[2] difficulty in identifying landmarks,[2-6] poor image quality,[7] and differences between operators[8,9] and techniques.[9,10]

Digitizing equipment has become more popular as more orthodontists acquire a computerized system for cephalometric data analysis. Only a few studies have dealt with digitizer accuracy.[11-13] The mean error of the equipment tested has been estimated to be around or below 0.1 mm,[11-13] but some of these digitizers were specifically designed for orthodontic purposes.[12] However, several of the commercially available digitizers (including the one tested in this study) were not specifically constructed for scientific research and the level of accuracy they offer may be unacceptable in clinical practice.

Part of the error of a digitizer can go unnoticed because of systematic distortion of the coordinates.[11] This problem, called nonlinearity, will not be revealed by repeated measurements of the same digitizing points, as the same amount of error will be introduced with each digitizing session. Only one study addressed this problem,[11] but the statistical method used was unable to calculate the magnitude and exact location of these distortions.

The objective of this study was to analyze the accuracy of a commercially available digitizer. In

**Abstract**

A commercially available digitizer has been tested for accuracy. The various areas of the digitizing tablet show degrees of precision that can differ for the x- and y-coordinates. The tablet has a mean absolute error of 0.016 mm for the x-coordinate and 0.09 mm for the y-coordinate. Error tended to increase toward the sides of the tablet. Partitioning the error into systematic and random components revealed that the x-error is due mainly to errors in linearity of the digitizer. When compared with the error involved in locating cephalometric landmarks, one can conclude that this type of digitizer is suited for orthodontic purposes. This article describes a new methodological approach to locating and measuring errors in linearity.

**Key Words**

Digitizer • Error • Linearity • Cephalometry

**Submitted:** March 1995 **Revised and accepted:** October 1995 Angle Orthod 1996; 66(6):433-440.

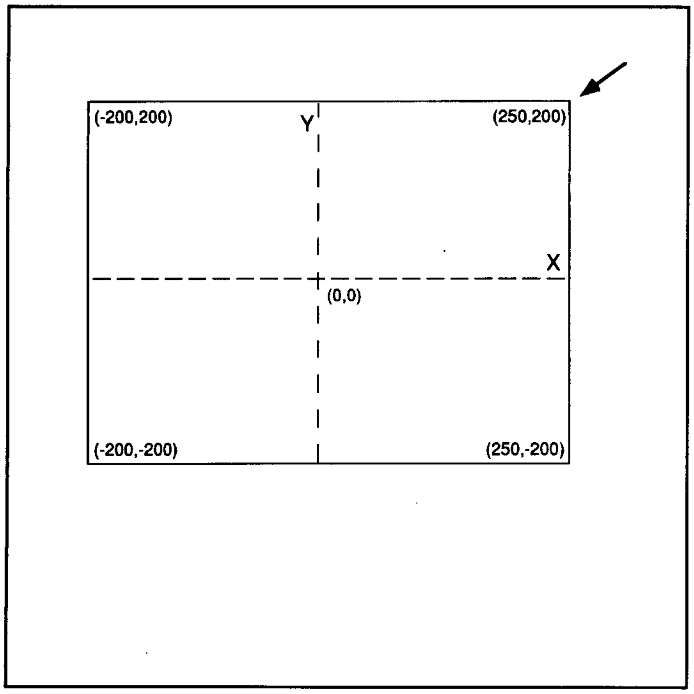

**Figure 1**

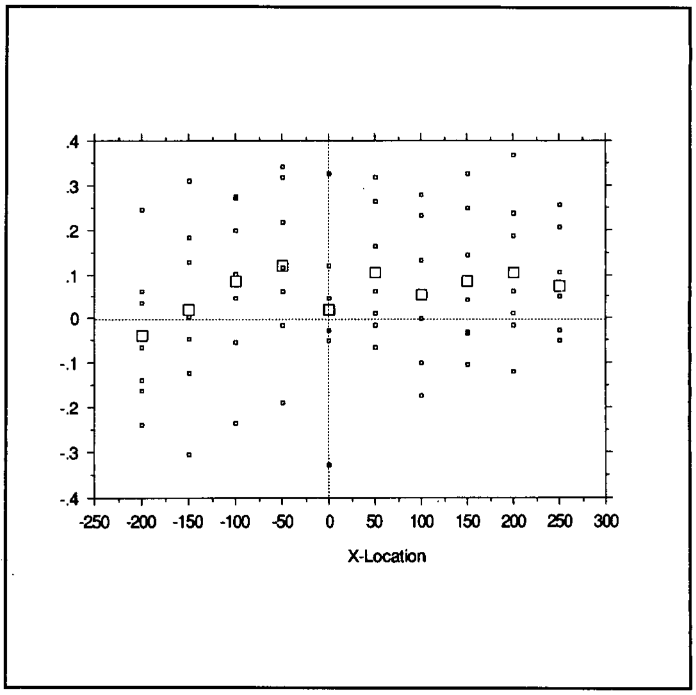

**Figure 2**

**Figure 1**
**Coordinate layout of digitizing tablet. Arrow at upper right corner denotes reader's position (point of view) when looking at 3-D plots (Figures 4 to 6).**

**Figure 2**
**Total true error in x-coordinate dependent on gridpoint's horizontal location on tablet (x-location). Small squares on a vertical line are x-errors of individual gridpoints with the same x-coordinate but different y-coordinate. Larger square denotes mean of those errors for this gridpoint group.**

addition, a new method is proposed to estimate errors in linearity of the digitizing tablet by partitioning the total error into estimates of its systematic and random components.

## Materials and methods

The digitizer used is a nontransparent Kurta XGT/ADB tablet (Kurta Corp, Phoenix, Ariz) connected to a Macintosh IIx computer (Apple Corp, Cupertino, Calif). The size of the digitizing area is about 12 inches by 18 inches. The manufacturer claims a resolution of the tablet of 1270 ppi (508 ppcm) and an accuracy of ±0.01" (.25 mm).

A sheet of paper with a calibrated millimeter grid printed on it was attached to the digitizer, covering the entire digitizing surface. An x,y coordinate system was set up on the calibrated grid, and the origin (0,0) was located on a grid point near the center of the digitizing area. The unit of measure on the coordinate axes was chosen to be 1 mm, so that the coordinates of the other gridpoints were equal to their distance in mm away from the origin. The theoretical (calibrated) coordinates of these gridpoints were recorded in 5 cm intervals, leading to 7 rows and 10 columns of gridpoints spanning the digitizing area. As such, 1 row contained 10 (x,y) gridpoints with the same y-values and different x-values, while 1 column contained 7 gridpoints with identical x-values and different y-values. Coordinate points located to the left or below (0,0) were assigned a negative x-value or y-value, respectively (Figure 1).

The digitizing software for the tablet was calibrated (scaled) to the paper axis system. As a consequence, if no errors were involved in the digitizing process, the coordinate readouts of the tablet would be identical to the theoretical coordinates of the paper calibration grid. The 70 gridpoints were digitized and their coordinates stored for statistical analysis. This process was repeated four times.

### Statistical analysis

For each digitizing session, the total abcis and ordinate digitizing error (total true error) and its absolute value (total absolute error) were calculated by subtracting the theoretical coordinate value from the digitized value. Subsequently, the mean error over the five digitizing sessions was calculated (mean total true error and mean total absolute error). Descriptive statistics of the mean total absolute error were used to describe overall performance of the tablet. The error estimate for a single measurement (root mean square error, RMSE) and a gridpoint's 95% confidence interval were calculated by means of Dahlberg's formula.[14]

The total true errors and mean total true error were used to screen for any systematic errors in the data and for partitioning the total error into its different components (systematic and random). To allow for visual inspection of any systematic trend in the data, the mean total true error was visualized by means of scattergrams and 3-D plotting, as done by Eriksen and Solow.[11] Statistical testing for systematic error (systematic true error) was done by a factorial repeated measures ANOVA (RMA).[15] The RMA had one within factor (digitizing sessions) and two between factors (x and y position on the tablet). In the event of an overall statistically significant F-value, post hoc testing was performed by means of a paired *t*-test with Bonferroni adjustment for the simple effects of the within factor. Statistical significance of the main effects was further investigated by post hoc testing of their simple effects by means of the Student-Neuman-Keuls procedure. The mean systematic true error (the error in linearity) was calculated by having the ANOVA-model estimate fitted values for the error. The mean random true errors were represented by the residuals of the ANOVA-model. Both error terms were visualized by 3-D plotting and statistically described. The predictive success of the ANOVA-model was described by the Omega Squared statistic ($\omega^2$).

To simplify semantics in the text to follow, total, systematic, and random error will denote the means of these errors over the five digitizing sessions unless otherwise stated.

**Table 1**
**Total absolute error of the digitizing tablet**

| | Mean | SE | SD | Min. | Max. |
|---|---|---|---|---|---|
| x-error | .156 | .011 | .096 | .024 | .368 |
| y-error | .093 | .005 | .043 | .010 | .260 |

**Table 2**
**Total true error of the digitizing tablet**

| | Mean | SE | SD | Min. | Max. |
|---|---|---|---|---|---|
| x-error | .065 | .020 | .167 | -.328 | .368 |
| y-error | -.050 | .010 | .081 | -.260 | .160 |

## Results

### Errors in x-direction

The error in the x-coordinate of the digitized grid points (horizontal error) will be considered first.

Total error: The total absolute error is described in Table 1. The tablet has a mean absolute error in the x-direction of 0.16 mm with a maximum of 0.37 mm near the lower right corner. The total true errors of the x-coordinate (tot.e,x) dependent on x-location on the tablet are illustrated in Figure 2. Almost all errors are situated within a band of -0.3 mm to 0.3 mm, with a mean of 0.07 mm (Table 2). Visual inspection of the graph reveals a possible linear tendency on the left side of the tablet, from x=-200 to x=-50. The same errors, however, plotted against their y-location on the tablet (Figure 3), clearly show a linear trend throughout the data: the error is positive in the lower half of the tablet, decreases toward the center line, and changes to increasingly negative values toward the upper border. A 3-D plot (Figure 4) simultaneously shows the total true error relative to the x- and y-axes of the tablet.

Calculation of the pooled standard deviation among the replicate measurements for each individual gridpoint revealed that the RMSE in x-direction for a single measurement varied from a minimum of 0.043 mm to a maximum of 0.225 mm, depending on the location on the tablet. A cumulative histogram revealed that about 75% of the RMSE values were smaller than 0.12 mm. As such, a 95% confidence interval can be established as x ± .24 mm for 75% of the gridpoints.

Systematic error: The ANOVA-table (Table 3) reveals a F-value that is significant for the different digitizing sessions (P<.001). Post hoc testing of the simple effects of this within factor proved the mean x-errors differ significantly from the earlier to later digitizing sessions. It also differed significantly from the hypothesized value of 0 for the last three digitizing sessions (P<.01). This may indicate some instability of the electromagnetic fields of the tablet over time or some systematic error introduced by the operator. The main effects of both between factors (x- and y-position on the tablet) were also significant (P<.001). The Student-Neuman-Keuls procedure at P<.01 revealed that for x-location, this was due mainly to the negative group mean error at x=-200 relative to the more positive values and the mean group error at x=-50 (Figure 2). The same post hoc procedure for the y-location grouping factor found significant differences between all but one couple of group means (between y=50 and y=100, Figure 3). None of the interaction effects were significant (Table 3). Calculation of $\omega^2$ (Table 4) shows that 85% of the variation in x-error can be explained by y-location, while 92% of the variance is explained by both factors.

The systematic error for each gridpoint was calculated and is displayed three-dimensionally in Figure 5.

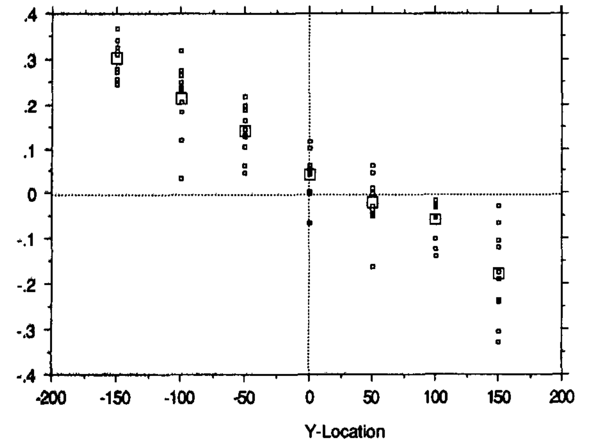

**Figure 3**

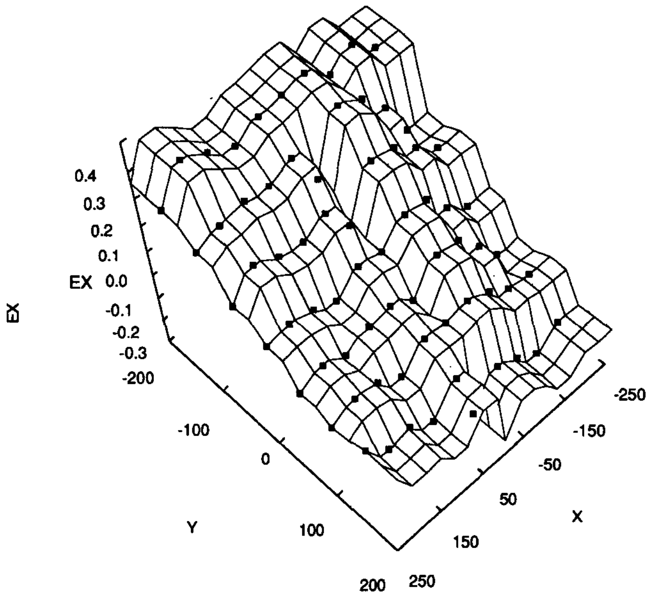

**Figure 4**

**Figure 3**
**Total true error in the x-coordinate dependent on vertical location on tablet (y-location). A vertical line of small squares denotes the x-errors of individual gridpoints with the same y-coordinate but different x-coordinate. A larger square denotes the mean of those errors for this gridpoint group.**

**Figure 4**
**3-D plot of total true error in x-coordinate relative to tablet position of gridpoints. The x- and y-axes of the tablet are reversed. This view is looking down from the upper right corner of tablet (see arrow in Figure 1).**

Residual (random) error: The random error was calculated and its 3-D plot (Figure 6) showed it to be quite small and relatively more important in the corners and along the left side of the tablet.

**Errors in y-direction**

Total error: The total absolute error in the y-coordinate of the gridpoints, with a mean of 0.09 mm (Table 1), was smaller than that of the x-coordinate. It had a maximum value of 0.26 mm near the upper left corner of the tablet. The total true error had a mean of -0.05 mm (Table 2), which was also smaller than that of the x-coordinate. Figure 7 gives a graphical overview of the total y-error dependent on the horizontal grid point location. The error seems to follow a sinusoidal curve from the left to the right side of the tablet. Figure 8 shows the error has a tendency to become larger near the upper and lower borders.

The RMSE in the y-direction for a single measurement varied from 0.003 mm to 0.16 mm. About 75% of the gridpoints have a 95% confidence interval of y ± .16 mm.

Systematic error: The ANOVA table (Table 5) revealed no significant difference among the five digitizing sessions, but testing of a null hypothesis of a total y-error equal to zero led to rejection of the hypothesis for four of the five sessions ($P<.001$). The main effects of x- and y-positions on the tablet were significant ($P \le .0001$). Student-Neuman-Keuls testing at $P<.01$ for x-location showed the positive errors on the right side of the tablet to be significantly different from the negative values on the left side (Figure 7). The same procedure for y-location found the error at the upper border of the tablet (y=150) to differ significantly from those of the middle region (Figure 8). No interaction effects were statistically significant (Table 5). The $\omega^2$ statistic indicated that the model was less efficient in explaining the variation in y-error than in x-error (Table 4), although 57% of the y-variation could be accounted for.

Residual error: As the contribution of the systematic error term to the total error was smaller, a relatively larger part of the error variation (about 43%) is due to random error. A scatterplot of the fitted versus the residual values for y-error shows the residuals to be of the same order of magnitude as the predicted values (Figure 9).

## Discussion

The error related to the digitizing procedure has been found to come from two sources.[11,13] First, there is an operator/hardware component comprising the accuracy of the crosshair pointer and the precision by which the researcher can pinpoint the gridpoints. Additionally, there is the error due to the physical characteristics of the tablet itself. Extreme care was taken in setting up the tablet to assure evenly easy operator access to the whole digitizing area of the tablet. In addition, each digitizing session was performed on a different day to avoid the introduction of systematic error by multiple repetitive measurements during one session. Nevertheless,

**Table 3**
**Repeated Measures ANOVA with total true x-error as dependent variable**

| Source | df | SS | MS | F-value | P-value |
|---|---|---|---|---|---|
| x-location | 9 | .767 | .085 | 7.664 | .0001 |
| y-location | 6 | 8.275 | 1.379 | 124.057 | .0001 |
| Subjects | 54 | .600 | .011 | | |
| Dig. sessions | 4 | .136 | .034 | 5.128 | .0006 |
| Dig. session* x-location | 36 | .273 | .008 | 1.144 | .2754 |
| Dig. session* y-location | 24 | .213 | .009 | 1.342 | .1393 |
| Dig. session* subjects | 216 | 1.430 | .007 | | |

**Table 4**
**Goodness of fit of the ANOVA-models by means of Omega Squared**

| | x-error | y-error |
|---|---|---|
| $\omega^2$ x-location | .069 | .402 |
| $\omega^2$ y-location | .850 | .163 |
| $\omega^2$ total | .919 | .565 |

introduction of a certain amount of operator error is possible. This could have been partially addressed by having different people digitize the gridpoints, so that intraindividual systematic error is minimized. Unavoidably, the error reported in this publication is the combination of operator and hardware limitations, and the data should be interpreted as such.

Another topic of concern is the representativity of the digitizer. The testing of one hardware product does not necessarily mean that all manufactured items of the same brand and type have identical characteristics. Ideally, a random sample of digitizers would have been tested, but this was not practical. This type of quality control is normally done by the manufacturer and is necessary in order to claim a certain accuracy for the digitizer brand. This paper assumes that the digitizer tested is representative of its type, but in the strict sense, the error data pertain only to this individual digitizer used by this individual operator. Ideally, every researcher or clinician interested in the amount of error of his or her digitizing equipment should test the tablet.

Error in orthodontic measurement has been divided theoretically into systematic and random error.[1] Systematic error (bias) is a reproducible kind of error, introducing a significant trend in the data.[16] The unpredictable part of the error is called erratic, accidental, or random error.[1,16] Systematic errors have usually been tested for by a paired *t*-test,[1,9] which does not allow for actual calculation of the magnitude of the systematic component. However, use of a general linear modeling procedure, such as ANOVA, allows for construction of a prediction model to estimate the systematic error term.[15] This is achieved by a least squares method that generates fitted values for the explainable part of the error. The residuals calculated from the model can then be viewed as the unexplained (random) part of the error. Additionally, in this study, a factorial ANOVA design allows the incorporation of horizontal and vertical tablet location as separate grouping factors. A significant F-value for one of these factors means that tablet location itself introduces a systematic error in the coordinate readout, which is in essence an error in linearity. Post hoc testing of significant differences within these factors (simple effects) and calculation of fitted values show statistically and graphically the tablet location and the magnitude of the linearity error. This is an advantage over a previous method of nonlinearity detection where the topical characteristics of linearity could only be described by graphical plotting.[11]

The total absolute error, calculated by summarizing the absolute values of the errors during the five digitizing sessions, gives a good overall description of tablet accuracy. The absolute value is used to avoid errors of different signs canceling each other out. Table 1 shows the mean values to be well within the claimed accuracy of the tablet manufacturer, but the maximum values (found in the corners) were considerably higher. Overall, the y-errors are smaller than the x-er-

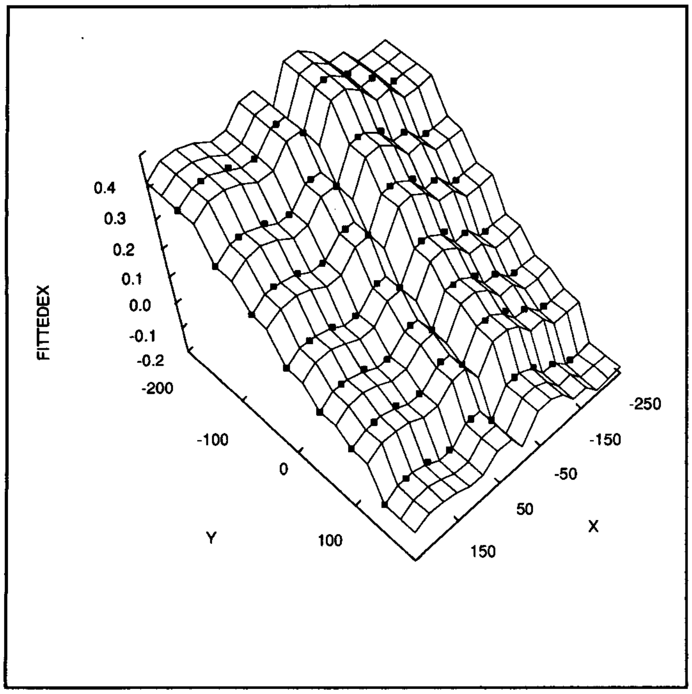

**Figure 5**

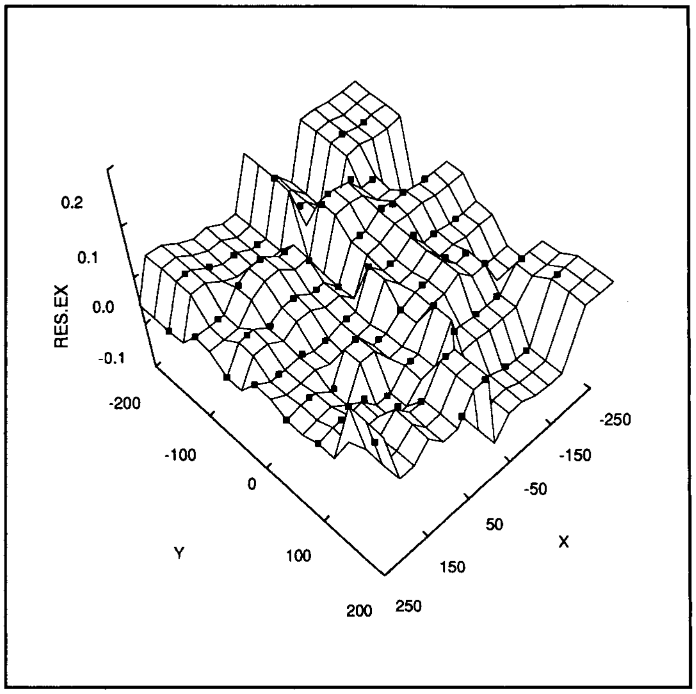

**Figure 6**

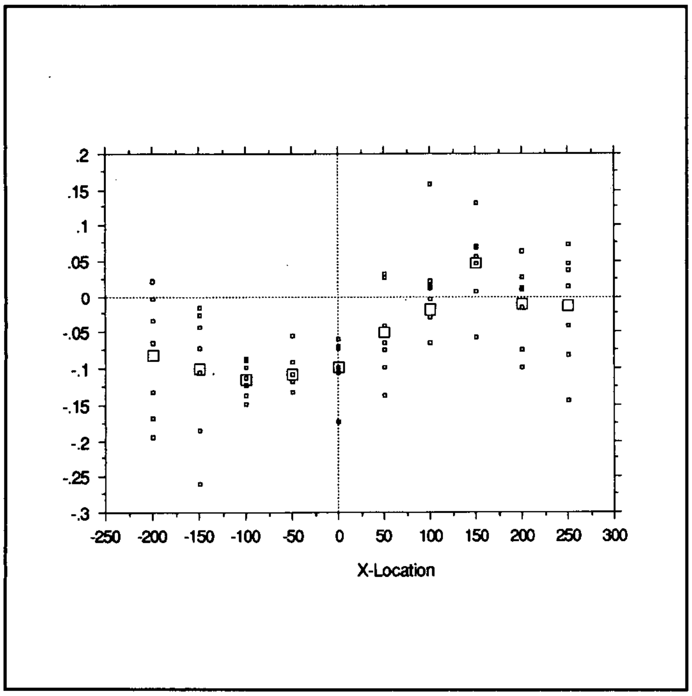

**Figure 7**

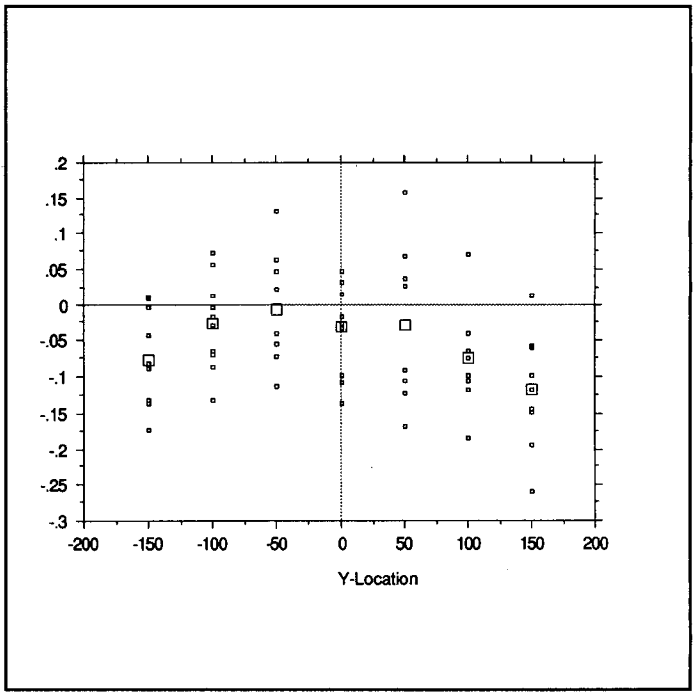

**Figure 8**

**Figure 5**
**3-D plot of systematic error in x-coordinate relative to tablet position of gridpoints. The x-and y-axes are reversed.**

**Figure 6**
**3-D plot of random error in x-coordinate relative to tablet position of gridpoints. The x-and y-axes are reversed.**

**Figure 7**
**Total true error in y-coordinate dependent on horizontal location on tablet.**

**Figure 8**
**Total true error in y-coordinate dependent on vertical location on tablet.**

rors. 3-D plots of the absolute errors (not shown) revealed the error to be highest near the borders of the tablet and in the corners. Consequently, accuracy can be maximized by centering the cephalogram or tracing on the digitizing surface.

Calculation of the total true (as opposed to absolute) error was necessary to detect any systematic linear trends in the data. The x-coordinate of the data suffered most from linearity error (Figures 2 and 3). A comparison of Figures 4 and

**Table 5**
**Repeated Measures ANOVA with total true y-error as dependent variable**

| Source | df | SS | MS | F-value | P-value |
|---|---|---|---|---|---|
| x-location | 9 | 1.044 | .116 | 8.495 | .0001 |
| y-location | 6 | .459 | .076 | 5.600 | .0001 |
| Subjects | 54 | .738 | .014 | | |
| Dig. sessions | 4 | .045 | .011 | 2.137 | .0773 |
| Dig. sessions* x-location | 36 | .157 | .004 | .830 | .7428 |
| Dig. sessions* y-location | 24 | .169 | .007 | 1.340 | .1404 |
| Dig. sessions* subjects | 216 | 1.134 | .005 | | |

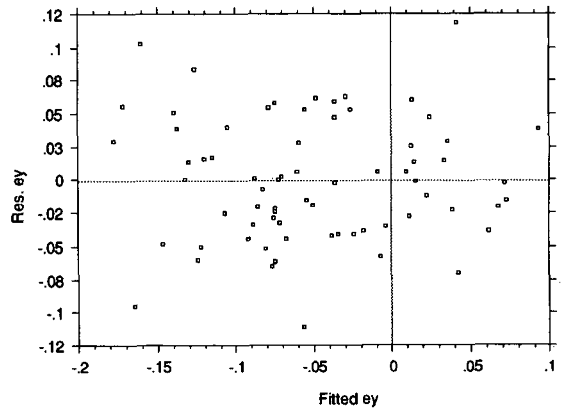

Figure 9

**Figure 9**
**Scatterplot of ANOVA-fitted values for y-error (systematic y-error) relative to unpredicted part of error (residual y-error).**

5 shows that the characteristics of the total error graph are due mainly to systematic error: the appearance as well as the magnitude of the linear error are nearly identical to those of the total error. This relative importance of the linear error allows for significant correction of the digitized values of any tablet point by subtracting the (interpolated) linear error from the x-coordinate readout. Although the y-coordinates showed some localized statistically significant linearity errors (Table 5 and Figures 7 and 8), the ANOVA-model indicated that almost half of the variation in y-error was random (Table 4). Given the already smaller total y-error (Table 2), one would gain little in trying to adjust for the systematic y-error.

The clinical importance of digitizing error is directly dependent on the level of accuracy that is needed for storing the coordinate points of the anatomical landmarks. The error involved in landmark identification itself has been reported to be related to the precision of the chosen landmark definition, the radiographic complexity of the region and the quality of the radiographic image.[3,4,7,17,18] Any combination of these factors tends to increase the landmark detection error beyond the limits of the digitizing error. Each landmark has its own specific scatterpattern of error[2,3,8] with the mean error ranging from 0.22 mm to as much as 1.06 mm.[19] As a consequence, it is clear that many landmarks cannot be located with a degree of precision matching that of the digitizer.

## Conclusions

Different areas of the digitizing tablet show degrees of precision that can vary for the x- and y-coordinates. Overall, the y-error is smaller and contains a relatively large random error component, while the total x-error is inflated by a significant error in linearity. This linearity error can be statistically shown and calculated by applying a factorial linear model to the data. If the clinician so desires, the digitizing data can be corrected for this error.

The mean absolute error of the tablet (0.16 mm for the x-coordinate, and 0.09 mm for the y-coordinate) falls well within the claimed accuracy of the tablet (0.25 mm), but errors are clearly larger near the borders. As such, digitizing near the center of the tablet helps reduce error. A majority of the gridpoints, 75%, have a 95% confidence interval of x ± 0.24 mm and y ± 0.16 mm.

The precision of this digitizer is clearly greater than the accuracy by which anatomical landmarks can be located. As a consequence, this digizer is suited for orthodontic use.

## Author Address:

Dr. Luc Tourne, DDS,MS
38 Steenweg
2800 Mechelen
Belgium
*Luc P.M.Tourne is in private practice in Mechelen, Belgium.*

## References

1. Houston WJB. The analysis of errors in orthodontic measurements. Am J Orthod 1983;83:382-390.
2. Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. Am J Orthod 1971;60:111-127.
3. Broch J, Slagsvold O, Rosler M. Error in landmark identification in lateral radiographic headplates. Eur J Orthod 1981;3:9-13.
4. Richardson A. An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. Am J Orthod 1966;52:637-651.
5. Savara BS, Tracy WE, Miller PA. Analysis of errors in cephalometric measurements of three-dimensional distances on the human mandible. Arch Oral Biol 1966;11:209-217.
6. Midtgard J, Bjork G, Linder-Aronson S. Reproducibility of cephalometric landmarks and errors of measurements of cephalometric cranial distances. Angle Orthod 1974;44:56-61.
7. McWilliam JS, Welander U. The effect of image quality on the identification of cephalometric landmarks. Angle Orthod 1988;48:49-56.
8. Stabrun AE, Danielsen K. Precision in cephalometric landmark identification. Eur J Orthod 1982;4:185-196.
9. Keeling SD, Cabassa SR, King GJ. Systematic and random errors associated with Johnston's cephalometric analysis. Br J Orthod 1993;20:101-107.
10. Savage AW, Showfety KJ, Yancey J. Repeated measures analysis of geometrically constructed and directly determined cephalometric points. Am J Orthod Dentofac Orthop 1987;91:295-299.
11. Eriksen E, Solow B. Linearity of cephalometric digitizers. Eur J Orthod 1990;13:337-342.
12. Bondevik O, Rosler M, Slagsvold O. The digital read-out system CM-1. An instrument for rational measuring on radiographic headplates and dental models. Eur J Orthod 1981;3:1-9.
13. McWilliam JS. Evaluation and calibration of x-y-coordinatograph used in cephalometric analysis. Scand J Dent Res 1980;88:496-504.
14. Dahlberg G. Statistical methods for medical and biological students. London: Allen and Urwin, 1940.
15. Howell DC. Statistical methods for psychology. Boston: PWS Publishers, 1987.
16. Hunter WS, Priest WR. Errors and discrepancies in measurement of tooth size. J Dent Res 1960;39(2):405-414.
17. Hurst RVV, Schwaninger B, Shaye R. Interobserver reliability in xeroradiographic cephalometry. Am J Orthod 1979;75:179-183.
18. Rossmann K, Wiley B. The central problem in the study of radiographic image quality. Radiology 1970;96:113-118.
19. Richardson A. A comparison of traditional and computerized methods of cephalometric analysis. Eur J Orthod 1981;3:15-20.