A new minimum zone method for evaluating straightness errors

S. T. Huang,* K. C. Fan,† and John H. Wu*
*Institute of Optical Sciences, National Central University, Taiwan, Republic of China
†Department of Mechanical Engineering, National Taiwan University, Taiwan, Republic of China

A new minimum zone method for straightness error analysis is proposed in this article. Based on the criteria for the minimum zone solution and strict rules for data exchange, a simple and rapid algorithm, called the control line rotation scheme, is developed for the straightness analysis of planar lines. Extended works on the error analysis of spatial lines by the least parallelepiped enclosure are also described. Some examples are given in terms of the minimum zone and least-squares. Finally, this easy-to-use method is illustrated by an example that demonstrates that, for a planar line, the minimum zone solution can even be found without the use of a computer.

Keywords: *Straightness; minimum zone method; least-squares method*

Introduction

The term "straightness error" is generally used to refer to many aspects of engineering quality, such as workpiece straightness, motion straightness, etc. In order to assess this quality, engineering measurements along a particular reference line using appropriate techniques are generally made. The interpretation of the measured data has been specified using many different standards, such as BS 308: part 3: "geometrical tolerance"[1] and ANSI Y14.5M.[2] These specifications are all based on the minimum zone concept appearing in ISO/R1101,[3] which specifies the form errors in a general scope. It states that an ideal geometrical feature must be established from the actual measurements such that the maximum deviation between the ideal and the actual measurement concerned is the least possible value. The peak-to-valley distance of the deviation data from the ideal geometrical feature thus established is taken to represent the form error. The orientation of the ideal feature can be regarded as the alignment error in setting the reference axis with respect to the measured axis.

Although the least-squares method,[4] due to its simplicity in computation and the uniqueness of the solution provided, is most widely used in industry for determining straightness, it provides only an approximate solution that does not guarantee the minimum zone value. Therefore, during the past decade, much research has been devoted to finding the minimum zone solutions for straightness error and other form errors using various methods. Some researchers applied the numerical methods of linear programming, such as the Monte Carlo method, the simplex search and spiral search used by Murthy and Abdin,[5] the revised simplex search with dual problem used by Chetwynd,[6] the minimax approximation algorithm proposed by Fukuda and Shimokohbe,[7] and the simplex search technique adopted by Shunmugam[8] for the comparison of linear and normal deviations of straightness errors (where the deviation is taken perpendicular to the ideal line). Most of these works can be extended to determine the minimum zone solutions of other form errors, such as flatness, roundness, etc. Another type of approach has been to find the enclosing polygon for the minimum zone solution, such as the eigen-polygon method proposed by Hong and Fan,[9] the convex polygon method presented by Lai and Wang,[10] and the convex hall theory given by Traband et al.[11] These methods are more or less similar in their computational conception, which dynamically shows the meaning of each search step from the distribution of the data points. A new algorithm, the MINMAX method proposed by Fan and Burdekin,[12] uses the concept of the rotations of enclosing lines with respect to a particular contact point at each data exchange step. This technique

Address reprint requests to S. T. Huang, Institute of Optical Sciences, National Central University, Taiwan, Republic of China.

reduces the separation of the enclosing lines step by step until the minimum zone is found. The concept is comprehensive in both the physical and engineering senses in regard to the geometrical form of the investigated data.

The methods mentioned above generally proceed initially with the random selection of data points and then follow with an iterative data exchange procedure. A longer computation time is naturally required by this kind of approach in order to reach the final minimum zone condition. Burdekin and Pahk[13] in their revision of the MINMAX method, called the enclose tilt technique, use the least-squares result as the initial condition for the analysis of the flatness error. This method, while providing an efficient method for reaching the minimum zone solution, deals with all the measured points, the so-called full field, during each data exchange process. This process can be improved by screening out unwanted data points, which makes the mathematical model simpler and the computation time even shorter.

We must recognize, in the first place, that all the algorithms so far developed for the minimum zone solution of straightness error can guarantee an exact and unique solution of the minimum zone value, which must be smaller than the least-squares value. The computation times will be different depending on the complexity of the mathematical model that each algorithm uses. From an engineering point of view, in practice, the ability to understand the physical meaning of the algorithm is more important than the computation time of that algorithm, because the computation of each algorithm is fast, even using a personal computer. In reality, some algorithms are indeed difficult to understand because they are purely numerical analyses. Therefore, the simpler and clearer the algorithm, the more readily it will be accepted by an inspector who needs to know the straightness error from measured data. Practically, the best algorithm should provide not only the simplest model for analysis, but should require the least computational use of the computer. The development of such an algorithm is the main objective of this work.

This article presents a new minimum zone method for the straightness analysis of any planar line or spatial line investigated. This method modifies Burdekin's method[13] and is applied to straightness analysis by rotations of the enclosing lines in "half field" only. The data exchange scheme starts with a 1-1 model, where one control point is on one control line and another control point is on the other control line, based on the least-squares result. It then continues with a 2-1 model, where two control points are on one control line and the third control point is on the other control line, using the strict rule of the control line rotation scheme (CLRS), which is developed based on the criterion of the minimum zone solution. With only a few steps of data exchange in 2-1 model iteration, the minimum zone solution can be easily obtained. In addition, because most investigated lines have straightness errors in two directions, i.e., vertical and horizontal in space, being spatial lines, the interpretation of the straightness error of a spatial line is also proposed according to the definition appearing in the ISO standard.[3]

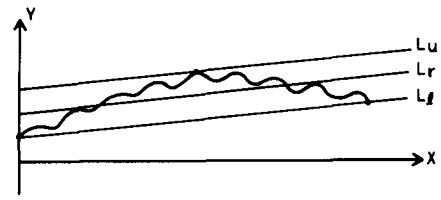

Figure 1 Minimum zone condition of the straightness error

Experimental tests on optical rails were conducted with an optical instrument capable of dual axis readings. Some examples are given including both planar lines and spatial lines. A handy method for planar line analysis is also illustrated without the use of a computer.

Straightness analysis of planar lines

Minimum zone criteria

The criteria for the minimum zone solution of straightness error have already been verified and adopted for use in numerous studies as indicated above.[6,7,12,13] Two conditions must be met in the final stages:

1. At least three points must be in contact with the two enclosing lines which are parallel.
2. These three points must lie on the enclosing lines in an upper-lower-upper sequence or a lower-upper-lower sequence.

Figure 1 illustrates the geometrical relationship of this minimum zone status. The distance between any two such enclosing lines defines the minimum zone of the straightness error. The contact points described here are called the control points and the corresponding enclosing lines are called the control lines for the purposes of this article. The mathematical models used to find this solution are described in the following sections.

Mathematical models

1-1 Model. Before searching for a best-fit line, one from which for a given set of data points the greatest deviations are at the minimum distance, it is best to find the least-squares line as the initial condition of the search, because in most cases it is close to the best-fit line.

Let the equation of the least-squares line be

$$Y = aX + b \qquad (1)$$

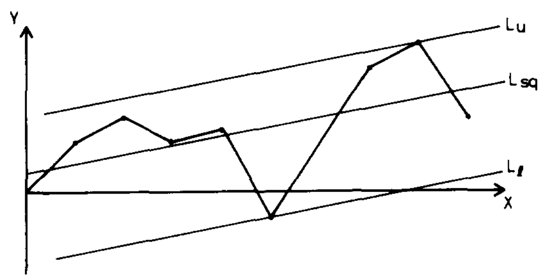

Figure 2 Construction of the 1-1 model from the least-squares line

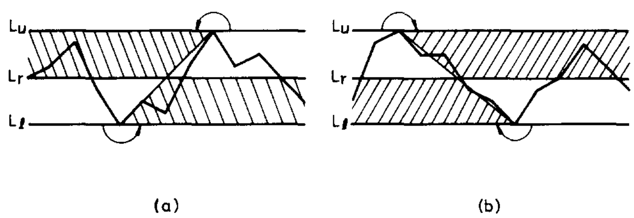

Figure 3 CLRS rule in half-field search.
(A) Lower-upper condition.
(B) Upper-lower condition

where a and b are coefficients that can be determined by the least-squares concept and variational principle. The deviation δY_i, or residual, of the straightness data from the least-squares line can then be defined as

$$\delta Y_i = Y_i - aX_i - b \qquad \text{for } i = 1, 2, \ldots, N \tag{2}$$

where N denotes the total number of data points. The highest of the data points with respect to this least-squares line is now defined as the upper control point, and the lowest point is defined as the lower control point. A 1-1 model is thus established in such a way that an upper control line is generated from the upper control point and a lower control line from the lower control point, with both lines being parallel to the least-squares line as shown in *Figure 2*. These lines actually enclose all of the data points.

2-1 Model. From the 1-1 model, two control points in association with two control lines are obtained by the least-squares method. In order to find the third control point conforming to the minimum zone solution, a strict rule of the CLRS is introduced. Here, each control line will rotate with respect to its corresponding control point in the direction that will most likely yield the result required by the second criterion of the minimum zone solution as indicated above. To meet this requirement, there are only two possible situations, depending on the occurrence of the control point and the determined direction of rotation, as illustrated in *Figure 3A and B.* If the space between two such control lines is defined as the full-field, during the rotation each control line will eventually find its own first new contact point, which must be located within the specified field. In other words, only the points within a specified quarter-field will be intersected by a particular control line for the determination of a new contact point. This CLRS process deals with data points in the half-field only, as shown by the shaded parts in *Figure 3.* In the computer algorithm, those unwanted points can be screened out automatically based on this concept. This should save half of the computational time required by a full-field search.

During the rotation of a particular control line, any point within the corresponding quarter-field may become the first contact point, depending on its position. Because each point within such a field will correspond to a rotation angle of the control line, the very first contact point must be the one having the smallest angle with respect to the control line. In the case of the search in the upper quarter-field, as shown in *Figure 4,* if we let E_{mu} be the deviation of the upper control point from the reference line, and E_i be the deviation of point i from the reference line, the angle of rotation of the control line from its initial position to the position as it contacts with point i is

$$\theta_i = \sin^{-1}\left(\frac{E_{mu} - E_i}{L_i}\right) \tag{3}$$

where L_i is the distance from the upper control point to point i. The very first contact point I within this quarter-field will be found by

$$\theta_I = \text{minimum}\,\{\theta_i\} \tag{4}$$

Similarly, in the lower quarter-field search, the angle of rotation of the lower control line from its initial position to the position where it contacts any point j in this field is

$$\theta_j = \sin^{-1}\left(\frac{E_{ml} - E_j}{L_j}\right) \tag{5}$$

where E_{ml} = deviation from the lower control point to the reference line, E_j = deviation from point j to the reference line, and L_j = distance from the lower control point to point j.

The first contact point J within this quarter-field will be found by

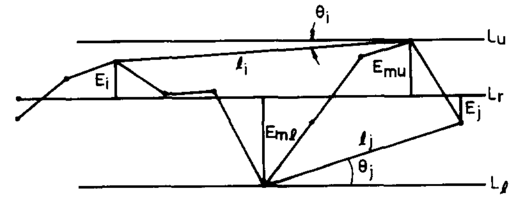

Figure 4 CLRS search from 1-1 model to 2-1 model

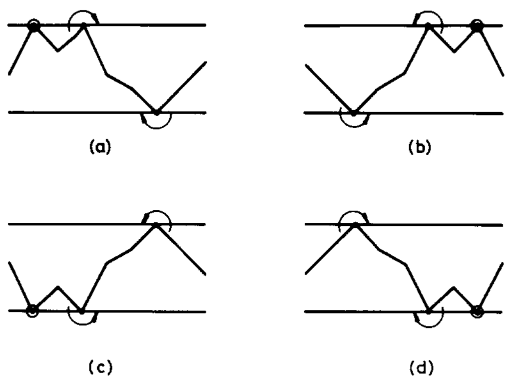

Figure 5 Four possible conditions requiring 2-1 model CLRS search

$$\theta_J = \text{minimum}\ \{\theta_j\}. \tag{6}$$

Although each control line may generate its own first point of contact during its own CLRS search, only one contact point is eventually needed to form the new 2-1 model. Therefore, selection between points I and J will be judged by the smaller value of θ_I and θ_J. If the final condition at the current stage meets the second criterion of the minimum zone solution, the job is done. However, if it is not, any of the four possible conditions shown in *Figure 5* may occur. Further search procedures will be required.

CLRS search in the 2-1 model. When the three control points of the current 2-1 model do not meet the minimum zone criterion, i.e., upper-lower-upper or lower-upper-lower, one of the control points must be discarded. It is obvious that this point must be the outside one on the two-point side, as circled in *Figure 5*. The remaining two control points will form a 1-1 model again, and the rule of CLRS for a 1-1 model is now applicable, pushing that discarded point inside the enclosing field. Equations (3) to (6) will be applied again to generate a new 2-1 model. This procedure will be iterated until the minimum zone criterion is reached.

Algorithm of the computer program. The algorithm of the computer program for the CLRS technique can now be summarized as follows:

1. Read the data (X_i, Y_i).
2. Get the straightness data with the least-squares method.
3. Construct the 1-1 model.
4. Apply CLRS rule to form a new 2-1 model.
5. Iterate the 2-1 model until the second minimum zone criterion is met.
6. Output the result.

Straightness analysis of spatial lines

Under most circumstances, straightness measurements address errors in two directions (vertical and horizontal) along a spatial line, such as the center line of a shaft, the guideway of a machine table, the moving axis of a linear stage, etc. These two error directions are usually treated separately in two orthogonal planes. Straightness errors are then evaluated and expressed in two directions.

An alternative interpretation of the straightness error of a spatial line is presented here based on the tolerance definition given by ISO/R1101.[3] According to this definition, the straightness error of any spatial line can be determined by two means: (1) the section $t_1 \times t_2$ of the smallest parallelepiped that can contain all the points, as seen in *Figure 6A,* or (2) the diameter of the smallest cylinder that can contain all the points, as shown in *Figure 6B.* The former problem can be solved by projecting data points onto two orthogonal planes on which the CLRS algorithm can be used. However, the latter one has a different structure, which is rather a spatial, or a cylindrical, problem. Because this is not directly related to the exchange algorithm of CLRS, this article addresses only the straightness error of the spatial line by parallelepiped enclosure.

The definition of minimum parallelepiped combines the results of two projected planar lines on two orthogonal planes, with minimum zone errors t_1 and t_2, respectively, to form an enclosing parallelepiped for all straightness data measured in two directions. Although this definition is clear, unfortunately, to our knowledge no mathematical expression of the straightness error of a spatial line has yet to be seen in any published report. Because this error value must be dependent on the cross-sectional size of the parallelepiped, in this article this type of straightness error is defined as the diagonal length of the rectangular cross-section of this parallelepiped, i.e.,

$$t_p = \sqrt{t_1^2 + t_2^2} \tag{7}$$

If we let a spatial line be directed in the X-direction with its straightness error measured in Y- and Z-directions simultaneously, the mean lines of the

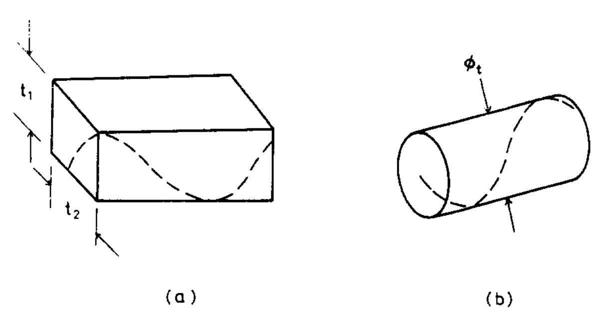

Figure 6 ISO definition of tolerance zone for spatial lines.
(A) Minimum parallelepiped enclosure.
(B) Minimum cylinder enclosure

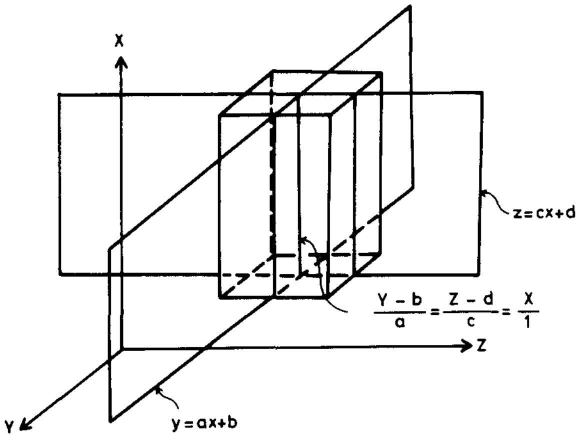

Figure 7 Intersecting two mean planes obtained from the projected lines to generate the parallelepiped axis

minimum zone solutions on two projected planes X-Y and X-Z, respectively, can be expressed as

$$\text{on X} - \text{Y plane,} \quad Y = aX + b \tag{8a}$$

$$\text{on X} - \text{Z plane,} \quad Z = cX + d \tag{8b}$$

where *a, b, c,* and *d* are coefficients that can be found by the CLRS method. In fact, Equations (8a) and (8b) can also be thought of as two plane equations in space, as shown in *Figure 7*. The intersecting line of these two planes represents the center line of the minimum enclosing parallelepiped, which can be expressed by the following form:

$$\frac{Y-b}{a} = \frac{Z-d}{c} = \frac{X}{1} \tag{9}$$

Having obtained this axis in space, all the measured data can be referred to this center line and the leveling work can be performed to remove the alignment error of the measuring axis with respect to the object axis.

Examples

There are quite a few instruments available on the market that have a capability for straightness measurements of planar lines, such as the straightedge with dial indicator, the laser interferometer, the electronic level, and the taut wire with sensor. There are also other instruments designed for the straightness measurements of spatial lines, such as the alignment telescope, the four-quadrant or dual-axis photodiode with collimated laser, and the dual-axis autocollimator.[14] Experimental work can easily be performed using appropriate equipment for any object being investigated. Some of the results of this type of work are presented below. The data points used for each of these examples are listed in the Appendix.

CLRS method for planar lines

In Example 1 the formal procedure for the minimum zone search was performed with a personal computer. The 1-1 model was first established from the least-squares line. The two control points thus obtained were points 6 and 9. Followed by only one CLRS process with the rotation rule, the minimum zone solution was readily obtained using three control points, 1, 6, and 9, as illustrated in *Figure 8*. Example 2 is a set of points from Murthy and Abdin.[5] Its straightness plot is given in *Figure 9*. *Table 1* shows the results of these two examples in terms of the least-squares and the minimum zone solutions.

An easy-to-use graphical CLRS method

The unique significance of the CLRS method when compared with all the other existing methods[4–13]

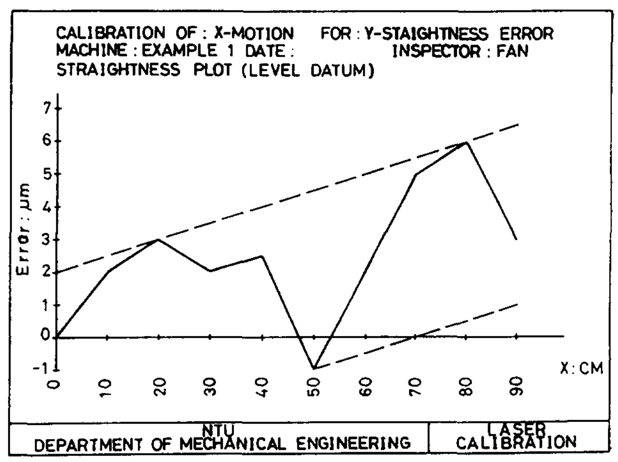

Figure 8 The minimum zone solution of example 1, the unleveled condition

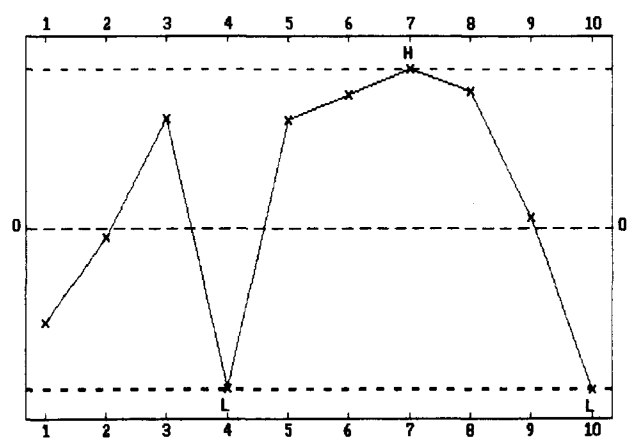

Figure 9 The minimum zone solution of example 2, the leveled condition

Table 1 Straightness errors of planar lines

	Example 1		Example 2	
	LSQ	CLRS	LSQ	CLRS
Coefficient				
a	0.0373	0.05	0.240	0.229
b	0.7727	−0.75	2.461	2.456
Error	5.882	5.5	0.913	0.880

LSQ, least-squares method; CLRS, minimum zone method.

is its strict rule of control line rotation. The least-squares line is only a reference for the construction of the 1-1 model. Because this method can efficiently provide the optimum strategy for data exchange, this reference line can even be any other line passing through the data points. The CLRS process can also be performed without the computer. *Figure 10* illustrates a series of procedures performed manually with a pair of triangles and a pen on graph paper with respect to the data points given in Example 1. *Figure 10A* shows a reference line drawn by linking two end points, which is deemed the simplest way to construct the reference line.

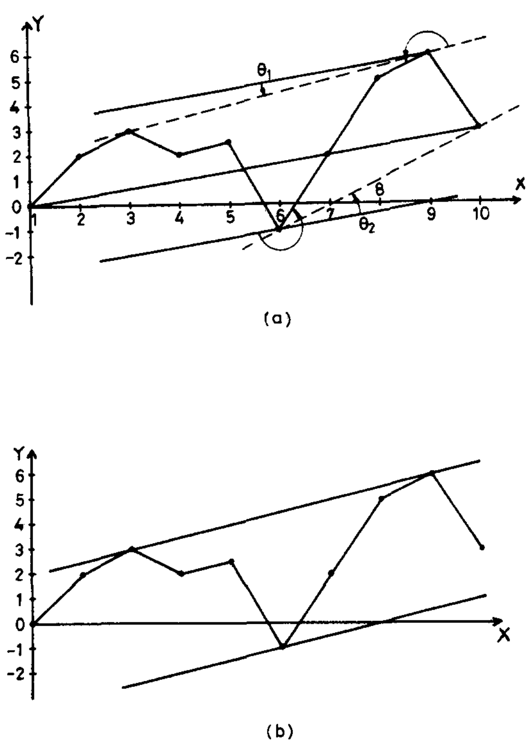

Figure 10 Graphical CLRS approach on the graph paper.
(A) CLRS search.
(B) Final result

The corresponding control points (6,9) and control lines (L_u, L_l) of the 1-1 model are determined accordingly. *Figure 10A* also shows two control lines that rotate around the corresponding control points in the directions prescribed by the CLRS rule. This rotation will result in two new contact points (3,10). Having compared the angles of rotations (θ_1, θ_2) of these two control lines, the contact point corresponding to the smaller angle (point 3) is thus selected as the new control point for the new 2-1 model. A judgment in accordance with the second criterion of the minimum zone solution is then made. In this example, the CLRS is now completed with the same result as in *Figure 8*. The minimum zone error can then be evaluated either from the diagram by approximation or by exact calculation using a computer.

From our experience, in most cases only one CLRS process is needed to reach the minimum zone solution. On very few occasions will one or two further CLRS processes be required to reach the final solution.

Straightness error for spatial lines

Two optical rails were measured using a collimated laser and a four-quadrant photodiode in the laboratory. Their data points are used in Examples 3 and 4, respectively. *Table 2* summarizes the analyzed minimum zone results compared with the least-squares solutions. A leveled three-dimensional graphical view of Example 3 is shown in *Figure 11* where the spatial points are noted by the symbol x and the projected points on two orthogonal planes are given by dots. It is clearly seen that on the X-Y plane the control points for minimum zone solution are points 4, 7, and 10, whereas on the X-Z plane there will be control points 1, 4, and 8. The straightness error of this spatial line can therefore be obtained by Equation (7).

Discussions and conclusions

This article proposes a concise and comprehensive data exchange algorithm, the CLRS, for the minimum zone solution of straightness error. It gives a simple and clear concept to which not only the minimum zone error can be guaranteed by the straightness criterion, but can be modified to operate without the use of a computer.

Although this article deals with residuals only in the normal to a nominal (instrument) direction, these residuals can be easily corrected to the normal to the reference line direction at the final stage if anyone is interested. However, this is not the major concern of this article, because the difference is relatively small if we remember that the errors are in microns and the step length of measurement is on the order of centimeters or inches.

Table 2 Straightness errors of spatial lines

	Example 3		Example 4	
	LSQ	CLRS	LSQ	CLRS
Coefficient				
a	0.765	0.783	−2.776E-04	−4.874E-04
b	−0.945	−1.75	−0.049	0.1035
c	−0.925	−0.957	−3.011E-03	−2.528E-03
d	−2.455	−2.413	0.4950	0.3849
Error				
t1	7.036	6.5	0.656	0.6129
t2	15.236	14.286	0.868	0.7697
tp		15.695		0.9839

t1, x-y plane; t2, x-z plane; tp, parallelepiped enclosure.

As indicated in the Introduction, the computing time required by the error analysis of any existing algorithm is less significant, because it must be short enough in comparison with the time consumed in the error measurements. The CLRS method, therefore, only highlights its features of simplicity (half-field search with a common equation), clarity (graphical explanation), and on-site applicability (manual operation).

An extended work of the CLRS method has been conducted to the minimum parallelepiped enclosure for spatial lines, and the new definition for this application, t_p, is proposed. However, the algorithm of minimum cylinder enclosure for spatial lines has not yet been found. This is regarded as a complicated case which might be investigated in different ways.

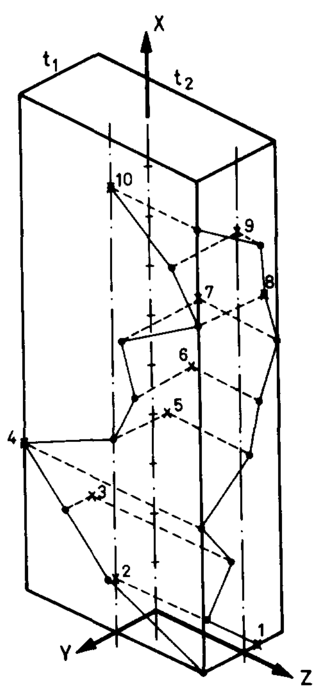

Figure 11 The leveled three-dimensional graphical plot of example 3

References

1 BS 308: Part 3. "Geometric tolerance," 1972
2 ANSI Y14.5M. "Dimensioning and tolerancing for engineering drawings," 1982
3 ISO/R1101. "Technical drawings—geometrical tolerancing," 1983
4 Miller, R. *Engineering Dimensional Metrology.* London: Edward Arnold Co., 1962
5 Murthy, T. S. R. and Abdin, S. Z. "Minimum zone evaluation of surfaces," *Int J Mach Tool Des Res* 1980, **20**, 123–136
6 Chetwynd, D. G. "Applications of linear programming to engineering metrology," in *Proc Instn Mech Eng* 1985, **199**, 93–100
7 Fukuda, M. and Shimokohbe, A. "Algorithms for form evaluation methods for minimum zone and least squares," in *Proc Int Symp Metrology for Quality Production,* Tokyo, 1984, pp. 197–202
8 Shunmugam, M. S. "Comparison of linear and normal deviations of forms of engineering surfaces," *Prec Eng* 1987, **9**, 96–102
9 Hong, J. T. and Fan, K. C. "An algorithm for straightness calculation from geometrical viewpoint," in Proc. 1st ROC-ROK Metrology Standard Symposium, Taiwan, 1986, 89–94
10 Lai, K. and Wang, J. "A computational geometry approach to geometric tolerancing," in *16th NAMRC,* 1988, 376–379
11 Traband, M. T. et al. "Evaluation of straightness and flatness tolerances using minimum zone," *Manufacturing Rev* 1989, **2**, 189–195
12 Fan, K. C. and Burdekin, M. "Straightness error analysis: an introduction to MINI-MAX1 method," in Proc 1st ROC-ROK Metrology Standard Symposium, Taiwan, 1986, 103–110
13 Burdekin, M. and Pahk, H. J. "The application of a microcomputer to the on-line calibration of the flatness of engineering surfaces," *Proc Instn Mech Engrs* 1989, **203**, 127–137
14 Burdekin, M. "Instrumentation and techniques for the evaluation of machine tool errors," *Technol Machine Tools* 1980, **5**, 1–37

Appendix: straightness data points

Example 1

No.	x(cm)	y(μm)
1	0	0
2	10	2
3	20	3
4	30	2
5	40	2.5
6	50	−1
7	60	2
8	70	5
9	80	6
10	90	3

Example 2

No.	x	y
1	1	2.428
2	2	2.891
3	3	3.445
4	4	2.931
5	5	3.895
6	6	4.196
7	7	4.497
8	8	4.662
9	9	4.545
10	10	4.303

Example 3

No.	x(cm)	y(μm)	z(μm)
1	0	−3	5
2	10	9	−12
3	20	15	−25
4	30	25	−38
5	40	29	−40
6	50	36	−48
7	60	42	−59
8	70	51	−62
9	80	59	−74
10	90	72	−88

Example 4

No.	x(cm)	y(mm)	z(mm)
1	0	0.41	0
2	25.4	0	0.124
3	50.8	−0.108	0.205
4	76.2	−0.17	0.306
5	101.6	−0.112	0.352
6	127	−0.068	0.387
7	152.4	−0.05	0.326
8	177.8	−0.15	0.248
9	203.2	−0.302	0.256
10	228.6	−0.286	0.1
11	254	−0.22	−0.068
12	279.4	−0.18	−0.262
13	304.8	−0.148	−0.558
14	330.2	−0.078	−0.74
15	355.6	−0.178	−0.899
16	381	−0.22	−0.956
17	406.4	−0.272	−0.942
18	431.8	−0.334	−0.928
19	457.2	−0.266	−0.88
20	482.6	−0.126	−0.98
21	508	0	−1.04
22	533.4	0.15	−0.83