

Author's Accepted Manuscript of paper

Rossi, A.; Lanzetta, M.: Optimal blind sampling strategy for minimum zone roundness evaluation by metaheuristics, Precision Engineering, vol. 37, 2013, pp. 241-247 (7), DOI:[10.1016/j.precisioneng.2012.09.001](https://doi.org/10.1016/j.precisioneng.2012.09.001).

Published by Elsevier, available on-line from  
[ScienceDirect.com](http://ScienceDirect.com)

## **Optimal blind sampling strategy for minimum zone roundness evaluation by metaheuristics**

*Andrea Rossi and Michele Lanzetta*

Department of Civil and Industrial Engineering  
University of Pisa, Via Diotisalvi 2, 56122 Pisa, Italy  
Email: lanzetta@unipi.it tel.: +39 050 2218122 fax: +39 050 2218140

### **Abstract**

The minimum zone tolerance is a non linear method to find a global solution to the roundness evaluation problem. Metaheuristics such as genetic algorithms, ant colony systems and particle swarm optimization concurrently process a set of solution candidates (chromosomes, ants, particles etc.) within a given search-space. Computation experiments carried out with an effective genetic algorithm have shown that the optimal sampling strategy providing sufficient accuracy at acceptable processing time represents a compromise between number of sample points and search-space size. An estimate of the neighborhood of the centroid containing the minimum zone center is given.

**Keywords:** minimum zone tolerance (MZT); roundness error; genetic algorithms; CMM

### **1. Introduction**

The growing complexity of shapes of manufactured parts and assembly tasks and the increase of performance demand to mechanical products requires high-speed inspection. Evaluation of form errors of machined parts is fundamental in quality inspection to verify their conformance to the expected tolerances. Performance of methods have been reviewed in [1].

Figure 1: MZ error  $E_{MZ}$ .  $c_1$  and  $c_2$  are possible locations of the centers of the two concentric circles.  $\Delta r_1$  and  $\Delta r_2$  are the differences in radii. If the minimal difference in radii  $\Delta r_2$  is the  $E_{MZ}$ ,  $c_2$  is the MZ center.

About here

Form tolerance is evaluated with reference to a Euclidean geometric feature, i.e. a circle in the case of roundness (also known as circularity). Roundness is a typical geometric form to be inspected as well as other typical forms such as straightness, flatness and cylindricity.

The most used criteria to establish the reference circle are: the least-squares method (LSQ), the maximum inscribed circle (MIC), the minimum circumscribed circle (MCC) and the minimum zone tolerance (MZT).

The use of a particular data fitting method depends on the required application, e.g. MIC and MCC can be used when mating is involved. The LSQ is one of the methods used by Coordinate Measuring Machines (CMM). It is efficient in computation and can be used with a large number of measured points, but the roundness error determined is larger than those determined by other methods, such as the MZT. Therefore, good parts can be rejected resulting in an economic loss. The MZT meets the standard definition of the roundness error, as reported in ISO 1101 [2]. It determines two concentric circles that contain the roundness profile and such that the difference in radii is the least possible value. Figure 1 shows two pairs of concentric circles that include the sample points centered respectively at  $c_1$  and  $c_2$  and where  $\Delta r_1$  and  $\Delta r_2$  are their difference in radii. Once the MZ center is found, the minimum zone error can be considered as the roundness error.

The MZT is a non linear problem and two approaches have been proposed in the literature: computational geometry techniques and solutions of a non linear optimization problem. The first approach is, in general, very computationally intensive, especially, when the number of data points is large. One of these methods is based on the Voronoi diagram [3]. The second approach is based on the minimization of the minimum zone error as a function of the MZ center, but the inconvenience is that this function has several local minima. Some examples are: the Chebyshev approximation [4], the simplex search / linear approximation [5] [6], the steepest descent algorithm [7], the particle swarm optimization (PSO) [8] [9], the simulated annealing (SA) [10], and genetic algorithms (GAs) [11] [12] [13] [14].

Xiong [15] develops a general mathematical theory, a model and an algorithm for different kinds of profiles including roundness where the linear programming method and exchange algorithm are used. As limaçon approximation is used to represent the circle, the optimality of the solution is however not guaranteed.

A strategy based on geometric representation for minimum zone evaluation of circles and cylinders is proposed by Lai and Chen [16]. The strategy employs a non-linear transformation to convert a circle into a line and then uses a straightness evaluation schema to obtain minimum zone deviations for the feature concerned. This is an approximation strategy to minimum zone circles.

M. Wang et al. [17] and Jywe et al. [18] present a generalized non-linear optimization procedure based on the developed necessary and sufficient conditions to evaluate roundness error. To meet the standards, the MZ reference circles should pass through at least four points of the sample points. This can occur in two cases: a) when three points lie on a circle and one point lies on the other circle (the 1-3 and the 3-1 criteria); b) when two points lie on each of the concentric circles (the 2-2 criterion). In order to verify these conditions the computation time increases exponentially with the dataset size. Gadelmawla [19] uses a heuristic approach to drastically reduce the number of sample points used by the min-max 1-3, 3-1 and 2-2 criteria.

Samuel and Shunmugam [20] establish a minimum zone limaçon based on computational geometry to evaluate roundness error; with geometric methods, global optima are found by exhaustively checking every local minimum candidate. Moroni and Petro [1] propose a technique to speed up the exhaustive generation of solutions (*brute force algorithm*), which starts with a single point and increases one sample point at each step in order to generate all the possible subsets of points, until the tolerance zone of a subset cover the whole dataset (*essential subset*).

A mesh based method with starting center on the LSC, where the convergence depends on the number of mesh cross points, representing a compromise between accuracy and speed, is proposed by Xianqing et al. [21].

The strategy to equally-spaced points sampled on the roundness profile is generally adopted in the literature. Conversely, in previous works the authors developed a cross-validation method for small samples to assess the kind of manufacturing signature on the roundness profile in order to detect critical points such as peaks and valleys [22] [23]. They use a strategy where a next sampling increasing the points near these critical areas of the roundness profile.

In [24], some investigations proved that the increase of the number of sample points is effective only up to a limit number. Recommended dataset sizes are given for different data fitting methods (LSQ, MIC, MCC, MZT) and for three different out-of-roundness types (oval, 3-lobing and 4-lobing). Similar works are [25] and [26] in which substantially the same results are given.

A sampling strategy depends on the optimal number of sample points and the optimum search-space size for best estimation accuracy, particularly with datasets that involve thousands of sample points available by CMM scanning techniques. In this paper, the sampling strategy problem tailored for a fast genetic algorithm to solve the MZT problem is addressed. To achieve more general

results, the sampling strategy used in this work can be defined as *blind* according to the classification in [27]. By sampling strategy not only the number and location of sample points on the roundness profile is addressed, but also their use by the data fitting algorithm [28].

Based on current experience, only few contributions are available in the literature regarding the sampling parameters, particularly with genetic algorithms. In [12] the search-space is a square of fixed 0.2 mm side, in [14] it is 5% of the circle diameter and center. In [11], the side is determined by the distance of the farthest point and the nearest point from the mean center. In [13] it is the rectangle circumscribed to the sample points. The optimal selection of the number of sample points and the search-space represent the main focus of current work.

## 2. Genetic algorithms for the MZT problem

To experimentally assess the sampling strategy with metaheuristics (such as genetic algorithms, ant colony systems, particle swarm optimization, taboo search etc.) a previously optimized genetic algorithm [14] has been selected. Genetic algorithms constitute a class of implicit parallel search methods especially suited for solving complex optimization or non-linear problems. They are easily implemented and powerful being a general-purpose optimization tool. Many possible solutions are processed concurrently and evolve with inheritable rules, e.g. the elitist or the roulette wheel selection, so to quickly converge to a solution, which is very close or coincident to the optimal solution.

Genetic algorithms maintain a population of center candidates (the *individuals*), which are the possible solutions of the MZT problem. The center candidates are represented by their *chromosomes*, which are made of pairs of  $x_i$  and  $y_i$  coordinates. Genetic algorithms operate on the  $x_i$  and  $y_i$  coordinates, which represent the inheritable properties of the individuals by means of genetic operators. At each generation the genetic operators are applied to the selected center candidates from current population in order to create a new generation. The selection of individuals depends on a fitness function, which reflects how well a solution fulfills the requirements of the MZT problem, e.g. the objective function.

Sharma et al. [29] use a genetic algorithm for MZT of multiple form tolerance classes such as straightness, flatness, roundness, and cylindricity. Because of the small dataset size (up to 100 sample points), there is no need to optimize the algorithm performance, by choosing the parameters involved in the computation.

Wen et al. [30] implement a genetic algorithm in real-code, with only crossover and reproduction operators applied to the population; thus in this case mutation operators are not used. The algorithm proposed is robust and effective, but it has only been applied to small samples.

A fast genetic algorithm with convergence speed greater than 0.1  $\mu\text{m}$  per 30 generations, within a selected stop condition, has been developed for large manufacturing samples and validated by certified software in [14]. The authors state that larger datasets require higher population size and not significantly affect the probability of crossover within a wide range. They conjecture that mutation is not a fundamental operator.

Table 1 lists all the parameters with their mechanism and value used by the data fitting algorithm proposed here. The optimal values of the genetic operators  $P_s$ ,  $P_c$  and  $P_m$  are taken from [14]. The genetic algorithm starts with a population of 70 center candidates ( $P_s$ ), randomly chosen in a search-space  $S_{r(x,y,\theta)}$  centered in  $C_n$  defined later in expression (2). At each generation the center candidates with their minimum zone reference circles and difference in radii are simultaneously evaluated for fitness by expression (1) also introduced later.

Table 1: Algorithm parameters and their description.

SAMPLING AND GENETIC PARAMETERS	SYMBOL	VALUES	COMMENT
<b>dataset size</b>	$n$	10÷10,000	Number of equally-spaced sample points on the roundness profile.
<b>search-space size [mm]</b>	$E_n$	0.1÷10	Radius of a circle where the initial population is randomly selected, centered in (1).
<b>population size</b>	$P_s$	70	Number of center candidates at each generation to be included in the search-space. The population density $\Delta$ can be estimated as  [ $P_s / \pi (E_n)^2$ ]
<b>selection</b>		elitist	Mechanism to select a center candidate for crossover operation. The center candidates are ordered for selection depending on their fitness function. The next generation includes the best chromosomes chosen between the set of offspring and the current population.

---

<b>crossover probability</b>	$P_c$	0.7	Mechanism to generate new center candidates by inheriting coordinates from parents. The number of parents selected for crossover is $P_c \times P_s$ . The arithmetic crossover mechanism used generates offspring as a component-wise linear combination of parents.
<b>mutation probability</b>	$P_m$	0.07	Mechanism to generate new center candidates by changing the coordinates with a random value. The number of the center candidates selected for crossover is $P_m \times P_s$ .
<b>stop criterion</b>	$G$	100	Mechanism to stop the genetic algorithm after $G$ generations with no improvements rounded off to the fourth decimal digit (0.1 $\mu\text{m}$ ).

---

### 3. Problem formulation

The minimum zone error  $E_{MZ}$  is the solution of the following optimization problem [14]:

$$\min \left[ \max_{\theta_i = i \times \frac{2\pi}{n}, i=1, \dots, n} r(x, y, \theta_i) - \min_{\theta_i = i \times \frac{2\pi}{n}, i=1, \dots, n} r(x, y, \theta_i) \right] \quad (1)$$

subject to  $(x, y) \in S_{r(x,y,\theta_i)}$

where  $\theta_i = i \times \frac{2\pi}{n}$ ,  $i=1, \dots, n$  are the angular positions of a number of equally-spaced points of the roundness profile  $r(x,y,\theta_i)$  of the reference circle of center  $(x,y)$  and radius  $R$ ;  $S_{r(x,y,\theta_i)}$  is the search-space, which is function of the sample points at  $\theta_i$  and their dataset size  $n$ .

The search-space is an area enclosed by the roundness profile where the center candidates of the initial population are selected for the data fitting algorithm. This area is rectangular because the crossover operator changes the  $x_i$  and  $y_i$  coordinates of the parents to generate offspring [11]. After crossover, the  $x_i$  and  $y_i$  coordinates of parents and offspring are located at the rectangle vertexes as shown in Figure 2.a and Figure 2.b.

Figure 2: Limit conditions to generate offspring by the crossover operator such that  $C_{MZ} \in S_{r(x,y,\theta)}$ .

a)  $C_{MZ}$  at the boundary of  $S_{r(x,y,\theta)}$ . b)  $C_{MZ}$  at the intersection of the boundary of  $S_{r(x,y,\theta)}$  and a principal axis of the reference system; c) position of centroids  $C$  and  $C_n$ , for increasing  $n$ :  $n1 < n2 < n3$  and position of  $C_{MZ}$  located at a distance smaller than  $E_C$  from  $C$  by (6).

About here

In order to solve the MZT problem, the search-space must include the global optima i.e. the MZ center,  $C_{MZ}$ , or at least it must be approached in order to find good solutions. Therefore we are interested to locate the center of the search space around  $C_{MZ}$  by a fast procedure.

The centroid  $C_n$  of  $n$  equally-spaced points:

$$C_n = \left( \frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i \right) \quad (2)$$

gives a fast estimation of the MZ center  $C_{MZ}$  [11] [12] [13] [14]. It can be considered itself as the center of  $S_{r(x,y,\theta)}$  if the size of the search space is conveniently large. Figure 2 shows the limit conditions such that  $C_{MZ}$  can be generated by crossover from parents:  $C_{MZ}$  must be located exactly at the boundary of the region  $S_{r(x,y,\theta)}$  of center  $C_n$  (Figure 2.a and Figure 2.b):

$$|C_n - C_{MZ}| = E_n \quad (3)$$

where  $E_n \in \mathfrak{R}$  is the search-space *size* (for example, the radius or the side of respectively circular and square features).

Only circular search-spaces are considered, neglecting that some center candidates are external to the feature (even if they fall outside the circular feature, if they are included in the circumscribed square). In fact:

*Lemma 1* – Consider  $S_{r(x,y,\theta)}$  as a circular feature of radius  $E_n$ . If the initial population of the genetic algorithm is included in  $S_{r(x,y,\theta)}$  then all center candidates of next generation are included in the square circumscribed to  $S_{r(x,y,\theta)}$ .  $\square$

■ The  $x_i$  and  $y_i$  coordinates of the center candidates of the initial population included in  $S_{r(x,y,\theta)}$  are also included in the square circumscribed to  $S_{r(x,y,\theta)}$  that has a side  $E_n$  and is centered to  $C_n$ . Therefore the  $x_i$  and  $y_i$  coordinates of the center candidates satisfy the following condition:

$$\begin{cases} C_n(x) - E_n \leq x_i \leq C_n(x) + E_n \\ C_n(y) - E_n \leq y_i \leq C_n(y) + E_n \end{cases} \quad (4)$$

For any two center candidates with coordinates  $(x_1, y_1)$  and  $(x_2, y_2)$  that verify expressions (4), the crossover generates offspring  $(x_1, y_2)$  such that:

$$\begin{cases} C_n(x) - E_n \leq x_1 \leq C_n(x) + E_n \\ C_n(y) - E_n \leq y_2 \leq C_n(y) + E_n \end{cases} \quad (5)$$

Therefore we infer that  $(x_1, y_2)$  and  $(x_2, y_1)$  themselves are included in the square circumscribed to  $S_{r(x,y,\theta)}$ .

□

The eccentric shape [31] is the major feature that maximizes the distance  $E_n$ . Figure 3 shows the worst-case achieved by considering a hypothetical roundness profile described by two concentric opposite arcs: the outer ( $A_R$ ) and the inner ( $A_r$ ) arcs of radii, respectively,  $R$  and  $r$ . An inset splits  $A_R$  into two identical arcs of length  $\alpha$ . When  $\alpha$  tends to zero, the two arcs become two semicircles with a degenerate set bounded to a singular point of coordinates  $(r, 0)$ .

Figure 3: Graphical proof of *Lemma 2* for  $-\frac{\pi}{2} \leq \alpha < \frac{\pi}{2}$ .

About here

*Lemma 2* – Increasing the number of equally-spaced points of the roundness profile ( $n$ ),  $C_{MZ}$  is trapped in a circular neighbor centered at the centroid  $C_n$  defined by:

$$E_n = (\sqrt{1 + \pi^2} - 1)^{-1} E_{C_n} \quad (6)$$

where  $E_{C_n}$  is the roundness error related to the centroid. □

■ To prove the lemma, a worst-case is considered: the feature of Figure 3 where either  $\alpha$  and  $r$  tend to zero.

For the boundary condition (3),  $E_n$  is equal to  $|C_n - C_{MZ}|$ . Let  $E$  be the distance between the centroid  $C$  and  $C_{MZ}$ . It results:

$$E = \lim_{n \rightarrow +\infty} E_n = \lim_{n \rightarrow +\infty} |C_n - C_{MZ}| = |(\lim_{n \rightarrow +\infty} C_n) - C_{MZ}| = |C - C_{MZ}| \quad (7)$$

□

*Evaluation of the minimum zone error*

According to ISO 1101, the  $C_{MZ}$  of  $n$  points sampled on the feature of Figure 3 is at (0,0), while the ratio  $\frac{r}{R}$  is greater than a given threshold  $t$ . While  $0 \leq \frac{r}{R} < t$ ,  $C_{MZ}$  is closer to  $C$  than (0,0).

Therefore for expression (7):

$$E = |C| \quad (8)$$

*Evaluation of the boundary condition (3) when  $r$  tends to 0*

The centroid of the outer and of the inner semicircles are on the X axis, respectively, at  $C_R = \frac{2 \cdot R}{\pi}$

and at  $C_r = \frac{2 \cdot r}{\pi}$ . From expression (8) it results that the distance  $E$  is:

$$E = \lim_{r \rightarrow 0} |C - C_{MZ}| = \lim_{r \rightarrow 0} \frac{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} R \cdot \cos \theta \cdot d\theta + \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} r \cdot \cos \theta \cdot d\theta}{\int_0^{2\pi} d\theta} = \frac{R}{\pi} \quad (9)$$

*Evaluation of  $E_c$*

Let  $RC_c$  and  $RI_c$  be the radii of, respectively, the minimum circumscribed and the maximum inscribed concentric circles centered at the centroid. By expression (9) it results:

$$RC_c = \sqrt{\left(\frac{R}{\pi}\right)^2 + R^2} = E\sqrt{1 + \pi^2} \quad (10)$$

$$RI_c = \frac{R}{\pi} = E \quad (11)$$

Hence, the roundness error  $E_c$  related to the centroid  $C$  is:

$$E_c = RC_c - RI_c = E(\sqrt{1 + \pi^2} - 1) \quad (12)$$

*Asymptotic condition*

For expressions (7) and (12):

$$\lim_{n \rightarrow +\infty} E_n = \lim_{n \rightarrow +\infty} (\sqrt{1 + \pi^2} - 1)^{-1} E_{C_n} \quad (13)$$

Therefore, increasing the number of sample points  $n$ , an  $n^*$  exist such that for  $n > n^*$  the thesis of the lemma is verified.  $\square$

*Lemma 2* has extensive practical applications, because  $E_{C_n}$  in expression (6) can be evaluated in linear function of the sampling data size; in fact, the centroid  $C_n$  is evaluated by expression (2), i.e. in closed form of the sampling data.

Figure 4 shows an experimental visualization of *Lemma 2* increasing  $n$ ;  $C_n$  converges to  $C$  for  $n > n^*$ .  $n^*$  is the optimum sampling size because it provides sufficient accuracy and minimum processing time to the algorithm.

Figure 4: Coordinates of  $C_n$  [mm] versus dataset sizes  $n$ , ranging from 10 to 10,000 points generated with certified software [32] with radius  $R=20$  mm and exact minimum zone error

$$E_{MZ}^*=0.06 \text{ mm.}$$

About here

#### 4. Computation experiments

The datasets used in computation experiments are generated with NPL Chebyshev best fit circle certified software [32].

The use of certified software has the following benefits:

- it produces randomly distributed error making the results not manufacturing signature-specific;
- the Chebyshev best-fit circle center  $C_{MZ}$  and the exact minimum zone error  $E_{MZ}^*$  are known, so it allows evaluating at each generation the error  $E_{MZ}$  of the algorithm in the estimate of  $C_{MZ}$ .

Several datasets are generated, with maximum residual deviation values from 0.01 to 0.09 mm, and sizes from 10 to 10,000 points. The maximum residual deviation is equal to half the exact minimum zone error  $E_{MZ}^*$  [33].

All datasets have center in the axes origin (0,0) and a radius  $R=20$  mm. The search-space  $S_{r(x,y,\theta)}$  is a circular feature centered in  $C_n$  according to *Lemma 1* with radius  $E_n$ .

The genetic algorithm has been executed 30 times for each experimental condition and the average  $E_{MZ}$  and the average computation time have been determined.

## 5. Results

As shown by computation experiments in Figure 5, the processing times of genetic algorithm tested with the configuration in Table 1 increase linearly with the dataset size and are barely affected by the exact minimum zone error  $E_{MZ}^*$ .

Figure 5: Average computation time of 30 runs of the genetic algorithm with the parameters in Table 1 using 3 GHz Pentium processor with increasing datasets size, different roundness errors and

$E_n=0.5$  mm.

About here

From Figure 6 it can be noticed that the minimum zone error  $E_{MZ}$  decreases with  $E_n$  while  $C_{MZ}$  is included in the search-space (1÷3 mm). Afterwards it starts increasing again because the density  $\Delta$  of the  $P_s$  center candidates decreases with increasing  $E_n$ .

In the range  $E_n=1\div3$  mm, by increasing the dataset size  $n$  up to 50÷100,  $E_{MZ}$  decreases and remains stable afterwards in the range 0.062÷0.063 mm, a good approximation of the exact minimum zone error  $E_{MZ}^*$  of 0.06 mm.

With small search-spaces ( $E_n=0.1\div1$  mm)  $E_{MZ}$  is very high and barely decreases with  $n$  because the exact minimum zone center may be outside the search-space.

The range for the sampling parameters  $n$  and  $E_n$  reported above provides the highest accuracy for the genetic algorithm in the search of the minimum zone error (lowest  $E_{MZ}$ ).

Figure 6: Minimum zone error  $E_{MZ}$  versus the number of sample points  $n$  and the search-space size  $E_n$ . The values of  $E_{MZ}$  are the average of 30 runs of the genetic algorithm with the parameters in

Table 1 and exact minimum zone error  $E_{MZ}^*=0.06$  mm.

About here

In order to optimize the algorithm speed the above range needs to be matched with the processing time in Figure 7.

Figure 7: Average computation time of 30 runs of the genetic algorithm with the parameters in Table 1 and  $E_{MZ}^*=0.06$  mm versus the number of sample points  $n$  and the search-space size  $E_n$ .

About here

In Figure 7 it can be noticed that the computation time increases both versus  $n$  and  $E_n$ . Consequently the optimal values for  $n$  and  $E_n$  are the minimal providing sufficient accuracy. The linear increase of the computation time versus  $n$ , while  $E_n$  is below 3 mm, complies with that of Figure 5. Beyond this limit, the density  $\Delta$  of the center candidates in the search-space is too low, therefore the genetic algorithm can be trapped in local optima, providing some faster but less accurate solutions. A population density  $\Delta=2.5$  chromosomes/mm<sup>2</sup> is the corresponding limit to  $E_n=3$  mm.

## 6. Discussion

It has been shown both theoretically, by a worst-case approach, and experimentally, using a genetic algorithm, that increasing the dataset size the  $E_{MZ}$  estimation accuracy increases, but also the computation time increases. The computation time increase in Figure 5 is not the only drawback of increasing the sample size  $n$ .

Figure 6 shows the beneficial effect of increasing  $n$  and the optimal range for  $E_n$ . However this graph cannot be considered alone, the effect of the same parameters on the processing time in Figure 7 should also be taken into account. For example, the average  $E_{MZ}$  (0.06398 mm) for  $n=75$  and  $E_n=2$  mm is close to that for  $n=100$ , with the same search-space (0.06209 mm, the minimum value achieved), but for  $n=75$  there is a computation time reduction from 2.297 to 0.836 s. This shows that achieving sufficient algorithm accuracy instead of the best one may produce significant practical benefits. Inversely, the lowest processing time (0.134 s) is obtained with the lowest values from Table 1 of  $n$  and  $E_n$  (respectively 10 and 0.1 mm), but the corresponding  $E_{MZ}$  is very high (4.66571 mm).

The effect of different combinations of  $n$  and  $E_n$  is summarized in Figure 8. An interpretation for the  $E_{MZ}$  increase (low algorithm accuracy) for the following four conditions is given.

- Low  $n$  and low  $E_n$ . Small datasets do not provide sufficient estimation accuracy. The optimal solution is not included in the search-space. Lowest processing time, the algorithm stops because it is trapped in local minima and does not decrease with subsequent generations.
- Low  $n$  and high  $E_n$ . The optimal solution is included in the search-space, but the small datasets do not provide sufficient estimation accuracy of the MZ center  $C_{MZ}$ . Increasing the population density  $\Delta$  is worthless. Premature convergence (low processing time), the algorithm stops because it is trapped in local minima.

- High  $n$  and low  $E_n$ . Larger datasets yield a better estimation of  $C_{MZ}$ , but the optimal solution is not included in the search-space. In addition  $n$  increases the computation time because each center candidate is compared with  $n$  sample points in order to assess its fitness.
- High  $n$  and high  $E_n$ . Larger datasets yield a better estimation of  $C_{MZ}$ , but larger search-spaces are dangerous because the algorithm is trapped in local minima. In this case the population density  $\Delta$  should be increased accordingly. The processing time is greatly affected by the lower convergence speed of the algorithm (MZT problem is not linear).

Figure 8: Possible causes and effects of different combinations of  $n$  and  $E_n$  on the algorithm accuracy and computation time. The optimal sampling parameters are  $n^*$  and  $E_{n^*}$ .

About here

The proposed sampling strategy is summarized in Figure 9, which enhances the operative aspects:

1. increasing the dataset size  $n$  until optimal or sufficient accuracy is achieved (sufficient accuracy depends on the manufacturing requirements);
2. finding a suitable value for the search-space size.

Once the optimal or suboptimal dataset size and the corresponding search-space are selected, the processing time is determined accordingly.

Figure 9: The proposed optimal sampling strategy.

About here

The proposed method can be considered as a guideline for assessing the optimal sampling parameters with different metaheuristics.

## 7. Conclusions

In this work the sampling strategy for the roundness evaluation of circular profiles using the MZT method has been optimized. In particular,

- it has been proven that an upper bound for the centroid to minimum zone center distance is given by  $(\sqrt{1+\pi^2}-1)^{-1} E_c$ ; consequently increasing the dataset size, this neighborhood of the centroid  $C_n$  is a good candidate for the search-space by metaheuristics;
- computation experiments with a genetic algorithm have shown a linear increase of computation time versus dataset size.

Based on theoretical considerations and extensive computation experiments, it can be concluded that there is an optimal value for the dataset size and the search-space providing the highest accuracy and lowest computation time. The population density  $\Delta$  seems a discriminating factor for the algorithm performance.

This pattern is helpful to take out indications on the optimal number of sample points and corresponding search-space size by metaheuristics in roundness evaluation for the whole class of optimization problems defined by equation (2) – the MZT problem. In addition, this sampling strategy can be defined as blind or not manufacturing signature-specific because sample points are equally-spaced and sample errors are randomly distributed.

## **Acknowledgements**

Computation experiments have been derived from Alessandro Meo and Luca Profumo's project work for the class of automation of production processes of the master degree in automation engineering at the school of engineering at Pisa University.

The authors would like to thank the anonymous referees for their insightful contributions to the theoretical analysis.

## **References**

- [1] Moroni, G. and Petro, S., Geometric tolerance evaluation: A discussion on minimum zone fitting algorithms, *Precision Engineering*, Volume 32, Issue 3, July 2008, 232-237, ISSN 0141-6359, DOI: 10.1016/j.precisioneng.2007.08.007.
- [2] International Organization for Standardization, Geneva, Switzerland, ISO 1101, Geometrical Product Specifications (GPS)–tolerances of form, orientation, location and run out, 2nd ed.; December 2004.
- [3] Roy, U. and Zhang, X., Establishment of a pair of concentric circles with the minimal radial separation for assessing roundness error, *Computer Aided Design*, 1992, 24(2), 161–168.
- [4] Dhanish, P.B. and Shunmugam, M.S., An algorithm for form error evaluation – using the theory of discrete and linear chebyshev approximation, *Computer Methods in Applied Mechanics and Engineering*, Vol. 92, 1991, 309–324.
- [5] Weber, T., Motavalli, S., Fallahi, B., Cheraghi, S.H., A unified approach to form error evaluation, *Journal of the International Societies for Precision Engineering and Nanotechnology* 26, 2002, 269–278.
- [6] Murthy, T.S.R. and Abdin, S.Z., Minimum zone evaluation of surfaces, *International Journal of Machine Tool Design Research*, 20, 1980, 123–136.
- [7] Zhu, L.M., Ding, H. and Xiong, Y.L., A steepest descent algorithm for circularity evaluation, *Computer Aided Design*, 2003, 35, 255–265.
- [8] Yashpal Kovvur, Hemant Ramaswami, Raj Bardhan Anand and Sam Anand, Minimum-zone form tolerance evaluation using particle swarm optimisation, *Int. J. Intelligent Systems Technologies and Applications*, Vol. 4, Nos. 1/2, 2008.

- [9] Mao, J., Cao, Y., Yang, J., Implementation uncertainty evaluation of cylindricity errors based on geometrical product specification (GPS), *Measurement*, 42(5), June 2009, 742-747.
- [10] Shakarji, C.M. and Clement, A., Reference Algorithms for Chebyshev and One-Sided Data Fitting for Coordinate Metrology, *CIRP Annals - Manufacturing Technology*, 53(1), 2004, 439-442.
- [11] Rohit Sharma, Karthik Rajagopal, and Sam Anand, A Genetic Algorithm Based Approach for Robust Evaluation of Form Tolerances, *Journal of Manufacturing Systems* Vol. 19/No. 1 2000.
- [12] Wen, X., Xia, Q., Zhao, Y., An effective genetic algorithm for circularity error unified evaluation, *International Journal of Machine Tools & Manufacture* 46, 2006, 1770-1777.
- [13] Yan, L., Yan, B., Cai, L., Hu, G., Wang, M., Research on roundness error evaluation of shaft parts based on genetic algorithms with transfer-operator, 9th International Conference on Electronic Measurement & Instruments, 2009. ICEMI '09, 2-362 - 2-366. doi:10.1109/ICEMI.2009.5274566.
- [14] Rossi, A.; Antonetti, M.; Barloscio M.; Lanzetta, M.: Fast genetic algorithm for roundness evaluation by the minimum zone tolerance (MZT) method, *Measurement*, vol. 44, n. 7, August 2011, ISSN 0263-2241, DOI: 10.1016/j.measurement.2011.03.031, 1243-1252 (10).
- [15] Xiong, Y.L., Computer aided measurement of profile error of complex surfaces and curves: theory and algorithm, *International Journal Machine Tools and Manufacture*, 1990, 30, 339-357.
- [16] Lai, J. and Chen, I., Minimum zone evaluation of circles and cylinders, *International Journal of Machine Tools and Manufacturing*, 1995, 36(4), 435-51.
- [17] Wang, M., Cheraghi, S.H. and Masud, A.S.M., Circularity error evaluation: theory and algorithm, *Precision Engineering*, 1999, 23(3), 164-76.
- [18] Jywe, W.-Y, Liu G.-H., Chen C.-K., The min-max problem for evaluating the form error of a circle, *Measurement*, 1999, 26, 273-282.
- [19] Gadelmawla, E.S., Simple and efficient algorithm for roundness evaluation from the coordinate measurement data, *Measurement*, 2010, 43, 223-235.
- [20] Samuel, G.L. and Shunmugam, M.S., Evaluation of circularity from coordinate and form data using computational geometric techniques, *Precision Engineering*, 2000, 24, 251-263.
- [21] Xianqing, L., Chunyang, Z., Yujun, X., Jishun, L., Roundness Error Evaluation Algorithm Based on Polar Coordinate Transform, *Measurement*, In Press, Accepted Manuscript, Available online 23 October 2010.
- [22] Rossi, A., A form of deviation-based method for coordinate measuring machine sampling optimization in an assessment of roundness, *Proc Instn Mech Engrs, Part B: Journal of Engineering Manufacture*, 2001, 215, 1505-1518.
- [23] Rossi, A., A minimal inspection sampling technique for roundness evaluation, In first CIRP International Seminar on PRogress in Innovative Manufacturing Engineering (Prime), Sestri Levante, Italy, June 2001.
- [24] Chajda, J., Grzelka, M., Gapinski, B., Pawłowski, M., Szelewski, M., Rucki, M., Coordinate measurement of complicated parameters like roundness, cylindricity, gear teeth or free-form surface, 8 International conference advanced manufacturing operations
- [25] Gapinski, B. and Rucki, M., Uncertainty in CMM Measurement of Roundness, AMUEM 2007 - International Workshop on Advanced Methods for Uncertainty Estimation in Measurement Sardagna, Trento, Italy, 16-18 July 2007.
- [26] Gapinski, B., Grzelka, M., Rucki, M., Some aspects of the roundness measurement with cmm, XVIII IMEKO World Congress Metrology for a Sustainable Development September, 17 - 22, 2006, Rio de Janeiro, Brazil.

- [27] Colosimo, B.M., Moroni, G. and Petro, S., A tolerance interval based criterion for optimizing discrete point sampling strategies, *Precision Engineering*, Volume 34, Issue 4, October 2010, 745-754, ISSN 0141-6359, DOI: 10.1016/j.precisioneng.2010.04.004.
- [28] Weckenmann, A.; Weber, H.; Eitzert, H.; Garmer, M.: Functionality-oriented Evaluation and Sampling Strategy in Coordinate Metrology. *Precision Engineering*, 17 (1995) 4, 244-252.
- [29] Sharma, R., Rajagopal, K. and Anand, S., A genetic algorithm based approach for robust evaluation of form tolerances, *Journal of Manufacturing Systems*, 2000, 19(1), 46–57.
- [30] Wen, X., Xia, Q. and Zhao, Y., An effective genetic algorithm for circularity error unified evaluation, *International Journal of Machine Tools and Manufacture*, 2006, 46, 1770–1777.
- [31] Whitehouse, D.J., *Handbook of surface metrology*, IOP Pub., 1994.
- [32] National Physical Laboratory (UK), Data Generator for Chebyshev Best-Fit Circle, <http://www.npl.co.uk/mathematics-scientific-computing/software-support-for-metrology/>, last accessed January 2011.
- [33] Goch, G. and Lübke, K., Tschebyscheff approximation for the calculation of maximum inscribed/minimum circumscribed geometry elements and form deviations, *CIRP Annals - Manufacturing Technology* 57, 2008, 517–520.