Measurement 127 (2018) 134–140

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

A novel hybrid trust region minimax fitting algorithm for accurate dimensional metrology of aspherical shapes

Yassir Arezki[a,b,*], Hichem Nouira[a,*], Nabil Anwer[b], Charyar Mehdi-Souzani[b]

[a] Laboratoire Commun de Métrologie (LCM), Laboratoire National de Métrologie et d'Essais (LNE), 1 Rue Gaston Boissier, 75015 Paris, France
[b] LURPA, ENS Cachan, Univ. Paris-Sud, Université Paris-Saclay, 94235 Cachan, France

ARTICLE INFO

Keywords:
Trust region methods
Aspheric shapes
Minimax
Chebyshev fitting
Minimization
Form error
Dimensional metrology
Ultra-high precision engineering

ABSTRACT

Ultra-high precision measuring machines enable to measure aspheric shapes with an uncertainty of few tens of nanometres. The resulting clouds of points are then associated to theoretical model at the same level of accuracy so as to obtain parameters that indicate about form error. Minimum zone (*MZ*), defined as the least value of peak to valley (*PV*), is widely used to assess form error. Least squares method (L_2) is often used to determine *MZ* but the resulting value is usually overestimated. For this reason, L_2 is replaced by L_∞ norm because it gives a more accurate value of *MZ* since it directly minimizes *PV*. Using L_∞ norm results in a non-smooth optimization problem and consequently its resolution becomes more challenging compared to L_2.

In this paper, a novel minimax fitting method for accurate metrology of aspheres and freeform based on a hybrid trust region algorithm (HTR) is proposed. To assess performance of the introduced method, it was compared to an available minimax fitting algorithm based on a smoothing technique: exponential penalty function (EPF). The choice of EPF is justified by superior performances in comparison to existing techniques. Comparison was conducted on reference data, data available in literature and data gathered form measurements of a real optical high quality asphere. Results show superiority of HTR over EPF in both returned *PV* values and execution time.

1. Introduction

Aspheres and freeform optics have replaced spherical components in several optical systems due to their superiority over classical (spherical) elements especially for eliminating spherical aberrations [1]. The emergence of new manufacturing techniques such as glass and plastic moulding and grinding as well as polishing methods expands its fields of application in medical imaging, lasers, astronomy, etc. [2].

Form quality of optical aspheres and freeforms is crucial to their performance and functionality. For this reason, form deviations must be tracked all over components' lifetime from design to operational use. Nowadays, available techniques allow manufacturing complex geometries and provide sub-micrometre-level corrections. On the other hand, form assessment of optical elements and data processing still a major issue [3]. Form assessment consists of determining whether form errors are within tolerance specifications. For complex shapes, aspherics for instance, data gathered from ultra-high precision CMMs must be treated in a way to give parameters that indicate about tolerance zone. One of these parameters is usually taken as the peak to valley (*PV*). Therefore, the least value of *PV* which corresponds to the minimum zone (*MZ*) must be determined (Fig. 1).

To determine the *PV*, deviations of data points from a reference surface must be determined *a priori*. There exist several ways to determine the reference surface but the one fitted according to a least squares (L_2) criterion is the widely used [3,4]. The main reason for using L_2 lies in simplicity when solving the resulting minimization problem compared to other criteria. Nevertheless, L_2 usually overestimates *MZ* which causes the rejection of a number of conforming parts. In another way, L_∞ criterion results in a direct minimization of *PV* and consequently returns the closest value of *MZ* to actual.

In this context, a European project 15SIB01-FreeFORM was launched in 2016 to develop reference L_∞ fitting algorithms and traceable metrology for aspheres and freeform optical lenses with below 30 nm accuracy [5].

In general, minimum zone determination problem could be mathematically formulated as follows:

$$\min_{\boldsymbol{x}} \phi(\boldsymbol{x}) \text{ where } \phi(\boldsymbol{x}) = \max_{1\leqslant i\leqslant m} f_i(\boldsymbol{x}) \text{ and } \boldsymbol{x} = \{\boldsymbol{T},\boldsymbol{s}\} \tag{1}$$

f_i is the Euclidean distance between the measured point (P_i) and its corresponding projection into the surface (Q_i), $\boldsymbol{x} \in R^n$ could be either the set of intrinsic shape parameters $\boldsymbol{s}$, or the motion parameters $\boldsymbol{T}$:

* Corresponding authors at: Laboratoire Commun de Métrologie (LCM), Laboratoire National de Métrologie et d'Essais (LNE), 1 Rue Gaston Boissier, 75015 Paris, France (Y. Arezki).
E-mail addresses: yassir.arezki@lne.fr (Y. Arezki), hichem.nouira@lne.fr (H. Nouira).

https://doi.org/10.1016/j.measurement.2018.05.071
Received 28 July 2017; Received in revised form 16 May 2018; Accepted 19 May 2018
Available online 21 May 2018

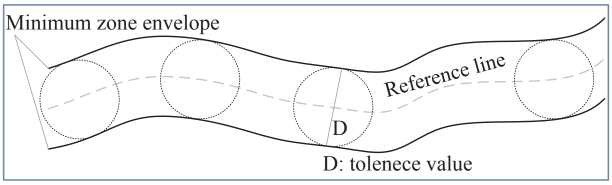

Fig. 1. Tolerance zone definition.

rotation and translation applied to $\{P_i\}$.

Even that L_∞ criterion gives a smaller value of *PV*, the formulated objective function is non-differentiable and the resulting problem is very difficult to solve since a wide range of derivative-based techniques could not be used.

The choice of a mathematical formulation to describe the aspheric is also crucial because it affects the obtained *MZ*. Although there exist several formulations to describe an aspheric lens: splines, Chebyshev polynomials, Zernike polynomials, etc. [6]. The one given by ISO 10110-Part 12:2007 [7] called monomial formulation is the most used. Its mathematical expression is presented in (2).

$$z(r) = \frac{r^2}{R\left(1 + \sqrt{1-(1+\kappa)\frac{r^2}{R^2}}\right)} + \sum_{m=0}^{M} a_{2m+4} r^{2m+4} \tag{2}$$

Where:

$z \rightarrow$ sag of surface
$r \rightarrow$ radial distance
$R \rightarrow$ radius of curvature
$\kappa \rightarrow$ conic constant
$a_{2m+4} \rightarrow$ monomial coefficients

The aperture size of the lens R_{max} defines the domain of r: $0 \leqslant r \leqslant R_{max}$ for which Eq. (2) is valid. The number of monomial terms M depends on the targeted accuracy. Despite its simplicity, this formulation represents some serious drawbacks especially those due to its numerical instability. Thus, when performing L_2 fitting, the resulting Gram matrix is usually ill-conditioned, which outcomes in less accuracy because of significant loss of digits. Other formulations were proposed to cope with these drawbacks. They consist of using orthogonal polynomials instead of monomials [8]. As consequence, the obtained Gram matrix is nearly diagonal and the resulting system is more stable.

This paper is structured as follows. In Section 2, an overview of minimum zone fitting methods is presented. In Section 3, the implementation of the hybrid trust region (HTR) algorithm is detailed. Validation of HTR against EPF is carried out on generated reference data as well as benchmark data in Section 4. In the last section, an investigation of a real case study of a measured high quality optical asphere is illustrated.

2. Literature review

Minimum zone determination for classical geometries such as lines, planes, circles and spheres has been extensively studied and different methods were developed [9]. Computational geometry techniques were used in [10–15] to determine minimum zone for straightness, flatness, circularity, cylindiricity and sphericity tolerance. This class of methods represent a major advantage since no derivative calculations are required. Furthermore, they can find the exact solution but their use is restricted to simple geometries and could not be extended to freeforms. Another free derivative method based on downhill simplex algorithm was proposed to determine straightness tolerance [16,17]. Genetic algorithms were also used for form error determination [18–20].

In regards to freeform shapes, many methods were developed for minimum zone assessment. A first approach makes use of L_p norm [21]. At each iteration the value of p is incremented and the corresponding L_p based smooth objective function is minimized using classical methods until a termination criterion is satisfied. This method suffers from serious instability especially when approaching the optimal solution because the resulting L_p based objective function becomes nearly non-differentiable. A heuristic method based on differential evolution algorithm (DE) was recently developed for freeforms [22]. This method performs poorly especially with large clouds of data points. Moreover, given results are not deterministic.

In order to make use of differentiation optimization techniques, the aggregation function method could be used. In [23,24], an exponential penalty function (EPF) is used to approximate the non-smooth objective function via a twice differentiable one. The resulting function could be minimized using Newton based method or any derivative-based optimization technique. This method gives good results but represents some instabilities due to exponential terms.

The minimum zone determination problem could be formulated as a nonlinear constrained problem. The formulation were detailed in [25], and a primal-dual interior point algorithm (PDIP) was implemented to solve the resulting problem. This method represents lower performances compared to EPF since minimization of the resulting Lagrangian function requires the resolution of large linear systems with ill-conditioned matrices.

3. Hybrid trust region algorithm (HTR)

The main idea of the hybrid trust region algorithm consists of performing either trust region step, line search step or curve search step according to the specific situation faced at each iteration [26,27]. It enables to avoid solving the trust region problem many times. For every iteration, a first stage consists of obtaining a trust region trial step d_k by solving the quadratic problem given in (3).

$$QP(x_k,\boldsymbol{B_k}): \begin{cases} \min_{(d,z)\in R^{n+1}} \frac{1}{2} < d,\boldsymbol{B_k} d > + z = M_k(d,z), \\ s.\ t. < \nabla f_i(x_k),d > -z \leqslant \phi(x_k) - f_i(x_k), i = 1,\dots,m \\ \|d\|_\infty \| \leqslant \Delta_k \end{cases} \tag{3}$$

where $\boldsymbol{B_k}$ is n by n symmetric positive definite matrix, Δ_k is the parameter defining the trust region domain, z is an introduced parameter depending on the first derivative of the objective function ϕ, ∇f_i is the gradient of the function f_i and $'' < .\,, . > ''$ denotes the dot product.

The trust region domain is defined using L_∞ instead of L_2 so as *QP* becomes an easily-solved quadratic problem. It should be mentioned that the proposed *QP* in (3) has always a solution since (0,0) lies inside the feasible domain. This problem could be solved using classical methods adapted to quadratic problems such as interior point method [28].

If the resulting trust region trial step d_k could not be accepted, a corrected step $d_k + \widetilde{d}_k$ is determined by solving the problem in (4).

$$\widetilde{QP}(x_k,\boldsymbol{B_k}): \begin{cases} \min_{(\widetilde{d},z)\in R^{n+1}} \frac{1}{2} < d_k + \widetilde{d},\boldsymbol{B_k}(d_k + \widetilde{d}) > + \widetilde{z} = \widetilde{M}_k(\widetilde{d},\widetilde{z}), \\ s.\ t. < \nabla f_i(x_k),\widetilde{d} > -\widetilde{z} \leqslant \phi(x_k + d_k) - f_i(x_k + d_k), i = 1,\dots,m \\ \|d_k + \widetilde{d}_\infty\| \leqslant \Delta_k \end{cases} \tag{4}$$

If neither the initial trust region step d_k nor the corrected step $d_k + \widetilde{d}_k$ could be acceptable in trust region scheme, a line search along d_k or a curve search is performed if d_k is a descent direction (the actual reduction $r_k > 0$ in (6)). Otherwise ($r_k \leqslant 0$), a curve search is used to find a step length t_k that verifies (5).

$$\phi(x_k + t_k d_k + t_k^2 \widetilde{d}_k) \leqslant \phi(x_k) - \alpha t_k \langle d_k,\boldsymbol{B_k} d_k \rangle \tag{5}$$

where $\alpha \in (0,1/2)$, d_k is the solution of (3) and $\widetilde{d}_k$ is the solution of (4). In the case $\|d_k\| \leqslant \|\widetilde{d}_k\|$, $\widetilde{d}_k$ should be taken to be 0. The implemented algorithm follows the next steps: