Precision Engineering 38 (2014) 499–505

Contents lists available at ScienceDirect

Precision Engineering

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

The uncertainty of radius estimation in least-squares sphere-fitting, with an introduction to a new summation based method

Kevin J. Cross, John W. McBride, Joseph J. Lifton *

Electro-Mechanical Research Group, Faculty of Engineering and the Environment, University of Southampton, Highfield SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 12 March 2013
Received in revised form 17 January 2014
Accepted 20 January 2014
Available online 31 January 2014

Keywords:
Least-squares
Sphere-fitting
Design of experiment
Monte Carlo
Uncertainty
Surface metrology
Profilometer

ABSTRACT

This paper considers the sensitivity of three sphere-fitting algorithms to real-world measurement errors. It pays particular attention to nominally spherical surfaces, such as those typically measured by tactile and optical profilometers, addressing the limitations of sensor gauge range and angular tolerance. A recently proposed linear circle-fitting algorithm is extended to a sphere-fitting algorithm and its performance compared to two long standing sphere-fitting algorithms; namely linear and non-linear least-squares. Sources of measurement error in optical profilometers are discussed, and user defined scan parameters are optimised based on the results of a designed experiment. The performance of all three sphere-fitting algorithms are tested on a sphere superimposed with varying degrees of surface irregularities in a Monte Carlo simulation; this study shows that both linear routines display a negative skewness in their radius error distribution. Finally, a method of predicting radius uncertainty is offered that considers the surface residual that remains after sphere-fitting and relates this to the radius uncertainty of the chosen algorithm.

1. Introduction

The contact lens and other closely linked optical industries have been considering the use of raster-scanning metrology systems for quality assurance since the early 90's [1]. A key objective to quality assurance is for every lens to pass under an instrument to determine its key parameters, for example, radius of curvature, form errors, and lens thickness. In reality, the complex nature of the moulding processes prohibits this approach; it is far simpler to adopt a batch sampling regime. Even with this very limited number of items to measure, the soft, flexible nature of the product, combined with the sources of error inherent in surface metrology instruments makes it difficult and expensive to measure the radius of curvature [2–4].

In previous studies [5,6], sample data has been acquired using the XYRIS 4000, a state-of-the-art surface metrology instrument manufactured by TaiCaan Technologies Ltd. (Southampton, UK). This instrument is custom engineered for applications such as spherical form analysis, adopting a granite gantry to maximise thermal and mechanical stability and using high-precision, 10 nm resolution motion stages coupled with a 10 nm resolution, confocal optical probe. It has been observed that when analysing 3D surface scans there is an error in the estimate of radius of curvature with all sphere-fitting algorithms [6]. In the case of a raster-scanning surface profiler, such as the XYRIS 4000, the probe's gauge range and angular tolerance limit the maximum scan area for nominally spherical objects; this measurement area defines the segment angle of the sphere that is considered by any post-measurement fitting algorithm. Sun et al. [7] considered the impact of the scan area on the algorithms used for calculating the underlying radius of curvature, investigating the performance of non-linear least-squares (NLLS) for small segment angles.

While NLLS has been demonstrated to generally produce the best-fit [6,8,9], it is an iterative algorithm with its solution-time being dependant on both the start estimation and the convergence criteria; furthermore, its speed is also strongly influenced by the size of the data-set, N, following an N^2 relationship. Ultimately, the systematic and random errors present in any real-world measurement system, coupled with the influence of small segment angles, may negate any error reductions from adopting an NLLS algorithm. The aim of this research is to revisit the influence of error sources on the standard least-squares algorithms, while offering a new, alternative algorithm with reduced computational overhead and hence speed benefits.

Section 2 extends a recently proposed linear circle-fitting algorithm to a sphere-fitting algorithm, while Section 3 discusses sources of measurement error in optical profilometers. In Section 4, the impact of three, user-defined scan parameters on the estimated radius are evaluated in a designed experiment. Section 5 compares the performance of the new sphere-fitting algorithm to the existing linear and non-linear least squares algorithms by superimposing

* Corresponding author. Tel.: +44 023 8059 5568; fax: +44 023 8059 301.
E-mail address: J.J.Lifton@soton.ac.uk (J.J. Lifton).

0141-6359/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.precisioneng.2014.01.004

varying degrees of Gaussian noise on a simulated, ideal spherical surface in a Monte Carlo simulation. In Section 6, the results of the Monte Carlo simulation are used to predict the uncertainty of the estimated radius for measured data.

2. Least-squares sphere-fitting

The principle of least-squares fitting is well understood, namely finding a surface that best-fits a data-set by minimising the sum of the square of the residuals [10]. The linear relationship of the parameters of a plane make this minimisation task easy, however, the more complex relationships in the equation of a sphere complicate the mathematics. The problem of fitting a sphere to a set of points in space is not one exclusive to metrology, but finds its place in archaeology, geography, and has been a topic of research since the early 60's [11].

Two popular least-squares approaches for the sphere-fit exist; the first relies on linearising the equation of a sphere (linear least-squares, LLS) and has been known since 1974 [12]. The second relies on an iterative approach, such as Gauss Newton, to find the minimum (non-linear least-squares, NLLS). Forbes [13,14] provides full details of the two approaches adopted for this study.

2.1. Summation least-squares

An alternative to these approaches for circle-fitting was identified by Bullock [15]; the mathematics are expanded here to a sphere-fit to supplement the former methods. This routine, referred to here as summation least-squares (SLS), makes use of domain shifting and substitution to simplify the equation of a sphere, resulting in a direct solution of a third-order simultaneous equation. This considerably reduces the computational cost when compared to the usual Jacobian matrix solver required by the alternatives.

Assume, a spherical surface of radius, r, comprising of n points at (x_i, y_i, z_i). The data can be translated into a new coordinate space (u, v, w) such that it has a new centre, (u_c, v_c, w_c), nominally the origin. With the centre as the origin, assumptions can be made about the distribution of the coordinates that allow the least-squares minimising function to be linearised.

Applying the following definitions:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i, \tag{1}$$

and

$$u_i = x_i - \bar{x}, \quad v_i = y_i - \bar{y}, \quad w_i = z_i - \bar{z} \tag{2}$$

For a least-squares fit, we wish to minimise the function:

$$s = \sum_i g(u, v, w)^2 \tag{4}$$

where

$$g(u, v, w) = (u_i - u_c)^2 + (v_i - v_c)^2 + (w_i - w_c)^2 - \alpha \tag{5}$$

and

$$\alpha = r^2 \tag{6}$$

Following Bullock's approach, $S(u, v, w, \alpha)$ is partially differentiated with respect to u_c, v_c, and w_c to find the minima.

Using the following nomenclature:

$$\sum_i u_i = S_u, \quad \sum_i u_i^2 = S_{uu}, \quad \sum_i u_i^3 = S_{uuu}, \quad \sum_i u_i v_i = S_{uv}, \quad \text{etc.} \tag{7}$$

The partial derivatives are simplified to:

$$u_c S_{uu} + v_c S_{uv} + w_c S_{uw} = \frac{S_{uuu} + S_{uvv} + S_{uww}}{2} \tag{8}$$

$$u_c S_{uv} + v_c S_{vv} + w_c S_{vw} = \frac{S_{vuu} + S_{vvv} + S_{vww}}{2} \tag{9}$$

$$u_c S_{uw} + v_c S_{vw} + w_c S_{ww} = \frac{S_{wuu} + S_{wvv} + S_{www}}{2} \tag{10}$$

Eqs. (8)–(10) represent a third-order simultaneous equation, which can be written in the matrix form, $Ax = b$:

$$\begin{bmatrix} S_{uu} & S_{uv} & S_{uw} \\ S_{uv} & S_{vv} & S_{vw} \\ S_{uw} & S_{vw} & S_{ww} \end{bmatrix} \begin{bmatrix} u_c \\ v_c \\ w_c \end{bmatrix} = \frac{1}{2} \begin{bmatrix} S_{uuu} + S_{uvv} + S_{uww} \\ S_{uuv} + S_{vvv} + S_{vww} \\ S_{uuw} + S_{vvw} + S_{www} \end{bmatrix} \tag{11}$$

This is solved for u_c, v_c, w_c by pre-multiplying both sides by A^{-1}.

To find the radius, r:

$$\alpha = r^2 = u_c^2 + v_c^2 + w_c^2 + \frac{S_{uu} + S_{vv} + S_{ww}}{n} \tag{12}$$

And translating the centres back into (x, y, z) space:

$$x_c = u_c + \bar{x}, \quad y_c = v_c + \bar{y}, \quad z_c = w_c + \bar{z} \tag{13}$$

On evaluating the time taken to process 100,000 data-sets comprised of 51×51, 101×101 and 201×201 data points for LLS and SLS, both algorithms' timings scale approximately linearly with the number of data points. Furthermore, the SLS algorithm is approximately twice as fast as LLS.

3. Sources of measurement error

Looking at the elements of a raster-scanning, coordinate measurement machine, such as the XYRIS 4000, it is relatively straightforward to identify sources of potential error. These are categorised into two areas, motion system errors and sensor errors.

Motion system errors may include:

- Stage resolution and accuracy – the smallest measurable step and the relationship between theoretical and real-world position;
- Stage run-out – while at a nominally linear velocity during scanning, there is an acceleration and deceleration phase;
- Stage pitch and roll – despite using high-end motion stages, the sample will still wobble microscopically as it moves;
- Thermal expansion – both the metal stages and the granite gantry expand and contract with changes in temperature;
- Non-orthogonal axes – alignment of the three different motion systems that combine to raster scan the sample and adjust the height of the sensor.

Sensor errors may include:

- Abbe Error – the sensor is unlikely to be aimed perfectly perpendicular to the measurement stages;
- Linearity – the response of the probe does not follow the actual displacement;
- Spot size – the finite focal spot-size causes an averaging effect;
- Sensor noise – random ripple on the output, even in steady-state conditions;
- Thermal response – temperature effects on air density;
- Angular tolerance – the slope of the sample may be too high to return light;
- Gauge range – the range of surface heights over which the sensor will operate.

In addition to the above errors, there are also a number of parameters that the user configures at the start of the measurement, all