Precision Engineering 54 (2018) 412–419

Contents lists available at ScienceDirect

Precision Engineering

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

Space-filling scan paths and Gaussian process-aided adaptive sampling for efficient surface measurements

Chengfei Yang[a], Chaoyang Peng[a], Yuhang Chen[a,*], Tingting Luo[b], Jiaru Chu[a]

[a] Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, Hefei, 230026, Anhui, China
[b] The 38th Research Institute of China Electronics Technology Group Corporation, Hefei, 230088, Anhui, China

ARTICLE INFO

Keywords:
Adaptive sampling
Surface measurement
Hilbert curve
Spiral scan

ABSTRACT

A method combining space-filling scan paths and adaptive sampling is proposed for surface measurements. Scan paths including a fractal Hilbert curve and a spiral pattern are mainly investigated. The adaptive sampling is based on iterative Gaussian process (GP) inference. Sampling positions are intelligently suggested along the scan path and the final sampled data are trained in a GP-model to reconstruct the entire topography. Simulations and experiments on different surfaces demonstrated the capability of the proposed method. When the special scan paths are employed alone, the required data amount is reduced to about 10%–13% of the uniform sampling and the relative error of surface reconstruction is within 10%. If the GP-aided adaptive sampling is further integrated, the data amount can be reduced to approximately 3%–4%. In addition, time-consumption in scanning is significantly eliminated. Compared with the raster scan, the integration of special scan paths and GP-aided adaptive sampling has several prominent advantages such as eliminating data amount, preserving surface reconstruction accuracy, maintaining a single-pass scan and saving time-cost. The measurement method has a potential application in situations where the efficiency is of critical importance.

1. Introduction

Measurement of surface topography is frequently involved in numerous science and engineering fields. In massive topography characterization, it is highly demanded that the methods could have the characteristics of high efficiency, small data amount and high accuracy. For scanning measurement techniques such as atomic force microscopy (AFM) and coordinate measuring machine (CMM), the scan rate may be not so satisfactory. Toward the main purpose of enhancing efficiency without sacrificing surface reconstruction accuracy, design of scan trajectory and optimization of sampling strategy have drawn considerable interests.

Two categories of general approaches have been investigated to improve the efficiency. The first one is the proper design of scan path. For example, the conventional scan in most scanning probe microscopy (SPM) instruments is in a raster manner with an equal interval. The scan rate is typically around 1 Hz, which means that it takes 1 s per scan line. For a common image with 256 × 256 pixels, the required time is then approximately 4.3 min. The image acquisition is time-consuming and it hinders the applications in either monitoring dynamic phenomena of a sample or massive topography measurements where much faster scanning is needed. One of the main factors limiting the scan rate is that the higher harmonic frequency components of the drive signal will excite the SPM scanner resonance and distort the measured results [1]. To overcome this barrier, several special scan patterns have been proposed. These scan paths including sinusoidal [2], spiral [3], cycloid [1] and Lissajous [4] patterns have been demonstrated to increase the allowable rate by one or two orders of magnitudes and they are easily implementable in practical instruments because no additional modifications of the hardware components are necessary [5]. In addition, the data amount may also be reduced as compared with the raster scan.

The other approach is based on undersampling or adaptive sampling. The most convenient way to scan the topography is at an equal interval. However, such a simple sampling strategy may lead to a large amount of redundant data, for instance, in the flat regions. To eliminate the unnecessary points, undersampling or adaptive methods can be adopted [6]. From determination of the local surface slope or curvature by interpolation of neighboring points, the sampling interval can be adjusted according to the topographic variation [7–9]. In case of a preknown surface model, the implementation of such an approach is much easier because the topographic change can be estimated in advance. AFM measurements based on compressed sensing principles [10,11]

This paper was recommended by Associate Editor Han Haitjema.

* Corresponding author.
E-mail address: chenyh@ustc.edu.cn (Y. Chen).

https://doi.org/10.1016/j.precisioneng.2018.07.011
Received 15 January 2018; Received in revised form 30 July 2018; Accepted 31 July 2018
Available online 03 August 2018

and random-walker scan patterns [12] have also been proposed. These methods are aimed to reduce the data amount as much as possible while keeping satisfactory surface reconstruction accuracy. The determined sampling locations are usually distributed in an irregular way. Though the necessary data amount could be significantly reduced, the total scan distance covering all the sampled positions may be even larger as compared with the conventional raster scan. Furthermore, additional probe approaching/retracting at each sampling position may be required. These factors lead to unwanted time-consumption on moving the probe or the sample. If the efficiency is of critical importance, approaches based on compressed sensing and random walker principles are much more suitable for discrete point-by-point measurements [13]. Based on a reasonable inference of the unknown surfaces, intelligent sampling methods employing kriging models [14] and Gaussian process (GP) regressions [15–17] have been developed recently for coordinate metrology and surface topography characterization. In fact, the kriging models there were founded on GP models and they were generally similar. The adaptive sampling demonstrated more distinctive performances than uniform sampling and other low-discrepancy sampling patterns [18].

A special scan path can improve the efficiency and adaptive sampling can reduce the data amount. The combination of the two approaches is assumed to be able to reduce the scan distance and the data density simultaneously. The measurement efficiency is expected to be improved and the surface topography reconstruction can keep the same order of accuracy level as dense uniform sampling. Based on this hypothesis, we integrate special scan patterns and adaptive sampling toward the main purpose of reducing data amount and time-cost for efficient surface topography measurement without sacrificing the accuracy. A Hilbert-curve path and a spiral path are employed for the demonstration purpose. The two paths can be considered as representatives of the scanning in a Cartesian coordinate system and a polar coordinate system, respectively. The integrated adaptive sampling is based on iterative GP-inference of the topography variation along the scan path and the most effective sampling position is suggested subsequently. The final sampled point clouds are used to reconstruct the entire surface via GP-modeling as well.

2. Methods

2.1. Scan trajectory

As a proof-of-concept, we integrate the GP-aided adaptive sampling with some special continuous scan patterns. Two types of scan paths, namely Hilbert curves and spiral patterns, are schematically illustrated in Fig. 1. The two-dimensional (2D) Hilbert curve is a space-filling fractal curve that passes each location in the scan area exactly once (see Fig. 1(a) and (b)). It can be generated recursively in a self-similar manner [19]. From the diagrams, a Hilbert curve of order n is composed of four Hilbert curves of order n-1, which are connected by three connector lines. The Hilbert curve enables an image-to-line mapping. For an arbitrary location having the coordinate (i, j), its previous and next sampling locations must be among the four coordinates (i-Δx, j), ($i+\Delta x$, j), (i, j-Δy) and (i, $j+\Delta y$) indicating that the scan path contains only horizontal and vertical scan steps. With a higher curve order, the data density in the same square area increases owing to the decreased intervals Δx and Δy. The orders of the Hilbert curves are respectively 3 and 4 in Fig. 1(a) and (b), for example. The data points in Fig. 1(b) are much denser than those in Fig. 1(a).

The spiral path starts from the center and it scans in both x and y directions simultaneously. In a polar coordinate system, the spiral path can be described as $r = vt$ and $\theta = \omega t$ with r the instantaneous radial distance at time t. Parameters v and ω are respectively the radial velocity and the angular velocity. Adjusting the radial interval (Δr) and the angular interval ($\Delta\theta$) per unit time step, the data amount within the scan area can be altered as schematically shown in Fig. 1(c) and (d). In the followed implementation, the path parameters are determined so that the data amount is roughly 20%–25% of the uniform sampling within the same scan area.

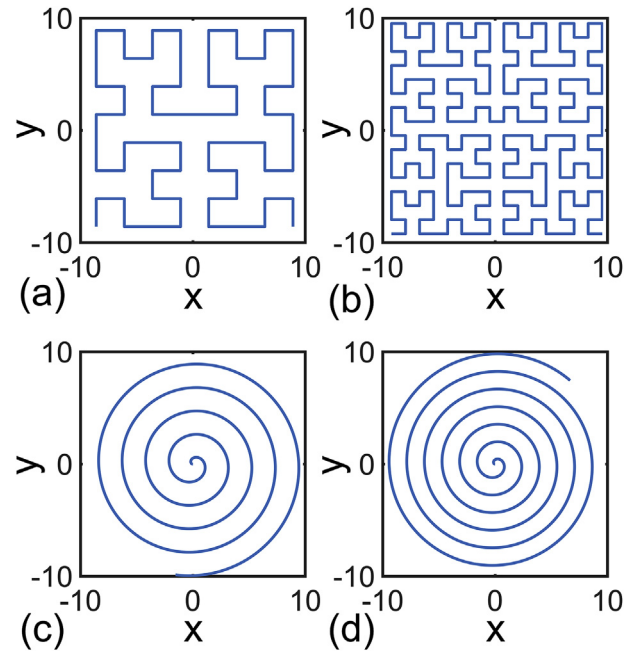

Fig. 1. Schematic illustration of the Hilbert-curve and the spiral scan. (a) (b) Hilbert curves with different orders. (c) (d) Spiral paths with different radial intervals.

2.2. Inference based on Gaussian process

GP-models have been intensively employed to infer unknown data by assuming that any subset follows a multivariate Gaussian distribution [20]. GP-inference is used here to guide the adaptive sampling along the scan path and to reconstruct the entire surface topography. Let us take the measurement of an arbitrary one-dimensional profile as an example. Given the current sampling points $z = \{z(x_1), ..., z(x_n)\}$, the main objective of GP-inference is to determine a best estimate of the profile height z_* at an unsampled position x_*.

A GP-model is characterized by a mean function and a covariance function, which can have many forms. The popular selections are that the mean function is zero everywhere and the covariance $k(x, x')$ is a squared exponential function [21],

$$k(x, x') = \sigma_f^2 \exp\left[-\frac{(x-x')^2}{2l^2}\right] \tag{1}$$

where σ_f^2 and l denote the covariance and the characteristic length, respectively. The followed covariance matrix can be then established,

$$\mathbf{k} = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \tag{2}$$

$$\mathbf{k}_* = [k(x_*, x_1) \ \ldots \ k(x_*, x_n)] \tag{3}$$

$$\mathrm{k}_{**} = k(x_*, x_*) \tag{4}$$

Because the surface data could be described as samples from a multivariate Gaussian distribution, we have,

$$\begin{bmatrix} z \\ z_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{k} & \mathbf{k}_*^{\mathrm{T}} \\ \mathbf{k}_* & \mathrm{k}_{**} \end{bmatrix}\right) \tag{5}$$

where the superscript T means the matrix transposition. The best estimate of height z_* is obtained as,