The International Journal of Advanced Manufacturing Technology (2023) 124:1341–1356
https://doi.org/10.1007/s00170-022-10576-7

CRITICAL REVIEW

Applications of data fusion in optical coordinate metrology: a review

Zhongyi Michael Zhang[1] · Sofia Catalucci[1] · Adam Thompson[1] · Richard Leach[1] · Samanta Piano[1]

Received: 11 March 2022 / Accepted: 22 November 2022 / Published online: 1 December 2022

Abstract

Data fusion enables the characterisation of an object using multiple datasets collected by various sensors. To improve optical coordinate measurement using data fusion, researchers have proposed numerous algorithmic solutions and methods. The most popular examples are the Gaussian process (GP) and weighted least-squares (WLS) algorithms, which depend on user-defined mathematical models describing the geometric characteristics of the measured object. Existing research on GP and WLS algorithms indicates that GP algorithms have been widely applied in both academia and industry, despite their use being limited to applications on relatively simple geometries. Research on WLS algorithms is less common than research on GP algorithms, as the mathematical tools used in the WLS cases are too simple to be applied with complex geometries. Machine learning is a new technology that is increasingly being applied to data fusion applications. Research on this technology is relatively scarce, but recent work has highlighted the potential of machine learning methods with significant results. Unlike GP and WLS algorithms, machine learning algorithms can autonomously learn the geometrical features of an object. To understand existing research in-depth and explore a path for future work, a new taxonomy of data fusion algorithms is proposed, covering the mathematical background and existing research surrounding each algorithm type. To conclude, the advantages and limitations of the existing methods are reviewed, highlighting the issues related to data quality and the types of test artefacts.

Keywords Metrology · Data fusion · Artificial intelligence · Machine learning

1 Introduction

1.1 Metrology and data fusion

The measurement of the external shape and surface texture of an engineered component is a ubiquitous activity in numerous fields from engineering to cultural heritage [1–3]. Optical coordinate metrology is the science and application of the measurement of the physical geometry of an object using instruments equipped with optical sensors [2]. Optical coordinate measurement technologies have been increasingly applied in the industry [4, 5]. A recent research theme in metrology is the development of data fusion pipelines for combining data acquired from multiple optical coordinate measuring systems.

✉ Zhongyi Michael Zhang
Zhongyi.Zhang@nottingham.ac.uk

1 Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, Nottingham, UK

As a broad range of processes and applications, data fusion, here defined as the combination of two or more datasets collected by multiple sensors, is used in various multidisciplinary subjects and research areas [6] [7], such as manufacturing [8], urban management [9], smart healthcare [10], and defence [11]. These research areas, combining data fusion with machine learning technologies, have driven researchers to explore more advanced algorithms and architectures [12]. Data fusion has gained increasing popularity because it allows improvement of measurement accuracy, expansion of measurement coverage (i.e. the dimensions that the measurement system is able to observe), or provision of more information than acquired using a single sensor system [13].

In this paper, we review research on data fusion algorithms proposed over the last 5 years, with a particular focus on the fusion of two or more datasets featuring three-dimensional (3D) coordinate information (i.e. point clouds) of engineered parts measured by optical instruments. Here, data fusion strictly refers to the registration of two or more 3D point clouds, and therefore, research on data fusion as a broader term, including its possible applications, is out of the scope of this review. Because of its popularity, numerous reviews have discussed the latest research in metrology for data fusion. For example, Pomerleau et al. [14] reviewed several iterative closest point (ICP) algorithm variations, used to register multiple point clouds of an object measured by an optical coordinate measuring system mounted on a robot; Wang et al. [15] provided a comprehensive introduction to the mathematical principles behind numerous data fusion techniques applied in surface metrology; Kong et al. [16] presented data fusion algorithms in manufacturing processes. The aforementioned reviews focus on data fusion algorithms proposed before 2017 and do not specifically introduce data fusion pipelines applied in the context of point cloud registration. Additionally, existing reviews do not include an in-depth discussion about how the mathematics of each type of algorithm influences its advantages and limitations in different measurement scenarios.

In this work, we investigated the latest proposed algorithms by observing their underlying mathematics, as we wanted to propose a method for selecting the most appropriate algorithm for an optical coordinate measurement task [17]. In the reviewing process, we categorised algorithms based on the same mathematical foundation into some categories, where algorithms in that category have similar advantages and limitations in measurement applications. With such a taxonomy, researchers can choose the most suitable algorithm for a given data fusion task. As we focus on the overall mathematics of a data fusion pipeline, individual techniques embedded in an algorithmic pipeline (such as Kalman Filters [18] and its variants, such as recursive filtering [19], particle filtering [20], inertial measurement unit [21] etc.) are not included in this discussion.

This review is structured as follows. In Section 1.2, we provide a general introduction to data fusion within the context of optical coordinate measurement of engineered parts. In Section 2, the methodology of collecting the information in this review is presented. In Section 3.1, the taxonomy of existing data fusion algorithms is introduced and explained. In this taxonomy, we classify the existing algorithms into three types: Gaussian process (GP), weighted-least-square (WLS), and machine learning algorithms. In Sections 3.2, 3.3, and 3.4, the mathematics behind each type of data fusion algorithms is presented and discussed. In Section 4, the latest research, including experimental and simulation results, for each algorithm type is presented. In Section 5.1, the benefits and limitations of each type of algorithm are discussed, in reference to the information presented in Section 4. In Section 5.2, we will particularly discuss the methodology of existing research (fashion in artefact selection and design). The information presented in this review is concluded in Section 6, and in Section 7, we propose an outlook for future research.

1.2 Data fusion: definitions

Data fusion was formally defined for the first time in 1987 by the Joint Directors of Laboratories (JDL) in the USA [22]. The initial definition in Data Fusion Lexicon, which was produced by the JDL, is as follows:

> "A process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats, and their significance. The process is characterized by continuous refinements of its estimates and assessments, and the evaluation of the need for additional sources, or modification of the process itself, to achieve improved results."

Later in 1991, the JDL adjusted the definition of data fusion, as a general technical term, to the following [23]:

> "A process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats as well as their significance."

Based on the two definitions given by the JDL, data fusion is today defined as a family of technologies that integrates and analyzes data collected by multiple sensors, employing a large variety of algorithms developed in recent years [24–26]. The most widely acknowledged definition of data fusion specifically within the context of metrology was given by Weckenmann et al. [27] in 2009, stated as follows:

> "Multi-sensor data fusion in dimensional metrology can be defined as the process of combining data from several information sources (sensors) into a common representational format in order that the metrological evaluation can benefit from all available sensor information and data. This means that measurement results can be determined, which could not – or only with worse accuracy – be determined solely on the basis of data from an individual source (sensor) only."

From this definition [16], the advantages of using data fusion in metrological contexts are clear. Particularly, the fusion of data from a multi-sensor system can be used to measure quantities that are not measurable using a single-sensor system [28]. Additionally, the potential benefits of data fusion for metrology include increased measurement coverage, increased data density, reinforced robustness to sensors and algorithms, and improved noise filtering and data accuracy. Many techniques for data fusion used in metrology have been proposed, the most common of which are discussed in the following sections.

2 Methodology

Based on the principles presented in Torres-Carrión et al. [29], the research questions and the methodology used for selecting literature sources are explained in this section.

2.1 Research questions

In this review, we identify the research gaps that have not been answered by existing reviews on data fusion in metrology. There are numerous existing reviews on data fusion and their applications, such as [6, 25, 30, 31] applied in general engineering and computer science contexts. In the specific field of coordinate metrology, reviews such as [15] explain how data fusion has been employed to fuse geometric data of engineered parts, particularly in the alignment of multiple measurements (i.e. registration of 3D point clouds). Although these reviews provide a comprehensive and systematic illustration of data fusion solutions, some observations can be drawn:

- Existing reviews present data fusion strictly from a computer science viewpoint, analysing the architecture and theory of general data fusion systems.
- Most of the applications of data fusion presented in existing reviews introduce scenarios in remote sensing, autonomous driving, object detection, and software engineering.
- The artefacts and 3D models used to test algorithms have not previously been reviewed in detail.

With a greater focus into the implications of data fusion for optical coordinate metrology, the following research questions are posed here:

1) What are the advantages of using data fusion for optical coordinate metrology?
2) Compared to other data fusion tasks (e.g. remote sensing, autonomous driving), what are the distinct features of data fusion in optical coordinate metrology?
3) How do existing algorithms differ in terms of the underlying mathematics, and how do these differences influence their performance?
4) As a relatively new technology, can machine learning be employed for data fusion tasks in coordinate measurement and can it perform better than existing algorithms?
5) Examining the methodology in testing an algorithm: are there any common trends in the geometries of artefacts chosen to test them, and what are the common geometric characteristics of the artefacts?

2.2 Selection of studies

This literature review was performed using various scientific databases, including Scopus, Google Scholar, Science Direct, and Research Gate. The searching keywords included the terms “data fusion”, “metrology”, “optical coordinate measurement”, “engineered part”, “point cloud registration”, “machine learning”, and “deep learning”. Publications dated after December 2016 were considered with the highest priority, but publications prior 2017 were also considered if they explained the theoretical foundations of a type of algorithms or introduced a new type of algorithms for the first time. Publications outside of the field of optical coordinate measurement were also quoted, if they demonstrated the characteristics of an algorithm which could be potentially beneficial to applications in optical coordinate measurement.

In this review, there are 112 referenced publications, among which 17 publications contain research on different data fusion algorithms particularly aimed at coordinate measurement, 71 publications are referred to for definitions of concepts and explanations of techniques, and 24 publications are referred to as data fusion algorithms applied to scenarios other than coordinate measurement.

2.3 Contents of the review

In this paper, we review the most common data fusion algorithms applied in the context of optical coordinate metrology, specifically how data fusion methods are employed in the processing of measurements acquired by optical coordinate measuring systems.

As mentioned in Section 1.1, in this review, we have classified data fusion algorithms into three categories, based on their mathematical basis. A new taxonomy is proposed in Section 3.1; then, the theoretical background of each category is discussed in Sections 3.2, 3.3, and 3.4, respectively. In Section 4, we present recent research on the application of each type of algorithms; for each category, one or two recent data fusion algorithms are discussed in detail, including their working pipelines, the processes of the experiments, the geometries of test artefacts, and the experimental results. Section 5 includes a discussion aimed at providing an understanding of the characteristics, advantages, and limitations of each type of data fusion algorithm. Additionally, the artefacts and models used in existing research for testing new algorithms are discussed. The geometrical characteristics of an artefact or model selected or designed for testing of a specific algorithm may affect the performance of the algorithm, especially when the part is symmetric, smooth, or regular.

3 Theoretical background

3.1 A new taxonomy of data fusion algorithms

Since the late 1980s, researchers have proposed many taxonomies classifying techniques for data fusion. One of the earliest and the most frequently quoted classifications is Luo and Kay's "data-feature-decision" three-layer classification [32], in which the authors classified data fusion algorithms into three types according to the level of analysis: data level, feature level, and decision level. This taxonomy was later elaborated into five fusion classes in Dasarathy's five-layer architecture [33]: in this taxonomy, the type of algorithm is classified according to which of the three levels, proposed in [32] the input and output data belong. A similar taxonomy is a Durrant-Whyte architecture [34], which consists of a data pre-processing level, data-refinement levels, and human–computer interactions. These taxonomies do not emphasise the differences in the mathematical basis for each algorithm; instead, the focus is placed on the structure of the data inputted into the fusion system or the connections between datasets or data fusion steps.

In this paper, we propose a new methodology for classifying data fusion algorithms, based on the mathematical principles that underpin existing data fusion algorithms. The mathematical basis of each algorithm classification is as follows:

- GP algorithms: need a GP "governing equation" to describe the calculation process and the geometric features of a surface. GPs are a mathematical tool used to express a stochastic process with Gaussian distribution equations. As such, a mathematical equation or a set of equations must be defined before applying the algorithm.
- WLS algorithms: designed to reduce the noise when applying data fusion by assigning weights to the measurement result, in the form of a matrix. The working principles of WLS algorithms are similar to those for GP algorithms and similarly require a governing equation to start the data fusion process.
- Machine learning algorithms: able to learn patterns in the input datasets, instead of being defined by mathematical rules in advance.

GP and WLS algorithms require pre-defined mathematical expressions before being implemented, while machine learning algorithms detect patterns in data autonomously. Therefore, the major difference between the first two types of algorithms and machine learning algorithms is that the computing processes used in the first two are user-dependent, while machine learning algorithms do not require manual input to initiate the data analysis process [35]. Due to this distinction, in this review, we propose classifying GP and WLS algorithms as "user-dependent algorithms", while machine learning algorithms, together with other artificial intelligence techniques, can be referred to as "user-independent algorithms". Here, we use the definition of machine learning defined by Eastwood et al. ("Machine learning can be thought of as a system which is not specifically programmed to solve a problem; it is instead told what problem to solve, given a set of training data, and then learns how best to solve the given problem on its own.") [35] to decide whether an algorithm should be classified as a machine learning algorithm. According to this definition, we consider any statistical learning algorithms relying on pre-programmed GP and WLS models as GP

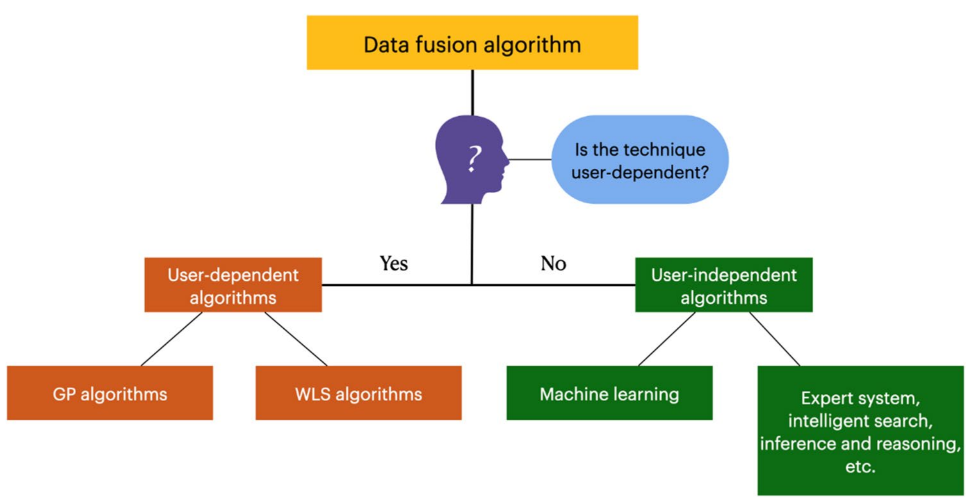

Fig. 1 Taxonomy of data fusion algorithms

and WLS algorithms, instead of machine learning. The proposed taxonomy is shown in Fig. 1.

3.2 Gaussian process algorithms

GP algorithms have been widely explored in the context of three-dimensional (3D) point clouds, i.e. sets of x, y, z positions in a 3D coordinate space [36]. GPs are a mathematical tool used to describe normally distributed stochastic processes that evolve in time according to probabilistic laws [37]. Each GP is a collection of random variables, any finite subset of which obeys a joint Gaussian distribution [38]. A GP is defined with the expression:

$$GP(\boldsymbol{X}) \sim N(\mu(\boldsymbol{X}), K(\boldsymbol{X}, \boldsymbol{X})) \tag{1}$$

where N represents a normal distribution function, and vector $\boldsymbol{X}$ indicates the locations of the data points collected by the sensor, expressed as

$$\boldsymbol{X} = \left[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\right], \tag{2}$$

where $\mu(\boldsymbol{X})$ is the mean function. $K(\boldsymbol{X}, \boldsymbol{X})$ is the covariance matrix, defined as

$$K(\boldsymbol{X}, \boldsymbol{X}) = \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & k(\boldsymbol{x}_1, \boldsymbol{x}_2) & k(\boldsymbol{x}_1, \boldsymbol{x}_3) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_n) \\ k(\boldsymbol{x}_2, \boldsymbol{x}_1) & k(\boldsymbol{x}_2, \boldsymbol{x}_2) & k(\boldsymbol{x}_2, \boldsymbol{x}_3) & \cdots & k(\boldsymbol{x}_2, \boldsymbol{x}_n) \\ k(\boldsymbol{x}_3, \boldsymbol{x}_1) & k(\boldsymbol{x}_3, \boldsymbol{x}_2) & k(\boldsymbol{x}_3, \boldsymbol{x}_3) & \cdots & k(\boldsymbol{x}_3, \boldsymbol{x}_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_n, \boldsymbol{x}_1) & k(\boldsymbol{x}_n, \boldsymbol{x}_2) & k(\boldsymbol{x}_1, \boldsymbol{x}_3) & \cdots & k(\boldsymbol{x}_n, \boldsymbol{x}_n) \end{bmatrix} \tag{3}$$

where $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is the covariance kernel function [38].

3.3 Weighted least-squares algorithms

Algorithms using weighted least-squares (WLS) methods were introduced by Forbes [39], specifically in the context of coordinate metrology. Forbes's work aimed at reducing the noise in data by applying weights to each dataset. WLS fusion is based on a linear measuring system [39], given by

$$z = H\boldsymbol{x} + \varepsilon \tag{4}$$

where $\boldsymbol{x}$ is an n-vector comprised of the model parameters to be measured; H is an $m \times n$ $(m > n)$ matrix of the measured points; z is an m-vector representing the measurement result, and ε is a noise vector independent from the collected data, given by $\varepsilon \sim (0, \sigma^2 \boldsymbol{I})$ [39]. Assume there is a sample set with K samples given by $\{z_k\}_{k \in K}$ with noise level ε, the model parameter vector $\boldsymbol{x}$ can be solved by minimising the weighted squares cost function.

$$\sum\nolimits_{k \in K} w_k \| z_k - H_k \boldsymbol{x} \|^2 \tag{5}$$

where w_k is designed scalar weights [40]. The fusion process is based on solving the model parameter vector $\boldsymbol{x}$ by forcing the partial differential equation of this weighted squares cost function to be zero. Existing research has so far focused on proposing different methods forcing this result. Additionally, researchers have been exploring new methods of designing weights w_k (see Section 4.2).

3.4 Machine learning algorithms

Machine learning has been an active research area in academia since the late 1950s, but has more recently become an industrial focus [35, 41, 42] because available computational resources have significantly increased in the past 2 decades [43]. As such AI, and particularly machine learning, has become an efficient tool for data-intensive research [44], especially in the context of optical coordinate measurement.

Eastwood et al. [35] define machine learning as follows:

> "Machine learning can be thought of as a system which is not specifically programmed to solve a problem; it is instead told what problem to solve, given a set of training data, and then learns how best to solve the given problem on its own."

According to this definition, the computer can be used to predict trends by learning patterns in data with preprogrammed logic. The central idea surrounding machine learning algorithms is to create an autonomous data processing system, unlike GP or WLS methods, where a manually defined mathematical description is required [45]. As such, machine learning algorithms are frequently used to solve problems that are difficult to model with predefined mathematical expressions [45, 46].

There are three categories of machine learning, with each category depending on how the original data is preprocessed: supervised learning, unsupervised learning, and semi-supervised learning [47]. These categories are as follows:

- Supervised learning: if the input datasets have been manually labelled, then supervised learning algorithms are used to learn such datasets. Two types of supervised learning are most frequently used: support vector machines (SVMs) and neural networks (NNs). SVMs are used to realise binary classification. NNs, also known as artificial neural networks (ANNs), learn certain parameters of a dataset by analysing the data with multiple layers of neurons (nodes), each of which has various statistical weights defined by the user [48].
- Unsupervised learning: in the case of unsupervised learning, the input datasets are unlabelled, meaning that the algorithms learn the patterns in the data without direct human input. Unsupervised learning algorithms recog-

nise the hidden patterns in a given dataset by clustering data points. It should be particularly noted that NNs and their variants can also be applied in an unsupervised learning case, particularly when used to detect or extract patterns in data [48–51]. A typical example of NNs used as an unsupervised learning technique is discussed in Section 4.3.

- Semi-supervised learning: semi-supervised learning algorithms process partially labelled datasets [48]. To our knowledge, in the context of optical coordinate measurement, no work has been published using semi-supervised learning for fusing datasets with incomplete labels. As such, a discussion of semi-supervised learning research is not included in this review.

In addition to machine learning, there are other fields under the broader area of AI, some of which are shown in Fig. 2, but the discussion of these areas is beyond the scope of the review.

4 The current state of the art

In the following sections, we present recent research on the application of each algorithm type introduced in Section 2. The advantages and limitations of each algorithm, together with the general problems presented in existing research, are examined.

4.1 GP algorithms

Research on GP algorithms shows a broad range of applications in optical coordinate measurement, separated into two branches: those involving fusion of datasets collected by multiple sensors (this is the most common branch) and those involving enhancement of the measurement processes to improve the quality of datasets.

Ji et al. [52] introduced the "adjustment model" to supplement GP methods. The adjustment model is designed to fuse inhomogeneous data: a low-accuracy dataset (here, a geometric dataset collected by an optical coordinate measurement system) with a high-accuracy dataset (here, a geometric dataset collected by contact coordinate measurement system). Before implementing the adjustment model, the external geometry of the measured surface is firstly predicted by applying GP to the low-accuracy dataset, forming a model of the surface coordinates. Then, the adjustment model is used to describe the difference between the model of the surface coordinates and the high-accuracy dataset. In this process, the high-accuracy dataset acts as the basis for correction.

In their paper, Ji et al. [52] chose two artefacts: an array of spherical holes (see Fig. 3) and a machined freeform surface. For each of these artefacts, a low-accuracy (LA) dataset and a high-accuracy (HA) dataset were collected by a contact and an optical coordinate measurement system. The authors attempted to fuse the two datasets of each of these two artefacts employing their adjustment model. Results showed that the GP method with the adjustment model could fuse the inhomogeneous measurements of the complex surfaces. Additionally, the fusion process only required a small portion of the HA dataset and one LA dataset, which improved the efficiency of the measurement and fusion processes.

In addition to the work presented by Ji et al., numerous GP algorithms have been proposed over the past 5 years. Ma et al. [53] developed a new method called the "fused Gaussian process" (FGP) with a two-component covariance structure. Their algorithm was designed to fuse large spatial datasets (e.g. remote sensing data). Yin et al. [54] introduced a similar GP regression algorithm specifically designed for multi-sensor systems for the measurement of complex surfaces. Experimental tests indicated that the developed

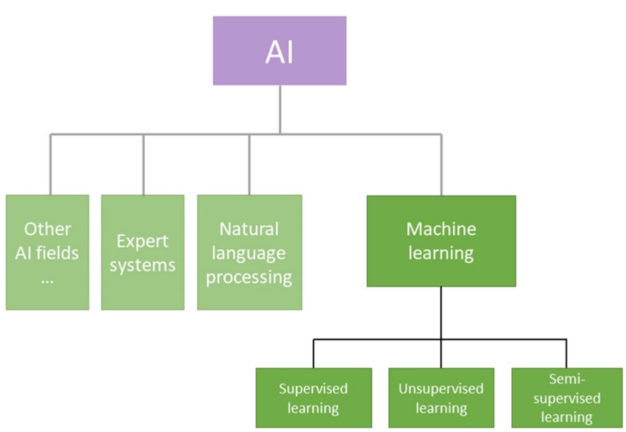

Fig. 2 Machine learning is a type of AI, together with other subjects such as expert systems and natural language processing. There are three categories of machine learning: supervised learning, unsupervised learning, and semi-supervised learning

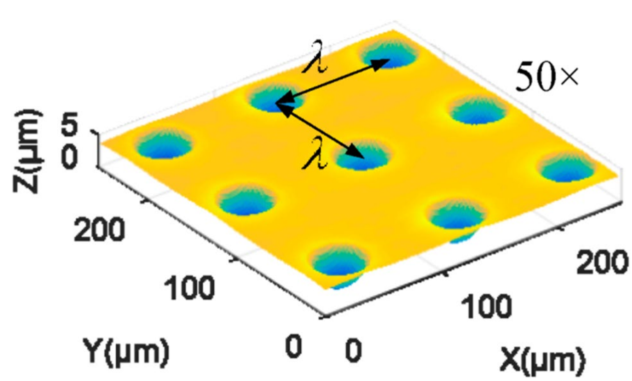

Fig. 3 Artefact comprising an array of spherical holes used in [52]

algorithm was able to perform intelligent sampling (i.e. autonomous selection of effective data points for analysis) when fusing datasets from various sensors. Chen et al. [55] proposed an adaptive sampling method based on GP inference. This algorithm was able to reduce the sampling density of the point cloud collected by a contact coordinate measurement system while having minimal impact on the accuracy of the surface reconstruction derived from multiple datasets.

In addition to GP data processing applications, there are also GP algorithms developed for enhancing the measurement process itself to improve the quality of the datasets. For example, Wang et al. [56] presented an approach for determining the contact positions of a coordinate measurement machine contact probe by implementing the GP model, which led to minimised surface uncertainty (surface variance) in experimental results. Yang et al. [57] introduced a new adaptive sampling technique based on GP inference, which allowed the scanning device to plan the sampling positions intelligently along the scan path (i.e. the sampling positions are chosen according to local geometric characteristics [55]). Their technique aimed at reducing the amount of data whilst saving time without sacrificing the accuracy of the fusion. The results presented in [56, 57] indicated that GP algorithms for planning the measurement process could improve the quality of the collected datasets which were to be fused at a later stage.

The outcomes of GP algorithms, such as those presented in [52–57], indicate that research into this type of algorithm has already reached relative maturity, and the feasibility and reliability of the algorithms have been explored with applications in various scenarios.

4.2 WLS algorithms

Many algorithms based on WLS theory have been proposed in the past few years, aiming at improving the accuracy and efficiency of data fusion. The most common WLS algorithms are those developed by Forbes [39] and Ren et al. [58] (see Fig. 4). The former introduced a general Bayesian approach in order to balance the noise parameters; however, this solution relies on the fitting accuracy of the linear surface model, and its application in multi-sensor fusion is limited [58]. To overcome this issue, the latter added a surface registration method into the general WLS algorithm. The results of both simulations and experiments indicated that their new method could improve the fidelity of the reconstructed surface, modelled using their experimental data and the algorithm.

Research on coordinate measurement has also highlighted the limitations of WLS algorithms in non-linear surface model and multi-sensor data fusion. Yu et al. [59] commented that, although the WLS method is capable of multi-sensor data fusion, showing a noticeable reduction of the measurement uncertainty; these algorithms are still unable to provide comparable accuracy to GP algorithms unless a large number of contact points have been measured, which limits the efficiency of the algorithms. Xiang et al. [60] pointed out that the performance of WLS algorithms is comparable to that of GP algorithms only when processing homogenous datasets, and they are not suitable to fuse datasets collected from large-scale surfaces (e.g. surfaces of major parts on the body of an aircraft or rocket [61]). Kong et al. [62] showed that WLS algorithms depend on linear approximations of the geometry, which means they are not ideal for measuring objects of highly complex geometry (e.g. objects with sharp geometrical changes or smooth surfaces with micro-structures embedded).

4.3 Machine learning algorithms

A general review on machine learning in data fusion was recently presented by Meng et al. [48]. In their review, the

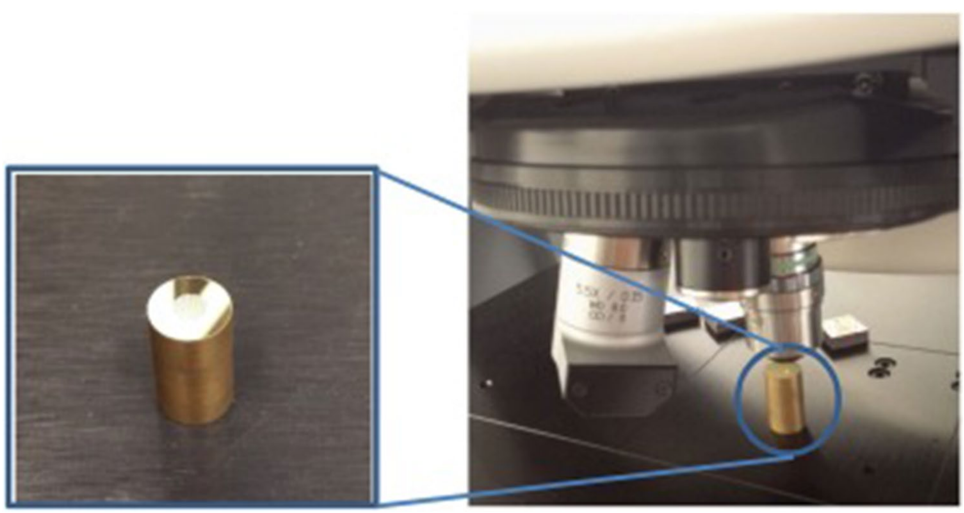

Fig. 4 Measurement of a sinusoidal structured surface using an optical surface topography measurement system. The data collected by the measurement system is used to test the WLS algorithm proposed by Ren et al. [58]

authors noted that most of the existing research is focused on stochastic or time-series data analyses, covering topics such as autonomous vehicles, the internet of things, and geographic information systems. In metrology, the most popular application is the fusion of 3D point clouds.

For a specific object under observation, multiple measurements may be acquired using more than one sensor to form a complete point cloud representing the observed object. Point clouds obtained using different sensors will have different angles of observation, coordinate systems, scales, etc. Abdelazeem et al. [63] noted that there is no apparent information in the measurement dataset indicating how the points in a point cloud are spatially related.

Registration is an important process before further fusion. The term "registration" is defined by Catalucci and Senin [54], based on the definition presented in ISO 10360 part 13, [64] as follows:

> "Registration is the process that brings multiple point clouds taken from observations of the same scene in their correct, relative position within a shared coordinate system."

The process of fusing multiple point clouds is then defined by Abdelazeem et al. [65], based on the definition presented in ISO 10360 part 13 [64], as follows:

> "Data fusion is the process of combining data from multiple sensors in order to obtain better 3D model of an object than that obtained from single sensor data."

An example point cloud that is the result of the fusion of multiple scans is shown in Fig. 5.

To apply machine learning to point cloud registration and fusion, Wang et al. [66] created a registration and fusion pipeline named "deep closest point" (DCP). This method is proposed as an alternative to the iterative closest point (ICP) algorithm, which is one of the most popular solutions for point cloud registration [66–68]. ICP is an iterative process that is employed to minimise the distance between two point clouds. A point cloud (identified as the reference or target) is kept fixed, while the second one (identified as the source or moving set) is transformed to best match the reference based on the rigid motion [63, 66]. The ICP algorithm can fail to reach the global minimum due to its non-convexity, i.e. a non-convex function may have multiple local minima in a certain range; a local minimum found in a certain range does not necessarily correspond to the global minimum [69]. The methods presented in [70–72], which are all variants of ICP supplemented with various statistical optimisation methods, were developed to address this issue. However, in some cases, these methods have not been proven effective.

The DCP model aims to provide a solution to the local minima problem with ICP and consists of three steps [66]:

1) Embed two individual point clouds into a common space and find the corresponding points in two clouds.
2) Create an attention module combined with a pointer generation layer to provide an initial matching, i.e. a probabilistic map from one point cloud to the other.
3) Extract the accurate alignment by analysing the results from step (2) with a differential singular value decomposition (SVD) layer.

Here, the attention module is a machine-learning mechanism that highlights the key data points to increase the accuracy of prediction [73]. A pointer is the core element of a deep-learning technique called "pointer networks", which uses attention as pointers to select input data for combinatorial matching [74]. In this research, pointer generation is a step used to expose matched pairs of points in two point clouds and create an initial matching [66]. The matrix of this matching is used to extract translation and rotation matrices for accurate alignment with the SVD technique [66].

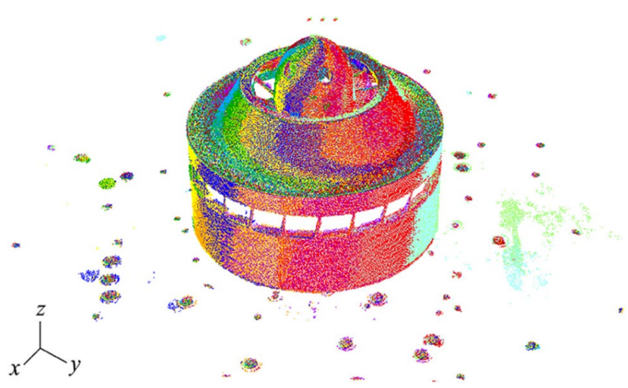

Fig. 5 Multiple point clouds registered in the same coordinate frame. Each individual point cloud is represented with a distinct colour (from [63])

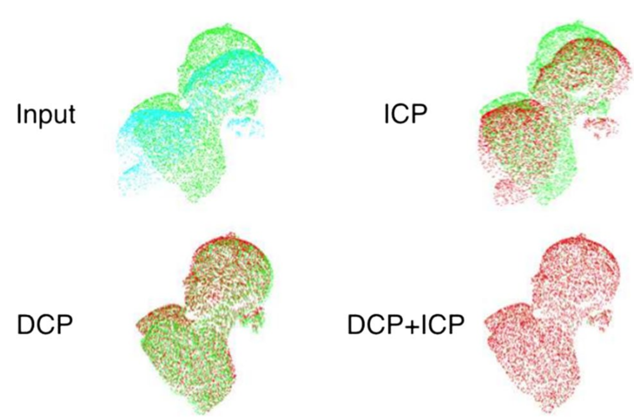

Fig. 6 The fusion of two point clouds by DCP and ICP [66]

Wang et al. [66] compared the outcomes of DCP and ICP algorithms for the registration of multiple point clouds. The results of this comparison are shown in Fig. 6. The comparison shows that ICP assisted by DCP can converge to the global optimum, and DCP can increase the accuracy of alignment when registering two point clouds with poor initial alignment [66]. Additionally, Wang et al. drew the conclusion that DCP is a capable algorithm for rigid registration tasks and can be used to replace ICP algorithms, considering its reduced registration errors.

Later research by Gojcic et al. [75] proposed an NN model, a popular technique in machine learning (see Section 3.4), to fuse two multi-view (i.e. measurement of spatial coordinates through registration and fusion of multiple single-view measurements in different locations and orientations of the optical sensor relative to the workpiece, as defined in ISO 10360 part 13 [64]) datasets into a single point cloud. The algorithm models the registration process using an end-to-end NN, whose accuracy is estimated by a specifically designed layer within the NN. In [75], Gojcic et al. defined the problem of aligning two point clouds as an iterative WLS problem, i.e. the 3D transformation matrices which adjust the orientations of two point clouds by iteratively refining by the NN with WLS method (see Section 3.3 for WLS). To demonstrate the advantages of their NN algorithm, Gojcic et al. tested it with three common benchmark datasets: 3Dmatch [76], Redwood [77], and ScanNet [78]. Compared to non-machine learning algorithms, the NN model was superior in terms of run time, rotational error, and translational error [75].

In addition to the research discussed above, Zhang et al. [79] gave a general introduction to machine learning algorithms for data fusion in optical coordinate measurement. In this review, Zhang et al. categorised all machine learning algorithms within the context of optical coordinate measurement into two types: machine learning as a step added to a traditional measurement pipeline and a complete substitute of traditional pipelines using machine learning technologies [35]. Research into both types has indicated that machine-learning models have the potential to surpass traditional methods in terms of effectiveness [35, 79].

5 Summary and discussion on existing research

5.1 Discussion on the state of the art

In this section, we will discuss the state of the art in user-dependent (i.e. GP and WLS algorithms) and user-independent (i.e. machine learning) algorithms.

In the studies on GP algorithms introduced in Section 4.1, researchers usually tested an algorithm by fusing a "high-accuracy" (HA) dataset with a "low-accuracy" (LA) dataset. The HA data were collected by micro-scale pointwise measuring sensors, such as coordinate measuring machines (CMMs) in the study by Yin et al. [54]. LA data were collected using optical inspection sensors such as fringe projection systems [52]. The data collection efficiency and the point density of an HA dataset are low due to the functionality of most pointwise measuring instruments; on the contrary, data collection for an LA dataset is rapid and its point density is high, but it cannot provide the coordinate information as accurate as an HA data [52, 54]. The benefit of registering an HA dataset with an LA dataset, therefore, is to combine the accuracy of the former and the efficiency of the latter.

Existing studies in GP algorithms usually chose one of the following two paths to fuse an HA with an LA dataset. The first path is to register two datasets by optimising the point-to-point distance in both datasets. In the GP algorithm research discussed in this review, the most typical method used in this path is presented by Ji et al. [52]. This method fuses two datasets by unifying the coordinate systems through optimisation of the point distances. The other path is to propose new sampling methods based on GP for data fusion. In this path, the LA dataset is first subsampled, then GP is applied to the subsampled dataset to reconstruct the surface geometry. Finally, the HA dataset is registered with the reconstructed surface. The key to this research path, therefore, is the proposition of new data sampling methods. The works by Yin et al. [54] and Chen et al. [55] are typical examples of this path: they both proposed intelligent adaptive sampling methods (i.e. sampling size and positions vary according to the local changes in surface geometry [80]) based on GP inference. Chen et al. particularly discussed scenarios where the measured surface has sharp geometrical variations.

Like the existing studies on GP algorithms, typical research on WLS algorithms such as [58] also fused an HA dataset with an LA dataset. However, as we have mentioned in Section 4.2, WLS is not generally useful for fusing complex geometric data because of its reliance on linear modelling [52].

Researchers in machine learning tend to test their algorithmic pipelines by fusing two or more datasets representing the same 3D object. For energy, Wang et al. [66] tested their DCP pipeline by fusing multiple pairs of point clouds, with each pair of point clouds represented a 3D object with approximately the same point densities. The test datasets in [75], as another example, were multiple point clouds representing a building interior. Unlike the datasets used for testing GP and WLS algorithms, these point clouds did not contain surface texture information, instead only representing the general shapes of the object or building interior. As such, existing research has not clarified whether machine learning is effective for registering point clouds containing tiny and dense surface textures to large, sparse point clouds. Moreover, as the benchmark datasets usually had similar densities, the performance of machine learning in registering dissimilar point clouds is also unknown. The fusion of dissimilar point clouds (e.g. small and dense point clouds to large and sparse point clouds) represents an area of future research.

5.2 Advantages and limitations of user-dependent and -independent algorithms

As the most popular type used in optical coordinate metrology, GP algorithms are relatively simple to implement and can be used for flexible nonparametric inference (i.e. inferring the unknown quantities in the data while making as few assumptions as possible) [81–83]. GP-based methods have these advantages because GP is the mathematical basis of many statistical learning algorithms [83]. However, most of the research on GP methods is limited to tests and applications performed on objects with simple geometries. Consequently, whether these algorithms are capable of dealing with multiscale complex surfaces is not yet clear [36, 54, 84–86]. Additionally, the implementation of GP algorithms simplifies the real modelling process into a set of GP equations [81]. In industry, however, the measurement and data fusion processes are more complex than a set of equations can describe and predict (e.g. due to the continuous change of environmental conditions) [87, 88].

As user-dependent algorithms, both GP and WLS algorithms rely on user-defined mathematical expressions to process the external geometry and surface topography data. Essentially, when applying these algorithms, it is assumed that the engineered part's geometry can be described with a set of GP or WLS equations. However, the measured part can have highly complex geometric features—more complex than those that GP and WLS algorithms are able to model. This limitation is particularly notable in the implementation of WLS algorithms, as presented in Section 4.2 by [59, 60, 62]. The surface geometries can be far more complex than the mathematical equations can model, which is the general limitation for user-dependent algorithms, as [87, 88] indicate.

As user-independent algorithms, machine learning algorithms define a model that allows learning of the patterns in the data autonomously after being trained with the training datasets, instead of using pre-defined mathematical expressions. Consequently, machine learning solutions potentially provide more flexibility than GP and WLS algorithms, particularly when measurement tasks cannot be modelled with specific equations.

Recent research has shown the potential of machine learning algorithms. The Elman ANN algorithm has been shown to be capable of nonlinear predictions in practical applications, for instance, in the determination of the position of an object in 3D space [89]. Fahmy et al. [90] indicate that machine learning-enhanced techniques can overcome data imperfection better than GP and WLS algorithms. Tong et al. [91] demonstrated that machine learning showed robustness against noise compared to GP algorithms. Shu et al. [92] demonstrated that a machine-learning model was robust in fusing datasets that had ambiguity and noise. A later work by Wang et al. [93] demonstrated the efficiency and stability of machine-learning algorithms with an object-tracking task. Alyannezhadi et al. [94] proposed a clustering algorithm to fuse datasets whose characteristics could not be identified by mathematical equations.

While the machine learning methods discussed were not all applied within the context of metrology, researchers have demonstrated that machine learning is capable of complex tasks and is robust against noise and data imperfections, which are important advantages for applications in metrology. One of the directions of future research is to explore machine learning models that can fuse dense point clouds. As Wang et al. [66] discuss, the latest machine learning models are only successful with point clouds of 500–5000 points; ideal machine learning models should be able to process up to 300,000 points.

Regarding the precision of these algorithms and evaluation of their contributions to measurement uncertainty, the published research has not included much discussion, instead only comparing new machine learning algorithm performance with other earlier algorithms, such as ICP and fast global registration (FGR) [95]. In optical coordinate metrology, comparison with non-machine learning algorithms using the same benchmark datasets is also rarely seen in the literature. Because of these issues, it is difficult to comment on methods for the evaluation of uncertainty and precision of existing machine learning algorithms, particularly when compared with non-machine learning algorithms.

To decrease errors, existing research on new machine learning methods commonly proposes iterative refinement or "Multiple Run" [66, 75, 96]. These experimental outcomes indicate that multiple runs could improve accuracy after initial registration. To our knowledge, researchers have not proposed more techniques particularly aimed at improving the uncertainty and precision of machine learning. As such, the examination of methods for evaluating the contribution to measurement uncertainty from these algorithms is a ripe area for future research.

The advantages and limitations of all types of algorithms are summarised in Table 1 using the taxonomy we proposed in Section 3.1. In this summary, user-dependent algorithms rely on user-defined mathematical models to learn the geometrical patterns in the input data and then fuse the input data based on the user-defined models. This feature makes

Table 1 Summary of the characteristics of all algorithms

Types	Method	Advantages	Limitations	Literature sources
User-dependent	GP	Simple to implement and can be used for flexible nonparametric inference [84, 85, 97]	Replying on user-defined mathematical models [97]; ineffective for fusing data of complex geometries[56, 86–88]	Ji et al. [52], Ma et al. [53], Yin et al. [54], Chen et al. [55], Wang et al. [56], Yang et al. [57]
	WLS	Simple to implement [39]	Low efficiency [59]; incapable of large-scale data fusion [60]; not ideal for fusing data of highly complex surface [62]	Forbes [39], Ren et al. [58]
User-independent	Machine learning	Does not require user-defined mathematical models, as it learns the patterns in the input data autonomously [98]; able to fuse data that cannot be described with user-defined mathematical models [94]; robust against noise in data fusion [92]; high effectiveness [79]	Applications particularly in external coordinate measurement are rare; it has not been tested with highly dense point clouds [66]	Wang et al. [66], Gojcic et al. [75], Kolanowski et al. [89], Fahmy et al. [90], Tong et al. [91], Shu et al. [92], Wang et al. [93], Alyannezhadi et al. [94]

user-dependent algorithms easy to deploy but also difficult to fuse data of complex geometries, because the surface geometries can be more complex than those that GP and WLS are able to describe, as [87, 88] indicate.

User-independent algorithms, such as machine learning, can recognise geometrical patterns without user-defined mathematical models. These algorithms use techniques, such as NNs, to detect the patterns autonomously. This feature makes user-independent algorithms effective in fusing data of complex geometries that are difficult to analytically model. However, as existing research is still relatively rare, more experimental studies are in need to prove the effectiveness and reliability of machine learning used in optical coordinate measurement, as stated in Section 5.1.

5.3 Discussion on methodology: test objects

As mentioned in Section 2.1, the geometries of the test artefact have rarely been discussed in existing reviews. In the literature collection process, we have noted that there are certain trends in the selections and designs of the test artefacts in existing research. In the research presented in Sections 4.1, 4.2, and 4.3, the artefacts or virtual surfaces that were used to test the algorithms usually had simple geometries. These artefacts and virtual surfaces generally exhibited the following characteristics: the virtual surfaces were defined using periodic mathematical patterns, i.e. the definitions of the surface used in each paper are all in form of $z = \sin(Ax) + \cos(By), A, B \in \mathbb{R}$. Similarly, artefacts used were either periodically patterned or had simple curvature. For example, artefacts with highly symmetrical geometries were frequently used for testing the newly proposed algorithms (see examples in Fig. 3, Fig. 5, and 7).

As discussed in Section 5.1, the geometries of engineered parts are likely to be more complex than the geometries of simulated surfaces and custom-designed artefacts in laboratories. Therefore, in addition to simulated surfaces and artefacts, common objects and engineered parts used in the industry should be employed in industrial and research activities to test different measurement techniques. For instance, coins represent inexpensive examples of metal freeform surfaces that can be used for the detection of defects in surface topography measurement [99].

Common objects used in practical circumstances, such as a coin, can challenge an algorithm more than a simulated surface or a specifically manufactured artefact. Taking a coin as an example, the patterns on it, e.g. motto, legend, and mint mark, display many convex and concave geometries, which are characterised by a wide dimensional range (see Fig. 8). As such, a coin can provide various opportunities to challenge the capability of a coordinate measurement technique, including data fusion algorithms used in this process [100]. Additionally, coins with different materials and worn surfaces lead to different optical reflectivity conditions, which can influence the data collected by optical sensors [101]. Therefore, a coin is also an effective object to test the robustness and stability of a data fusion algorithm in coordinate measurement processes.

In terms of geometric complexity, a simulated surface may also display highly complex geometry if its mathematical expressions have numerous items, as Eastwood et al. [102] and Todhunter et al. [103] indicated in their work. If a 3D model is constructed with such polynomials and inputted into an additive manufacturing system, a highly complex artefact can be manufactured for testing data fusion algorithms. However, the more complex the surface, the higher the computational cost.

Another popular type of object used for metrological research is aspherical lenses [104, 105] (see Fig. 9). Aspherical lenses have become increasingly common in industry because of the progress of manufacturing technology; therefore, they are frequently used in medical imaging, optical systems, astrophysics, lithography, automotive, and

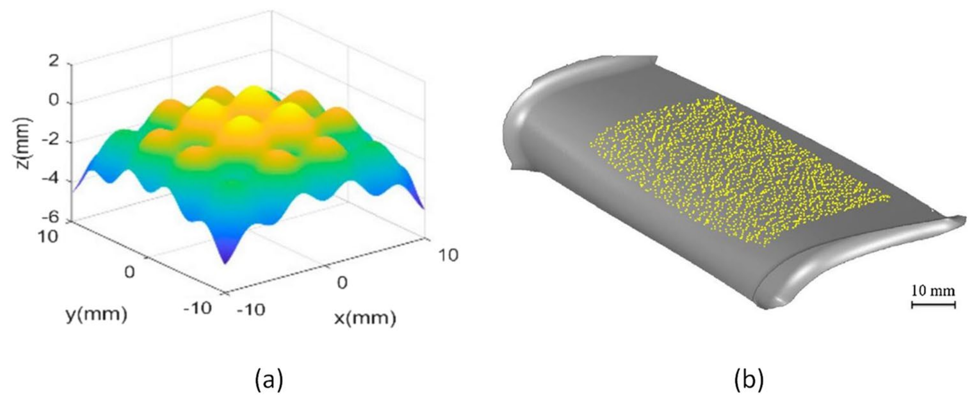

Fig. 7 Test examples of recently developed algorithms: **a** shows a simulated surface generated by a mathematical equation in form of $z = \sin(Ax) + \cos(By), A, B \in \mathbb{R}$ (from [38]), and **b** is an artefact with relatively smooth geometry, with the yellow dots representing the data points collected by the measurement instrument (from [40])

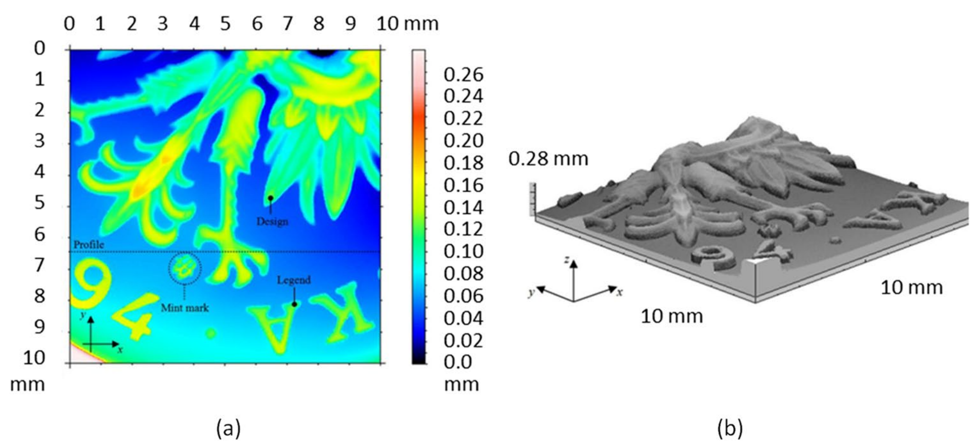

Fig. 8 An example of the surface profile of a part of a coin (1 zł, Poland) adapted from [100], showing complex geometric features: **a** the colour map of height, in which the authors mark geometric features including design, legend, and mint mark. The dot line marked with “Profile” is where the authors of [100] used for further research in their work on surface proliferation. **b** 3D model of the same area on the coin, rendered with data shown in **a**

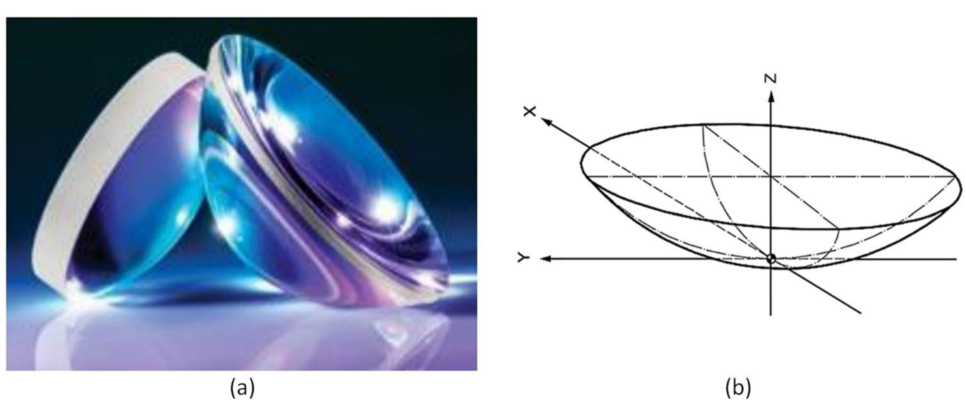

Fig. 9 **a** A photo of two example aspherical lenses and **b** schema of an asphere [104, 110]

metrology [106]. The broad range of applications in high-precision optical fields demands better techniques for coordinate measurement in manufacturing, designing, and testing aspherical lenses [107, 108]. As such, coordinate measurement of an aspherical lens challenges the accuracy of data fusion algorithms used in this process [109].

6 Conclusions

In this review, we have discussed a broad array of data fusion algorithms used for optical coordinate measurement proposed since 2017 and have defined a new taxonomy for the classification of existing algorithms based on their mathematical backgrounds: user-dependent algorithms (GP methods and WLS methods) and user-independent algorithms (machine learning algorithms).

By reviewing research on each class of algorithm, we have seen that user-dependent algorithms are relatively easy to implement, as these algorithms assume that the geometry of a surface can be described by a series of mathematical equations. However, these algorithms are not capable of modelling highly complex surfaces. Recent research on user-independent algorithms represented by machine learning and its derivatives such as deep learning is relatively scarce, but the results have already demonstrated its potential and value for further exploration.

In addition to the algorithm review, the problems with experimental methods used in existing research are also discussed. Researchers generally use a virtual surface plus an engineered object to test the algorithms. The problem with virtual surfaces is that they are usually created using simple mathematical expressions, consisting of a sine and a cosine term. The problem with engineered artefacts is that there is commonly a lack of complex geometric features on their surfaces, e.g. sharp changes in height.

7 Future work

As introduced and discussed in Sections 2 and 4, research on GP algorithms has reached maturity with many outcomes in the past 2 decades in a broad range of application scenarios. Research on WLS algorithms is relatively rare, because researchers have realised their limitations and disadvantages compared with GP algorithms. Compared with user-dependent algorithms, research on machine learning for data fusion applications specifically within the context of metrology is not frequently seen. As such, potential directions for further exploration on machine learning algorithms are:

1) Machine learning algorithms for the fusion of datasets measured from parts with highly complex geometries should be developed. Additionally, existing algorithms should be tested with highly complex artefacts or industrial objects, e.g. coins, aspherical lenses, and aspherical mirrors.
2) As Wang et al. [66] discussed, research should focus on the applications of machine learning algorithms to the fusion of highly dense point clouds (i.e. clouds having more than 5000 points) and inhomogeneous datasets, i.e. datasets with the difference in point quantity, point density, features, and particularly, dimensions.
3) Research should investigate how to apply the advantages of machine learning, which have been shown in other application scenarios such as autonomous vehicles and object recognition to optical coordinate measurement in manufacturing. A direction that is worthy of investigation is locating a dense point cloud representing a sub-section of an object within a larger but less dense point cloud covering the whole of an object (e.g. a small field-of-view surface texture measurement within a wider co-ordinate measurement). The latest algorithms, such as PointNet [111] and its variances (e.g. PointNet++ [112]), are able to recognise a small point cloud within a larger point cloud. However, this architecture can only determine what the object is and cannot provide the precise location of the recognised object in a larger point cloud. Moreover, PointNet and its variances can recognise an object in a scenario only when the objects in the scenario have been manually labelled, i.e. they cannot recognise a pattern in a point cloud scenario if this pattern does not have any label. As such, future research should explore how to let a program autonomously recognise an object or pattern within a larger point cloud without labelling any object.

Acknowledgements The authors would like to thank the UKRI Research England Development (RED) Fund for funding this work via the Midlands Centre for Data-Driven Metrology.

Author contribution Resources: Zhongyi Michael Zhang, Sofia Catalucci, Adam Thompson.

Supervision: Richard K Leach, Samanta Piano.

Writing—original draft: Zhongyi Michael Zhang.

Writing—review and editing: Zhongyi Michael Zhang, Sofia Catalucci, Adam Thompson, Richard K Leach, Samanta Piano.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

1. Leach RK (2014) Fundamental principles of engineering nanometrology. Elsevier, Amsterdam
2. Leach R K 2020 Terms, definitions and standards In: Leach R K Advances in Optical Form and Coordinate Metrology (IOP Publishing), Chap. 1
3. Saha S, Foryś P, Martusewicz J, Sitnik R 2020 Approach to analysis the surface geometry change in cultural heritage objects Proc. ICISP 2020, Marrakesh, Morocco, Jun. 3–13
4. Carmignato S, Voltan A, Savio E (2010) Metrological performance of optical coordinate measuring machines under industrial conditions *CIRP Ann. - Manuf*. Technol 59:497–500
5. Shimizu Y, Chen LC, Kim DW, Chen X, Li X, Matsukuma H (2021) An insight into optical metrology in manufacturing *Meas*. Sci Technol 32:042003
6. Castanedo F (2013) A review of data fusion techniques *Sci*. World J 2013:704504
7. Leach R K 2020 *Advances in optical surface texture metrology* (IOP Publishing)
8. Diez-Olivan A, Del Ser J, Galar D, Sierra B (2019) Data fusion and machine learning for industrial prognosis: trends and perspectives towards Industry 4.0 *Inf*. Fusion 50:92–111
9. Liu J, Li T, Xie P, Du S, Teng F, Yang X (2020) Urban big data fusion based on deep learning: an overview *Inf*. Fusion 53:123–133
10. Dautov R, Distefano S, Buyya R (2019) Hierarchical data fusion for Smart Healthcare. J Big Data 6:19
11. Zahra S R 2021 Securing the defense data for making better decisions using data fusion In: T J Saleem, M A Chishti Big Data Analytics for Internet of Things (Wiley), Chap. 14
12. Fourati H (2017) Multisensor data fusion: from algorithms and architectural design to applications. CRC Press, Boca Raton
13. Llinas J, Hall D L 1998 Introduction to multi-sensor data fusion Proc. - IEEE Int. Symp. Circuits Syst., Monterey, USA, Feb. 537–40
14. Pomerleau F, Colas F, Siegwart R (2015) A review of point cloud registration algorithms for mobile robotics *Found*. Trends Robot 4:1–104
15. Wang J, Leach RK, Jiang X (2015) Review of the mathematical foundations of data fusion techniques in surface metrology *Surf*. Topogr Metrol Prop 3:023001
16. Kong L, Peng X, Chen Y, Wang P, Xu M (2020) Multi-sensor measurement and data fusion technology for manufacturing process monitoring: a literature review *Int*. J Extrem Manuf 2:022001
17. Leach R K Advances in optical form and coordinate metrology (IOP Publishing)
18. D'Errico GE (2012) À la Kalman filtering for metrology tool with application to coordinate measuring machines *IEEE Trans*. Ind Electron 59:4377–4382
19. Amamra A, Aouf N, Stuart D, Richardson M (2016) A recursive robust filtering approach for 3D registration *Signal*. Image Video Process 10:835–842
20. Sandhu R, Dambreville S, Tannenbaum A (2010) Point set registration via particle filtering and stochastic dynamics *IEEE Trans*. Pattern Anal Mach Intell 32:1459–1473
21. Chen Z, Li Q, Li J, Zhang D, Yu J, Yin Y, Lv S, Liang A (2022) IMU-aided registration of MLS point clouds using inertial trajectory error model and least squares optimization. Remote Sens 14:1365
22. White F E 1987 *Data fusion lexicon* (San Diego: the data fusion subpanel of the joint directors of laboratories, technical panel for C3)
23. White F E 1991 *Data fusion lexicon* (San Diego: the data fusion subpanel of the joint directors of laboratories, technical panel for C3)
24. Zhang Z, Blum R S 1999 A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application *Proc. IEEE, Aug.* 1315–26
25. Esteban J, Starr A, Willetts R, Hannah P, Bryanston-Cross P (2005) A review of data fusion models and architectures: towards engineering guidelines *Neural Comput*. Appl 14:273–281
26. Starr A, Desforges M 1998 Strategies in data fusion - sorting through the tool box *Proc. EuroFusion98, Great Malvern, UK* 85–92
27. Weckenmann A, Jiang X, Sommer KD, Neuschaefer-Rube U, Seewig J, Shaw L, Estler T (2009) Multisensor data fusion in dimensional metrology *CIRP Ann*. Manuf Technol 58:701–721
28. Xu B J, Willomitzer F, Yeh C K, Li F, Gupta V, Tumblin J, Walton M, Cossairt O 2019 3D surface measurement and analysis of works of art *Conf. Rec. Asilomar Conf. Signals. Syst. Comput., Pacific Grove, USA, Nov.* 1779–82
29. Torres-Carrion P V, Gonzalez-Gonzalez C S, Aciar S, Rodriguez-Morales G 2018 Methodology for systematic literature review applied to engineering and education *IEEE Global EDUCON, Santa Cruz de Tenerife, Spain, Apr.* 1364–1373
30. Dong J, Zhuang D, Huang Y, Fu J (2009) Advances in multi-sensor data fusion: algorithms and applications. Sensors 9:7771–7784
31. Chen G, Liu Z, Yu G, Liang J (2021) A new view of multisensor data fusion: research on generalized fusion *Math*. Probl Eng 2021:1–21
32. Luo R C, Kay M G 1989 Multisensor integration and fusion in intelligent machines and systems IEEE Trans. Syst. Man Cybern. 901–31
33. Dasarathy BV (1997) Sensor fusion potential exploitation-innovative architectures and illustrative applications *Proc*. IEEE 85:24–38
34. Luo RC, Yih CC, Su KL (2002) Multisensor fusion and integration: approaches, applications, and future research directions *IEES*. Sens J 2:107–19
35. Eastwood J, Sims-Waterhouse D, Piano S 2020 *Machine learning approaches advances in optical form and coordinate metrology* In: Leach R K *Advances in Optical Form and Coordinate Metrology* (IOP Publishing), Chap. 6
36. Colosimo BM, Pacella M, Senin N (2015) Multisensor data fusion via Gaussian process models for dimensional and geometric verification *Precis*. Eng 40:199–213
37. Hida T, Hitsuda M (1993) Gaussian processes. R.I., American Mathematical Society, Providence
38. Ren MJ, Cheung CF, Xiao GB (2018) Gaussian process based bayesian inference system for intelligent surface measurement. Sensors (Switzerland) 18:4069
39. Forbes AB (2012) Weighting observations from multi-sensor coordinate measuring systems *Meas*. Sci Technol 23:025004

40. Wang J, Pagani L, Leach RK, Zeng W, Colosimo BM, Zhou L (2017) Study of weighted fusion methods for the measurement of surface geometry *Precis*. Eng 47:111–121
41. Nilsson N J 1965 Learning machines: foundations of trainable pattern-classifying systems (McGraw-Hill Companies)
42. Samuel AL (1959) Some studies in machine learning using the game of checkers *IBM*. J Res Dev 3:210–229
43. Alippi C, Ferrero A, Piuri V (1998) Artificial intelligence for instrument & measurement applications *IEEE Instrum*. Meas Mag 1:9–17
44. Halevy A, Norvig P, Pereira F (2009) The unreasonable effectiveness of data *IEEE Intell*. Syst 24:9–12
45. Liu S, Zhang L, Yan Z (2018) Predict pairwise trust based on machine learning in online social networks: a survey *IEEE*. Access 6:51297–51318
46. Wei L, Luo W, Weng J, Zhong Y, Zhang X, Yan Z (2017) Machine learning-based malicious application detection of android *IEEE*. Access 5:25591–25601
47. Jing W, Kang J, Liu M (2018) Mining taxi trajectories for most suitable stations of sharing bikes to ease traffic congestion *IET Intell*. Transp Syst 12:586–593
48. Meng T, Jing X, Yan Z, Pedrycz W (2020) A survey on machine learning for data fusion *Inf*. Fusion 57:115–129
49. Lin K C, Lin C H, Lin V C 2009 A planar multiband antenna with parasitic-element design for multistandard mobile terminals *IEEE AP-S Internat. Symp. (Digest)* 1–4
50. Julisch K (2003) Clustering intrusion detection alarms to support root cause analysis *ACM Trans*. Inf Syst Secur 6:443–471
51. Völker C, Shokouhi P 2015 Data aggregation for improved honeycomb detection in concrete using machine learning–based algorithms *Internat. Symp. NDT-CE, Berlin, Sept.*
52. Ji D, Liu Q, Bai M, Sun P (2020) A multisensor data fusion method based on gaussian process model for precision measurement of complex surfaces. Sensors (Switzerland) 20:278–293
53. Ma P, Kang EL (2020) A fused Gaussian process model for very large spatial data. J Comput Graph Stat 29:479–489
54. Yin Y, Ren MJ, Sun L (2017) Dependant Gaussian processes regression for intelligent sampling of freeform and structured surfaces *CIRP Ann. - Manuf*. Technol 66:511–514
55. Chen Y, Peng C (2017) Intelligent adaptive sampling guided by Gaussian process inference *Meas*. Sci Technol 28:105005
56. Wang X, Qian X 2018 Gaussian process model for touch probing ASME 2018 13th Internat. MSEC, College Station, Texas, Jun. MSEC2018–6548, V002T07A003
57. Yang C, Peng C, Chen Y, Luo T, Chu J (2018) Space-filling scan paths and Gaussian process-aided adaptive sampling for efficient surface measurements *Precis*. Eng 54:412–419
58. Ren MJ, Sun LJ, Liu MY, Cheung CF, Yin YH, Cao YL (2017) A weighted least square based data fusion method for precision measurement of freeform surfaces *Precis*. Eng 48:144–151
59. Yu Z, Wang T, Wang P, Tian Y, Li H (2019) Rapid and precise reverse engineering of complex geometry through multi-sensor data fusion *IEEE*. Access 7:165793–165813
60. Xiang B, Li Y, Chen G, Liu X, Yang W (2020) Multi-source integrated fusion for surface measurement *Int*. J Adv Manuf Technol 109:1815–1823
61. Zhou G, Li Y, Liu C, Hao X (2018) A posture adjustment optimization method of the laser inspection device for large complex surface parts. Proc Inst Mech Eng Part B J Eng Manuf 13:2375–2385
62. Kong LB, Ren MJ, Xu M (2017) Development of data registration and fusion methods for measurement of ultra-precision freeform surfaces. Sensors 17:01110
63. Catalucci S, Senin N 2020 State-of-the-art in point cloud analysis In: Leach R K Advances in Optic*al Form and Coordinate Metrology* (IOP Publishing), Chap. 2
64. ISO 10360 part 3 2021 Geometrical product specifications (GPS) — acceptance and reverification tests for coordinate measuring systems (CMS) — Part 13: Optical 3D CMS (International Organization for Standardization)
65. Abdelazeem M, Elamin A, Afifi A, El-Rabbany A 2021 Multisensor point cloud data fusion for precise 3D mapping Egypt. J. Remote Sens. Sp. Sci. 835–44
66. Wang Y, Solomon J 2019 Deep closest point: learning representations for point cloud registration *Proc. IEEE Int. Conf. Comput. Vis., Soeul, South Korea, Oct.-Nov.* 3523–3532
67. Besl PJ, McKay ND (1992) A method for registration of 3-D shapes *IEEE Trans*. Pattern Anal Mach Intell 14:239–256
68. Segal A V, Haehnel D, Thrun S 2009 Generalized ICP Proc. Robot.: Sci. Syst., Seattle, US, Jun.-Jul. 21
69. Boyd S, Vandenberghe L 2004 Convex functions In: Boyd S, Vandenberghe L Convex Optimization (New York: Cambridge University Press), Chap. 1
70. Fitzgibbon A W 2003 Robust registration of 2D and 3D point sets Image Vis. Comput. 1145–53
71. Rusinkiewicz S, Levoy M 2001 Efficient variants of the ICP algorithm Proc. Internat. Conf. 3DIM, 3DIM, Quebec City, Canada, May-Jun., 145–152
72. Yan J, Yin X C, Lin W, Deng C, Zha H, Yang X 2016 A short survey of recent advances in graph matching Proc. ACM ICMR, New York, US, May-Jun., 167–74
73. Zhang H, Zhang Q, Shao S, Niu T, Yang X (2020) Attention-based LSTM network for rotatory machine remaining useful life prediction *IEEE*. Access 8:132188–132199
74. Vinyals O, Fortunato M, Jaitly N 2015 Pointer networks Proc. NIPS., Montreal, Canada, Jan. 2692–2700
75. Gojcic Z, Zhou C, Wegner J D, Guibas L J, Birdal T 2020 Learning multiview 3D point cloud registration Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Seattle, Jun. 1756–1766
76. Zeng A, Song S, Nießner M, Fisher M, Xiao J, Funkhouser T 2017 3DMatch: learning local geometric descriptors from RGB-D reconstructions Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Honolulu, Jul. 199–208
77. Choi S, Zhou Q Y, Koltun V 2015 Robust reconstruction of indoor scenes Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern. Recognit., Boston, US, Jun. 97–104
78. Dai A, Chang A X, Savva M, Halber M, Funkhouser T, Nießner M 2017 ScanNet: richly-annotated 3D reconstructions of indoor scenes Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Honolulu, Jul. 5828–5839
79. Zhang Z, Dai Y, Sun J (2020) Deep learning based point cloud registration: an overview. Virtual Real Intell Hardw 2:222–246
80. Wang J, Jiang X, Blunt LA, Leach RK, Scott PJ (2012) Intelligent sampling for the measurement of structured surfaces *Meas*. Sci Technol 23:085006
81. Rasmussen CE, Williams CKI (2008) Gaussian processes for machine learning. MIT Press, Cambridge
82. Park C, Huang JZ, Ding Y (2011) Domain decomposition approach for fast Gaussian process regression of large spatial data sets. J Mach Learn Res 12:1697–1728
83. Corder G W, Foreman D I 2014 *Nonparametric statistics: an introduction* In: Corder G W, Foreman D I *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach* (Hoboken, New Jersey: John Wiley & Sons, Inc), Chap. 1
84. Song X, Jiang X, Gao J, Cai Z (2019) Gaussian process graph-based discriminant analysis for hyperspectral images classification. Remote Sens 11:2288
85. Dumas A, Echard B, Gayton N, Rochat O, Dantan JY, Van Der Veen S (2013) AK-ILS: an active learning method based on Kriging for the inspection of large surfaces *Precis*. Eng 37:1–9

86. Richardson RR, Osborne MA, Howey DA (2017) Gaussian process regression for forecasting battery state of health. J Power Sources 357:209–219
87. Lázaro-Gredilla M, Titsias MK, Verrelst J, Camps-Valls G (2014) Retrieval of biophysical parameters with heteroscedastic Gaussian processes *IEEE Geosci*. Remote Sens Lett 11:838–842
88. Ghaffari Jadidi M, Valls Miro J, Dissanayake G (2018) Gaussian processes autonomous mapping and exploration for range-sensing mobile robots. Auton Robots 42:273–290
89. Kolanowski K, Świetlicka A, Kapela R, Pochmara J, Rybarczyk A (2018) Multisensor data fusion using Elman neural networks. Appl Math Comput 319:236–244
90. Fahmy M S, Atiya A F, Elfouly R S 2008 Biometric fusion using enhanced SVM classification Proc. 4th IIH-MSP 2008, Harbin, China, Aug. 1043–1048
91. Tong W G, Li B S, Jin X Z, Yang Y Q, Zhang Q 2006 A study on model of multisensor information fusion and its application Proc. 2006 ICMLC, Dalian, China, Aug. 3073–3077
92. Shu H, Wang Y, Jiang J 2007 Multi-rada data fusion algorithm based on K-central clustering Proc. FSKD, Haikou, China, Aug. 4406311
93. Wang H, Liu T, Bu Q, Yang B 2016 An algorithm based on hierarchical clustering for multi-target tracking of multi-sensor data fusion Chinese Control Conference, Chengdu, China, Jul. 5106–5111
94. Alyannezhadi MM, Pouyan AA, Abolghasemi V (2017) An efficient algorithm for multisensory data fusion under uncertainty condition. J Electr Syst Inf Technol 4:269–278
95. Zhou Q Y, Park J, Koltun V 2016 Fast global registration Lecture Notes *Proc. ECCV, Amsterdam, The Netherlands, Oct.* 9906 LNCS 766–82
96. Agamennoni G, Fontana S, Siegwart R Y, Sorrenti D G 2016 Point clouds registration with probabilistic data association *IEEE Int. Conf. Intell. Robots Syst., Daejeon, Korea (South) Nov*. 4092–4098
97. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press, Cambridge
98. Ongsulee P 2018 Artificial intelligence, machine learning and deep learning *Int. Conf. ICT Knowl. Eng., Bangkok, Thailand, Nov.* 17
99. Ekberg P, Su R, Leach R (2017) High-precision lateral distortion measurement and correction in coherence scanning interferometry using an arbitrary surface *Opt*. Express 25:18703–18712
100. Kapłonek W, Sutowska M, Ungureanu M, Çetinkaya K (2018) Optical profilometer with confocal chromatic sensor for high-accuracy 3D measurements of the uncirculated and circulated coins. J Mech Energy Eng 2:181–192
101. Parra Escamilla GA, Kobayashi F, Otani Y (2017) Three-dimensional surface measurement based on the projected defocused pattern technique using imaging fiber optics *Opt*. Commun 390:57–60
102. Eastwood J, Newton L, Leach R, Piano S (2022) Generation and categorisation of surface texture data using a modified progressively growing adversarial network. Precis Eng 74:1–11
103. Todhunter L, Senin N, Leach R, Lawes S, Blateyron F, Harris P 2018 A programmable software framework for the generation of simulated surface topography Conf. Proc. 18th euspen, Venice, Italy, Jun. 138400
104. Arezki Y, Zhang X, Mehdi-Souzani C, Anwer N, Nouira H (2018) Investigation of minimum zone assessment methods for aspheric shapes *Precis*. Eng 52:300–307
105. Arezki Y, Nouira H, Anwer N, Mehdi-Souzani C 2018 A novel hybrid trust region minimax fitting algorithm for accurate dimensional metrology of aspherical shapes Meas. J. Int. Meas. Confed. 134–140
106. Karow H H 2004 Fabrication Methods for Precision Optics (Hoboken, NJ: Wiley-Interscience)
107. Wang Z, Qu W, Yang F, Tian A, Asundi A 2017 Absolute measurement of aspheric lens with electrically tunable lens in digital holography Opt. Lasers Eng. 313–318
108. Shao G, Hai R, Sun C (2020) 3D printing customized optical lens in minutes *Adv*. Opt Mater 4:1901646
109. Adams D, Ament S (2018) Understanding aspheric lenses: key specifications and their impact on performance *Opt*. Photonik 13:60–63
110. ISO 10110 Part 12 2007 Optics and photonics – preparation of drawings for optical elements and systems
111. Qi C R, Su H, Mo K, Guibas L J 2017 PointNet: deep learning on point sets for 3D classification and segmentation Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Honolulu, Jul. 652–660
112. Qi C R, Yi L, Su H, Guibas L J 2017 PointNet++: deep hierarchical feature learning on point sets in a metric space Adv. Neural Inf. Process. Syst. 5105–5114