Mathematical Biosciences and Engineering

MBE, 20(8): 14260–14280.
DOI: 10.3934/mbe.2023638
Received: 13 April 2023
Revised: 04 June 2023
Accepted: 11 June 2023
Published: 28 June 2023

http://www.aimspress.com/journal/MBE

Research article

Multiview intelligent networking based on the genetic evolution algorithm for precise 3D measurements

Yujing Qiao[1,*], Ning Lv[1] and Baoming Jia[2]

[1] School of Mechanical Engineering, Yangzhou Polytechnic College, Yangzhou 225009, China
[2] Intelligent Machine Institute, Harbin University of Science and Technology, Harbin 150080, China

* **Correspondence:** Email: qiaoyujing@sina.com.

Abstract: The use of multi-visual network 3D measurements is increasing; however, finding ways to apply low-cost industrial cameras to achieve intelligent networking and efficient measurement is a key problem that has not been fully solved. In this paper, the multivision stereo vision 3D measurement principle and multivision networking process constraints are analyzed in depth, and an intelligent networking method based on the genetic evolution algorithm (GEA) is proposed. The genetic operation is improved, and the fitness function is dynamically calibrated. Based on the visual sphere model, the best observation distance is assigned as the radius of the visual sphere, and the required constraints are fused to establish an intelligent networking design of the centering multivision. A simulation and experiment show that the proposed algorithm is widely feasible, and its measurement accuracy meets the requirements of the industrial field. Our multiview intelligent networking algorithms and methods provide solid theoretical and technical support for low-cost and efficient on-site 3D measurements of industrial structures.

Keywords: 3D measurement; multiview intelligent networking; genetic evolution algorithm; visual sphere model; constraint condition

1. Introduction

3D vision measurement technology uses the type of image information obtained by a visual sensor as the carrier. The geometric information of the object to be measured in the space is calculated, and then the target object is reconstructed, and its 3D morphology information is restored [1–4]. High-

precision 3D scanners have been widely used for large and complex structures based on the need for fast and noncontact measurements. However, the measurements can only be performed offline and require a long time. To enable online panoramic 3D measurements of small industrial structures, a high-speed 3D profilometer has been developed. The ultrafast industrial speed can reach 64 kHz, and materials with different reflective intensities are simultaneously measured with high speed, precision and stability. However, the measurements are expensive and have a limited application range. It is difficult to measure large complex structures in three dimensions [5]. Although visual inspection systems provide excellent accuracy within their field of view, they are not reliable in verifying the spacing between two entities located in different fields of view. In addition, a visual inspection system requires the installation of multiple cameras on fixtures dedicated to an application, project or component. This lack of flexibility is a major disadvantage, because if certain parameters of the tested component change, the position of the camera may become askew, and the fixtures may need to be replaced. Therefore, a fixed, rigid, permanent, expensive system suitable for small variations and mass production is well suited to the emerging needs of intelligent manufacturing. To date, there is no perfect solution for achieving flexible online 3D inspection of industrial components.

Multivision networking measurement technology is an intelligent networked measurement method that has been developed in recent years. Some of the problems described above were solved, and it is one of the important development directions for modern industrial measurement [6–8]. The method can be used to accurately measure the 3D topography of industrial parts with multiple sizes at a low cost. Using this method, the visual sensors are placed at different measuring positions, and the object to be measured is photographed from different angles. Multiple viewpoint positions are required for the formation of a camera imaging network. In the actual measurement environment, the surface topography curvature of the part is not singular, which makes the best measurement position and azimuth angle of each visual sensor different. Based on the above issues, a visual measurement network planning problem arises. However, the determination of the position points in the current visual measurement network is mainly restricted by the measurement accuracy [9–11], and this is somewhat incomplete. Therefore, it is necessary to extend stereo vision measurement technology to the topography measurement engineering practice through intelligent algorithms.

Scholars have conducted a substantial amount of research on camera pose optimization and camera distribution methods for multivision 3D measurement technology. In the initial research on multivision measurement, optimizations were mainly based on full coverage regarding imaging with certain constraints [12,13]. For example, in reference [14], the visibility constraints and coverage are considered, and the measurement cameras are classified and screened. Then, the penalty function is used to continuously optimize the screening, and the optimal visual measurement network is obtained. However, this measurement network does not take into account the influence of the shielding area and observation accuracy, so it does not have complete universality. In [15], in the process of multivisual networking, only the attitude of each node is considered, and the constraints between the nodes are not considered. Therefore, the obtained network has serious redundancies. Chen et al. proposed a photogrammetric network planning method based on a genetic algorithm (GA), which in turn was based on the high measurement accuracy requirements of a large trough concentrator [16]. This method used the 3D reconstruction uncertainty of the measurement target point as the network planning evaluation standard and combined the constraints of the actual measurement environment to conduct a simulation analysis to obtain the optimized simulation measurement network. The simulation results are verified by 3D measurement experiments.

The various visual measurement methods described above attempt to predict the visibility of the target points and the feasibility of the camera station to optimize the initial imaging network without using CAD models [17]. There are known disadvantages; all of these technologies need a good initial network expert (i.e., prior knowledge) [18]. Some network design strategies are suitable for coordination and are used for specific object features or texture changes of the 3D reconstructions from multiple views. Their method considers the difference in the image texture in the overlapping area, and by adjusting the feature vectors of the 3D points perpendicular to the camera view direction, the feature values are minimized to ensure the accuracy of the 3D reconstructions. However, the constraints related to the measurement range are not considered, and the self-occlusion area is not applicable.

There are many constraints within the measurement range, such as the image size and resolution constraints, number and distribution of the feature points, camera field of view, depth of field, the imaging and size overlapping area of the target working area [19–24]. These range constraints are related to another set of visibility constraints, including the incidence angle, occlusion and self-occlusion area. The design of a multiview visual network is characterized by the use of multiple parameters, multiple constraints, a large amount of computation and so on. These characteristics determine that the search for an absolute exact optimal solution requires a very large amount of computation. Theoretically, this is a complex optimization strategy problem, which is difficult to describe theoretically and requires a highly technical implementation. Therefore, the efficiency of the algorithm must be considered in the implementation. If the above factors are considered in the planning of a visual measurement network, and intelligent algorithms extend stereo vision measurement technology to the topography measurement engineering practice. Then, in the industrial vision measurement field, the accuracy and reliability of the measurements will make a qualitative leap.

The feasibility of a GA solution in the search space has broad possibilities, and the population search process has significant advantages. Moreover, the algorithm can search for the global optimal solution in parallel, and it can be selected and applied in combination with other intelligent algorithms for different engineering problems. In this paper, the multiview stereovision 3D measurement principle and the multivision networking constraints are analyzed, and the parameters to be optimized are obtained. To seek an accurate position and pose of the camera in the vision measurement network, it is proposed that the genetic evolutionary algorithm (GEA) is integrated into the network design. In this algorithm, the genetic operation step is improved, and the fitness function is dynamically calibrated. Then, the centering networking method is designed for the workpiece structure, resulting in easy centering. The spherical model is adopted, and the optimal observation distance is taken as the spherical radius. The required constraint conditions are considered, and the optimal centering camera network is iteratively selected by GEA to implement precise 3D measurements.

2. The 3D measurement principle based on multiview stereo vision

To achieve full measurement coverage, numerous measurement modules must be arranged in space to form a multiview stereo vision measurement network. Based on the binocular vision 3D measurement principle, the multivision 3D measurement model is further derived. The 3D measurement principle is shown in Figure 1.

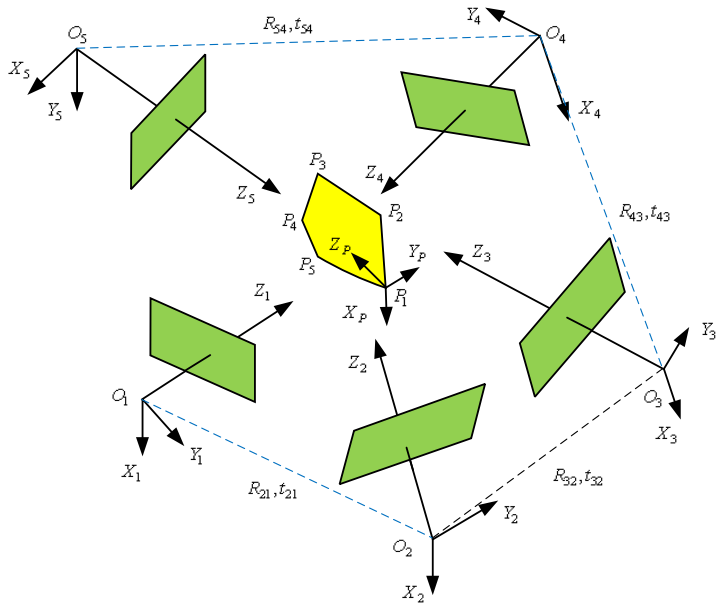


Figure 1. The 3D measurement principle in multiview stereoscopic networking.

The cameras of each measuring viewpoint are set as $O_1, O_2, \cdots, O_n$. For the convenience of calculation, the camera coordinate system, $(O_{c1} - X_{c1}Y_{c1}Z_{c1})$, of Camera 1 overlaps with the world coordinate system, $(O_{w1} - X_{w1}Y_{w1}Z_{w1})$ analogously with the 3D measurement principle of binocular stereovision. The characteristic points on the object to be measured are $P_i(i = 1,2,3\ldots)$. The coordinate of P_i in the coordinate system of Camera 1 is (x_{i1}, y_{i1}, z_{i1}), and the coordinate system of P_i in the coordinate system of Camera 2 is (x_{i2}, y_{i2}, z_{i2}). The rotation matrix between the coordinate systems of Cameras 1 and 2 is R_{21}, and the translation matrix is t_{21}. Therefore, the transformation formula for the coordinate system of Cameras 1 and 2 is shown in Eq (1).

$$\begin{bmatrix} x_{i1} \\ y_{i1} \\ z_{i1} \end{bmatrix} = R_{21} \begin{bmatrix} x_{i2} \\ y_{i2} \\ z_{i2} \end{bmatrix} + t_{21} \tag{1}$$

By analogy, the coordinates of P_i in the coordinate system of Camera n-1 are $(x_{i(n-1)}, y_{i(n-1)}, z_{i(n-1)})$, and the coordinates of P_i in the coordinate system of Camera n are (x_{in}, y_{in}, z_{in}). The transformation relation between the coordinate systems of Cameras n-1 and n is shown in Eq (2).

$$\begin{bmatrix} x_{i(n-1)} \\ y_{i(n-1)} \\ z_{i(n-1)} \end{bmatrix} = R_{21} \begin{bmatrix} x_{in} \\ y_{in} \\ z_{in} \end{bmatrix} + t_{n(n-1)} \tag{2}$$

The coordinate transformation matrix of the camera is obtained from Eqs (1) and (2), as shown in Eq (3).

$$K_{n(n-1)} = \begin{bmatrix} R_{n(n-1)} & t_{n(n-1)} \\ 0 & 1 \end{bmatrix} \tag{3}$$

According to the above derivation, any camera coordinate system in the measurement network of a multiview stereo vision can be expressed as $O_{cj} - X_{cj}Y_{cj}Z_{cj}(2 \le j \le n)$, and the spatial 3D coordinates of its characteristic point P_i are shown in Eq (4).

$$\begin{bmatrix} X_{Wi} \\ Y_{Wi} \\ Z_{Wi} \\ 1 \end{bmatrix} = K_{j1} \begin{bmatrix} x_{ij} \\ y_{ij} \\ z_{ij} \\ 1 \end{bmatrix} \tag{4}$$

The 3D information of any object to be measured in space can be handled through integrating the binocular and multiview stereo vision 3D measurement principle.

3. Multiview intelligent networking method based on GEA

Visual networking has many constraints, a large amount of computation, complex parameters and requires interdisciplinary knowledge. Based on these characteristics, the stability and operational efficiency of algorithm technology must be considered in networking design implementation, however, it the method of finding the global optimal solution is difficult to use in traditional mathematical modeling. This paper improves on GA to solve the visual networking design problem.

3.1. Implementation process of GEA

According to Darwinian evolution theory, a genetic algorithm (GA) can automatically find a solution model for the theoretical global optimal solution [25]. Genetic operations include three methods: selection, crossover and mutation. GEA is a new calculation model that evolved based on the GA mechanism. In addition to having all the characteristics of GA, GEA has different improvement methods for different genetic operation problems, and can also conduct individual optimizations for multiple objectives [26]. The specific implementation process is as follows:

Step 1. Input the total number of individuals *N*, the maximum number *G* of iterations, the cross probability $C \in [0.4, 0.99]$, the mutation probability $M \in 0, 0.05$ and the intensity factor α. The parameters of fitness function threshold *Fit* and the discrete precision *epsDiscrete* are set according to the measurement accuracy requirements. The *N* algorithm individuals x_i (i = 1, 2, …, *N*) are randomly initialized.

Step 2. Since each function to be solved has an interval constraint, the operation process uses binary and is coded according to the iterative precision. N coded individuals are uniformly and randomly generated, and the fitness function is calculated for each individual.

Step 3. Implement genetic manipulation, selection, crossover and mutation.

Step 4. Record the initial generated optimal individual x_{op} and its fitness function $f(x_{op})$.

Step 5. When the number of iterations $k < G$, consider whether a large mutation is needed. The mutation operation in the genetic operation is used to avoid the "precocity" of the algorithm. However, since, in the operation, the mutation probability is usually small, it is necessary to iterate a large

progeny group before the mutation of new individuals can occur, which greatly increases the amount of computation.

Define:

$$k < G/2\&\left(f_{ave} - f_{min}()(fmin_{max}\&r < M_{max})\right) \tag{5}$$

where f_{ave} is the average fitness value, f_{min} is the minimum fitness value and f_{max} is the maximum fitness value. If Eq (5) is satisfied, there is no need for large variations. If Eq (5) is not satisfied, all the individuals, except for the best ones, need to conduct large genetic mutation operations.

Step 6. The fitness function after large mutations is dynamically calibrated. In the algorithm operation process, the fitness function of each individual may be slightly different, resulting in the selection operation weakening or even disappearing, which is not conducive to the stability of the algorithm. In this paper, the fitness function is calibrated dynamically to ensure the relative difference of the individual fitness function.

The formula $f(x_i) = f(x_i) - fk_{min}$ is used to calculate the individuals of the optimal fitness function and replace the worst individual, where Q is the initial adjustment value, $q \in [0.9, 0.999]$ represents the constringency coefficient, and $\mathrm{f_{ave}}$, f_{min} and f_{max} are calculated again.

Step 7. Equation (6) is used to calculate the selection probability and cumulative probability:

$$p_i = f(x_i)/\sum_{i=1}^{N} f(x_i),\ P_i = \sum_{j=1}^{i} f\left(x_j\right) \tag{6}$$

Step 8. *N* offspring are produced. When $P_{i-1} < r < P_i$, for each offspring *i*; based on the idea of "survival of the fittest" in GEA, the roulette algorithm is used to select the paternal line in the GA, and then the maternal line is randomly selected.

The calculation of adaptive genetic crossover probability is shown in Eq (7):

$$prC = \begin{cases} \dfrac{f_{\max} - f}{f_{\max} - f_{ave}} C_{\min}, & f \geq f_{ave} \\ C, & f < f_{ave} \end{cases},\ f = max\left(f_{father}, f_{mother}\right) \tag{7}$$

If $r < prC$, a cross operation is performed. The probability of adaptive genetic mutation is calculated as shown in Eq (8):

$$prC = \begin{cases} \dfrac{f_{\max} - f}{f_{\max} - f_{ave}} C_{\min}, & f \geq f_{ave} \\ C, & f < f_{ave} \end{cases} \tag{8}$$

If $r < prM$, perform the mutation operation. All the individuals, except the optimal one, are replaced by the next generation.

Step 9. Update the iterative optimal individual x_{opt}, and record the iterative number k of offspring. When $f_{ave}/f\left(x_{opt}\right) \geq thFitness$, the fitness function exceeds the threshold, the algorithm converges and the global optimal solution is obtained. At the end of the iteration, the optimal visual networking method is obtained by decoding.

3.2. Multivision measurement networking method

For the structural parts with easy centering, take the center of the target to be measured as the center of the sphere to build the visual sphere model, and combine the GEA and visual constraints. The theoretical global optimal camera networking method is obtained, and in this paper, it is called peering-type networking.

3.2.1. Visual sphere model and parameters to be optimized

First, the target structure to be measured is centered (refer to the literature [27]); the visual sphere model is established based on this. In peering-type networking, the distance between the camera and the object to be measured is fixed in the measurement process, and the optical axis of the camera is toward the center of the sphere, so the camera networking is on the surface of the spherical model with a fixed radius, as shown in Figure 2. In the visual sphere model, each viewpoint represents a camera station point, and the angle on the camera projection plane is called the azimuth angle α. The included angle on the vertical plane is called the elevation angle β, where $\alpha \in [0,2\pi)$, $\beta \in [-\pi/2,\pi/2]$. D is the equivalent focal length of the camera. In the visual sphere model, the camera is fixed on a spherical surface with a radius of D, and then the spatial position and orientation coordinates of camera i can be expressed as (α_i, β_i, D). The external parameter vector of the camera can also be further estimated from the above information, where the rotation vector $R = [\alpha \quad \pi - \beta \quad \pi/2]^T$ and the translation vector $t = [0 \quad 0 \quad D]^T$. Assuming m cameras are given, the parameters to be optimized for the multiview measurement network are shown in Eq (9):

$$x = [x_1, x_2, \cdots, x_m]^T, \quad x_j = \left[\alpha_j, \beta_j\right]^T \tag{9}$$

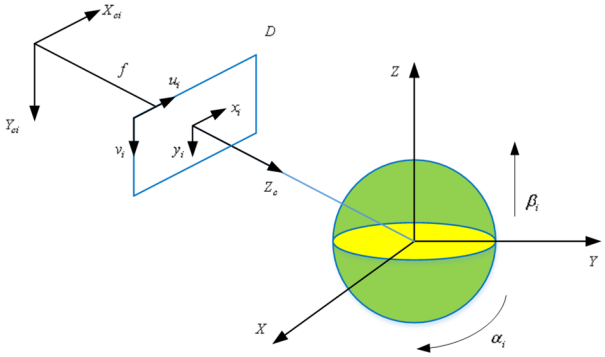


Figure 2. Visual sphere model.

A reasonable camera spatial position is determined by $I = [x, y, z, \alpha, \beta, D]$, which has 6 parameters to be optimized. In addition, the constant parameters that affect the camera's position mainly include the focal length f and the baseline B. Theoretically, the spatial position coordinates (x, y) of the camera remain unchanged. Increasing z can increase the number of measurable points of the camera and improve the networking coverage. However, increasing z continuously will reduce the spatial resolution of the camera, resulting in the surface details of the target to be measured being ignored, and reducing the networking accuracy. Therefore, it is crucial to reasonably select the optimal

distance of the camera. According to the conclusion of Literature [28] in Section 3 (this paper contains our team's the research work), the change in the baseline distance between the two cameras will directly lead to a change in the camera's spatial pose coordinates (α, β, D), thus affecting the measurement accuracy of the image depth. The depth information measurement error and baseline variation curve can be obtained from Eq (10) [28], as shown in Figure 3.

$$\Delta z = \frac{\sqrt{2}z}{f}\left(\frac{z}{B} + \frac{B}{4z}\right) \tag{10}$$

where f is the focal length and B is the baseline.

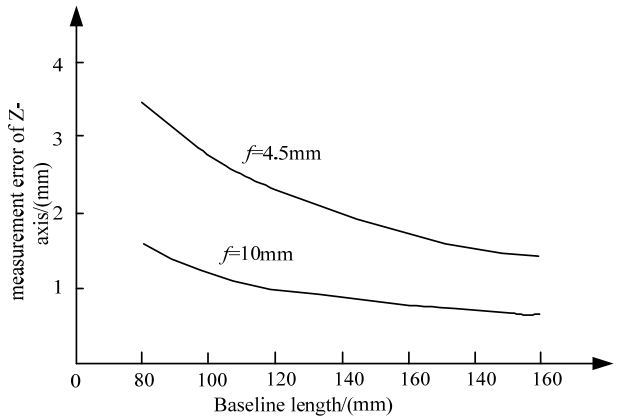


Figure 3. Variation curve of depth information measurement error and baseline.

According to the above analysis, in terms of system accuracy, any two vision sensors in the measurement network can be combined into a binocular stereo vision system, and the baseline distance B can be calculated through the Cartesian coordinates of the viewpoint, so increasing the baseline distance B of the binocular vision in a certain range can effectively reduce the measurement error of the system. As shown in Figure 2, the changes in the azimuth angle and elevation angle are the key parameters to control the direction of the camera's optical axis, which directly determines the camera's spatial position and orientation coordinates. It is necessary to analyze the relationship between the two factors and the accuracy of the measurement network. In the actual binocular measurement process, to meet the requirements of the measurement task, the two azimuth angles, α_1 and α_2, may be different, so it is necessary to analyze the influence of the two azimuth angle changes on the measurement error. Based on the mechanism analysis results of the impact of the projection angle on the measurement accuracy in Literature [28], taking $[0, \pi/2]$ as an example, the change curve is shown in Figure 4, and the other three quadrants can be obtained based on the mutual complement. Δ represents the measurement error (unit: mm). Figure 4 shows that with an increasing α_1 and α_2, the measurement error Δ also increases.

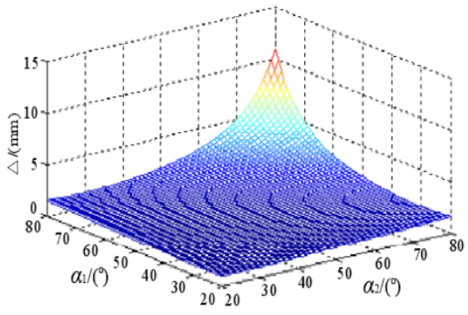


Figure 4. Variation curve of the elevation angle.

In the same way, it can be deduced that the change curve of the relationship between the Elevation Angles β_1 and β_2 and the measurement error is shown in Figure 5. With the increase of the Elevation Angles β_1 and β_2, the measurement error Δ first decreases, tends to be stable, and then increases. Figure 5 only shows the change curve between $\beta \in [0, \pi/2]$. The curve trend of $\beta \in [-\pi/2, 0]$ and $\beta \in [0, \pi/2]$ is the same, so the interval is $[-70°, -10°]$. In binocular stereo vision, the Azimuth Angle and Elevation Angle simultaneously restrict the accuracy of the measurement network.

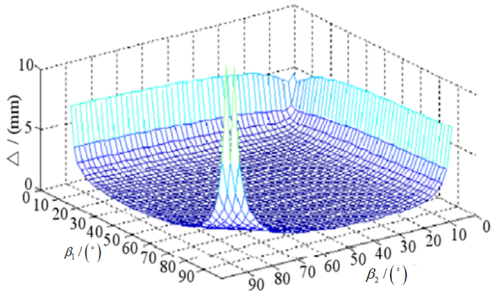


Figure 5. Variation curve of the elevation angle.

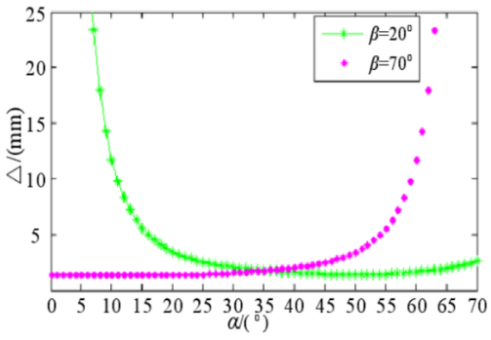


Figure 6. The system accuracy varies with the angle between the optical axis and the baseline.

To select a reasonable range of azimuth angles, based on the range of the interval, two complementary azimuth angles α_1 and α_2 are arranged, and a two-dimensional change curve is drawn, as shown in Figure 6. At this time, it is easy to observe the change in measurement error in the

azimuth range. As shown in Figure 6, the measurement error Δ is evenly distributed within the interval $[25°, 45°]$ of α with a small error value. Therefore, the optimal interval of measurement error Δ within $\alpha \in [0,2\pi)$ is $[25°, 45°]$, $[115°, 135°]$, $[205°, 225°]$ and $[295°, 315°]$.

3.2.2. Target model to be measured and constraint conditions

The visual measurement process is affected by various kinds of constraints. The major influencing factors include the multiview measurement principle, the structural parameters of the camera itself, the measurement environment, etc. To establish a multiview measurement network, it is necessary to analyze and screen out the constraints of factors with significant influence weights from the complex constraints and establish a mathematical model. This process will directly affect the efficiency of the algorithm in the process of network design.

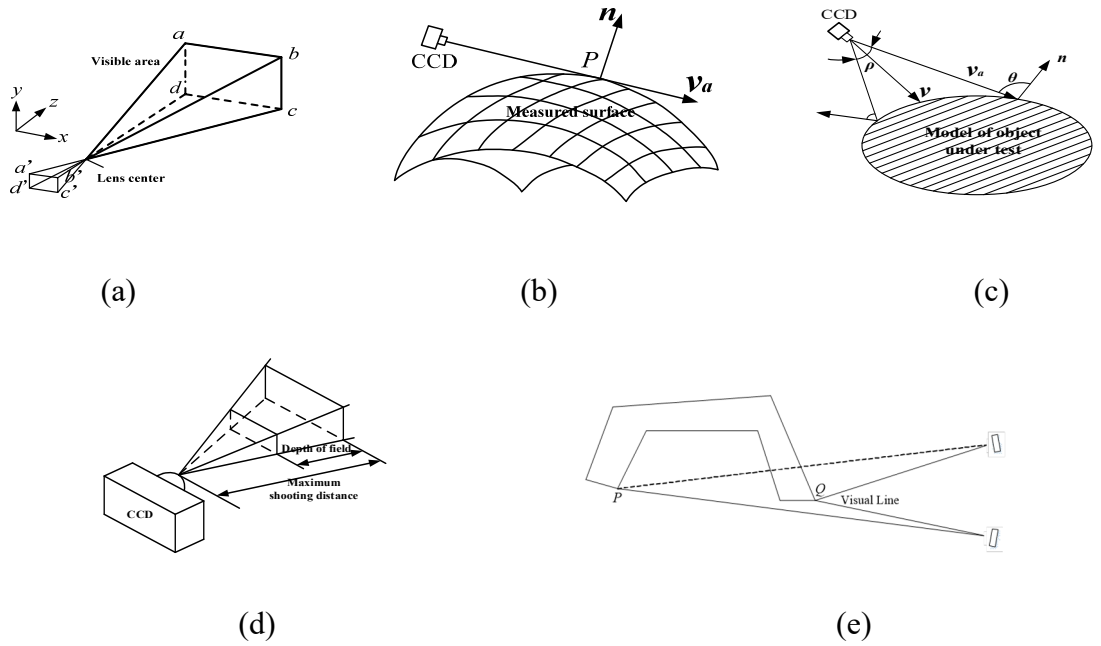


Figure 7 The constraints in the multiview networking process. (a) Field angle constraint; (b) Visibility constraint; (c) Constraints on incident angle constraint; (d) Depth of field constraint; (e) Coview constraint.

The main constraints are summarized as the following: the field angle constraint, visibility constraint, incidence angle constraint, depth of field constraint, common view constraint, curvature constraint and empty set constraint.

1) Field angle constraint. The field angle is defined as a measure of the imaging range in a camera system. It refers to the angle between two rays formed from the camera lens and along the boundary of the maximum object that can be contained by the lens. The larger the field of view angle, the larger the camera's field of view. If the angle between the boundary of the target to be measured and the lens exceeds the field of view angle, the part beyond the field of view angle will not be imaged in the camera, as shown in Figure 7(a). The visual area within the pyramid region represents the visual area, and the mathematical formula to constrain the pyramid region is the field of view angle constraint, as

shown in Eq (11).

$$g_1: \|\vec{v}\| \cdot \|\vec{v}_a\| \cos(\rho/2) - \vec{v}\vec{v}_a \leq 0 \tag{11}$$

where $\vec{v}$ is the unit direction vector of the camera's optical axis, $\vec{v}_a$ is the optical axis of the camera and ρ represents the field angle of the camera.

2) Visibility constraint. As shown in Figure 7(b), $\vec{n}$ is the normal vector at measurement target point *P*. If the angle between the normal vector $\vec{n}$ and the direction vector of the sensor viewpoint is equal to 90°, then the part on the right cannot be collected. The mathematical model of the visibility constraint is shown in Eq (12).

$$g_2: \vec{n} \cdot \vec{v}_a \leq 0 \tag{12}$$

3) Constraints on incident angle constraint. As shown in Figure 7(c), in the actual visual measurement process, to reduce the pixel error value of the image points after two-dimensional imaging, it is not expected that the position of some viewpoints is coplanar with the surface of the measured object. θ_{max} is the maximum acceptable angle between the camera viewpoint direction vector $\vec{v}_a$ and the target point normal vector $\vec{n}$, which is defined as the incident angle. The constraint conditions of the incident angle are shown in Eq (13).

$$g_3: \pi - \cos^{-1}\frac{\vec{n}\vec{v}_a}{\|\vec{n}\|\|\vec{v}_a\|} - \theta \leq 0 \tag{13}$$

4) Depth of field constraint. A schematic diagram of the depth of field is shown in Figure 7 (d), where ΔL_1 represents the foreground depth, and ΔL_2 represents the back depth of the field. The object to be measured has a distance between the front and back depth of field, and the imaging is clear. Therefore, camera imaging should be constrained within this area, which is called a depth of field constraint, and the constraint formula is shown in Eq (14).

$$g_4: \begin{cases} \frac{F\delta L(L-f)}{f^2+F\delta(L-f)} - \Delta L_1 \leq 0 \\ \Delta L_2 - \frac{F\delta L(L-f)}{f^2-F\delta(L-f)} \leq 0 \end{cases} \tag{14}$$

where *F* is the camera aperture, δ is the diameter of the allowable dispersion circle, *L* is the shooting distance and f is the focal length of the camera.

5) Coview constraint. In the process of multiview visual measurement, any point *P* in space is observed by at least two cameras at the same time to attain high observation accuracy, as shown in Figure 7(e). Limited by the structure of the target under test, there is self-occlusion or mutual occlusion, and the *P* point can only be observed by one camera, so the spatial 3D coordinates of the *P* point cannot be obtained. Therefore, it is necessary to constrain each area to meet the multicamera coview condition, which is called a coview constraint, as shown in Eq (15):

$$g_5: \Omega_d = \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_n \tag{15}$$

6) Curvature constraint. According to the analysis in Section 3.2.1, the measurement accuracy of the visual measurement network is closely related to the camera spatial pose (α, β, *D*). The higher the surface complexity of the object structure to be measured, the greater the curvature is, and the more information it contains, which requires higher camera pose requirements. Therefore, the constraints

need to be carried out based on the actual situation, and are called curvature constraints. The mapping relationship is shown in Eq (16).

$$g_6: (\rho_i) \rightarrow (\alpha, \beta, D) \tag{16}$$

where ρ_i is the curvature of the object structure corresponding to the camera at the *i*-th station point.

7) Null set constraint. To ensure the independence of each view in the multiview measurement network and avoid the view overlap increasing the calculation amount, it is also necessary to limit the overlap part Λ of each view, which is called a null set constraint. Equation (16) is as follows:

$$g_7: \Phi = \Lambda_1 \cap \Lambda_2 \cap \cdots \cap \Lambda_n \tag{17}$$

3.2.3. Fitness function

In the GEA, the fitness function is used to represent the superiority and inferiority of each individual to determine the chance that it can be inherited. In the visual measurement networking design process, it represents the measurement networking accuracy. For the point cloud of the target to be measured, under the above constraints, it is reasonable to measure when it will be visible by two or more cameras. The parameters involved in visual network design are multidimensional, and the measurement accuracy of the point clouds are effectively estimated by using the covariance.

The line of sight intersection equation under reasonable measurement can be expressed in the nonhomogeneous form $Ax = b$, and the object point coordinates obtained from the intersection conditions are expressed in the normal equation $A^T AX = A^T b$, which can be converted into $X = (A^T A)^{-1} A^T b$, and Eq (18) can be obtained:

$$X = f(x) \tag{18}$$

where *f* represents the mapping relationship, and *x* represents the parameters related to *A* and *b. Then*, Eq (19) is obtained as follows.

$$x = [x_1, x_2, \cdots, x_m]^T \tag{19}$$

The covariance matrix of the image point is shown in Eq (20).

$$\Sigma_x = diag(\Sigma_{x1}, \Sigma_{x2}, \cdots, \Sigma_{xm}), \ \Sigma_{xj} = diag\left(\sigma_{xj}^2, \sigma_{yj}^2\right) \tag{20}$$

According to the *Monte Carlo* simulation method [26], if the covariance of an image point in the point cloud cannot be obtained, that is, let Σ_{xj} be the unit matrix, the Gaussian noise covariance of this point's coordinate is "1 pixel". The covariance matrix of object points obtained from the above analysis is shown in Eq (21).

$$\Sigma_x = J\Sigma_x J^T = \sum_{j=1}^{m} J_j \Sigma_{xj} J_j^T, \ J_j = \left[\frac{\partial f}{\partial x_j}, \frac{\partial f}{\partial y_j}\right] \tag{21}$$

where $J' = \left[{J'_1}^T, {J'_2}^T \cdots, {J'_m}^T\right]^T$, J'_j is the first derivative matrix of the image point on the *j*-th camera to the measurement marker point, which is 2 rows and 3 columns, and the covariance matrix of the image point is shown in Eq (22).

$$\Sigma_x = diag(\Sigma_{x1}, \Sigma_{x2}, \cdots, \Sigma_{xm}),\ \ \Sigma_{xj} = diag\left(\sigma_{xj}^2, \sigma_{yj}^2\right) \tag{22}$$

If a mark point on the object structure to be measured is visible, its covariance matrix can be calculated by Eq (21), and the three diagonal elements of the covariance matrix are represented by $\sigma_{xi}, \sigma_{yi}, \sigma_{zi}$. Then, the measurement accuracy of the object point can be obtained as the maximum error of the three coordinates in the space, as shown in Eq (23).

$$\sigma_i = max\{\sigma_{xi}, \sigma_{yi}, \sigma_{zi}\} \tag{23}$$

If the object point is not visible, the measurement accuracy is defined as $\sigma_i = 10\kappa$, and κ is the worst spatial resolution of all the measuring cameras. The measurement marker points on the target structure to be measured are classified into visible points and invisible points, which are numbered n_1 and n_2, respectively. Then, the improved individual fitness function table is shown in Eq (24):

$$fit = -\max_{i=1,\cdots,n_1}\{\sigma_i\} - 10\kappa n_2 \tag{24}$$

According to genetic knowledge, individuals with poor fitness will be gradually eliminated. By analogy, in the measurement networking design, the larger the number of invisible measurement marks n_2, the worse the fitness will be, and this kind of networking design method will eventually be eliminated.

3.2.4. Implementation of GEA in intelligent networking

An input operation is required before the GEA is implemented that includes the structure model of the target to be tested, the selection of constraints, the number and type of cameras, the value range of each parameter in the visual sphere model of the camera, the fitness function threshold and the measurement accuracy threshold.

The target structure to be measured is discretized into a uniform point cloud, the coordinates of the center point $[x_0, y_0, z_0]^T$ and the unit normal vector $[n_1, n_2, n_3]^T$ of the triangular grid area are calculated, and the radius r of the visual sphere is determined. The parameters to be optimized are coded and decoded, and the fitness function is calculated.

The algorithm's operation flow chart is shown in Figure 8. First, the structure of the target to be measured is input, as well as the number and type of cameras. The parameters $x_j = [\alpha_j, \beta_j]^T$ of the camera pose that are the azimuth angle and the elevation angle, the rotation vector $R = [\alpha_j, \pi - \beta_j, \pi/2]^T$ and the translation vector $t = [0,0,D]^T$ of the camera external parameter matrix are encoded by binary code to generate the new individuals. To perform subsequent genetic operations, the individuals are divided into subpopulation A and subpopulation B. According to the constraint conditions, the fitness functions $f(A)$ and $f(B)$ corresponding to subpopulations A and B are calculated, respectively, to obtain the optimal individual x_{op} of the two subpopulations initially generated. Except for the optimal individual, all the other individuals are replaced by the next generation. The fifth step in Section 3.1 considers whether large mutations, adaptive crossovers and mutations are needed. The above steps are repeated, update the iteration optimal individual x_{opt} and record the iteration subalgebra k. When $f_{ave}/f(x_{opt}) \geq Fit$, the fitness function exceeds the threshold, the algorithm converges, the global optimal solution is obtained and the iteration ends. The

optimal individual is decoded, and the networking method of the visual measurement network is obtained.

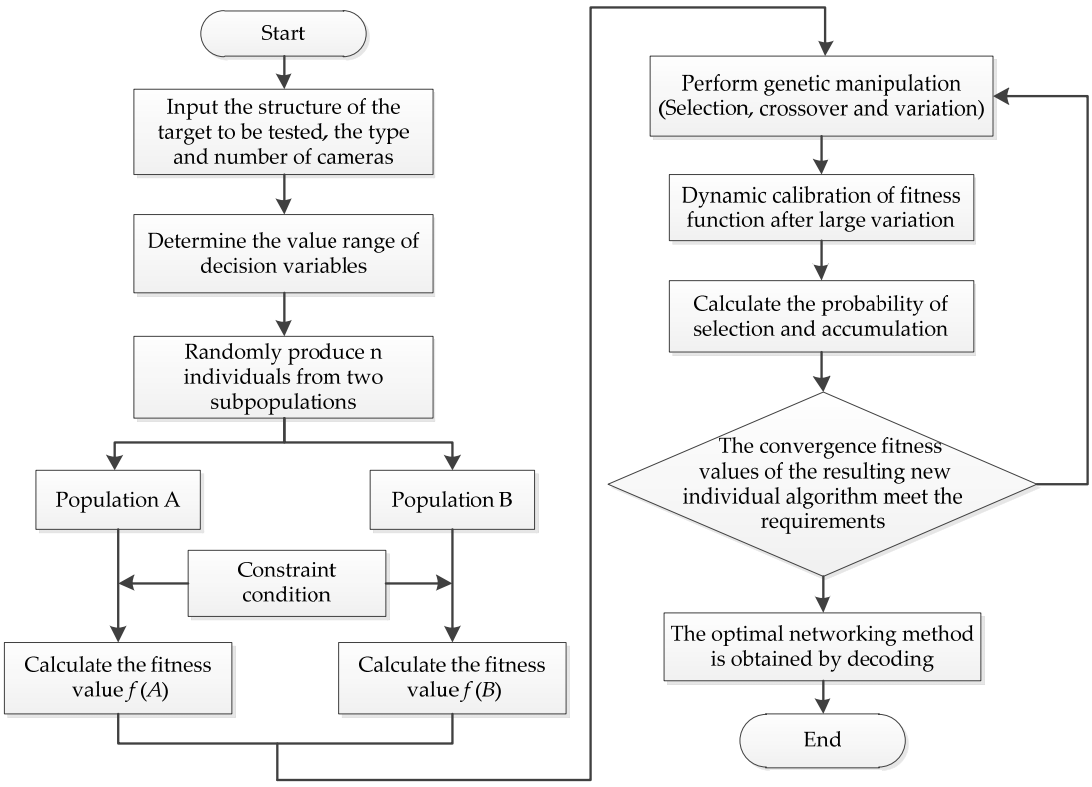


Figure 8. The operation flow of GEA.

4. 3D measurement experiment based on intelligent networking

4.1. Simulation experiment analysis

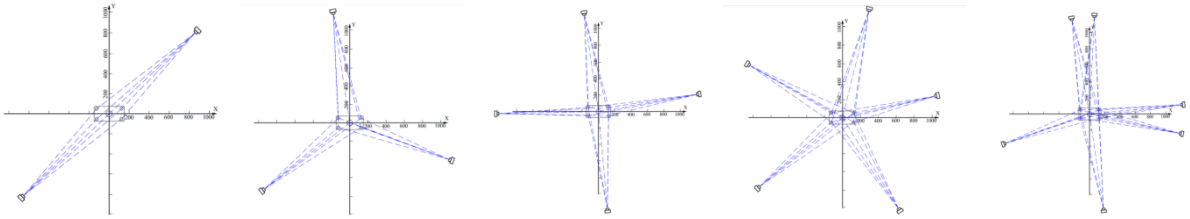
To verify the correctness of the centering networking method, a simulation environment is set up, and the conditions required by the GEA are given. A rectangular plane with one side length 298 × 148 mm is set on the *XOY* plane, and then the normal vector of the plane is $[0, 0, 1]^T$. Taking the center of the plane and four corners within the plane, that is, there are a total of five marked circles, the diameter of the large center circle is $\phi_1 = 59\ mm$ and the diameter of the four small circles is $\varphi = 29$ mm. For the convenience of calculation, the resolution of the camera is defined as 2000 × 2000 pixels, the equivalent focal length is 3000 mm and the radius of the visual sphere is 1500 mm. Then, the spatial resolution of the camera is 0.5 mm/pixel, the range of the azimuth angle is $\alpha \in [0, 2\pi)$ and the range of the elevation angle is $\beta \in [-\pi/2, \pi/2]$. Different numbers of cameras are used for the networking method, and 100 individuals are selected for iteration until the algorithm converges. The camera pose information of each networking camera is shown in Table 1, and the camera pose information is shown in Figure 8. Figure 9(a) is the result of two-camera networking, Figure 9(b) is the result of three-camera networking, Figure 9(c) is the result of four-camera networking, Figure 9(d) is the result of five-camera networking and Figure 9(e) is the result of six-camera networking.

According to the analysis in Section 3.2.1, the optimal range of the azimuth angle of the visual

sphere network is $[10°, 70°]$. In the case of binocular vision, the azimuth angle optimal interval $[25°, 45°]$ should be considered. All the parameters of the visual sphere obtained by the above simulation experiments fall into the optimal interval, and the fitness is small when -0.063; that is, the maximum error of the 3D coordinates in space is 0.063 mm, and the absolute value decreases with increasing camera numbers, which proves the rationality of the centering networking method in this paper. Notably, the acceleration of the fitness reduction decreases gradually, indicating that the fitness will become stable when the number of cameras continues to increase, and the benefit of measurement networking will begin to decrease because of the measurement costs.

Table 1. Camera pose information of each network.

Number of cameras		2	3	4	5	6
Parameters of the visual sphere network	α	43.52° 224.15°	99.52° 217.64° 340.06°	10.15°, 98.17° 180.89°, 275.34°	13.47°, 76.56° 151.22°, 219.94° 301.68°	5.09°, 87.40° 101.06°, 198.24° 279.32°, 348.36°
	β	38.13° 38.13°	36.56° 37.82° 39.22°	35.71°, 37.62° 35.71°, 37.62°	45.24°, 35.71° 37.62°, 36.67° 37.62°	36.67°, 36.67° 37.62°, 39.52° 37.62°, 37.62°
Fitness	Fit	-0.230	-0.142	-0.097	-0.086	-0.063



(a) (b) (c) (d) (e)

Figure 9. Camera pose information in multivisual networking. (a) The result of two-camera networking; (b) The result of three-camera networking; (c) The result of four-camera networking; (d) The result of five-camera networking; (e) The result of six-camera networking.

4.2. 3D measurement experiment of multiview networking for a real workpiece

The experimental platform is composed of an upper computer, mechanical arm system and camera system. The upper computer adopts a Windows 10 x 64-bit control system with 8.00 GB memory for the model ““Inter” (R)”Core” (‘TM’)”i5-4200H CPU@2.80GHz “ “. The upper computer is used to control the motion of the manipulator, take pictures of the target to be measured and extracts its three-dimensional position information. A six-axis manipulator is used to accurately and quickly move the camera system to each viewpoint position of the network. The layout of the experimental platform is shown in Figure 10.

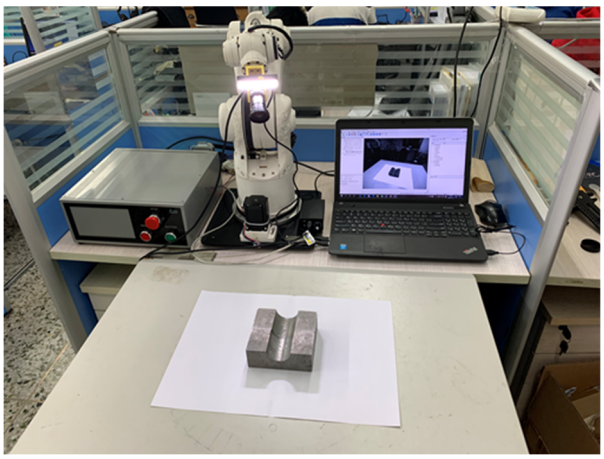
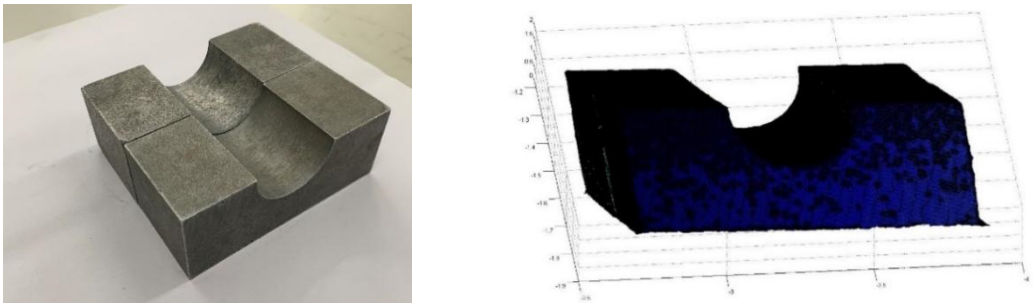


Figure 10. Layout of the experimental platform.

The workpiece to be tested is shown in Figure 11(a), and the 3D CAD model structure diagram of the model to be tested is obtained through the motion recovery structure. Then, the structure diagram is gridded through MATLAB, as shown in Figure 11(b).



(a) (b)

Figure 11. The actual measurement target used to verify the algorithm. (a) The workpiece to be tested; (b) 3D model diagram.

In the experiment, the measurement distance is set to 300 mm, that is, the radius of the visual sphere is 300 mm. The range of the known camera azimuth angle is $\alpha \in [0,2\pi)$, and the range of the elevation angle is $\beta \in [-\pi/2,\pi/2]$. To ensure the best imaging, based on the optimal analysis of the azimuth angle and elevation angle in Section 3.2.1, the settings are as shown in Eqs (25) and (26).

$$\alpha \in [25°,45°] \cup [115°,135°] \cup [205°,225°] \cup [295°,315°] \tag{25}$$

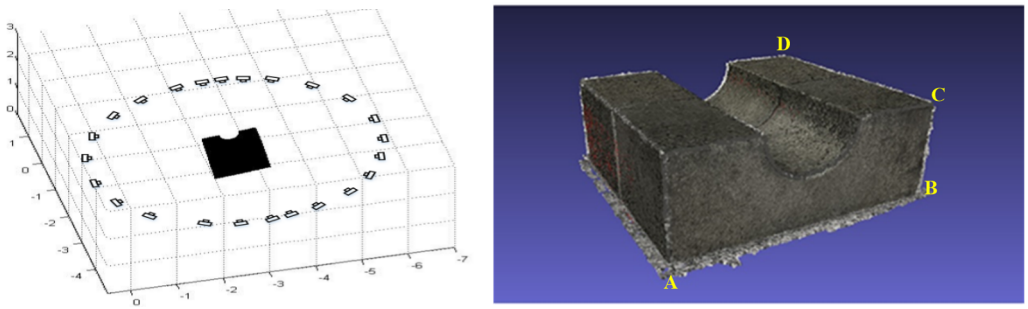
$$\beta \in [10°,\ 70°] \cup [-70°,\ -10°] \tag{26}$$

Different numbers of cameras were used for the networking method. After running the GEA many times until convergence, the performance of the comprehensive fitness function under different camera

numbers draws the fitness of 10, 15, 20, 22, 24, 26, 28 and 30 cameras, as shown in Table 2. The centering networking method with 24 cameras as an example is shown in Figure 12(a), and the dense 3D reconstruction results obtained under this networking method are shown in Figure 12(b).

Table 2. Adaptability of concentric networking experiment.

Number of cameras	10	15	20	22	24	26	28	30
Fitness	-5.158	-3.037	1.231	-1.182	-1.154	-1.088	-1.021	-0.983



(a) (b)

Figure 12. Centering networking and 3D reconstruction results of 24 cameras. (a) Pose diagram of the 24 cameras centering networking method; (b) The dense 3D reconstruction result.

4.3. Precision analysis of measurement

The reconstruction measurement results are analyzed based on the photographic geometry idea, and the feasibility and accuracy of the core networking method are analyzed quantitatively. To obtain the mapping scale relationship between the actual size and the reconstruction size, the calibration plate is reconstructed and measured. Three squares, A, B and C, with rich reconstruction information are selected, and three intersections are recorded as $A_1A_2A_3$, $B_1B_2B_3$ and $C_1C_2C_3$, respectively. The results are shown in Figure 13.

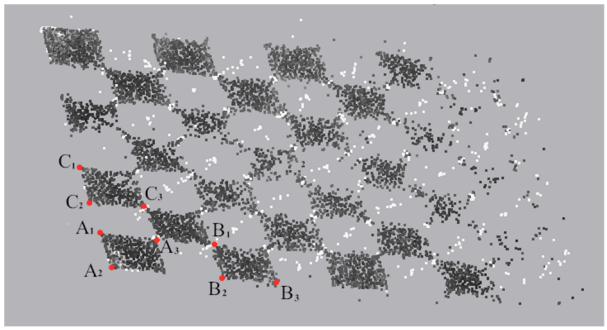


Figure 13. Calibration board point cloud.

The mapping scale Factor s represents the ratio between reconstruction size $L_{reconstruction}$ and the

actual size L_{true}, as shown in Eq (27). The scale factor remains unchanged under the same camera. The error between the reconstruction size and actual size can be compared and analyzed by using the scale factor calculated by the calibration plate, and the error calculation is shown in Eq (28). Thus, a reconstructed coordinate system with the pixel size in units is constructed. The size information is shown in Table 3.

$$s = \frac{L_{\text{reconstruction}}}{L_{true}} \tag{27}$$

$$e = |dr - dt/s| \tag{28}$$

In the formula, $L_{\text{reconstruction}}$ is the reconstruction size, L_{true} is the actual size, e is the size error, dr is the actual size of the target to be measured and dt is the reconstruction size after photographic geometric mapping.

Table 3. Calibration board size information table.

	A_1A_2	A_1A_3	B_1B_2	B_2B_3	C_1C_2	C_2C_3
Actual size	25 mm	25 mm	25 mm	25 mm	25 mm	25 mm
Rebuild dimensions	0.2450 mm	0.2369 mm	0.2381 mm	0.2421 mm	0.2383 mm	0.2355 mm
Mapping Scale s	0.0098	0.009476	0.009524	0.009684	0.009532	0.00942

The shape of the object measured by the centering network is regular, and the size can be used as the evaluation target. The average of each mapping ratio in Table 3 is calculated to obtain s = 0.009573. Then, the reconstructed dimensions and errors of the target under the centering networking method are shown in Table 4.

Table 4. Comparison of dimensional errors.

	The length AB	The width CD	The height BC	Diameter of inner hole
Actual size	120 mm	108 mm	46 mm	44 mm
Rebuild dimensions	1.1470 mm	1.0235 mm	0.4379 mm	0.4178 mm
Error e	0.1839 mm	0.1085 mm	0.2568 mm	0.3564 mm

Table 4 shows that the error between the reconstructed size and the actual size under the centering networking method is controlled within half a millimeter. Compared with the expensive noncontact measurement system, the experimental layout is simple, which saves significant costs and meet the needs of industrial topography measurements. The 3D reconstruction image is clear and has a fine texture, which meets the needs of human-computer interaction in industrial measurements. The correctness and feasibility of the centering networking method designed in this paper are verified.

5. Conclusions

The intelligence of multivision networking enables 3D measurement technology based on vision to implement low cost and real-time measurements. In this paper, a multivision intelligence networking method based on a genetic evolution algorithm is proposed. By using the proposed method, centering multivision intelligence networking is established. GEA is a heuristic algorithm, and each operation

may have different results, so it is necessary to deploy it as many times as possible and select the best individual according to the fitness function. In the centering networking method, the camera position coordinates and pose coordinates are designated as the parameters to be optimized in combination with the visual sphere model. Considering the constraints encountered in the camera measurements, the fitness function is improved to reduce the computation amount, and ultimately, an optimal centering networking method is obtained.

The following work further studies the intelligent visual networking measurement of irregular shapes based on the method proposed in this paper, and is called, scattered networking. It is expected that scattered networking will implement real-time 3D measurements of complex shapes and large structures with multivision flexible networking.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by Yangzhou City, 2021, leading talents of the "Green Yang Jinfeng project" (Innovation in Colleges and Universities) (Grant no.YZLYJFJH2021CX044) and the Science and Technology Talent support project of Jiangsu Province, China (Grant no. FZ20211137).

Conflict of interest

The authors declare no conflict of interest.

References

1. Y. Tang, L. Li, C. Wang, M. Chen, W. Feng, X. Zou, et al., Real-time detection of surface deformation and strain in recycled aggregate concrete-filled steel tubular columns via four-ocular vision, *Rob. Comput. Integr. Manuf.*, **59** (2019), 36–46. https://doi.org/10.1016/j.rcim.2019.03.001
2. Y. Wang, V. C. Guizilini, T. Zhang, Y. Wang, H. Zhao, J. Solomon, Detr3d: 3d object detection from multi-view images via 3d-to-2d queries, in *Conference on Robot Learning,* PMLR, (2022), 180–191. https://doi.org/10.48550/arXiv.2110.06922
3. Y. Tang, Z. Huang, Z. Chen, M. Chen, H. Zhou, H. Zhang, et al., Novel visual crack width measurement based on backbone double-scale features for improved detection automation, *Eng. Struct.*, **274** (2023), 115158. https://doi.org/10.1016/j.engstruct.2022.115158
4. M. Chen, Y. Tang, X. Zou, K. Huang, L. Li, Y. He, High-accuracy multi-camera reconstruction enhanced by adaptive point cloud correction algorithm, *Opt. Lasers Eng.*, **122** (2019), 170–183. https://doi.org/10.1016/j.optlaseng.2019.06.011
5. Available from: https://www.keyence.com.cn/products/measure/laser-2d/.
6. R. Chen, S. Han, J. Xu, H. Su, Point-based multi-view stereo network, in *Proceedings of the IEEE/CVF international conference on computer vision*, (2019), 1538–1547.

7. A. Chen, Z. Xu, F. Zhao, X. Zhang, F. Xiang, J. Yu, et al., Mvsnerf: Fast generalizable radiance field reconstruction from multi-view stereo, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, (2021), 14124–14133.
8. R. Weilharter, F. Fraundorfer, Highres-mvsnet: A fast multi-view stereo network for dense 3d reconstruction from high-resolution images, *IEEE Access*, **9** (2021), 11306–11315. https://doi.org/10.1109/ACCESS.2021.3050556
9. S. Shuai, Y. Ling, S. Liang, H. Zhong, X. Tian, C. Liu, et al., Research on 3D surface reconstruction and body size measurement of pigs based on multi-view RGB-D cameras, *Comput. Electron. Agric.*, **175** (2020), 105543. https://doi.org/10.1016/j.compag.2020.105543
10. T. Wu, J. H. Liu, S. Liu, P. Jin, H. Huang, W. Liu, A measurement method of free-form tube based on multi-view vision for industrial assembly, *Assem. Autom.*, **40** (2020), 553–564. https://doi.org/10.1108/AA-05-2019-0087
11. L. Yang, Q. Huang, X. Song, M. Li, C. Hou, Z. Xiong, Girth measurement based on multi-view stereo images for garment design, *IEEE Access*, **8** (2020), 160338–160354. https://doi.org/10.1109/ACCESS.2020.3021019.
12. M. Saadat-Seresht, F. Samdzadegan, A. Azizi, M. Hahn, Camera placement for network design in vision metrology based on fuzzy inference system, in *XXth ISPRS Congress*, **16** (2004), 88.
13. Y. Cai, X. Qin, Q. Liu, S. Zhang, Study on high-precision inspection multi-vision cameras space layout, in *2010 Second International Conference on Communication Systems, Networks and Applications*, **1** (2010), 168–171. https://doi.org/10.1109/ICCSNA.2010.5588677
14. B. S. Alsadik, M. Gerke, G. Vosselman, Optimal camera network design for 3D modeling of cultural heritage, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, **3** (2012), 7–12. https://doi.org/10.5194/isprsannals-I-3-7-2012
15. H. Chen, K. Ding, C. Qin, P. Zhang, M. Xiao, Optimization design of camera pose for network base on multi-view stereo vision, in *Journal of Physics: Conference Series*, **2170** (2022), 012018. https://doi.org/10.1088/1742-6596/2170/1/012018
16. J. Chen, M. Dong, Photogrammetric network design for large-scale trough concentrator surface measurement, *Renewable Energy Res.*, **34** (2016), 1671–5292. https://doi.org/10.13941/j.cnki.21-1469/tk.2016.03.006
17. Y. Qiao, S. Tan, J. Jiang, A multi-vision measurement network planning strategy, *Acta Opt. Sin.*, **38** (2018), 216−223. https://doi.org/10.3788/AOS201838.0515005
18. Y. Qiao, B. Jia, J. Jiang, J. Wang, Networking method of multi-view stereo-vision measurement network, *Infrared Laser Eng.*, **49** (2020), 20190492. https://doi.org/10.3788/IRLA20190492
19. E. Eriksson, G. Dán, V. Fodor, Coordinating distributed algorithms for feature extraction offloading in multi-camera visual sensor networks, *IEEE Trans. Circuits Syst. Video Technol.*, **28** (2017), 3288–3299. https://doi.org/10.1109/TCSVT.2017.2745102
20. M. Ren, L. Wang, J. Zhao, Z. Tang, Viewpoint planning of surface structured light scanning for complex surface parts, *Chin. Opt.*, **16** (2023), 1–14. http://dx.doi.org/10.37188/CO.2022-0026
21. R. Mahony, T. Hamel, J. Trumpf, An homogeneous space geometry for simultaneous localisation and mapping, *Annu. Rev. Control*, **51** (2021), 254–267. https://doi.org/10.1016/j.arcontrol.2021.04.012
22. N. D. M. Phan, Y. Quinsat, S. Lavernhe, C. Lartigue, Scanner path planning with the control of overlap for part inspection with an industrial robot, *Int. J. Adv. Manuf. Technol.*, **98** (2018), 629–643. https://doi.org/10.1007/s00170-018-2336-8

23. I. D. Lee, J. H. Seo, Y. M. Kim, J. Choi, S. Han, B. Yoo, Automatic pose generation for robotic 3-D scanning of mechanical parts, *IEEE Trans. Rob.*, **36** (2020), 1219–1238. https://doi.org/10.1109/TRO.2020.2980161
24. P. Osiński, J. Markiewicz, J. Nowisz, M. Remiszewski, A. Rasiński, R. Sitnik, A novel approach for dynamic (4d) multi-view stereo system camera network design, *Sensors*, **22** (2022), 1576. https://doi.org/10.3390/s22041576
25. S. Mirjalili, Evolutionary algorithms and neural networks, in *Studies in computational intelligence*, Berlin/Heidelberg, Germany, Springer, **780** (2019).
26. A. Slowik, H. Kwasnicka, Evolutionary algorithms and their applications to engineering problems, *Neural Comput. Appl.*, **32** (2020), 12363–12379. https://doi.org/10.1007/s00521-020-04832-8
27. R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.
28. Y. Qiao, S. Gao, Y. Pi, Optimization method of node pose for measurement network based on multi-view stereo vision, *Laser Optoelectron. Prog.*, **56** (2019).

AIMS Press