RESEARCH ARTICLE

A Novel Global Calibration Approach for Multi-Viewpoint 3D Vision Measuring System

PENGFEI SUN[1], **HAISHU TAN**[2,3], **(Member, IEEE), AND FUQIANG ZHOU**[1]

[1]School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China
[2]School of Physics and Optoelectronic Engineering, Foshan University, Foshan 528000, China
[3]Ji Hua Laboratory, Foshan 528000, China

Corresponding author: Haishu Tan (tanhaishu@fosu.edu.cn)

ABSTRACT To unify the local 3D point clouds for the measured object with big size or complex structure to a global coordinate system, we usually need to adopt global calibration approach to calibrate the transformation relationship between the global binocular vision sensor (GBVS) and measuring binocular sensor (MBS) in different viewpoints in the 3D Vision Measuring System (VMS). The field of view (FOV) of the two global cameras in GBVS determines the measurement range of the 3D VMS in the global calibration process. In the traditional 3D VMS of single-station multi-viewpoint, the measurement range is limited because of the two global cameras in the GBVS are needed to have a public FOV. Then, the global calibration process is also limited. And the global calibration process is complex and easy to cause error accumulation when a laser tracker is used to replace the GBVS in the 3D VMS. In order to solve this problem, we proposed a novel global calibration approach based on multi-directional target (MDT) for a novel designed 3D VMS. It can achieve the function of global calibration for the designed 3D VMS once the MDT appears in the FOV of the global camera. Finally, we adopt two global cameras which don't need to have a public FOV to form GBVS. It can further expand the measurement range of the 3D VMS. And the global calibration efficiency can also be improved at the same time. The feasibility and effectiveness of the proposed global calibration algorithm are verified through experiments. The experimental results show that the accuracy of the proposed global calibration algorithm can reach up to 0.08 mm after optimizing. It has broad application prospects and practical theoretical research value.

INDEX TERMS Multi-viewpoint, global calibration, global binocular vision sensor, 3D vision measuring system, transformation matrix.

I. INTRODUCTION

With the rapidly development of 3D measurement technology [1], 3D reconstruction technology [2], [3] has been widely used in workpiece flaw detection [4], bioengineering [5], cultural relic protection [6], virtual reality [7] and other fields. The 3D VMS [8], [9], [10] is often used for 3D reconstruction of the measured objects. A 3D VMS usually consists of projector and MBS. Projector projects structured light to the measured object. The MBS acquires the structured light image. Then, a 3D reconstruction of the measured object is completed by processing the structured light image.

However, one scan can only complete the local 3D reconstruction of the measured object when the measured object has a large size or complex structure. So, it is necessary to unify the local 3D point clouds obtained from different viewpoints into the global coordinate system [11] in order to complete the overall 3D reconstruction for the measured object. Therefore, we can see that the global calibration of the transformation relationship between local 3D point clouds obtained from multiple viewpoints by the MBS is very important.

Since the local 3D point clouds is obtained by the MBS, the global calibration of the transformation relationship between the local 3D point clouds is also the global calibration of the MBS in different viewpoints. According to the different

---

principles of global calibration, the global calibration approach is mainly divided into the following five types: 1. Global calibration with common region. We must make sure that there is a common region between the adjacent local 3D point clouds [12], [13]. The transformation relationship between the adjacent local 3D point clouds can be calculated according to the common region. Then, the global calibration can be completed by the calculated transformation relationship. This approach needs to deal with a large amount of point clouds information, which is time-consuming. And the global calibration function cannot be completed when the adjacent point clouds have no common regions. 2. Global calibration with marker. Paste marker [14], [15] on the measured object, and then, we can calibrate the transformation relationship of the local point clouds according to the marker. The pasted marker will contaminate the surface of the measured object. And when the aim of the 3D reconstruction of the measured object is for defect detection, it is will result in failure to detect the defect if the marker is coincidental pasted on the defect of the measured object. Therefore, this approach is not suitable for defect detection of the measured object. And it is also not suitable for the measured objects that are prone to expansion or corrosion. 3. Global calibration with the platform. Put the measured object on a moving platform, and the transformation relationship between the local point clouds of the measured object [16], [17] can be provided by the moving platform. Then, the global calibration can realize by the provided transformation relationship between the local point clouds. This method strictly depends on the positioning accuracy of the moving platform, and the process of human-computer interaction is complex. It is only suitable for the measurement of small object. 4. Global calibration with the mechanical arm. The measuring head is rigidly connected to the mechanical arm [18], [19], and the pose of the local point clouds at each viewpoint can be provided by the mechanical arm. Then, the global calibration can be completed by the provided transformation relationship between the local point clouds. This approach strictly depends on the accuracy of the mechanical arm and needs to calibrate the mechanical arm in advance. The global calibration is inability completed when the scanning range exceeds the working range of the mechanical arm. 5. Global calibration of single-station multi-viewpoint [20], [21], [22]. In this method, a single-station multi-viewpoint 3D VMS is needed. The system is usually composed by a GBVS (composed by two global cameras with public FOV) and a measuring head (composed by positioning marker and MBS). The positioning marker is rigidly connected with the MBS. The MBS is used to collect the image of the measured object. The GBVS is used to collect the image of the positioning marker Synchronous. Take the positioning marker as intermediate medium, then, the global calibration of the transformation relationship between the MBS and GBVS can be realized with the positioning marker. This method has the advantages of high accuracy and flexible usage. It is suitable for the measurement the objects with various size. Therefore, the global calibration of single-station

multi-viewpoint approach has been widely used in the field of 3D reconstruction. The scholars all over the world have also carried out in-depth research on it.

Barone [20] et al proposed a multi-viewpoint shape measurement method based on 3D optical measuring system. This method needs to set up a series number of reverse reflection spheres with known transformation relationship on the 3D optical measuring system. It needs to build a standard model in advance for the reverse reflection spheres. Compare the standard model with the reverse reflection sphere route which moving in space to complete the global calibration. It is required different spatial triangles which formed by every three reverse reflection spheres in the process of global calibration. So, at least three reverse reflection spheres must be captured each time, and a GBVS (composed by two global cameras with public FOV) must be used to complete the global calibration. The measurement range is limited and the structure of the 3D optical scanner is complex. The calibration condition is harsh. Shi [21] et al proposed a large-scale 3D measuring system based on LED. This method adopts LED as the medium to complete the global calibration. But, the light of LED will be distorted in the process of propagation due to the reflection and scattering of light. This will result in a decrease of the accuracy for the feature points extracted from the image. Thereby reducing the accuracy of global calibration. It also needs a GBVS (composed by two global cameras with public FOV) to complete the global calibration. Zhou [22] et al proposed a global calibration approach based on multidimensional combined collaboration target (MCCT). This approach completes the global calibration by collecting the image of the reflection sphere on the MCCT by the laser tracker. The multiple reflection spheres on the MCCT need strict constraints for distance and angle. The structure of the MCCT is complex. It also needs a 3-coordinate measuring machine to calibrate the geometric relationship between the reflection spheres in advance. The calibration process is easy to cause error accumulation, and the usage of the scene is limited. It also has a high cost.

In summary, the traditional single-station multi-viewpoint 3D VMS mainly has two disadvantages: 1. The two global cameras which constitute the GBVS to collect images of the positioning marker must have a public FOV, which will lead to the reduction of the measurement range of the system. 2. The measurement range is extended when a laser tracker is used to capture images of the positioning marker. But a reflector needs to be installed on the measured object to reflect the laser to the laser tracker. The Laser is easy to be affected by the external interference factors such as illumination, vibration and so on. Then, the laser tracker may not receive the correct laser information because of the external interference factors. Besides that, the maintenance cost of the laser tracker is very high. In order to overcome these limitations, a 3D VMS is designed in this paper. We also propose a global calibration approach aiming at the designed 3D VMS. With the proposed approach, we can use two global cameras without the public FOV to form the GBVS to expand

the measurement range of the designed 3D VMS. Our method is flexible to use and the system designed by this paper is low cost. It is can be suitable for a variety of scenes.

The organization of this paper is as follows: Section II introduces the overall measurement model of the designed 3D VMS. And we illustration the proposed global calibration algorithm in Section III. The experiments and results are given to verify the feasibility and effectiveness of the proposed global calibration approach in Section IV. In section V, the advantages and disadvantages of the designed 3D VMS and the proposed global calibration approach are discussed.

II. MEASUREMENT MODEL

The designed 3D VMS is composed by GBVS (composed by left global camera and right global camera) and measuring head. The 3D VMS is shown in Fig. 1 (a). The measuring head is shown in Fig. 1 (b).
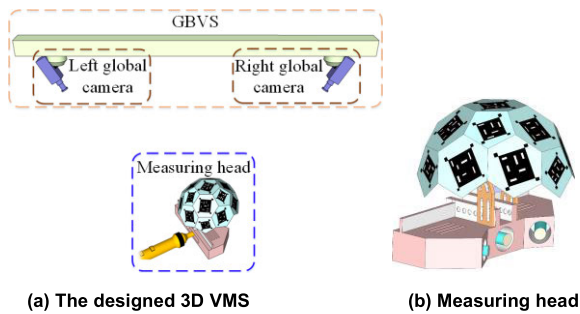

(a) The designed 3D VMS   **(b) Measuring head**

FIGURE 1. Schematic diagram of the designed 3D VMS.

The designed measuring head is composed by MBS, MDT, and connector [23]. The morphologies of the MBS, MDT, and connector are shown in Fig. 2, where (a) represents the MBS for collecting local 3D point clouds of the measured object, (b) represents the MDT which is used for the GBVS to collect images and (c) represents the connector for rigidly connecting the MBS and MDT.
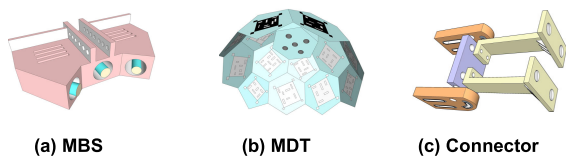

(a) MBS   **(b) MDT**   **(c) Connector**

FIGURE 2. The main components of the measuring head.

The measurement model of the designed 3D VMS is shown in Fig. 3. The scanning head is controlled by manual or manipulator to move around the measured object and the standard target. The MBS collects the local 3D point clouds information of the measured object and the standard target synchronously. At the same time, we keep the pose of the GBVS unchanged and the GBVS collects MDT images at each viewpoint. The local 3D point clouds information collected by the MBS can be unified into the coordinate system of the GBVS by using the MDT as intermediate medium.

Then, the function of global unity of the local 3D point clouds can be realized. We need to define some symbols and the coordinate systems in advance in order to describe the measurement model of the designed 3D VMS. First, we establish the left global camera coordinate system as global coordinate system $o_g - x_g y_g z_g$ and establish the right global camera coordinate system $o_r - x_r y_r z_r$. The transformation matrix from $o_r - x_r y_r z_r$ to $o_g - x_g y_g z_g$ is called $\boldsymbol{T}_g^r$. The MBS coordinate system and MDT coordinate system is established as $o_s - x_s y_s z_s$ and $o_t - x_t y_t z_t$ respectively. Then, the transformation matrix from $o_s - x_s y_s z_s$ to $o_t - x_t y_t z_t$ is called $\boldsymbol{T}_t^s$. We make $o_{sn} - x_{sn} y_{sn} z_{sn}$ represent the coordinate system of MBS at viewpoint n ($n = \{1, 2, \ldots, N | N \in N^+\}$). Then, at viewpoint n, the transformation matrix from $o_t - x_t y_t z_t$ to $o_g - x_g y_g z_g$ is $\boldsymbol{T}_g^{t_n}$, and the transformation matrix from $o_t - x_t y_t z_t$ to $o_r - x_r y_r z_r$ is $\boldsymbol{T}_r^{t_n}$.
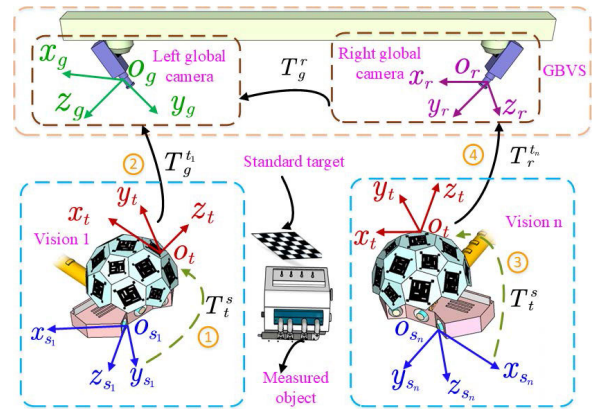

FIGURE 3. Measurement model of the 3D VMS.

At the viewpoint n, we can calculate $\boldsymbol{T}_g^{t_n}$ by Equation (1) when only the right global camera can capture the image of the MDT. When the left global camera can capture the MDT image directly, we can directly calculate $\boldsymbol{T}_g^{t_n}$ according to the PNP principle [24].

$$\boldsymbol{T}_g^{t_n} = \boldsymbol{T}_r^{t_n} \cdot \boldsymbol{T}_g^r \tag{1}$$

where, the transformation matrix $\boldsymbol{T}$ is composed of rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{t}$, as shown in (2).

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0} & 1 \end{bmatrix} \tag{2}$$

Then, we can solve the transformation matrix $\boldsymbol{T}_g^{s_n}$ from $o_{sn} - x_{sn} y_{sn} z_{sn}$ to $o_g - x_g y_g z_g$ at the viewpoint n by (3). $\boldsymbol{T}_g^{s_n}$ represents the result of the global calibration. $\boldsymbol{T}_g^{s_n} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_n}$ or $\boldsymbol{T}_g^{s_n} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_r^{t_n} \cdot \boldsymbol{T}_g^r$ can be chosen when the left global camera or right global camera is used to collect the MDT images respectively.

$$\boldsymbol{T}_g^{s_n} = \begin{cases} \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_n}, \, left \\ \boldsymbol{T}_t^s \cdot \boldsymbol{T}_r^{t_n} \cdot \boldsymbol{T}_g^r, \, right \end{cases} \tag{3}$$

At viewpoint n, $\boldsymbol{P}^{s_n}$ represents the 3D coordinates of the feature points on the measured object in $o_{sn} - x_{sn} y_{sn} z_{sn}$.

Then, we can transform $\boldsymbol{P}_g^{s_n}$ to $o_g - x_g y_g z_g$ to calculate $\boldsymbol{P}_g^{s_n}$ by (4).

$$\boldsymbol{P}_g^{s_n} = \boldsymbol{T}_g^{s_n} \cdot \boldsymbol{P}^{s_n} \tag{4}$$

Through all the above process, we can know that it is necessary to calculate $\boldsymbol{T}_g^{s_n}$ to transform $\boldsymbol{P}^{s_n}$ to $o_g - x_g y_g z_g$ under the premise of $\boldsymbol{P}^{s_n}$ is known. The function of global unified for the local 3D point clouds of the measured object in different viewpoints can be realized with the calculated $\boldsymbol{T}_g^{s_n}$. Therefore, the solution of $\boldsymbol{T}_g^{s_n}$ is the key in the process of global unification for the local 3D point clouds. The process of calculating $\boldsymbol{T}_g^{s_n}$ is called global calibration for the designed 3D VMS. We can complete the preliminary solution of $\boldsymbol{T}_g^{s_n}$ by (1), (2) and (3). In Fig. 3, we can get $\boldsymbol{T}_g^{s_1}$ by ① and ② at viewpoint 1. In the same way, $\boldsymbol{T}_g^{s_n}$ can be calculated by ③ and ④ at viewpoint n. However, the accuracy of $\boldsymbol{T}_g^{s_n}$ calculated at the preliminary may be low. This will result in the low accuracy of the global calibration. So, we optimize $\boldsymbol{T}_g^{s_n}$ with the following measurement method.

First of all, we make $\boldsymbol{P}_d^{b_n}$ ($d = \{1, 2, \ldots, M | M \in N^+\}$, M represents the total number of the feature points) represent the coordinates of the feature points in the standard target in $o_{s_n} - x_{s_n} y_{s_n} z_{s_n}$, where d represents the dth feature points in the standard target. Then, we can calculate $\boldsymbol{P}_d^{g_n}$ by transferring $\boldsymbol{P}_d^{b_n}$ to $o_g - x_g y_g z_g$ by (5).

$$\boldsymbol{P}_d^{g_n} = \boldsymbol{T}_g^{s_n} \cdot \boldsymbol{P}_d^{b_n} \tag{5}$$

We calculate $E_g^{b_n}$ According to (6). $E_g^{b_n}$ represents the average distance between the corresponding $\boldsymbol{P}_d^{g_n}$ and $\boldsymbol{P}_d^{g_{n+1}}$. $\boldsymbol{P}_d^{g_{n+1}}$ represents the coordinates of feature points in the standard target in $o_g - x_g y_g z_g$ at viewpoint $(n+1)$.

$$E_g^{b_n} = \frac{1}{M} \sum_{d=1}^{M} \left\| \boldsymbol{P}_d^{g_n} - \boldsymbol{P}_d^{g_{n+1}} \right\|_2 \tag{6}$$

When $\boldsymbol{P}_d^{b_n}$ and $\boldsymbol{P}_d^{b_{n+1}}$ remain unchanged, we can know that, the closer $E_g^{b_n}$ is to 0, the higher the accuracy of $\boldsymbol{T}_g^{s_n}$ we can get. Therefore, we propose an optimization algorithm by minimizing $E_g^{b_n}$ to optimize $\boldsymbol{T}_g^{s_n}$ to improve the overall accuracy of the global calibration according to this principle.

III. GLOBAL CALIBRATION ALGORITHM

The designed algorithm of global calibration consists of two parts: 1. Main algorithm. 2. Optimization algorithm. The main algorithm is used to complete the initial solution of $\boldsymbol{T}_g^{s_n}$. And the optimization algorithm is used to improve the accuracy of the initial $\boldsymbol{T}_g^{s_n}$.

A. MAIN ALGORITHM

The main algorithm consists of two parts: 1. Calibrate the transformation matrix $\boldsymbol{T}_t^s$. 2. Calibrate the transformation matrix $\boldsymbol{T}_g^{s_n}$ at different viewpoints.

1) CALIBRATION BETWEEN MBS AND MDT

Through (3), we can know that it is necessary to calculate $\boldsymbol{T}_t^s$ before calculating $\boldsymbol{T}_g^{s_n}$. We use optical measurement method

to calculate $\boldsymbol{T}_t^s$ in this paper. The measurement model of $\boldsymbol{T}_t^s$ is shown as Fig. 4. In Fig. 4, The binocular vision sensor (BVS) is composed of left camera and right camera. The intermediate plane target is placed in front of the MBS. The MBS and the left camera of the BVS collect the intermediate plane target image respectively. And the right camera of the BVS captures the MDT image synchronously.
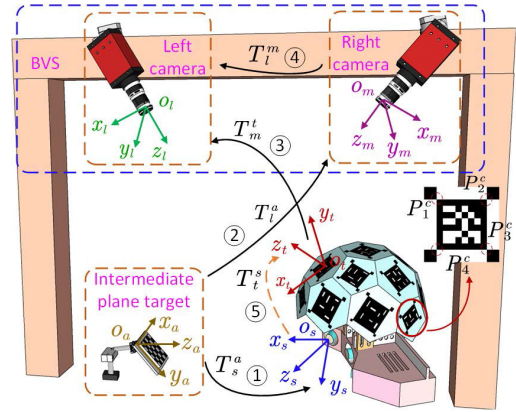

FIGURE 4. The measurement model of T_t^s.

We establish intermediate plane target coordinate system $o_a - x_a y_a z_a$, the right camera of the BVS coordinate system $o_m - x_m y_m z_m$, the left camera of the BVS coordinate system $o_l - x_l y_l z_l$ and the image coordinate system $o - uv$. The 3D coordinates of the feature points in the intermediate plane target in $o_a - x_a y_a z_a$ is $\boldsymbol{P}_a = (X_a, Y_a, Z_a)$. The 2D coordinates of the feature points in the intermediate plane target in $o - uv$ corresponding to MBS and right camera of the MBS are $\boldsymbol{p}_a^s = (x_a^s, y_a^s)$ and $\boldsymbol{p}_a^l = (x_a^l, y_a^l)$ respectively. We can calculate the transformation matrix $\boldsymbol{T}_s^a$ from $o_a - x_a y_a z_a$ to $o_s - x_s y_s z_s$ and $\boldsymbol{T}_l^a$ from $o_a - x_a y_a z_a$ to $o_l - x_l y_l z_l$ with PNP principle after correct matching $\boldsymbol{P}_a$ and $\boldsymbol{p}_a^s$, $\boldsymbol{P}_a$ and $\boldsymbol{p}_a^l$. Show as ① and ② in Fig. 4. Then, the transformation matrix $\boldsymbol{T}_s^l$ from $o_l - x_l y_l z_l$ to $o_s - x_s y_s z_s$ can be solved by (7).

$$\boldsymbol{T}_s^l = (\boldsymbol{T}_s^a)^{-1} \cdot \boldsymbol{T}_l^a \tag{7}$$

The MDT consists of multiple coplanar small targets (CSTs) and each CST has four feature points. In Fig. 4, we can see that the feature points on a single CST in the MDT are $\boldsymbol{P}_h^c = \{P_1^c, P_2^c, P_3^c, P_4^c\}$. We extract the 2D coordinates of the feature points in the image collected by the right camera. Make N_1 represent the total number of feature points on the MDT, $\boldsymbol{p}_i^m (i = \{1, 2, \ldots, N_1 | N_1 \in N^+\})$ represent the ith feature point in MDT. The corresponding 3D coordinates $\boldsymbol{P}_i^m$ in $o_t - x_t y_t z_t$ can be calculated by the following 4 steps according to the reference [23]:

① The camera collects MDT images from multiple angles. And we can extract the 2D coordinates $\boldsymbol{p}_h^c$ of the feature points in MDT images. Based on $\boldsymbol{p}_h^c$ and the known side length of the CST, the transformation

relationship between the coordinate system of all CSTs and the camera coordinate system is solved.

② We need calculate the angle θ between the plane normal vector of a CST and the camera optical axis. Then, the CST whose θ is too large should be eliminated. After that, the images are reordered to ensure that there is at least one same CST in two adjacent images. Then, transform $\boldsymbol{P}_h^c$ to the corresponding camera coordinate system to get $\boldsymbol{P'}_h^c$.

③ According to the ICP principle, all $\boldsymbol{P'}_h^c$ can be transformed to the global coordinate system. Then, the initial value of $\boldsymbol{P}_h^{c_t}$ can be calculated.

④ Reproject $\boldsymbol{P}_h^{c_t}$ to the corresponding images to get the reprojection points $\boldsymbol{p}_h^{c_t}$. Then, we can calculate the reprojection error between $\boldsymbol{P}_h^{c_t}$ and $\boldsymbol{p}_h^{c_t}$. After that, we can establish the optimization objective function according to the distribution relationship between θ and reprojection error. Then, $\boldsymbol{P}_i^m$ with high accuracy can be got after the global optimization.

After get $\boldsymbol{p}_i^m$ and $\boldsymbol{P}_i^m$, we can calculate the transformation matrix $\boldsymbol{T}_m^t$ from $o_t - x_t y_t z_t$ to $o_m - x_m y_m z_m$ according to PNP principle, show as ③ in Fig. 4. And then, the transformation matrix $\boldsymbol{T}_l^t$ from $o_t - x_t y_t z_t$ to $o_l - x_l y_l z_l$ can be solved by using the principle of binocular vision calibration. After that, we can calculate the transformation matrix $\boldsymbol{T}_l^t$ from $o_t - x_t y_t z_t$ to $o_l - x_l y_l z_l$ by (8).

$$\boldsymbol{T}_l^t = \boldsymbol{T}_m^t \cdot \boldsymbol{T}_l^m \qquad (8)$$

At last, we can calculate the transformation matrix $\boldsymbol{T}_t^s$ from $o_s - x_s y_s z_s$ to $o_t - x_t y_t z_t$ by simultaneous (7) and (8), shown as (9).

$$\boldsymbol{T}_t^s = \left(\boldsymbol{T}_l^t \cdot \boldsymbol{T}_s^l\right)^{-1} \qquad (9)$$

2) CALIBRATION BETWEEN MBS AND GBVS

We need calculate $\boldsymbol{T}_g^{t_n}$ to get $\boldsymbol{T}_g^{s_n}$ after solving $\boldsymbol{T}_t^s$ with (3). The process of solving $\boldsymbol{T}_g^{t_n}$ is as follows: Firstly, calibrate the parameters of GBVS to calculate $\boldsymbol{T}_g^r$ by using Zhang's calibration method [25]. At viewpoint n, when only the right global camera of the GBVS can capture the MDT image, we can calculate $\boldsymbol{T}_g^{t_n}$ according to (1). When the left global camera of GBVS can capture the MDT image, we can directly calculate $\boldsymbol{T}_g^{t_n}$ with PNP principle.

In fact, the process of unifying $\boldsymbol{P}^{s_n}$ to $o_g - x_g y_g z_g$ is the process of registering the point clouds of previous and next viewpoint for the MBS. So, it is only necessary to research the process of unifying $\boldsymbol{P}^{s_n}$ at the previous and next viewpoints to $o_g - x_g y_g z_g$. According to the above analysis, we can know that the global calibration can be completed by researching the process of calculating the transformation matrix $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ from $o_s - x_s y_s z_s$ to $o_g - x_g y_g z_g$ in viewpoint 1 and viewpoint 2. We can get four cases when solving $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ according to the situation of using the left global camera or right global camera when the GBVS captures the MDT pictures, as shown in Table 1.

TABLE 1. Four cases of global calibration for two viewpoints.

Case	1		2		3		4	
Viewpoint	1	2	1	2	1	2	1	2
Camera	left	left	left	right	right	left	right	right

In case 1, when both of viewpoint 1 and 2 adopt the left global camera to collect MDT images, $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ can be solved respectively, as shown in (10).

$$\begin{cases} \boldsymbol{T}_g^{s_1} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_1} \\ \boldsymbol{T}_g^{s_2} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_2} \end{cases} \qquad (10)$$

In case 2, when viewpoint 1 adopts the left global camera and viewpoint 2 adopts the right global camera, $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ can be solved respectively with (11).

$$\begin{cases} \boldsymbol{T}_g^{s_1} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_1} \\ \boldsymbol{T}_g^{s_2} = \boldsymbol{T}_t^s \cdot \left(\boldsymbol{T}_r^{t_2} \cdot \boldsymbol{T}_g^r\right) \end{cases} \qquad (11)$$

In case 3, when viewpoint 1 adopts the right global camera and viewpoint 2 adopts the left global camera, $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ can be solved respectively by (12).

$$\begin{cases} \boldsymbol{T}_g^{s_1} = \boldsymbol{T}_t^s \cdot \left(\boldsymbol{T}_r^{t_1} \cdot \boldsymbol{T}_g^r\right) \\ \boldsymbol{T}_g^{s_2} = \boldsymbol{T}_t^s \cdot \boldsymbol{T}_g^{t_2} \end{cases} \qquad (12)$$

In case 4, when both of viewpoint 1 and 2 adopt the right global camera, we can get $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ respectively by (13).

$$\begin{cases} \boldsymbol{T}_g^{s_1} = \boldsymbol{T}_t^s \cdot \left(\boldsymbol{T}_r^{t_1} \cdot \boldsymbol{T}_g^r\right) \\ \boldsymbol{T}_g^{s_2} = \boldsymbol{T}_t^s \cdot \left(\boldsymbol{T}_r^{t_2} \cdot \boldsymbol{T}_g^r\right) \end{cases} \qquad (13)$$

Through all the above process, we can solve $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ in all cases. Then, the main algorithm of the global calibration proposed in this paper can be realized.
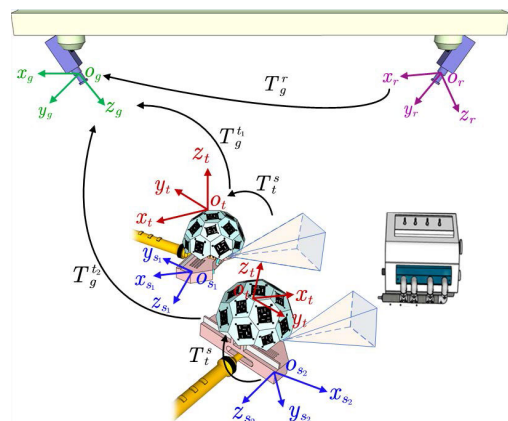

FIGURE 5. Measurement model of standard target for two viewpoints.

B. OPTIMIZATION ALGORITHM

This section proposes an optimization algorithm to optimize the main algorithm. We achieve the optimization by dealing with the feature points on the standard target according to the proposed measurement model.

The MBS collects the standard target images from viewpoint 1 and 2 respectively. And the GBVS collects the MDT images synchronously. $P_d^{b_1}$ and $P_d^{b_2}$ in two viewpoints are both unified to $o_g - x_g y_g z_g$ by using the MDT images as the intermediate medium. The measurement model of standard target for two viewpoints is shown in Fig. 5.

According to the condition of case 1 in Table 1, we combine (5), (6) and (7) to get (14). It can be seen from (14) that the value of $E_g^{b_1}$ can be reduced and the accuracy of global calibration can be improved by an optimal combination of T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$ when $P_d^{b_1}$ and $P_d^{b_2}$ remain unchanged. Therefore, we analyze the process of solving T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$ to find out the optimal combination of the three transformation matrixes.

$$E_g^{b_1} = \frac{1}{M} \sum_{d=1}^{M} \left\| P_d^{b_1} \cdot T_t^s \cdot T_g^{t_1} - P_d^{b_2} \cdot T_t^s \cdot T_g^{t_2} \right\|_2 \quad (14)$$

The MDT images are all needed in the process of solving T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$. Because of the multi-directional visibility of the MDT, the image of the MDT collected by the camera may contain multiple CSTs with different codes. And the normal plane vectors of the CSTs are different. Assume that the angle between the normal plane vector of the CST and the optical axis of the camera which collects the image is θ ($0^0 \le \theta \le 90^0$). Then, the closer θ is to 0, the higher accuracy of the extracted 2D feature points is. The larger θ is, the lower accuracy of the extracted 2D feature points is. The wrong feature points may be extracted when θ is almost close to 90^0. This will result in errors in the calculated transformation matrix and failing to provide data support for the subsequent 3D reconstruction of the measured object. According to the above analysis, we can see that there may be multiple CSTs in an MDT image and this will lead to the accuracy of the 2D feature points varied in a large range. Furthermore, it affects the solution of the transformation matrix from the MDT coordinate system to each camera coordinate system. Therefore, we can confirm the optimal combination of T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$ by processing the feature points in the MDT images.

First, we calculate each θ in all MDT images. Then, arrange θ from small to large in the corresponding image respectively to get θ^L. We get θ_1, θ_2 and θ_3 by setting the value in θ^L as the threshold in turn. θ_1, θ_2 and θ_3 correspond to the angle of the CST in the three MDT images which used to solve the initial values of T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$, respectively. The CSTs with angles greater than θ_1, θ_2 and θ_3 in the image are moved, and only the feature points on the remained CSTs are processed. The three MDT images for solving T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$ are defined as the first MDT image, the second MDT image and the third MDT image respectively. The number of remaining feature points in one MDT image should be guaranteed to be no less than 4. After that, the feature points with large distortion in the image can be eliminated by further processing the image. We define the various symbols that will be used later by the following process.

1) PROCESS THE FIRST MDT IMAGE

The remaining feature points in the first MDT image are processed when solving T_i^s. The feature point $\tilde{p}_i^m$ which P_i^m reprojected to the image captured by the right camera of the BVS are calculated according to (15). Wherein, the rotation matrix R_m^t and the translation vector t_m^t are solved by (2).

$$\tilde{p}_i^m = R_m^t \cdot P_i^m + t_m^t \quad (15)$$

We make $p_i^m = (x_i^m, y_i^m)$ and $\tilde{p}_i^m = (\tilde{x}_i^m, \tilde{y}_i^m)$. Then, the reprojection error E_i^m between p_i^m and $\tilde{p}_i^m$ can be solved by the (16).

$$E_i^m = \sqrt{\left(x_i^m - \tilde{x}_i^m\right)^2 + \left(y_i^m - \tilde{y}_i^m\right)^2} \quad (16)$$

$E_{i'}^{m1}(i' = \{1, 2, \ldots, N_1 | N_1 \in N^+\})$ can be got after arranging E_i^m from large to small. Then, calculate $P_{i'}^{m1}$ and $p_{i'}^{m1}$ corresponding to $E_{i'}^{m1}$. Where $p_{i'}^{m1}$ represents the 2D coordinates of the i'^{th} feature point in $o - uv$. $P_{i'}^{m1}$ represents the 3D coordinates corresponding to $p_{i'}^{m1}$ in $o_t - x_t y_t z_t$. E_1^{m1} is the largest value in $E_{i'}^{m1}$. The 3D and 2D feature point corresponding to E_1^{m1} is P_1^{m1} and p_1^{m1} respectively. Then, only the feature points corresponding to the small reprojection errors are left to recalculate the transformation matrix with higher precision by remove P_1^{m1} and p_1^{m1}.

Calculate $P_{i'-1}^{m1}$ and $p_{i'-1}^{m1}$ respectively by (17). $P_{i'-1}^{m1}$ represents the remaining 3D feature points after removing P_1^{m1} from $P_{i'}^{m1}$. $p_{i'-1}^{m1}$ represents the 2D feature points after removing p_1^{m1} from $p_{i'}^{m1}$.

$$\begin{cases} P_{i'-1}^{m1} = P_{i'}^{m1} - P_1^{m1} \\ p_{i'-1}^{m1} = p_{i'}^{m1} - p_1^{m1} \end{cases} \quad (17)$$

Recalculate the rotation matrix R_m^{t1}, translation vector t_m^{t1} and transformation matrix T_m^{t1} according to PNP principle by matching $P_{i'-1}^{m1}$ and $p_{i'-1}^{m1}$ one by one. Combine (7), (8) and (9) to get (18) to recalculate the transformation matrix T_{t1}^s from $o_s - x_s y_s z_s$ to $o_t - x_t y_t z_t$.

$$T_{t1}^s = \left(T_m^{t1} \cdot T_l^m \cdot \left(T_s^a \right)^{-1} \cdot T_l^a \right)^{-1} \quad (18)$$

Then, combine (14) and (18) to get (19) to recalculate the average distance $E_{g1}^{b_1}$ between $P_g^{b_1}$ and $P_g^{b_2}$.

$$E_{g1}^{b_1} = \frac{1}{M} \sum_{d=1}^{M} \left\| P_d^{b_1} \cdot T_{t1}^s \cdot T_g^{t_1} - P_d^{b_2} \cdot T_{t1}^s \cdot T_g^{t_2} \right\|_2 \quad (19)$$

Set up the threshold β. We adopt the T_{t1}^s, $T_g^{t_1}$ and $T_g^{t_2}$ as the optimal combination if $E_{g1}^{b_1} < \beta$. When $E_{g1}^{b_1} > \beta$, we begin to process the second MDT image to solve $T_g^{t_1}$.

2) PROCESS THE SECOND MDT IMAGE

At viewpoint 1, we extract the 2D coordinate of the feature points in the remaining CSTs in the second MDT image. N_2 represents the total number of feature points in this MDT image. $p_j^{l_1}(j = \{1, 2, \ldots, N_2 | N_2 \in N^+\})$ represents the jth feature point. Calculate the 3D coordinate $P_j^{l_1}$ in $o_t - x_t y_t z_t$

according to the codes in CST which $\boldsymbol{p}_j^{l_1}$ located. Calculate the rotation matrix $\boldsymbol{R}_g^{t_1}$, translator vector $\boldsymbol{t}_g^{t_1}$ and transformation matrix $\boldsymbol{T}_g^{t_1}$. Then, the 2D coordinate $\tilde{\boldsymbol{p}}_j^{l_1}$ which $\boldsymbol{P}_j^{l_1}$ reprojects to $o - uv$ can be solved by (20).

$$\tilde{\boldsymbol{p}}_j^{l_1} = \boldsymbol{R}_g^{t_1} \cdot \boldsymbol{P}_j^{l_1} + \boldsymbol{t}_g^{t_1} \tag{20}$$

Set up $\boldsymbol{p}_j^{l_1} = \left(x_j^{l_1}, y_j^{l_1}\right), \tilde{\boldsymbol{p}}_j^{l_1} = \left(\tilde{x}_j^{l_1}, \tilde{y}_j^{l_1}\right)$. The reprojection error $E_j^{l_1}$ between $\boldsymbol{p}_j^{l_1}$ and $\tilde{\boldsymbol{p}}_j^{l_1}$ can be realized by (21).

$$E_j^{l_1} = \sqrt{\left(x_j^{l_1} - \tilde{x}_j^{l_1}\right)^2 + \left(y_j^{l_1} - \tilde{y}_j^{l_1}\right)^2} \tag{21}$$

$\boldsymbol{E}_{j'}^{l_1 1}(j' = \{1, 2, \ldots, N_2 | N_2 \in N^+\})$ is got by arrange $E_j^{l_1}$ from large to small. Calculate $\boldsymbol{P}_{j'}^{l_1 1}$ and $\boldsymbol{p}_{j'}^{l_1 1}$ corresponding to $\boldsymbol{E}_{j'}^{l_1 1}$. $\boldsymbol{p}_{j'}^{l_1 1}$ represents the 2D coordinate of the j'^{th} feature point in $o - uv$ corresponding to $E_{j'}^{l_1 1}$. $\boldsymbol{P}_{j'}^{l_1 1}$ represents the 3D coordinate in $o_t - x_t y_t z_t$ corresponding to $\boldsymbol{p}_{j'}^{l_1 1}$. $E_1^{l_1 1}$ is the largest value in $\boldsymbol{E}_{j'}^{l_1 1}$. The 3D and 2D feature point corresponding to $E_1^{l_1 1}$ is $\boldsymbol{P}_1^{l_1 1}$ and $\boldsymbol{p}_1^{l_1 1}$ respectively. Only the remaining feature points with high accuracy can be used to recalculate the higher accuracy transformation matrix after removing $\boldsymbol{P}_1^{l_1 1}$ and $\boldsymbol{p}_1^{l_1 1}$.

Calculate $\boldsymbol{P}_{j'-1}^{l_1 1}$ and $\boldsymbol{p}_{j'-1}^{l_1 1}$ by (22). $\boldsymbol{P}_{j'-1}^{l_1 1}$ represents the 3D feature point after removing $\boldsymbol{P}_1^{l_1 1}$ from $\boldsymbol{P}_{j'-1}^{l_1 1}$. $\boldsymbol{P}_{j'-1}^{l_1 1}$ represents the corresponding 2D feature point after removing $\boldsymbol{p}_1^{l_1 1}$ from $\boldsymbol{p}_{j'}^{l_1 1}$.

$$\begin{cases} \boldsymbol{P}_{j'-1}^{l_1 1} = \boldsymbol{P}_{j'}^{l_1 1} - \boldsymbol{P}_1^{l_1 1} \\ \boldsymbol{p}_{j'-1}^{l_1 1} = \boldsymbol{p}_{j'}^{l_1 1} - \boldsymbol{p}_1^{l_1 1} \end{cases} \tag{22}$$

We recalculate the rotation matrix $\boldsymbol{R}_g^{t_1 1}$, translation vector $\boldsymbol{t}_g^{t_1 1}$ and transformation matrix $\boldsymbol{T}_g^{t_1 1}$ according to the PNP principle by matching $\boldsymbol{P}_{j'-1}^{l_1 1}$ and $\boldsymbol{p}_{j'-1}^{l_1 1}$ one by one. Then, the average distance $E_{g2}^{b_1}$ between $\boldsymbol{P}_g^{b_1}$ and $\boldsymbol{P}_g^{b_2}$ can be obtained by (23).

$$E_{g2}^{b_1} = \frac{1}{M} \sum_{d=1}^{M} \left\| \boldsymbol{P}_d^{b_1} \cdot \boldsymbol{T}_{t1}^s \cdot \boldsymbol{T}_g^{t_1 1} - \boldsymbol{P}_d^{b_2} \cdot \boldsymbol{T}_{t1}^s \cdot \boldsymbol{T}_g^{t_2} \right\|_2 \tag{23}$$

We adopt $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1 1}$ and $\boldsymbol{T}_g^{t_2}$ as the optimal combination if $E_{g2}^{b_1} < \beta$. We begin to process the third MDT image at viewpoint 2 when $E_{g2}^{b_1} > \beta$. The process of solving $\boldsymbol{T}_g^{t_2 1}$ is similar to the process of solving $\boldsymbol{T}_g^{t_1 1}$. At viewpoint 2, $\boldsymbol{p}_{k'-1}^{l_2 1}(k' = \{1, 2, \ldots, N_3 | N_3 \in N^+\})$ represents the 2D coordinate of the k'^{th} feature point in $o - uv$ after removing the feature points with the largest reprojection error. N_3 represents the total number of the feature points in the third image. Then, we can recalculate the average distance $E_{g3}^{b_1}$ between $\boldsymbol{P}_g^{b_1}$ and $\boldsymbol{P}_g^{b_2}$ with (24).

$$E_{g3}^{b_1} = \frac{1}{M} \sum_{d=1}^{M} \left\| \boldsymbol{P}_{b_1} \cdot \boldsymbol{T}_{t1}^s \cdot \boldsymbol{T}_g^{t_1 1} - \boldsymbol{P}_{b_2} \cdot \boldsymbol{T}_{t1}^s \cdot \boldsymbol{T}_g^{t_2 1} \right\|_2 \tag{24}$$

Set $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1} 1$ and $\boldsymbol{T}_g^{t_2} 1$ as the optimal combination when $E_{g3}^{b_1} < \beta$. If $E_{gn_4}^{b_1} < \beta$, the remaining feature points in the three MDT images are iterated in turn until the optimal combination of $\boldsymbol{T}_{t,n_1}^s, \boldsymbol{T}_g^{t_1,n_2}$ and $\boldsymbol{T}_g^{t_2,n_3}$ can be found. We take $E_{gn_4}^{b_1}$ to represent the average distance between $\boldsymbol{P}_g^{b_1}$ and $\boldsymbol{P}_g^{b_2}$. $\boldsymbol{T}_{t,n_1}^s, \boldsymbol{T}_g^{t_1,n_2}$ and $\boldsymbol{T}_g^{t_2,n_3}$ represent the transformation matrix after iterating the corresponding feature points respectively. n_1, n_2 and n_3 represent the number of loops needed to solve each transformation matrix respectively. n_4 represents the total number of loops needed to obtain the result. Their range of values is shown in (25).

$$\begin{cases} n_1 = \{1, 2, \ldots, N_1 \mid N_1 \in N^+\} \\ n_2 = \{1, 2, \ldots, N_2 \mid N_2 \in N^+\} \\ n_3 = \{1, 2, \ldots, N_3 \mid N_3 \in N^+\} \\ n_4 = \{1, 2, \ldots, (N_1 + N_2 + N_3 - 9)\} \end{cases} \tag{25}$$

To sum up, the process of finding the optimal combination of $\boldsymbol{T}_t^s, \boldsymbol{T}_g^{t_1}$ and $\boldsymbol{T}_g^{t_2}$ mainly consists of the following steps:

① Select θ_1, θ_2 and θ_3 in turn from $\boldsymbol{\theta}^L$ to remove the CSTs in the corresponding image which θ is greater than θ_1, θ_2 and θ_3. Then, we can get the feature points with higher accuracy.

② Calculate $\boldsymbol{T}_{t1}^s$ by the remaining feature points to find the combination of $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1}$ and $\boldsymbol{T}_g^{t_2}$. If $E_{g,n_4}^{b_1} < \beta$, we stop the iteration, otherwise, go to the step ③.

③ At viewpoint 1, we calculate $\boldsymbol{T}_g^{t_1 1}$ to look for the combination of $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1 1}$ and $\boldsymbol{T}_g^{t_2}$. If $E_{g,n_4}^{b_1} < \beta$, we stop the iteration, otherwise, go to the step ④.

④ At viewpoint 2, we calculate $\boldsymbol{T}_g^{t_2 1}$ to find the combination of $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1 1}$ and $\boldsymbol{T}_g^{t_2 1}$. If $E_{g,n_4}^{b_1} < \beta$, we stop the iteration, otherwise, go to the step ⑤.

⑤ At viewpoint 2, process the remaining feature point $\boldsymbol{p}_{k'-1}^{l_2 1}$ in the third MDT image. Similar to the process of solving $\boldsymbol{T}_g^{t_2 1}$, solve $\boldsymbol{T}_g^{t_2,n_3}$ in turn until finding the combination of $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1 1}$ and $\boldsymbol{T}_g^{t_2,n_3}$ to make $E_{g,n_4}^{b_1} < \beta$. Then we can stop the iteration and start to process the remaining feature point $\boldsymbol{p}_{i'-1}^{l_1 1}$ in the second MDT image to calculate $\boldsymbol{T}_g^{t_1 1}$. If $E_{g,n_4}^{b_1} > \beta$ when the number of remaining feature points $N\left(\boldsymbol{p}_{k'-n_3}^{l_2,n_3}\right) < 4$, we need to go to step ⑥.

⑥ At viewpoint 1, adopt the method that is similar to the process of solving $\boldsymbol{T}_g^{t_1 1}$ to solve $\boldsymbol{T}_g^{t_1,n_2}$ in turn. Repeat step ⑤ every time to calculate $\boldsymbol{T}_g^{t_1,n_2}$ until finding the combination of $\boldsymbol{T}_{t1}^s, \boldsymbol{T}_g^{t_1,n_2}$ and $\boldsymbol{T}_g^{t_2,n_3}$ to make $E_{g,n_4}^{b_1} < \beta$. Then the iteration can be stopped. We start to process the remaining feature points $\boldsymbol{p}_{i'-1}^{l_1 1}$ in the second MDT image to calculate $\boldsymbol{T}_{t1}^s$. If $E_{g,n_4}^{b_1} > \beta$ when the number of remaining feature points $N\left(\boldsymbol{p}_{j'-n_2}^{l_1,n_2}\right) < 4$, go to step ⑦.

⑦ Calculate $\boldsymbol{T}_{t,n_1}^s$ in turn with the similar process of solving $\boldsymbol{T}_{t1}^s$. Repeat the step ③~⑥ in turn every time to calculate $\boldsymbol{T}_{t,n_1}^s$ until confirming the combination of

T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$ to make $E_{g,n_4}^{b_1} < \beta$. Then the iteration can be stopped. if $E_{g,n_4}^{b_1} > \beta$ when the total number of the remaining feature points $N\left(p_{i'-n_1}^{m,n_1}\right) < 4$, go to step ⑧.

⑧ Arrange all $E_{g,n_4}^{b_1}$ from small to large to calculate E. Find the combination of T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$ that corresponding to first value E^1 in E. This combination is the optimal combination.

Algorithm 1 Global Calibration Algorithm

Input 3D coordinates $\hat{P}_i^m$, $\hat{P}_j^{l_1}$, $\hat{P}_k^{l_2}$ and the corresponding 2D coordinates
$\hat{p}_i^m$, $\hat{p}_j^{l_1}$, $\hat{p}_k^{l_2}$ of the feature points in MDT image

Output The optimal combination of T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$

1. T = Initation(P) // calculate initial matrix
2. $E_g^{b_1} \leftarrow T_t^s, T_g^{t_1}, T_g^{t_2}$ // begin optimizing if $E_g^{b_1} \leftarrow \beta$
3. θ = CalculateAngle(P) // calculate angle
4. $P_i^m, p_i^m, P_j^{l_1}, p_j^{l_1}, P_k^{l_2}, p_k^{l_2} \leftarrow \theta_1, \theta_2, \theta_3$ // calculate remain feature points
5. **for** $i = 0$ to size (n_1) **do**
6. $T_{t,n_1}^s \leftarrow \{P_i^m, p_i^m\}$ // calculate transformation matrix
7.   **for** $j = 0$ to size (n_2) **do**
8. $T_g^{t_1,n_2} \leftarrow \{P_j^{l_1}, p_j^{l_1}\}$
9.   **for** $k = 0$ *to* size (n_3) **do**
10. $T_g^{t_2,n_3} \leftarrow \{P_k^{l_2}, p_k^{l_2}\}$
11.   **if** $E_{g,n_4}^{b_1} < \beta$
12.   **output** $\left\{T_{t,n_1}^s, T_g^{t_1,n_2}, T_g^{t_2,n_3}\right\}$
13.   **break**
14.   **if** $N\left(p_{k'-n_3}^{l_2,n_3}\right) < 4$ // calculate $T_g^{t_2,n_3}$
15.   **break**
16.   **if** $N\left(p_{j'-n_2}^{l_1,n_2}\right) < 4$ // calculate $T_g^{t_1,n_2}$ with remaining feature points
17.   **break**
18.   **if** $N\left(p_{i'-n_1}^{m,n_1}\right) < 4$ // calculate T_{t,n_1}^s with remaining feature points
19.   **Break**
20. $E \leftarrow E_{g,n_4}^{b_1}$ // arrange $E_{g,n_4}^{b_1}$ from small to large
21. $\left\{T_{t,n_1}^s, T_g^{t_1,n_2}, T_g^{t_2,n_3}\right\}$ // find optimal combination corresponding to E^1
22. **return** $\left\{T_{t,n_1}^s, T_g^{t_1,n_2}, T_g^{t_2,n_3}\right\}$

The optimal $T_g^{s_n}$ in the previous and next viewpoint can be solved after finding the optimal combination of T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$. Then, the function of global calibration for case 1 in Table 1 can be realized. In the same way, we can calculate the optimal $T_g^{s_n}$ in the corresponding cases through the principle of calculating T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$ when the usage of the GBVS belongs to other cases in Table 1.

After all the above process, the global calibration for the designed 3D VMS can be realized by the proposed main algorithm and optimization algorithm. We use the pseudo code to describe the proposed global calibration approach for the designed 3D VMS more intuitively. The pseudo code of the proposed global calibration algorithm is shown in

Algorithm 1, where $\hat{P}_i^m$, $\hat{P}_j^{l_1}$, $\hat{P}_k^{l_2}$ and $\hat{p}_i^m$, $\hat{p}_j^{l_1}$, $\hat{p}_k^{l_2}$ represent the 3D and 2D coordinates respectively in the MDT images to calculate T_t^s, $T_g^{t_1}$ and $T_g^{t_2}$.

IV. EXPERIMENT

In this section, the proposed main algorithm and optimization algorithm of global calibration are verified by real experiments. This section is divided into two parts with experimental device introduction and algorithm verification. The experimental device introduction part mainly introduces the designed 3D VMS and the corresponding parameters. In the algorithm verification part, the feature points in the images are processed to solve the optimal combination of T_{t,n_1}^s, $T_g^{t_1,n_2}$ and $T_g^{t_2,n_3}$ corresponding the four cases in Table 1 to get the high accuracy of $T_g^{s_n}$ in the different viewpoints.

A. EXPERIMENTAL DEVICE INTRODUCTION

The designed 3D VMS is shown in Fig. 6. The intermediate plane target and standard target we select are both the same chessboard target. Its material is glass. The effective area of the chessboard target is 70 mm × 80 mm, and there is 7 × 8 feature points on the effective area. The cell length is 9 mm, and the processing accuracy is 0.002 mm. The polyhedron skeleton of the MDT is made by 3D print and the CSTs are made by film. The processing accuracy of CST is 0.002 mm, the side length is 48 mm. The interior of the CST is coded according to Tag36h10 in AprilTags.
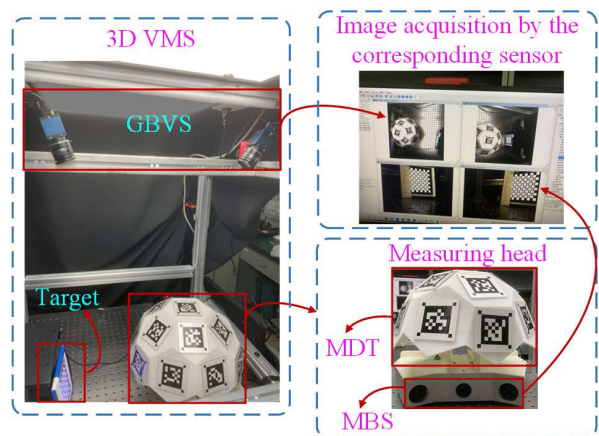

FIGURE 6. Schematic diagram of the experimental device.

The models of the left global camera and right global camera of the GBVS are all ME2P-2621-15U3M, and the image resolution is 5120 pixels × 5120 pixels. The models of the cameras in the MBS are ME2P-1230-23U3M, and the image resolution is 4096 pixels × 3000 pixels. The lens models of the cameras are all V1228-MPY. The cameras and lens are all made by Daheng. The parameters and distortion coefficients of the GBVS and the MBS are shown in Table 2 and Table 3 respectively, where r_r^l and t_r^l represent the rotation vector and translation vector from the left camera to the right camera

TABLE 2. Internal parameters and coefficients of the cameras.

Internal Parameter ($pixel$)		f_x	f_y	u_0	v_0	k_1	k_2	p_1	p_2	k_3
GBVS	Left	5011.82	5001.83	2576.21	2593.05	-0.11	0.12	-9.94e-05	2.59e-05	0.33
	Right	4973.87	4963.61	2482.01	2614.41	-0.15	0.39	-6.33e-04	6.39e-04	-0.43
MBS	Left	4877.39	4877.64	2103.75	1529.26	-0.06	0.28	1.63e-04	7.61e-04	-0.21
	Right	4921.93	4921.62	2058.25	1572.30	-0.05	0.36	4.46e-04	1.09e-03	-0.52

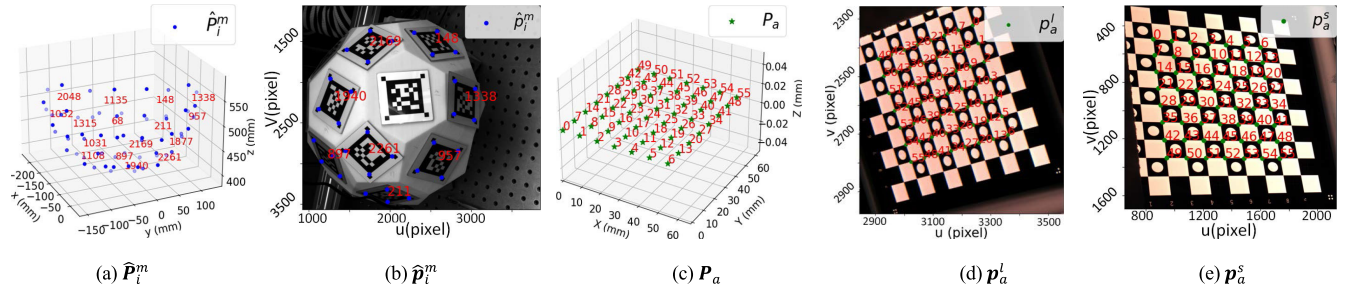(a) $\hat{P}_i^m$ (b) $\hat{p}_i^m$ (c) P_a (d) p_a^l (e) p_a^s

FIGURE 7. Feature points in the images to calculate T_t^s.

respectively. We adopt the same set of cameras to form the GBVS and BVS.

B. DALGORITHM VERIFICATION
We complete the verification of the main algorithm and the optimization algorithm of the global calibration in this section.

1) VERIFY THE MAIN ALGORITHM
We calculate T_t^s to verify the main algorithm. First, we should collect the images. The process of collecting images is as follows: Place the measuring head and intermediate plane target on the optical platform. The intermediate plane target is about 300 mm from the front of the MBS in the measuring head. The BVS is placed above the intermediate plane target and measuring head and it is fixed on the aluminum frame on the optical platform. We keep the relative pose of the BVS, measuring head and intermediate plane target unchanged. The right camera of the BVS can collect the image of the MDT at a distance about 0.55 m. The left camera of the BVS synchronously collects the image of the intermediate plane target at a distance about 1m. At the same time, the left camera of the MBS collects the intermediate plane target image. Process images in turn and extract the corresponding 2D feature points $\hat{p}_i^m$, p_a^l and p_a^s, as shown in (b), (d) and (e) of Fig. 7. Unify the 3D feature points on the intermediate plane target to $o_a - x_a y_a z_a$ to solve P_a, as shown in Fig. 7. (c). According to the algorithm proposed by the reference [23], the 3D feature points on the MDT are unified to $o_t - x_t y_t z_t$ to solve $\hat{P}_i^m$, as shown in Fig. 7. (a).

Calculate T_m^t, T_l^a and T_s^a in turn according to PNP principle by matching $\hat{p}_i^m$ and $\hat{P}_i^m$, p_a^l and P_a, p_a^s and P_a one by one. The transformation matrix T_m^l from $o_l - x_l y_l z_l$ to $o_m - x_m y_m z_m$ can be solved by the values in Table 3. Then T_t^s can be solved by simultaneous (7), (8) and (9).

TABLE 3. External parameters of the cameras.

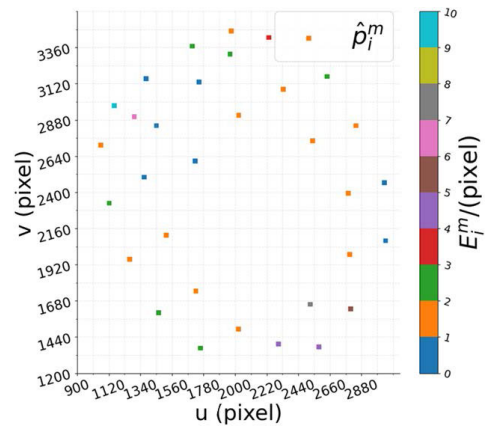| External Parameter (mm)		x	y	z
GBVS	r_r^l	-0.10	-0.91	-0.48
	t_r^l	369.76	-106.72	122.26
MBS	r_r^l	-0.01	0.87	-7.92e-04
	t_r^l	-194.74	0.81	89.64

FIGURE 8. The corresponding distribution of $\hat{p}_i^m$ and E_i^m.

Calculate E_i^m according (15) and (16). The corresponding distribution of $\hat{p}_i^m$ and E_i^m is shown as Fig. 8. We can see that the maximum value in E_i^m is greater than 10 $pixels$, which is too large and seriously reduces the accuracy of global calibration. Therefore, it is necessary to eliminate the feature points corresponding to the error which is relatively larger in the subsequent optimization algorithm.

We should solve $T_g^{s_n}$ by processing the images collected by the GBVS and MBS at each viewpoint after solving T_t^s. Then, we need to collect the MDT images by the designed 3D VMS. The process of collecting MDT images is as follows:

Place the measuring head and standard target on the optical platform. The standard target is placed about 300 mm from the front of the MBS which is in the measuring head. The GBVS is placed above the standard target and measuring head and it is fixed on the aluminum frame of the optical platform. The working distance of the left global camera and right global camera of the GBVS are both 0.55 m~1 m. MBS collects the standard target image at viewpoint 1, and the GBVS synchronously collects the MDT image. We move the measuring head to viewpoint 2 under the condition of keeping the pose of the GBVS and the standard target unchanged. Then, the MBS collects the standard target image again, and the GBVS collects the MDT image synchronously. After all the above processes, we can obtain MDT images at two viewpoints. Next, we begin to process the feature points in the MDT images to verify the proposed main algorithm of the global calibration approach.

At viewpoint 1, we can extract the feature points p_1^l and p_1^r in the standard target images collected by the left camera and right camera of the MBS, as shown in Fig. 9 (a) and Fig. 9 (b) respectively. At view point 2, we can extract the feature points p_2^l and p_2^r in the standard target image collected by the left camera and right camera of the MBS, as shown in Fig. 9 (c) and Fig. 9 (d) respectively.
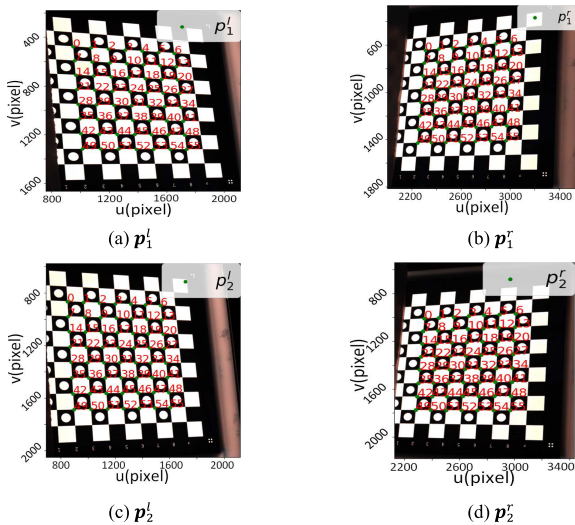

FIGURE 10. $P_d^{b_1}$ and $P_d^{b_2}$ in viewpoint 1 and 2 respectively.

FIGURE 11. $\hat{p}_j^{l_1}$ and $\hat{p}_j^{l_2}$ of two viewpoints in case 1 in Table 1.

FIGURE 12. Corresponding distribution between $P_d^{g_n}$ and E_d.

FIGURE 9. p_1^l, p_1^r, p_2^l, p_2^r in viewpoint 1 and 2 respectively.

Then, we can get $P_d^{b_1}$ and $P_d^{b_2}$ with the relationship of p_1^l and p_1^r, p_2^l and p_2^r based on the triangulation principle of binocular stereo vision. As shown in Fig. 10, where (a) and (b) represent $P_d^{b_1}$ and $P_d^{b_2}$ respectively. We can see from Fig. 10 that the MBS can collect corresponding 3D point clouds information at viewpoint 1 and viewpoint 2 respectively. We can verify the effectiveness of the main algorithm in the global calibration approach with these 3D point clouds.

Next, we should process the MDT images captured by the GBVS. First, for the case 1 in Table 1, the 2D feature points $\hat{p}_j^{l_1}$ and $\hat{p}_j^{l_2}$ in the MDT images at two viewpoints are extracted, as shown in Fig. 11, where (a) and (b) respectively
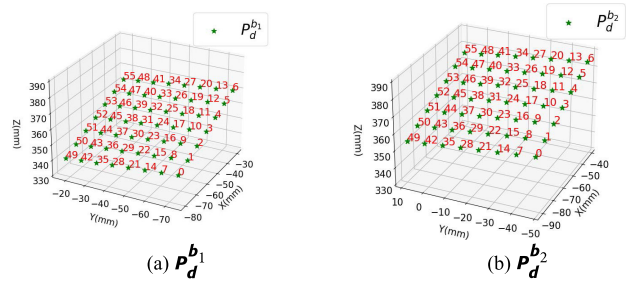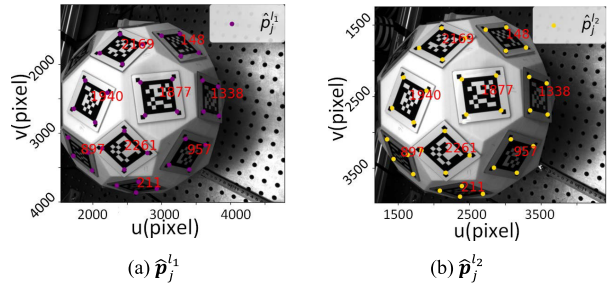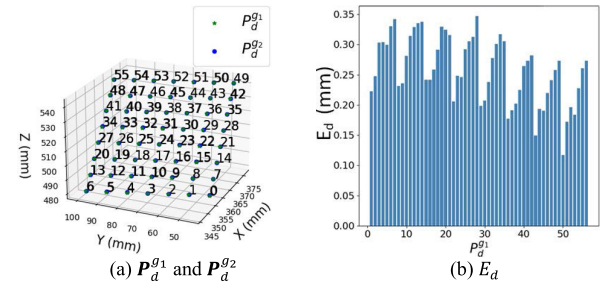
represent $\hat{p}_j^{l_1}$ and $\hat{p}_j^{l_2}$. Calculate $T_g^{t_1}$ and $T_g^{t_2}$ in turn according to PNP principle by matching $\hat{p}_j^{l_1}$ and $\hat{P}_i^m$, $\hat{p}_j^{l_2}$ and $\hat{P}_i^m$ one by one. Then $T_g^{s_1}$ and $T_g^{s_2}$ can be calculated by (10).

$P_d^{b_1}$, $P_d^{b_2}$, $T_g^{s_1}$ and $T_g^{s_2}$ can be calculated through all the above process. And then $P_d^{g_1}$ and $P_d^{g_2}$ can be calculated by (5), as shown in Fig. 12 (a). We take E_d as the distance error of the corresponding feature points after subtracting the corresponding point $P_d^{g_1}$ and $P_d^{g_2}$. The corresponding distribution between $P_d^{g_1}$ and E_d is shown in Fig. 12 (b). The value of $E_g^{b_1}$ is 0.26mm calculate by (14). In the same way, the images under case 2, 3 and 4 in Table 1 are processed in turn and the calculated $E_g^{b_1}$ is 1.45mm, 1.25mm and 0.36mm respectively. The results show that the proposed method can solve T_t^s, $T_g^{s_1}$ and $T_g^{s_2}$. This can provide data support for unifying the feature points of the measured objects under different viewpoints to $o_g - x_g y_g z_g$. It proves the feasibility of the proposed main algorithm.

2) VERIFY THE OPTIMIZATION ALGORITHM
According to the principle of obtaining the corresponding distribution of $\hat{p}_i^m$ and E_i^m, we calculate the corresponding
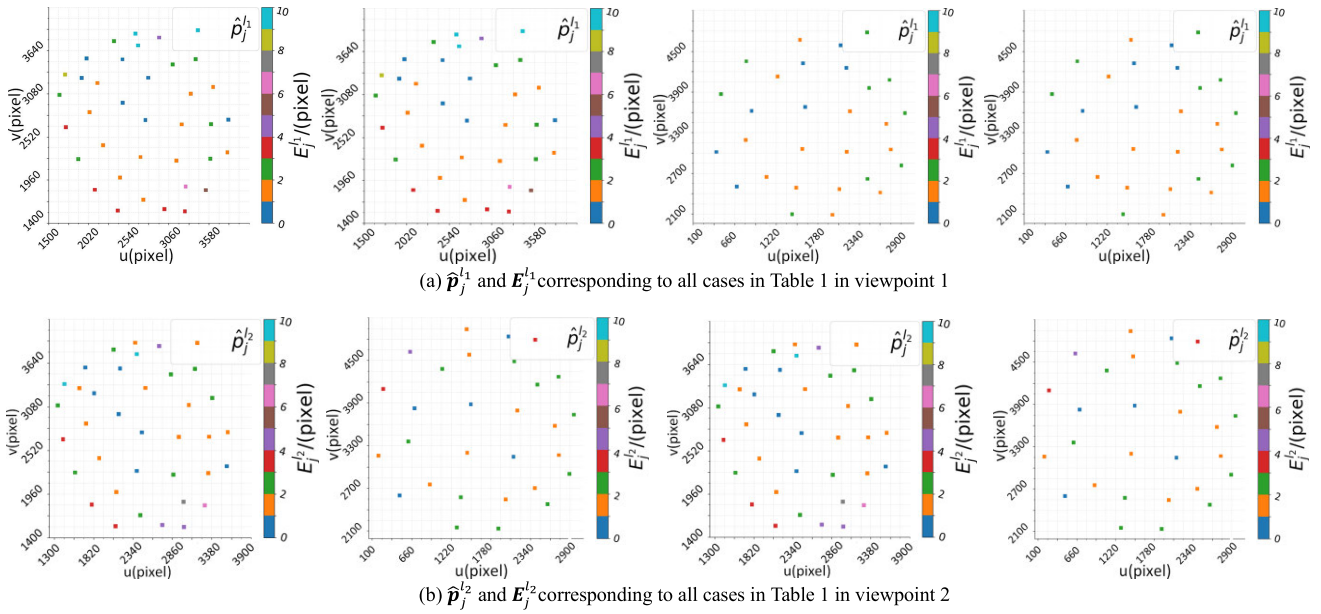
(a) $\hat{\boldsymbol{p}}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$ corresponding to all cases in Table 1 in viewpoint 1

(b) $\hat{\boldsymbol{p}}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$ corresponding to all cases in Table 1 in viewpoint 2

FIGURE 13. $\hat{\boldsymbol{p}}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$, $\hat{\boldsymbol{p}}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$ corresponding to all cases in Table 1 in two viewpoints.

TABLE 4. θ_1, θ_2, θ_3 and $E_{g,n_4}^{b_1}$ after optimization in all cases.

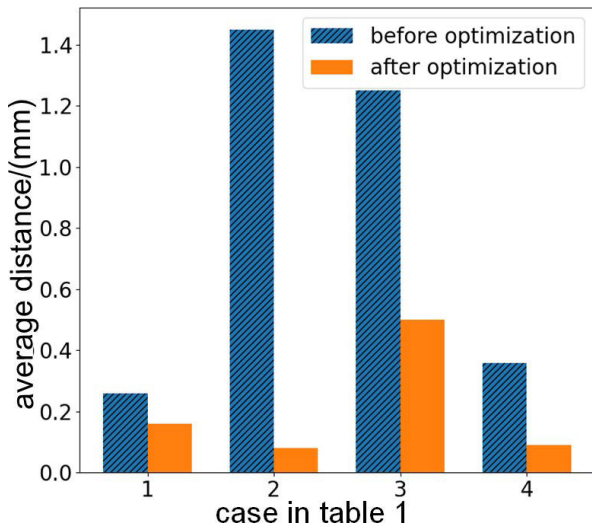| Case	Threshold			Average Distance $E_{g,n_4}^{b_1}$/mm
	θ_1	θ_2	θ_3	
1	34.00°	40.00°	40.15°	0.19
2	34.00°	41.00°	52.00°	0.08
3	34.00°	41.00°	41.00°	0.50
4	34.00°	32.00°	3200°	0.09

FIGURE 14. Comparison of average distance before and after the optimization algorithm.

distribution of $\hat{p}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$, $\hat{p}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$, as shown in Fig. 13, where (a) and (b) respectively represent the corresponding distribution relationship between feature points and

reprojection errors under the corresponding conditions in viewpoint 1 and 2. We can see that both $\boldsymbol{E}_j^{l_1}$ and $\boldsymbol{E}_j^{l_2}$ exceed $10\,pixels$, which will result in the low accuracy of $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$. This will further reduce the accuracy of global calibration. Therefore, in the optimization algorithm proposed in this section, the accuracy of the global calibration algorithm is improved by successively removing the feature points corresponding to the excessive reprojection errors in turn in the MDT images.

Set $\beta = 0.1$ mm, and then, run the optimization algorithm proposed in Algorithm 1, the calculated θ_1, θ_2, θ_3 and $E_{g,n_4}^{b_1}$ are shown in Table 4. And the values of the average absolute error calculated under the four cases are shown in Fig. 14. It can be seen from Fig. 14 and Table 4 that the accuracy in all cases in Table 1 has been improved significantly after optimizing. It means that the optimization effectiveness is significant. Moreover, according to the optimization algorithm, the accuracy of the designed 3D VMS can reach up to 0.08 mm when the CST corresponding to the angle threshold in the MDT image and the feature points in the MDT image with large reprojection errors are all eliminated. This can prove that the measurement accuracy of our designed 3D VMS is high enough to measure most of the object with large size or complex structure by the proposed global calibration approach in this paper.

After optimization, the corresponding distribution of $\boldsymbol{p}_i^m$ and $\boldsymbol{E}_i^m$, $\boldsymbol{p}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$, $\boldsymbol{p}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$ is shown in Fig. 15. It can be seen from Fig. 15 that the minimum reprojection error calculated by the remaining feature points in the MDT image can be less than 0.03 pixel after optimization. So, the calibration accuracy of $\boldsymbol{T}_{t,n_1}^s$, $\boldsymbol{T}_g^{t_1,n_2}$ and $\boldsymbol{T}_g^{t_2,n_3}$ is significantly improves because of the greatly reduced of the reprojection error. This further proves the feasibility of our proposed optimization
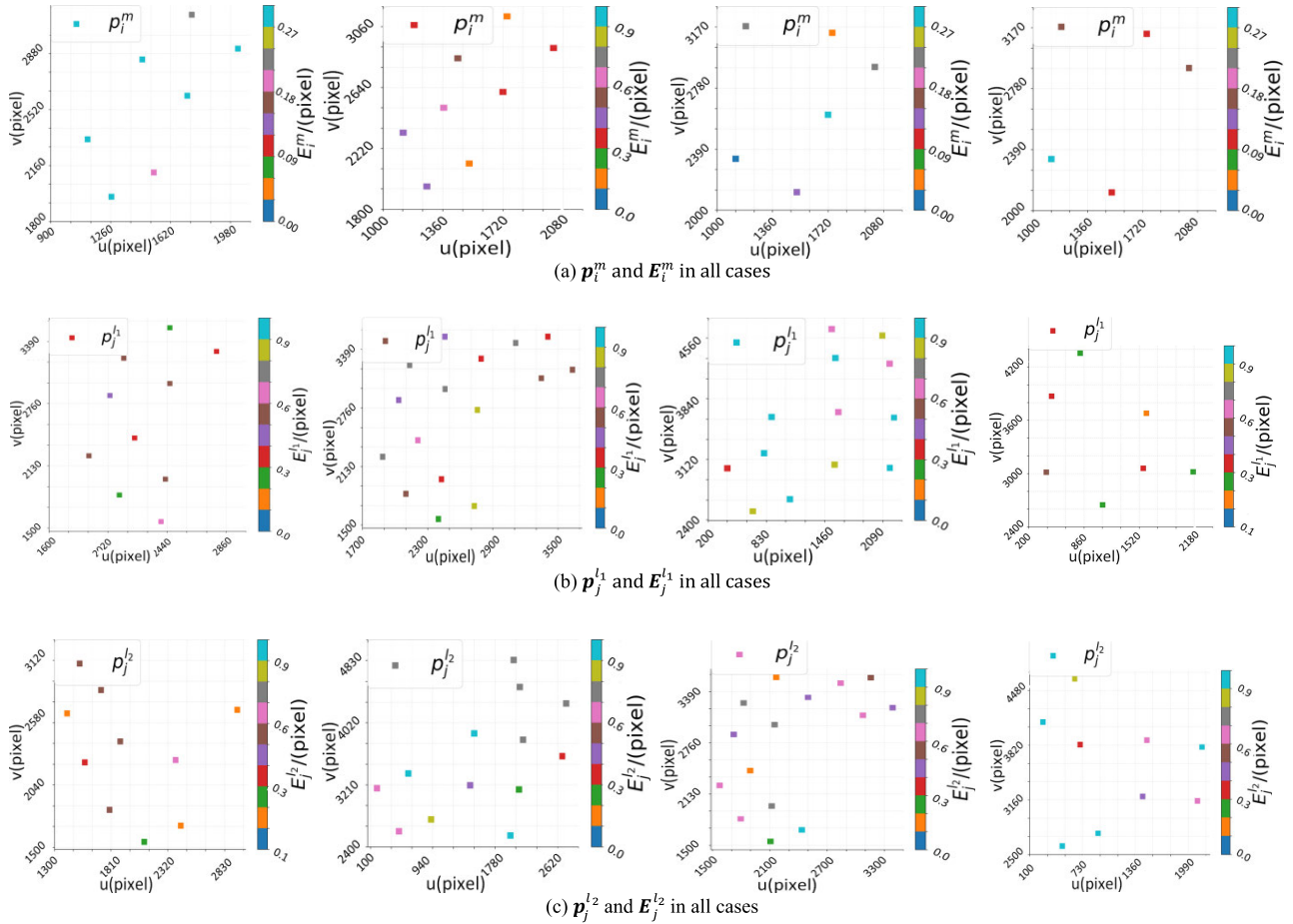
(a) $\boldsymbol{p}_i^m$ and $\boldsymbol{E}_i^m$ in all cases

(b) $\boldsymbol{p}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$ in all cases

(c) $\boldsymbol{p}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$ in all cases

FIGURE 15. Distribution of $\boldsymbol{p}_i^m$ and $\boldsymbol{E}_i^m$, $\boldsymbol{p}_j^{l_1}$ and $\boldsymbol{E}_j^{l_1}$, $\boldsymbol{p}_j^{l_2}$ and $\boldsymbol{E}_j^{l_2}$ in all cases after optimization.

(a) case 1   (b) case 2   (c) case 3   (d) case 4
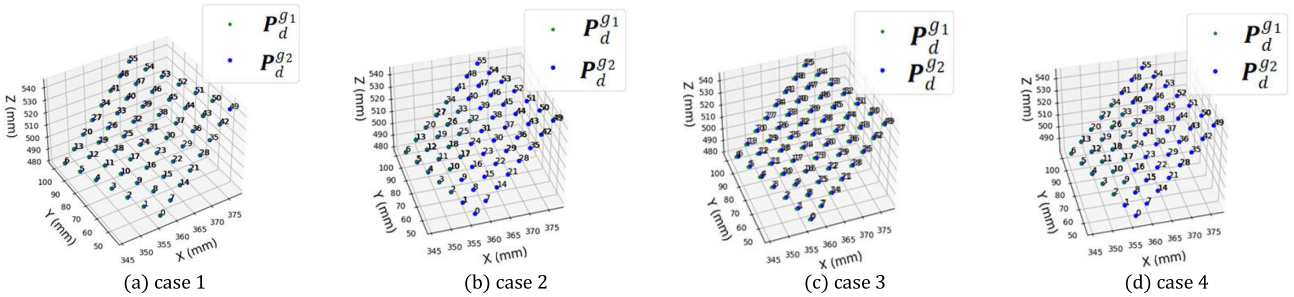
FIGURE 16. $\boldsymbol{P}_d^{g_1}$ and $\boldsymbol{P}_d^{g_2}$ in all cases in Table 1.

algorithm. Then, $\boldsymbol{P}_d^{b_1}$ and $\boldsymbol{P}_d^{b_2}$ in all cases in Table 1 can be unified into $o_g - x_g y_g z_g$ to get $\boldsymbol{P}_d^{g_1}$ and $\boldsymbol{P}_d^{g_2}$ by the optimized $\boldsymbol{T}_g^{s_1}$ and $\boldsymbol{T}_g^{s_2}$ with (5). $\boldsymbol{P}_d^{g_1}$ and $\boldsymbol{P}_d^{g_2}$ in all cases in Table 1 is shown as Fig. 16. We can see from Fig. 16 that the 3D point clouds in different viewpoints can be unified into the global coordinate system with the calculated transformation matrix between MBS and GBVS.

To sum up, we verify the proposed global calibration approach for the designed 3D VMS with real experiment in this Section. And we calculate the optimal combination of

$\boldsymbol{T}_{t,n_1}^s$, $\boldsymbol{T}_g^{t_1,n_2}$ and $\boldsymbol{T}_g^{t_2,n_3}$ by using the remaining high accuracy feature points in the MDT images after removing some invalid CSTs. The calculated global calibration result can provide data support for the subsequent global unification of the local 3D point clouds collected from different viewpoints for the measured object.

V. CONCLUSION
In this paper, a 3D VMS is designed. This system adopts the MDT as intermediate medium to unify the local point

clouds data of the measured object to the global coordinate system. We also proposed a global calibration approach aiming at the designed 3D VMS. Compared with the traditional global calibration for the 3D VMS, this approach has two advantages: 1. We can use only one global camera to achieve the function of global calibration. Through redundant design philosophy, we adopt two global cameras without a public FOV to form a GBVS to expand the measurement range of the designed 3D VMS. 2. We proposed an optimization algorithm to improve the accuracy of the global calibration. It can be seen from the real experiment that the accuracy of the proposed algorithm can reach up to 0.08 mm. This strongly proves the effectiveness of the proposed global calibration approach.

However, the algorithm proposed in this paper heavily relies on the accuracy of the feature points on the MDT images. So, how to extract the feature points with high accuracy is the focus of the follow-up research process.

REFERENCES

[1] F. Miao, J. Kang, Z. Sun, L. Feng, C. Peng, and B. Wu, "Calibration and measurement method based on optical lenses for large-scale 3-D precise measurement," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[2] Y. Gu, J. Lv, J. Bo, B. Zhao, Y. Chen, J. Tao, Y. Qin, W. Wang, and J. Liang, "Joint dense 3D reconstruction method for endoscopic images of weak texture scenes," *IEEE Access*, vol. 9, pp. 138254–138266, 2021.

[3] C. Sui, K. He, C. Lyu, and Y. Liu, "Accurate 3D reconstruction of dynamic objects by spatial–temporal multiplexing and motion-induced error elimination," *IEEE Trans. Image Process.*, vol. 31, pp. 2106–2121, 2022.

[4] J. Gao, F. Li, C. Zhang, W. He, J. He, and X. Chen, "A method of D-type weld seam extraction based on point clouds," *IEEE Access*, vol. 9, pp. 65401–65410, 2021.

[5] W. Xia, G. Ameri, D. Fakim, H. Akhuanzada, M. Z. Raza, S. A. Shobeiri, L. McLean, and E. C. S. Chen, "Automatic plane of minimal hiatal dimensions extraction from 3D female pelvic floor ultrasound," *IEEE Trans. Med. Imag.*, vol. 41, no. 12, pp. 3873–3883, Dec. 2022.

[6] E. Liu, X. Cheng, X. Cheng, T. Zhou, and Y. Huang, "Application of three-dimensional laser scanning in the protection of multi-dynasty ceramic fragments," *IEEE Access*, vol. 8, pp. 139771–139780, 2020.

[7] S. Chen, B. Duinkharjav, X. Sun, L. Wei, S. Petrangeli, J. Echevarria, C. Silva, and Q. Sun, "Instant reality: Gaze-contingent perceptual optimization for 3D virtual reality streaming," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 5, pp. 2157–2167, May 2022.

[8] X. Liu, H. Madhusudanan, W. Chen, D. Li, J. Ge, C. Ru, and Y. Sun, "Fast eye-in-hand 3-D scanner-robot calibration for low stitching errors," *IEEE Trans. Ind. Electron.*, vol. 68, no. 9, pp. 8422–8432, Sep. 2021.

[9] X. Xiang, H. Jiang, G. Zhang, Y. Yu, C. Li, X. Yang, D. Chen, and H. Bao, "Mobile3DScanner: An online 3D scanner for high-quality object reconstruction with a mobile device," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 11, pp. 4245–4255, Nov. 2021.

[10] J. An and E. Kim, "Novel vehicle bounding box tracking using a low-end 3D laser scanner," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3403–3419, Jun. 2021.

[11] Y. Lu, W. Liu, Y. Zhang, H. Xing, J. Li, S. Liu, and L. Zhang, "An accurate calibration method of large-scale reference system," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6957–6967, Sep. 2020.

[12] L. Han, S. Gu, D. Zhong, S. Quan, and L. Fang, "Real-time globally consistent dense 3D reconstruction with online texturing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 3, pp. 1519–1533, Mar. 2022.

[13] Y. Cai, M. Cao, L. Li, and X. Liu, "An end-to-end approach to reconstructing 3D model from image set," *IEEE Access*, vol. 8, pp. 193268–193284, 2020.

[14] M. Babu, P. Franciosa, and D. Ceglarek, "Spatio-temporal adaptive sampling for effective coverage measurement planning during quality inspection of free form surfaces using robotic 3D optical scanner," *J. Manuf. Syst.*, vol. 53, pp. 93–108, Oct. 2019.

[15] I. Goda, G. L'Hostis, and P. Guerlain, "In-situ non-contact 3D optical deformation measurement of large capacity composite tank based on close-range photogrammetry," *Opt. Lasers Eng.*, vol. 119, pp. 37–55, Aug. 2019.

[16] Ş. Ozan and Ş. Gümüştekin, "Calibration of double stripe 3D laser scanner systems using planarity and orthogonality constraints," *Digit. Signal Process.*, vol. 24, pp. 231–243, Jan. 2014.

[17] H. Shang, C. Liu, and R. Wang, "Measurement methods of 3D shape of large-scale complex surfaces based on computer vision: A review," *Measurement*, vol. 197, Jun. 2022, Art. no. 111302.

[18] S. Yin, Y. Ren, Y. Guo, J. Zhu, S. Yang, and S. Ye, "Development and calibration of an integrated 3D scanning system for high-accuracy large-scale metrology," *Measurement*, vol. 54, pp. 65–76, Aug. 2014.

[19] E. Cuesta, B. J. Alvarez, S. Martinez-Pellitero, J. Barreiro, and H. Patiño, "Metrological evaluation of laser scanner integrated with measuring arm using optical feature-based gauge," *Opt. Lasers Eng.*, vol. 121, pp. 120–132, Oct. 2019.

[20] S. Barone, A. Paoli, and A. V. Razionale, "Shape measurement by a multi-view methodology based on the remote tracking of a 3D optical scanner," *Opt. Lasers Eng.*, vol. 50, no. 3, pp. 380–390, Mar. 2012.

[21] J. Shi and Z. Sun, "Large-scale three-dimensional measurement based on LED marker tracking," *Vis. Comput.*, vol. 32, no. 2, pp. 179–190, Feb. 2015.

[22] Z. Zhou, W. Liu, Y. Wang, B. Yu, X. Cheng, Y. Yue, and J. Zhang, "A combined calibration method of a mobile robotic measurement system for large-sized components," *Measurement*, vol. 189, Feb. 2022, Art. no. 110543.

[23] P. Sun, F. Zhou, J. Sun, and L. Wang, "High-accuracy three-dimensional measurement based on multi-directional cooperative target with weighted SfM algorithm," *Measurement*, vol. 172, Feb. 2021, Art. no. 108955.

[24] B. Zhou, Z. Chen, and Q. Liu, "An efficient solution to the perspective-n-point problem for camera with unknown focal length," *IEEE Access*, vol. 8, pp. 162838–162846, 2020.

[25] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 666–673.

PENGFEI SUN received the B.S. and M.S. degrees in measuring, testing technologies, and instruments from Henan Polytechnic University, Henan, China, in 2012 and 2015, respectively. She is currently pursuing the Ph.D. degree with the School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing, China. Her current research interests include machine vision and precision measurement.

HAISHU TAN (Member, IEEE) received the B.S. and Ph.D. degrees in optical engineering from Tianjin University, China, in 1994 and 1998, respectively. He is currently a Professor with the School of Physics and Optoelectronic Engineering, Foshan University, China. His research interests include computer vision, photoelectric measurement, and optical metrology.

FUQIANG ZHOU received the B.S., M.S., and Ph.D. degrees in instrument, measurement, and test technology from Tianjin University, China, in 1994, 1997, and 2000, respectively. In 2000, he joined the School of Automation Science and Electrical Engineering, Beihang University, China, as a Postdoctoral Research Fellow. He is currently a Professor with the School of Instrumentation Science and Opto-Electronics Engineering, Beihang University. His research interests include precision vision measurement, 3-D vision sensors, image recognition, and optical metrology.

● ● ●