



HAL
open science

Dynamic 3D Scanning Based on Optical Tracking

Han Jiangtao, Yao Longxing, Yang Long, Zhang Zhiyi

► **To cite this version:**

Han Jiangtao, Yao Longxing, Yang Long, Zhang Zhiyi. Dynamic 3D Scanning Based on Optical Tracking. 19th International Conference on Entertainment Computing (ICEC), Nov 2020, Xi'an, China. pp.434-441, 10.1007/978-3-030-65736-9_39 . hal-03686003

HAL Id: hal-03686003

<https://inria.hal.science/hal-03686003v1>

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Dynamic 3D Scanning Based on Optical Tracking^{*}

Han Jiangtao¹, Yao Longxing¹, Yang Long¹, and Zhang Zhiyi^{*1}

Northwest A&F University, College of Information Engineering
Yangling, Shaanxi, China
zhangzhiyi@nwafu.edu.cn, hanjiangtao1995@163.com

Abstract. In order to solve the problem that traditional binocular vision 3D scanning requires multiple scans and then registration, a dynamic 3D scanning method based on optical tracking is proposed. The first part is monocular optical scanning, which uses laser stripes as the active light source to achieve the function of 3D scanning. The second part is monocular optical tracking, which realizes the calculation and tracking of the position and posture of the scanning device, and instantly converts the point cloud data obtained by the scanning camera to the tracking camera coordinate system, realizing real-time data registration during the dynamic scanning process. The experimental results show that this method can achieve dynamic scanning to obtain the point cloud information of the object. The accuracy of the scan is about 1mm in the direction of the **XOY** coordinate plane of the coordinate system and about 0.08mm in the direction of the z -axis. At the same time, the scanning time can be saved by 50% compared with the traditional binocular vision 3D scanning method.

Keywords: Dynamic 3D scanning · Optical tracking · Real-time registration.

1 Introduction

The 3D scanning technology can scan the shape of the space object to obtain 3D point cloud data on the surface of the scanned object. At present, most of the 3D scanning equipment that uses structured light or laser as the active light source needs to perform multiple scans at different viewpoints, and then registration each part of the data to obtain complete data of the object. This application process, because it does not solve the problem of dynamic scanning, not only makes its application scenarios extremely limited, but also the time consumption and accuracy loss of subsequent registration affected its ease of use.

Since the 1990s, camera calibration methods based on specific targets have made successive breakthroughs. Classic five-point algorithm, six-point algorithm, calibration algorithm [9, 11], self-calibration algorithm, and Bundle Adjustment

^{*} Supported by “the Fundamental Research Funds for the Central Universities”, and “National Natural Science Foundation of China No.61702422”

algorithm [2, 6] have appeared one after another. In 2003, Andrew Davison proposed Visual SLAM based on monocular camera, published the groundbreaking MonoSLAM [3], and opened the era of visual positioning—Visual SLAM [1]. Recent research shows that using LCD (Liquid Crystal Display) to achieve camera calibration [4], the accuracy can reach 0.018 pixels. Zhiyi Zhang and Lin Yuan put forward a method to realize 3D scanning by monocular vision combined with laser stripes [10]. In addition, the best paper of CVPR 2019 [8] starts with Shape-from-X, and can obtain millimeter-level accuracy. In the classical calibration algorithm, Zhang Zhengyou algorithm needs several images of position and posture transformation to get accurate results, and it is difficult to obtain position and posture information with only one image. The Tsai’s two-step calibration algorithm can be used for reference [7, 5].

Different from the existing methods, the method proposed in this paper decomposes the traditional 3D scanning system based on laser stripes and binocular vision into two parts: one part is a scanning system based on laser stripes and monocular vision, and the other part is based on optical tracking system for monocular vision.

2 Methodology

2.1 Overall system description

The dynamic scanning system designed in this paper is mainly composed of two parts, as shown in Fig.1.

The first part is a monocular optical tracking system, consisting of a tracking camera. It mainly implements the calculation and tracking of the position and posture of the scanning device. First, calibrate the position and posture of the scanning camera in the tracking camera coordinate system, and then the position and posture between the scanning camera and the tracking target in the tracking camera coordinate system should be calibrated.

The second part is a monocular optical 3D scanning system, consisting of a scanning camera, tracking target and laser. Using laser stripes as the active light source, the center of the light strip is first extracted to obtain sub-pixel-level light strip information, and then the light plane parameters are calibrated by repeatedly calibrating the calibration board. Finally, the light strip information and light plane parameters are used to calculate the point cloud data of a single image.

Finally, the point cloud information obtained by the monocular 3D scanning system, using the parameter information calibrated by the monocular optical tracking system, instantly converts the scanned data to the system with the optical tracking device as the boundary coordinate, to realize the data registration in the dynamic scanning process.

2.2 Monocular optical tracking system

In this part, the optical tracking of 3D dynamic scanning is mainly realized. By calculating the position and posture between the tracking target and the

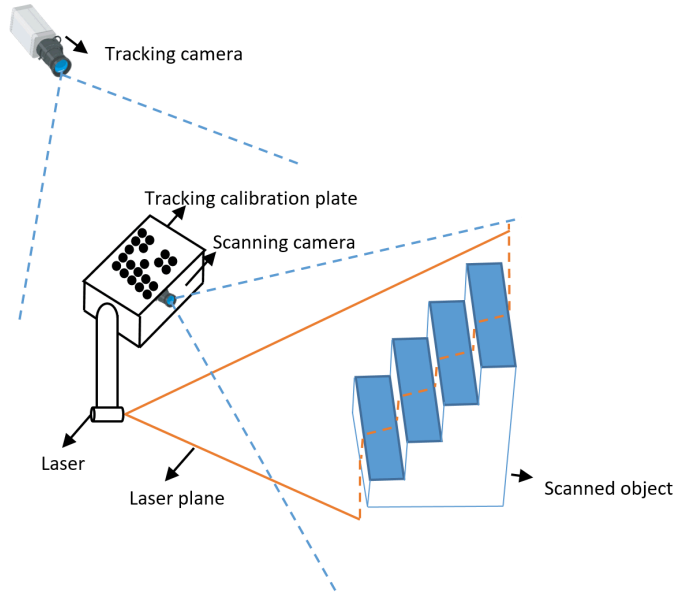


Fig. 1. 3D dynamic scanning model

scanning camera, the point cloud information obtained by the monocular 3D scanning system is instantly converted to the tracking camera coordinate system to complete the dynamic scanning. This part mainly includes the position and posture calibration of the scanning camera and the tracking target and the conversion from a single image to the tracking camera coordinate system.

Scanning camera and tracking target position and posture calibration

In order to ensure the 3D point cloud information obtained from each image during the dynamic scanning process, and implement the conversion from the scanning camera coordinate system to the tracking camera coordinate system, the key step is to calibrate the position and posture between the scanning camera coordinate system and the tracking target. In this part, the scanning camera and the tracking camera are used to obtain the position and posture of the plane of the calibration board at the same time multiple times, so as to calibrate the position and posture of the scanning camera and the tracking target. The transformation relationship model of position and posture calibration is shown in Fig.2.

It can be seen from Fig.2 that the tracking camera can simultaneously obtain the calibration point coordinates of the tracking target $\mathbf{P}_{tt} = [x_{tt}, y_{tt}, z_{tt}]^T$ and the calibration target $\mathbf{P}_{tc} = [x_{tc}, y_{tc}, z_{tc}]^T$, and the scanning camera can obtain the calibration point coordinates $\mathbf{P}_{sc} = [x_{sc}, y_{sc}, z_{sc}]^T$ of the calibration target. Since the entire dynamic scanning system is under the same world coordinate system, the position and posture of the calibration target in the world coordinate system are determined and unique. The position and posture of the scanning

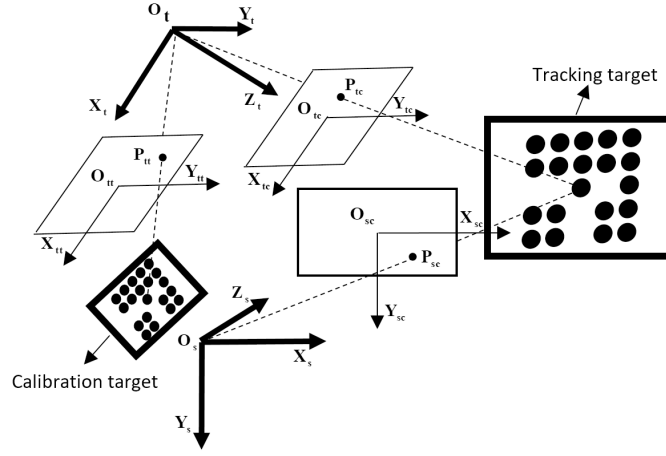


Fig. 2. Scanning camera and tracking target position and posture calibration model

camera coordinate system under the tracking camera coordinate system can be calculated using different calibration point coordinates under the scanning camera coordinate system and the tracking camera coordinate system.

First, calculate the translation transformation \mathbf{T}_{sc} between \mathbf{P}_{tc} and \mathbf{P}_{sc} , then move \mathbf{P}_{tc} and \mathbf{P}_{sc} to the origin of the coordinate system, and use the least square method to calculate the rotation transformation \mathbf{R}_{tc} between the two, as shown in equation(1).

$$\mathbf{P}_{tc} = \mathbf{R}_{tc} \mathbf{P}_{sc} + \mathbf{T}_{tc} \quad (1)$$

The coordinate of the scanning camera under the tracking camera can be obtained by using the rotation and translation transformation, as shown in equation(2). The calibration board used in this paper is composed of a set of circular calibration points in an ordered sequence. The direction of the x -axis and y -axis of the calibration plate can be determined from only one image, so the sequence of characteristic calibration points of the \mathbf{XOZ} coordinate plane can be used instead of the position and posture of the coordinate system. It should be noted here that since the rotation and translation transformation from the scanning camera coordinate system to the tracking camera coordinate system was previously calculated, the translation operation needs to be reversed when the scanning camera position and posture is reversed by the tracking camera.

$$\mathbf{P}_{scaCam} = \mathbf{R}_{tc} \mathbf{P}_{traCam} + \mathbf{R}_{tc}^{-1} \mathbf{T}_{tc} \quad (2)$$

Where $\mathbf{P}_{traCam} = [x_i, 0, z_i]^T$ is the calibration points under the tracking camera coordinate system, and $\mathbf{P}_{scaCam} = [x_i, y_i, z_i]^T$ is the calibration points under the scanning camera coordinate system.

The position and posture of the tracking target in the tracking camera coordinate system can be directly calculated. Therefore, after determining the

position and posture of the scanning camera in the tracking camera coordinate system, the final calibration of the scanning camera and the tracking target can be performed. Here, it is necessary need to calculate the rotation and translation matrix \mathbf{RT}_{ts} from the tracking target \mathbf{P}_{tt} to the scanning camera \mathbf{P}_{scaCam} , as shown in equation(3).

$$\mathbf{P}_{scaCam} = \mathbf{RT}_{ts} \mathbf{P}_{tt} \quad (3)$$

Conversion from single image to tracking camera coordinate system

After determining the position and posture between the scanning camera and the tracking target, the point cloud information obtained by each scan can be converted to the tracking camera coordinate system. First, the tracking camera obtains the 3D coordinates of the tracking target in the tracking camera coordinate system, and uses the inverse operation of equation(3) to calculate the corresponding 3D coordinate position and posture of the scanning camera coordinate system in the tracking camera coordinate system. Then it is necessary to determine the rotation and translation transformation \mathbf{RT}'_{ts} from the tracking coordinate system to the scanning coordinate system, as shown in equation(4).

$$\mathbf{RT}'_{ts} = [\mathbf{P}_{scaCam} \mathbf{P}_{traCam}^T] [\mathbf{P}_{traCam} \mathbf{P}_{traCam}^T]^{-1} \quad (4)$$

After applying the calculated rotation and translation transformation to the point cloud \mathbf{P}_{s_obj} obtained under the scanning camera coordinate system, the 3D coordinates of the scanning object point cloud under the tracking camera coordinate system \mathbf{P}_{t_obj} can be obtained, as shown in equation(5).

$$\mathbf{P}_{t_obj} = \mathbf{RT}'_{ts} \mathbf{P}_{s_obj} \quad (5)$$

3 Experimental results

In this paper, C++ is used as the programming language, and OpenCV 3.0 is used as the image processing library. Windows 10 with 8 GB memory is used as the development platform for the system. The tracking camera is Logitech HD Pro Webcam C920 with a resolution of 1920*1080. The scanning camera is Huagu Power WP-UT530 with a resolution of 2592*2048.

3.1 Error assessment

In this part uses a 3D printer to print out a standard model, and scans the standard model to obtain the point cloud information of the model, so as to analyze the error of the dynamic scanning model. The error assessment model used in this part is a staircase model. As shown in Fig.3(a), this staircase model has four steps, and the height and width of each step set to 20mm. The point cloud data obtained by scanning the front steps of the stairs is shown in Fig.3(b).

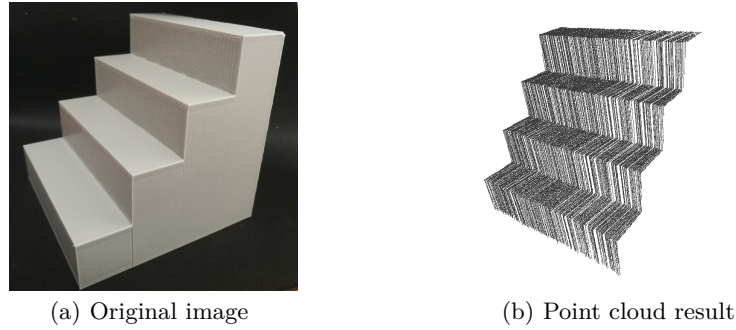


Fig. 3. Error assessment model

For the obtained point cloud data, the front model consists of 8 faces with 4 steps. After accurate measurement with vernier calipers, the distance between the horizontal and vertical planes of the real model of each step and the corresponding plane of the next step can be obtained with an accuracy of 0.01mm. Use the measured point cloud data to fit 8 planes respectively, calculate the distance between adjacent steps and compare with the actual distance. The comparison results are shown in Table 1.

Table 1. Distance error measurement

Compare plane	True distance/mm	Scan fitted distance/mm	Distance error/mm
Plane 1 and 3	19.88	18.6261	-1.2539
Plane 2 and 4	20.02	19.2738	-0.7462
Plane 3 and 5	19.82	18.6382	-1.1818
Plane 4 and 6	19.94	19.1493	-0.7907
Plane 5 and 7	19.82	18.3960	-1.4240
Plane 6 and 8	19.94	18.9861	-0.9539
Average	19.90	18.8449	-1.055

3.2 Point cloud details

In order to verify the effectiveness and feasibility of the method in this paper, the bronze teapot and doll printed by 3D printer was used as measurement models to show the obtained point cloud data and details. As shown in Fig.4, each model consists of the original image, the overall point cloud image and the detailed enlarged image. It can be seen from the images that each model can realize the dynamic scanning of the scanned object.

However, in addition to the scanning method, the material and color of the scanned object will also affect the scanning result. For the matte material, the red laser has a reflection property similar to Lambertian, so the point cloud

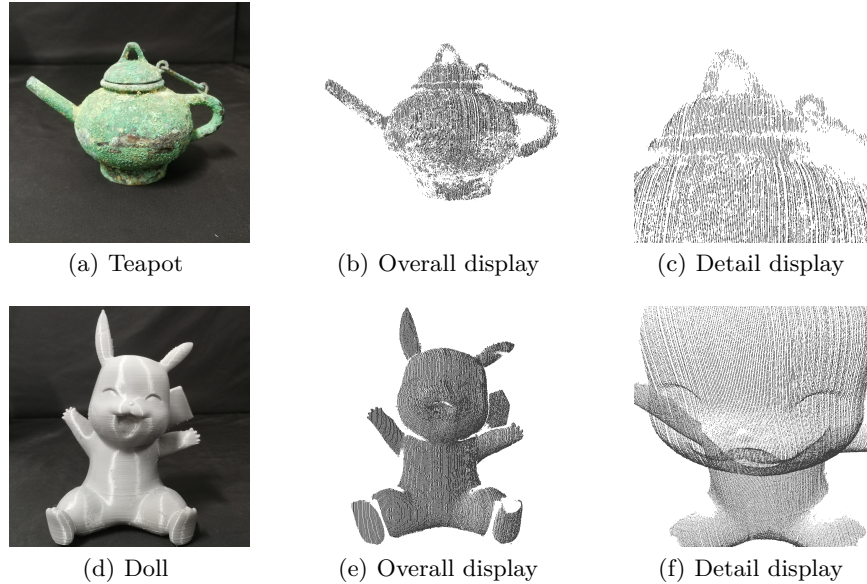


Fig. 4. 3D dynamic scanning results

data obtained by scanning is better. As shown in Fig.4(b), it is the scan result of the teapot. Due to the age of the bronze teapot model, the surface is corroded to produce patina (basic copper carbonate, $\text{CuCO}_3 \cdot \text{Cu}(\text{OH})_2$), and dark spots formed after the patina is peeled off. When the laser light strip is projected to these parts, the red laser light is absorbed, and the R channel of the corresponding pixel point on the image cannot reach the threshold, so that the surface is mottled. The edge of the teapot lid has a thickness of about 3mm, and there is a ring with an inner diameter of about 2mm next to the teapot lid. Both details can be obtained with clear point cloud data. For the doll model of Fig.4(e), it can be seen that the eyes of the doll in the original picture have a depth of about 1mm. It can be seen from the enlarged image that the dynamic scanning method proposed in this article can obtain the details of the eye part well. It can be seen from the point cloud details that the dynamic scanning model proposed in this paper can convert the scanned data to the direct coordinate system with the optical tracking device in real time during the dynamic scanning process, which can avoid the registration process of multi-view scanning data in the later traditional method, and support interrupted scanning.

4 Conclusion and Outlook

The experimental results show that the method proposed in this paper can scan the two groups of objects to obtain the point cloud information of the scanned objects, with an accuracy of about 1mm. The method proposed in this paper

can convert the scanned point cloud into the tracking camera coordinate system in real time by separating scanning and tracking, which not only avoids the registration process of multi-viewpoint scanned point cloud in the traditional method, but also allows intermittent scanning operations. However, it should be pointed out that in the scanning range, it is affected by the tracking device's field of view range and the accuracy of tracking target board acquisition on the tracking camera coordinate system. How to achieve precise control of optical tracking equipment, how to expand the field of view of the tracking camera is the main content to be studied next.

References

1. Bloesch, M., Czarnowski, J., Clark, R., Leutenegger, S., Davison, A.J.: Codeslam - learning a compact, optimisable representation for dense visual slam. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2560–2568 (2018)
2. Davison: Real-time simultaneous localisation and mapping with a single camera. In: Proceedings Ninth IEEE International Conference on Computer Vision. pp. 1403–1410 vol.2 (2003)
3. Davison, A.J., Reid, I.D., Molton, N.D., Stasse, O.: Monoslam: Real-time single camera slam. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(6), 1052–1067 (2007)
4. Jiang, J., Zeng, L., Chen, B., Lu, Y., Xiong, W.: An accurate and flexible technique for camera calibration. *Computing* **101**(4), 1–18 (2019)
5. R, T.: An efficient and accurate camera calibration technique for 3d machine vision. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. pp. 364–374 (1986)
6. Tang, Q.H., Zhang, Z.Y.: Camera self-calibration based on multiple view images. *Computer Engineering and Science* (2017)
7. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal on Robotics and Automation* **3**(4), 323–344 (1987)
8. Xin, S., Nousias, S., Kutulakos, K.N., Sankaranarayanan, A.C., Narasimhan, S.G., Gkioulekas, I.: A theory of fermat paths for non-line-of-sight shape reconstruction. In: Proceedings of (CVPR) Computer Vision and Pattern Recognition (June 2019)
9. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), 1330–1334 (2000)
10. Zhang, Z., Yuan, L.: Building a 3d scanner system based on monocular vision. *Appl Opt* **51**(11), 1638 (2012)
11. Zhengyou Zhang: Flexible camera calibration by viewing a plane from unknown orientations. In: Proceedings of the Seventh IEEE International Conference on Computer Vision. vol. 1, pp. 666–673 vol.1 (1999)