Predictive Display for Telerobot Under Unstructured Environment

Qing Wei, NaiLong Liu and Long Cui

Abstract In this paper, the issue that traditional virtual reality (VR) system can’t be used in unstructured environment is addressed, and a novel predictive display method based on 3D scene reconstruction online is proposed. In this method, the virtual environment is modeled and reconstructed online by the structured light, so the consistency with the real world can be ensured; An adaptive random sample consensus (ARANSAC) algorithm is proposed to denoise the point cloud when modeling the environment, which adjusts the parameters of the standard RANSAC algorithm adaptively; A rigid contact model based haptic rendering algorithm is adopted to generate the force feedback directly from the manipulator’s dynamics, which eliminates the time delay and security problem introduced by the feedback force and increases the frequency of force feedback; Moreover, a low frequency filter is designed to keep the virtual force smooth. Finally, the effectiveness of the proposed method under unstructured environment is demonstrated by experiments.

Keywords Telerobot  VR  Unstructured environment  Scene reconstruction  Force feedback
1 Introduction

Telerobotics, perhaps one of the earliest aspects of robotics, and has been used in many areas, like space exploration, underwater exploration and nuclear environment. When teleoperating a robot, the perception is very important for the teleoperator [1]. However, affected by the random communication delay, the interactive information is delayed, deficient and scrambled. In order to overcome the time delay issue, many control strategies have been proposed [2–10], and the predictive display method has been reputed as the most promising solution to tackle big time delay. But, the predictive display technology relies heavily on the accuracy of the virtual model, so how to apply it to unstructured environment become a serious problem.

In Mars exploration, NASA built the 3D virtual environment by the received stereo images [11], but the accuracy, instantaneity and robustness of this method can’t be insured. A point cloud based geometry modeling method is proposed by Fan Qiang [12], the algorithm of this method doesn’t require complex calculation, so the instantaneity can be insured, and this method has been used in many virtual reality systems [13–15]. But it does not mention the environment dynamics and the point cloud processing algorithms have not been deeply discussed, and the lose of contact force feedback is dangerous in contact tasks. In [3], a virtual force computing method is proposed, and the dynamic parameters of the remote environment is identified by the sliding-average least-square algorithm. But the algorithm can not been used in variable time delay condition, and the safety can’t been insured when the slave environment is hard. Moreover, the collaborative expression of predictive vision and force is also an important and urgent problem should be solved.

In order to solve the aforementioned issues, a novel predictive display method based on 3D scene reconstruction online is proposed in this paper. In this method, the geometry model of the environment is constructed by point cloud in real time, and an adaptive random sample consensus (ARANSAC) algorithm is proposed to denoise the point cloud. On the force interaction, the manipulator’s dynamics which can be calculated in advance is used to calculate the virtual force. The proposed method eliminates the time delay and security problem introduced by the feedback force and improve the adaptability of the algorithm in unstructured environment. In order to realise the collaborative expression of predictive vision and force, the physical simulation and graphic simulation are calculated in different thread, so the haptic and graphical refresh rate meet the requirements of continuous perception. The teleoperation system constructed in this paper was shown in Fig. 1.
In this paper, we use structured light to get the point cloud of the fusion reactor. The structured light calculates the depth data by triangulation principle, so the processing method is very simple, and has a high real time capability. But affected by light and other factors, the original point clouds are normally quite noisy, so denoising and smooth operation are very important. In our experiment, the reconstructed model consists of three parts: fusion reactor (a part of the spherical), beryllium tiles (plane) and obstacles (No contact task, don’t need denoising). The model of the point clouds need denoising are known, and an adaptive random sample consensus (ARANSAC) algorithm is proposed in our work to perform this operation.

2.1 A-RANSAC and Point Cloud Denoising

The standard random sample consensus (RANSAC) algorithm is proposed by Fischler [14], and proceeds as follows: first, samples of points are drawn uniformly and at random from the input data set. Each point has the same probability of selection. For each sample of points, a model hypothesis is created by computing model parameters from it. In the next step, the quality of the model is evaluated on the whole input data set. A cost function, which evaluates the model quality by calculating a distance threshold $\delta$, is applied to identify the inliers. This process is repeated and the model with the largest number of inliers is stored, and terminated when the probability $p$ that at least one of the random samples of points is free from outliers becomes high enough. Assume that $w$ is the proportion of inliers in the data, $n$ is the sample size and $k$ is the number of iterations required for RANSAC.
Then, we can get:

\[ k = \log \left( \frac{1 - p}{1 - w^n} \right) \]  

In the standard RANSAC, \( w \) and \( \delta \) are two important parameters. However, their values are estimated, if the \( w \) is too large, maybe we can’t get the right model; In turn, if the \( w \) is too small, the instantaneousity of the algorithm will be reduced. On the other hand, if the threshold \( \delta \) is too large, the denoising result maybe not good; In turn, choose the threshold \( \delta \) too small will cause atrophy of the target model.

In order to reduce the randomness of parameter estimation, an adaptive random sample consensus (ARANSAC) algorithm is proposed. In our algorithm, the \( w \) and \( \delta \) are adjusted based on the point cloud and least square method. First of all, identify the original mode through least square method, and the \( \delta \) is calculated as (3).

\[
\delta = \sqrt{\frac{\sum_{i=1}^{n} (d_i - d_m)^T (d_i - d_m)}{n-1}}
\]

where \( d_i \) is the distance of point I to the original model, and \( d_m \) is the average of \( d_i \). The number of inliers \( n_i \) can be get based on the original model and \( \delta \), and the \( w \) is calculated as \( w = n_i/n \). Then the algorithm proceeds as the standard random sample consensus (RANSAC) algorithm.

In order to verify the effectiveness of the proposed algorithm, a contrast matlab simulation is conducted, and the result was shown in Fig. 2 (the data is a part of the point cloud of the fusion reactor). From the result we can see, compared with the standard random sample consensus (RANSAC) algorithm, our algorithm not only denoises the point cloud, but also keeps more surface details. What more, the randomness of the parameter estimation is eliminated.
2.2 Scene Reconstruction

The point cloud scanned by the structured light is in the camera coordinates, but the robot is in the world coordinates. So the point cloud should be converted to the world coordinates before reconstruction. In our system, the structured light camera fixed at the end of a holder, the holder has three DOFs, which ensures the structured light can scan all the workspace of the robot. The joints of the holder can be obtained through the controller, and the transformation matrix from the camera coordinates to the world coordinates is shown as (4).

\[
\begin{pmatrix}
  c_1 c_2 c_3 - s_1 s_3 & s_1 c_3 + c_1 c_2 s_3 & c_1 s_2 & l \cdot s_1 \\
  -c_1 s_3 - s_1 c_2 c_3 & c_1 c_3 - s_1 c_2 s_3 & -s_1 s_2 & l \cdot c_1 \\
  -s_2 c_3 & -s_3 c_3 & c_2 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

where, \( s_i = \sin \theta_i, c_i = \cos \theta_i, i = 1, 2, 3 \), is the joints of the holder, and \( l \) is the long of the holder.

As the point cloud has been denoised, we use the ball-pivoting algorithm to convert it to a triangle mesh. The vertex normal is the average of the triangle normals that sharing this vertex. The framework of the mesh data is fixed as vertex-normal-index, which decreases the number of logic judgment when loading the model, and improves the real-time performance of the system.

3 Force Predictive Display

Force information is an effective supplement to the vision, particularly in the contact tasks, where 70% of the perceptual information are get from force. However, influenced by the communication delay, the feedback force from the slave always be delayed, which dislocates the force feedback and the prediction graphics. The dislocation not only makes the operator fatigue but also threaten operation safety, so the real time force feedback is also very important in the predictive display system.

A great number of research on virtual force calculation have been reported, for example, mass-damp-spring model, artificial potential field model and virtual fixture method. However, those methods can’t be used in our system directly. Because the environment is unstructured and the task is contact task with rigid bodies, so even very small delay or modeling errors would cause large contact forces.

Considering the deformation and penetration would not happen in rigid collision, a novel mass-damp-spring model based virtual force calculation method is proposed. The dynamic parameters of the environment is instead by manipulator’s, and the calculation model as follows:
\[
\tau = M(\Theta) \dot{\Theta} + V(\Theta, \dot{\Theta}) + G(\Theta) \\
F = J^{-T}(\Theta) \cdot [\tau - G(\Theta)]
\]  

where, \( \tau \) is the joint torque, \( M(\Theta) \) is the mass matrix of the manipulator, \( V(\Theta, \dot{\Theta}) \) is the coriolis and centrifugal, \( G(\Theta) \) is the gravity terms and \( F \) is the contact force and torque. \( J(\Theta) \) is manipulator’s jacobian matrix when the collision occurs. As the manipulator would stop after collision, so the inverse of the jacobian matrix only calculates once in a collision course, and the frequency of force feedback increases. The virtual manipulator uses the same velocity planning algorithm as the real manipulator, \( \dot{\Theta} \) and \( \dot{\Theta} \) is calculated based on the joints when the collision occurs, the target joints and the manipulator’s velocity planning algorithm.

4 Experiments

Our experiment platform is shown in Fig. 3, consists of a stereoscopic projection system, a structured light camera (3D CaMega CP-300), a holder, a 6DOFs robot (Staubli TX60L), a master manipulator (Force Dimension Omega.3) and a control server. The communication channel between the master and slave is internet.
4.1 Geometry Reconstruction

The slave environment consists of a fusion reactor and three beryllium tiles, as Fig. 4. In order to improve the instantaneity of the system, only a part of the environment (in the workspace of the robot) need to be reconstructed.

First, move the holder to a suitable pose by the master controller; then give a scanning instruction to the structured light camera; after the scanning finished, the point cloud and the joints of the holder are send to the virtual reality system to reconstruct the virtual scene. The coordinate transformation result of the point cloud and the reconstruction result is shown in Fig. 5, and Fig. 6 shows the reconstruction result when a complex shape obstacles suddenly appeared.

The reconstruction experiments indicate that the proposed method can reconstruct the virtual scene under unstructured environment on line.

4.2 Virtual Force Feedback

When calculating the virtual force, the inertia tensor of the robot links was calculated by numerical integration based on their 3D model. Due to the visual residue phenomenon, the operator will feel the video continuous when its refresh rate is...
more than 24 Hz. However, if we want the operator feel a continuous force, the refresh rate of the feedback force should more than 200 Hz. In order to improve the frequency of the force feedback, a separate thread is created for physical simulation. The GUI refresh rate and the physical simulation frequency are shown as follows:

From Fig. 7 we can see that both the GUI refresh rate and the force feedback rate meet the continuous perception demand of human.

When calculating the virtual contact force, the first and second order differential of the joints is used, so it’s very sensitive to joint jitter. However, the jitter of the operator is inevitable when operating the master manipulator, so the contact force calculated from the dynamical model jittering heavily, as Fig. 8a.

In order to keep the feedback force smooth, a low pass filter is designed, and the cutoff frequency of the filter is determined by the FFT result of the virtual force, as Fig. 8b. And the virtual feedback force after filtered is shown in Fig. 8c. Restricted by the flexible and response time of the operator, the shaking value of the contact force is bigger at the moment of the collision occurs, and the value is bigger when the contact force is big, as Fig. 8d. Because of the algorithm proposed in this paper doesn’t need the dynamics of the environmental, and the geometry is reconstructed on line, so it can be used in unstructured environment.
This article examines the predictive method under unstructured environment, and a novel predictive display method based on 3D scene reconstruction online is proposed. In this method, the slave environment is monitored by camera, and the virtual environment is modeled and reconstructed online, so the consistency with the real world can be ensured. For denoising the point cloud, an adaptive random sample consensus (ARANSAC) algorithm is proposed, which adjusts the parameters of the standard RANSAC algorithm adaptively; On the force calculation, the rigid contact model and the manipulator’s dynamics are used to calculate the virtual force, which eliminates the time delay and security problem introduced by the feedback force and increases the frequency of force feedback; In our program, the physical simulation and graphic simulation are calculated in different thread, which ensures the GUI refresh rate and the force feedback rate are meet the continuous.

Fig. 8 a Force calculated from the dynamical model. b FFT result of the contact force. c Contact force after filtered. d Shaking value of the contact force

5 Conclusion

This article examines the predictive method under unstructured environment, and a novel predictive display method based on 3D scene reconstruction online is proposed. In this method, the slave environment is monitored by camera, and the virtual environment is modeled and reconstructed online, so the consistency with the real world can be ensured. For denoising the point cloud, an adaptive random sample consensus (ARANSAC) algorithm is proposed, which adjusts the parameters of the standard RANSAC algorithm adaptively; On the force calculation, the rigid contact model and the manipulator’s dynamics are used to calculate the virtual force, which eliminates the time delay and security problem introduced by the feedback force and increases the frequency of force feedback; In our program, the physical simulation and graphic simulation are calculated in different thread, which ensures the GUI refresh rate and the force feedback rate are meet the continuous.
perception demand of human, and a low frequency filter is designed to keep the virtual feedback force smooth. At last, the effectiveness of the proposed method is confirmed by experiments.

References