The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

# Viewpoint Planning for Automated 3D Digitization using a Low-cost Mobile Platform

Sijian Zhang, Gangfeng Yan, Weihua Sheng

*Abstract*— **This paper presents a low-cost, simple automated mobile platform for 3D environmental digitization. Compared with other customized 3D digitization platforms, all parts in our platform are commercial off-the-shelf. In order to build a 3D model of a salient target, we present a new viewpoint planning method for fast 3D digitization. Within a predefined accuracy, our method can determine the minimum overlap between two consecutive scanning images to speed up the digitization process. This guarantees the next scanning image can be merged to the previous one properly. The results from both simulation and experiments show the effectiveness of our viewpoint planning method. Our tests of this mobile robot system demonstrate the feasibility of 3D digitization based on a low-cost platform.**

*Index Terms*— **LRF sensor, 3D digitization, mobile robot, ICP**

## 1 INTRODUCTION

Obtaining environmental information is very important for robotic applications, such as search and rescue. 3D environmental modeling can help the robot navigate and avoid obstacles. On the other hand, automated 3D digitization can help blind people sense the surrounding obstacles. Compared with optical cameras, Laser Range Finder (LRF) sensors can be used in dark and smoky environments, therefore they are widely used in 3D digitization.

Some existing 3D digitization platforms are equipped with expensive LRF sensors, computers and mobile robots. For example Nüchter *et al*. provide a heavy and complicated *Ariadne* robot for 3D scanning experiments [1]. Allen *et al*. provide a mobile site modeling robot for 3D urban environments modeling [2]. *AEST* (Autonomous Environmental Sensor for Telepresence) is developed to reconstruct 3D interiors from reality [3]. Due to the complexity and cost, it is not very easy to duplicate their platforms. This prevents other researchers from improving the platforms and the algorithms for 3D digitization. Therefore it is necessary to provide a simple and low-cost platform for 3D digitization.

This paper presents a solution to build a low-cost automated 3D digitization platform. Our platform consists of an iRobot [4] equipped with a URG-04LX 2D LRF sensor [5] and a compact computer eBox2300 [6]. It can find out a salient target, move to different viewpoints and create a 3D model of the target.

To digitize large targets, the platform should use multiple viewpoints and be able to decide where the next viewpoint is. A scanning image at one viewpoint can only cover one part of the whole model. In order to create a 3D model of large targets, scanning images from different viewpoints should be merged together. The Iterative Closest Point (ICP) algorithm [7, 8] is an effective method to register two images taken from two different viewpoints. However it may fail when there is insufficient overlap between the two images [9]. On the other hand, if the amount of overlap is too much, the digitization process will significantly slow down. A viewpoint planner should consider these two factors to finish 3D digitization within the predefined error tolerance as fast as possible. Therefore, in this paper, we also present a viewpoint planning algorithm that enables the platform to analyze the current scanning scene data, so that the proper overlapping area can be identified between the current scan and the next scan.

The paper is organized as follows. Section 2 introduces some existing work related to 3D digitization. Section 3 describes the hardware and software components of the 3D digitization platform. Section 4 presents the basic data processing for 3D digitization. Section 5 provides the viewpoint planning algorithm. The experimental results are shown in Section 6. Section 7 concludes the paper and shows our future work.

## 2 RELATED WORK

Many researchers have worked on 3D environment modeling. Nüchter *et al*. [1] develop the *Ariadne* robot equipped with the *AIS* 3D LRF sensor. There are three laser scanners, two computers and a 250 kg robot in this platform. It is powerful, but also expensive and heavy. In order to reduce the cost, a few groups try to replace the 3D scanner with a 2D scanner. Nüchter *et al*. [10] use a standard 2D scanner and a servo motor to form a 3D LRF sensor. When the sensor conducts yawing scanning, it results in the maximal possible field of view. Ueda *et al*. [11] provide the similar mechanical setup but use a smaller and lighter LRF sensor. These platforms are not built from commercial off-the-shelf parts. Therefore it is not easy to duplicate them.

In order to build a 3D model of a real object, scanning

Sijian Zhang, is with the College of Electrical Engineering, Zhejiang University, Hangzhou, Zhejiang 310027 China (e-mail: zsjzju@yahoo.com.cn).

Gangfeng Yan, is with the College of Electrical Engineering, Zhejiang University, Hangzhou, Zhejiang 310027 China (e-mail: ygf@zju.edu.cn).

Weihua Sheng, is with the School of Electrical and Computer Engineering ,Oklahoma State University, Stillwater, OK, 74078 USA (e-mail: weihua.sheng@okstate.edu).

images from various views need to be aligned. A few groups use invariant features for ICP registration, which can increase the chance of success [12]. Surmann *et al*. [13] use a scan matching algorithm based on the ICP algorithm to register the 3D scans in a common coordinate system and localize the robot. Some groups extract the invariant feature from different sensed views to ensure that the ICP algorithm can find out the correspondent point-pairs. For example Besl *et al.* compute curvatures in range images [14]. Sadjadi *et al.* detect moment invariants in three dimensions [15]. Burel *et al.* propose a framework for deriving 3D rotational invariant features [16]. However, the algorithm for invariants is complicated. Different kinds of invariant features need different extraction algorithms. This requires powerful computation sources. Furthermore not all objects have the same kind of invariant features. The scanning system should select the proper type of invariant features before deriving invariants.

Automated 3D digitization not only needs to register different images taken from different viewpoints but also needs to solve the Next Best View (NBV) problem [17]. Roberts *et al.* provide a method that aims to minimize the number of viewpoints while ensuring that the selected viewpoints be close to the best viewpoints [18]. There is a tradeoff between digitization accuracy and speed. Overlap constraint must be satisfied during viewpoint planning. Pito applies a simple non-linear function to decide how much of the object's surface would be resampled from the current position [17]. However the choice of the threshold for this function is dependent on the features of the object. An effective method should be used to describe the complexity of the object surface, so that the scanner can reduce the rescanning area and minimize the number of viewpoints.

## 3 HARDWARE AND SOFTWARE OF THE 3D DIGITIZATION PLATFORM

We developed a low-cost mobile robot using commercial off-the-shelf (COTS) parts for 3D scanning. As shown in Fig.1, the robot consists of a URG-04LX 2D LRF sensor, an eBox2300 compact computer, an iRobot and a rechargeable battery. The total cost of this robot is around 2,500 US dollars.

The URG-04LX is a small and light LRF sensor. Its high accuracy, high resolution and wide angle provide a good solution for 3D scanning. Its maximum detection distance is 4000 mm and its scanning angle is 240 degrees. The iRobot provides a platform for easy integration and creation of new robot applications. There are two RS-232 serial ports on the eBox2300: one is used to communicate with the LRF sensor, the other is used to control the iRobot. As shown in Fig.2, the eBox2300 controller acquires 2D distance information from the LRF sensor and controls the motion of the iRobot. The robot can continuously operate for about 2 hours with one fully charged nickel hydride battery.

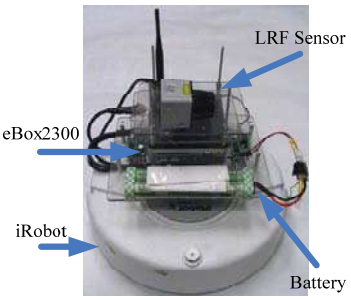

Fig. 1. The 3D scanning robot

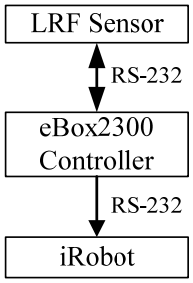

Fig. 2. The functional blocks for the robot

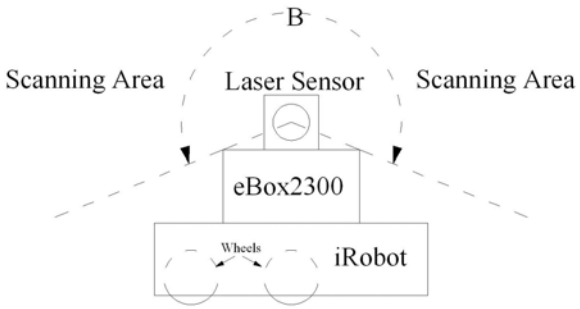

Fig. 3. The left view of the scanning area

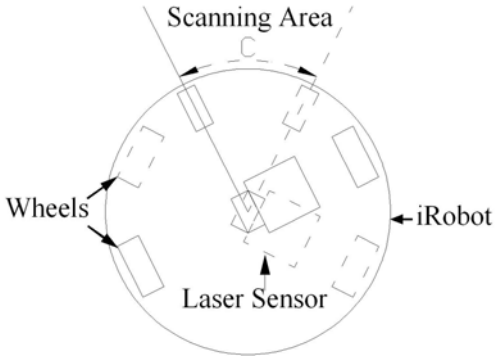

Fig. 4. The top view of the scanning area

The LRF sensor is mounted on the top of the eBox2300. As shown in Fig.3, the scanning plane of the sensor is vertical. Fig.4 shows the top view of the platform. With autogyration of the iRobot, a servo motor for rotating the LRF sensor is not necessary. The sensed view of the platform is *B* degree vertical and *C* degree horizontal. The maximum scanning angle *B* can be up to 240 degree and the maximum scanning angle C can be up to 360 degree. There are some vibrations of the LRF sensor when the robot is rotating. We assume that the resulted drift distance and vibration are small and can be ignored.

Fig.5 shows the software flow chart of 3D digitization. The controller configures the LRF sensor (e.g. baud rate, scanning area) in the initialization. After receiving all data in one scan, the controller checks if there is any data loss during the communication. With the distance data from the LRF sensor and orientation data from the iRobot, the controller can build the 3D image at one viewpoint and register the image with the

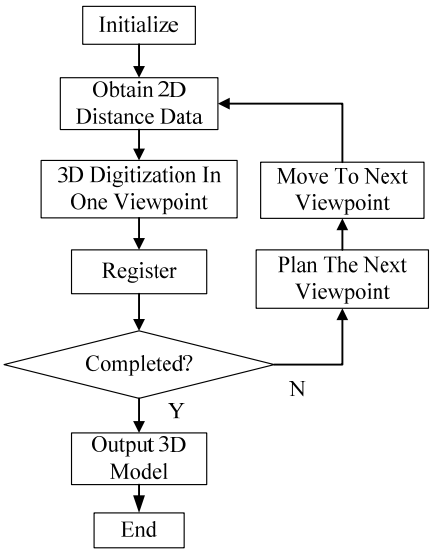


Fig. 5 The software flow chart for 3D digitization.

previous scanning images. The current scanning image is analyzed by the controller to find out the overlapping area and plan the next scanning position and robot orientation ( $x, y, \theta$ ). Then the controller commands the robot to move to the planned viewpoint. When the robot completes the 3D digitization, a 3D model of the target is created.

## 4 Data Processing for 3D Digitization

This section describes how to process the raw data.

### 4.1 Gap Filling

Raw data should be processed to improve the quality of the 3D model. The measured distance from the LRF sensor is generally greater than the minimum detection distance (20 mm). However, as shown in Fig.6a, there are some gaps in the raw image. The data read from the LRF sensor is zero in these directions. These zero points are caused by shiny curved surfaces. The laser light emitted by the LRF sensor is reflected to certain unknown directions. Therefore the LRF sensor can not receive any light in these areas and then treats the distance as zero.

Curve fitting such as linear interpolation is applied to fill these gaps. The result is shown in Fig.6b, which clearly fits the bad points and removes the gaps.

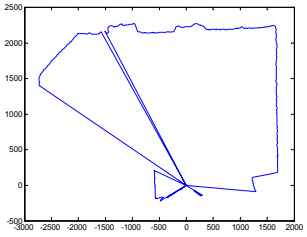

Fig. 6 a) The raw image.

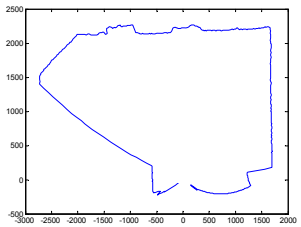

Fig. 6 b) The processed image

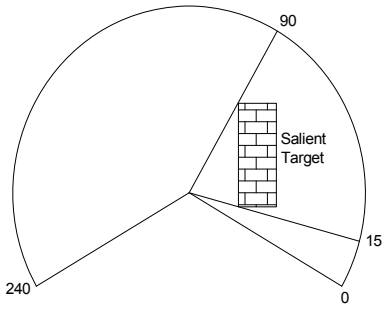


Fig. 7 Extracting salient targets

### 4.2 Target Extraction

In many applications, only salient targets rather than the surrounding environment are of interest. Therefore, algorithms such as edge detection are applied to remove the background information and some spare error points. After this processing the salient target is extracted. Fig.7 illustrates the idea of background removing. In the 240 degree detection area of the LRF sensor, only 75 degree scanning area is taken into account.

## 5 Viewpoint Planning

As an automated 3D digitization robot, it should have the capability to plan the next viewpoint. In general, the problem of viewpoint planning is to make a tradeoff between two goals: minimizing the amount of overlapping area to speed up the digitization process and maintaining the accuracy of registration. There is overlap constraint between two consecutive scans. If this constraint cannot be satisfied, the next image may fail to register with the current scanning image. In general, large overlapping areas contain more features than their corresponding subsets. Therefore large overlapping areas can increase the probability of successful registration. However, in order to speed up the digitization, the next scanning image should cover as little overlapping area as possible. Additionally, less overlapping areas makes the ICP algorithm run faster.

This paper provides a new algorithm for viewpoint planning. First, it generates a set of candidate viewpoints based on the selected overlapping areas in the current scanning image and virtual scanning images are generated for these viewpoints. Then the regular ICP algorithm is applied to simulate the registration between the current scanning image and the virtual scanning image. The performance of the registration will be evaluated, which determines the goodness of the corresponding viewpoint. We call this algorithm the Simulated ICP (SiICP) based viewpoint planning algorithm.

Given two independently acquired sets of 3D points, $M$ (model set) and $D$ (data set), the ICP algorithm works to find a rotation $R$ and a translation $T$ which minimizes the following mean square error (MSE):

$$E(R,T) = \frac{1}{N}\sum_{i=1}^{N}\left\|m_i - (Rd_i + T)\right\|^2 \quad (1)$$

Generally, when the overlapping area between two

scanning images to be merged is more characteristic, the MSE in the merged overlapping area will be less. Therefore the MSE can be treated as a performance index of the candidate viewpoint. Therefore, our problem is to find out the viewpoint from which the LRF sensor can cover the maximum unscanned area while the evaluated MSE in this viewpoint is within the predefined error tolerance.

During the planning stage, the platform does not move to the candidate viewpoints to obtain the scanning images. Instead, virtual scanning images are created and registered, which gives the name of *Simulated ICP*. Hence, there are two problems to solve. One is how to decide the candidate viewpoints and the other is how to generate a virtual scanning image from each of the candidate viewpoints.

As shown in Fig.8, let $(x_1, y_1, \theta_1)$ represent the position and orientation of the LRF sensor at the current viewpoint $O_1$ and $(x_2, y_2, \theta_2)$ represent the position and orientation of the LRF sensor at the next viewpoint $O_2$. Given $(x_1, y_1, \theta_1)$ and the LRF sensor scanning segment $S_1$, when the overlapping segment $S_2$ at viewpoint $O_1$ is fixed, the position and orientation of the next viewpoint $(x_2, y_2, \theta_2)$ can be calculated based on the geometry of the current scanning image. Therefore the remaining problem is to generate a virtual scanning image which contains the same overlapping area as the current image.

The virtual scanning image can be generated from the overlapping subset of the current image. As shown in Fig.9, let $P$ represent the current image, $Q$ represent the virtual overlapping subset. Then,

$P = \{l_1, l_2, \ldots, l_n\}$, where $l_i$ is one scanning line, $n$ is the total number of scanning lines in the current scanning image.

$Q = \{l_1, l_2, \ldots, l_{nQ}\}$, $n_Q < n$, where $n_Q$ is the number of simulated overlapping scanning lines

Assume that two images to be aligned have been roughly preregistered, so that the relative rotation $R$ and translation $T$ between these images depend on the motion error of the mobile platform. We only present one dimension of rotating, because the robot obtains 3D images in yawing pattern. Let $Q''$ represent the virtual scanning image Therefore,

$$R = \begin{bmatrix} \cos(\theta_u) & -\sin(\theta_u) & 0 \\ \sin(\theta_u) & \cos(\theta_u) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; T = \begin{bmatrix} x_u \\ y_u \\ 0 \end{bmatrix} \tag{2}$$

$$Q'' = R \times Q + T + N_{noise}$$

From our experience the uncertainty in translation $(x_u, y_u)$ is random and proportional to the distance $D$ between the current viewpoint and the next viewpoint. Similarly assume that the uncertainty in rotation angle $\theta_u$ is

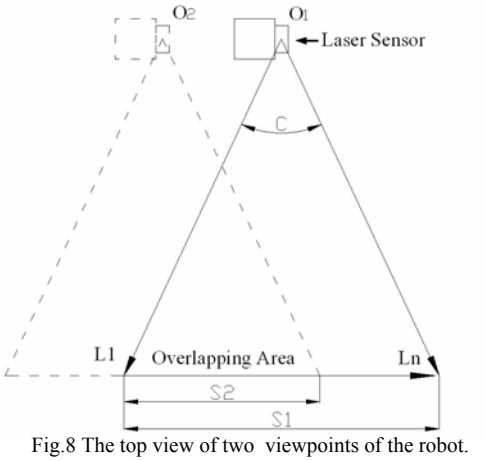

Fig.8 The top view of two viewpoints of the robot.

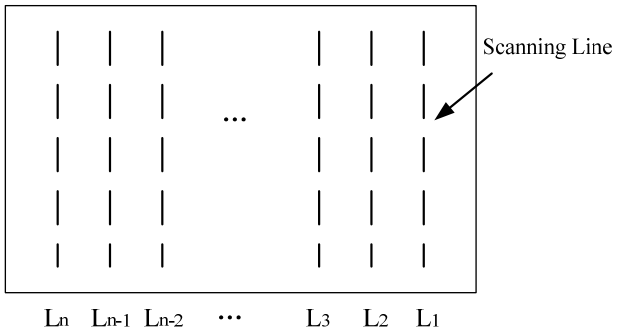

Fig. 9 The scanning image consists of *n* scanning lines.

random and proportional to the orientation difference between $\theta_1$ and $\theta_2$, $(x_u, y_u, \theta_u)$ can be written as follow:

$x_u \propto N(0, (\lambda \Delta x)^2)$, where $\Delta x = x_1 - x_2$

$y_u \propto N(0, (\lambda \Delta y)^2)$, where $\Delta y = y_1 - y_2$

$\theta_u \propto N(0, (\gamma \Delta \theta)^2)$, where $\Delta \theta = \theta_1 - \theta_2$

where $\lambda$ and $\gamma$ are constants. The regular ICP is applied between the current image $Q$ and the virtual scanning image $Q''$ to calculate the minimum MSE

$$E(R', T') = \min\left(\frac{1}{n_Q} \sum_{i=1}^{n_Q} (R' \times q''_i + T' - q_i)^2\right) \tag{3}$$

If the performance index $E$ is greater than the predefined accuracy, we increase the number of the overlapping scanning lines $n_Q$. Otherwise we decrease $n_Q$.

Overall, the essential idea for the SiICP is to find out the minimum $n_Q$ that the correspondent MSE $E \le \varepsilon$, where $\varepsilon$ is the predefined accuracy.

$$\begin{cases} n_Q = \arg\min_{n_Q} E \\ E \le \varepsilon \end{cases} \tag{4}$$

In real registration between two images, the sizes of the overlapping area from two images can not be exactly the same. Therefore $Q''$ is rewritten as

$$Q' = \{l_1, l_2, \ldots, l_{nQ'}\}, \text{ where } n_{Q'} = \alpha \cdot n_Q, \alpha < 1 \tag{5}$$

$$Q'' = R \times Q' + T + N_{noise} \tag{6}$$

The worst scenario is that the scanned surface is featureless, so that the minimum $E$ is larger than the predefined accuracy. As a default method, our algorithm will set the overlapping segment to be 80% of the scanning segment $S_2 = 0.8S_1$ to make the robot move to $O_2$.

The pseudo code of our viewpoint planning algorithm is as follows.

*Viewpoint Planning Algorithm*

1. *For $n_Q = \eta \cdot n$, $\eta$ is from 0.2 to 0.8.*
    a. *Select a subset $Q$ from $P$ as a model.*
    b. *Select a subset $Q'$ from $Q$ (eq. (5)).*
    c. *Calculate the position of the next virtual viewpoint based on the selected model $Q$.*
    d. *Generate a random rotation matrix $R$ and a translation vector $T$ (eq. (2)).*
    e. *Generate a virtual scanning image $Q''$ from $Q'$ (eq. (6)).*
    f. *Run the ICP between $Q$ and $Q''$ to calculate the minimum mean square error $E$ (eq. (3)).*
    g. *If $E$ is less than the predefined accuracy $\varepsilon$, then exit.*
2. *Move to the next viewpoint.*

## 6 Experimental Result

To demonstrate the usefulness of the low-cost mobile robot and the effectiveness of our algorithm for viewpoint planning, we present test results from a simulated 3D digitization in Matlab and from a real 3D digitization experiment.

### *6.1 3D Digitization Simulation*

Fig. 10 is the top view of the outline of simulated 3D model. The dimension of the model is about $40 \times 30 \times 50$ cm and there are totally 60,000 points. This model consists of some typical surfaces (e.g. sine surface and flat surface). Since flat surfaces do not contain much constraint, it allows two surfaces to slide against each other when they are merged together. Vertexes and gaps have more constraint, so that they contain more features.

Fig. 11 is the top view of the result of the 3D digitization. Fig. 12 (a) shows the 3D view of the model and Figure.12 (b) shows the digitization result. The star points show the position of all viewpoints. From Fig.11 we can see that if the surface is characteristic enough (e.g. with large curvatures or gaps), the next viewpoint is far from the current viewpoint. Otherwise the distance between the current and the next viewpoint is closer. The predefined accuracy in this test is 1cm, and every registration error (MSE) is under this predefined accuracy.

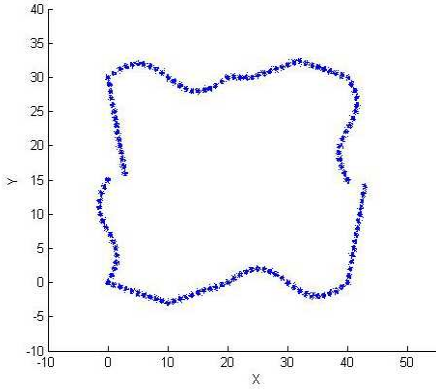


Fig. 10 The simulated 3D model (top view).

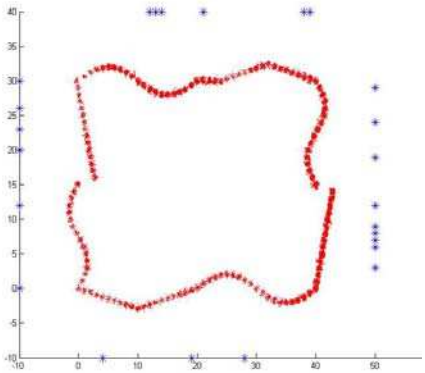

Fig.11 The simulation result of the 3D digitization (top view).

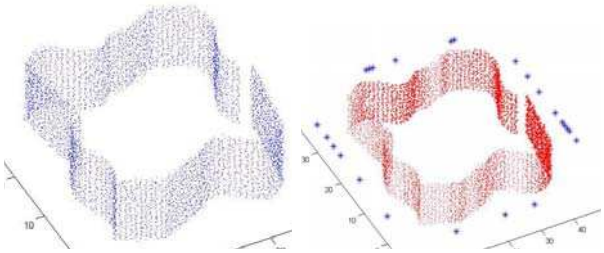

Fig. 12 a) The simulated 3D model (3D view) .

Fig. 12 b) The simulation result (3D view).

### *6.2 Real 3D Digitization*

The proposed method has also been tested on real targets. Fig.13 shows the photo of the scene. The target to be scanned consists of three boxes. In this experiment, the mobile robot only scans the front side.

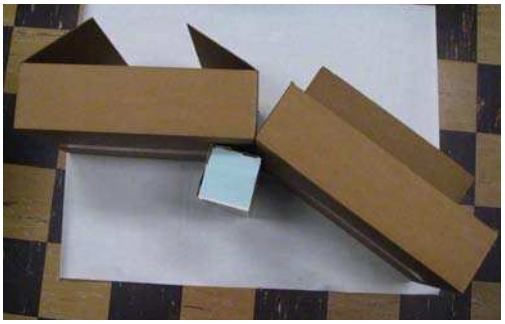

Fig. 13 The real scene for 3D digitization

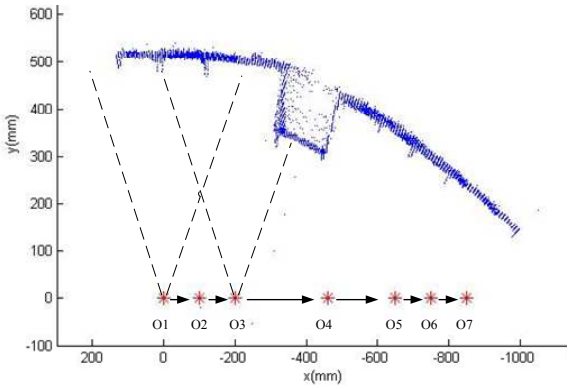

Fig.14 The 3D model acquired with seven viewpoints (top view).

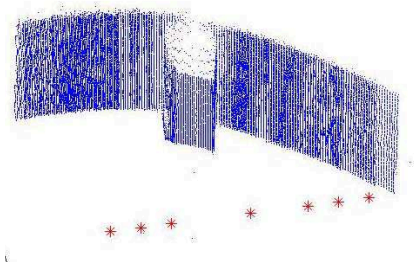
Fig.15 The 3D scanned model in another angle view.

Fig. 14 is the top view of the result of the 3D digitization while Fig. 15 shows the result in another view angle. The predefined accuracy in this experiment is 10mm, which is the same as the accuracy of the LRF sensor. As shown in Fig. 14, there are totally 7 viewpoints in this experiment which are represented by star points. The dashed lines show the scanning area of viewpoint $O_1$ and $O_3$, respectively. Since the surfaces acquired in viewpoint $O_1$ and $O_2$ are very smooth, they do not contain sufficient features. The distance from $O_3$ to $O_4$ is obviously longer than the distance between $O_1$ and $O_2$. This is because the LRF sensor in $O_3$ can cover the edge of the small box, which causes the robot to move more distance.

## 7 Conclusion and Future Work

This paper presents a low-cost automated mobile platform for digitizing 3D targets. The system consists of an iRobot, a small and light 2D LRF sensor and a compact computer eBox2300. Our platform is cheap and compact. Compared with other existing platforms, all parts in our platform are commercial off-the-shelf. There is no customized part, so that it is easy to duplicate our platform.

To address the viewpoint planning problem, an algorithm based on Simulated ICP is presented as well. Based on the scanned image from the current viewpoint, the algorithm can find out the minimum overlapping area while ensuring the image acquired from the next viewpoint can be registered within the predefined accuracy.

The aim of future work is to improve the accuracy of the 3D digitization algorithm. Furthermore, based on the scanned image the algorithm can predict the shape of the surface to be scanned. This can speed up the 3D digitization process.

## Acknowledgment

This project is supported by the DoD ARO DURIP grant 55628-CS-RIP.

## References

[1] Andreas Nüchter *et al*, "Planning Robot Motion for 3D Digitalization of Indoor Environments," in *Proc. of the 11th International Conference on Advanced Robotics (ICAR)*, 2003.

[2] P. Allen *et al*, "AVENUE: Automated Site Modeling in Urban Environments," In *Proceedings of the third International Conference on 3D Digital Imaging and Modeling* (3DIM 2001), Quebec City, Canada, May 2001.

[3] V. Sequeira *et al*, "Automated 3D reconstruction of interiors with multiple scan-views," In *Proceedings of SPIE, Electronic Imaging 1999, The Society for Imaging Science and Technology /SPIE's 11th Annual Symposium*, San Jose, CA, USA, January 1999.

[4] iRobot Corporation. http://www.irobot.com

[5] HOKUYO Automatic CO.,LTD. http://www.hokuyo-aut.jp

[6] DoNetNuke Corporation. http://www.embeddedpc.net/eBox2300/tabid/110/Default.aspx

[7] P. Besl, N. McKay, "A method for registration of 3-D shapes," *Trans. PAMI*, 14(2), 1992.

[8] Y. Chen, G. Medioni, "Object modeling by registration ofmultiple range images," *Proc. IEEE Conf. on Robotics and Automation*, 1991.

[9] Kok-Lim Low, Anselmo Lastra, "Predetermination of ICP Registration Errors And Its Application to View Planning," in *Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM),* 2007.

[10] Andreas Nüchter *et al*, "3D Mapping with Semantic Knowledge," in *Proceedings of the RoboCup International Symposium*, Osaka, Japan, June 2005.

[11] Tatsuro Ueda *et al*, "Mobile SOKUIKI sensor system: Accurate Range Data Mapping System with Sensor Motion," *International Conference on Autonomous Robots and Agents*, Dec. 2006.

[12] Gregory C. Sharp, Sang W. LEE, David K. Wehe, "ICP registration using invariant features," *PAMI 24*, 1 (2002), 90.102.

[13] Hartmut Surmann *et al*, "An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments," *Robotics and Autonomous Systems*, vol. 45, pp. 181–198, 2003.

[14] P.J. Besl and R.C. Jain, "Invariant Surface Characteristics for 3D Object Recognition in Range Images," *Computer Vision Graphics and Image Processing.* vol. 33, no. 1, pp. 33-80, Jan. 1986.

[15] F.A. Sadjadi and E.L Hall, "Three-Dimensional Moment Invariants," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 2, no. 2, pp. 127-136, Mar. 1980.

[16] G. Burel and H. Henocq, "Three-Dimensional Invariants and Their Application to Object Recognition," *Signal Processing*, vol. 45, no. 1, pp. 1-22, July 1995.

[17] Pito, R., ''A sensor-based solution to the next best view problem,'' *Proc. ICPR '96*, 1996, pp. 941-945.

[18] D.R. Roberts, A.D. Marshall, "Viewpoint selection for complete surface coverage of three dimensional objects," in *Proceedings of the British Machine Vision Conference (BMVC),* University of Southampton, UK, 1998, pp. 740 – 750.