A 3-D point clouds scanning and registration methodology for automatic object digitization

Liang-Chia Chen, Dinh-Cuong Hoang, Hsien-I Lin and Thanh-Hung Nguyen

Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan; Graduate Institute of Automation Technology, National Taipei University of Technology, Taipei, Taiwan; Graduate Department of Mechanics of Materials & Structures, Hanoi University of Science & Technology, Hanoi, Vietnam

ABSTRACT
The article presents a robot 3-D scanning system for generation of 3-D point clouds of an object by using multi-view 3-D scanning and novel data registration. Our approach mainly comprises two important elements in the determination of next best probe pose and multiple-view point clouds registration. A novel technique is proposed to register 3-D object scene with overlapped or stacked condition. Under this scenario, conventional registration methods such as the iterative closest point algorithm usually fail to converge to a global minimum when a good initial estimate for image registration does not exist. Our proposed technique uses a 3-D scanner to be mounted on a six degree of freedom-articulated industrial robot. It keeps moving probe continuously in the working range against the object and autonomously varying the probe with various gestures required for complete object scanning and for achieving best 3-D sensing accuracy. The robot scanning path is determined through a proposed algorithm using information from the latest scanning data and registered result of the object. The developed method has been verified through experimental tests for its feasibility test. It confirms that the registration accuracy with one standard deviation can be controlled within 0.323 mm when the objects underlying reconstruction are in range of hundreds of millimeters.

KEYWORDS
robot vision; point clouds; object reconstruction; 3-D image registration; next-best-view planning

1. Introduction
Recently, model reconstruction technology using range scanners has been widely applied in many applications, such as 3-D printing, reverse engineering, rapid prototyping, and medical prosthetics. According to various sensing principles, the existing techniques can be generally classified into two main categories, namely semi-automatic and fully automatic scanning techniques. In a semi-automatic system, the model of a 3-D object is reconstructed by manually guiding the range scanner sensing over the surface of the object.\[1,2\] The effectiveness of this scanning system highly depends on the skills of the user and the scanning process is time-consuming. To enhance the efficiency, the 6-axis robot arm integrated with 3-D scanners has emerged as a technical developing trend for 3-D surface scanning for objects having arbitrary or complex geometry. The critical problems currently remained in developing the automated systems are mainly in accurate data registration and automatic robot path planning for complete object scanning.

Firstly, the registration of point clouds is one of the most important steps in processing data from the scanner. Conventionally, the iterative closest point (ICP) algorithm, proposed by Besl and McKay,\[3\] has been widely used for multi-view point cloud registration. The approach iteratively detects the closest point pairs between two scenes and generates a transformation using the unit quaternion technique. If the initial transformation estimate between scenes is provided, then the ICP can effectively recover the rigid body transformation that aligns two point clouds. It converges monotonically to a local minimum, which may or may not be the global minimum, and all points in a point cloud are assumed to have correspondence in the other cloud. Since the ICP is based only on a local search algorithm to recover correspondence between two point clouds and it minimizes the sum of square distances between possible corresponding points, its iterative process converges slowly and tends to fall into local minima. Therefore, fine registration requires the good
initial estimated transformation between two set of point clouds having sufficient level of data overlapping. To establish a rough initial estimate of the motion between pairs of consecutive scanning views, there were many existing algorithms, known as coarse alignment. [4–6] As a general case, to compute the initial pose estimate, the geometric distances between corresponding 3-D points existing in different views are minimized. The most common correspondences are points, curves, and surfaces. Point signature [4] is a point descriptor to search for correspondences, which describes the structural neighborhood of a 3-D point by the distances between neighboring points and their corresponding contour points. Johnson and Hebert proposed the concept of a spin image [5] to create for an oriented point at a vertex in the surface, which is a 2-D histogram containing the number of points in a surrounding supporting volume. More recently, Tombari et al. [6] presented a 3-D descriptor called the Signature of Histograms of Orientations (SHOT) that encodes information about the surface within a spherical support structure. Under some assumptions, the coarse registration techniques reviewed above have demonstrated their individual feasibility for computing an initial estimation of the rigid motion between two clouds of 3D points. However, there still remains one significant common problem among them, in which the existing methods mainly rely on the geometric characteristics of the point clouds based on regional or local geometric information, which could easily lead to less discrimination or sensitivity in dealing with object geometry having symmetrical or plural repeating surface contours.

As critical as the registration, the robot path planning is referred to as the next-best-view planning and has been addressed by several researchers.[7–12] Maver [7] selected the next-best-view based on the degree of sensing occlusion on object surface. However, error-free positioning was assumed and viewing constraints on view overlap for data registration was ignored. Pito [8] presented an algorithm which solves ‘Next Best View’ (NBV) problem by the visibility analysis of the object and range sensor through casting visibility images onto a positional space. The approach used a view overlapping constraint but did not address the impact of pose error on view planning. Recently, Kriegel et al. proposed a method called Next-Best Scanning (NBS) planning for autonomous 3-D modeling by iteratively detecting the surface trend of the boundary regions representing in a triangle mesh. [9] Holes detection, 6-DOF search space, pose error, and constraints of view overlap were all taken into account in the method. However, according to the experimental results, generated models still lacked a sufficient point density or completeness. Therefore, the NBV planning for best scanning efficiency and object coverage still remains as a challenge to be further explored.

In the research, an automated robot scanning system with a proposed methodology is developed to address the technical challenges being remained. The system mainly consists of a NTU-developed optical 3-D measuring probe (briefed NTU-3D probe) as attached on a fixed position of an industrial 6-axis robot that can provide proper location and orientation for 3-D full-field area sensing. The robot path is defined by iteratively selecting a next best robot pose based on the latest scanned and registered data. To produce a complete model, a surface map is reconstructed from the unorganized point cloud by the acquisition and registration steps. [13–16]

2. Overview of the automated 3-D scanning methodology

The system consists of a NTU-3D probe, a six degree of freedom articulated industrial robot, a dedicated robot controller and a control computer, which communicates with the robot controller and the probe. The working principle of the robotic 3-D scanning system is that the optical probe is attached on the robot end effector, in which the probe is controlled and scanned continuously by the robot in the working area and simultaneously posed with different gestures in the process of scanning. Shown in Figure 1, initially, the robot moves to a pre-defined position for the first probe scan. By doing so, the system uses the acquired image to determine the initial object orientation. According to the processed information, the position and gesture of robot for next view are further estimated and employed to generate the robot path for next movement. This process is repeated until the whole object is scanned. In the scanning process, the system controls the robot position and gesture to continuously scan and register point clouds of consecutive scenes of the scanned object in real time.

3. Developed algorithms for robot control and data processing

3.1. Preprocessing of point clouds

Since the speed of data processing strongly depends on the number of points in the point clouds, it is necessary to remove as many irrelevant cloud points as possible. This task mainly includes three tasks of outlier removal, down sampling, and plane segmentation.

Because of measurement errors, the data acquired from 3-D scanner potentially may contain undesired sparse outliers. For this issue, outlier removal is performed through the computation of the distribution of the mean distance
from a query point to its nearest neighbors in the input data-set.\cite{17}

Similarly, downsampling is an efficient solution to reduce the size of point clouds. It is important that the global shape of the scene is preserved during the data downsampling. By specifying the minimum distance between two sets of point clouds, the number of points can be adequately reduced. Figure 2 illustrates the concept of downsampling; the point cloud is decreased from 480,000 points to 40,000 points in the case, where the major geometric contour and structure of the scanned object are contained.

Meanwhile, for plane segmentation, the RANSAC algorithm\cite{18} is adapted to segment the planar background. RANSAC is used to estimate the parameters of a mathematical model from a set of observed data, which contains outliers. All the points belonging to the ground can be segmented by applying the following four major processing steps:

1. To select three random points \( \{ p_i, p_j, p_k \} \) from the point cloud \( P \).
2. To compute the plane that fits these points \( ax + by + cz + d = 0 \).
3. To determine the number of inliers. That is, for each point \( p \) in the point cloud \( P \), to compute the distance to the plane. If the distance is less than the specified threshold, the point is taken as an inlier.
4. To repeat this process by saving the plane parameters which produce all the inliers.

\subsection{3.2. Coarse registration of point clouds}

To obtain the transformation matrix between the consecutive views or coarse registration, which serves as a good initial estimate for the ICP algorithm for further refinement of the transformation matrix, the position and orientation of the optical must be known in the scanning process. Based on the current robot position, transformation of robot coordinates can be performed to determine the transformation matrix since the probe attached on the robot end effector. According to this requirement, three coordinate systems as shown in Figure 3 are established as follows:

\[ \mathbf{O}_{w}X_wY_wZ_w \] defined as the world coordinates, for convenience computation, is usually registered with the world coordinates with the robot base coordinates \( \mathbf{O}_{R}X_RY_RZ_R \). The probe or camera coordinates \( \mathbf{O}_{C}X_CY_CZ_C \) is also registered with the robot wrist coordinate which is implicit in the kinematic model of the robot with its motion parameters \( (x, y, z, R_x, R_y, R_z) \). The coordinate transformation matrix between the robot base coordinates and the probe coordinates is generally given as:

\[
\begin{bmatrix}
  x_C \\
  y_C \\
  z_C
\end{bmatrix} =
\begin{bmatrix}
  r_1 & r_2 & r_3 \\
  r_4 & r_5 & r_6 \\
  r_7 & r_8 & r_9
\end{bmatrix}
\begin{bmatrix}
  x_R \\
  y_R \\
  z_R
\end{bmatrix} +
\begin{bmatrix}
  t_x \\
  t_y \\
  t_z
\end{bmatrix}
\]

where \( r_1 \rightarrow r_9 \) are the components of the rotation matrix \( R \) with the translation vector \( (t_x, t_y, t_z)^T \).

\subsection{3.3. Overlapping detection and refined registration}

Following the above operation, the transformation matrix established in the previous section is used to initially align point clouds between two consecutive scans with different views. Although the parameters of this transformation matrix can be obtained in the initial registration, the accuracy may not be satisfactory due to various robot positioning errors. Therefore, a refine registration like ICP algorithm is performed for precise surface reconstruction between the registered point clouds. However, by applying ICP algorithm, the point clouds of View \( i \) can be registered in its best alignment with the point cloud of View \( i + 1 \) if...
and only if their overlapping area of views can be accurately detected and used as the input to the ICP.

Illustrated in Figure 4, the overlap detecting process is detailed in the flowchart. Assuming that $P$ and $Q$ are sets containing $n$ points and $m$ points for View $i$ and View $i+1$, respectively, in order to collect points in set $P$ which belong to the overlapping area, Kd-Tree algorithm is employed to determine the $k$ nearest points $N_k$ in set $Q$ to each point $p_i$. [19] If the following condition is satisfied, $p_i$ belongs to the overlapping region $P_{\text{overlap}}$:

$$\sum_{j=0}^{k} d\left(p_i, N_k^j\right) \leq r_{\text{thresh}}$$

(2)

where $r_{\text{thresh}}$ represents the threshold distance:

$$r_{\text{thresh}} = \text{Max} \left\{ \frac{\sum_{j=0}^{k} d\left(p_i, M_k^j\right)}{k}, i = 1, 2, \ldots, n \right\}$$

(3)

where $M_k^i$ is the set $k$ nearest neighbors of $p_i$ in $P$.

Similarly, set $Q_{\text{overlap}}$ can be also obtained by comparing its local mean distance of each point $q_i$ in $Q$ to the $k$ nearest points in $P$ and again to the $k$ nearest neighbors in $Q$. Then, following this, the ICP algorithm [3] is then employed to obtain the refined transformation. The algorithm for refine registration is illustrated in Figure 5.

### 3.4. Robot path planning

In order to reach the entire surface of the scanned object, multiple scans with various views are normally needed. In the study, the developed system successively completes a 3-D model from unknown objects by iteratively selecting a next robot pose based on the previously acquired data. A 3-D scan is performed by moving the robot along a continuous scan path and obtaining 3-D point clouds data, which is synchronized with the robot poses. The initial scan path is a pre-defined path in the free space, while subsequent one is planned based on its previous scanned data. This process mainly includes four steps as follows:

1. To check whether the point clusters are located on the boundary area of the image frame not?
2. If the point clusters are located on the boundary area of frame, the gravity point of each cluster is calculated. The closest gravity point to the probe will become the next robot position $P (x, y, z, R_x, R_y, R_z)$. 

---

**Figure 3.** Transformation of robot coordinates.

**Figure 4.** Kd-tree-based overlap points detection algorithm.

**Figure 5.** Algorithm for refine registration.
is accomplished on smoothed depth map using the method proposed by Suzuki.[21]

When all the surfaces of the object are scanned, automatic hole detection is performed.[22] A boundary probability is computed for every point based on its neighborhood collection and the angle criterion. Then, the points are classified into boundary and interior points. Finally, for each hole, a boundary loop is extracted. Figure 6 shows an example for hole detection on two samples with two real object holes and a hole caused by missing data being detected.

4. Experiments and evaluation

4.1. Optical probe setup

In this study, we tested the system on two different kinds of optical probes developed at NTU. One is the phase shifting probe (NTU-3D Probe 1) that uses digital fringe projection method by projecting a sinusoidal fringe pattern onto the object, grabbing the image of the object illuminated with projected fringes, and then calculating the height information of the object through phase shifting,[23] phase unwrapping, and phase-height transformation processes (Figure 7).

Another 3-D scanner (NTU-3D Probe 2) is based on laser using random speckle pattern projection principle and triangulation theory.[24,25] Figure 8(a) and (b) displays the speckle image and the acquired 3-D point clouds, respectively. The measured field of view is approximately $147 \times 110 \times 80$ mm$^3$.

4.2. Experimental results and discussion

To demonstrate the feasibility of the proposed system, two objects including a 3-D printing hammerhead and a toy-face mask were scanned and verified. The size of the hammerhead is around $110 \times 35.5 \times 21$ mm$^3$. The size of the toy-face mask is $118 \times 79 \times 46$ mm$^3$. To estimate the computation cost of the scanning process, the experiments are processed on a computer with a Core i7 processor having 3.40 GHz and 8 GB RAM. The hammerhead was located on a fixed table, which allows for viewing the object from several positions around the object. However, not all scan views can be reached. The object was scanned by the speckle probe and the system terminated after 8 scans. Each scan contains 20 k points. The total time for 3-D model generation of the Hammerhead was approximate five minutes. This includes the time for moving the robot while scanning, preprocessing, data registration, and generating robot path. The robot motion and coordinate transformation normally consume most of the operation time. Figure 9(a) shows

![Figure 6](image1.png)

**Figure 6.** Example of holes detection: (a) Two real object holes detected on the hammer sample; (b) A hole caused by missing data on a top part of a toy model due to missing scan.

![Figure 7](image2.png)

**Figure 7.** Example of measurement using the NTU-3D Probe 1: (a) Acquired fringe image of object; (b) Acquired 3-D point clouds of object.

![Figure 8](image3.png)

**Figure 8.** Example of measurement using the NTU-3D Probe 2: (a) Acquired speckle image of 3-D printed hammerhead. (b) Acquired 3-D point clouds of hammerhead.

(3) Otherwise, the boundary points of current object point clouds are detected. Based on the detected boundary, the next robot position $P(x, y, z, R_x, R_y, R_z)$ is defined to scan the object with the next view angle.

(4). When the whole surface of the object is assumed to be scanned, in which $P = NULL$, any hole remaining in the point clouds are detected for continuous scan until no hole is detected or the hole is identified as a non-measured geometric feature of the object.

In Step 3, the boundary detection can be achieved using the method proposed in.[20] An orthogonally projected depth map is established by projecting the 3-D point clouds along the optical imaging axis. Then, the smoothness of the depth map is enhanced through employing the morphology operations. Eventually, boundary estimation
the scanned hammerhead model with some overlapping point clouds while Figure 9(c) illustrates scanned models after overlapped data was removed, resulting in a triangular mesh with 10,201 triangles.

Similarly, the toy-face mask was scanned using the phase shifting probe and accomplished within around four minutes. Four scans were implemented with 480 k points per scan that are diminished to 40 k points after down sampling. The scanned toy-face mask results are shown in Figure 10 with 3-D point clouds (c), a mesh with 33,156 triangles (d).

The performance of the registration can be estimated based on the distance from each overlapping point in the point clouds P of View i to the closest overlapping point in the point clouds Q of View i + 1. Let \( d_i \) be the distance between a point \( p_i \) in P and its closest neighbor \( q_i \) in Q. Then, the mean distance \( \mu \) and standard deviation \( \sigma \), which are used to evaluate the performance of the registration, are computed as follows:

\[
\mu = \frac{1}{n} \sum_{i=1}^{n} d_i
\]  

\[
\sigma = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \mu)^2}{n - 1}}
\]

In these experiments, the mean deviation in data registration for Hammerhead was 0.369 mm and for toy-face mask was 0.325 mm, respectively. The standard deviation for Hammerhead was 0.323 mm and for toy-face mask was 0.317 mm, respectively.

5. Conclusions

In this study, a robotic surface digitization system consisting of an industrial robot and a 3-D optical probe, for accurate surface reconstruction of an object is presented. The key step of the developed method relies on the probe pose and position information which is used as a robust initial estimate for the ICP algorithm in image registration. The system can autonomously scan 3-D surface by
iteratively determining a next probe pose based on the existing acquired geometry without losing any crucial surface data. More importantly, the kd-tree-based overlap points detection algorithm is also developed to detect crucial overlapping scanned data. The experimental results have demonstrated that the registration accuracy with one standard deviation can be controlled within 0.323 mm when the objects underlying reconstruction are in range of hundreds of millimeters.

Disclosure statement
No potential conflict of interest was reported by the authors.

References
[1] Bodenmoller T, Hirzinger G. Online Surface Reconstruction from Unorganized 3D-Points for the DLR Hand-guided Scanner System. Proceedings of 2nd International Symposium on 3D Data Processing, Visualization and Transmission; 2004, Sep 6–9, pp. 285–292. Thessaloniki: IEEE.
[2] Strobl KH, Sepp W, Wahl E, Bodenmoller T, Suppa M, Seara J, Hirzinger G. The DLR Multisensory Hand-Guided Device: The Laser Stripe Profiler. In: Proceedings of ICRA; 2004, 26 April–1 May, pp. 1927–1932. New Orleans, LA: IEEE.
[3] Besl PJ, McKay HD. A method for registration of 3-D shapes. IEEE Trans. Pattern Recognit. Machine Intelligence. 1992;14:239–256.
[4] Chua CJR. Point signatures: a new representation for 3d object recognition. Int. J. Comput. Vision. 1997;25:63–85.
[5] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Transa. Pattern Recognit. Machine Intelligence. 1999;21:433–449.
[6] Tombari F, Salti S, Di Stefano L. Unique Signatures of Histograms for Local Surface Description. In: Proceedings of the 11th European Conference on Computer Vision Conference on Computer Vision: Part III; 2010, September 5–11, p. 356–369. Hersonissos: Springer Berlin Heidelberg.
[7] Maver J, Bajcsy R. Occlusions as a guide for planning the next view. IEEE PAMI. 1993;15:417–433.
[8] Pito R. A solution to the next best view problem for automated surface acquisition. IEEE PAMI. 1999;21:1016–1030.
[9] Kriegel S, Rink C, Bodenmuller T, Narr A, Suppa M, Hirzinger G. Next-best-scan planning for autonomous 3d modeling. Intelligent Robots and Systems (IROS); 2012. IEEE/RSJ International Conference on, (Oct., p. 2850–2856. Vilamoura: IEEE.
[10] Blaer P, Allen PK. Data acquisition and view planning for 3-d modeling tasks. Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2007 Oct 29–Nov 2; San Diego, CA.
[11] Wong M, Dumont C, Abidi MA. Next Best View System in a 3-D Object Modeling Task. In IEEE CIRA. 1999 Nov; Monterey, CA. p. 306–311.
[12] Trummer M, Munkelt C, Denzler J. Online Next-Best-View Planning for Accuracy Optimization Using an Extended E-Criterion. In: IEEE ICPR. 2010 Aug; Istanbul, Turkey. p. 1642–1645.
[13] Hoppe H, DeRose T, Duchamp T, McDonald J, Stuetzle W. Surface Reconstruction from Unorganized Points. Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques; 1992; Chicago, IL, p. 71–78.
[14] Levin D, Brunnett G, Hamann B, Müller H, Linsen L. Geometric modeling for scientific visualization. Berlin, Heidelberg, New York: Springer-Verlag; 2004, p. 37–49.
[15] Alexa M, Behr J, Cohen-Or D, Fleishman S, Levin D, Silva CT. Computing and rendering point set surfaces. IEEE Trans. Visualization Comput. Graphics. 2003;9:3–15.
[16] Guennebaud G, Gross M. Algebraic Point Set Surfaces. Proceedings of the International Conference on Computer Graphics and Interactive Techniques. 2007; San Diego, CA.
[17] Rusu RB, Marton ZC, Blodow N, Dolha M, Beetz M. Towards 3D Point cloud based object maps for household environments. Rob. Auton. Syst. 2008;56:927–941.
[18] Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM. 1981;24:381–395.
[19] Arya S, Mount DM, Netanyahu NS, Silverman R, Wu A. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. J. ACM. 1998;45:891–923.
[20] Chen L-C, Nguyen T-H, Lin S-T. Viewpoint-independent 3-D object segmentation for randomly stacked objects using optical object detection. Measurement Sci. Technol. 2015;26:105202.
[21] Suzuki S, Abe K. Topological structural analysis of digitized binary images by border following. Comput. Vision, Graph. Image Process. 1985;30:32–46.
[22] Bendels GH, Schnabel R, Klein R. Detecting holes in point set surfaces. J. WSCG. 2006;14:1–3.
[23] Gåsvik kJ. Optical metrology. 3rd ed. Chichester: Wiley; 2002, p. 186–204.
[24] McNeill SR, Sutton MA, Miao Z, Ma J. Measurement of surface profile using digital image correlation. Exp. Mech. 1997;37:13–20.
[25] Dai HJ, Su XY. Shape measurement by digital speckle temporal sequence correlation with digital light projector. Opt. Eng. 2001;40:793–800.