Trajectory Planning for Surface Following with a Manipulator under RGB-D Visual Guidance

Danial Nakhaeinia, Raouf Fareh, Pierre Payeur, and Robert Laganière
School of Electrical Engineering and Computer Science
University of Ottawa
Ottawa, ON, Canada
[dnakhaei, rfareh, ppayeur, laganier]@uottawa.ca

Abstract — **This paper introduces a manipulator robot surface following algorithm using a 3D model of vehicle body panels acquired by a network of rapid but low resolution RGB-D sensors. The main objective of this work is to scan and dynamically explore regions of interest over an automotive vehicle body under visual guidance by closely following the surface curves and maintaining close proximity to the object. The work is motivated by applications in automated vehicles inspection and screening in security applications. The proposed path planning strategy is developed based on a perception-modeling-planning-action approach. Raw data rapidly captured by a calibrated network of Kinect sensors are processed to provide the required 3D surface shape of objects, normal measurements, orientation estimation, and obstacle detection. A robust motion planning method is designed that relies on this information, resulting in a safe trajectory that is planned to follow and explore the curved surfaces while avoiding collision with protruding components of the vehicle. The feasibility and effectiveness of the proposed method is validated through experimental results with a 7-DOF manipulator navigating over automotive body panels.**

Keywords—vehicle security screening, path planning, RGB-D sensors, surface following, robot visual guidance.

I. INTRODUCTION

Detecting explosive and other dangerous or prohibited substances in an automated and efficient manner became a critical asset to ensure the security of people and properties worldwide. Public institutions such as government buildings, research centers, military bases, airports, and critical infrastructures such as power plants or communication centres, require vehicle or individual screening systems at their periphery. While very efficient technologies exist to detect the presence of minute amounts of dangerous material particles, the process of collecting them safely and efficiently, without direct human intervention, remains a challenge. This research develops efficient and automated sampling procedures through the use of adaptive robotic technologies driven by multi-modal sensors to fully automate the particles collection while making it safer for the operators. The technology automatically adjusts the screening procedure to the diversity of shapes and sizes that characterize vehicles of various types, brands and sizes. It aims at providing authorities with pre-event screening mechanisms for prohibited substances that are versatile, safe for inspectors, easy to use, and permit automated scanning of large vehicles in time critical applications.

Different types of sensors have been employed for 3D surface reconstruction, such as: active sensors [1], passive camera [2], online images [3], or from unordered 3D points [4, 5]. A combination of a 3D laser sensor and a line-base SLAM algorithm is used in [6] to produce 2D line maps of highly cluttered indoor environments. An important direction of research is to replace commonly used 2D laser range sensors by 3D perception. A stereo vision system is presented in [7] for the extraction of the three-dimensional structure of a scene using depth and image information, providing good results but at the price of a significant computational load. Also classical stereo vision is prone to important feature matching constraints which often lead to low density depth maps. Unfortunately, most of the sensors used in previous works are not capable to capture color and a depth image simultaneously, and require a considerable amount of time to perform range acquisition over an entire surface. To overcome these limitations, several attempts have been made to capitalize on the recently introduced RGB-D Microsoft Kinect sensor, which can generate high density depth maps and corresponding color images in a fraction of a second. In the context of vehicle security screening considered here, which imposed tight time constraints, this technology can be beneficial. However, typical vehicles size largely exceeding the field of view of any RGB-D sensor, a multi-view vision system is required to rapidly acquire and reconstruct a full scale model of a vehicle under inspection. The compromise for speed of acquisition is a relatively low resolution of the depth maps, which must be dealt with at the path planning stage.

Three-dimensional data collected by sensors have often been used to support robot path planning. Similarly, several approaches have been proposed to solve the path planning problem, including cell decomposition techniques, skeletons, and potential fields [8, 9]. In addition, soft computing approaches, such as neural networks [10], fuzzy logic [11], and evolutionary algorithms [12], have been developed. Using sensory information, various path planning strategies were proposed such as: coverage path-planning [13, 14], grid [15], and optimal search strategies [16]. Where a priori information is available, such as sensor models, environmental conditions, and prior measurements, optimal search strategies are preferred. When the problem consists of detecting a target in a region of interest, probabilistic deployment approaches [17, 18] can be used to compute a search path based on the probability of finding a target in every unit bin of a discretized workspace.

This work introduces a fully automated vision-guided global path planning method adapted to explore and scan regions of interest over a vehicle body using a rapidly reconstructed 3D model of the vehicle. A calibrated network of Kinect sensors is used to rapidly acquire color and depth information which is processed to generate a 3D model of the vehicle, referenced to the base frame of a manipulator. Regions of interest that support the vehicle screening operation are extracted from the 3D model before being further processed to plan a path and guide the manipulator to scan the surface of the vehicle while closely following its aesthetic curves.

The paper is organized as follows: Section II gives an overview of the proposed approach. Section III presents the design procedure and details the implementation of the path planning strategy. Experimental results are shown in section IV, and a final discussion concludes the paper in Section V.

II. PROBLEM FORMULATION

Multiple RGB-D sensors are selected and organized in a distributed sensor network to achieve full coverage and rapid 3D profiling of automotive vehicles [19]. The main goal of the vision stage is to support the motion of a manipulator's end effector to operate in close proximity of a vehicle while safely accommodating its curves and inherent surface obstacles, such as outside mirrors or door handles, in order to perform an inspection task. Figure 1 shows the layout of the proposed automated vehicle screening station. The vehicle is stopped in front of a set of Kinect RGB-D sensors while a manipulator is moved aside on a linear track to avoid creating any occlusion during the acquisition phase. The color and depth information is collected over a 180 degrees view of the vehicle (one side) by five calibrated Kinect sensors within only a few seconds. Rapid acquisition is essential to ensure high inspection rates. For the other side of the vehicle, the same setup can be duplicated. Finally, a textured 3D model of the vehicle is obtained from the information given by the RGB-D imaging system.
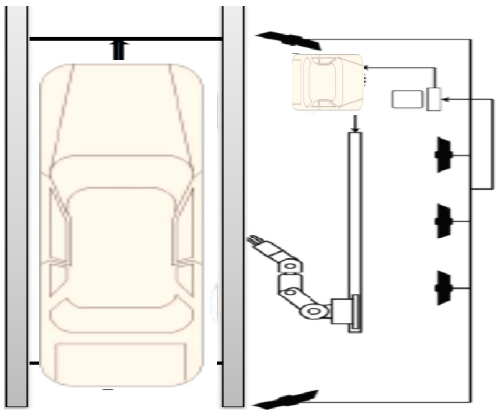

Figure 1. Vehicle screening station.

III. DESIGN OF THE PATH PLANNER

Four steps are performed to control the robotic manipulator using the textured 3D model. First, the network of Kinect sensors is used to capture color images of the scene and the depth of each pixel in the workspace. The data is processed to extract the 3D information over each region of interest.

Second, the regions of interest are defined according to critical parts of the vehicle that require inspection. Third, the data is processed and analyzed to plan a trajectory which ensures the entire scanning of each region of interest. Finally, the normal directions of the vertices in the triangular mesh of the surface model are calculated along the desired path to estimate the local surface orientation that will determine the orientation of the end effector along the path. As a result, a set of points are computed in the operational space (Cartesian coordinates) to guide the robot for scanning and exploring each region of interest while following the curves and remaining close to the object. The series of end-effector's positions and orientations are then mapped to the joint space using inverse kinematics and sent to the controller to perform the surface following operation (Fig. 2).
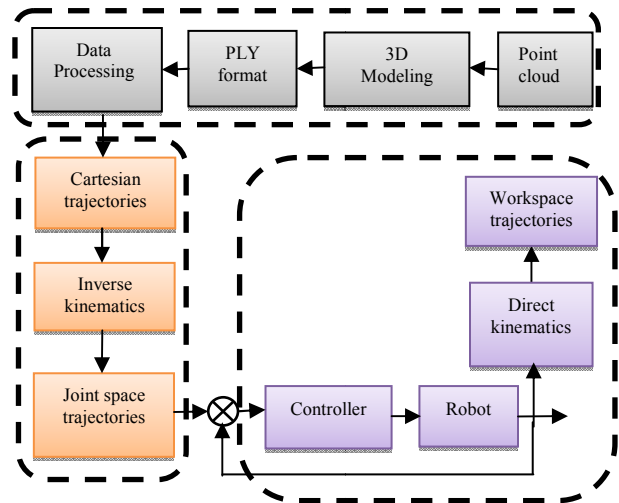

Figure 2. System design algorithm.

A. Vision stage

In the first step a set of five Kinect sensors organized as a calibrated distributed sensor network [19] are used to simultaneously collect color and depth information over one side of the vehicle under screening. Sections of the depth information on one side of the vehicle are aligned and merged into a global RGB-D model representing the bodywork panels. Three Kinect sensors are positioned in a single row (sensor's baseline) with their viewing axes perpendicular to the bodywork and two extra sensors are used for collecting data over the front and back parts of the vehicle. The raw information is stored in a textured 3D point cloud form. Given that the vision stage is also calibrated with the robot's base reference frame during the setup procedure, all the points are defined by coordinates in the Cartesian space with respect to the robot's base reference frame, which allows for path planning and guidance of the manipulator from the visual stage. Using the color information from the RGB-D model, a collection of reference parts (wheels, door handles, roof, windows, mirrors, etc) over the bodywork of the vehicle are automatically extracted using a previously proposed visual detector of vehicle parts [20]. The detector makes use of Haar-like features and a classifier trained over a training set of images containing various types and models of vehicles.

A statistical analysis of the location of significant vehicle features allows for the recognition of the main reference parts, as shown in Fig. 3, from which regions of interest that are relevant for security screening can be defined over the vehicle, e.g. surroundings of door handle, trunk seam. These areas are later selected as priority regions for the robot to closely inspect by close imaging or contact.
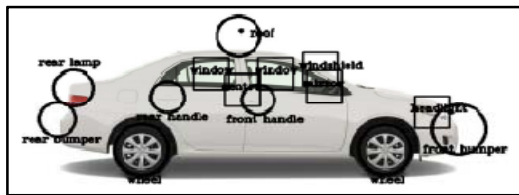

Figure 3. Detection of parts of interest over a car.

B. 3D modelling

In the second stage, the shape over selected regions is reconstructed using the depth information from the global RGB-D model. Figure 4 illustrates the procedure for 3D reconstruction of regions of interest over the vehicle lateral panels. The locations of the detected regions of interest in the color images are mapped on the RGB-D model of the vehicle to provide 3D shape and location information about sections of the bodywork panels, which is ultimately used to navigate the manipulator within the regions. 3D points contained within each of the selected regions of interest are extracted as a separate point cloud. The coordinates of the points provide the width and height of a 3D bounding box enclosing the desired areas. A fixed parameter (here 30 cm) is added and subtracted to the average depth of the points to define the bonding box depth. Next, a triangle mesh is built over the respective groups of point cloud that correspond to each region of interest detected over the surface and these are stored in a PLY format file. A triangle mesh of a door model is shown in Fig. 5.
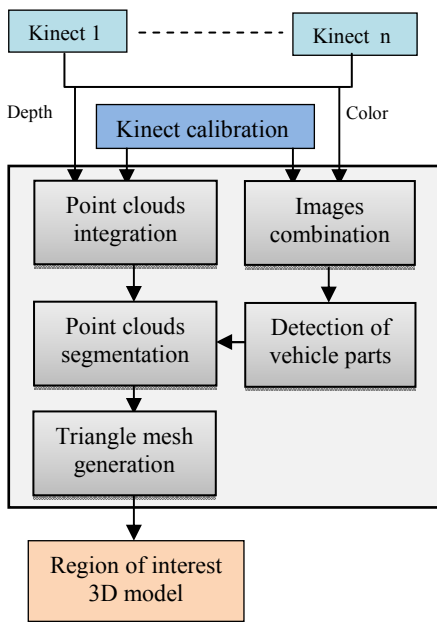

Figure 4. 3D reconstruction of a region over the vehicle lateral panel.

Figure 5. Triangle mesh of a door model.

C. Path planning for surface following

The objective of the robot path planning is to allow the manipulator's end effector to efficiently scan a given region of interest previously identified and modeled in 3D. For that purpose, the triangle surface mesh of the region is first processed to extract its vertex list and face list. Next the planning of the trajectory for scanning the whole region of interest is executed. Then the calculation of the normal to each vertex that is part of the planned trajectory is performed. Finally the orientation of the surface area that will be scanned is estimated to compute the orientation of the end-effector at each location while meeting the desired points coordinates over the path.

3D Data Processing: The 3D data information over regions of interest is used to plan a path that will closely follow the surface curves of the body panel, ensuring the integrity of the vehicle while permitting close visual inspection with an eye-in-hand camera or contact for particles collection. The 3D surface mesh of a region of interest contains a collection of vertices and faces which are called the vertex list and the face list. The vertex list (Table1 shows a set of five vertices as an example) is a list of (x, y, z) triplets for vertices with respect to the robot base frame. The face list (Table 2) contains the number of vertices that form the face in the first column, followed by the vertex indices that compose each face. The vertex list and face list are used jointly to plan a path for the manipulator and simultaneously form the 3D triangle mesh of reference to estimate the local normal and orientation of the surface which are also used to control the orientation of the manipulator (Fig. 6).
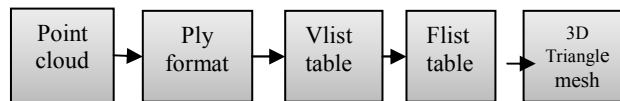

Figure 6. PLY data processing.

Table 1: _Vertex list_.

x	y	z
902.709	1549.53	136.944
902.662	1549.36	133.364
902.828	1545.94	137.118
902.616	1549.18	129.783
902.57	1549.01	126.202

Table 2: *Face list.*

n_p	P_1	P_2	P_3
3	0	1	2
3	1	3	2
3	1	4	3
3	4	5	3
3	4	6	5

Trajectory planning: In order to scan the region of interest, the trajectory of the end effector is planned using the 3D model defined in the previous section. As shown in Fig. 7, it is assumed that each zone of interest can be bounded in a rectangular box delimited by a set of points at the edge of the surface (Y_{min}, Y_{max}, Z_{min}, Z_{max}), corresponding respectively to the minimum and maximum values of the Y and Z coordinates in the vertex list. To scan and explore the whole region, the end-effector center position at each step is defined by a set of points according to the end-effector coverage zone, which corresponds to the size of the tool attached on the robot to perform the screening operation. In this work the manipulator is considered to cover a circular area with radius of r = 50 mm.

First a start point is defined as the closest vertex to the upper-left corner of the rectangular bounding box. Since the vertices are not stored in a specific order in the list, the start point is determined by computing the distance of all vertices from the upper left corner. The vertex with the minimum distance is selected as the start point. Then the robot moves horizontally towards the next point which is within a distance 2r from the previous point along the Z-axis until it reaches to the surface edge on the right side of the rectangular bounding box (Z_{max}). The surface edge is identified by checking the neighbour vertex at each step. Once the neighbour vertices are no longer within the rectangular area, the robot moves vertically down and the next position of the end-effector is the point which is within a distance 2r from the previous location of the robot along the Y-axis. Then the robot again moves horizontally but in the opposite direction until it reaches the surface edge on the left side of the rectangular bounding area (Z_{min}). The process continues until the robot has scanned the entire area.
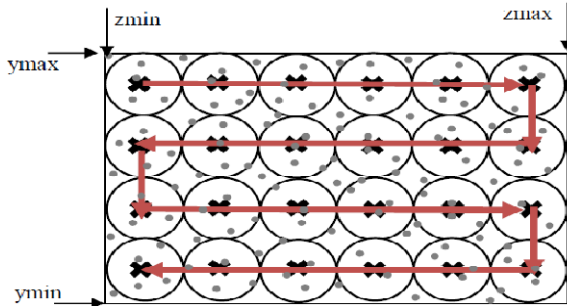

Figure 7. Trajectory planning over the region of interest.

Normal calculation: The set of points forming a path for the robot defined in the previous step determine the positions for the end effector to reach in Cartesian coordinates. However, to ensure proper alignment of a close inspection camera or accurate contact over the vehicle bodywork, the local orientation of the end effector must match that of the normal to the surface of the vehicle at every point. This is especially important given the typical aesthetic curves and significant changes of orientation that exist over any type of vehicle. Therefore, it is required to calculate the normal to the surface and to estimate the local surface orientation in order to compute the proper end-effector orientation and ensure precise surface following. Using the vertex list and the face list defined earlier, a triangle mesh of the surface is formed. The normal of each triangle is computed as the cross product between the vectors representing two sides of the triangle. The probability of having a normal in one direction is the same as having it in the opposite direction, which depends on the cross product order (Fig. 8). To solve this issue, the direction of the normal is imposed to be in the same direction for all triangles, which corresponds to the orientation of the surface of the vehicle when viewed from the outside.

The following equations define the normal vector, N, calculated from a set of three vertices, A, B, C, point coordinates:

$$N = AC \times AB \tag{1}$$

where $N = [N_x \quad N_y \quad N_z]$ and

$$N_x = AC_y \cdot AB_z - AC_z \cdot AB_y \tag{2}$$

$$N_y = AC_z \cdot AB_x - AC_x \cdot AB_z \tag{3}$$

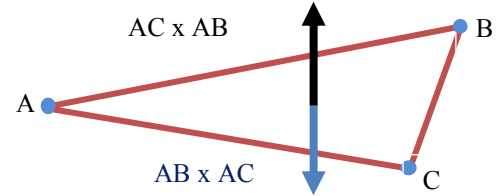$$N_z = AC_x \cdot AB_y - AC_y \cdot AB_x \tag{4}$$

Figure 8. Triangle normal vector calculation.

The resulting normals are then normalized such that the length of the edges does not come into account. Since the planned path is a set of vertices from the vertex list, it is desired to calculate an orientation corresponding to each vertex. A vertex normal is the normalized average of all the triangle normals that contain that vertex. However, since the data represented by the Kinect tends to be noisy and discretized, especially on depth measurements, the normal tends not to be accurate enough to properly estimate the orientation in all locations. In order to smooth out the normal estimate, the vertex normal is calculated for all the local vertices which are covered by the end-effector, as defined by its radius, r, at each step. The average of the involved vertex normals is then calculated and normalized to represent the overall direction of the surface in the local neighborhood of the corresponding vertex along the planned trajectory. The normal vector is finally used to calculate the rotation matrix and compute the RPY angles that define the

surface orientation with respect to the robot base frame. The orientation of the end effector for a given location along the trajectory is set as the opposite to the local surface orientation vector such that the vehicle panel surface and the robot end effector directions are opposite to each other.

D. Action

As a final stage, the Cartesian space trajectory which defines the position and orientation for the end-effector, as determined in the previous section, is transformed to the joint space of the robot via its inverse kinematics. Then, the joint space trajectory is sent to the controller to perform the surface following over the region of interest on the vehicle. The entire procedure is repeated for every region of interest extracted from the RGB-D model of the vehicle.

IV. EXPERIMENTAL RESULTS

To validate the feasibility of the proposed path planning method and to assess the accuracy of the end-effector motion in accordance with the surface of automotive panels, experiments are carried out with a 7-DOF CRS F3 manipulator which consists of a 6-DOF robot arm mounted on a linear track. The latter provides an additional axis of motion and an expanded working area. The resulting redundancy is also useful to prevent some singularities. In these experiments a 3D model of a car door (Fig. 9) is constructed using the color and depth information collected with the network of Kinect sensors (here involving 3 lateral sensors). It is desired to closely scan the whole door with robot except for the window area since glass does not pick well with Kinect sensor technology. The considered region of interest over the door is stored as a 3D triangular mesh.
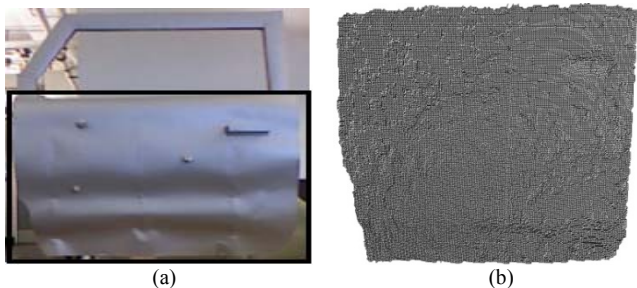

(a)   (b)

Figure 9. a) Region of interest over a car door, b) 3D triangular mesh.

The start point at the upper-left corner of the surface is detected (Fig. 10a). Then the robot follows the planned trajectory while remaining in close proximity to the surface (a safety distance of 50 mm is considered between the end-effector and the surface) and tracking the surface curves (Fig. 10b and 10c). Figure 11a shows the robot trajectory in Cartesian space which successfully tracks the desired path and scan the whole surface smoothly. The blue line shows the desired path (y_d, z_d) and the red dashed line shows the actual robot trajectory (y,z). Note that the desired path is planned based on 3D model of the door collected by the vision sensors, as shown in Fig. 9b, which ensures the proper alignment of the planned trajectory over the surface of the panels. Therefore, the error, (e_y, e_z), between the desired path and the robot trajectory (Fig. 11b) represents

the end effector evolution against the actual position and shape of the panel. The set safety margin is not considered here for clarity.
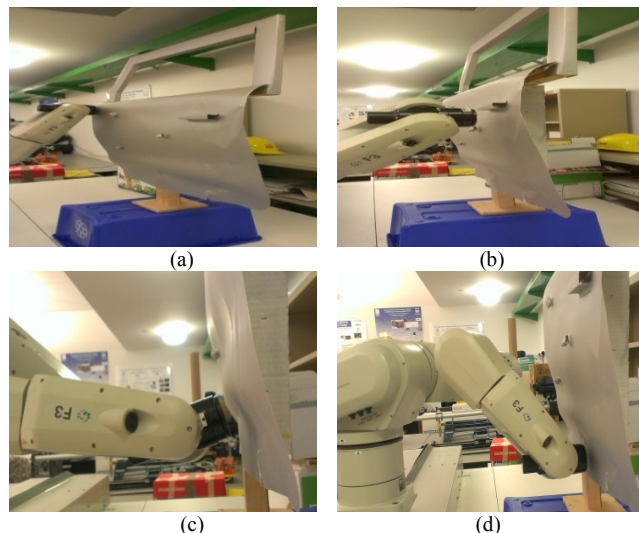

(a)   (b)

(c)   (d)

Figure 10. Robot performance at following the panel's curved surface and matching its position and orientation.
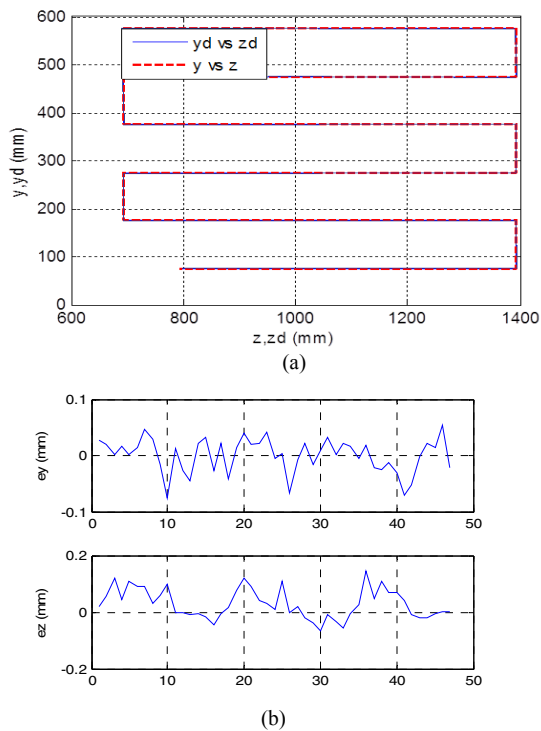

(a)

(b)

Figure 11. a) Cartesian space tracking trajectory, b) tracking error in the Cartesian space.

Figure 12a shows the robot performance at following the desired path in the joint space, that is the desired value, q_{di}, of the joints versus the real joints value, q_i, after each step. As shown in Fig. 12b, the error in the joint space is also very small. Despite the low resolution 2.5D data provided by the network of Kinect sensors, the region of interest over the car door is fully scanned while its aesthetic curves are accurately followed.
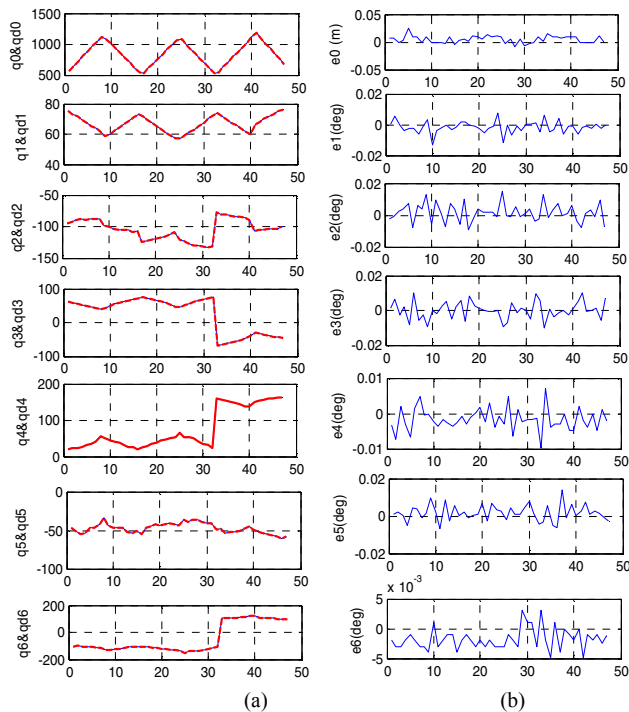
Figure12. a) Joint space tracking, and b) tracking error in the joint space.

V. CONCLUSION

This work presented a global trajectory planning method to automatically screen large objects such as automotive vehicles using rapid RGB-D acquisition and 3D modeling of the object to support close surface shape following with a manipulator. A network of Kinect sensors distributed around the object collects color images and depth information over the lateral panels of a vehicle to reconstruct the 3D model of the object. The model is then used to automatically select regions of interest over the vehicle for scanning. In spite of the inherent low resolution of Kinect sensors, the 3D data provides sufficiently accurate spatial information about the bodywork of the vehicle to plan a precise trajectory for the robot to closely interact with the automobile panel while performing a close inspection procedure. The proposed method is experimentally validated on a 7-DOF CRS F3 manipulator. The experimental results show that the robot successfully scans the whole region of interest and closely follows the curvatures on the surface of an automotive vehicle body panel. Future work will involve performing similar operation over various areas of a vehicle, including door handles and trunk that exhibit more pronounced orientation variations, therefore making the proposed approach appropriate to closely follow any smooth surface shape within the identified regions of interest.

REFERENCES

[1] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, "The digital Michelangelo project: 3D scanning of large statues," *27th Annual Conf. on Computer Graphics and Interactive Techniques*, pp. 131-144, 2000.

[2] P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, "Real-time visibility-based fusion of depth maps," *IEEE 11th Intl Conf. on Computer Vision*, pp. 1-8, 2007.

[3] J.-M. Frahm, P. Fite-Georgel, D. Gallup, T. Johnson, R. Raguram, C. Wu, "Building Rome on a cloudless day," European Conf. on *Computer Vision*, Springer (Ed), pp. 368-381, 2010.

[4] M. Kazhdan, M. Bolitho, H. Hoppe, "Poisson surface reconstruction," *4th Eurographics Symp. on Geometry Processing*, 2006.

[5] K. Zhou, M. Gong, X. Huang, B. Guo, "Data-parallel octrees for surface reconstruction,", *IEEE Trans. on Visualization and Computer Graphics,* vol. 17, pp. 669-681, 2011.

[6] O. Wulf, K. O. Arras, H. I. Christensen, B. Wagner, "2D mapping of cluttered indoor environments by means of 3D perception," *IEEE Intl Conf. on Robotics and Automation*, pp. 4204-4209, 2004.

[7] M. Z. Brown, D. Burschka, G. D. Hager, "Advances in computational stereo," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 25, pp. 993-1008, 2003.

[8] A. Hourtash, M. Tarokh, "Manipulator path planning by decomposition: algorithm and analysis," *IEEE Trans. on Robotics and Automation,,* vol. 17, pp. 842-856, 2001.

[9] E. Rimon, D. E. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Trans. on Robotics and Automation,* vol. 8, pp. 501-518, 1992.

[10] I.-K. Jung, K.-B. Hong, S.-K. Hong, S. C. Hong, "Path planning of mobile robot using neural network," *IEEE Intl Symp. on Industrial Electronics*, pp. 979-983, 1999.

[11] H. Surmann, J. Huser, J. Wehking, "Path planning for a fuzzy controlled autonomous mobile robot,", *IEEE Intl Conf. on Fuzzy Systems*, pp. 1660-1665, 1996.

[12] C. Hocaoglu, A. C. Sanderson, "Planning multiple paths with evolutionary speciation,", *IEEE Transactions on Evolutionary Computation,* vol. 5, pp. 169-191, 2001.

[13] E. U. Acar, H. Choset, Y. Zhang, M. Schervish, "Path planning for robotic demining: Robust sensor-based coverage of unstructured environments and probabilistic methods," *Intl Journal of Robotics Research,* vol. 22, pp. 441-466, 2003.

[14] H. Choset, "Coverage for robotics–A survey of recent results," *Annals of Mathematics and Artificial Intelligence,* vol. 31, pp. 113-126, 2001.

[15] X. Liao, L. Carin, "Application of the theory of optimal experiments to adaptive electromagnetic-induction sensing of buried targets," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 26, pp. 961-972, 2004.

[16] J. R. Spletzer, C. J. Taylor, "Dynamic sensor planning and control for optimally tracking targets," *Intl Journal of Robotics Research,* vol. 22, pp. 7-20, 2003.

[17] C. Hofner, G. Schmidt, "Path planning and guidance techniques for an autonomous mobile cleaning robot," *Robotics and Autonomous Systems,* vol. 14, pp. 199-212, 1995.

[18] M. Qian, S. Ferrari, "Probabilistic deployment for multiple sensor systems," *Smart Structures and Materials*, pp. 85-96, 2005.

[19] R. Macknojia, A. Chávez-Aragón, P. Payeur, R. Laganière, "Calibration of a network of Kinect sensors for robotic inspection over a large workspace", *IEEE Workshop on Robot Vision*, pp. 184-190, Jan. 2013.

[20] A. Chávez-Aragón, R. Laganière, and P. Payeur, "Vision-based detection and labelling of multiple vehicle parts," IEEE *Intl Conf. on Intelligent Transportation Systems*, pp. 1273-1278, 2011.