Active Vapor-Based Robotic Wiper

Takuya Kiyokawa*, Hiroki Katayama, Jun Takamatsu, and Kensuke Harada

Abstract—This paper presents a method for estimating normals of mirrors and transparent objects challenging for cameras to recognize. We propose spraying water vapor onto mirror or transparent surfaces to create a diffuse reflective surface. Using an ultrasonic humidifier on a robotic arm, we apply water vapor to the target object’s surface, forming a cross-shaped misted area. This creates partially diffuse reflective surfaces, enabling the camera to detect the target object’s surface. Adjusting the gripper-mounted camera viewpoint maximizes the extracted misted area’s appearance in the image, allowing normal estimation of the target surface. Experiments show the method’s effectiveness, with RMSEs of azimuth estimation for mirrors and transparent glass approximately 4.2° and 5.8°, respectively. Our robot experiments demonstrated that our robotic wiper can perform contact-force-regulated wiping motions to clean a transparent window, akin to human performance.

I. INTRODUCTION

For robots to be useful in home environments, they must recognize a wide range of objects. The robot must identify target object surfaces accurately to avoid damaging fragile items. Recognizing the plane normal of mirrors and transparent objects is challenging due to specular reflection or light transmission on non-Lambertian surfaces. To enable robots to clean windows and mirrors, a system that identifies and wipes these surfaces without specific machines [1], [2] is needed.

This study proposes using water vapor to enhance the visibility of mirrors and transparent surfaces for plane normal estimation. This is the first attempt to develop a robotic wiper system based on the Active Vapor method, shown in Fig. 1. The system creates a misted area by spraying water vapor with an ultrasonic humidifier attached to a robot arm gripper. The robot then wipes the surfaces based on the estimated plane normal with regulated contact force. Experiments demonstrated that the robotic wiper accurately estimated plane normals for mirrors and transparent glass and performed effective wiping. The method estimated the azimuth angle with errors less than 4.2° for mirrors and 5.8° for transparent glass.

II. ACTIVE VAPOR-BASED NORMAL ESTIMATION

A. Estimating Normal Direction from Misted Cross-Shape

The misted area appears different from its surroundings due to its blurred, diffusely reflected nature. By shaping the misted area into a cross, we can estimate the normal direction from the line segment lengths, making precise boundaries unnecessary. A vapor-covered line segment suffices for normal estimation, suggesting traditional image segmentation methods are suitable.

The azimuth angle \( \theta \) and elevation angle \( \phi \) can be estimated based on the tilt angle geometry of the target plane and the cross-shaped misted area. As both angles are calculated similarly, the azimuth angle calculation is exemplified. Fig. 2 (a) illustrates the azimuth angle geometry. If the vertical and horizontal axis lengths in the cross-shaped misted area are \( 2l \), determined by the robot arm trajectory length, the angle value is calculated accordingly.

\[
\theta = \pm \cos^{-1} \frac{2l}{L_{\text{azimuth}}}.
\]

Fig. 1: Robotic window-wiping system based on Active Vapor. The system has three functions: spraying water vapor, grasping an object, and capturing an image.

Fig. 2: Normal estimation method. (a) shows an overview of the gripper motion and geometries between the gripper and the misted plane. The top part of (b) shows the reflection surface that appears on the target plane. The bottom part of (b) shows the geometry of the cross-shaped misted area viewed from the viewpoint 1.
This solution is not uniquely determined. It is assumed that the calculated misted area length $L_{azimuth}$ is maximized when the target plane is perpendicular to the camera’s image-capturing direction. The azimuth angle is estimated using this geometric constraint.

First, by adjusting the camera viewpoint with the robot arm after water vapor spraying, we identify the camera direction $i$, where $L_{azimuthi}$, $(i = 1, 2, .., n)$ reaches its maximum, with $n$ representing the number of viewpoint adjustments made by the robot arm. Second, the angle between the estimated and initial directions before adjustment is determined as the estimated azimuth $\theta$. As depicted in Fig. 2 (b), top, $L_{azimuth}$, is derived by first finding the intersections of the major axes with the misted area’s contours, then measuring the distance between these intersections.

With the estimated elevation angle $\phi$, the normal $N$ at the sprayed plane’s initial position is calculated as

$$N = (\cos \phi \cos \theta, \ cos \phi \sin \theta, \sin \phi). \ (2)$$

B. Spraying Water Vapor

Fig. 3 outlines the proposed method. Water vapor is sprayed as the robot gripper nears the target plane. The robot arm moves the vapor to create a cruciform misted area. This cross-shaped mist helps determine azimuth and elevation angles from the line lengths.

The robot arm executes reciprocating linear motions up/down and left/right from the initial position to form the misted cross. Post-spraying, the robot arm moves along arcs at a speed ensuring the arc movement is completed in less than $T_{completion}$[s], satisfying the following equation:

$$\frac{r \beta}{v_{wrist}} < T_{completion}, \ (3)$$

where $r$ [mm] is the distance to the target plane, $\beta$ [rad] is the rotation angle of the robot gripper when estimating the normal after completing water vapor spraying, and $v_{wrist}$ [mm/s] is the circumferential speed at that time.

III. Wiping System Implementation

A. Active Vapor System

Fig. 1 (a) displays the robot gripper with an Active Vapor sensing system. A two-fingered gripper (SAKE Robotics, EZGripper) serves as the end-effector, allowing the robot to grasp a wiping tool like a washcloth or sponge. A small ultrasonic humidifier atomizes the water, and the vibrator’s ultrasonic waves cause the water to be ejected from the donut-shaped tip. An RGB camera, mounted at the bottom of the gripper, captures the target plane misted by the vapor. The camera operates at HD resolution (1280 x 720) and 30 fps. All devices are controlled by a single computer.

The camera captures the target plane, and the misted area is isolated using GrabCut [3]. In the wiping experiments, the misted area’s shape, identified as a cross after GrabCut processing, is used to estimate the normal of the misted surface.

B. Contact-Regulated Robotic Wiper

Stable contact with the target object’s surface is necessary to effectively wipe off dirt. Instead of directly controlling the contact force, we regulate it within an empirically determined range with minimum and maximum normal force thresholds of 3.0 and 8.0 [N]. In our regulation process, the normal force $F_n$ is calculated as $-f_z$, representing the contact force in the z-axis direction.

Experiments utilized a contact-force-regulated robotic wiping system. The contact force $f_z$ is the forward force when the robot approaches the target, derived from the torque sensors in the arm’s six joints. Force values were obtained using iiwa_stack1, the ROS package for the KUKA LBR iiwa 14 R820 robot arm used in our experiments.

IV. Experiments

To assess the effectiveness of the Active Vapor method, we conducted plane normal estimation for both a mirror and a transparent glass window. The experimental setup, depicted in Fig. 1 (b), involved attaching the sensor system to a seven-degree-of-freedom robot arm. An automatic rotating stage adjusted the target plane’s azimuth angle. We used a 3mm thick mirror and glass, as the system cannot vaporize surfaces thicker than 3mm. Only the azimuth angle was estimated, varying by $-20^\circ$, 0$^\circ$, and $20^\circ$, with three trials per pattern per object. Mist generation was challenging beyond 100mm from the target, so $r$ was set to 100mm. From this distance, the robot moved linearly with three vertical and horizontal reciprocations.

We performed ten trials to measure the time until the mist disappeared. Room temperatures were 23°C, 26°C, and 22°C, respectively. The misted area vanished within approximately 6 seconds. Based on these results, a circular arc trajectory for the gripper was designed to complete within $T_{completion} = 6s$. The robot gripper speed was set to 50mm/s to ensure the spraying and image capture were completed before the mist dissipated, exceeding the minimum calculated speed of 17mm/s.

1https://github.com/IFL-CAMP/iiwa_stack
A. Plane Normal Estimation Accuracy

We trained the GrabCut extraction model using 100 images from the same experimental environment. The RMSE of the azimuth estimation from $-20^\circ$ to $20^\circ$ was $4.2^\circ$ for the mirror and $5.8^\circ$ for the transparent glass window. These results indicate that the proposed Active Vapor-based sensing system, when attached to the robot, can accurately estimate the plane normal.

The smaller estimation error on the mirror surface is likely due to the laboratory setting, where few characteristic objects were in the background of transparent objects, making it challenging to extract the misted area. Conversely, the mirror surface reflected the robot gripper or arm, providing multiple objects with visual characteristics, allowing GrabCut to learn the misted area as the foreground based on refraction.

B. Evaluating Window Wiping

1) Wiping Accuracy: The wiping performance was evaluated by manually painting water-based blue ink on a predefined window area of approximately 150mm by 20mm. The robot performed three round trips along a straight line (Fig. 4 (b)). The remaining ink area was measured to determine performance, defined as $A_{\text{unwiped}}$ [mm$^2$]. The relationship between window-wiping quality and azimuth angle estimation error was verified. The unwiped area $A_{\text{unwiped}}$ [mm$^2$] is calculated as:

$$A_{\text{unwiped}} = A_{\text{initial}} \times \frac{N_{\text{final}}}{N_{\text{initial}}},$$

where $A_{\text{initial}}$ [mm$^2$] is the initial ink-stained area (5000 mm$^2$), and $N_{\text{initial}}$ and $N_{\text{final}}$ are the pixel counts of the ink-stained area before and after wiping, respectively.

The end effector moved along a straight line (Fig. 4 (b)). The ink-stained area was extracted using GrabCut. We evaluated human wiping and robot wiping with a $5.8^\circ$ error, comparing the wiped area $\alpha$ [%] for both cases.

$$\alpha = \frac{(A_{\text{initial}} - A_{\text{unwiped}})}{A_{\text{initial}}} \times 100.$$  \hspace{1cm} (5)

Fig. 5 presents the experimental results. With an azimuth angle estimation error of $5.8^\circ$ for the transparent glass window, the mean $\alpha$ across three trials was 45.2%. The regulated contact force of the robot gripper enhanced the accuracy of the robot’s wiping motions on the ink-stained surface, improving $\alpha$ to 52.2%. When performed by a human, the mean $\alpha$ was 65.1%. Thus, using an estimation error of $5.8^\circ$, approximately 52.2% of the ink can be removed.

2) Wiping Efficiency: We evaluated the efficiency of the proposed wiping system by measuring the reduction in time for wiping operations. We tested normal estimation with various spraying and image-capturing motions in a shorter duration. Our experiments indicated that a one-way spraying motion under 3 seconds typically fails, with failure defined as a misted area of nearly zero, confirmed by visual inspection. While results vary with camera frame rate, image-capturing motions under 3 seconds generally fail, when using a camera with a 30 fps frame rate.

3) Wiping Effectiveness: To validate the proposed wiping system’s effectiveness, its accuracy was compared. Since accurately extracting misted areas based on the reflected scene and background was challenging, we tested the normal estimation’s robustness on various backgrounds with a transparent window.

We compared the extraction accuracy of misted areas across different backgrounds for a transparent glass window, using indoor and outdoor scenes. Indoor (room) and outdoor (garden and sky) images were sourced from the ETH3D Stereo Benchmarks [4], and outdoor videos from the VIRAT Video Dataset [5]. We used three images of different indoor scenes, three outdoor scenes, and three video frames of outdoor scenes. Fig. 6 depicts the image examples used in the experiment.

Fig. 7 compares misted area extraction across different
To enhance normal estimation, automated robot gripper and the target surface is crucial for accurate and efficient wiping motions. Leidner et al. [6] automated robot wiping using whole-body motion control, and Luo et al. [7] allowed users to adjust the end-effector’s force setpoint. Although force control is beyond this study’s scope, these studies inform flexible wiping motion. Precise contact force control can increase motion time. Future work will focus on control methods and re-estimation of the plane normal to reduce motion time.

3) Temperature: The saturated water vapor content determines an adequate misted area for normal estimation. We tested drying times at room temperatures of 20°C, 25°C, and 30°C with 65% humidity. The average drying times were 92s, 74s, and 68s, respectively, with minor, statistically insignificant differences.

4) Improved Normal Estimation: An algorithm to analyze reflections and background for optimal spraying is advisable due to the difficulty of extracting misted areas. Discussing a method for generating the minimum trajectory before the misted areas dry is also warranted. This study used a cross-shaped misted area, which is time-consuming; future work should optimize motions to save time. While this study focuses on plane normal estimation, normal estimation for curved surfaces using light-field features [8] is a promising future direction.

VI. CONCLUSIONS

This study proposes a method to estimate the normals of mirrors and transparent objects that are challenging to recognize with a camera. A sensor system on a robot arm wipes the target object, and water vapor is sprayed to create a cross-shaped misted area, which is extracted using GrabCut to estimate the plane normal. Experiments showed an average error of 4.2° for mirrors and 5.8° for transparent glass windows.

Future work will enhance the estimation system, test various home objects, and develop a contact-controlled robotic window-cleaning system robust to normal estimation errors.

REFERENCES

[1] T. Miyake, H. Ishihara, R. Shoji, and S. Yoshida, “Development of small-size window cleaning robot a traveling direction control on vertical surface using accelerometer,” in Proc. IEEE Int. Conf. Mechatron. Autom., 2006, pp. 1302–1307.
[2] N. Imaoka, S.-g. Roh, N. Yusuke, and S. Hirose, “SkyScraper-I: Tethered whole windows cleaning robot,” in Proc. IEEE/RSJ Int. Conf. Intell. Rob. Syst., 2010, pp. 5460–5465.
[3] C. Rother, V. Kolmogorov, and A. Blake, “GrabCut - interactive foreground extraction using iterated graph cuts,” ACM Trans. Graph., vol. 23, no. 3, pp. 309–319, 2004.
[4] T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger, “A multi-view stereo benchmark with high-resolution images and multi-camera videos,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2017, pp. 2538–2547.
[5] S. Oh, A. Hoogs, A. Perera, N. Cuntoor, C.-C. Chen, J. T. Lee, S. Mukherjee, J. K. Aggarwal, H. Lee, L. Davis, E. Swears, X. Wang, Q. Ji, K. Reddy, M. Shah, C. Vondrick, H. Pirisavash, D. Ramanan, J. Yuen, A. Torralba, B. Song, A. Fong, A. Roy-Chowdhury, and M. Desai, “A large-scale benchmark dataset for event recognition in surveillance video,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2011, pp. 3153–3160.
[6] D. Leidner and M. Beetz, “Infering the effects of wiping motions based on haptic perception,” in IEEE Proc. Int. Conf. Humanoid Robots, 2016, pp. 461–468.
[7] L. Lu and J. T. Wen, “Human-directed robot motion/force control for contact tasks in unstructured environments,” in Proc. IEEE Int. Conf. Autom. Sci. Eng., 2015, pp. 1165–1170.
[8] D. Tsai, P. Corke, T. Peynot, and D. G. Dansereau, “Refractive light-field features for curved transparent objects in structure from motion,” IEEE Robot. Autom. Lett., vol. 6, no. 4, pp. 6923–6930, 2021.