8th IEEE International Conference on Automation Science and Engineering
August 20-24, 2012, Seoul, Korea

Window Detection for Gondola Robot Using A Visual Camera *

Dong Yeop Kim, Jongsu Yoon, Hwadong Sun, and Chang-Woo Park

Abstract— **Maintenance of building façade is a risky operation for human labor. We have researched a robot system that substitutes for human. This robot system requires much intelligence for sensing and control. In this paper, we propose window detection methodologies as a sensing theme. It is essential for robotic façade maintenance. We have two approaches. One is the connected-component labeling, and the other is histogram of row and column direction.**

I. Introduction

Among a lot of roles of a robot, we have focused on robots working at dangerous places. For example, a robot-mounted excavator can be deployed instead of human workers. The operators just need stay in a warm or air-controlled area while enjoying some cookies and drinks. In this paper, we show another application of this approach. It is a building façade maintenance robot, especially for high rising buildings such as Empire States building and Marina Sands Bay hotel. Workers who maintain building façade use ropes to hold their body at 50 m above from the ground. It is very risky. Many accidents have occurred, and many people have died. Even though there is no accident, the horror of high place without steady supports makes the workers hurt on their mentality. Because of it, the labor cost for building façade workers is very high. Few people want to work at this dangerous job.

As shown in [1, 2], there are many types of building maintenance robots. We categorized them in two types. First is a built-in type robot. It moves along rails on building façade. The rails provide steady support for robots, and the robot can be specialized at the building. However, this type seldom gets a general design for various building. The rails for built-in type can harm the designers' concept of building, and precise installation of rails is another difficult problem for construction. Second, a gondola-typed building façade robot shows other strong points. It can accommodate itself to building that is not prepared for robots. A gondola hanger on the rooftop of the building is just needed for operation. Therefore, a service company operating some gondola-type building façade robots is expected as a new business model.

Our mission is the sensor systems and control methodology [3, 4]. The goal is the remote control and automation of the gondola robot. This concept is essential for the robot system. In this paper, we proposed another sensing theme. It is window detection. For painting or cleaning of building façade, windows are the exception of operation. If paint is applied to windows, it will require another mission that is more difficult and more sensitive. Cleaning tools for reinforced concrete walls may harm windows and cause glass cracks. As a solution, we have researched the visual sensing algorithms. In chapter 3, we show two methods. One is the connected-component labeling [5], and the other is the histograms of row and column direction.

* The work presented in this paper was funded by BMRC (Building-Façade Maintenance Robot Research Center), supported by Korea Institute of Construction and Transportation Technology Evaluation.

Dong Yeop Kim is with the Korea Electronics Technology Institute (KETI), Bucheon-si, Gyeonggi-do, Republic of Korea (corresponding author to provide phone: +82-32-328-2859; fax: +82-32-621-2866; e-mail: sword32@keti.re.kr).

Hwadong Sun is the Korea Electronics Technology Institute (KETI), Bucheon-si, Gyeonggi-do, Republic of Korea (e-mail: elysun80@gmail.com).

Chang-Woo Park is with the Korea Electronics Technology Institute (KETI), Bucheon-si, Gyeonggi-do, Republic of Korea (e-mail: drcwpark@keti.re.kr).

II. Robot System for Building Maintenance

A. Scenario

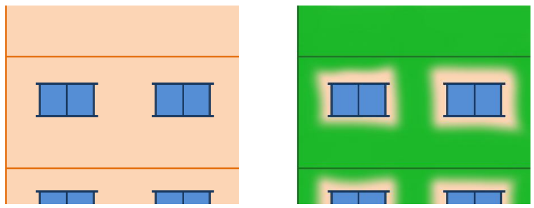
In order to assist understanding of the proposed methods, our scenario need be ascertained. Following scenario is applied to the gondola system, but the built-in system can also be applied easily.

After the gondola arrives on a site, the gondola hanger is installed on the rooftop. Wire ropes are hanged down to ground, and the gondola cage is secured with it. Then, the sensor, control and tool system execute initial check. With the guarantee of the safety check system, the gondola cage arises to the top of the building façade. Painting or cleaning begins after the gondola cage moves downward. Detail procedures about painting and cleaning vary according to the tool mechanism. In our system, the painting tool is an array of spray guns which has an actuator system enabling movement inside of gondola cage. Therefore, we accept the stop-and-go strategy for the downward movement of the gondola cage. The gondola cage holds its position while the procedure of tool is executed. After its completion, it moves downward as far as the tool covers.

B. Concepts for the Building Façade Maintenance Robot

For this scenario, some concepts and assumptions are required. First, the operation of the building façade maintenance is prohibited with over speed wind because of its safety. This is identical with human-based operation. Additionally, it helps the precision and robustness of sensor and control system. Most of all, we can trust that the gondola cage keeps close adhesion using its suctions fans. Based on this assumption, we can apply sensing and control strategies with fixed distance from the façade robot to the building façade.

Second, the safety zone for painting is allowed. It is unpainted areas around window frames. Human workers will fill up this vacancy through the open window. It means that extremely precise control is not required. Many fault-tolerant techniques can be used. Fig. 1 is the concept of the safety zone of painting. As shown on Fig. 1 (b), the shape of the safety zone is not regular.



(a) (b)

Figure 1. Concept of the safety zone of painting. (a) An excample of building façade with typical windows. (b) Green paint is sprayed on the wall. There are unpainted area (light orange color) as the safety zone of painting.

III. Window Detection

Painting on window and applying cleaning tools for reinforced concrete on the window glass are unwanted situations for the building façade maintenance robot. Therefore, we need the safety zone of painting. In order to make the robot recognize the safety zone, the robot should detect windows. In remote control situation, the window detection can help operators through stopping spray guns even though the operator's mistake or unintended command let them go over windows.

There are many possible ways to detect windows. For example, aluminum detection sensors and visual cameras can extract window frames, and the contrast of piercing ability of ultra sonic and infra red ray can detect window glass. In this paper, we focused on the visual camera based on the color difference of building façade and window frames.

Fig. 2 is the installation guide of a visual camera in accordance with the above scenario. Because the gondola cage goes down while the core operation, the field of view of the camera covers the building façade under the cage. Suction fans hold the close contact between the robot and the façade, and the depth from camera to the façade is also fixed. The camera watches the place the robot will cover next time.

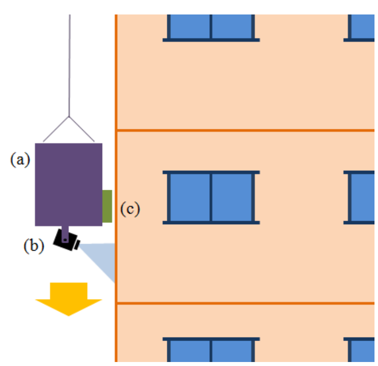


Figure 2. Installation of a visual camera in accordance with the scenario. (a) A gondola cage. (b) A visual camera. (c) A suction fan for the gondola cage.

With above system, the image from the camera is processed with following steps. First, the color of the window frame is filtered, and the image is converted to the binary image. Second, the dimensional information on image plane is extracted from the converted binary image using connect-component labeling technique or histogram approach. Finally, the information on image plane is transformed to the camera coordinate.

A. Binary Image

Visual cameras provide information in RGB (Red, Green and Blue) space. Among that information, the window frame should be extracted. The key is that the operator of the building façade robot knows the color of the window frame. However, the one indentified color cannot describe every window frame even though the color of window frame is the same. The luster on the window frame alters the RGB value on the image. Illuminations also influence the value.

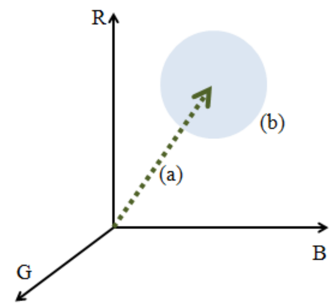


Figure 3. RGB space with three axis of R (red), G (green) and B (blue). (a) The vector of the window frame color. (b) The margin for the luster and illumination condition.

Fig. 3 shows the RGB space as three-dimensional space. Fig. 3 (a) is the representative color vector of the window frame. This color may be different of the actual color of window frame. It is the color selected by operators as the color can describe the window frame under the luster and illumination condition. Fig. 3 (b) is the tolerance for the condition. The colors inside this sphere space are considered as the window frame color. Even though Fig. 3 has only one vector as the representative color, plural RGB vectors can be set. This concept can be shown as

$$d_{ij} = \min_k \{ \sqrt{(r_{ij} - r_k)^2 + (g_{ij} - g_k)^2 + (b_{ij} - b_k)^2} \} \tag{1}$$

d_{ij} is the distance from the representative color vector $\begin{bmatrix} r_k & g_k & b_k \end{bmatrix}$ to the pixel color $\begin{bmatrix} r_{ij} & g_{ij} & b_{ij} \end{bmatrix}$ on i th row and j th column. Because the representative color is not unique, the representative color has the index k. Therefore, the binary image base on (1) is

$$v_{ij} = \begin{cases} 255 & if\, d_{ij} > d_{thres} \\ 0 & if\, d_{ij} < d_{thres} \end{cases} \tag{2}$$

d_{thres} is the threshold distance to decide whether the color distance of pixel d_{ij} is close enough to regard as the window frame. v_{ij} is the pixel value of the binary image on on i th row and j th column. It means that d_{thres} is the radius of Fig. 3 (b). This binary image is a kind of pepper-and-salt image based on the window frame color. The distribution of the salt pixel is tuned by the number of the representative color vectors and the threshold distances.

B. Connect-component Labeling

After this binary image is converted, there may be some noises, such as the similar color on the wall. This situation is severe if the window frame is silver of white and the luster on the wall is vivid.

As a first solution, the salts on the binary image is converted to the connect-component labeling [5]. This algorithm makes labels on the blob of white pixels. Then, the following step is to check the size of the blob. As shown in Fig. 2, the gondola cage keeps contact to the building façade. The size of the window frame is also expected easily. Smaller or bigger blobs are excluded.

Base on the close contact of the gondola cage and the building façade, the depth from the camera to the façade is also fixed. Therefore, the window frame on the image plane can be converted to the camera coordinate information.

$$\begin{aligned} X &= \frac{Z_{xy}}{f_x}(x - c_x) \\ Y &= \frac{Z_{xy}}{f_x}(y - c_y) \end{aligned} \tag{3}$$

x and y are image coordinates of the pixel. c_x and c_y are the image center in pixel. Z_{xy} is the depth from the camera to the façade, and it is fixed by the kinematics of the gondola cage and camera. In scenario of Fig 2, the depth varies in accordance with the image coordinate because the direction of camera is not perpendicular to the façade, but it is fixed. A look-up table can be used to get this varying depth information.

C. Histogram

Histogram is a stochastic approach that there are many salt pixels of the binary image on the direction of the window frame. This concept is valid because the salt pixel in (2) is choose because the pixel has the similar color with the window frame. In order to simplify the stochastic, we assume that the window frame is aligned to row and column direction of the image. If the pose sensing and control is accomplished well, it is also possible. If not, image distortion for alignment is a trivial image processing.

The row histogram is built through moving the salt pixels to the left along the row direction, and the column histogram is along the column direction. Detail examples are shown in latter chapter. Because the operator knows the size of the window frame, it is simple to set the threshold for the length of histogram to determine whether the histogram is the window frame or not. Then, through applying the size of other direction, we can select the gap of histogram to determine the outside of the window frame. For example, at the row histogram, assume that the first histogram longer than the threshold is the upper edge of the window frame. From this edge, the histogram is checked with threshold and the gap to the lower edge of the window. This approach is very useful for windows that have a security grille. Many security grills have the same color with the window frame, so there are many histogram bars over the threshold. By applying the height of the window frame, the real window frame can be detected.

After finishing the histogram process, the lower, upper, left and right edges of the window frame are fixed, and the cross points are the corners of the window frame. By applying the same approach in (3), the histogram approach also can be concluded the dimensional information of the camera coordinate.

IV. Experiment Results

To apply the proposed methods to the real window images, we gather various samples of window frames with a webcam[1] and an ultra sonic range finder[2] in Fig. 4. 640 x 480 pixel images are captured, and one centimeter resolution depth information is recorded in synchronized time.

Figure 4. A webcam and an ultra sonic range finder to gathering real window frame images and their depth information.

A. Connect-component Labeling

For connect-component labeling approach, a window which has the window frame with different color is selected. As shown in Fig. 5 (a), the wall is painted with white, and the window frame is brown. The pixel on $(132, 399)$ of row and column space is selected as the representative color vector of the window frame. The pixels within 60 (distance of RGB space) are set as salt in the binary image at Fig. 5 (b). On this binary image, the window frame is visible, and the pixels of the wall are almost excluded.

[1] http://www.microsoft.com/hardware/en-us/p /lifecam-cinema/H5D-00001

[2] http://www.robot-electronics.co.uk/htm/srf02techSer.htm

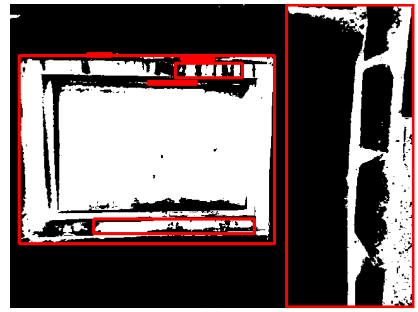
(a)

(b)

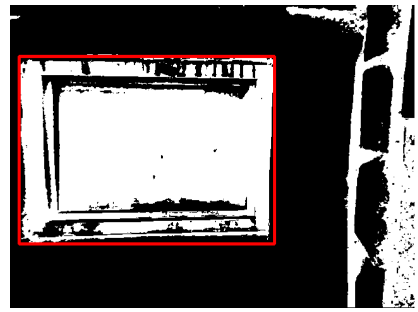
Figure 5. (a) Sample image for the connect-compoenet labeling. (b) The binary image with the representative color vector of the window frame.

Fig. 6 (a) is the output of the connect-component labeling. All of the separated blobs are boxed with the red lines. Because they are not filtered, a lot of blobs are extracted. Even in the window area, there are plural blobs. In accordance with the size of the blobs, the noise blobs are excluded. The relative size between the image and the blob is used for this experiment because these parameters are intuitional and easy to be tuned. We just excluded blobs smaller than 50 percent and bigger than 90 percent of the camera image. Fig. 6 (b) is the output of the filtered connect-component labeling, and Fig. 6 (c) is the blob boundary is overlaid on the input RGB image.

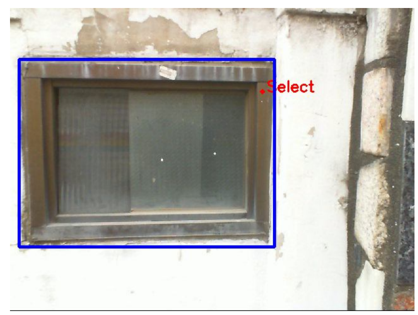
Many license plate recognition algorithms use the ratio of horizontal and vertical length of the objects [6], Fig. 5 is not needed it because the binary image is really clean. However, in other building facades, this approach can be applied.



(a)



(b)



(c)

Figure 6. (a) The output of the connect-component labeling. (b) The blobs are filtered 50 percent smaller and 90 percent bigger than image size. (c) The window frame is overlaid with the window fram blob boundary.

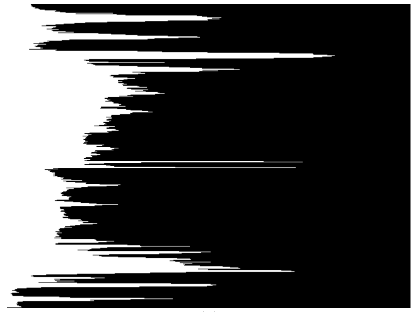
B. *Histogram*

Fig. 7 (a) is the input image for the histogram approach. It has the security grille and its color is metal gray. In order to improve the binary image, two color vectors are input as the representative color vector of the window frame. Their pixels are $(80,304)$ and $(361,555)$. The pixels within 40 (distance of RGB space) are set as salt in the binary image at Fig. 7 (b). Because there are more representative vectors, the margin decreases.

(a)

(b)

Figure 7. (a) Sample image for histogram. (b) The binary image with the two representative color vectors of the window frame.



(a)

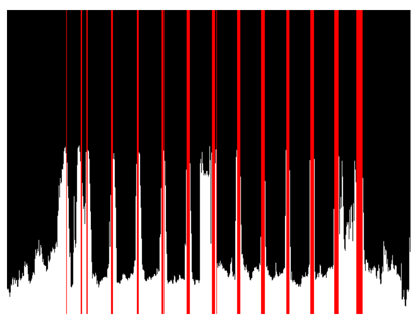
(b)

Figure 8. (a) Row histogram of the binary image. This is a sifted image of salt pixels to the left. (b) Column histogram.

Fig. 8 (a) is the row histogram. The occurrences of the salt pixels are accumulated as a histogram bar. Actually, this is a shifted image of the salt pixel to the left. There are three groups of histogram which have large value. They are the upper and lower edge of the window frame and the center support of the safety grille. Other salt pixels caused by luster of the wall and left and right edges of the window frame are distributed for overall histogram. Fig. 8 (b) is the column histogram. The safety grille is a type that has vertical bars. Therefore, the histogram shows the similar heights of value as large as the height of the window. Note that the gaps of histogram are identical to the gap of safety grille bars.

Fig. 9 (a) and (b) are the threshold applied histogram for row and column, respectively. The red lines means the histogram bar is over the threshold. The threshold of the row histogram is 450, and the other is 213. The gaps of the red lines coincide to dimension of the window frame and the safety grille.

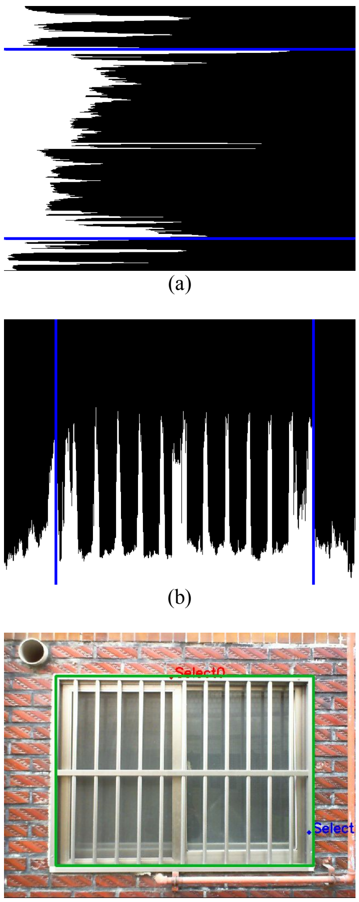
(a)



(b)

Figure 9. Histogram bars larger than threshold are checked with red lines.

As the next step, the first and the last red lines are selected as the edge of the window frame. However, not only the sequence is applied, but the dimension information of the window frame and safety grille is also applied. For example, if there is a bigger histogram than threshold on column histogram but it is farther than the gap of the safety grille, the histogram will be discarded as a noise histogram bar. Fig. 10 (a) and (b) shows the histogram bars concluded as the edges of the window frame. With cross points of these edges, the window frame box is overlaid on Fig. 10 (c).



(a)

(b)

(c)

Figure 10. (a) The histogram bar concluded as the horizontal edge of the window frame on row histogram. (b) The vertical edge of the window frame on column histogram. (c) The detetected window area is boxed with horizontal and vertical edges with green color.

C. Dimensional Information on Camera Coordinate

In this experiment, the depth information is acquired through the ultra sonic range finder. However, on the real gondola system, the depth information can be measured through the kinematics of the gondola cage.

On the center of Fig. 11, the ultra sonic data and the image center in pixel are shown. With the focal lengths $f_x = 506.8\ pixel$ and $f_y = 394.0\ pixel$, the camera coordinate values are extracted in accordance with (3), and the values are displayed in Fig. 11.

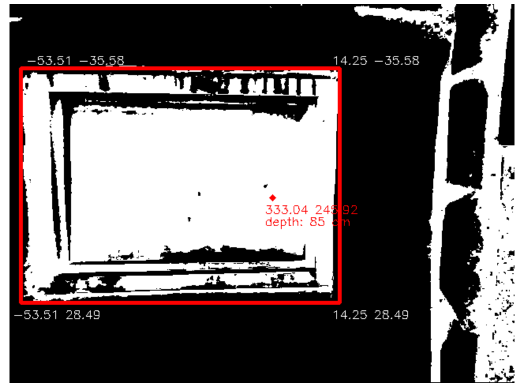


Figure 11. Applying depth information and extracting three diemnsional information in the camera coordinate.

V. Conclusion

The building façade cleaning robot system is a valuable system to release human workers from risky working environments. As a component of this system, we have researched the sensor and control systems. In this paper, we proposed two window detection methods. First method is to apply the connect-component labeling, and second one use stochastic through histogram approach. A simple window and a window with a safety grille are processed as experiments.

As future works, we will apply these methods on around view monitoring system, and solve partially shown window problem.

Acknowledgment

The work presented in this paper was funded by BMRC (Building-Façade Maintenance Robot Research Center), supported by Korea Institute of Construction and Transportation Technology Evaluation

References

[1] Gambao, E., Hernando, M., (2006) "System for a Semi-automatic Façade Cleaning Robot", in *Proc. of the 23rd International Symposium on Automation and Robotics in Construction (ISARC),* 2006, pp. 406-411.

[2] Zhang, H., Zhang, J., Zong, G., Wang, W., Liu, R., (2009) "Sky Cleaner 3: a real pneumatic climbing robot for glass-wall cleaning", *Robotics & Automation Magazine, IEEE*. 2006, 13 (1): 173-183.

[3] H. Sun, D. Y. Kim, J. H. Kwon, B.-S. Kim and C.-W. Park, "The position and orientation measurement of gondola using a visual camera", in *Proc. of the 2011 International Symposium on Automation and Robotics in Construction (ISARC 2011)*, September 2011, pp. 103-106.

[4] D. Y. Kim, J. H. Kwon, H. Sun, B.-S. Kim and C.-W. Park, "Rope attribute measurement system for gondola type facade robot", in *Proc. of the 2011 2nd International Conference on Construction and Project Management (ICCPM 2011)*, July 2011, pp. 693-697.

[5] H. Samet and M. Tamminen, "Efficient component labeling of images of arbitrary dimension represented by linear bintrees", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, pp. 579-586. July 1988.

[6] E. R. Lee, P. K. Kim and H. J. Kim, "Automatic recognition of a car license plate using color image processing", in *Proc. of Image Processing, 1994. ICIP-94., IEEE International Conference*, November 1994, Vol. 2, pp. 301-305.