Door Recognition Principles in GITA’s Machine Vision

Nanxuan Zheng\textsuperscript{1,\textit{a}}, Fangyu Cui\textsuperscript{1,\textit{b}} and Yamching Hui\textsuperscript{1,\textit{c}}

\textsuperscript{1}University of California, Los Angeles, Los Angeles of California, USA

Email: \textsuperscript{a}archznx@g.ucla.edu; \textsuperscript{b}fangyucui@g.ucla.edu; \textsuperscript{c}huiyamching@g.ucla.edu

Abstract. Nowadays, in the traditional door or object recognition system, people use GluonCV and Tensorflow to help teach machines millions of pictures, and then base on the information in the database collected from the pictures to do the door recognition. While in the new proposal, a new door recognition system based on a machine carrier called GITA is proposed. In the proposal, the recognition elements comprise motionless elements, including door boundary, handle/bar, and ratio of body width and door width, and loco-motor elements, which include people’s gesture, change of door depth and change of illumination. In order to prove it, several different tests based on seven doors of different types are done.

1. Introduction
GITA, a kind of carrier robots which created by PFF Company in 2017, will become a quite popular pet robot in the future, because it can follow people and free people’s hands when they are walking. However, when people are passing through doors, the GITA sometimes will stop working. And one of the most important reasons is that the GITA cannot correctly recognize the door and will hit it after people getting through.

In traditional door recognition system, machines have to learn millions of pictures at first\cite{1-2}. Then, architectural elements, like handle and panel, will be separately recognized\cite{3}. After combining all these together, we will get the result. Different from traditional door recognition system, the new recognition system includes new principles. As is shown in Figure 1, the new system includes six fundamental elements, door boundary, handle/bar, ratio of body width and door width, people’s gesture, change of door depth and change of illumination.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.jpg}
\caption{The relationship between the six elements(principles) we propose.}
\end{figure}
2. Machine vision of GITA
Different from people’s vision, the GITA’s machine vision can be more like a pet view. The height of the depth camera in GITA is about 530mm. As a result, in GITA’s machine vision, the picture and the information gotten can be totally different from people’s view.

3. Motionless recognition elements
The motionless recognition elements are composed of door boundary, handle/bar, ratio of body width and door width. They are all from architectural door elements, like door frame and handle/bar. Here they will be explained separately accompanied by the tests we have done. The specific conclusion will be drawn through these tests.

3.1. Door boundary
Door boundary mainly comes from door frame, one of the architectural elements. Because of the difference between the wall and the door, as well as the structure of door frame, it can be obviously shown in Intel D435 Depth Camera that the door frame, door panel and wall will have different color due to the difference of distance. As is shown in Figure 2, the left picture is from Intel D435 Depth Camera. The red part means it is far from camera, while the blue part represents it is close to camera. Also, the right picture is the post-processed one in the software Roborealm, and the door boundary is obviously extracted from the picture in the left.

Figure 2. Door Boundary shown in Intel D435 Depth Camera.

3.1.1. The tests. In order to prove the first principle, several tests based on some common doors from different distance are done on campus. Here are the tests and the information gotten from Intel D435 Depth Camera. Meanwhile, the post-processed pictures are also included inside. Figure 3 is about the tests. The tests are done on three kinds of doors, including seminar room, shop and entrance doors on campus. The distance represents the distance between the camera and the door. Here, the minus means that the camera is outside. The monochromatic pictures are those processed in the software Roborealm, and the door boundary is obviously highlighted.

Figure 3. The tests done on campus.
3.1.2. The conclusion. After analyzing all the tests, we find that, no matter what the material it is, the door boundary is obvious in Intel D435 Depth Camera because of the door frame. In a word, the door boundary should be one of the fundamental elements to help recognize the door in GITA’s vision.

3.2. Handle
The door handle is also related to one of the door architectural elements, which is called handle or bar. Because of the height of GITA’s camera, if the GITA is quite close to door, it cannot see the door handle. It means that we need to test whether it can be used as an element to help GITA recognize the door and the approximately available distance.

In fact, in daily life, there are lots of different kinds of “handle”. And then we try to collect most of different “handles”, which will do good to the research. The following pictures are about this part, and different kinds of handles are shown in Figure 4. Here, the first two rows are the pictures from the Internet, and the third row include the pictures collected from the campus. Here, the bars, handles and armrests are all included.

![Figure 4. Different kinds of door handles.](image)

3.2.1. The tests. In order to prove the second principle, as is shown in Figure 5, several tests are done on some common doors at the experiment site. Here are the tests and the information collected through Intel D435 Depth Camera. And the post-processed pictures from the Roborealm software are also included in the chart.

![Figure 5. The tests done in campus about the handles.](image)
3.2.2. *The conclusion*. After the experiments, we can draw the conclusion that the door “handle” is also one of the elements used in the door recognition system. The best distance for recognition should be 2–2.3 m. In addition, the bars or handles are quite common on the doors and obvious for the GITa’s camera to recognize.

3.3. *Ratio of body width and door width*

Different from the first two motionless elements, ratio of body width and door width is not quite related to traditional architectural elements. We would like to use a picture and an equation to help explain this motionless principle. Figure 6 just shows the ratio of body width and door width in the grid x-y axis.

![Figure 6](image)

**Figure 6.** The x-y axis built to evaluate the ratio of body width and door width.

\[ R = \frac{x_4 - x_1}{x_3 - x_2} \]  

(1)

\[ R \]  ratio of body width and door width  

\[ x_4 - x_1 \]  body width  

\[ x_3 - x_2 \]  door width

In the Depth Camera and Roborealm software, based on the information of door boundary and people whom GITa is following, the machine can position people’s body width points and door’s width points, just like the blue points and red points in the figure above. At these points, they will be shown as \((x, y)\), which can be shown in Figure 7.

![Figure 7](image)

**Figure 7.** The point in the Roborealm software. And it is shown as \((x, y)\).

3.3.1. *The tests*. Here are the tests done at the experiment site, which is shown in Figure 8. After processing them in the Roborealm software, the ratio results that are collected from every experiment can be used to help get the range of the ratio where most results are inside.
3.3.2. The conclusion. As is shown in Figure 9, here are the results of the tests. Most of them are between 0.44 and 0.60. So, in the future, the ratio of body width and door width can be regarded as a part of the recognition system.

![Figure 8](image_url) [Figure 8. The results of the experiments (The red line represents body width, and the black one represents door width. The distance represents the distance from the camera, and the “–” represents that GITA is outside).]

Figure 9. The result of the ratio of body width and door width. The figure here shows most results are in between 0.44 and 0.60.

4. loco-motor recognition elements
The loco-motor recognition elements include people’s gesture, change of door depth, and change of illumination. They are not traditional architectural door elements, like door frame. Instead, they are definitely related to the process of people’s passing through the door. Meanwhile, they will be explained separately accompanied by all the tests done on campus. Finally, after the explanation, the conclusion can be summarized according to the tests and analysis.

4.1. People’s gesture
In this part, the process of people’s passing through the door will be focused on. As GITA is always following people, it can easily capture the movement of people’s gesture when people try to open the door. In order to simply the gesture of opening the door, we use a software called Posenet to analyse the original video and then extract the movement of people’s skeleton, just like Figure 10.
Then, in order to conclude the common principles of people’s gesture when people are opening the door, several tests of different doors are done first. As is shown in Figure 11, we also extracted some clips of opening door process from famous movies or teleplays to get multiple samples.

Through the analysis of these materials, it is found that there are three principles of people’s gesture when opening the door. The first one is that people always turn around to close the door after passing, the second one is that the skeleton will become smaller during the passing process, and the last one is that people’s hand will move to open the door. If a gesture meets all the three principles, we can recognize that it is the gesture of people’s opening the door.

4.2. Change of door depth
As is known to all, when the GITA is following people and people need to go through the door, the GITA’s camera will see the process of people’s opening the door and then closing the door. In the
depth camera, the process can be more obvious. When the door depth is changed, the color of the edge of the door panel also changes in the GITA’s depth camera, just like Figure 12.

![Image of color changes in depth camera](image.png)

**Figure 12.** Different color represents different distance from the camera. It shows the relationship between the specific colors and the distances.

Then, the specific distance data can be extracted from the depth camera. In this way, we can build mathematical model which can be used to evaluate in the door recognition in the future. Here are the three obvious tests done on campus, which is shown in Figure 13. Just focusing on the color of the door panel edge, we find that, the process mentioned before can be quite obvious.

![Image of three tests](image.png)

**Figure 13.** The three tests done at our experiment site.

After all these tests, we can draw the conclusion that, the change of door depth should be like a quadratic function, which has an equation similar to \( f(x) = ax^2 + bx + c \).

However, because of some factors related to the specific doors, we will cannot get the exact equation to help do the door recognition, but can only get the approximately tendency that the door depth will become larger and then become smaller, or first become smaller then become larger. There are two tendencies because there are two directions to open the door. And the result is shown in Figure 14.
Figure 14. The result shows the approximate relationship between the door depth and the color in the process of people’s going through the door, which is based on the tests.

4.3. Change of illumination
Since the camera of GITA uses infrared ray to detect light environment it can see better in the dark environment. The difference of illumination during the door opening process can be used as the basis for the GITA to judge whether the front door is a door or not. As is shown in Figure 15, the picture of the same door differs when it is opened and closed. We use Roborealm to deal with these images. The white area of the images shows information lost. It means that process of opening doors does bring about a difference in light.

Figure 15. The two photos of the same door in the depth camera and we use Roborealm to deal with these images.
Then, several experiments are done based on five different doors. Meanwhile, we record the data of the change of illumination, which is shown in Figure 16.

| CLASSROOM | ENTRANCE | SHOP | ROBOT ROOM | OFFICE |
|-----------|----------|------|------------|--------|
| CLOSED    |          |      |            |        |
| 42 lux    | 18 lux   | 26 lux | 82 lux     | 34 lux |
| OPENED    |          |      |            |        |
| 85 lux    | 79 lux   | 61 lux | 89 lux     | 72 lux |

**Figure 16.** Use the luxmeter to get the data of illumination change and the specific illumination difference in the process of opening a door.

**Figure 17.** The figure is about the change of illumination. And it shows that, when people open the door, the illumination difference is most likely larger than 30lux.

After the analysis, it is found that, most of the data of illumination difference is larger than 30lux, which is shown in Figure 17. So, we would like to use the change of illumination as another element in the door recognition system. If the GITA detects that the change of illumination is more than 30lux in a short time, the person it follows is likely to open a door.

5. Conclusion
Object recognition is one of the most important parts of machine vision system, and the door recognition is also a part of the object recognition[4-5]. The door recognition can be especially important for the GITA, as it is a robot which follows people always. If it can recognize the door, then the GITA can follow people everywhere and will not get lost.

Usually, the door recognition will focus on teaching machines millions of pictures or separately recognizing architectural elements like handles and the panel. Instead, based on architectural elements of the door and the process of people’s going through the door[1,6], we propose six new elements of
the door recognition system, including the door boundary, the handle/bar, the ratio of body width and door width, people’s gesture, change of door depth, and change of illumination.

And after our test, we find that, these six principles are available for most common doors in our daily life, even including the glass door. Then, Figure 18 just shows our one of the tests which we have done, and the result shows that the six principles still work well enough although there is a lot to improve.

![Figure 18](image)

**Figure 18.** Here are the two results of our final experiments. They are done in the Roborealm software to help simulate our principles and our final conclusion.

Finally, in the future, when the machines like GITA can recognize the door well, then people may free their hands totally, even when passing through the door.

**Acknowledgments**

We would like to express our gratitude to those who provide us assistance during the research. We acknowledge the help from our professor, Greg Lynn, who has given me suggestions on loco-motor elements. Also, he also provides us with GITA, the robot we do research on. And we also want to acknowledge the assistant from our instructors, Marta Nowak and Kaiho Yu, who provide us with suggestions during our research, and also the Intel D435 Depth Camera. Meanwhile, we will acknowledge Weifan Weng, who gives us advice on the whole frame and outline. Finally, we want to acknowledge every juries whom are invited to our technology seminar final review, as you all give us some suggestions and also mention some problems we need to improve!

**References**

[1] Rodrigo Verschae and Javier Ruiz-del-solar 2015 Objects Detection: Current and Future Directions *Frontiers in Robotics and AI* pp: 1-7

[2] Wei Chen, Ting Qu, Yimin Zhou, Kajjian Weng, Gang Wang and Guoqiang Fu 2014 Door recognition and deep learning algorithm for visual based robot navigation *In proceedings of the 2014 IEEE International Conference on Robotics and Biomimetics* IEEE

[3] Adrian Llopart, Ole Ravn and Nils. A. Andersen 2017 Door and Cabinet Recognition Using Convolutional Neural Nets and Real-Time Method Fusion for Handle Detection and Grasping *In proceedings of 2017 3rd International Conference on Control, Automation and Robotics* pp: 144-9 IEEE

[4] M. Turk and A. Pentland 1991 Eigenfaces for recognition *Journal of Cognitive Neuroscience* 3(1):71-86, 1991

[5] H. Murase and S. K. Nayar 1995 Visual learning and recognition of 3-D objects from appearance *International Journal of Computer Vision* 14:5-24, 1995

[6] P. Belhumeur and D. Kriegman 1998 What is the set of images of an object under all possible illumination conditions *International Journal of Computer Vision* 28(3):1-16, 1998