Target Detection Based on Moment Invariants

Jiwu Wang Masanori Sugisaka

Department of Electrical and Electronics Engineering, Oita University, Oita 870-1192, Japan
(Tel : +81-97-554-7831; E-mail: wang_jiwu@hotmail.com , msugi@cc.oita-u.ac.jp)

Abstract: Perceptual landmarks are an effective solution for a mobile robot realizing steady and reliable long distance navigation. But the prerequisite is those landmarks must be detected and recognized robustly at a higher speed under various lighting conditions. This made image processing more complicated so that its speed and reliability can not be both satisfied at the same time. Color based target detection technique can separate target color regions from non-target color regions in an image with a faster speed, and better results were obtained only under good lighting conditions. Moreover, in the case that there are other things with a target color, we have to consider other target features to tell apart the target from them. Such thing always happens when we detect a target with its single character. On the other hand, we can generally search for only one target for each time so that we can not make use of landmarks efficiently, especially when we want to make more landmarks work together. In this paper, by making use of the moment invariants of each landmark, we can not only search specified target from separated color region but also find multi-target at the same time if necessary. This made the finite landmarks carry on more functions. Because moment invariants were easily used with some low level image processing techniques, such as color based target detection and gradient runs based target detection etc, and moment invariants are more reliable features of each target, the ratio of target detection were improved. Some necessary experiments were carried on to verify its robustness and efficiency of this method.

Keywords: **Perceptual landmarks, Moment invariants, Navigation, image processing**

1. INTRODUCTION

For a mobile robot used in practical applications, it is necessary to make it move reliably at a higher speed in a long distance navigation. But because of unavoidable slip and some incorrigible drift errors of sensors, the self-location in a map is nearly impossible so that reiterative and reliable long distance navigation is more difficult to be realized. It impeded the further applications of a mobile robot. Perceptual landmarks provide a solution to such problems. Compared with other sensors, a CCD video camera is cheap and always provides stable visual information. However, it also brings new difficulties: how to extract necessary and robust data used for navigation from each image. Because visual effect is heavily affected by lighting conditions and present algorithms are not powerful enough to deal with any situation especially if we need a higher processing speed, we can only try to extract some necessary and specified targets with special features. This is partial reason that landmarks get more attentions. But it does not mean that we can easily get landmarks' information robustly in all cases. This is because some problems can not be avoided, such as lighting conditions, selecting target objects from a filtered image, etc. And the robustness of an image processing system on landmarks was directly determined by solutions of those problems[1,2]. In this paper, based on the research of application of our mobile robot, some image processing methods of different landmarks were given, and the reliable solution to select one or more targets from one filtered image were also discussed.

In our research, the landmarks were divided into two groups. One is the continuous landmark that controls the navigation route of a robot. Another is some sign landmarks that control robot to perform specified commands at some necessary spots. Because their functions, shapes, sizes, and work environments are different, we need to use different image processing technique to improve landmark recognition capability. For example, because the continuous landmark will work under different lighting conditions, its image processing method should be less affected by lighting. But for a sign landmark, because they can be put any optimized selective position, we can ignore the influence of lighting. Although such classification was seemed simple, it can effectively improve the quality and speed of image processing even with image processing techniques.

In the following sections, we introduced color based target detection method for various sign landmarks and gradient runs based method for continuous landmarks. Their functions were to separate target regions from non-target regions in an input image. Then we gave the method to calculate moment invariants of each isolated objects. Thus we could find specified target from filtered target regions with their moment invariant thresholds. Noteworthiness, using moment invariants, we could not only find only target but also find multiple targets at the same time in an input image. It was helpful to make use of sign landmarks effectively.

2. COLOR BASED TARGET DETECTION METHOD

The main idea of a color based target detection method is to separate target color regions from non-target color regions in an image based on color thresholds of a target object. And the key point is that necessary features of a target object were kept in the separated color regions[3,5]. For a static image, we can adjust thresholds manually. But for robot applications, generally we will scan each captured image dynamically with the same thresholds. Because lighting conditions are changed when robot is moving, it is necessary to make those thresholds insensitive to lighting conditions when we want to get satisfactory results. The simplest method is only used color component as a threshold. Because the triple components (R, G, B) in a common RGB image represents not only color but also luminance, and it is luminance that will vary under different lighting conditions, it is not appropriate to directly use the triple components (R, G, B) as thresholds.

One of solutions is to determine such thresholds in the chromatic color space. Chromatic colors, also known as pure colors in the absence of luminance, are defined by a normalization of the triple components (R, G, B). A pixel in the chromatic color space can be defined by two components (r, b).

$$
\begin{aligned}
r &= R/(R+G+B) \\
g &= G/(R+G+B) \qquad (1) \\
b &= B/(R+G+B)
\end{aligned}
$$

And the third component is redundant because of $r+g+b=1$.

The problem of target color thresholds method is that it is not convenient to determine appropriate thresholds, especially for some colors with unobvious characters, such as being not pure red or yellow color etc. In such case as well as for general applications, a good solution is to build a target color model. And each image was separated with target color regions and non-target color regions based on the target color model. Because the target color model was built based on sample target color, it is more robust in most cases. The target color model was mainly built based on statistics information of one target color in various lighting conditions. From sample images, we extracted target color manually, and calculate the mean values and covariance of sample colors in the chromatic color space:

$$
\bar{x} = \frac{\sum_n x}{n} \qquad (2)
$$

$$
C = \frac{\sum_n (x-\bar{x})(x-\bar{x})^T}{n-1}
$$

where, x is a pair of normalization components (r, b) for each pixel in the chromatic color space, $\bar{x}$ is the mean value and C is the corresponding covariance.

Then whether the pixel in an input image belonged to target color or not can be determined by following formula

$$
P_x = \exp[-0.5(x-\bar{x})C^{-1}(x-\bar{x})] \qquad (3)
$$

Thus, the input image will be expressed with target color regions and non-target color regions.

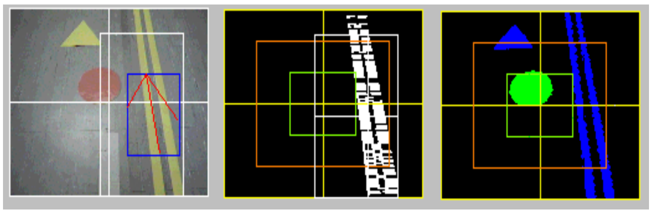
In the Hue Saturation Value (HSV) color space, it can separate the hue (color) from the saturation (the concentration of color) and the brightness. Therefore, we can also use the hue value as a threshold parameter to separate target color regions from non-target color regions in order to reduce the effect of the ambient lighting. Moreover, the target color model in HSV color space can be built as the same as in chromatic color space. Experiment results showed they gave similar robustness for landmarks recognition.

3.1 First part

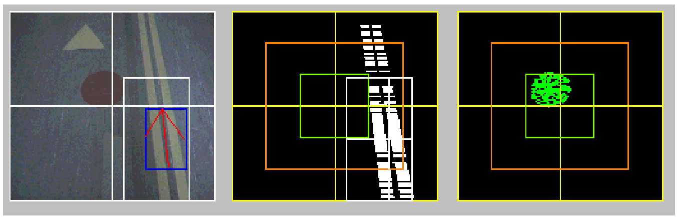
3. GRADIENT RUNS BASED TARGET DETECTION METHOD

Although it became more robust to separate target color regions from non-target color regions in an image in chromatic color space or HSV color space, its reliability still depended more on lighting conditions, especially the background color. It means it is rather difficult to further improve the robustness of the color based target detection technique. Moreover, it is not so robust enough for all colors. Gradient run technique gave another solution. Generally, there are color gradients everywhere in an image. And if in common cases, the influence of lighting on color gradients is very small, especially after equalization. In fact, with gradients, we can easily find the edges of any object (target or non-target) in an image. This is because there are distinctive gradients in such place.

Therefore, if some landmarks can be easily detected with their edge information or geometric information in relation to its edge, gradient run detection is a better solution. Its robustness was illustrated in figure 1.



a) The normal indoor lighting environment



b) The indoor environment after turn off the lighting

Fig. 1 A comparison of landmark extraction under different lighting conditions

In figure 1, the white lines were extracted with gradient run method. We found no matter it is lighting or not, the double yellow guideline can be extracted accurately. This is because the gradient of the edges of the guidelines was kept nearly same in both cases. The blue lines in figure 1 a) were extracted with color based target detection method. But when we turned off the light, the guideline can not be found. This is because at this moment the color information of the guidelines was not enough to be detected.

On the other hand, it is more robust to find edges when there are gradients. That means it can only be applied to some special landmarks with special edges, and such landmarks can be recognized with these edge information easily. Therefore, it is not appropriate for recognition of circle and triangle in figure 1 with this method. This is because we can not get the whole landmark area only with their edges robustly.

Generally speaking, gradient run method is more robust for some landmarks especially when the lighting is different at each spots. But the landmarks should be easily recognized with their edge information. Therefore, for some important landmarks, we can design them with some special gradients and edges so that it is more robust to be extracted. And It also filtered some noises when these landmarks were extracted with special thresholds of gradients.

4. CALCULATION OF MOMENT INVARIANTS

After each image was scanned with color-based filters or gradient runs based filters, one or more separated objects will be kept. Then we have to decide which object is a target. There are many algorithms that can deal with such problems, but moment invariants is one of the easiest and more robust method for shape detection. Actually moment invariants represent the geometry information of each object. Most of important, these variables nearly keep constants when the size of a target was changed in an image. At the same time, with moment invariants, multi-target can be recognized in one image[6,7].

Before the calculation of moment invariants, each isolated object in that image should be labeled first. This is because each filtered image in general will contain more than one object. Here we assume that each object is isolated from the others in the image, that means only distinctive objects were considered. In order to improve the image processing speed, we only keep objects with their area larger than thresholds.

We scan the filtered image, row by row, top to bottom, left to right. If we find the value of one pixel is not zero, the 8-neighbour pixels of that pixel will be checked whether they were labeled. If one of the pixels in its neighbour was labeled, that pixel will have the same label as its neighbour. Otherwise, it will be marked with a new label number. After all the pixels were labeled with this method, a second scan through the image is required to clean up the label equivalences, giving each connected component in the image a unique label. Then this labeled image will be input for moment invariants calculation.

When considering moment invariants, in Hu's paper the first significant publication was given. The moment invariants gave a way to describe and possibly identify an object, and were defined by

$$I_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^m y^n f(x, y)\,dxdy$$

and the discrete form for programming is

$$I_{mn} = \sum_x \sum_y x^m y^n f(x, y)$$

For simplification, all boundary pixels and all pixels inside the boundary were set to 1 whereas others were set to 0. That means $f(x, y)$ in the discrete form always is 1.

On the other hand, central moments are invariant to the location of the images, and useful for practical application. They were defined as follows:

$$J_{mn} = \sum_x \sum_y (x - \bar{x})^m (y - \bar{y})^n$$

where $\bar{x} = \dfrac{I_{10}}{I_{00}}, \bar{y} = \dfrac{I_{01}}{I_{00}}$

If only for seven moment invariants, the central moments are calculated up to the third order. And these moment invariants can be derived from the normalized central moment. The normalization of the central moments as:

$$I^c_{mn} = \frac{I_{mn}}{I^p_{00}}, \; p = \frac{m+n}{2} + 1$$

Then we can calculate the seven moment invariants which are invariant to rotation and changes in scale. But in some practical applications, only the first moment invariant was considered.

$$MI_1 = I^c_{20} + I^c_{02}$$

Because MI was unaffected by the orientation and scale of the objects in an image, it can be used for classification. On the other hand, these MI are normally very small so that they are difficult to be applied directly in classification, some transformations (such as log transformation etc) were done to overcome this problem.

In figure 2 we give an example to determine the center of the hole on the top of a gasoline can with different color. The results showed we could detect the hole and calculate its center with moment invariants after some other image processing such as edge detection etc. Most of important, it is very robust. In figure 2, we marked the center with the same parameters in the image processing.

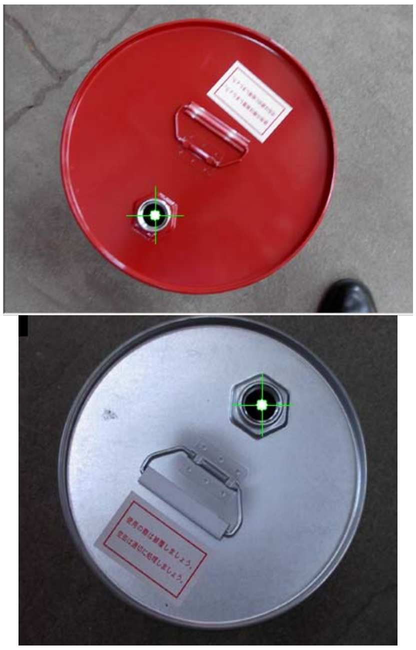
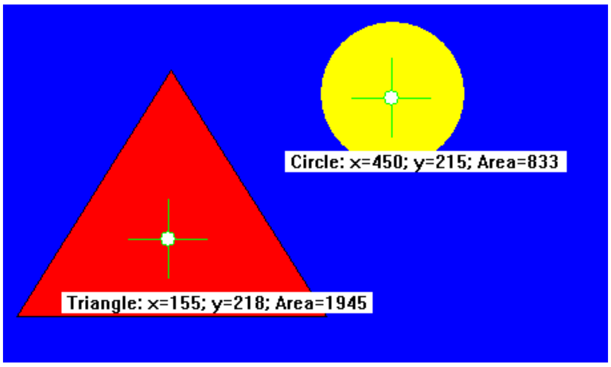


Fig. 2 An example to determine the hole center with moment invariants

5. EXPERIMENTS

It is obviously more robust if we use two different target detection methods than that with only one method. In our applications of target detection in robot navigation, the separated objects with color based target detection technique or gradient runs detection technique, were determined whether they belong to specified landmarks by their moment invariants. That means we determine a landmark based on both its color information and geometric information. Though it is possible that there is a non-target object with the similar color and size as a target object, the probability is very low.

Another solution is that we try to use landmarks with different colors from the background. The function of a landmark is used for robot realizing self-location and performing specified commands. It is more helpful if we can use an appropriate landmark at a specified location. On the other hand, in order to make use of the landmarks more efficiently, the robot should have the ability to multiple landmarks (such as different size and colors) in an image at the same time. As we known, MI was unaffected by the orientation and scale of each landmark in an image. If we put many triangle landmarks with same color but different size and orientation in one image, their MI values satisfy the same thresholds. In such case, we can only give different commands with their amounts. A more common application is to combine landmarks with different color and size, as shown in figure 3.



In figure 3, there is a red triangle landmark and a yellow circle landmark. Although there are only two types of landmarks, it will give different information if they were put in different sequences.

Although MI nearly keeps constant when the size of a target was changed, it will be more robust if the template was used. This is because pattern recognition is based on the extracted features. With the same template, the classification of the discovered features will keep consistency. It is necessary to remove the effects of scale changes, rotations and translations for a successful pattern classification. But here the template is different from the template matching, in which it is more complicated to build a pattern template for better recognition results. In our application, the template is only a empty frame, and the threshold was given by MI. It made pattern recognition simple and robust, and reduce time for image processing.

6. CONCLUSION

Perceptual landmark is an effective solution to realize reliable self-location in long distance navigation. In order to improve recognition robustness of landmarks under different lighting conditions, we made two types of landmarks: continuous landmarks for navigation route and sign landmarks for specified commands. Thus, based on their difference in functions, shapes, sizes and working spots, two types of low level image processing were developed to separate target regions from non-target regions in an input image: gradient runs based target detection technique for continuous landmarks, and color based target detection technique for sign landmarks. Experiments showed such classification was effectively to improve image processing quality and speed. Most of important, it could reduce the influence of lighting conditions on image processing.

On the other hand, it is difficult to realize high recognition ratio only filtered with a single feature of landmarks because it is unavoidable to possibly have a similar target in an image. Considering the moment invariant of each target shape to be nearly constants no matter the changes of size and orientations, it was used to find the target from the filtered target regions in an input image. Moreover, different from a pattern template, we used only a template to further improve the robustness of target recognition.

Finally, the above technique was used in the long distance indoor navigation of our mobile robot. All navigation strategies were based on the information gotten from the CCD video camera.

REFERENCES

[1] M.Sugisaka, Wang Xin and J. Lee, “Artificial brain for a mobile vehicle”, Applied mathematics and computation, Vol. 111, 2000.

[2] J. Wang, M. Sugisaka, Study on visual-based indoor navigation for an alife mobile robot, ICSEng2002, Las Vegas, USA, 2002, August 6-8 : 390-396.

[3] Cathy-Ann Radix, Andrey Loukianov, Masanori Sugisaka, “Evaluating motion on the alife robot prototype”, Proceedings of the 32nd ISR (International Symposium on Robotics), 19-21 April 2001.

[4] Peter I.Corke, Malcolm C.Good, “Dynamic effects un visual closed-loop systems”, IEEE Transactions on Robotics and Automation, Vol. 12, No. 5, October 1996.

[5] Jiwu Wang, M.Sugisaka, “Study on a color based target detection technique”, Proc. of the seventh Int. Symp. On artificial life and robotics (AROB 7th’02), Beppu, Oita, Japan. 16-18 January, 2002

[6] S.O.Belkasim, M.Shridhar and M.Ahmadi, “Pattern recognition with moment invariants: a comparative study and new results”, Pattern recognition, Vol. 24, No. 12, 1991

[7] Alireza Khotanzad, Yaw Hua Hong, “Rotation invariant image recognition using feature selected via a systematic method”, Pattern recognition, Vol. 23, No. 10, 1990

[8] J. Wang, M. Sugisaka, Improvement on the image processing for an automous mobile robot with an intelligent control system, International conference on control, automation and system (ICCAS2001), 2001: 437-440