'94 KACC (1994. 10. 17 ~ 20)

A Watching System of a person entering a restricted area by Image Processing

* Tadashi AKIMOTO, * Kageo AKIZUKI, ** Mamoru KAWAMURA
※ Department of Electrical Engineering, Waseda University,
3−4−1 Okubo Shinjuku, Tokyo, 169, Japan
※※ Japan Radio Co, Ltd. 1−1 Shimorenjaku 5, Mitaka, Tokyo, 181 Japan

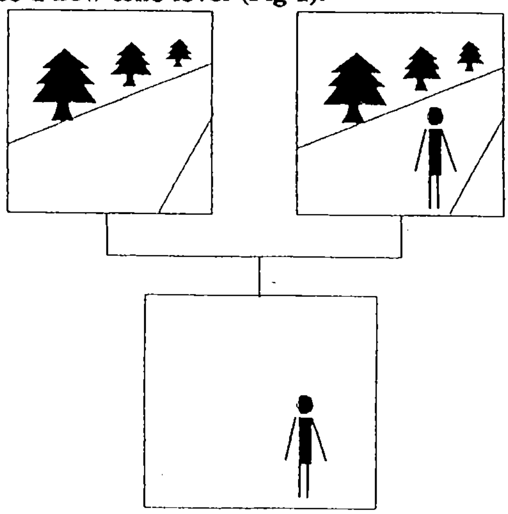
Abstract: For a safety supervision, watching a restricted area so that no one go there is very important. This has been mostly accomplished by people. They keep an eye on many monitors at onece for a long time. It, however, is too simple and boring to concentrate it for a long time. So it's worth while to construct a watching system by image processing. And the system we made is now actually working at a certain hydroelectric power plant and some other restricted areas in Japan.

1 Introduction

The biggest difference between the object, human beings in this case, and the background is whether it's moving or not. So we take the background image in advance, which doesn't have any moving object. And we use it all through the image processing. The image is called "the base image". Then we get a real time image through a video camera in every t seconds. The t depends on the fastness of the moving object we want to recognize. In this study, t is chosen as two seconds. We call that image "the object image".

We use the 512(sides) by 512(ends) digital images, which are black and white images with 256 tone levels (from 0 to 255). Zero means white and 255 means black. We use black and white images, because it's easier than using color images and they have enough information for our aim.

To detect the moving objects such as human beings, we subtract the base image from the object image. That is, we subtract the tone value of the base image from that of the object image, take the absolute value, let it be a new tone level (Fig 1).



$$S(i,j) = | R(i,j) - B(i,j) |$$

S : The subtracted image
R : The real−time image
B : The base image

Fig 1 The way to get the subtracted image

In this image, when there is nothing but background, the tone value of that area has a small value, ideally zero. Since they are almost same value. However when something is there, the new tone value of that area we get by the subtraction has a large value, since there is a big differnce between the moving object and the background. But we often can't get an exact zero for the value of the picture element where somethig is on because of an effect of wind, a change of brightness. And if we use the image which has a surface of water or some kind of liquid, reflection of a sunlight can be one of those causes. Because it keeps flickering. From a view point of getting the moving object, these picture elements with low tone level are noise picture elements which we need to get rid of.

2 Noise reduction

The noise picture elements with lower tone levels give us the way to remove those noise. Fig 2 shows the ratio of the remaining picture elements depending on a various threshold in the binarization.

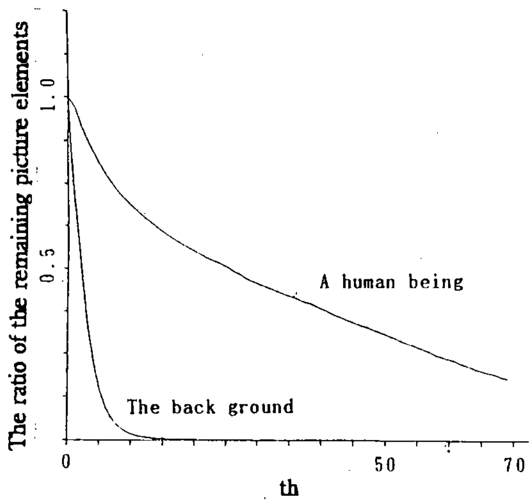


Fig 2 The ratio of the remaining picture elements

If we take 15 as the threshold and binarize the image, smaller tone noise can be removed. The most important thing is we have to keep a shape of the moving object as perfect as possible. That is more important than getting rid of noises. Because once we loose the information of the object we need, we can get them back no longer. The figure 3 is one of the binarized image.

If $|R(i,j)-B(i,j)| \geq th$ then $Bn(i,j)=255$

If $|R(i,j)-B(i,j)| < th$ then $Bn(i,j)=0$

Bn : The binarized image

th : The threshold

Fig 3 The binarized image

As you can see, there are some noise we can't remove only by the binarization. So we have to find another method which can take the object from noises.

In most cases, noises are much smaller than the moving object. As you will see later, we have some exceptions. But we are going to tell you the way we take care of those exceptions later, and explain the way we get rid of smaller noises first.

We use the process of shrinking and expanding to get rid of the remaining noises after the binarization. The figure 4 shows what those processes are.

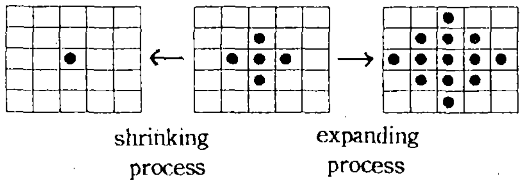


Fig 4 The shrinking & expanding process

The process of shrinking is the process which makes each object in the image smaller by one picture element. Doing this process several times, mostly once or twice, we can get rid of smaller objects. In this case, we do it twice. Then, to restore the sizes of the objects which are remaining in the image, we use the process of expanding, which makes each object in the image larger by a picture element. We have to do it same times as the process of shrinking. The figure 5 shows the effect of those process.

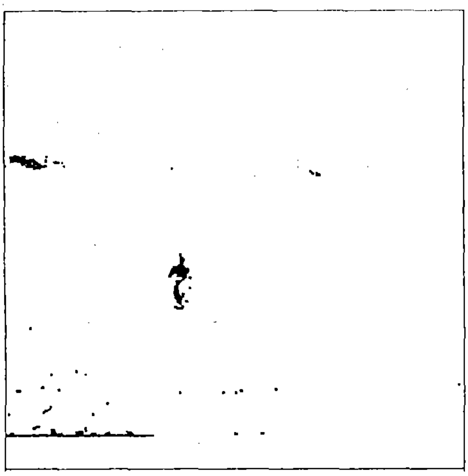


Fig 5 The effect of shrinking & expanding process (This image is from Fig 3.)

3 Recognition

We are ready for the part of recognition. If one of the object in the image is a human being, it consists of about eight hundred picture elements. The number of picture elements depends on the range of the area the camera covers. Eight hundred is for the image we are handling. It may apper that we could recognize a human being only by checking the number of picture elements. It, however, is very difficult. The following table (Fig 6) shows the result of the test with twenty two images under the condition that we consider there is someone in the area when the number of the picture elements of the object in the image is in the range,

$$700 \leq N \leq 1000$$

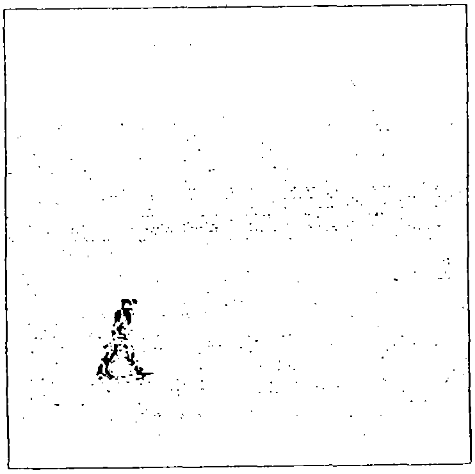
(N stands for the number of the picture elements)

Times of recognizing a human being	15 (68%)
Times of missing a human being	3 (14%)
Times of regarding a noise as a human being	4 (18%)

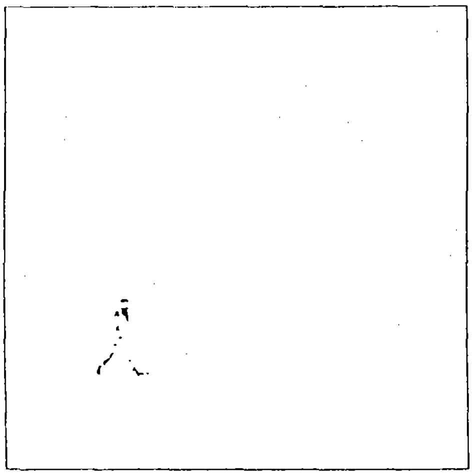
The "miss" means that a human being entering the restricted area wasn't recognized.

Fig 6 The result of the checking test only by the number

We have two kinds of misjudgements, especially, "miss" is more undesirable. Because we can't give him any alarms when the "miss" occurs. So another way is required. This mistake is caused by lack of the picture elements of a human being. To avoid making this mistake, the easiest way is to expand the range of the condition, which, however, ends up to consider more noises as a human being. Another way is to reset the threshold a little bit higher, which, however, allows more noise appear as shown in the figure 7.



th=18



th=52

Fig 7 Same images with different thresholds

To make matters worse, a noise, which is as large as a human being, occurs sometimes. In previous chapter, where we indicated most noises are smaller than the object, we mentioned the exceptions. And those noises are exceptions. Setting the threshold higher gives that kind of noise more chances to appear.

4 Modified noise reduction

Let's get back to the image just after noise reduction process (Fig 5 again). Then you'll notice that most noises is small and isolated. A human being and other objects are largeer than noises. That's the point we use here. We take a black picture element in the image one by one, put a 5 by 5 mask so that the black picture element comes to the center of the mask, and count the black picture elements in the mask. If the number of the black picture elements is small, we may consider the center black picture element is isolated or belongs to a small object. That means the picture element in the center of the mask is noise. So we remove that picture element. If the number of the black picture element in the mask is large, the center black picture element must belong to the large object. Then we keep it. The figure 8 shows one of the images (from Fig 3) after these processes.

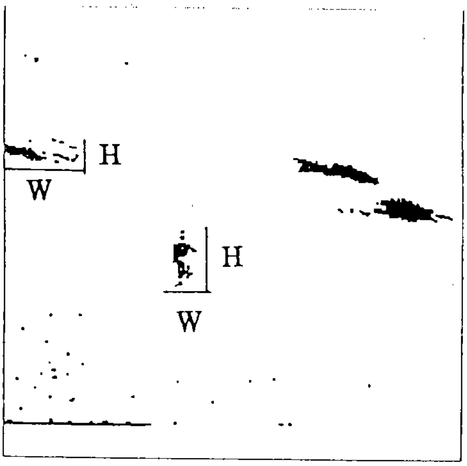
Fig 8 The image removed isolated picture elements

This process is as effective as the

shrinking and expanding process from the view of noise reduction. What's more, this process can retain the shape of the object better than the shrinking and expanding process. But unfortunately, none of this process and shrinking and expanding process can remove the noise which consists of a large number of picture elements. And we can't get a satisfactory result only by checking the number of picture elements of the object. So we have to set another condition for the better recognition.

5 Recognition with the additional condition

We have to find a difference between a human being and noises other than the number of picture elements. Then we took thier shapes as the additional condition. That is, we take a height to width ratio as the additonal condition as shown in the figure 9.



$r=H/W$

Fig 9 The height to width ratio

In our experience, a noise can't satisfy these two conditions (the number of the picture elements and the height to width ratio) at the same time. For instance, the river gose horizontally, which gives us a horizontal noise. On the other hand, shape of a human being is vertical. Even when some noises satisfy both conditions at one because of the angle of the video camera, it's possible to reset the camera and change the angle.

Now we regard the object as a human being only when both conditions are satisfied. The following table (Fig 10) shows how effective the additional condition is.

The additional condition is as follows.

If $520<N<980$, $0.8<r<2.7$ then consider the object as the human being

Times of a recognizing a human being	22 (100%)
Times of missing a human being	0 (0%)
Times of regarding a noise as a human being	0 (0%)

Fig 10 The result of the test with the additional condition

There is no misjudgements. And this is exactly what we are using now. We will show the results of our system which is used in a certain hydroelectlic power plant in Japan in the next chapter.

6 Applied results

The figure 11 shows the surroundings where our system is applied.

In this area, There is a river, and sometimes puddles of water. The surge tank has a lot of water and releases its water periodically. If someone is near the tank when the tank releases the water, it can be very dangerous. So we have to recognize him, and alarm him. That's why our watching system of a person entering a restricted area is required.

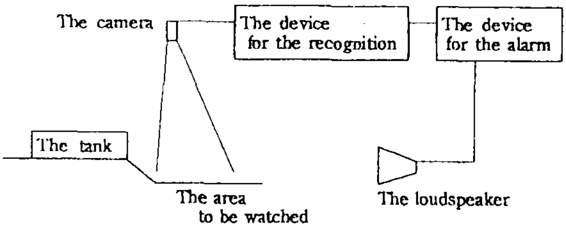


Fig 11 The area to be watched

The figure 12 shows the flow chart of the system.

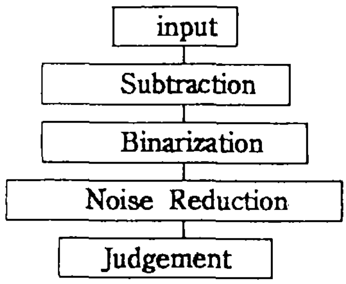


Fig 12 The flow chart

We use this system only in the daytime. When someone enters the area, the system have to recognize him and gives an alarm. And we have to avoid missing a human being when he is actually there.

In the actual application, when the real−time image has only the background, the system takes two seconds to go through one cycle shown in the figure 12.

When the real−time image has any objects other than the background, the system needs four seconds. Taking four seconds, however, isn't a problem. Because no one can pass through the area within four seconds. That means this system is sufficient for our purpose.

We had tested the effectiveness of the system for fourty five days before the actual application to the watching the restricted area. The test took place in the area where the system would be applied. The table shown in the figure 13 is its result.

Times of a person entering the area	81
Times of recognizing a human being	81 (100%)
Times of missing a human being	0 (0%)

Fig 13 The results of the test

The darker it gets, the more the system misses a human being when he is in the area. When the average of the tone levels gets down to about sixty, the system starts to miss in recognition of a human being.

But this system is used in daytime as we explained before. In this test, even when a person wore clothes in a similar color to the background, the system could recognize him. The test proved that the system can recognize a human being in spite of the river and puddles. And it is now actually working there. The results are satisfactory enough.

7 References

[1] コンピュータ画像処理
安居院 猛・中嶋 正之

[2] 画像工学の基礎
安居院 猛・中嶋 正之

[3] 画像処理による侵入者認知警報システム
川村 守

[4] Digital Picture Processing
Translated by Makoto Nagao