Wearable Detecting System of Approaching Dangerous Persons

Takamasa Nakano and Hirohito Yamada
Graduate School of Engineering, Tohoku University
Sendai, Japan

Abstract—**A wearable detecting system for approaching human have been investigated. Approaching humans can be detected within 15 m distance range. However detecting speed was very slow for the current system. We need more improvement for the detecting algorithm.**

Keywords-component; formatting; style; styling; insert (key words)

I. INTRODUCTION

Recently violent crime tends to increase. Among such violent crimes, violent crimes will cause getting people who do not know involved, it is possible that suddenly come down to us. To deal with such crimes, working on crime prevention and post process around the world for example installing a surveillance camera in the city . Furthermore, not only state but individuals install a security camera at home and attach the drive recorder to the car. However, such things have restrictions that cannot be carried because it fixed in private space. So our study aim to be cheap and portable wearable device.

We proposed a new security wearable device assisting human blind spots and helping us due to equip with object detection function and add a program that can detect approaching of humans. Wearable devices are cheap and light, so it is a convenient machine about individuals. But, model is required less computational load because wearable devices are inferior compared to normal PC. Furthermore, it required calculation accuracy calculation speed.

In this paper, we analyzed using SSD from several object detection models, as it detected a person from the image. Using the analysis information, we have written a program which detect approaching person, and installed it in a wearable device and we verified how much time it takes after confirming that it is operating correctly.

II. APPROACH DETECTION PROGRAM OVERVIEW

There are many ways to detect approaching people. For example, how to detect distance from difference with two stereo camera, ToF method that measures the distance from the time of the light returning from the laser emission. But, such devices are heavy and difficult to carry, and have a problem of becoming expensive. So, we think that it can solve those problems and is very convenient wearable devices are able to detect approaching people using only one program. Figure 1 is the flow chart of the written program. The program is reading five images, and detecting people and their coordinates. When analysis is complete, wearable devices check the position of the target person using change Bounding Box (BBox) area. If person approach in all five images, warning with alarm.

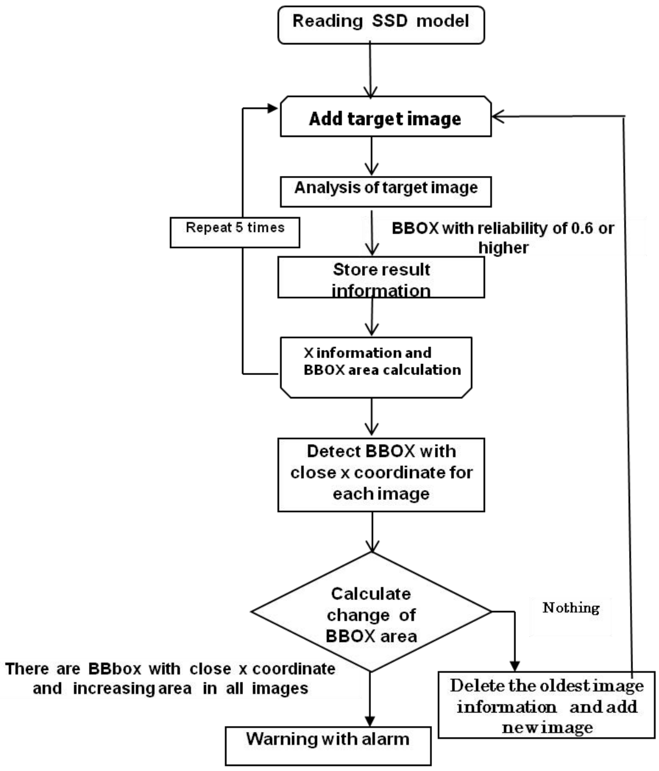

Figure 1. Program flowchart

III. IDENTIFICATION OF PERSON APPROACH

A. Approach detection at long and short distances

Next, run the created program, verify that the area change is actually calculated correctly. This time, actually measured the distance, taken 3120 × 4160 pixels image, analyzed it. Figure 2 shows detection at a short distance. The image on the left is 7 m from the observer and the image on the right is 6 m from the observer.

Figure 2. 7 m (left) and 6 m (right) distance image

It can be seen from these images that people can be detected at a short distance. And when comparing the size of those areas, 208250 pixels in the left image, it was increased by 297567 pixels in the right image. It was found from the result that the person could be detected at a short distance, and the area increased correctly according to the distance. Next, check if the same thing can be done at a long distance. Figure 3 shows the image that is farther from the observer than before, the image on the left is 14 m from the observer, and the image on the right is 13m from the observer.

Figure 3. 14 m (left) and 13 m (right) distance image

It can be seen from these images that people can be detected at a long distance. The size of the area was 53196 pixels for the left image and 57816 pixels for the right image, and the area increased correctly according to the distance even at a long distance.

IV. Approach Detection in Multiple People

In figure2 and 3, the information was surely the same person because there is only one person in the image. But, when multiple people are detected in the image, whether the person is in the next image, and the machine cannot determine which BBOX belongs to the same person when the person is in the next image. This time we are thinking of shooting at short intervals, so that location information doesn't change much. Therefore, judgment whether the same person or not is made by using the x coordinate of BBOX as a reference for each image.

To make it easier to see that the same person is recognized, named person 1, person 2, person 3 because there are three people in the image. Figure 4 and 5 show the same person, but images taken at different times. However, the same person can be distinguished as person 2, 1, 3 from the left.

Figure 4. detection image 1 of multiple people

Figure 5. detection image 2 of multiple people

V. Introduction to Wearable Device

A. Normalized size and analysis time

Let's actually introduce these programs to wearable devices. Experimental equipment used are below.

Object detection model ・・・ SSD

Wearable device ･･･ Raspberry pi3 Model B+

TABLE I. RASPBERRY PI3 MODEL B+ PERFORMANCE

CPU	Broadcom BCM2837B0, Cortex-A53 (ARMv8) SoC
CPU clock	1.4 GHz
Memory	1 GB
Power rating	DC 5 V
Current consumption	1.5 A
External dimensions	86(W) × 57(D) × 17(H) mm
Weight	50 g

Camera ･･･ Raspberry pi Camera Module V2

TABLE II. RASPBERRY PI CAMERA MODULE V2 PERFORMANCE

Interface	CSI
IC	Sony IMX219
Still image resolution	8 megapixels
Resolution/frame rate	1080p (1920 × 1080)/ 30 fps
	720p (1280 × 720)/ 60 fps
	480p (640 × 480)/ 90 fps
External dimensions	25(W) × 24(D) × 9(H)mm
Weight	5 g

There are several object detection models. For example Yolo, SSD, Faster-RCNN. Faster R-CNN achieved 7 FPS at 73.2% mAP. SSD achieved 59 FPS at 74.3% mAP, compared with 45 FPS at 63.4% mAP for Yolo in VOC2007 test. SSD is a model that can be detected very quickly and accurately. So, we used SSD. When executed figure 1 program chart in raspberry pi3 model B+, the time to calculate whether it is approaching by comparing the positional relationship after reading and analyzing 5 images is 86 seconds. Normalized size 400 pixels. This is not a practical time. Therefore, the result information is deleted only for the oldest image, the latest four images are saved, and a new image is added for analysis. Because this allows to analyze just one image, it can be calculated much faster than before. Normalized size is 400 pixels same as before, calculations that took about 86 seconds can now be calculated in about 14 seconds. This time is getting closer to a practical time. We thought that it might approach the target time of 5 seconds depending on the normalized size, and examined the relationship between the normalized size and the analysis time.

It can be seen from figure 8 that the analysis time of the image increases in a quadratic function as the normalized size increases. From the current situation, it was found that if the normalized size is 200 pixels, the target speed of 5 seconds can be achieved. The reason we set a target time of 5 seconds is that because the walking distance of men and women is different by 1m in 5 seconds. The reason why we haven't examined more than 700 pixels is that the memory installed in Raspberry pi3 Model B + is 1 GB, and if we specify a size of 700 pixels or more, it will run out of memory.

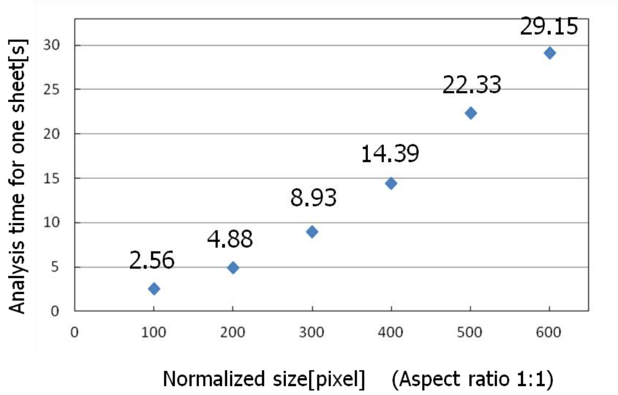

Figure 6. Change in analysis time relative to normalized size

B. *Recognition rate and distance*

We know the change in analysis time relative to normalized size, if it doesn't actually recognize people at its normalized size, it doesn't make sense. Therefore, we investigated how many 300 images can be recognized with a normalized size of 100-600. Figure 9 below shows the distance and the recognition rate. In this experiment, even if it was recognized as a person, it was decided that it could not be recognized unless the BBOX was accurately matched to the person shown in the photo.

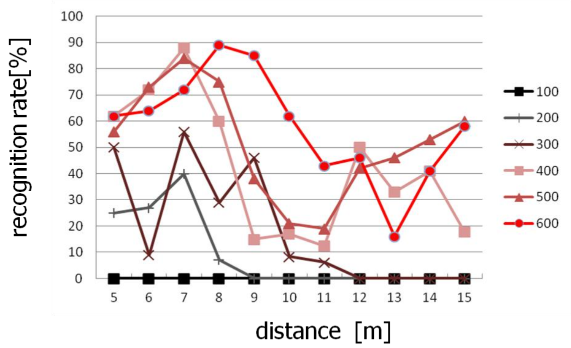

Figure 7. Recognition rate relative to distance

Looking at figure 7, it can be seen that the greater the normalized size, the higher the recognition rate is for distant distances. 100 pixels could not be recognized in all 5-15 m. 200 pixels could recognize distances as close as 5-8 m, but could not recognize after 9 m. 300 pixels could recognize farther distances and could recognize up to 11 m. 400-600 pixels could recognize all distances of 5-15 m. Based on the above, we aim to detect in the range of 5 to 15 m, so we think desirable that the normalization size should be at least 400 pixels.

VI. CONCLUSION

Prototype of a program to detect approach was made, and it was possible to cope with realistic situations by dealing with multiple people, and the program itself worked without problems. The image analysis program was expected to be very heavy for Raspberry pi. Actually, normalization with 200 pixels can finish the calculation without any problems. If this is the case, it may be possible to consider actual operation. However, if the recognition rate is considered, 400 pixels or more is desirable. But, the time required for 400 pixels is 14 seconds, which is longer than the target time. So we have to think about how to shorten the analysis time so that it can fit within the specified time even with a large size.

REFERENCES

[1] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedv, Scott Reed, Cheng-Yang Fu, Alexander C. Berg “SSD: Single Shot Multi Box Detector”