A Detecting Method of Human at the Rear of Player UsingRaspberry Pi

Yoshitsugu Yamaguchi [1] and Akira Suganuma [2]
[1]Advanced Production and Information Systems Engineering Course,
National Institute of Technology, Ariake College,
[2]Life and Environmental Science Course, National Institute of Technology, Ariake College
150 Higashihagio-cho, Omuta, Fukuoka 836-8585, Japan

Abstract

Recent years, along the IT development, the image processing has also been used in various places, for example, a car collision prevention device, a face authentication system, and so on. In this study, we have used the image processing as a human body detection.For the purposes of this study, we have determined that the "look behind the man without looking behind."As an application of this study, we thought enemy discovery system of the Airsoft game, detect system of stalking, and so on.

Keywords: Image Processing,Human Detection

1. Introduction

Unlike the human interpretation, the computer interprets the image as an array of mere numbers. Therefore, the computer cannot identify human "as a human" in input frame.

Therefore, in this study, from a few thousand of the human image and the other incorrect image, we have produced a specimen of the human body feature value.By applying this specimen to the input frame,this system tries to detect human body in the input frame.

2. Feature descriptor for human detection

As the de facto standard feature amounts used for person detecting,HOG(Histogramof Oriented Gradients) is known among image processing engineer. This is a representation of the luminance gradient of the local area as a histogram. It is possible to express a rough shape because of a high-dimensional descriptor.

Calculation algorithm of the feature descriptor is generally following.

(1) To calculate "Brightness gradient"

The gradient strength $m(u,v)$ and the gradient direction $\theta(u,v)$are derived at each pixel in the frame by the following equations:

$$m(u,v) = \sqrt{\big(I(u+1,v) - I(u-1,v)\big)^2 + \big(I(u,v+1) - I(u,v-1)\big)^2}$$

$$\theta(u,v) = \tan^{-1}\frac{I(u,v+1) - I(u,v-1)}{I(u+1,v) - I(u-1,v)}$$

(2) To make histograms by cells

A gradient direction histogram is generatedin the each cell area (consists of 5*5 pixels).

(3) To normalization by blocks

A normalization is conducted in the each block area (consists of 3*3 cells).

3. Creating cascade filesfor human detection

We describe the preparation method of discriminator for determining the presence or absence of a person.The method takes the statistics of the feature from the sample image, using "Adaboost" which is a kind of learning algorithm.The adaboost algorithm consists of four steps as following:

(1) Preprocessing
  It uses N leaning samples.
  It preparesLabels corresponding to the class.
(2) Initializing the weight of learning sample
(3) For t=1,2,... ,T (T times leaning)
  (3-1) For m=1,2,…,M (M weakclassifiers)
    (3-1-1) Selectinga candidate of weak classifiers
    (3-1-2) Calculatingan error rate
  (3-2) selectingthe lowest error rate weak classifier on its candidate
  (3-3) Calculatingthe weight oftheweak classifier
  (3-4) Refreshingthe weight of the leaning samples
  (3-5) Normalizingthe weight of the leaning samples
(4) Constructingthe strong classifier

We have definedthe purposeto use the collect statistics as "human specimen" for discovery of the human in a frame. It is multiple generated as a weak classifier with a weight. In the human body detection, we have created a "cascade classifier" which is connected multiple the weak classifiers. The outline of the cascade classifier is shown in Fig. 1. The frame captured by video camera is divided into partial images. The images are checked in one after another classifier. Each classifier calculates the similarity for human detection. When the similarity calculated by one classifier is smaller than the threshold value, the partial image is discarded.Therefore, only the partial image,

whose similarity is larger than the threshold value in all classifier, is extracted as what a person is reflected.

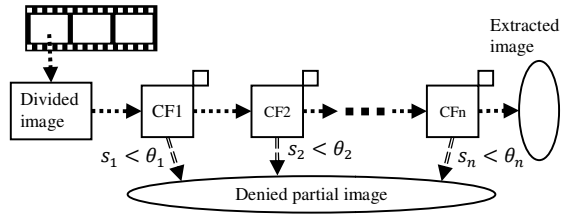

Fig.1:Cascade classifier

It checks the frame from the beginning of the weak classifiers, continues the inspection until the end of the weak classifiers. As a result, it can fast reject of inapposite images and can precise acquisition of the correct image.

It should be noted that, the machine-learning was learning under the following conditions.

- Positive samples:5369
- Negative samples:2100
- Commonattributes:standing, (almost)whole body
- Accuracy rate:0.95
- Incorrect rate:0.18
- Number of strongclassifier:20
- Type of learning algorithm: adaboost

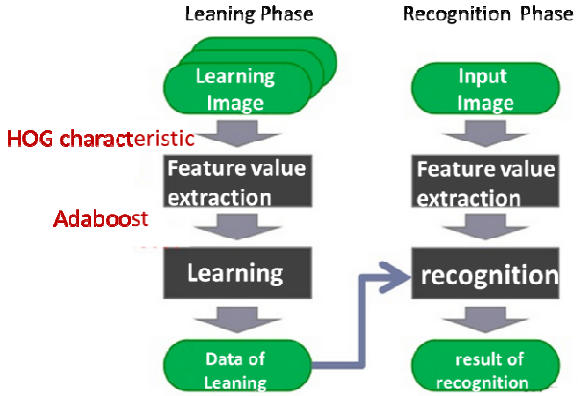

Fig.2:Image of Leaning and Recognition Phase

4. Design of our system

The main purpose of this study, is "todetect the human behind without looking back."To achievethis purpose, we constructedthe human body detection system with "Raspberry-pi", which is the small-size computer. A video camera of this system is attached on the user's head. At this time, the camera is assumed to face the rear. Based on the processing results on Raspberry-pi, it causes the feedback to the user by voice.

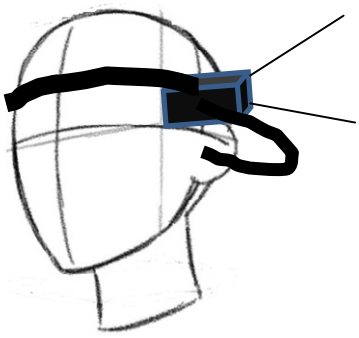

Fig.3:The wearing image of the system

5. Accuracy and problem

As the main result of the human detection using our system, we observed the feature of the following.

- It tends to make small non-detection in the frame which the person is reflected.
- It tends to detect the wood-lattice pattern-iron pillar erroneously as human beings

Fig4:Example offalse positives on this system

In the former, the failure is likely to appear in the case where the lower body is hidden. As a reason, it is conceivable that the learning image is limited to a view of the nearly whole body. We are planning to also consider such an exception in the future.

The latter failure is usually resolved in a way that make the sophisticated and complex image processing. The processing power of Raspberry-pi is, however, lower than one of the desktop PC, so it is not well suited for the latter way.

To increase the accuracy efficiently using low load technique, we devised the following process, and tried using a combination of a plurality of each.

➢ Three-part processing

An input image is divided into three parts: "view," "sky-line,"and "ground". This process is the way to denying the detection frames of other than the peripheral line-of-sight direction.If the device wearer is a normal standing way, there are not almost human body at the end of input image(upper left,upper

right,below left, and below right). Thus, this process was help discarding erroneous frames.

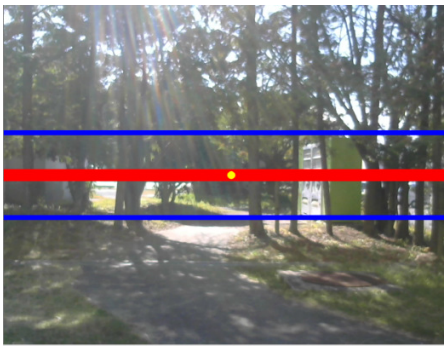

Fig.5:Three-part processing ("view" is between two blue lines)

- Ω edge matching

As a pre-preparation,we create a characteristic edge specimen which is a Ω form(edge of along shoulder-head-shoulder). Our system compares the body edges using Ω edge specimen to the input image which is binarized and calculates the similarity. The frame whose similarity is smaller than the threshold value is discarded in this process.

Fig.6:Example of shape of Ω edge

The phenomenon of detecting a non-human body almost disappeared with this process. It is not, however, flexible to the attitude change, because appearing strong limitation to the conditions of the human body detection.

Fig.7:Non-effective example frames

In addition, this solution technique is difficult to demonstrate the effect in the painful situation taken by the human body of the edge.Besides, there was also to pick up an extra edge in image with the situation of strong contrast.Improvement of these shortcomings has become a future work.

- Skin color detection

This methoddetects flesh color area of human body(face,arms,and legs), and checks the shape of the vector of the center of gravity of the skin color region. It calculatesthe similarity of human body shape. Itdecides true or false for the input frame with the similarity.

Fig.8:Detected skin color(converted blue color)

If skin color is present in the background, this process does not work well. If the skin color as noise on the outside of the human body, the noise can be removed doing "expansion" processing and "contraction" processing. If skin color is present in a wide range,using the labeling process, which is a process of identifying the area, we will plan the process that denying area having over threshold.

6. Conclusion

This time, we have developed a human body detector using Raspberry-pi which is a very small computer. At present, this system operates well if all conditions are good. However, background, attitude, and lighting are likely to change, it often does not work well.

It may be difficult to make a universal detector on the Raspberry-pi. The first reason is that Raspberry-pi doesn't have much power for processing complex algorithms. The next reason is that background, attitude, and lighting have a few limiting factor(ex:threshold of color factor).However, by limiting the location and the date and time that the use, by performing the adjustments according to the background and the illumination of the place, we believe that there is a good chance.

As future challenges, we think about the following things.

- To improvement of the operating speed of when multiple people are exist in the input frame.
- To reduction of the non-detection rate.
- To increase of the detection rate.
- To consider application of this system.
- To develop application system in line with considered plan.

References

[1] N.Dalal, B.Triggs: “Histograms of oriented gradients for human detection,” Proc.of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp.886-893,2005.

[2] Yuji YAMAUCHI, Takayoshi YAMASHITA, HironobuFUJIYOSHI: “Human Detection Based on Statistical Learning from Image,” IEICE Journal Vol J96-D No.9 pp.2017-2040, 2013, (in Japanese).

[3] Ryousuke MUTOU, Kazutaka SIMADA, and Tutomu ENDOU: “Hand Shape Recognition based on SVM and online learning with HOG,” Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.16, No.5, pp.1-3, 2010.