International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.12 (2015), pp.217-224
http://dx.doi.org/10.14257/ijmue.2015.10.12.22

An Improvement to the Algorithm of Human Body Detection in Video Surveillance

Jingjing Yang, Benzhen Guo and Xiao Zhang*

School of Information Science and Engineering, Hebei North University, Zhangjiakou, Hebei, China
*r78z-yang@126.com,*r78z@foxmail.com*

Abstract

*Human body detection in video surveillance is an important research point, which has been devoted for many years by researchers. But it can't detect human body with the feature of HoG and classifier of SVM. This paper proposes an improvement measure to the traditional algorithm of human body detection in order to detect human body in video surveillance in real time. Through the Integral images to speed up the computation of features of HoG and HoD under the condition of Kinect, It can ' t detect human body in video surveillance in real time. As human body detection using HoG features can't meet the real-time requirement in the current time. This paper abandon the time-consuming processing of computing features of HoG with three line interpolation and Gauss filter in terms of traditional algorithm of human body detection, and This paper introduces the notion of Integral images to speed up the computation of the features of HoG and HoD with the help of Kinect. As a result, Our improved algorithm can get a better time performance than ever, which can get the average detection rate of 30~40 FPS, current performance getting promotion greatly than ever, under the condition of CPU 3.10Ghz, RAM 4 GB, 640*480 detection window.*

Keywords: *Computer Vision; Video Surveillance; Detection of Human Body; Kinect; HoG; HoD*

1. Introduction

The technology of human body detection is able to automatically detect static pictures or continuous human body frames in video surveillance to find out the human body, through which we can analyze the behavior of the detected human body and further make intelligent judgments. But because of the technology involved in several technical subjects, such as digital image processing, pattern recognition, computer vision and artificial intelligence *etc.* it also reflects the difficulty in studying the technology. For example, different clothes, different postures and different weather will have a great impact on result of the detection [1]. The technology of Human body detection is still a big challenge for in research and engineering. But the research of the technology has been becoming a hot research area of rapid development in recent years, and the application of this technology is wider than ever. More and more researchers at home and abroad participate in the activity, and they put forward many new methods and new technologies, greatly improving the detection accuracy. The technology has a wide range of applications, such as residential intelligent monitoring system, intelligent transportation, robots, human-computer interaction, the public security surveillance and control system *etc*. [2].

There are many researchers from abroad. In 2011, Luciano Spinello put forward using Kinect for the detection of human body. Author used HoD (Histogram of Oriented Depth)

* Corresponding Author

ISSN: 1975-0080 IJMUE

features from depth data of Kinect Depth detector and HOG (Histogram of Oriented Gradient) features from RGB data of Kinect RGB detector, constructing Combo-HoD detector combined with the probability output SVM when detecting human body indoor, it could get the detection rate of 30FPS under the condition of GPU; As Luciano Spinello's algorithm was achieved under GPU, detection rate in CPU in general still can't get the running effect in real time [3].

There are also many researchers at home. For example, there are many researchers in SUPCON, in Zhejiang province. The technology is imbedded in the project of the research of robot. And there are researchers in Zhejiang dahua Technology Company. The technology is imbedded in the project of intelligent vehicle.

In this paper we propose an improvement measure to the traditional algorithm of human body detection in order to detect human body in video surveillance in real time. Through Integral images to speed up the computation of features of HoG and HoD under the condition of Kinect, We can detect human body in video surveillance in real time. As traditional human body detection using HoG features can't meet the real-time requirement in the current time. We abandon the time-consuming processing of computing features of HoG with three line interpolation and Gauss filter in terms of traditional algorithm of human body detection, and we introduce the notion of integral images to speed up the computation of the features of HoG and HoD with the help of Kinect. As a result, we can get a better time performance than ever. Our algorithm can get the average detection rate of 30~40 FPS, current performance getting promotion greatly than ever, under the condition of CPU 3.10Ghz, RAM 4 GB, 640*480 detection window.

2. Related Work

The technology of Human detection is to detect human body from static images or continuous video frames. From the development of the technology of the human body detection up to now, it can be divided into two categories: the traditional detection techniques (non statistical learning technique) and statistical learning technology of the human body detection [4].

As for the traditional detection techniques, The first step is to detect moving objects with some specials techniques, such as the method of background difference, the method of difference of adjacent frames, the method of optical flow, and the second step is to analyze and judge the results of extracted foreground from the first step [5]. Namely according to the shape and proportion of each components in prior human, or a series of rules of the composition of human skin, people analyze and match the underlying characteristics to determine whether there is the human body in the detection window. The judgment part of the detection of Human body is usually used after extraction of the moving target. The result of traditional technology isn't very good, easily affected by complex environment and object occlusion, more susceptible to interference from low-level features noise, which reduced the detection effect. The typical and traditional techniques of the detection of human consist of: the method of image segmentation and the method of skin color. The method of image segmentation is put forward by America professor Shi Jianbo, born in China, and he used underlying characteristics of the images to segment out the human body and background. The method of skin color is to transform RGB color space of normal color to HSV color information, which can reduce the impact from illumination on human body detection [6].

As for the technology of human body based on statistical learning, it was firstly put forward in a doctoral thesis by Dalal. Dalal present the features of human body using histograms of oriented gradients (HOG), and then use the support vector machine (SVM) to train and classify the features of human body through the features from above, to determine whether there is a human body in the detection window. The advantages of this method are detecting human body in static images efficiently and the rate of false

detection is very low; the shortcomings are the slow efficiency of detection and the lack of real time [7]. Later, Qiang used the notion of integral image to speed up the detection rate on human body, but the effect was not very good. Although the detection speed had been greatly improved, but the real effect was still not up to the real-time requirement. At the end of 2011, Luciano Spinello used the Kinect for human detection with GPU, and achieved the effect of real-time detection. But in the normal condition of CPU, detection still can’t meet the requirements of real-time, in this paper our algorithm is based on the previous work, and we put forward the notion of integral image to accelerate the rate of human detection based on Kinect, realizing it with CPU instead of GPU. The use of this technology makes human detection technology obtain good development. In the normal condition of CPU our improved algorithm can achieve the effect of real-time effect [8].

In outdoor conditions, as for the non-rigidness of human body framework, the performance of detection results are often vulnerable to external factors, for example, body postures and light conditions. In 2001, Paul Viola *et al.* put forward Haar-Like feature, with the help of classification algorithm of Adaboost, to detect objects [9]; And in 2005, Dalal & Triggs introduced a kind of HoG (Histogram of Oriented Gradient) feature, with the classification algorithm of SVM (Support Vector Machine) to detect human body. The effect of detection proved good, however, the speed was still very slow. And it can’t meet the need of the real time requirement [10]; In 2006, Qiang Zhu then presented a detection algorithm, that combines IHoG (HoG with Integral Image) feature and Adaboost learning and classification approach, putting forward an accelerated way to calculate HoG feature embedded in the technology of integration images. The rate of correct detection had been greatly improved, but the rate of false detection was no better than the former’s, and also its training consumed a long time [11].

Dalal & Triggs defines the size of sliding detection window 64*128 pixels, each block 16*16 pixels, each block 4 cells. The size of each cell is 8*8 pixels. When detecting or training, the sliding window slides by 8 pixels from left to right or from top to bottom. The way of computing the feature of traditional HOG is repeated to a certain degree, that means blocks overlap each other every 8 pixels. Each cell contributes to a 9-bin Histogram of Oriented Gradients (HoG), then the algorithm adopts three linear interpolation, as well as gauss filter a 36-D vector stands for each block and the vector is normalized by the standard of an L2 unit length. From above, we can conclude that each 64*128 sliding window contains 7*15 blocks, forming a total of 3780 features .After the computation of HoG and the achievement of the vector of 3780-D, linear SVM classifier can train or classify the feature. Through the analysis of the computation of the traditional HoG feature, we conclude that the process using three line interpolation and Gauss filter cost most of time. The rest of the time is much less than that of process before. There several steps to form the final HoG feature: Firstly, we get the input image form the video; Secondly, we shall normalize the image in the respect of color; Thirdly, we compute the gradients of the image using the vector [-1,0,1]; Fourthly, we go on with the process of three line interpolation and Gauss filter, which occupies most of the time when we compute the HoG feature. The rest of steps are the process of training and classifying with the technology of classifier SVM and the process of detection of human body. The calculation process of traditional HoG features is presented in Figure 1:

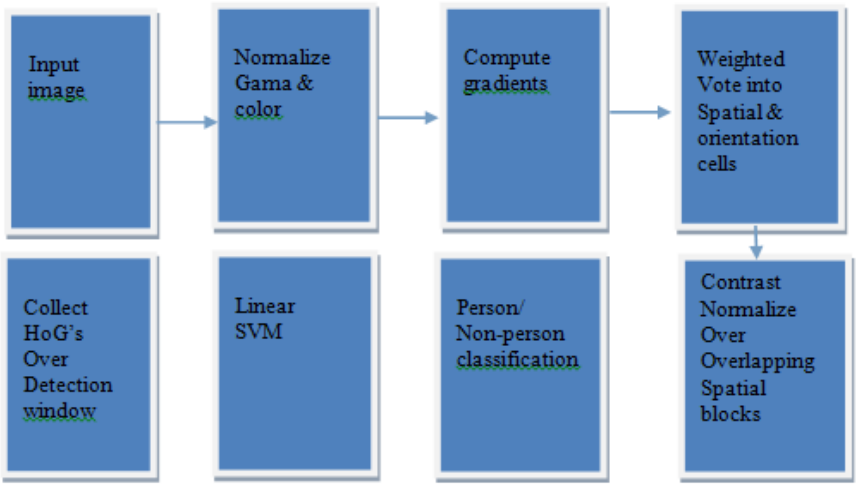

Figure 1. The Process of Traditional HoG Feature

3. Algorithm Realization and Comparison

In 2003, Viola proposed Haar wavelet rectangle feature the feature has a good impact on the presentation of the structure of characteristic directions. For a detection window of 24*24 pixels, there are about more than 16000 kinds of Haar features in its internal, but not all of these Haar features are effective representations of face feature information, we shall choose those features which can effectively present the characteristics of face feature and choose better algorithm to select Haar features. It means that we shall choose those rectangular which can effectively express human face information box. There are two kinds of rectangular template: black and white. The presentation of characteristics of Haar rectangular is the minus of the sum of a rectangular pixels and the sum of the other rectangular pixels. There are two ways to present the features of Haar rectangular box. One is presentation of horizontal and vertical rectangles, the other is presentation of rectangles with 45 degree direction. For the former way of presentation, There are five ways to present this kind of feature: the minus of the sum of left rectangular pixels and the sum of right rectangular pixels; the minus of the sum of top rectangular pixels and the sum of down rectangular pixels; The presentation of this kind of Haar feature is the difference of the sums of three rectangle where : there is a large rectangular box with three small rectangles, the presentation of the kind of Haar features is minus of the sum of both ends of the rectangular pixel and the 2*sum of the middle rectangle and the absolute value is the final feature; the presentation of four rectangles is that there is a large rectangular with four small rectangles and final result is the minus of the sum of two rectangles in the diagonal direction and the sum of the other two rectangles in the diagonal direction; center surround characteristics is absolute value of the minus of 8* the sum of the center rectangular and the sum of the rest of rectangles. For the latter way of presentation, after proposing several kinds of Haar features by Viola, Lienhart put forward a notion of 45 degree feature, which is the extension of the above several features. The new features can describe the contour features of the edge better. This feature can be presented by three ways, two rectangles with 45 degree angle, three rectangles with 45 degree angle, centered on rectangle with 45 degree angle. The way of computation of this feature is as same as that of computation of horizontal and vertical rectangles.

Integral images is firstly applied to the detection of human face by Paul Viola. Through the using of integral images, with four kinds of rectangular area, He

accelerated the speed of the computation of Haar-wavelet features. As for the defect of the only four kinds of rectangular area, Porikli extended them to any arbitrary shape of rectangular area [12] later, Qiang Zhu used the notion of integral images to the application of human detection with the cascade HOG features, however the training stage consumed much time. Inspired by the idea of Qiang's, the notion of the integral images is applied to the detection of human body.

After introducing the notion of integral image, We think of the detection of the human body using Kinect, and the process of the algorithm is almost as same as that of Dalal-Triggs', such as the size of detection window, block and cell, the step of striding and the computation of features of HoG and HoD. The HoD feature is obtained by the same way as the HoG feature, and the difference is the resource of source data. We apply the notion of integral images to the computation of the features of HoG and HoD. Gradient direction is then quantified into 9 direction maps, each map standing for intensity value of each gradient direction. 9 integral images can be calculated rapidly by the 9 maps from the above processing when the image is firstly inputted in the preprocessing stage. These 9 integral images can be used as features for HoG or HoD in a constant time when we need features of a certain area. The complexity of the computation of features from 9 direction maps is constant, which is much faster than any other algorithm in the time efficiency. Through the result below, we can conclude that the introduction of integral image is very important and the time efficiency of the improved algorithm is much better than any other algorithm similar. Through the introduction of the notion of integral image, we speed up the computation of the feature of HoD and HoG so that we can get the feature just as the way we look up for a Table in a constant time. The complexity of the achievement of integral image is O(N), but as soon as it is computed, it is convenient to compute the feature of HoD and HoG. And the time performance is constant, reducing the time-consuming stage. The application of integral image is presented in Figure 2:

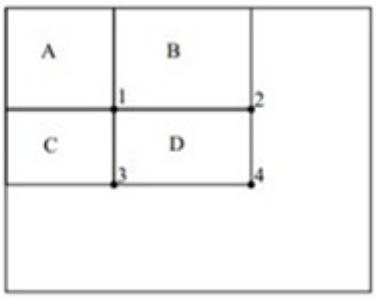

Figure 2. The Calculation Process of Integral Image the Value of Integral Image at Location 1 is the Sum of Pixels in Rectangle A, the Value at Location 2 is A+B, at Location 3 is A+C, and at Location 4 is A+B+C+D. The Sum within D, which presents the Gradient Sum of a Relevant Direction, can be Computed as 4+1-(2+3)

It consumed a lot of time to compute the features of HoG and HoD with the process of three linear interpolation and gauss filter in the traditional computation of the feature of HoG and in the process of detecting human body by Luciano Spinello using kinect. As for the latter algorithm, the detection can achieve real-time effect under the condition of GPU. This paper applies the notion of the Qiang's integral images to the calculation of the features of HoG and HoD, and we drop the traditional process of three linear interpolation and gauss filter. Just like the picture

below, there are several steps to detect human body. And the processing is parallel using the Kinect with the RGB detector and Depth detector. Firstly, we get the inputted images; Secondly, we normalize the color with the process of gama transformation; Thirdly, we compute the gradients of 9 angle bins, which is different from the traditional computation of the feature of HoG and is our main solution to the problem of real-time requirement; and the rest of the steps are similar with the traditional process of the computation of HoG feature. After introducing the notion of the integral image, we observe that time performance has been improved greatly. And then the features of HoD and HoG are then sent to train probabilistic linear SVM separately [13], determining whether the content in a detection window is positive or negative. The figure below shows the process of the whole improved algorithm of human body detection based on Kinect using integral images:

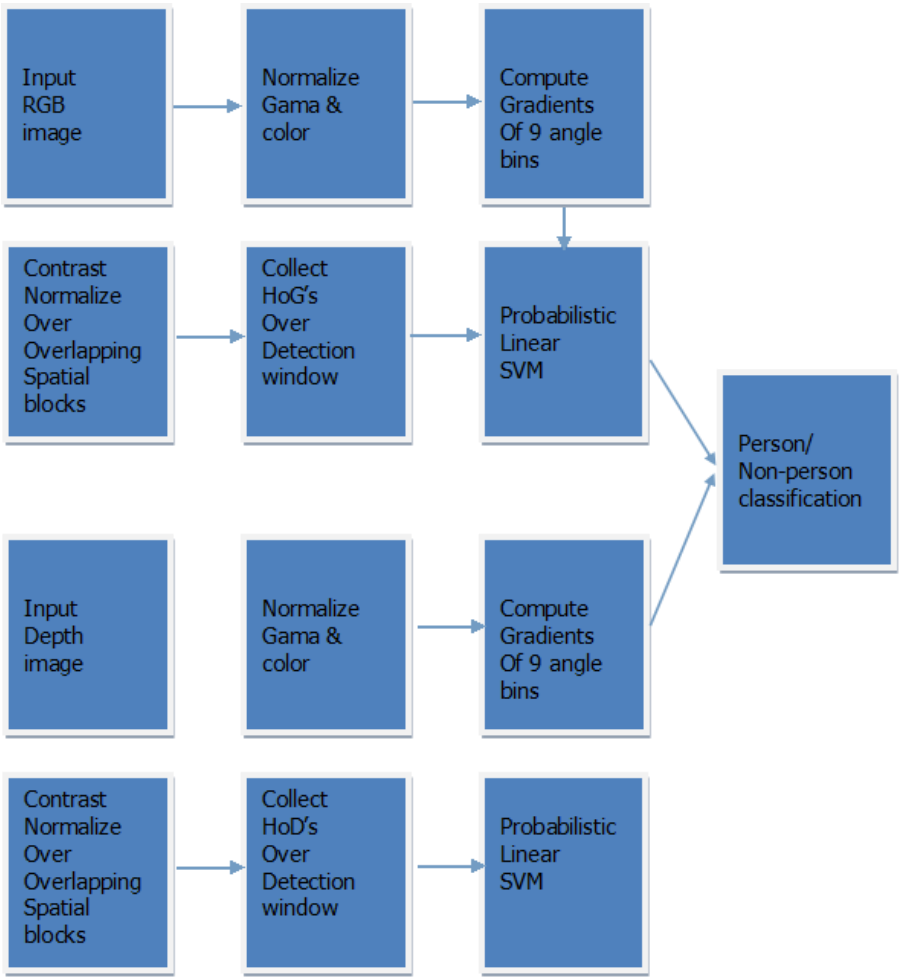

Figure 3. The Process of Improved Algorithm Using Integral Images

4. Experiment Results and Analysis

In this paper, The data frames of Kinect are collected indoors, and these frames are run separately on the traditional algorithm of the HoG features presentation of Dalal's, Luciano Spinello's algorithm under the condition of CPU and the improved algorithm on the basis of Luciano Spinello's algorithm using integral images to speed up the calculation of the features of HoG and HoD, the results of three operations are presented in Table 1 below, and the result picture of the detection of our algorithm is presented in Figure 4:

Figure 4. The Result Picture of Our Improved Algorithm

Table 1. The Realization and Comparison of Three Algorithms

	Window size	Basic frequency (Ghz)	RAM(G)	Frame number	Running time(s)	FPS	Miss rate (FPPW=10-4)
Dalal's algorithm using HoG features	640*480	3.10	4	1083	3629.05	3.36	0.34
The luciano's algorithm running on cpu	640*480	3.10	4	2*1083	93.256	11.48	0.476
The improved algorithm of luciano's using integral images	640*480	3.10	4	2*1083	27.201	39.82	0.377

According to the results of operations and comparison of results. On one hand, we can learn that the traditional algorithm of Dalal Triggs's can get 3.36FPS, on the other hand, the running result of Luciano Spinello's algorithm using CPU is 11.48FPS, those time performance is 3.4 times than the traditional Dalal's performance, and our algorithm's performance is also 11.86 times faster than the traditional Dala's performance and is 3.48 times faster than the algorithm of Luciano's using CPU. Miss rate of our algorithm is exactly as good as the other two algorithms, however, detection performance of our paper is greatly improved. And the result of our improved algorithm is presented, which can detect multi-scale human body. As the improved algorithm is mainly about the time performance, we only show the result of our improved algorithm.

5. Conclusion

In this paper, human detection algorithm based on Kinect applied integral images to speed up the calculation of the features of HoG and HoD. Experimental results show that the detection performance of our algorithm is greatly improved as for the better detection performance and the better time performance, and our algorithm can be used to the indoors human body detection in video surveillance with Kinect.

Acknowledgment

Foundation item: The Major Programs of Hebei North University (No.ZD201301), The Youth Foundation of the Education Department of Hebei Province(No.QN2015225) and Engineering Research Center of Population Health Information, Hebei Province.

References

[1] L. Spinello and K. O. Arras, "Intelligent Robots and Systems", **(2011)**, pp. 3838-3843.
[2] M. Luber, L. Spinello and K. O. Arras, **(2011)**, pp. 3844-3849.
[3] L. Spinello and K. O. Arras, "Intelligent Robots and Systems", **(2011)**.
[4] B. Babenko, M. H. Yang and S. Belongie, "Pattern Analysis and Machine Intelligence", vol. 33, **(2011)**, pp. 1619-1632.
[5] C. Aeschliman, J. Park and A. C. Kak, "Computer Vision and Pattern Recognition", **(2010)**, pp. 1371-1378.
[6] H. T. Nguyen, Q. Ji and Arnoid W. M. Smeulders, "Pattern Analysis and Machine Intelligence", vol. 29, **(2007)**, pp. 52-64.
[7] M. C. Ho, C. C. Chiang and Y. L. Chen, "ICIG", **(2007)**, pp. 524-529.
[8] B. Babenko, M. H. Yang and S. Belongie, "Computer Vision and Pattern Recognition", **(2009)**, pp. 983-990.
[9] P. Viola and M. Jones, "Rapid Object Detection using a Booster Cascade of Simple Features", CVPR **(2001)**.
[10] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", **(2005)**.
[11] Q. Zhu, S. Avidan, M. C. Yeh and K. T. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients", **(2006)**.
[12] R. Lienhart and J. Maydt, "Image Processing", **(2002)**.
[13] H. T. Lin, C. J. Lin and R. C. Weng, "A Note on Probabilistic Outputs for Support Vector Machines", **(1999)**.

Authors

Jingjing Yang, received his Master Degree from Lanzhou Jiaotong University in 2012. He is currently a lecturer of the school of information science and engineering at Hebei North University. His research interests are in the field of image processing and internet of things.

Benzhen Guo, a lecturer in the School of Information Science and Engineering, Hebei North University, China. His research interests are in the field of medical information and internet of things.

Xiao Zhang, a Professor in the School of Information Science and Engineering, Hebei North University, China. His research interests are in the field of medical information and internet of things.

