

Study of multiscale detection in near distance image for numbers of people in elevator car

Xu Bin^{a*}, Yang Jing^b, Xu Feng^c, Guo Mangmang^d

NO.5 South Jinhua Road, Xi'an, Shaanxi, China

^axbuingoaix@163.com, ^byjzhd@163.com, ^cxf422256093@163.com

^d1329321963@qq.com

Key words: near distance image; multiscale; region image; overlapped region

Abstract: It is necessary to have a video surveillance of the quantity of people in construction elevator to take down the emergence of casualty like death in group. Considering the relatively tiny spacing in the car and complex background, the results of existing people counting algorithm are not satisfactory. This paper proposes a machine learning recognition method of multiscale ROI image processing for the working conditions of object occlusion and uncertain image size. Multiscale image change according to region is to improve accuracy to object recognition. Then overlapped processing method is used in border area targets. The results of experiment show that the algorithm for object recognition is better in the condition of object occlusion and uncertain imagesize, and greatly reduces runtime and improves accuracy for the condition of complex background in elevator.

Introduction

It is necessary to have a video surveillance of the quantity of people, because the overload car of construction hoist and elevator and vehicles will be lead to the emergence of casualty like death in group. Counting for numbers of people in elevator car is often used image recognition technology. Literature[1,2,3] using frame difference and background modeling to identify the moving object, but more difficult to identify and track moving targets in the workingcar condition of vibration and tiny spacing and crowing. Literature [4] proposes a counting algorithm for the numbers of people based on improved Hough transform which is used to detect head contour. Literature [5] proposes algorithm for counting numbers of people using the feature of Histograms of Oriented Gradients based on method of machine learning. Both of these methods for people counting in simple background elevator car have achieved some results. But existing algorithms have large errors in the working condition of complex background and poor edge detection, and object occlusion in close-up image of a tiny space. Considering unknown size object in close-up image, Literature [6] proposes a method using multiscale detection window to scan image, and literature [7] used a multi-tiered scale scanning to recognize object. Both of these methods have low efficiency and need many times processing to image recognition. In the pedestrian recognition, literature [8] used a partition matching technology of multiscale database, and established different scales of target template, and matching. Construction hoist is the important vertical transport in construction site. Considering the characteristic of complex background and uncertain objective image size in near

distance image, in counting numbers of people for construction hoist car case, this paper proposes a machine learning recognition method of multiscale ROI image processing. It aims to greatly reduce runtime and improves accuracy for the condition of complex background in hoist.

Multiscale ROI image processing

Multiscale processing for ROI. Target detection method based on machine learning is to use training samples for classifier, then scanning image to detect target according to a certain algorithm. In this process, a window will be missed if its ratio of foreground pixels is less than a certain threshold value. Five targets (helmet) is not the same size within the construction elevator car shown in Figure 1. Target in the detection window size proportion is uncertain if using the same size detection window for scanning images. Suitable features of target in the detection window with high effective pixel ratio can be extracted shown in Figure 2(a). Target in detection window accounted for less pixels and more background information will cause lower detection accuracy in the imaging smaller areas shown in Figure 2(b). Target 4 and 5 in Figure 1 in adhesion can appear in a detection window as shown in Figure 2(c). Classifier training samples generally do not contain more than one target, so Figure 2 (c) will appear undetected or false detection.



Figure 1 Construction of the elevator car interior scene

In this paper, the multiscale ROI image processing method is to detect the image to be detected in different areas, in order to ensure the foreground image pixels in detection window have a certain proportion. The image is divided into A, B, C three local area to be detected according to the size of the target image varies with the position of the image, in Figure 1 as an example, circular target in area A is 88x88 pixels, and circular target in area B is 64x64 pixels, circular target in area B is 40x40 pixels. In the process of scanning, the target of figure 1 will be a different situation as shown in Figure 3.



suitable (b) Smaller (c) Adhesions

Figure 2 the same detection window in different imaging helmet

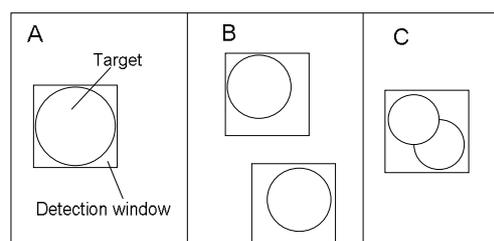
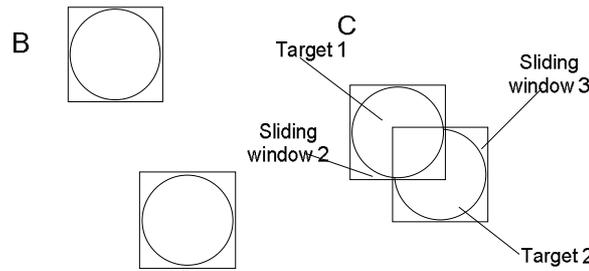


Figure 3 multiscale region image processing

We take a local amplification multiscale processing to B and C zone in order to ensure more pixels of foreground in that window, and enlarge 1.37 times and 2.21 times respectively as shown in Figure 4. After the target in area B enlarged, target pixels ratio have been increased and background pixels have been decreased in the window to be detected, then recognition rate can be improved. Two adhesions targets will not appear in a window simultaneously after target amplification for area C. So two targets will be identified respectively. This method solve a certain degree of adhesion target identification.

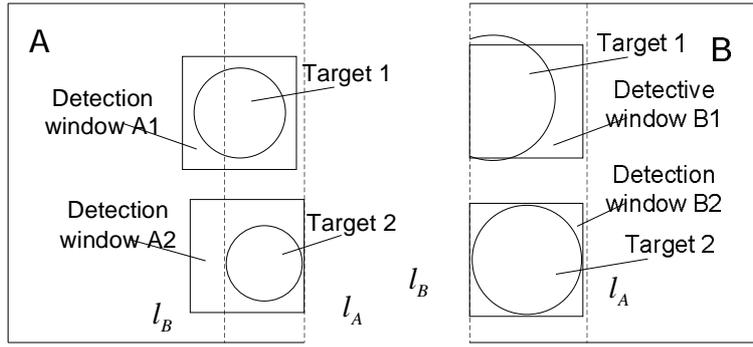
overlap processing of area boundary. According to the above principle of multiscale and divided



(a)Area B multiscale detection (b)Area C multiscale detection

Figure 4 Regional multiscale detection

area image processing, there will be a case of a target across the two adjacent regions. If taking a simple line to divide target area, a target across the two adjacent regions will produce defective target image in image processing of multiscale and divided, and lead to undetected. According to the principle of sliding window scanning detection, this paper take a method called overlap processing of area boundary. The border of area A is l_A , and the border of area B is l_B , as shown Figure 5. There had an overlap between border l_A and border l_B . The overlap will be treated as part of the A and B, respectively. If the detection window size is $m \times m$ pixels, image in area B will be amplified k times in multiscale processing and overlap in the original image pixels are m/k . The target 1 and 2 in Figure 5 appears in the area of the overlap. The foreground of target 1 in the detection window position in the center can be detected. But target 2 in the detection window A2 have a small occupation ratio, and its position slants to right to cause the change of the characteristic vector. So target in the detective window cannot be detected. The target 1 and 2 of overlap region in area B will be amplified in multiscale processing in scanning recognition of area B. Target 1 cannot extract the appropriate and effective characteristics to identify itself, because it cannot be contain very well after amplification. Target 2 after amplification in detective window can extract effective feature vector. Therefore, the target in overlapping area can be appropriate detected according to the identification method of overlap region.



(a) A regional processing (b) B regional processing

Figure 5 Overlapping area detection

In order to avoid image Mosaic phenomenon in multiscale processing in some degree and raise the speed of the algorithm at the same time. The bilinear interpolation method is adopted to amplification when processing the image with the method of multiscale and divided area image processing, such as formula (1). The coordinate values of a target pixel divided by the scaling factor respectively. And the floating point coordinates $(i+u, j+v)$ of the original image are mapped by the reverse transformation , where u and v are floating point numbers in the interval after taking over coordinates. So the target pixel value $f(i+u, j+v)$ can be calculated by the formula (1).

$$f(i+u, j+v) = (1-u)(1-v)f(i, j) + (1-u)v f(i, j+1) + u(1-v)f(i+1, j) + uvf(i+1, j+1) \quad (1)$$

The pixel value of the original coordinates (i, j) is $f(i, j)$, and i and j are nonnegative integers. Target pixel values of resized image $f(i+u, j+v)$ is composed of four adjacent pixels in the original image according to the different weights to calculated.

Statistics of the number of people in elevator car based on multiscale image processing

Considering the construction site personnel must wear safety helmet and the helmet have three-dimensional characteristic, this paper combine Local Binary Pattern (LBP) of pyramid texture feature [9] with liner support vector machine (SVM) classifier [10], and take the method of multiscale region image processing to identify the safety helmet. The overall flow chart shown in Figure 6. Firstly, we need to train the classifier offline. Then we select all kinds of positive and negative samples in the area of different from different angles, and process with multiscale processing before extract the texture feature of LBP from those samples to train the SVM classifier . Then it is the detection of real-time images. According to different field condition of elevator, the image is manually partitioned, and region is processed by multiscale image processing. And then the sliding window is scanned with a step size of 6 unit pixels, and we extract LBP Pyramid feature from the scanning window. So we can use the trained classifier to identify and count the number of targets.

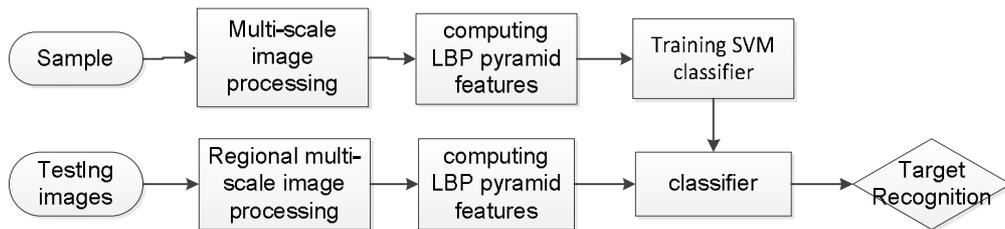


Figure 6 Flowchart of elevator car for identification number based on multiscale image

Experimental results and analysis

For the case of an elevator car shown in Figure 7, the targets have different location and then its pixels are 88x88, 64x64 and 40x40 respectively. So the detected image is divided into three dimensions area as shown in Figure 7 in blue, green and red trichromatic rectangle domain of A, B and C respectively.

In the condition of normal working condition the shape of the helmet for passenger is varied and its angle is varied as well. In this paper, the positive and negative training samples are selected as shown in Figure 8. The positive and negative sample is composed of 2716 positive samples and 5359 negative samples, which contains three different sample of different pixel target. In order to avoid the feature vector dimension is too large or too small, this experiment selects the target size of 64x64 pixel as a reference. The different scale samples are normalized, and the LBP Pyramid feature vector is calculated, and the SVM is used to train the classifier.



Figure 7 Multiscale partition detection

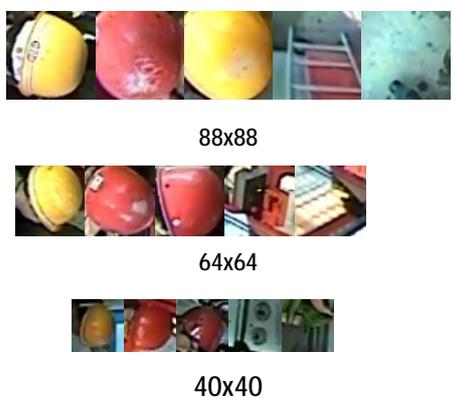


Figure 8 Samples for training

The safety helmet is identified in the working condition of elevator car using the trained SVM classifier. The Figure 9 shows the results of a variety of working conditions, which have targets in different areas. In Figure 9, the target appears in different regions of the whole car. There are also the case of helmet in the overlapping area, as shown in Figure 9(a) target 1 and 4, and in Figure 9(b) target 2, 4 and 5, and in Figure 9(c) target 2, 3 and 4. Even if there is an image adhesion, as shown in Figure 9(b) target 4 and 5, and in Figure 9(c) target 3 and 4. It is obviously that the algorithm has a better recognition accuracy for the above all kinds of cases. In a large number of tests are also found in the case of false recognition, which a false detection is shown when the background texture of the passenger is similar to that of a helmet as shown in Figure 9(b).



(a) Example 1 Of test results (b) Example 2 Of test results(c) Example 3 Of test results

Figure 9 Detection renderings

At the scene of the construction elevator, we conducted extensive testing. We have tested in work personnel boarding peak periods and normal working time on the construction site in the experiments. This article is extracted 468 different condition photos. This algorithm is an error of 92% with the above area multiscale recognition algorithms. So the algorithm recognition rate in such complex conditions is relatively high. This article puts forward the regional multiscale processing method comparing with the method of literature [7] layered scanning, the two algorithms can identify the target image size not sure and have the almost same identification accuracy for the construction elevator working condition of Figure 7. Based on the multiscale processing of the three regions, literature [7] adopted a three-layered scanning to identify different scales, Table 1 shows the comparison of the two algorithms' rate. This multiscale region image processing algorithm of this article is nearly 2 times faster and better in real time.

Table 1 The rate of detection algorithm comparison

method	The classifier training time	Detection time
Layered scanning	4.2minutes	602ms
This paper	4.2 minutes	243ms

conclusion

For complex background of elevator car, small space, close image adhesions and size uncertain conditions, we use machine learning recognition technology. This paper proposes a multiscale regional image processing method that can effectively solve the complex background and close image adhesion and uncertain size target recognition. The overlap processing of area boundary solves the target recognition for boundary region, which runs faster. Through field testing, multiscale region image processing of the proposed algorithm has high detection efficiency, runs faster, and is suitable for complex background, image recognition for complex background or adhesion objectives and the size of uncertain targets.

Acknowledgements

This work was financially supported by Shaanxi Provincial Department of Education Industrialization Fund Project (2013JC25) and Shaanxi Province Science and Technology Research and Development Program funded project(2014TG-01).

References

- [1] GAO Mei-feng, LIU Di. Moving object detection based on consecutive blocks frame difference and background subtraction[J]. *Application Research of Computers*, 2013(1):299-302.
- [2] Chen Mingsheng, Liang Guangming, et al. Fast moving object detection method using temporal-spatial background model[J]. *Journal of Image and Graphics*, 2011,16(6):1002-1007.
- [3] Jin Haiyan, Xiong Qingyu, Wang Kai, Shi Weiren. Study of a Counting Method for the number of People in the Elevator based on Image Processing Technology[J]. *Chinese Journal of Scientific Instrument*, 2011,32(6):161-165.
- [4] YING Jun, LI Qiang. Counting algorithm for numbers of people in elevator car based on improved Hough transform[J]. *Computer Engineering and Applications*. 2014,50(19):165-168.
- [5] ZHOU Xun, TAO Qing-chuan. Research on the Statistical Methods of People Number in Elevator Based on HOG[J]. *Modern Computer*, 2014(1):42-45.
- [6] GUO Lie, GAO Long, ZHAO Zongyan. Pedestrian Detection and Tracking Based on Automotive Vision[J]. *Journal of Southwest Jiaotong University*. 2012,47(1):19-25.
- [7] ZHANG Yanmin. Pedestrian Detection Based on Machine Learning[D]. Shijiazhuang:Hebei Normal University, 2011.
- [8] Jorge García, Alfredo Gardel, et al. Directional people counter based on head tracking[J]. *IEEE Transactions on Industrial Electronics*, 2013,60(9):3991-4000.
- [9] Wang Wei, Huang Feifei, Li Jianwei, Feng Hailiang. Face Description and Recognition by LBP Pyramid[J]. *Journal of Computer-Aided Design&Computer Graphics*, 2009,21(1):94-100.
- [10] Vapnik V N. The nature of statistical learning theory[M]. New York:Springer,1995.