

Small digital recognition in gravel aggregate production scenario

Ling Li, Kai Liu^a, Fei Cheng and Long Li

School of Computer Science and Technology, Xidian University, Xi'an, China

Keywords: Small Digital Recognition, Dissimilarity Characteristic, Cascade Identification, Structure Characteristics, HOG

Abstract. The production process of aggregate industrial is monitored by the identification of the digital. In the video, the digital character whose size is less than 12×15 (The width of the digital image is 12 pixels, and the height is 15 pixels.) is defined as the small digital. The recognition effect is evaluated by the recognition rate and speed. Because of the small digital of pixels in the video, the numerical recognition algorithm based on template matching is not ideal for the small digital recognition. In order to solve this problem, this paper presents a special number of cascaded recognition process. Firstly, the SVM classifier based on HOG features is used to recognize the number of images, and the special digital of the identification results is based on the two level recognition. Compared with the traditional template matching method, this method can improve the respective average recognition rate of the small digital in production and shipment of two scenarios by 27.36% and 5.05%.

1 Introduction

Numerical Recognition is one of the most active research topics due to its wide range of applications such as surveillance, License plate recognition and meter digital recognition. The numerical recognition is early based on template matching. Kim K proposed a recognition method based on learning ^[1], Bin M and Bai Q used the recognition method based on Neural Network for the meter ^{[2][3]}. Qin X also used a method based on the neural network for the recognition license plate. However, the neural network algorithm is relatively complex. Wu J, Liu Yang proposed a robust digital recognition method with resisting noise ^[4]. Cao Weijiong, Li Yujie proposed a digital recognition methods based on support vector machine ^[5].

In aggregate and other building materials industry, the factories through the numerical recognition of the camera and surveillance control the process, and have explicit requirements for the recognition rate and recognition speed. The average recognition rate of the traditional template matching method is only 69.99% and the recognition method based on neural network recognizes 5 digits per second. Thus,

the recognition rate of the template matching method is lower, the speed of the method base on neural network is slower. Because of the camera's angle and low image resolution, the robust recognition method with noise is not ideal in aggregate production industry, based on the recognition method based on support vector machine, there are errors in the process of similar numbers. The previous methods are not still satisfactory due to the low resolution, fuzzy pixels in the industrial scenarios, therefore, we use a new method based on cascade classification of special digits. Under the condition of industrial speed, ours method improves the recognition rate and has more significance in industrial application.

2 Background

AS shown in Fig.1. In the process, the factory need to strictly control the amount of raw materials, mixing time, temperature, humidity and wind, etc. In the shipment, the staffs access to the number on board electronic scale of the surveillance to control the transport of freight.

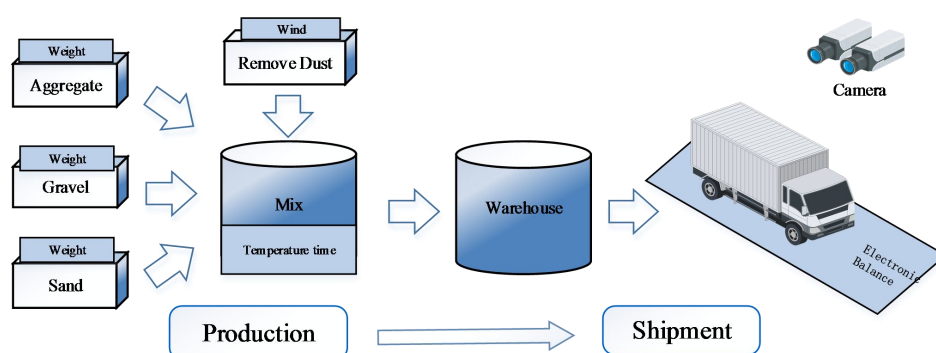


Figure 1. Process of Aggregate Production Line

Aggregate production line mainly monitors the production and shipment of two scenarios, and requires speed of a frame per second. Fig.2 shows the local schematic diagram of the aggregate plant in surveillance video. Fig.2(a) is a group of small digits from the production scenario, whose a single digital is 8×13 . In the video, the numbers located in the middle region are accurate to the units, the digital precisions are after the decimal point 3 in the left region and right region. The numbers on both sides of image, the decimal point is always appear, so it can be outputted directly without being recognized. Fig.2(b) is the local schematic diagram of surveillance video from the shipment scenario, which a single digital is 11×13 . In addition to the angle of camera and light, the low pixels of image block, the low resolution and the fuzzy image also bring difficulties to recognition. In this case, due to the low recognition rates, the traditional template matching methods [1-5] are not useful. In this paper, cascade recognition method [6] can improve the recognition rate of small digits. The primary recognition adapts to SVM classification based on HOG features, the secondary recognition adapts matching method based on different features.



Figure 2. Local Schematic Diagram of Aggregate Plant in Surveillance

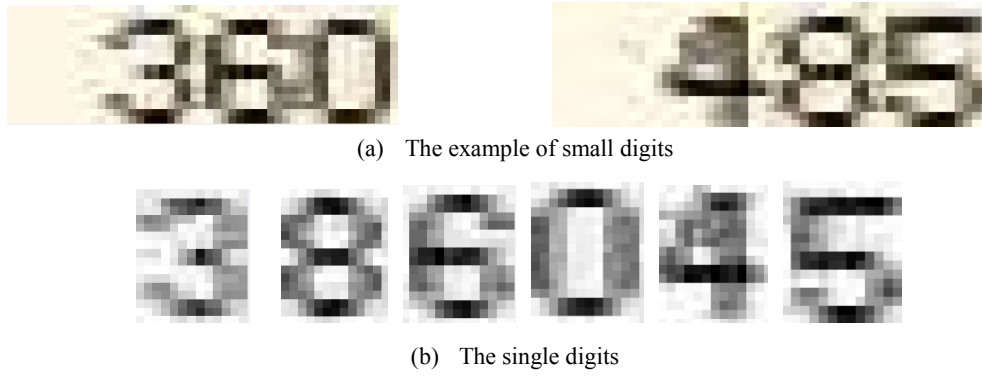


Figure 3. The Example of Digits from Production

As shown in Fig.3(a), there are two digitals in the video of production scenario. The one is "360", the other is "485". As shown in Fig.3(b), there are the single digital samples that the digits "360" and "485" consist of. Every digit is similar and fuzzy, thus, it is not easy to match only with the pixels. The histogram of oriented gradients (HOG) ^{[7][8]} features of "0", "1", "2", "4" and "7" have obvious difference, however, the similarity between HOG of "3" and HOG of "8" is 0.86, the similarity between HOG of "5" and HOG of "8" is 0.76, likewise "6" and "9" are similar to "8". Thus, it is not completely reliable to recognize the digits by HOG.

3 Small Digital Cascade Recognition

3.1 Pre-processing

Input video is processed, because the positions of the digitals are fixed in the frame. Firstly, the digital image block is obtained according to the coordinate position. Secondly, the image block is segmented by horizontal and vertical projection. This step gets some single digital image blocks. Thirdly, the size of a single digital image is normalized. In this paper, the pixel point in the digital lines is defined as a target point, and the point in the background is defined as background point. In order to make the size of training samples uniform, and the target points are not distributed on the edge of the images, normalized dimension of image is turned into 13×16 pixels. Finally, the image blocks are transformed into gray images.

3.2 Primary Recognition Based on Support Vector Machine

We use HOG as the feature descriptor of primary recognition. The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is then the concatenation of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing. Firstly, the original image is processed convolution-operator by using the gradient operator to get the horizontal gradient component. Then the original image is processed convolution-operator by using the gradient operator to get the vertical gradient component, shown as Formula (1), and the gradient of the point in the image is shown as Formula (2).

(1)

(2)

Where, are respectively the point in the horizontal direction gradient, the vertical direction gradient and the pixel value. The gradient magnitude and gradient direction of point are respectively shown as (3) and (4):

$$(3)$$

$$(4)$$

In this method, the size of scanning window is set to 48×64 . The size of block is called and set to 16×16 , its step is 8×8 . The size of cell is called and set to 8×8 . The histogram channels are evenly spread over 0 to 360 degrees. Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. For each cell, we calculate histograms of all pixel points to get a 9 dimensional feature vector. Every four adjacent cells compose of a block, all features are connected to a 36 dimensional feature vector. We move the block using the step to scan the window and connect all the features of blocks to a 1260 dimensional feature vector, the calculation process is shown as Formula (5) and Formula(6). There is the HOG feature of image, the whole process is shown as Fig.4. The image after the pretreatment is normalized into 48×64 , extracted its HOG feature.

$$(5)$$

$$(6)$$

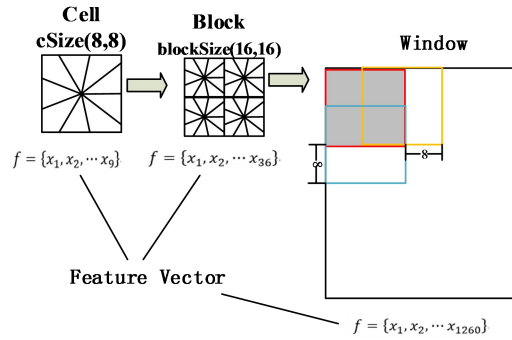


Figure 4. Process of Extracting HOG Feature

Where and are respectively feature vector dimensions of block and window, , and respectively represent width and height of window, width and height of the block.

Support vector machine (SVM) [6][7] is a supervised learning model with associated learning algorithms that used for pattern recognition, classification and regression. The nonlinear classifiers were proposed that the original finite-dimensional space be mapped into a much higher-dimensional space. The scenario of small numerical recognition belongs to not linearly separable case, by defining them in terms of a kernel function to solve the problem. We chose Radial Basis Function (RBF) is the kernel function of SVM. We use the SVM based on HOG as the Primary Recognition.

3.3 Cascade Recognition Based on Different Features

The experiment found that the SVM method recognizes "0", "1", "2", "4" and "7" more accurately, these five digits are defined as general digits, however for "3", "5", "6", "8" and "9" have more errors, these digits are easy to be confused, then called them special digits. In this paper, firstly, the recognition results are obtained by SVM, if the result is a general digital, system directly output the result, otherwise this digital image is recognized by the cascade recognition based on different features. As shown in Fig.5, one digital image is recognized as a special digital image using SVM, then it need

to further be recognized. In order to extract the different features of the image, the image is processed by binarization, the gray value of the image is only 0 and 255 two cases. Then system uses the Gauss denoising method to eliminate the black noise and white noise generated by binarization. Secondly, image expansion and image sharpening ensure that the image has enough target points, and the distribution of the target points is more appropriate. Finally, system extracts the different features of the image into cascade recognition based on different features.

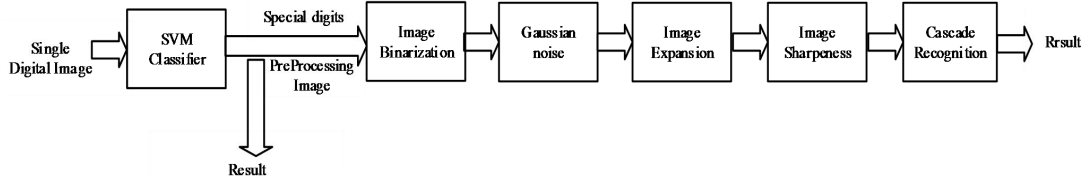


Figure 5. Process of Cascade Recognition

3.3.1 The Number of Special Regional Target Points

After image binarization, is the pixel value of the point, when value of point is 0, the point is a target point, if the value is 255 then the point is a background point.

According to the distribution of special digital pixel, as shown in Fig.6, in vertical direction the digit is evenly divided into three regions, in horizontal direction, is evenly divided into two regions. Thus, the digit is divided into six regions, upper-left region, upper-middle region, upper-right region, the lower-left region, lower-middle region and lower-right region ^{[9][10]}. Because there is little difference between lower-right region, upper-middle region, lower-middle region of the special digit., we only count target points of the upper-left region, upper-right region and lower-left region. The number of target points of these three regions is one part of feature, in addition to the total number of target points is another feature.

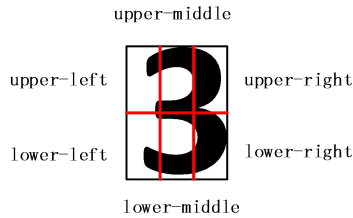


Figure 6. Special Region

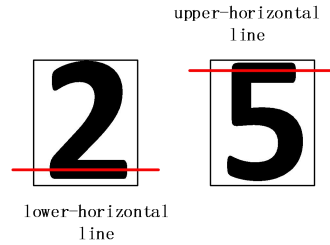


Figure 7. Horizontal Line

3.3.2 Horizontal Line Feature

By analyzing the digital structure, we can found that some digits have horizontal and vertical lines feature. According to the position of horizontal line, it is divided into upper-horizontal line and lower-horizontal line. System respectively begins scanning in sequence in the horizontal direction and the vertical direction. The adjacent several horizontal lines or vertical lines are regarded as one line, as shown in Fig.7, the digit "2" has a lower-horizontal line, "5" has a upper-horizontal line.

Only "5" in special digits contains the horizontal line, so as long as this determines whether the digit contains the feature, these digits are divided into two categories, "5" or others. The horizontal line feature is defined as in the horizontal direction it meet the Formula (8)

(8)

Where, is consecutive number of black pixels in the horizontal direction, and respectively are the width and height of a digital image, when that the there is horizontal line in the digit.

3.3.3 Hole Feature

With the digital structural characteristics, we add upper-hole feature and lower-hole feature to further distinguish the digits. According to the relative position of the center point and center line in the hole to define the upper-hole or the lower-hole. When the longitudinal coordinate of center point is bigger than the center line, the hole belongs to upper-hole, otherwise it is the lower-hole. As shown in Fig.8, "8" has a upper-hole and a lower -hole, "6" has a lower-hole, "9" has a upper-hole, "3" and "5" don't have holes.

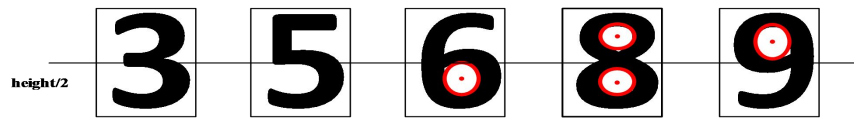


Figure 8. Hole Feature

3.3.4 Feature Vector

The different features of special digits consist of the number of special regional target points, horizontal line features and hole features. As shown in Table 1, the regional target pixel points contains of upper-left, upper-right, the lower-left and the whole region of image. When the digital does not contain horizontal line, the value of horizontal line feature is "0" otherwise, the value is "1". When the digital does not contain upper-hole (lower-hole), the value of upper-hole (lower-hole) is "0", otherwise, the value is "1". Different feature vectors of Special digits are shown in Table 2. In order to further improve the accuracy, our method matches features according to the weight^[12]. When special regional target points feature, horizontal line feature, upper-hole feature and lower-hole feature all occupy 25%, the result of experiment is the best. According to the weight matching, the method selects the digit with highest similarity as the result.

Table 1. Feature Vector

Feature Vector						
Total Target points	Upper-left Region	Upper-right Region	Lower-left Region	Horizontal Line	Upper-hole	Lower-hole

Table 2. The Feature Vector of Special Digits

Digit	Feature Vector
3	[68 17 20 7 0 0 0]
5	[77 19 22 12 1 0 0]
6	[77 15 17 19 0 0 1]
8	[76 16 18 18 0 1 1]
9	[61 14 15 8 0 1 0]

4 Experimental Results

In this experiment, the CPU is Intel (R) Core (TM) i7-4790, the Clock Speed is 3.60GHZ. The software environment is Visual Studio2013 and OpenCV2.49. The production and shipment of two scenarios

respectively are selected 300 frames to recognize by template matching method(TM), SVM and ours method. The single digital recognition rate and processing speed per frame are the evaluating indicators for the experiment. Table 3 and Table 4 show the number of recognition for the three methods in the two scenarios, the data in the tables represents times of every digit is successfully recognized. Fig.9(a) and Fig.9(b) to show the recognition rate of the three methods in two scenarios.

Table 3. The Result of Production Scenario

	0	1	2	3	4	5	6	7	8	9
Total digits	672	1548	828	1092	467	696	672	569	612	541
TM	623	1548	801	1013	455	634	515	564	604	519
SVM	672	1548	828	1041	467	659	531	568	607	512
Ours Method	672	1548	828	1092	467	696	661	569	612	541

Table 4. The Result of Shipment Scenario

	0	1	2	3	4	5	6	7	8	9
Total digits	936	839	589	654	815	1010	665	774	246	543
TM	912	793	546	174	815	485	54	774	193	212
SVM	936	839	589	401	815	957	187	774	246	498
Ours method	936	839	589	541	815	1010	605	774	246	543

As shown in Table 3,in the production scenario, the digit "3" appears 1092 times, and it is successfully recognized 1013 times by the template matching method, 1041 times by the SVM method and 1092 times by ours method. In the shipment scenario, the digit "9" appears 543 times, the template matching method only succeeds in recognizing it 212 times, in comparison, the recognition effect of SVM method has been increased, recognizes successfully 498 times. However, "9" is easy to be confused with the other 4 special digits, so the cascade recognition based on the different characteristics can correct the error to recognize all the digit "9".

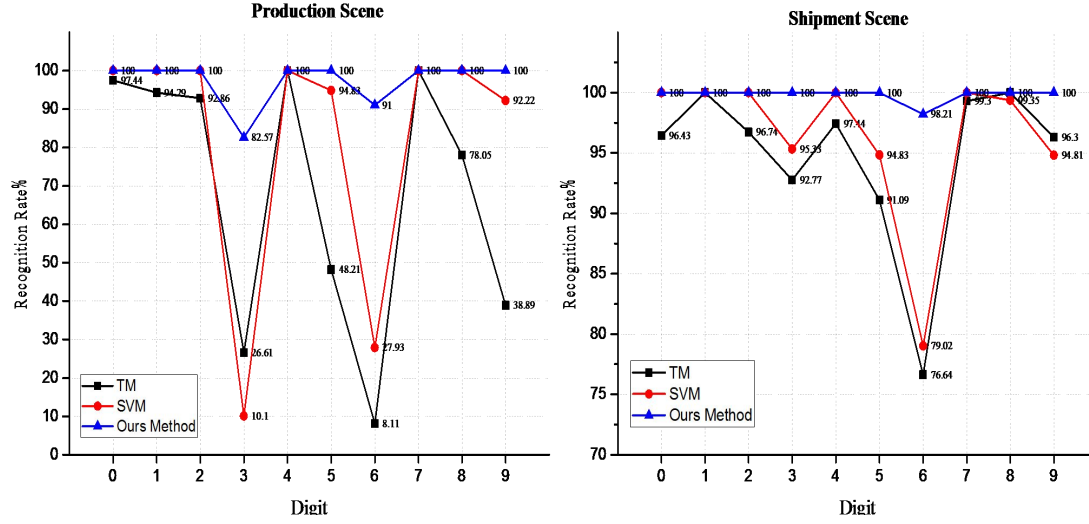


Figure 9. Recognition Rate of Production and Shipment

As shown in Fig.9, ours method has achieved a better recognition rate in production scenario. Specifically, the recognition rate of "4", template matching is 97.44%, while recognition rates of the SVM and ours method rise to 100%. The recognition rate of digital "6", template matching is 76.64%, SVM is 79.02%, and ours method rise to 98.21%. For the respective average recognition rate, the template matching method is 56.99%, the SVM method is 82.51% and ours method is 97.35%. In shipment scenario, the recognition rate by this method is significantly higher than template matching method. Specifically, the recognition rate of "2", template matching is 92.86%, SVM and this method rise to 100%. The recognition rate of "6", template matching is only 8.11%, SVM is 27.93%, and our method rises to 91%. For the respective average recognition rate, the template matching method is 94.67%, the SVM method is 96.33% and ours method is 99.82%.

In the small digital recognition of industrial video, the method based on SVM is obviously superior to the traditional template matching method. The SVM based on HOG features with cascade recognition of special digits corrects the recognition mistakes of confusable digits.

Table 5. Speed of Results

	Production Scenario		Shipment Scenario	
	Single Digit	Single Frame	Single Digit	Single Frame
TM	1.7	192.5	1.4	74.8
SVM	4.3	310.6	3.5	96.1
Ours method	6.1	361.4	4.2	134.3

On the one hand, the industrial aggregate system requires speed to keep 1 frame per second to identify all the numbers in the frame. As shown in Table 5, the unit of time is millimetre, in the speeds of a single digit and 1 frame, although ours method is a bit larger than the template matching method, it still meets the requirement of the scenario. On the other hand, the system requires in the accuracy rate of digital recognition not be less than 80%, then the template matching method and the SVM method are not satisfied. So we comprehensively consider the requirements of recognition rate and speed, then

ours method has achieved a good result in the small numerical recognition of cameras and surveillance from aggregate production industry.

5 Conclusion

In the aggregate production, we present a new robust method for small numerical recognition of the industrial video. Ours method succeeds in improving the recognition rate of small numbers. In addition, this method also achieves good results in vehicle-mounted system and the concrete industrial production. Compared with the template matching and the SVM classification, this method has a better effect on recognition of the small numbers for industrial scenarios, such as aggregate production, and has a valuable industrial application.

Reference

1. Kim K K, Kim K I, Kim J B, et al. Learning-based approach for license plate recognition[C]// Neural Networks for Signal Processing X, IEEE Signal Processing Society Workshop. IEEE, 2000:614-623 vol.2.
2. Bin M, Xiangbin M, Xiaofu M, et al. Digital Recognition Based on Image Device Meters[C]// Intelligent Systems (GCIS), 2010 Second WRI Global Congress on. IEEE, 2010:326-330.
3. Bai Q, Zhao L, Zhang Y, et al. Research of automatic recognition of digital meter reading based on intelligent image processing[C]// Computer Engineering and Technology (ICCET), 2010 2nd International Conference on. IEEE, 2010:V5-619-V5-623.
4. Wu Jin, Liu Yang, A Robust Method for Handwriting Digital Recognition[J]. WuHan: WuHan University of Science and Technology (Natural Science Edition), 2004,27(4): 406-408.
5. Cao Weijiong, Li Yujie. Digital Modulation Recognition Based on SVM [J]. Electronic Measurement Technology, 2009,2(32): 24-27.
6. Tu Z Y, Ning F H, Yu W J. Study about Recognition of Digital Meter Dial Reading Based on SVM[J]. Applied Mechanics & Materials, 2014, 615:194-197.
7. Wen Y, Lu Y, Yan J, et al. An Algorithm for License Plate Recognition Applied to Intelligent Transportation System[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(3):830-845.
8. Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection". Int. Conf. on Computer Vision & Pattern Recognition, 2005, volume 2, pages 886-893.
9. Ni Guibo, Liang Xiaozun. The Method of Printed Figures Based on Structure[J]. Software Guide. 2010,9(5): 67-68.
10. Chen Aibin, The Printed Number Character Recognition Based on Features[J]. Computing Technology and Automation. 2011,9(30): 105-108.
11. Liu Yangyang, Ji Qichun. Digit Recognition Based on Improved Template Matching Method[J]. Industrial Personal Computer, 2010(23): 76-77.