International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.1 (2016), pp.1-10
http://dx.doi.org/10.14257/ijsip.2016.9.1.01

A Road Vehicle Detection Algorithm Based on Compressive Sensing

Yiqin CAO[1], Xiaoci ZHOU[1] and Xiaosheng HUANG[2]

[1] *School of Software, East China Jiaotong University, Nanchang 330013, China*
[2] *School of Information Engineering, East China Jiaotong University, Nanchang 330013, China*
E-mail:745852546@qq.com

Abstract

With the aim to solve the problems of large amount of image data transmission and low accuracy of the initial background image extracted in traditional vehicle detecting system, this article proposes a road vehicle detecting algorithm based on compressive sensing. Image signals are sparse in a wavelet basis and the Gaussian random measurement matrix is adopted to compress videos, which reduce the amount of image data transmission. This article uses the proposed improved initial background extracting method and selective background updating method to obtain the initial background image and background updating which improves the accuracy of the initial background image. The vehicle detection and selective reconstruction of foreground image of vehicle are achieved by integrated background subtraction and the orthogonal matching pursuit algorithm. Through many experiments in video monitoring of real scenes, the article proves the correctness and efficiency of the algorithm. It not only improves the accuracy of the initial background image extracted but also reduces the amount of image data transmission and power consumption as well as the price of video transmission.

Keywords: *Compressive sensing; Vehicle detection; Orthogonal matching pursuit algorithm; Background subtraction*

1. Introduction

The vehicle detection is the key and core of the traffic monitoring system which is an important part of transportation as it is a crucial pillar of the national economy. There are several vehicle detection methods which are commonly used at present as follows: circular magnetic induction coil detection, ultrasonic testing (UT), microwave radar detection and video-based vehicle detection. Video detection technology is becoming the most dominant and potential detection method in the traffic monitoring system with the functions of video camera and computer to mimic the human eye. Video-based vehicle detection method needs to transport large amount of video images captured to video processing center first, the transmission bandwidth is a major challenge, thus it is hopeful that we can do reliable detection by merely transporting a small amount of data which contain enough information.

In recent years, we can achieve this goal with the emergence of the compressive sensing (CS) [1] theory. The surveillance cameras are transmitted a small amount of the measurements of video compression instead of pixel values by applying CS technology. Because the number of measurements are less than the total number of pixel values, the amount of data transmission and the requirement of bandwidth are reduced, thus reducing network costs which contribute to a broad application prospects. Video-based vehicle detection algorithm can be divided into the following categories [2-5]: Optical flow, frame difference, background subtraction, edge detection and motion vector detection. Because

ISSN: 2005-4254 IJSIP

of background subtraction is the easiest way to achieve and can fully segment the moving targets, it is suitable for terminal node processing of traffic cameras. After the measurements of surveillance video obtained, the pixel values of video frame image is unknown, thus the traditional background subtraction can't be used directly in vehicle detection. There is a method about reconstruction background difference image directly based on CS described in [6], theoretical analysis and experimental results have verified their feasibility, but the effect of the result is inferior in its reconstruction. In the field of the motion detection based on CS, it has published many related articles [7-11]. A ground moving target detection method based on CS proposed in [7], The article main research the problem of conventional Space-Time Adaptive Processing (STAP) hard to obtain sufficient independent and identically distributed samples in airborne radar for ground moving target detection. [8] studies target detection based on sparse representation in the infrared image, which is mainly for pedestrians detection at night. Due to the nature of radar signal, infrared image signal and video signal are different, so these two methods can't be used directly in video detection. [10] applies CS technology to detect targets in traffic video monitoring, it is mainly for the contradiction between an increasing demand of road traffic monitoring and the limit of network bandwidth. The application of CS can reduce the requirements of transmission bandwidth, but the problem is that the step size has a big impact for the resolution capability of different speed of moving target in its background model established.

Based on the above problems of large memory requirement, long transmission time and background modeling and update in the process of vehicle detection, this article proposes a road vehicles detection algorithm integrated CS theories and wavelet analysis. The algorithm uses CS to achieve sparse and reconstruction of vehicle image and detecting vehicle combined with improved background subtraction method. This article adopts the db3 wavelet to sparse image according to the characteristics of video image first, uses Gaussian random measurement matrix to observe and combines background model and the orthogonal matching pursuit(OMP) algorithm to detect and selective reconstruct foreground image of vehicle.

2. The Compressive Sensing Theory

The traditional Nyquist sampling theorem requires that the sampling rate of the signal must reach more than twice the signal bandwidth which can recover the original signal well, therefore the corresponding hardware devices are facing great pressure. This process of compression after high-speed sampling is also a waste of sampling resources.

As for Candes, Terres and Tao with the proposal of a new theoretical framework in CS, the theory advocates that if signal is compressible or sparse on a transform domain, an measurement matrix can be used which is irrelevant with sparse basis, the projection of high-dimensional signal after sparse transformation to a low-dimensional space, it can accurate or approximate to reconstruct the original signal from these small amount of low-dimensional projection by solve optimization problem.

2.1. Sparse Representation

Suppose that we have an image x of size $N_1 \times N_2$ and we vectorize it into a column vector x of size $N \times 1(N = N_1 N_2)$ by concatenating the individual columns of x in order. The n th element of the image vector x is referred to as $x(n), n \in [1,2,\cdots,N]$. Let us assume that the basis $\Psi = [\psi_1,\cdots,\psi_N]$ provides a K -sparse representation of x :

$$x = \sum_{n=1}^{N} \theta(n)\psi_n = \sum_{l=1}^{K} \theta(n_l)\psi_{n_l} = \Psi\theta \tag{1}$$

Meanwhile transform coefficients θ obtained

$$\theta = \Psi^T x \tag{2}$$

Where $\theta(n)$ is the coefficient of the n th vector basis $\psi_n(\psi_n : N \times 1)$ in the formula(1) and the coefficients indexed by n_l are the K -nonzero entries of the basis decomposition, among $K << N$; θ is a $N \times 1$ column vector which containing the K -nonzero elements. Different kinds of basis can sparse representation of natural images, these mainly include wavelet basis, cosine basis, Gabor frames and Curvelet *etc*. This article adopts db3 wavelet to sparse image, which characterizes is construct simple, achieve rapid and have a lower computation complexity, can convert the vast majority of the natural images and videos to a small amount of a non-zero value.

2.2. Measurement Matrix

In the CS framework, removing the middle process of collecting all the N pixels in the traditional data collection system, collect image vector in the form of compressed, namely adopt an measurement matrix $\Phi : M \times N, (M < N)$ which irrelevant with the orthogonal basis Ψ to perform compression and measurement to the signal x :

$$Y = \Phi\theta = \Phi\Psi^T x \tag{3}$$

It has proved in the CS that the larger no correlation among measurement matrix and sparse matrix, the better in reconstruction results. Gaussian random matrix just has these advantages, it is almost irrelevant with any sparse signal and can satisfied RIP (Properties of Restricted Isometric) in larger probability, therefore, this article gets the desired measurement matrix Φ by selecting the size of $M \times N$ from Gaussian random matrix.

2.3. Reconstruction Algorithm

Signal reconstruction is a problem to find the optimal solution of underdetermined equation in CS. Candes has pointed out in [12], If the original signal x is K-sparse, in addition, Y and Φ meet certain conditions, then the signal x can reconstruct from the measurements Y by solving the problem of optimal norm (the number of non-zero elements in the signal):

$$\hat{\theta} = \arg\min \left\|\Psi^T x\right\|_0 \ s.t.\ A^{CS} x = \Phi\Psi^T x = Y \tag{4}$$

Meanwhile, Donoho has proved in [5], the above problem is an NP-hard problem and needs to calculate all possible non-zero values in x , it is unable to solve, and therefore, the problem (4) generally translates into a simple approximate form in practical applications:

$$\hat{\theta} = \arg\min \left\|\Psi^T x\right\|_0 \ s.t.\ \left\|Y - \Phi\Psi^T x\right\|_2^2 \le \delta \tag{5}$$

Where δ is a minimal constant.

The commonly used reconstruction algorithm can be divided into three categories [13] at present: greedy pursuit algorithm, convex relaxation algorithm and combination algorithm. Because of the simple structure and a small calculation, greedy pursuit algorithm has caused for concern. OMP algorithm is its mainstream, so the article using OMP algorithm to reconstruct the foreground image of vehicle.

3. Road Vehicle Detection Algorithm based on CS

As the problems of large amount of image data transmission and low accuracy of the background image extracted in the traditional vehicle detection system, the article applies background subtraction based on CS to detect vehicle. After the application of CS technology, the system is transmitted a small amount of measurements instead of pixel values of images, so it can greatly reduce the amount of image data transmission. Due to the level of accuracy of the initial background image would affects update time of the background image, it is necessary to obtain a higher accuracy of the initial background image. Therefore, on the basis of the traditional average method, this article applies alternative and the idea of average seek times and proposes an improved method of extracting the initial background image and correspondingly background updating.

The algorithm framework works as follows: the first background image obtained by the improving initial background image extraction method first; then apply sparse basis and measurement matrix to obtain the measurements of the input video image and the initial background image, with a view to the ever-changing external environment and updating the measurements of the background image in real time; next, the measurements of the video frame image and the background image are subtracted to obtain the measurements of the differential image, in determining whether the moving target, the difference threshold should be determine according to the measurements of the background image and video frame image updated in real time; finally, use OMP algorithm to reconstruct foreground image of vehicle as needed, thus performing the next step.

3.1. Adaptive Background Modeling

In the traditional background modeling, the average method and Gaussian background modeling method are commonly used. The average method has a simple algorithm structure and low computational complexity with a wide range of applications. Although Gaussian background modeling method has a better modeling result than the average method, but the complex structure of high computational complexity, it has a larger limitations in its applications. So this article researches extraction and updates the initial background image based on the average method.

The idea of the average method is that frame images are averaging in direct way in the initial background image. However, there are a lot of blur of moving objects in the initial background image obtained by this method, the accuracy of image is low and with a big error, in addition, more number of video frame images are needed. In addition, factors such as changes in lighting, camera shake and the shadow will change the background, in order to improve the robustness and accuracy of the algorithm, for shortcomings in the average method, we propose an improved initial background modeling methods based on alternative and the idea of average seek times. Modeling idea is as follows: average values of several frame images obtained by the average method first, then calculating average differences of several frame images and use those average differences to remove the pixel value, which have a larger changes in those frame images, finally, averaging again and as pixel values of the original background image.

Specific modeling steps are as follows:

Calculating the average value $Mean(x, y)$ of N-frame images for obtaining the initial background image

$$Mean(x, y) = \left(\sum_{i=0}^{N} I_i(x, y) \right) / N \tag{6}$$

Calculate the mean difference $MD(x, y)$ of N-frame images

$$D(x,y)=\sum_{i=0}^{N}\left|I_i(x,y)-Mean(x,y)\right| \tag{7}$$

$$MD(x,y)=D(x,y)/N \tag{8}$$

Replace the pixel value which have a lager changes in the current frame image with the pixel value of the next frame image

If

$$\left|I_i(x,y)-Mean(x,y)\right|>\mu MD(x,y) \tag{9}$$

Then

$$I_i(x,y)=I_{i+1}(x,y) \tag{10}$$

Where $i=0,1,2,\cdots,N$, set $\mu=2.5$ (2.5 is the empirical value).

Calculating average values of N-frame image after replacement and as pixel values of the initial background image

$$B_0(x,y)=\left(\sum_{i=0}^{N}I_i(x,y)\right)/N \tag{11}$$

The method in extraction of the initial background image, only needs about 10 frames video images and a high accuracy of initial background image can be obtained, it improves the shortcomings of the average method in above described.

3.2. Background Update Policy

The traditional background update policy is designed for the pixel values of image, as the measurements of image are obtained after using CS, therefore, to change design object to the measurement in design update policy. As known from [6], the difference image could be reconstructed by using the measuring difference of the current frame image and the background image, therefore the background image also can updated by those measuring differences. Suppose that X_n, B_n respectively the current input image and background image, y_{tn} is the measurement of X_n, y_{bn} is the measurement of B_n, the measurement y_{bn+1} of the $n+1$ moment background image B_{n+1} is calculated by the following formula:

$$y_{bn+1}(i)=\begin{cases}\alpha y_{bn}(i)+(1-\alpha)(y_{tn+1}(i)-y_{tn}(i)), & there\ is\ moving\ in\ i\\ y_{bn(i)}, & there\ is\ non-moving\ in\ i\end{cases} \qquad i=1,2,\ldots\ldots,M$$

(12)

Where α means learning speed of the model, it is a constant and meet $0\le\alpha\le1$, its reciprocal means a time constant in decay process and is a experience in generally; M means the number of measurements; i means the position of the corresponding measurement; y_{b0} is the measurements of the initial background image.

At position i, if the condition meets the following formula, it indicates that there is a moving target in this position.

$$\left|y_{tn+1}(i) - y_{tn}(i)\right| > T_{n+1}(i) \tag{13}$$

Where $T_{n+1}(i)$ is a threshold updated in real time, the updating policy is expressed by the following formula.

$$T_{n+1}(i) = \begin{cases} aT_n(i) + (1-a)\left(\left|y_{tn}(i) - y_{bn}(i)\right|\right), & there\ is\ moving\ in\ i \\ T_n(i), & there\ is\ non-moving\ in\ i \end{cases} \qquad i = 1,2,\ldots .M. \tag{14}$$

Where a is a constant which is close to 1, the size of the threshold $T_{n+1}(i)$ can adjusted by change the value of a.

Videos are transferred into the system of vehicle detection in the form of a frame by framing after collection, the system will automatically update the background according to the movement of target and make the background model adapt to environmental changes.

4. Experiment and Analysis

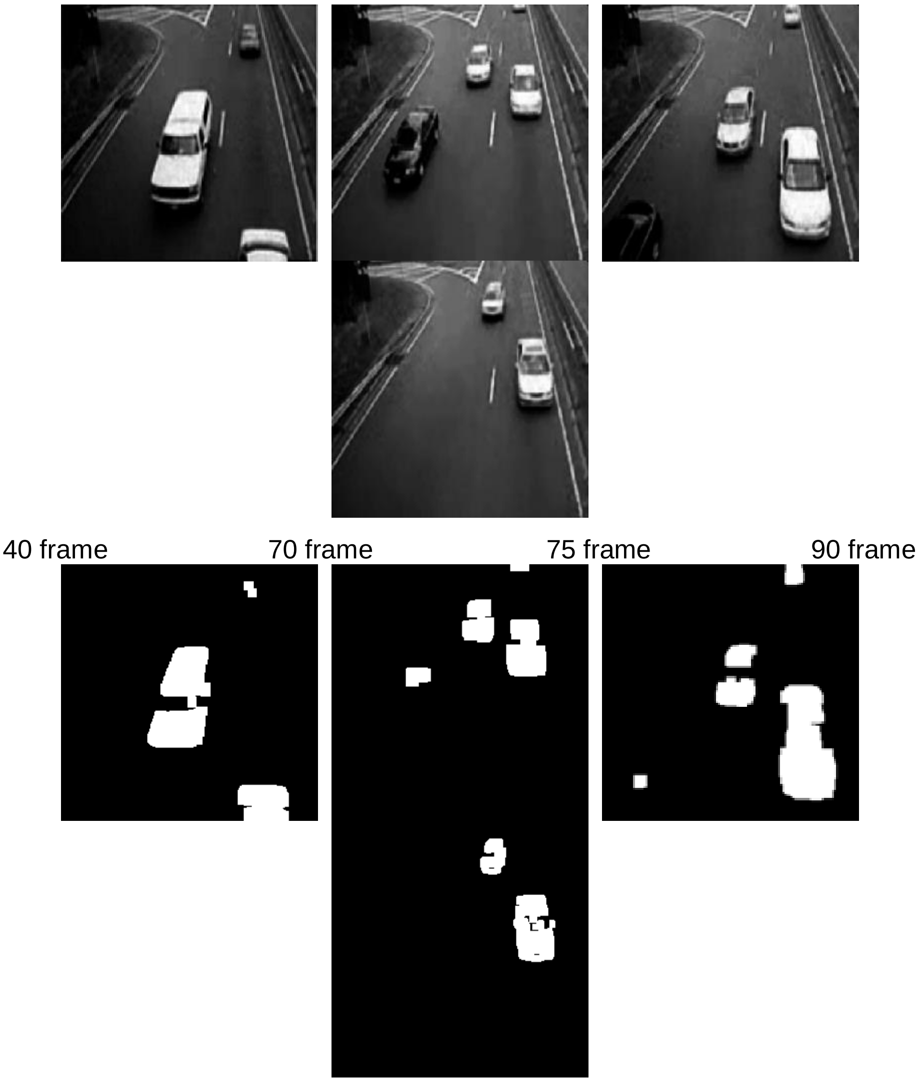
In order to verify the effectiveness of the algorithm, the resolution of a traffic surveillance video of real scene is 256×256 were tested, which includes multiple vehicles and interferes with complex background such as shade, leaves and road, as well as the scenes of various vehicles newborn and disappeared.

Figure 1 is the initial background image which is obtained by three different methods, Figure 1(a) is the result of the average method, as it can be seen from the image, a large number of blur of vehicles in it, the image accuracy is low and has a large errors, and needs more number of images to extract the initial background image; Figure 1(b) is the result of [14], from the result, this method has a greater improvement than the average method, but the accuracy of the initial background image isn't very high; Figure 1(c) is our method, the result shows that our method has some improvements than [14], and we use just six images in obtaining the initial background image, so less images required in obtaining the initial background image and the updating time of the background is greatly reduced.

The average method   document[14]   our method

Figure 1. The Initial Background Image Obtained by Different Methods

Figure 2 is the result of vehicle detected in different frame image, the first two lines are the result of the original video frame images, the next two lines are the result of vehicles detected by integrated CS theories and background subtraction after morphological processing. In Figure 2(a), there are three cars on the road driving down from the top, the algorithm can be detected. There are four cars in Figure 2(b), although a car just enters this frame image, the algorithm can still be accurately detected. Figure 2(c) has four cars and Figure 2(d) has two cars, the algorithm also can be clearly detected. You can also seen from Figure 2, the contour and size of the bright colored vehicle detected are similar to the original image, the error is small, but the contour and size of the dark colored vehicle detected are vary considerably from actual, this is because the color of vehicle close to the color of road. Although the result of the dark colored vehicle detected worse than the result of the bright colored vehicle detected, but the total number of vehicles traveling on the road can be accurately detected by this algorithm.

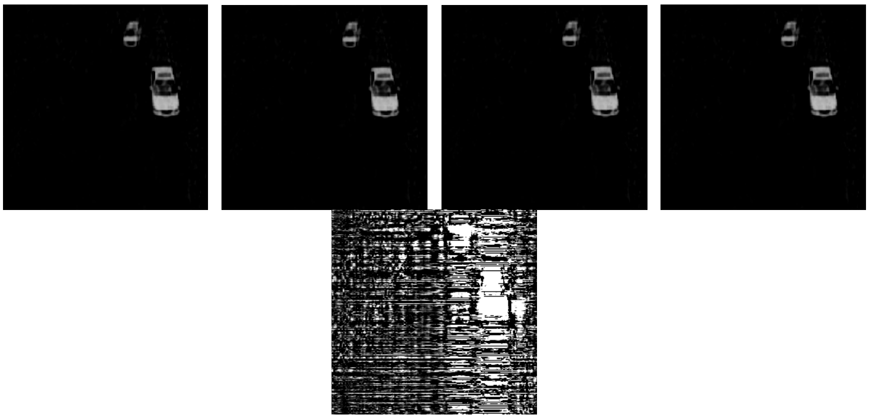


(a) (b) (c) (d)

Figure 2. The Result of Vehicles Detected in Different Frame Images

Figure 3 is the result of the 90th frame image by four kinds of reconstruction algorithm. It can be intuitively seen from Figure 3, the reconstruction result of iterative hard threshold(IHT) is the worst, the results of the other three are almost and with the result of conventional background subtraction method has no difference in the human visual. However, due to different structures of these four algorithms, the reconstruction time is in difference, the running time of these four algorithms is shown in Table 1. Peak Signal to Noise Ratio (PSNR) is an objective standard for image evaluation, it is the mean square error between the original image and the processed image respect to the value of (2 ^ n-1)

^ 2 (the square of the maximum signal, n is the number of bits of per sample). In order to measure the quality of the processed image, it usually takes the PSNR as a reference to measure whether a program is satisfactory, the PSNR values of images by these four algorithms reconstructed are shown in Table 1.



Conventional back-ground subtraction　　OMP　　CoSaMP　　SP　　IHT

Figure 3. The Reconstruction Results of Four Reconstruction Algorithms

Table 1 shows that the reconstruction times and the PSNR values of four reconstruction algorithm which can be seen from Table 1, the reconstruction time of IHT is minimum, but its reconstruction result is the worst, the PSNR values of the other three reconstruction algorithms are similar, it describes that their reconstruction results are quite. In these three reconstruction algorithms, the time of the OMP algorithm required is the least, in summary, the global performance of the OMP algorithm is optimal of the four.

Table 1. The Running Time and the PSNR Values of Four Kinds of Reconstruction Algorithm

algorithm	时间（s）	PSNR
OMP	19.001542	10.2717
CoSaMP	122.362654	10.2714
SP	61.058543	10.2716
IHT	2.157865	-10.3688

5. Conclusions

The article detected vehicle based on CS theory, integrated background subtraction, the OMP algorithm and local time equalizer adaptive background model. Based on the db3 wavelet basis and the Gaussian random measurement matrix to obtain the measurements, the method is mainly to solve the problem of the traditional video coding which requires a large amount of calculation. It also can reduce the tasks of hardware processing in the video collect terminal and the amount of data transmitted over the network as well as the more computing transfer to the back-end information processing center. By detecting video of multiple vehicle to obtain its accuracy and analysis error appears, experimental results show that the proposes method can extracted the foreground region under the complicated traffic environment, and can more quickly complete the reconstruction of foreground image of vehicle with the realization of a project.

Acknowledgements

This research was financially supported by the National Natural Science Foundation under Grant No.61365008, Science and Technology Support Program of Jiangxi under Grant No.20123BBE50093.

References

[1] D. L. Donoho, "Compressed sensing[J]", IEEE Transactions on Information Theory, vol. 52, no. 4, **(2006)**.
[2] R. Jain, K. P. Sankar and C. V. Jawahar, "Interpolation based tracking for fast object detection in videos[C]", Computer Vision, Pattern Recognition, Image Processing and Graphics(NCVPRIPG), 2011 Third National Conference on. IEEE, **(2011)**.
[3] Z. Minghan and L. Dayong, "Moving objects detection and tracking based on two consecutive frames substruction background model[J]", Computer Measurement & Control, vol. 14, no. 8, **(2006)**.
[4] P. Wang, C. Shen, N. Barnes and H. Zheng, "Fast and robust object detection using asymmetric totally corrective boosting[J]", IEEE Digital Object Identifier, vol. 23, no. 1, **(2012)**.
[5] Ng H F. Automatic thresholding for defect detection[J], Pattern Recognition Letters, vol. 27, no. 14, **(2006)**.
[6] V. Cevher, A. Sankaranarayanan, MF Duarte, D. Reddy, RG Baraniuk and R. Chellappa, "Compressive sensing for background subtraction[M]", //Computer Vision-ECCV 2008, Spring Berlin Heidelberg **(2008)**.
[7] W. Weiwei, L. Guisheng, Z. Shengqi and X. Jingwei, "A ground moving target indication method based on compressive sensing[J]", Journal of Electronics & Information Technology, vol. 34, no. 8, **(2012)**.
[8] S. Gao, I. Kim and S. T. Jhang, "Sparse Representation based Target Detection in Infrared Image[J]", International Journal of Energy, Information and Communications, vol. 4, no. 6, **(2013)**.
[9] R. Yao and Y. Zhang, "Compressive sensing for small moving space object detection in astronomical images[J]", Journal of Systems Engineering and Electronics, vol. 23, no. 3, **(2012)**.
[10] L. Fenlan, P. Zhuotao and Z. Zhemin, "Objects detection algorithm of traffic video monitoring based on compressed sensing[J]", Journal of Shantou University(Natural science), vol. 28, no. 2, **(2013)**.
[11] H. Jiang, W. Deng and Z. Shen, Surveillance video processing using compressive sensing[J], Inverse Problems and Imaging, vol. 6, no. 2, **(2012)**.
[12] E. CANDES, "Compressive sampling[C]", //Proc of International Congress of Mathematicians, Ziirich, Switzerland: European Mathematical Society Publishing House, **(2006)**.
[13] S. Guangming, L. Danhua, G. Dahua, L. Zhe, L. Jie and W. Liangjun, "Advances in theory and application of compressed sensing[J]", Acta Electronica Sinica, vol. 37, no. 5, **(2009)**.
[14] Z. Yang, "A new Algorithm of Background Image Extraction and Update in the Vehicle Detection System", Multimedia Technology (ICMT), **(2011)** November 5238-5241; International Conference on Hangzhou, China.

Authors

Yiqin CAO, He is a Professor, master, His main research areas: image processing, pattern recognition.

Xiaoci ZHOU, She is master, Her main research areas: image processing, pattern recognition.

Xiaosheng HUANG, He is an Associate master, Doctor, main research areas: image processing.