

Vehicle Detection Research Based on USILTP Operator

Fei Li*, Xiaohong Zhang, Chenqiu Zhao & Meng Yan
College of Software Engineering, Chongqing University, Chongqing, China

ABSTRACT: This paper presents a uniform SILTP operator based on the SILTP operator. In the vehicle detection, SILTP can solve the problems caused by the change of sunshine, the shadow of vehicle and the noise in the surrounding environment. But the algorithm has high dimensionality which can lead to error because of the deviation of the texture characteristics. The USILTP operator can reduce the dimensionality of the detection data which adapts to the problems caused by illumination variations and the noise in the surrounding environment. First, the method uses the SILTP operator to extract the vehicle image texture characteristics and reduce the dimensionality of the detection data, and then it uses the Gauss mixture model to do background modeling, and uses the texture characteristics of the new image to update background dynamically. At last, it gets the vehicle by contracting with the background model. It has been proved that this detection algorithm has a good performance with the test of the vehicle on public roads.

Keywords: local texture characteristics; vehicle detection; uniform SILTP operator

1 INTRODUCTION

The vehicle detection technology plays an important part in the research topic which applies computer vision in the intelligent traffic monitoring system. It is a guarantee to obtain the information of the traffic and a key to making the traffic management more intelligent. In the intelligent traffic monitoring system, the vehicle flow and speed are obtained in the traffic video information collected by many cameras. Only if separating the vehicle from the scene that camera sees, can we obtain more information about the vehicle. How to detect the vehicle accurately in the complex dynamic backgrounds such as illumination variations becomes the key to the detection algorithm.

The famous traditional methods for detecting vehicle are background subtraction^[1] method, frame difference method^[2] and optical flow method. The frame difference method extracts the motion region of the image based on the time difference scheme among two or three neighboring image sequences. The background subtraction method approximates the pixel value of the background image with the parameter model of the background and detects the motion region by comparing the current frame with the background image. The pixel regions which have large differences are considered to be the motion region and the pixel regions which have little differences are considered to be the background region. The optical flow method estimates the playground with the spatial and temporal gradient of the image frame, detects and segments the motion region from the background with the playground. However, these traditional methods have several problems. For example, the illumination variations and the angle of detections can affect the appearance of the vehicle.

The LBP^[3-6] operator can extract the texture of the image and turn it into the histogram. With the histogram, we can compare different images. The LBP operator does not work very robustly on flat image areas such as sky, where the gray values of the neighboring pixels are very close to the value of the center pixel.

This paper presents a uniform SILTP operator based on the Scale Invariant Local Ternary Pattern^[7] operator. The USILTP operator is more invariant than the SILTP operator if the illumination is suddenly changed from darker to brighter or vice versa. Through the experiment on the vehicle data set of the CDNET, the USILTP operator has a good performance.

2 FEATURE EXTRACTION BASED ON SILTP

2.1 SILTP operator

Either global or local, the illumination variations often cause sudden changes of gray scale intensities of neighboring pixels simultaneously, which would be approximately a scale transform with a constant factor. The SILTP operator can handle the illumination variations with its characteristics. The SILTP operator encodes it as shown in Equation 1:

$$\text{SILTP}_{N,R}^T(X_c, Y_c) = \bigoplus_{k=0}^{N-1} S_T(I_c, I_k) \quad (1)$$

(X_c, Y_c) is the coordinate of the pixel C in the image. I_c is the gray intensity value of the center pixel, I_k are that of its N neighborhood pixels which are equally spaced on a circle of radius R; \bigoplus denotes the concatenation operator of binary strings; T is a scale factor indicating the comparing range; S_T is a piece-

*Corresponding author: 449795275@qq.com

wise function defined in Equation 2.

$$S_T(I_c, I_k) = \begin{cases} 01, & \text{if } I_k > (1+T)I_c \\ 10, & \text{if } I_k < (1-T)I_c \\ 00, & \text{otherwise} \end{cases} \quad (2)$$

The SILTP encodes it with two bits as each comparison can result in one of three values. The scale invariance of SILTP operator can be easily verified. The SILTP operator is illustrated in Figure 1.

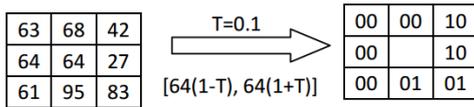


Figure 1. SILTP operator

The pixel is calculated by $S_T(I_c, I_k)$ and the result is a double-digit binary number. In order to bring into correspondence with eight-bit binary numbers, we get the binary numbers in the directions of up, down, left and right. The advantage of SILTP operator lies in three folds. First, the SILTP operator can make one more comparison than the LBP for each neighbor which is computationally efficient. Second, the SILTP operator is robust to local image noises within a range. When the region is darker and contains more noises, the LBP operator would be affected much while the SILTP operator is tolerable. At last, the scale invariance property makes the SILTP operator robust to illumination variations.

2.2 USILTP operator

The SILTP operator can produce different binary patterns. With the circle region whose radius is R contains P sampling sites, the SILTP operator can produce 2^p patterns. Obviously, with the increase of the sampling sites, the quantity of the binary patterns is increased rapidly. For example, there are $2^{20} = 1,048,576$ binary patterns with the 20 sampling sites in 5×5 regions. It is disadvantageous for the extraction or recognition of the texture and the access of information with so many binary patterns. Meanwhile, it is also harmful to indicate the texture. For example, we often use a histogram to indicate the information of the images, and too many patterns will lead to the large amount of data and the sparse histogram. We should reduce dimensions of the original SILTP pattern so that the data can be used to present the image more accurately.

The USILTP operator based on the SILTP operator makes the model of SILTP operator become a ring. The binary numbers corresponding with the USILTP operator start from 0 to 1 or 1 to 0, and the numbers can be changed at most twice. Then the algorithm pattern is called a Uniform pattern. The method which checks whether a pattern is a uniform pattern or not is

that whether the sum of the pattern and the one comes from the pattern moving a single bit is less than or equal to 2. If the result is less than or equal to 2, the pattern is a uniform pattern. For example, 00000000, 01110000 and 11001111 are all uniform patterns. 00000000 doesn't change, 01110000 changes twice and 11001111 changes twice, too. 11001001 which changes four times and 01010011 which changes six times are not uniform patterns. The uniform pattern can reduce the quantity of the result from 2^p to $p(p-1)+2$. P which stands for the number of the sampling sites we obtain from the images. With the improvement of the SILTP, the variety of the binary patterns is reduced a lot and doesn't lose any information. The dimension of the feature vector is reduced that reduces the effect of the noise.

The USILTP operator compresses the texture information and reduces the dimension of the feature vector. It can adapt to the noises in the environment.

3 VEHICLE DETECTION

3.1 Image preprocessing

In the process of image acquisition, the noise comes into being because of the sensor. So we must preprocess the images before send them to the operator. With the image preprocessing, we can obtain the texture features of the image accurately. We use two-dimensional discrete wavelet transform to preprocess the image. The two-dimensional discrete wavelet is used for image filtering. It decomposes images and makes them take up less space for storage. And it doesn't affect the extraction of the texture features in the image.

$X[m, n]$ is the original image, m stands for the image position of pixels on the vertical direction, and n stands for the image position of pixels on the horizontally direction. First, we do high-pass, low-pass and frequency reduction processing in the n direction. The high frequency part of the original image is filtered by the low-pass filter $g[n]$ with the low frequency of the image left. And the low frequency part of the original image is filtered by the high-pass filter $h[n]$ with the high frequency of the image left. Because the output sequence has the same length with the input sequence in the two filters, the result is twice the length of the original image. Then, we can reduce half of the original sampling rate without losing any information. The steps of the two-dimensional discrete wavelet transform are shown in the Figure 2.

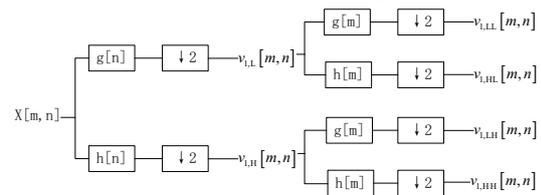


Figure 2. Two-dimensional discrete wavelet transform steps

We obtain the low frequency $v_{1,L}[m,n]$ of the image with low-pass filter $g[n]$, and acquire the high frequency $v_{1,H}[m,n]$ of the image with low-pass filter $h[n]$. The process is shown in Equation 3 and Equation 4.

$$v_{1,L}[m,n] = \sum_{k=0}^{K-1} x[m, 2n-k]g[k] \quad (3)$$

$$v_{1,H}[m,n] = \sum_{k=0}^{K-1} x[m, 2n-k]g[k] \quad (4)$$

Process $v_{1,L}[m,n]$ and $v_{1,H}[m,n]$ are along m direction for high-pass, low-pass and frequency reduction. The process is shown in Equation 5, Equation 6, Equation 7 and Equation 8:

$$v_{1,LL}[m,n] = \sum_{k=0}^{K-1} v_{1,L}[2m-k,n]g[k] \quad (5)$$

$$v_{1,HL}[m,n] = \sum_{k=0}^{K-1} v_{1,L}[2m-k,n]h[k] \quad (6)$$

$$v_{1,LH}[m,n] = \sum_{k=0}^{K-1} v_{1,H}[2m-k,n]g[k] \quad (7)$$

$$v_{1,HH}[m,n] = \sum_{k=0}^{K-1} v_{1,H}[2m-k,n]h[k] \quad (8)$$

$v_{1,LL}[m,n]$, $v_{1,HL}[m,n]$, $v_{1,LH}[m,n]$ and $v_{1,HH}[m,n]$ respectively stand for the low-pass component, the horizontal details component, the vertical details component and the details component in the direction of 45°. Then we obtain the image $v_{1,LL}[m,n]$ after two times' low-pass filter. And we use the USILTP operator to obtain the texture features of image $v_{1,LL}[m,n]$.

3.2 Background modeling

The background modeling^[8-11] is the most important part of any background subtraction algorithm. The aim of background modeling is to create a statistic model to represent the background. In the complicate environment such as illumination changes, swaying vegetation, rippling water and flickering monitors, it is difficult to model the background. In this paper, the camera is assumed to be nonmoving. We model each pixel of the background identically. The background model updates procedure for one pixel and the procedure is identical for each pixel.

After we obtain the feature information of the vehicle, we can model the background. We consider the feature vectors of a particular pixel over time as a pixel process. The USILTP histogram computed over a circular region of radius R around the pixel is used as the feature vector. The user himself sets the value of R. The background is made of a group of adaptive USILTP histograms which is $\{\vec{n}_0, \dots, \vec{n}_{k-1}\}$.

K is also selected by the user. In order to let the k model histograms be summed up to one, each model histogram has a weight between 0 and 1. The weight of the kth model histogram is denoted by w_k . We donate the USILTP histogram of the given pixel computed from

the new video frame by \vec{h} . In the first step of processing, \vec{h} is compared to the current k model histograms using a proximity measure. The formula of the histogram intersection is Equation 9 which is shown as follows:

$$\bigcap(\vec{a}, \vec{b}) = \sum_{n=0}^{N-1} \min(a_n, b_n) \quad (9)$$

\vec{a} and \vec{b} are the histograms, and N is the number of the histogram bins. The meaning of the formula is that it calculates the common part of two histograms. The advantage of the calculation is that it explicitly neglects features which only occur in one of the histograms. The complexity of the calculation is very less because it requires very simple operations. The threshold for the proximity measure TP is defined by user and it is between 0 and 1. We compare the histogram \vec{h} of the new image coming from the camera with the histogram $\{n_0, \dots, n_{k-1}\}$ of the background model. If the proximity is below the threshold TP for all model histograms, the model histogram with the lowest weight is replaced with m. And we give the new histogram a low initial weight. In our experiment, we give TP the value of 0.01.

If matches are found, we will make more processing in the experiment. The best match is selected to become the model histogram with the highest proximity value. The best matching model histogram is adapted with the new data by updating its bins as Equation 10 which is shown as follows:

$$\vec{n}_k = a_b \vec{h} + (1 - a_b) \vec{n}_k \quad a_b \in [0, 1] \quad (10)$$

The a_b is a user-settable learning rate. And we update the weights of the model histogram as the Equation 11 which is shown as follows:

$$w_k = a_w M_k + (1 - a_w) w_k \quad a_w \in [0, 1] \quad (11)$$

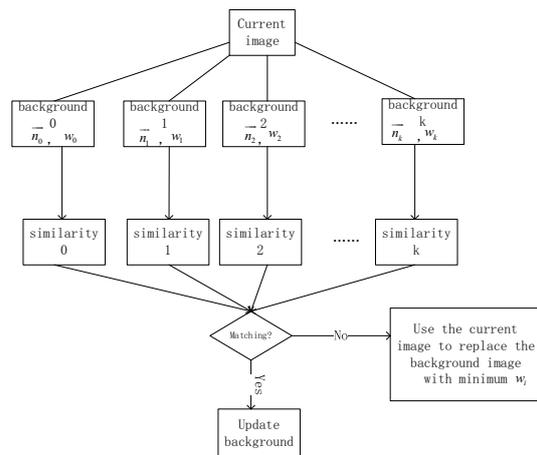


Figure 3. Process of background update

The process of background update is shown in Figure 3.

a_w is also a user-settable learning rate. M_k is 1 for the best matching histogram and 0 for the others. The learning rate parameters a_b and a_w control the adaptation speed of the background model. If the learning rate is bigger, the adaptation is faster. All the model histograms can't become the background model. We can confirm whether a model histogram is a background model or not with its sustainability. And the weight of the model histogram decides the sustainability of the model histogram. When the weight is bigger, the model histogram has higher probability as a background histogram. We sort the model histograms in decreasing order based on their weight, and choose B histograms as the background histograms. It is shown in Equation 12:

$$w_0 + \dots + w_{B-1} > T_B \quad T_B \in [0, 1] \quad (12)$$

T_B is a threshold defined by user.

3.3 Foreground detection

The foreground detection begins before the update of the background model. We compare the histogram h with the current histograms B . If proximity for at least one background histogram is higher than the threshold TP , the pixel is marked as foreground.

4 EXPERIMENTAL RESULT AND ANALYSIS

We perform our experiment on the Windows 8 system, its memory is 4G and its CPU is 64-bit. The experiment parameters are shown in Table 1. We use Matlab2011b to code and test, and use the car images on the CDNET (<http://wordpress-jodo.in.dmi.usherb.ca/>) to test the USILTP operator. The experiment result is shown in Figure 2-4.



Figure 4. Original image

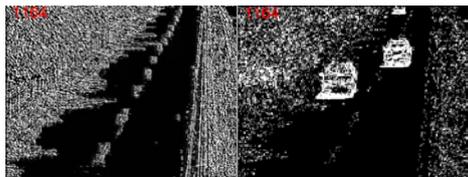


Figure 5. Background image and foreground image in LBP operator

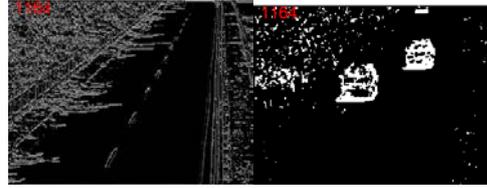


Figure 6. Background image and foreground image in SILTP operator

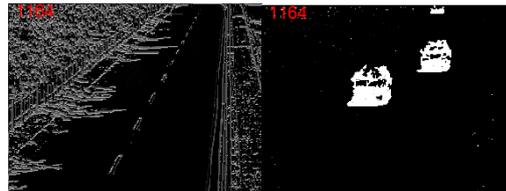


Figure 7. Background image and foreground image in USILTP operator

Table 1. Experiment parameter

Sequence	R	k	a_b	a_w	T_p	T_B
1100-1197	3	3	0.01	0.01	0.6	2

Figure 4 shows the original image. Figure 4 shows the background image and foreground image extracted from the original image by the LBP operator and the Gaussian mixture model. From Figure 5, we can make a conclusion that the LBP operator cannot extract the texture features from the original image accurately because of the influence of all kinds of noises in the environment. There are still lots of pixels which belong to the background images exist in the foreground image. Figure 6 shows the background image and the foreground image extracted from the original image by SILTP operator and the Gaussian mixture model. The noise in foreground image is less than the noise in the foreground image of Figure 4. But it still cannot satisfy us. Figure 7 shows the background image and the foreground image extracted from the original image by the USILTP operator and the Gaussian mixture model. The image texture features extracted by the USILTP operator are better than the LBP operator and the SILTP operator. The results of both background image and foreground image are perfect. The result proves that the USILTP operator has a strong adaptability, and it can accurately identify the points which move periodically in the background image. The USILTP operator can separate the foreground image from the background image thoroughly.

In order to compare the accuracy of each algorithm, we use FP (False Positive) to be the rule to show the result. We choose the images which begin from the 1100th to the 1197th, and calculate the value of FP.

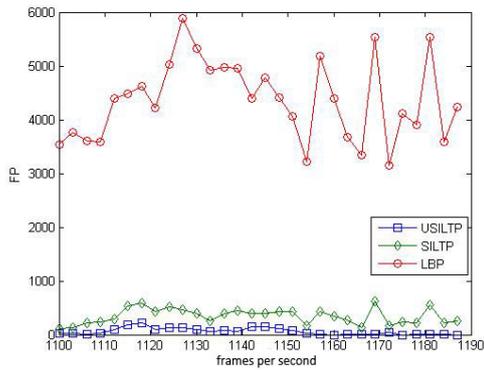


Figure 8. Experimental comparison

From the Figure 8, it can be inferred that the value of FP in the USILTP operator is lower than other algorithm. The LBP algorithm has the worst result because there are noises in the environment and the LBP cannot handle it as well as other algorithms. The value of FP in the Gaussian Mixture algorithm is larger than the SILTP algorithm as the shadow which exists in the environment is considered as the foreground. The SILTP algorithm has a good performance in reducing the noises, but it cannot catch up with the USILTP algorithm.

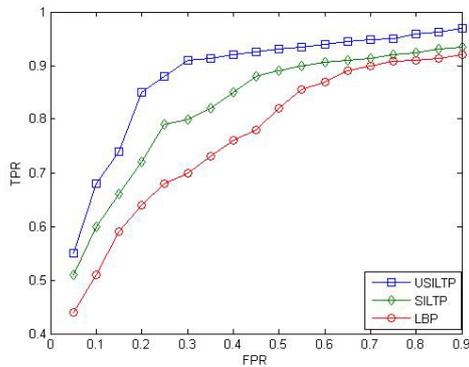


Figure 9. Detection efficiency of LBP, SILTP, and USILTP

Figure 9 shows the superiority of the USILTP operator. Compared with the LBP operator and the SILTP operator, the USILTP operator has a better performance in the test. Figure 7 makes a quantitative evaluation with the ROC curve. We use FPR (False Positive Rate) as the x-axis and the FPR stands for the proportion that the number of background pixels is detected as the foreground pixels of the total number of pixels. We use TPR (True Positive Rate) as the y-axis and the TPR stands for the proportion that the vehicle pixel points detected correctly in the total number of vehicle pixel points. And we use TP (True

Position) to express the threshold of ROC curve made by the three operators. The result of the test shows that the USILTP operator is better than the other two operators.

5 CONCLUSIONS

In this paper, the authors have demonstrated a USILTP algorithm which can be used in vehicle detection. It is robust against the sudden changes of gray scale intensities of neighboring pixels such as illumination variations. And it has a better performance than the SILTP operator in the vehicle detection.

ACKNOWLEDGEMENT

This paper is sponsored by the shape of the semantic structure and classification study (No.61173131).

REFERENCES

- [1] Liu H, Hou X. 2012. Moving detection research of background frame difference based on Gaussian model. *Computer Science and Service System, International Conference on. IEEE*, pp: 258-261.
- [2] Yongtian L, Zhijian J, Junwei G, et al. 2014. Robot fish detection based on a combination method of three-frame-difference and background subtraction. *Control and Decision Conference, Chinese. IEEE*, pp: 3905-3909.
- [3] Hegenbart S, Uhl A. 2014. A scale-adaptive extension to methods based on LBP using scale-normalized Laplacian of Gaussian extrema in scale-space. *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE*, pp: 4319 - 4323.
- [4] Huang Y, Ou Z, Hsieh H, et al. 2011. A rapid texture-based moving object detection method. *Control Conference (ASCC), 2011 8th Asian. IEEE*, pp: 1205-1209.
- [5] Qin Y, Tang Y. 2011. Dynamic texture recognition based on multiple statistical features with LBP/WLD. *Computer Science and Network Technology (ICCSNT), 2011 International Conference on. IEEE*, pp: 957-960.
- [6] Nguyen D T, Ogunbona P, Li W. 2011. Human detection with contour-based local motion binary patterns. *Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE*, pp: 3609-3612.
- [7] Liao S, Zhao G, Kellokumpu V. 2010. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. *Computer Vision and Pattern Recognition, IEEE Conference on. IEEE*, pp: 1301-1306.
- [8] Yin H, Yang H, Su H, et al. 2013. Dynamic background subtraction based on appearance and motion pattern. *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on. IEEE*, pp: 1-6.
- [9] Wang H, Sang N, Yan Y. Real-time tracking combined with object segmentation. *Pattern Recognition (ICPR)*,

2014 22nd International Conference on IEEE, 24-28 Aug. 2014, pp: 4098-4103.

- [10]Deng G, Guo K. 2014. Self-adaptive background modeling research based on change detection and area training. *Electronics, Computer and Applications, 2014 IEEE Workshop on*, pp: 59 - 62.
- [11]Yuan G, Gao Y, Xu D, et al. 2011. A moving objects detection method based on a combination of improved local binary pattern texture and hue. Springer Berlin Heidelberg, pp: 261-268.