

Detection of Object Carried Using Spatio-temporal Pattern and Local Directional Pattern Descriptor

Han Su^{1,2}

1 Key Lab of Virtual Reality and Visual Computing of
Sichuan

Sichuan Normal University
Sichuan, Chengdu 610066

2 College of Computer Science
Sichuan Normal University
Sichuan, Chengdu 610101

Wenjie Wang^{2,3}

2 College of Computer Science

Sichuan Normal University
Sichuan, Chengdu 610101

3 Chengdu University
Sichuan, Chengdu 610106

Abstract—A novel method of carried object detection based on silhouette width and spatio-temporal analysis is proposed in this paper. The width vector of silhouette is calculated to characterize the spatial feature and the width image is adopted to preserve the temporal features. The width image which is considered as the texture image represents the walking sequence of pedestrian in a grey-level image. For analyzing the texture features, the local directional pattern code is used to encode the local properties and the local texture pattern descriptor is applied to capture the global characters. Chi-square distance equation is exploited to measure the difference between descriptors. Experimental results show that the width image is an effective and efficient representation, the local directional pattern descriptor is robust and insensitive to noise, and our method is effective.

Keywords—*carrying object detection; silhouette; the width image; spatio-temporal analysis; local texture pattern descriptor*

I. INTRODUCTION

As an important part of video surveillance system, carried object detection is intended to determine whether the pedestrian is carrying, the carrying type of the walking pedestrian, and the interaction of the carrier. Compared with other part of video surveillance system, it is challenging and difficult, but it also has a vast application prospect. At present, the research is still in the primary stage.

The earlier research on carried object detection is carried out by I. Haritaoglu, R. Cutler, D. Harwood, and L.S. Davis. They combine periodic motion estimation with static symmetry analysis of the silhouettes and propose a method to detect the backpack carried[1]. Moreover, they build a real time surveillance system which is named as W4, carried object detection is discussed[2]. They construct appearance models for silhouettes. The method can recognize events between people and objects[2]. [3] is based on silhouette and divided the silhouette into four sub-region, compute the width series and their corresponding autocorrelation functions in 2002. In [4], the authors use three different Gabor filter based multi resolution representations to present a multi-resolution

representation of human gait, and propose the general tensor discriminant analysis (GTDA) to reduce the ill-posed problem. A. Branca et al. represent the moving objects by using wavelet representation[5]. [6] models dynamic shape deformations based on kinematics manifold embedding and decomposable generative models by kernel map and multilinear analysis. Dima Damen et al. create the temporal template which is aligned by Iterative Closest Point algorithm and match it with exemplar templates[7]. Later, a fuzzy c-means algorithm[8] is proposed to improve [7] by Yuan Bo et al. In [9], the models are built by topological features. [10] gets the generic shape properties and builds the geometric shape models instead of using pre-trained object class models or solely relying on protrusions. [11] uses star skeletonization technique with the adaptive centroid point to present features, and it works on the luggage and a backpack carried type. Recently, [12] uses ratio histogram, [13] proposes a learning method with minimal supervision to obtain the candidate image and a kernel SVM classifier is used to select the plausible object region. The body main axis is present in [14].

Like these methods, our method is also based on silhouette and symmetry analysis. Obviously, the area of silhouette, the centroid and shape symmetry will change when a person carry an object. Capturing these features can classify carrying object. In this paper, the reference axis is used to present the features. But different from the axis which computed by PCA, the main axis and star skeleton method etc., the computation of the reference axis in our paper depends on the shoulder part which is not sensitive to carry. For a given sample, the center line is extracted for calculating the deviation. Another important difference is that we capture the key frames for judgment instead of all silhouettes in a gait period. It makes the features more effective.

Unlike these methods, Tobias Senst, Rubén Heras Evangelio, Volker Eiselein, and et al. use gray value information instead of segmentation masks and present features by periodicity dependency descriptor[15]. They try to avoid the prerequisite such as segmentation and track. Moreover, they adopt the HOG descriptor directly on the

FTLE field based on Lagrangian dynamics [16]. They also use we use a Gaussian mixture motion model to detect[17]. Our method adopts the grey-level image to present the feature and analyzes the image directly as [15]. But different with their method[15], the feature image which is the grey-level image is not the silhouette image, is a kind of dynamic property image which is calculated from the walking sequence and combine the spatial and temporal features. Our earlier research on gait recognition and human motion analysis, we proposes the similar images to describe dynamic patterns[18]. The feature images which capture both the shape structure information of each frame and dynamic properties of gait sequence are based on the width of silhouette. It is an effective approach for describing the spatial and temporal pattern. In this paper, the feature image is considered as the texture image. According to the properties of the feature image, local directional pattern(LDP) descriptor is applied to analyze the features.

This paper is structured as follows: In Section II we present and demonstrate our method. Then, in Section III we evaluate the performance of the width image and the local texture descriptor and discuss its results. Finally, we conclude our works in Section IV.

II. THE PROPOSAL METHOD

The method of carried detection will encounter many problem in the field of image processing and pattern recognition, such as image segmentation, tracking, image understanding and so on. These problems are difficulties in the fields. The research on carried detection is still in the initial stage, and there are still many problem to be solved. In this paper, we focus on detecting object carrying and the carrying type. Like other method, we assume that there is only one pedestrian in the observation area, no occlusion between pedestrians, and the camera view is perpendicular to the walking direction of pedestrians.

The overview of our method is shown in Fig. 1. It is composed of preprocessing, feature representation by the width image, feature analysis by local directional pattern descriptor, and classification.

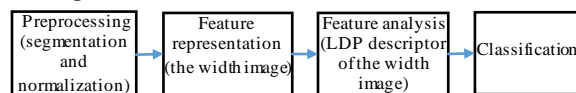


Figure 1. The overview of carried object detection system.

A. Preprocessing

To segmentation, the Codebook method [19] is used to model the background and segmentation. Fig. 2 shows the segmentation result by Codebook. In Fig. 2(a), some holes are still in the image. To amend it, the morphological operators are adopted. Fig. 2(b) is the corrected result.



Figure 2. The segmentation result by Codebook and morphological operators.

Normally, the 2D affine transformation matrix is computed through comparing the rotation, shearing, reflection and orthogonal projection and scaling relationship between the reference image and the incoming one. In our case, only the shape size including the height and width of silhouettes should be normalized. To calculate the scaling parameters by comparing the height and width of the reference image and the incoming one, the affine transformation matrix is got. But for carrying object, the width of silhouette which is an important factor of detection will be effected by the transformation matrix. The width information is greatly changed by normalizing. The result will be wrong, so the height is only element to compute the transformation matrix in this paper.

Surely, the width of silhouette also need to be corrected. The way we normalize the width is depended on the height of silhouette. Like other method, the main idea is normalizing the width according to the same scale ratio of silhouette height. The reference silhouettes are collected in training progress. Only the walking pedestrians without carrying are chosen to be the reference subjects. Each image of the reference sequence is segmented, and the bounding box is used to bind the silhouette. The bounding box is decided by the topmost, bottommost, leftmost and rightmost pixels of silhouette. There are many different height values of bounding box corresponding to the different silhouette shape. The reference height H_R is equal to the quotient between the sum of the all bounding box height and the number of bounding box. In fact, H_R is the average value of the height of all bounding box which is in training process.

Comparing H_R and the silhouette height of one incoming image which is named as H , the ratio is equal to the quotient of H_R and H . Thus, the transformation matrix is ready. At same time, this ratio is used to normalize the width. The width will change the same ratio with the height of silhouette. The bounding box has the different width because of the different silhouette width. The variance of width not the absolute values is used in the feature analysis step, so normalizing width by the same ratio with height is the right way to get the final result.

B. Spatio-temporal Feature

In this paper, we focus on detection of carried object based on silhouette. For silhouette sequence, the silhouette shape and the dynamic information of silhouette shape are the key characters and easy to extract. How to extract the shape and the dynamic diversification of shape, present them and analyze the salient property between the carried object status and no carry status by them is the key problem. We try to describe the silhouette shape in single frame to represent the spatial feature of pedestrian, and capture the shape continuous changes on the time line to represent temporal feature. The spatio-temporal pattern include both the spatial feature and temporal features.

In a broad sense, walking with carrying object is one kind of special gait. Among many gait recognition methods based on silhouette, the methods using spatial and temporal analysis is the main trend method. Especially, some researches represent gait in a single image while preserving temporal information such as MHI(motion-history image), MEI(motion-energy image), GEI(Gait energy image), and et al[20,21,22]. MEI presents where

motion has occurred in a gait image sequence by a binary image and MHI represent show motion in the image is moving by the grey image. [22] extracts the binary silhouette and consider the average silhouette image as GEI which preserving temporal information. These methods work well on gait recognition.

To compare with such features, silhouette width is key feature which varies considerably in some walking status after carried. For example, when the people is carrying a shoulder bag, the width of middle part of body don't periodical change as normal. When the people is carrying a luggage, the lower part width is always big and the variance is small between two silhouettes in whole walking cycle. In our earlier research on gait analysis and human action recognition, the silhouette width are used to present gait and action. We name it as the periodic sequence width images which record the 3D information by a 2D image. In this paper, the silhouette width and its variance are basic features. We also use this idea to present the feature of the carried object. The details as follows.

The first step is to describe the spatial feature in a frame. That is to compute the silhouette width for every silhouette in bounding box which are normalized before. The silhouette width in this paper is defined as the distance between the leftmost pixel and the rightmost pixel of silhouette in same row. Fig. 3 shows the silhouette width, width vector and the variances. In Fig. 3(a), P1 and P2 are the leftmost pixel and rightmost pixel on the edge of silhouette in the i th row of the bounding box K, P3 and P4 are the leftmost pixel and rightmost pixel in the j th row. x_{ki} which is the silhouette width in the i th row is equal to the difference of vertical coordinate between P1 and P2, x_{kj} is the difference between P3 and P4. There is one silhouette which is belong to one of the gait status in the bounding box K. Computing the width in all rows of the bounding box K, we can get $X_k = (x_{k1}, x_{k2}, \dots, x_{ki}, \dots, x_{kj}, \dots, x_{kn})$ which is the silhouette width vector. It is shown in Fig. 3(b). The blue curve in Fig. 3(b) is the width vector X of Fig. 3(a). In Fig. 3, the horizontal axis is the width value, and the vertical axis is the position of the silhouette in the bounding box. Considering the periodicity of walking, we can compute a series of width vector according to the cycle. $X = (X_1, X_2, \dots, X_k, \dots, X_m)$ is obtained from one sequence of a given walking pedestrian. Each X_i is one of curve in the Fig. 3(c). X is named as the width vector sequence. Fig. 3 (c) and (d) are the width vector curves from different carrying type, the curves in Fig. 3(c) is the backpack carrying, and the luggage carrying is shown in

Fig. 3(d). Comparing the Fig. 3 (b) and (c), their width at the same status are dramatically different because the carrying type. The range of width value is also different. Obviously, the silhouette width of chest part is affected by the backpack, the lower part width broaden by the luggage and he period of the width is changed.

Similar to the method that we mention in the first few paragraphs, we represent the width vectors in a single image. Fig. 4 shows how the vectors convert to be the grey-level images. These grey-level images record the change of 2D shape with time. The images is named as the width image.

The second step of presenting temporal pattern is to convert the width vectors to an image. Each width value of width vector is corresponding to one pixel, and a width vector is transformed to one of columns in the grey-level image. Fig. 4 (a) is the curves which represent the width vectors, the curves is calculated with the timeline of silhouette. But in Fig. 4 (a), the order of the curves is difficult to preserve. However, the order is important because the variance between two width vectors at random are insignificant for detection. The order of curves is the order of frames in the sequence. According to the order of frames, the width vectors convert to the gray value successively. Fig. 4 (b) is the grey-level image which transform from (a). The brightness of each pixel $Q(x,y)$ in Fig. 4(b) or (c) is equal to the width value on y th row in x th frame. The width value is bigger, and the point is brighter. The brightness on the red line in Fig. 4(c) describes the width variance at the same row in different frames. The red line is in the range of lower part body and around the limb. For a given walking sequence without carry, the width is periodical. Especially, owing to the limb swing, the width variances of lower part body are dramatic. At the same time, the brightness dramatically changes with that of the width. The value of each pixel records the shape structure information of each frame and dynamic properties. The carry type of the subject in Fig. 4 (b) and (c) are backpack and luggage. Obviously, the brightness periodically changes excluding the back part of body in Fig. 4 (b). The brightness value of limb part is larger than normal in Fig. 4(c). The cycle of that part changes unreasonably also. It means the luggage expands the width significantly.

The width image is 2D image. But actually, the images which show in Fig. 4 maintain the 2D silhouette shape and its changes on the timeline. The way to keep the spatio-temporal pattern in the gray image is simply and

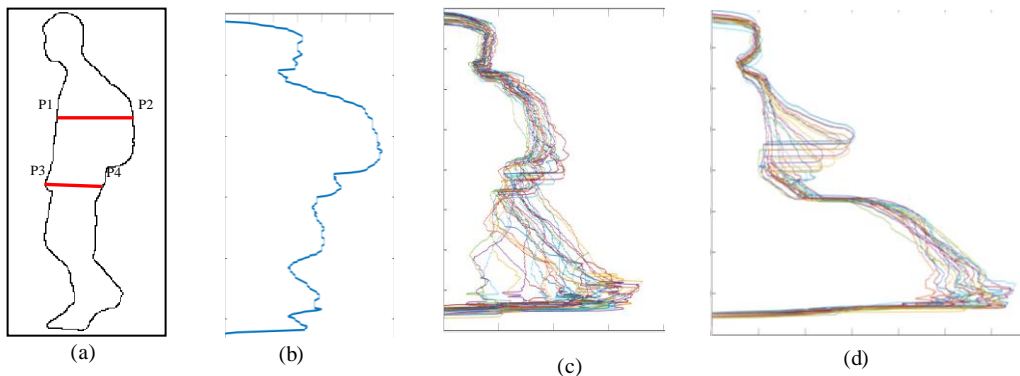


Figure 3. The silhouette width and width vector.

effectively. The width image describe the holistic and spatiotemporal pattern precisely and preserve the useful feature for analysis. Moreover, the grey-level image is robust for self-occlusion and noise problem.

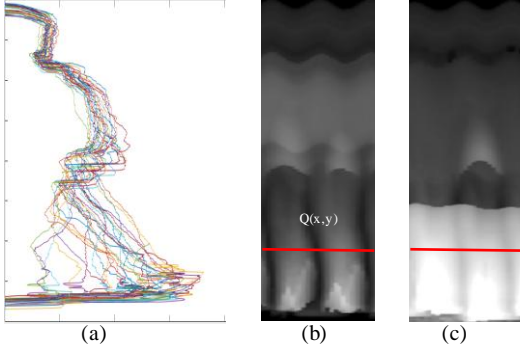


Figure 4. The width images of the subject who is carrying backpack and luggage respectively.

C. Feature analysis using LDP descriptor

The width image is regarded as a texture image in our method. The next step is to analyze the width and recognize different carry type from them using local texture pattern descriptor. Among the texture descriptors, LDP descriptor contains the fine detail of images, such as edges, corners and other texture features. Analyzing the property of the width images, we choose the advanced LDP descriptor to extract features.

1) LDP

LDP is proposed by Taxkeed Jaid, Md.Hasanul Kabir and et al. in [24,25,26]. LDP which is similar to LBP (local Binary Pattern) is a binary code. For each pixel of a given image, an eight bit binary code is assigned. The code generates by comparing the edge responses value. The generating process of LDP descriptor is as follows.

The first step is to calculate the eight directional edge response from the edge mask for every pixel. The edge masks include Kirsch, Prewitt, Roberts, Canny, Sobel, and so on. The original method chooses Kirsch mask, some advanced methods use Prewitt and et al. In this paper, we use Kirsch mask to calculate the directional edge response value. There are eight direction operators in Kirsch masks. It is named as M_0, M_1, \dots, M_7 . Fig. 5 shows the eight operators. Each pixel of the width image, eight edge response values are computed with M_0, M_1, \dots, M_7 .

$$\begin{array}{cccc} \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} & \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\ (a)M_0 & (b)M_1 & (c)M_2 & (d)M_3 \\ \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \\ (e)M_4 & (f)M_5 & (g)M_6 & (h)M_7 \end{array}$$

Figure 5. Kirsch edge response masks

The second step is to code according to comparing the edge response values. An example shows in Fig. 6. Fig. 6(c) is an image patch, the size of the center block is 3×3 , we will calculate the edge response for the center pixel V . We use the red rectangle to refer the center block which is calculated. Fig. 6(a) is the position of edge response value, (b) is the bit position of encoding, (d) is the process of generating the binary code, (e) is the directional response

value of eight direction, and (f) is the bit code of (e). We get eight mask values for V with eight directional masks. Reference the step of [25], the response values are denoted as m_0, m_1, \dots, m_7 . Then, the response values is sorted by their absolute values, and the ranks are listed in Fig. 6(d). For LDP, the rank of the response values are different, the important of that are different. There is an important parameter k in encoding LDP descriptor. The function of k is to decide the number of the prominent directions. Equation (1) [27] is the encoding rule.

$$LDP_k = \sum_{i=0}^7 c_i \times 2^i \quad (1)$$

$$c_i = \begin{cases} 1 & m_i - m_k \geq 0 \\ 0 & m_i - m_k < 0 \end{cases} \quad (2)$$

where m_i is the response value. In Fig. 6, k is equal to 3, the value of m and m is the k top values, their code bit are set to 1, and other bit are set to 0. The LDP code is 10010010 and the decimal value is equal to 146 which is shown in the last row of Fig. 6(d).

In fact, the main idea of this step is to choose the potential directional edge for every pixels and encode it. [25] demonstrates that LDP is more stable than LBP when the image is in non-monotonic illumination variation and random noise condition.

The last step is to generate the LDP histogram. This step is similar to other local texture patterns. The histogram depends on the block size which we explain in next paragraph.

2) Feature extraction by LDP descriptor

In our method, the code of each pixel in the width image is generated according to the steps in 2.3.1. The width image is converted to the encoded image. Fig. 7 explains the process of generating LDP code and LDP histogram of the width image.

We normalize the width image to the same size $R \times C$. The bilinear interpolation algorithm is used to resize the image size. In our method, an advanced LDP [28] is adopted. The difference between the original LDP with the advanced one is in the second step. When the response values are sorted, the original one uses the absolute of the values, and the advanced one directly use the values. Fig. 7(a) shows one normalized width image, (b) is the LDP image of (a) using the advanced LDP. The advanced LDP only encodes the positive directions which can fully exploit the local features.

As we know, the different carrying type cause the area that width is affected different. Easy to compare the different carrying type, the LDP image is divided into blocks (shown in Fig. 7(b)). The divide rule in our method is that the LDP image is divided into 8×8 blocks in one normalized cycle. The numbers of row is determined by the carrying type and the minimum size of carried object that we can detect, that of column is according to the key status in one walking cycle. Normally, there are 8 different key status in a period. For each block, one feature histogram is extracted. To get the histogram of the LDP image, equation 3 [27] is used. Each pattern is accumulated into one bin.

$$H(\tau) = \sum_{x=1}^{R/8} \sum_{y=1}^{C/8} f(LDP_k(x, y), \tau) \quad (3)$$

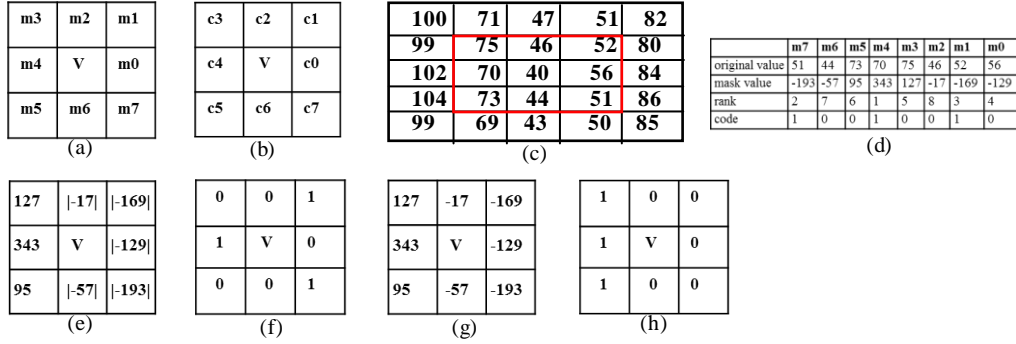


Figure 6. An example of describing and encoding by LDP.

$$f(LDP_k(x, y), \tau) = \begin{cases} 1 & LDP_k(x, y) = \tau \\ 0 & otherwise \end{cases} \quad (4)$$

Equation (4) is also reference [27]. Fig. 7(c) is the histogram of each blocks which calculated over Fig. 7(b). There are 64 histograms of the width image in one cycle. The histograms which are the LDP descriptor are the features for detection.

3) The similarity measure

To compare the difference between two width images, the distance of their LDP descriptors should be measured. That distance is equal to the distance of the corresponding histograms. The Chi square distance is adopted to calculate the distance of histograms. The equation is defined as

$$\chi^2(B_1, B_2) = \sum_i \frac{(B_1(i) - B_2(i))^2}{B_1(i) + B_2(i)} \quad (5)$$

where B_1 and B_2 are histograms of two block in two LDP images, i refers to the different bins in the histograms. B_1 and B_2 are corresponding to the same block in two different images.

III. EXPERIMENTS

Our method is tested on two databases. We adopt the nearest neighbor classifier and leave-one-out cross validation technique to classify. The key parameter k of LDP chooses 3.

First our method is evaluated on the dataset B of CASIA Gait Database. There are 124 subjects in the dataset. The videos were taken in 11 perspectives in three different conditions which are walking, dressing overcoat and carrying object. The result of the horizontal view sequence is 89%. In dataset B, the carried type is backpack. For our method, the performance is not good. The reason is that the width of chest part has no typical periodicity, and the brightness corresponding to the width value is the only factor for analysis. However, the silhouette cannot be normalized by the width. The individual difference of silhouette affects the result directly. The specificity of

chest part is in both two phases.

Second the method is tested on our database not the public database of PETS2006 because the database focuses on the detection of left luggage in complex scene. There are 32 volunteers and 4 carried type in our database. The volunteers are consisted of 14 females and 18 males, the height is from 148cm to 186cm, and the weight is firm 47kg to 85kg. The extensive body shapes are involved in. The distance between the subjects and camera is about 12 meters and 3 different views for data collection. The result on our database is shown in Table 1.

TABLE I. CONFUSION MATRIX ON OUR DATASET IN HORIZONTAL VIEW

	No carrying	Backpack	Handbag	Shoulder bag	Luggage
No carrying	0.79	0.19	0.02	0	0
Backpack	0.17	0.83	0	0	0
Handbag	0.02	0	0.98	0	0
Shoulder bag	0	0	0	1	0
Luggage	0	0	0	0	1

It is clearly, our method is effective for representing the spatio-temporal feature by the width image and analyzing the features based on LDP descriptor. Similar to the first test, the performance of backpack carried is not good as other type. Fig. 8 shows the width image and their LDP pattern with different carried type, (a), (b), (c), and (d) are the width image of backpack, handbag, shoulder bag, and luggage carried, respectively.

IV. CONCLUSIONS

In this paper we have proposed a method for detecting carried object based on spatio-temporal pattern using LDP descriptor. We adopted the width image to represent the feature of walking sequence and the silhouette of each

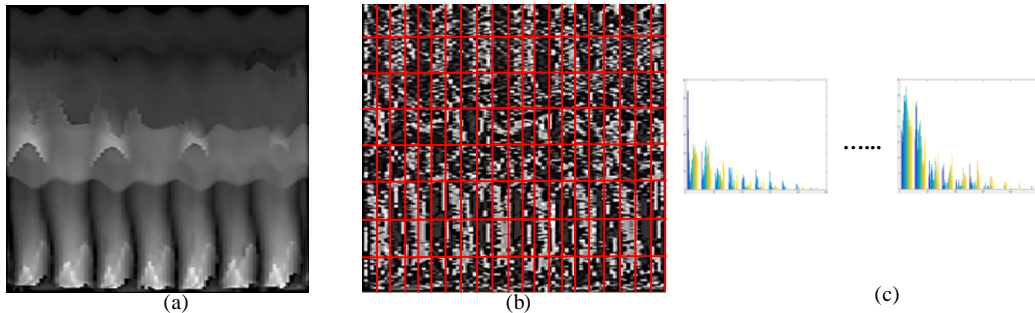


Figure 7. Generating LDP descriptor for the width image.

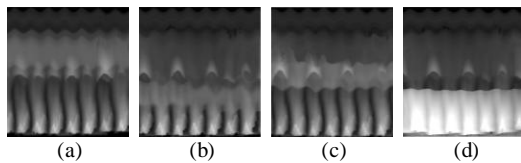


Figure 8. The width images and their LDP pattern with four different carried types.

frame. The width vector of silhouette is the basic feature element. The width image converts a set of width vectors into be a grey-level image. The brightness in the image is equal to the value of width value after normalized. The width image not only describe the spatial feature, but also capture the temporal information. Then, we use LDP descriptor to represent the width image to be the feature image which is divided into blocks. The number of block is decided by the walking cycle and the minimum size of object that we can detect.

Future improvements to this method might be discuss the view angle and multi-pedestrians with multi-camera.

ACKNOWLEDGMENT

This research has received funding from the Key Project of Natural Science Foundation of Sichuan Provincial Education Department (Grant No.14ZA0029), Visual Computing and Virtual Reality Key Laboratory of Sichuan Province Foundation (Grant No.KJ201419) and 251 key talent project of Sichuan Normal University. We would like to thank the Institute of Automatic, Chinese Academic of Science for providing the database and the volunteers who take part in and help us to creating our database.

REFERENCES

- [1] I. Haritaoglu, D. Harwood, and L. Davis, "Hydra: Multiple People Detection and Tracking Using Silhouettes," *Proc. IEEE Workshop Visual Surveillance*, 1999
- [2] I. Haritaoglu, D. Harwood, and L.S. Davis, "W4 : Real-Time Surveillance of People and Their Activities," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, Aug. 2000
- [3] 9 Javier Lamar-León, Raul Alonso Baryolo, Edel B. García Reyes, Rocío González-Díaz, "Gait-Based Carried Object Detection Using Persistent Homology," *Proceedings of CIARP 2014*, pp.836-843, 2014
- [4] 13 Radu Dondera, Vlad I. Morariu, Larry S. Davis, "Learning to Detect Carried Objects with Minimal Supervision," *Proceedings of CVPR Workshops 2013*, pp. 759-766, 2013
- [5] 4 D. Tao, X. Li, S.J. Maybank, and W. Xindong, "Human Carrying Status in Visual Surveillance," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, 2006.
- [6] 10 Aryana Tavanai, Muralikrishna Sridhar, Feng Gu, Anthony G. Cohn, David C. Hogg, "Carried Object Detection and Tracking Using Geometric Shape Models and Spatio-temporal Consistency," *Proceedings of ICVS 2013*, pp. 223-233, 2013
- [7] 3 C. Benabdelkader and L. Davis, "Detection of People Carrying Objects: A Motion-Based Recognition Approach," *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pp.378-384, 2002.
- [8] 5 A. Branca, M. Leo, G. Attolico, and A. Distanti, "Detection of Objects Carried by People," *Proc. Int'l Conf. Image Processing*, vol. 3, pp. 317-320, 2002
- [9] 7 D. Damen and D. Hogg, "Detecting Carried Objects from Sequences of Walking Pedestrians," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2012
- [10] 8 Bo Yuan ; Inst. of Inf. Sci., Beijing Jiaotong Univ., Beijing, China ; Qiuqi Ruan ; Gaoyun An, "Carried object detection in short video sequences," *Proceedings of ICSP 2014*, pp. 1311 - 1316, 2014
- [11] 6 Chan-Su Lee, Ahmed M. Elgammal, "Carrying Object Detection Using Pose Preserving Dynamic Shape Models," *Proceedings of AMDO 2006*, pp.315-325, 2006
- [12] 11 Rawin Chayanurak, Nagul Cooharajanane, Shin'ichi Satoh, and Rajalida Lipikom, "Carried Object Detection using Star Skeleton with Adaptive Centroid and Time Series Graph," *In Proc. 10th IEEE Int'l Conf. Signal Processing*, 2010
- [13] 12 C.H. Chuang, J.W. Hsieh, L.W. Tsai, S.Y. Chen, and K.C. Fan, "Carried Object Detection Using Ratio Histogram and Its Application to Suspicious Event Analysis," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 19, no. 6, pp. 911-916, June 2009
- [14] 14 Yue Qi, Guo-Chang Huang, and Yun-Hong Wang, "Carrying Object Detection and Tracking Based on Body Main Axis," *International Conference on Wavelet Analysis and Pattern Recognition*, 2007
- [15] T. Senst, R. Heras Evangelio, V. Eiselein, M. Patzold, and T. Sikora, "Towards detecting people carrying objects: A periodicity dependency pattern approach," *International Conference on Computer Vision Theory and Applications (VIS-APP 10)*, 2010
- [16] Tobias Senst, Alexander Kuhn, Holger Theisel, Thomas Sikora, "Detecting People Carrying Objects Utilizing Lagrangian Dynamics," *Proc. 9th IEEE Int'l Conf. Advanced Video and Signal-Based Surveillance (AVSS)*, 2012
- [17] Tobias Senst, Rubén Heras Evangelio and Thomas Sikora, "Detecting People Carrying Objects based on an Optical Flow Motion Model," *IEEE Workshop on Applications of Computer Vision (WACV)*, 2011, pp. 301-306
- [18] Han Su, Guoyue Chen, "A new method of gait recognition independent of view angle," *In Proc. Of ICMLC 2010*, 3091-3096.
- [19] Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, Larry Davis, "Real-time foreground-background segmentation using codebook model," *Real-Time Imaging*, vol.11, no.3, 2005, pp.172-185, doi: 10.1016/j.rti.2004.12.004.
- [20] Bobick, A., Davis, J. "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no.3, pp. 257-267, 2001.
- [21] Weinland, D., Ronfard, R., Boyer, E., "Free viewpoint action recognition using motion history volumes," *Computer Vision and Image Understanding*, vol. 104, no.2, pp. 249-257.
- [22] Han Su, Guoyue Chen, "A new method of gait recognition independent of view angle," *In Proc. Of ICMLC 2010*, pp.3091-3096.
- [23] Han J, Bhanu B, "Individual recognition using gait energy image," *IEEE Trans Pattern Anal Mach Intell.* 2006,28(2), pp.316-22.
- [24] Taskeed Jabid, Md. Hasanul Kabir, Oksam Chae, "Local Directional Pattern (LDP) - A Robust Image Descriptor for Object Recognition," *AVSS 2010*, pp.482-487
- [25] Taskeed Jabid, Md. Hasanul Kabir, Oksam Chae, "Facial expression recognition using Local Directional Pattern (LDP)," *ICIP 2010*, pp.1605-1608
- [26] Taskeed Jabid, Md. Hasanul Kabir, Oksam Chae, "Gender Classification Using Local Directional Pattern (LDP)," *ICPR 2010*, pp.2162-2165
- [27] Hasanul Kabir, Taskeed Jabid, and Oksam Chae, "Local Directional Pattern Variance (LDPv): A Robust Feature Descriptor for Facial Expression Recognition," *The International Arab Journal of Information Technology*, 2012, 9(4), pp.382-391 July
- [28] WU Wei, JIANG Yanxia, ZHANG Zhe, "The Face Recognition Algorithm Based on Improved Local Directional Pattern," *Opto-Electronic Engineering*, 2014,41(12), pp.72-77
- [29] Adin Ramirez Rivera, Jorge A. Rojas Castillo, Oksam Chae, "Local Directional Texture pattern image descriptor," *Pattern Recognition Letters*, 2015,51, pp.94-100