

Automatic Recognition of the Micro-Expressions of Human Torso Posture in the Environment of Interrogation

Haopeng Wang^{1,2}, Xianying Feng^{2,*}

¹Shandong Police College, Jinan, China

²Key Laboratory of High Efficiency and Clean Mechanical Manufacture, Ministry of Education, Shandong University, Jinan, China

* Corresponding author: Xianying Feng

Keywords: Human torso; Micro-expressions; Posture recognition; 6 points human skeleton model; Local binary pattern.

Abstract. In order to improve trial efficiency and realize the automatic recognition of the micro-expressions of criminal suspects in the interrogation process, according to the characteristics of the human torso micro-expressions posture, A 6 points human torso model which is used to describe the human torso micro-expressions posture is proposed. An automatic recognition algorithm of the micro-expressions of human torso posture is designed. First, the original image of the human torso posture is converted to a local binary pattern image. The edge of the human torso is detected on the local binary pattern image by the canny operator. Then, the human torso posture characteristics are described and calculated by 6 points human skeleton model. Compared with the traditional 20 points human skeleton model, the key points of the left and right chest are added to the skeleton model, and the key point of waist is expanded to about two points. The 6 points model is more suitable for describing the micro-expressions characteristics of human torso posture. Finally, the different postures are classified and recognized by support vector machine. Experimental results indicate that the recognition results of the 8 different human torso pose are uniform, and are all in the 83.1~92.5%. The average correct recognition rate of 8 different human torso posture is 87.9%. The average correct recognition rate of the recognition algorithm using the traditional 20 points human skeleton model is only 67.3%. Compared with the traditional 20 points human skeleton model, the 6 points human torso model can be more accurate and reasonable description of the characteristics of the human torso micro-expressions posture. The algorithm can satisfy the requirements of human torso micro-expressions posture automatic recognition, and effectively judge the different postures.

1. Introduction

Psychological stress micro-reaction (Jiang, 2014) is the effective stimulation of the moment shown by effective stimulus is not controlled by the thinking of the real reaction, also referred to as micro-reactions. In the course of the trial observation of micro-reaction of suspects can infer the true mentality and feelings of criminal suspects, can greatly enhance the efficiency of interrogation (Dong, 2012).

Human micro reaction is mainly reflected in 3 aspects, such as facial expression, limb movements and trunk posture (Ekman, 1974). As the micro-reaction has the characteristics of low frequency, short duration and small amplitude of motion, using artificial recognition to identify the micro-reaction is not only difficult to identify, more misidentification and identification of leakage, but also need to identify people for a longer time to identify training. So the research on the automatic recognition of micro-reaction using image processing technology can improve the accuracy of micro-reaction recognition. Shreve in the United States (Shreve, 2011), Polikovsky in Japan (Polikovsky, 2009), and Fu Xiaolan in China (Liu, 2015) all in-depth studied the automatic identification of micro-responses, but most studies focused on the two aspects of facial expressions and movements of limbs, torso posture is relatively small. During the interrogation, the suspects usually keep their seated position, the

limbs are restrained and the torso movement is subtle. It is difficult to describe the characteristics of the torso posture with the traditional 20-point model.

According to the characteristics of human torso stress micro-reaction posture, in this paper, a 6-point model is proposed to describe the posture of human torso stress micro-reaction, and a related automatic recognition algorithm is designed.

2. The Structure of Image Inspection System and Its Working

2.1 The common micro - reactions of body torso posture in sitting posture

The body torso discussed in this article includes the shoulder, chest, abdomen and buttocks. Sitting posture, the human body posture common micro-reactions can be summarized as seven types, the action characteristics and emotional meaning as shown in Table 1.

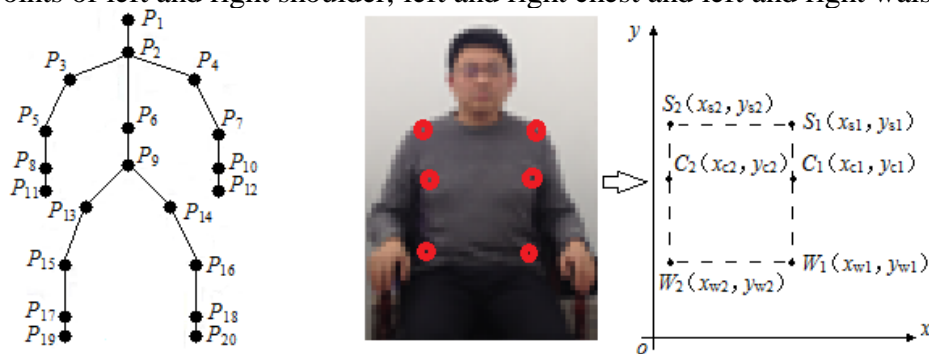
Table.1 Common micro-expressions posture of human torso

	Name	Motion characteristics	Emotional meaning
1	Ventral veto	Tilt away from a person or thing	Escape, negative, opposition
2	Ventral pretage	Leaning toward a person or thing	Intimate, positive, agree
3	Trunk bending	bend over	Flattery, respect, humility
4	Trunk extension	Backward stretch	Comfortable, relaxed, negative
5	Chest	Torso straight, chest slightly convex	Show off, arrogant, provocative
6	Shrug	The trunk does not move, shoulder lift up	Suspicion, fear, tension
7	Shoulder contraction	Shoulder slowly lift up the top	Negative, cautious, submissive

Through the description of the action characteristics in Table 1, it can be seen that the micro-reactions of human torso in sitting position are mainly concentrated in the three parts of the shoulder, chest and waist, and often is not a part of a separate action, but the three parts of the linkage. After the end of some micro-reaction action will maintain a certain time trunk posture, the other micro-action range of these actions are very small, strong covert.

2.2 Establishment of human torso model

The current 20-point human skeleton model is shown in Figure 1 (a). Which involves the human body has four points of the torso, respectively, right shoulder P_3 , left shoulder P_4 , waist P_6 and hip P_9 . According to the analysis of 2.1, the human body's micro-reaction will involve torso twisting action and chest action, only 20-point in the model of P_3 , P_4 , P_6 and P_9 is difficult to describe the characteristics of these movements more accurately. The 6-point model of human torso is established, which is shown in Figure 1(b), according to the characteristics of human torso micro-reaction and extracting six points of left and right shoulder, left and right chest and left and right waist.



(a) The traditional 20 points human skeleton model (b) The 6 points human torso model of this paper
Figure 1. Human torso model

In Fig.1(b), the image of the lower left corner as the origin of the establishment of coordinate system, coordinate system S_1 and S_2 for human torso rectangular outline of the upper left corner and the upper

right point, representing the human body left and right shoulder key points; W_1 and W_2 are the lower left corner of the human body's rectangle outline and the lower right corner, representing the human body around the waist of the key points; C_1 and C_2 on behalf of the body of the left and right chest points, but these two points cannot be directly calibrated by the human body image, in this paper, we use the torso and the lower edge corner of the left and right arm joints instead of indicates.

2.3 Selection of human torso posture

When the human torso pose recognition is performed, the position attributes x and y and the time attribute t of each key point are recorded frame by frame, and the relative positions of six key points and the rate of change were calculated to identify the characteristics of different trunk pose. According to the description of the action characteristics in Table 1, the main body torso features selected in this paper are as follows:

- 1) Six key points of the location features, that is, six points of the vertical and horizontal coordinates;
- 2) The relative distance characteristics of six key points, including l_{s1s2} , l_{c1c2} , l_{w1w2} , l_{s1c1} , l_{s2c2} , l_{c1w1} , l_{c2w2} , l_{s1w1} , l_{s2w2} ;
- 3) The angle between the six key points, including α_{s1s2} , α_{c1c2} , α_{w1w2} , α_{s1c1} , α_{s2c2} , α_{c1w1} , α_{c2w2} , α_{s1w1} , α_{s2w2} ;
- 4) Other geometric features between the six key points, including $S_{s1c1c2s2}$ and $R_{s1c1c2s2}$, $S_{c1w1w2c2}$ and $R_{c1w1w2c2}$, $S_{s1w1w2s2}$ and $R_{s1w1w2s2}$;
- 5) The rate characteristic of the change of all geometric features in the above four items.

3. Extraction of human torso region

3.1 Human Foreground Object Extraction

Commonly used foreground object extraction methods are background difference method, image sequence modeling method, the body shape modeling method and the depth of information method and so on (Kim, 2015). Interrogation room is a relatively closed independent environment, the background is fixed, the camera position angle is fixed, and the illumination change is small. Therefore, this paper uses the background difference method to extract the human foreground target.

3.2 Torso region extraction

Firstly, the whole human foreground object image is transformed into local binary pattern (LBP) image. Then, the canny operator is used to detect the edge of the LBP image to obtain the trunk and limb contour. Finally, according to the proportion of the human body parts the relationship experience value determines the torso region.

Canny operator is an edge detection operator based on the optimization algorithm (Hu, 2015). The idea is to denoise the original image $f(x, y)$ by using the two-dimensional Gaussian function $G(x, y)$, and get the smooth image $I(x, y)$ and the gradient magnitude $M(x, y)$ and the direction $O(x, y)$ are obtained by the finite difference calculation, such as formula

$$G(x, y) = \sqrt{-(x^2 + y^2) / 2\sigma^2} / 2\pi\sigma^2 \quad (1)$$

$$I(x, y) = G(x, y) \otimes f(x, y) \quad (2)$$

$$g_x(x, y) = G_x(x, y) \otimes I(x, y) \quad (3)$$

$$g_y(x, y) = G_y(x, y) \otimes I(x, y) \quad (4)$$

$$M(x, y) = \sqrt{g_x^2(x, y) + g_y^2(x, y)} \quad (5)$$

$$O(x, y) = \arctan[g_x(x, y), g_y(x, y)] \quad (6)$$

Where, x, y for the pixel coordinates; σ is the distribution of Gaussian function parameters; \otimes that convolution.

Canny operator has the advantages of relatively complete edge detection, strong continuity and high precision, but it is easy to detect the pseudo-edge when it is disturbed by the local noise. There are

many patterns of clothes, especially some regular stripes, the gradient of the square pattern, which are similar to human torso edges. This paper uses LBP image to solve this problem.

The local binary pattern (Yang, 2014) is a way to express the whole image characteristics through the frequency of the gray-value structure of the local area pixels. The gray image of a regular pattern of stripes and squares can be transformed into a LBP image, which can play a role in smoothing images, thereby reducing the production of pseudo-edge, as shown in Figure 2.

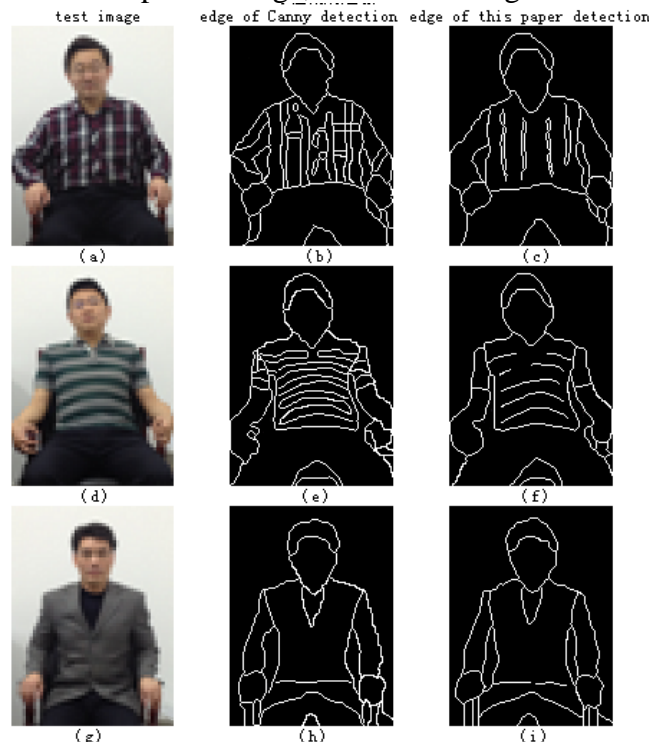


Figure 2 The edge detected by Canny operator and the edge detected by this paper

It can be seen from the figure that, for the test image (g), as the clothes pattern is very simple, the Canny operator has little difference from the results of the method designed in this paper, but the test images (a) and (d) are regular Square pattern and striped pattern have great influence on Canny edge detection, which results in a large number of pseudo-edges. The method proposed in this paper has fewer pseudo-edges and less influence on subsequent human torso division.

For a human image that depicts the edges, the human body torso and limbs can be distinguished according to the length and width ratio of each part of the human body and the proportion relation between the parts, so as to calibrate the six key points of human torso.

4. Human torso posture recognition

Support vector machine (SVM) is a supervised classification method. SVM has many advantages over traditional methods of machine learning, such as small sample size, non-linearity and high dimension. SVM has been widely used in human pose recognition (Song, 2014). In this paper, SVM is used as the classifier of the trunk posture, and seven SVM classifiers are used to combine them in one pair (ie, seven classifiers are constructed for seven human torso postures. With the remaining categories to distinguish), and the radial basis function (RBF) is used as the kernel function of the classifier.

In this paper, the main steps of human torso pose identification algorithm are as follows:

- 1) Image preprocessing, noise removal and so on;
- 2) After segmentation of the whole human foreground target area by background difference method, the image is binarized (LBP), and the edge of the LBP image is detected by Canny operator. The body and limb contours are obtained, the relationship between the parts of the empirical value of the torso area and calibration of six key points;

- 3) Record the x , y and t attributes of the six key points in the video image frame by frame, and calculate the position, geometry, and rate of change of the video image as features to identify different torso postures;
- 4) SVM classifier is used to classify and identify different pose.

5. Experiment

Experimental hardware environment for the Intel Core i3-3240 3.4GHz CPU desktop computer, 4GB of memory, the use of C # algorithm development, operating environment for Windows 7.

In this paper, ten different subjects were selected, each recording eight kinds of torso posture video (including the normal sitting position and seven micro-reaction posture in Table 1), recording ten segments of video (320×240 , 25frame/s), each video 3s, a total of 800 video samples. All samples were categorized according to the test subjects and tested using a one-off method.

In order to validate the effectiveness of the torso pose identification algorithm designed in this paper, three different algorithms are used to compare. The first one is the human torso pose recognition algorithm named HTPRA_6LC; the second one is based on the algorithm designed in this paper, the traditional 20-point human body model is used instead of the 6-point model described in this paper, named HTPRA_20LC; the three algorithms are also based on the algorithm designed in this paper, when extracting the edge of the torso, the original image is not transformed into LBP image, and is directly detected by Canny operator, named HTPRA_6C. The recognition results of the three algorithms are shown in Figure 3.

1	84.4	4.9		6.1	3.3			1.3
2	6.8	83.1	4.5				5.6	
3		3.9	88.7				5.8	1.6
4	9.9			85.3	4.8			
5	3.3			5.1	91.6			
6						89.3	10.7	
7			3.2			8.4	88.4	
8	4.4	3.1						92.5
	1	2	3	4	5	6	7	8

(a) Confusion matrix of HTPRA_6LC recognition results

1	67.6	15.4	1.7		7.2			8.1
2	11.9	61.6	16.4			2.6	4.2	3.3
3	8.6	19.3	62.8				9.3	
4	21.2			69.1	9.7			
5	12.4	3.5	5.6	14.5	56.7			7.3
6				6.1		76.5	17.4	
7	4.4		3.7	5.2		13.5	73.2	
8	5.2	8.6			15.4			70.8
	1	2	3	4	5	6	7	8

(b) Confusion matrix of HTPRA_20LC recognition results

1	77.8	8.2		5.6	2.1	1.3	1.8	3.2
2	4.5	79.5	6.4			1.8	2.3	5.5
3		13.4	73.2			7.3	6.1	
4	8.8	2.6		74.7	13.9			
5	9.3	3.5		7.1	76.4			3.7
6				4.6	1.5	85.5	8.4	
7	3.2		10.5			4.2	82.1	
8	6.1	3.8	4.7		2.2	1.4		81.8
	1	2	3	4	5	6	7	8

(c) Confusion matrix of HTPRA_6C recognition results

Note: 1 Ventral denial; 2 Ventral fronting; 3 Trunk bending; 4 Trunk extension; 5 Chest out; 6 Shrug; 7 Shoulder contraction; 8 normal sitting position

Figure 3 Confusion matrix of 3 algorithms recognition results

From Figure 3 (a), it can be seen that the HTPRA_6LC algorithm designed in this paper has an average recognition rate of 87.9% for eight different gestures, and the recognition results are fairly uniform, ranging from 83.1 to 92.5%. The recognition rate of ventral veto and ventral prepositions was lower than 85%, because these two kinds of gestures were smaller in their own motion range and their features were not obvious. Figure 3 (b) is HTPRA_20LC algorithm recognition results, the average correct recognition rate of 67.3%, far below the design of the algorithm, because the traditional 20-point human body model only shoulder, waist and hip key points, lack of chest key points, so shoulder action obvious shrug, shoulder contraction and other posture correct recognition rate is high, respectively 76.5% and 73.2%, chest recognition accuracy of only 56.7%. Figure 3 (c) is the result of HTPRA_6C algorithm recognition, the average correct recognition rate of 78.9%, lower than the algorithm designed in this paper, because the HTPRA_6C algorithm to determine the human torso edge using only Canny operator, The algorithm transforms the original image into LBP image and then uses Canny operator to determine the edge, which reduces the noise interference. However, the HTPRA_6C algorithm uses the 6-point human torso model proposed in this paper, so its correct recognition rate is higher than the HTPRA_20LC algorithm using the traditional 20-point human body model.

6. Conclusion

In this paper, the 6-point human skeleton model for human torso pose identification is proposed. In the extraction of human torso region, the original image is firstly transformed into LBP image and the edge is detected by canny operator. Finally, the experimental results show that the proposed method can identify the different pose effectively, and the average correct recognition rate of the human torso pose identification algorithm is 87.9% for 8 different pose. In the follow-up study, we will further analyze the characteristics of the posture of stress micro-reaction of body sitting torso, and find out the features that can describe the characteristics of these actions more accurately, and further improve the recognition accuracy.

Acknowledgements

Supported by Natural Science Foundation of Shandong Province of China (No. 2016ZRB03001) .
Supported by Science and Technology Project of Shandong Province of China (No. J16LN16).

References

- [1] Dong, N. 2012. Application of Micro-facial Expression & Reactions in Crime Investigation. *Journal of Henan Police College* 21(3): 52-54.
- [2] Ekman, P. & Friesen W V. 1974. Nonverbal behavior and psychopathology. *Contemporary Theory and Research* (1): 203-232.
- [3] Hu, H. P. & Zhu, M. 2015. Fast stereo matching based on seed pixel propagation. *Optics and Precision Engineering* 23(3): 887-894.
- [4] Jiang, ZH. Y. 2014. The application of micro-expressions in the trial. *The Chinese Procurators*, (18): 59-64.
- [5] Kim, Jaewoong. & Lee, Sukhan. 2015. Extracting major lines by recruiting zero-threshold canny edge links along sobel highlights. *IEEE Signal Processing Letters* 22(10): 1689-1692.
- [6] Liu, Z. Q. & Yin, J. Q. & Zhang, L. 2015. Human action recognition based on Kinect data principal component analysis. *Optics and Precision Engineering* 23(10): 702-711.
- [7] Polikovsky, S. & Kameda, Y. & Ohta, Y. 2009. Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor. *3rd International Conference on Imaging for Crime Detection and Prevention* London, United kingdom: ICDP.
- [8] Shreve, M. & Godavarthy, S. & Goldgof, D. 2011. Macro-and micro-expression spotting in long videos using spatio-temporal strain. *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops* Santa Barbara, CA, United states: FG, 2011: 51-56.
- [9] Song, J. M. & Zhang, H. & Gao, Z. 2014. Human action recognition based on multi-spatio-temporal features. *Journal of Optoelectronics Laser* 25(10): 2009-2017.
- [10] Yang, L. P. & Gu, X. H. 2014. Relative gradient histogram features for face recognition. *Optics and Precision Engineering* 22(1): 152-159.